

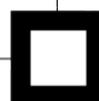
Editoriale

I protagonisti del futuro digitale dell'Italia sono già fra noi

Il Congresso Nazionale AICA tenuto a Fisciano a metà settembre nello straordinario campus dell'Università di Salerno sul tema: "Frontiere Digitali: dal Digital Divide alla Smart Society" ha dimostrato chiaramente come, pur nel contesto della più grave crisi economica registrata in Italia dal dopoguerra, stiano operando i nuovi protagonisti del futuro digitale che attraverso nuovi modi di imprenditorialità e di innovazione tecnologica aprono nuove frontiere per un nuovo ciclo di sviluppo aperto ai giovani e ricco di straordinarie opportunità.

Il Congresso di AICA ha inteso fare luce su questo mondo in grande fermento al fine di costruire attorno a questa realtà nascente un sistema paese che non ostacoli, ma supporti queste forze, consentendo libertà di futuro.

La sede di Salerno in un campus che ospita 40.000 studenti è stata scelta con il particolare obiettivo di portare nel Sud il dibattito e la ricerca di soluzioni offerte dalle tecnologie digitali per favorire la nascita di nuove competenze, nuove opportunità di lavoro e di impresa in questa grande parte d'Italia che ha troppo sofferto e soffre di crescente disimpegno e scarsa attenzione da parte delle scelte della politica e degli operatori economici, mentre possiede le grandi risorse, gli asset che guidano e determinano il futuro, cioè cultura, scienza, intelligenza e creatività.



Il Congresso ha dato la parola a tanti nuovi artigiani digitali che grazie alla rete ed alle tecnologie disponibili a costi limitati, quali smart phone e tablets e le nuove forme di software componibile per lo sviluppo di apps, stanno determinando processi di democratizzazione dell'accesso mobile in rete ed una straordinaria diffusione senza confini dei contenuti intelligenti di conoscenza, con l'utilizzo di grandi masse di informazioni, di Big Data, di Open data.

Vi sono oggi segnali non più deboli che vengono dai giovani, dalla nuova imprenditoria innovativa, che opera, con risorse finanziarie limitate, sulla frontiera dell'incrocio creativo tra bit e atomi, come nei Fab Lab, definiti la nuova bottega rinascimentale.

Dall'utilizzo delle stampanti 3D nasce la coniugazione tra competenze digitali e capacità manifatturiere che sono proprie della tradizione imprenditoriale e artigiana del nostro paese.

Nuove imprese, nuove iniziative nascono in Italia nel totale silenzio da parte dei media e delle istituzioni e sono create da giovani innovatori, che uniscono competenze digitali con skill manifatturieri, i "makers", mettendo assieme bit e atomi, con hardware open source, come il modello Arduino, lavorando su software e hardware libero. Abbiamo visto al Congresso come la diffusione di stampanti 3D a basso costo stia modificando anche la grande fabbrica attraverso processi di "digital fabrication revolution".

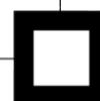
Sta nascendo finalmente anche in Italia una nuova generazione di artigiani digitali che stanno creando nuovi layer, nuovi snodi nella "value chain" digitale con le nuove forme di manifattura e con la straordinaria proliferazione di sviluppatori di apps, in un clima di open innovation, di scambio di conoscenza, di crossmedialità innovativa.

Abbiamo incontrato a Salerno giovani competenti e motivati il cui obiettivo è la creazione e valorizzazione di talenti e di opportunità d'impresa, non più un drammatico parcheggio senza speranza di futuro delle giovani generazioni, l'asset più strategico che invece l'Italia sta sprecando e regalando ad altri, così perdendo futuro.

In AICA siamo convinti, e quanto abbiamo visto al Congresso di Salerno ce lo ha confermato, che si debba e si possa invertire il regresso in corso nel paese, ma occorre non perdere tempo.

L'intervento di Francesco Sacco, in rappresentanza dell'Agenda Digitale, ci ha dato il segno di una maggiore consapevolezza anche a livello politico per quanto riguarda il processo di digitalizzazione delle amministrazioni e delle nuove relazioni in rete delle istituzioni con cittadini e imprese e dobbiamo augurarci che questo processo di fondamentale cambiamento non si arresti per l'ennesima volta.

L'Italia digitale non è impiego passivo di tecnologie, come è avvenuto sinora, ma una nuova Italia trasformata dall'integrazione intelligente delle tecnologie di rete nel suo tessuto profondo.



Il Congresso ha posto l'accento sulle competenze e sulle nuove professionalità, a cui AICA intende dare il suo pieno apporto per la costruzione di competenze ed imprenditoria digitali che sono "l'energia elettrica" del nuovo ciclo di sviluppo e di innovazione per l'Italia.

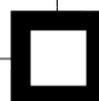
Per dare nuove speranze a tanti giovani, favorendo lo sviluppo di competenze tecnologiche e imprenditoriali assieme, per la generazione di startups e nuova imprenditorialità, per costruire risorse professionali di livello europeo ed internazionale da parte della scuola e nel lavoro attraverso le certificazioni di competenze che AICA da sempre basa su modelli europei ed ora nel nuovo contesto disegnato dall'European Competence Framework (e-cf).

Abbiamo di fronte a noi la possibilità di costruire nuove relazioni sociali, nuove forme di lavoro e di attività, di lavoro partecipato in cui ciascuno può e deve essere padrone delle sue competenze e dei suoi obiettivi di vita, attraverso nuove forme di imprenditorialità e di comunità sociali, come le smart cities dove deve prevalere la qualità di vita, il rispetto e la crescita delle persone piuttosto che l'esasperazione di una crescita puramente finanziaria-quantitativa che si accompagna sempre ad una polarizzazione della ricchezza ed un crescente divide sociale.

La costruzione di una società della conoscenza condivisa, una società che realizza in modo intelligente e creativo le opportunità offerte dalle tecnologie digitali non è una utopia, ma un percorso molto concreto e praticabile, purchè in questa Italia si comprenda e si intenda procedere senza ipocrisie e conservatismi, superando la difesa di interessi limitati e di diritti falsamente acquisiti, senza la sindrome del rinvio ed il provincialismo chiuso che non può che portare al declino in un mondo che invece intende essere sempre più aperto ed in cui le parole chiave sono open society, open innovation.

Il Congresso si è concluso in una prospettiva di costruttiva speranza da parte dei giovani protagonisti della nuova Italia digitale che sono già tra noi ed a cui dobbiamo dare fiducia ed a cui dobbiamo affiancare un efficace sistema-paese.

Bruno Lamborghini
Presidente AICA



Modelli virtuali nell'ambiente reale:

Sinergie tra stampanti 3D e scanner 3D

Luca Iuliano – Enrico Vezzetti

Il continuo ridursi della vita utile degli oggetti di uso comune ha spinto il contesto industriale a doversi dotare di soluzioni a supporto dello sviluppo prodotto sempre più efficienti in termini di tempo e costo. Oggi nelle prime fasi di “concepts”, le valutazioni relative alle prestazioni del prodotto vengono realizzate attraverso l'utilizzo di soluzioni informatiche in grado di “simulare” la realtà, ossia i fenomeni fisici ad essa correlati, piuttosto che attraverso prove su oggetti reali. La presenza di soluzioni tecnologiche, 3D scanner e stampanti 3D, in grado di creare uno stretto rapporto tra il reale e il “virtuale” rappresentano oggi proprio l'elemento chiave di questa tendenza che fornisce al “prototipo virtuale” un ruolo chiave nel processo di innovazione di prodotto.

Keywords: Virtual Prototype, Rapid Prototyping, Additive Manufacturing, 3D Scanner, Reverse Engineering

1. Introduzione

Lo sviluppo di nuovi prodotti e la ricerca di soluzioni sempre più innovative tende a sfruttare sempre più il “virtuale”, piuttosto che il “reale” grazie a ridotti tempi di risposta e a costi di utilizzo significativamente contenuti. L'attenzione a soluzioni informatiche in grado di “simulare” la realtà trova anche significativa attenzione grazie anche alla disponibilità di modelli tridimensionali interattivi, i così detti “prototipi virtuali” (Fig.1), in grado di facilitare l'interazione dell'operatore con l'oggetto rendendo queste valutazioni notevolmente intuitive.

Per mettere a frutto queste sue potenzialità il modello virtuale deve potersi integrare con l'ambiente reale in modo rapido ed efficiente. Per raggiungere questo obiettivo è quindi necessario dotarsi, non soltanto di soluzioni software 3D (CAD/CAM/CAE/...) in grado di manipolare i modelli stessi, ma di sistemi per "materializzare" l'oggetto virtuale e crearne copie fisiche da un lato, e di soluzioni in grado di tradurre gli oggetti reali in modelli virtuali per poterli elaborare ed apportare miglioramenti dall'altro. Se il passaggio dal virtuale può essere garantito dall'impiego di "stampanti tridimensionali", in grado di realizzare oggetti fisici tridimensionali con un sistema di stampa per strati, il passaggio da reale a virtuale può essere ottenuto attraverso gli "Scanner 3D" che operando set di misurazioni sono in grado di tradurre la sagoma dell'oggetto reale in un modello virtuale.

Attraverso un processo iterativo che vede più passaggi da virtuale al reale e viceversa il processo di sviluppo prodotto riesce quindi a convergere ad una soluzione ingegnerizzata e spendibile sul mercato con tempi e costi ben più contenuti rispetto a quello che si potrebbe ottenere con un approccio più tradizionale.



Figura 1
Il prototipo virtuale nello sviluppo prodotto

2. Dal virtuale al reale: le stampanti 3D

Le stampanti 3D, meglio conosciute in ambito industriale come sistemi di prototipazione rapida o macchine additive, lavorano costruendo l'oggetto per sovrapposizione di strati. Partendo dal modello "virtuale" il software di gestione della stampante 3D, una volta deciso l'orientamento dell'oggetto sulla tavola di lavoro, viene suddiviso in strati di spessore costante e specifico (Fig.2)

Ciascuno strato ha di norma uno spessore compreso tra 0.05 e 0.25 mm. Ne consegue che per creare uno spessore di un centimetro siano necessari da 80 a 200 strati. La presenza di strati così sottili permette al sistema di realizzare dettagli con elevata accuratezza (Fig.3).

Una volta che il software di gestione della stampante 3D ha suddiviso l'intero oggetto in una serie di strati a spessore costante, la stampante 3D si occuperà di sovrapporli attraverso una traslazione verticale della tavola di lavoro. Le strategie normalmente utilizzate per la costruzione degli strati sono due: raster oppure vettoriale. Nel caso della logica raster il software di gestione della stampante 3D gestisce ogni strato come un'immagine scomposta in pixel. I pixel nei quali sono presenti parti solide dell'oggetto saranno codificati come "accessi", mentre quelli in cui ci sono parti cave verranno identificati come "spenti". Attraverso questa logica, acceso o spento, il sistema sarà in grado di comprendere quali zone devono essere "solidificate" e quali invece dovranno restare disaggregate perché

successivamente al processo di costruzione degli strati dovranno essere ripulite, poiché nell'oggetto finale è prevista una parte cava.

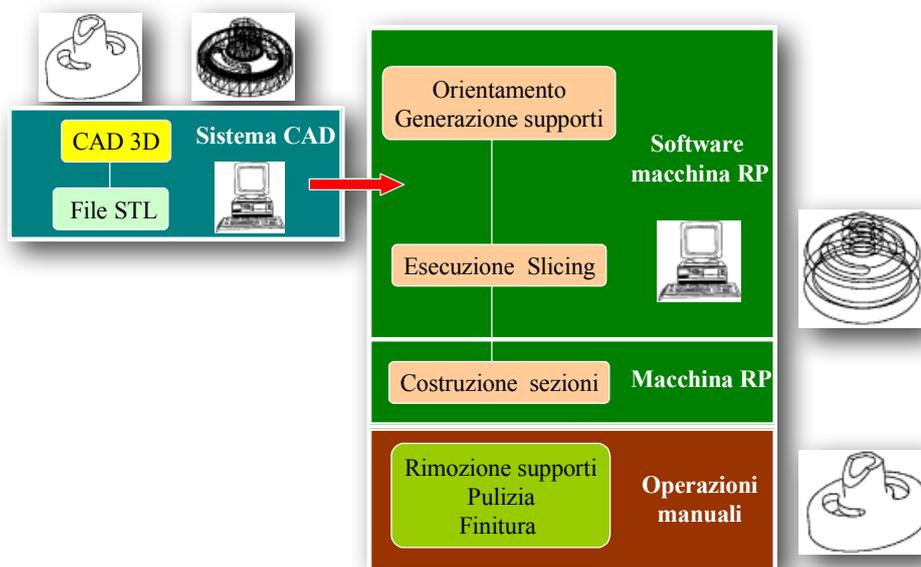


Figura 2
Il processo di Additive Manufacturing

Oltre alla logica raster è possibile adottare una logica vettoriale in cui i bordi dell'oggetto vengono convertiti in comandi di movimento della stampante 3D interpretabili dall'unità di controllo. A differenza del sistema raster nella logica vettoriale il sistema oltre ad eseguire i comandi di movimento in corrispondenza dei bordi, procede nel creare una fitta trama tra i bordi in modo tale da realizzare le porzioni dell'oggetto che non presentano cavità ma continuità di materiale.

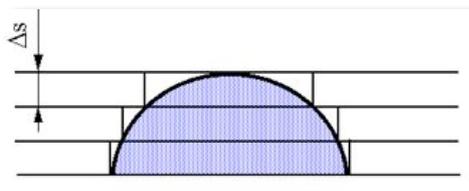


Figura 3
La suddivisione in strati

Indipendentemente dalla logica di costruzione adottata il processo di stampa tridimensionale si conclude con una fase di post trattamento che, a seconda della tipologia di processo utilizzato, potrà consistere in una semplice rimozione del materiale di costruzione non aggregato, o una finitura superficiale per rimuovere eventuali imperfezioni dell'oggetto finale, fino ad operazioni di stabilizzazione in forno oppure attraverso l'utilizzo di additivi specifici in grado di rendere l'oggetto finale più resistente o esteticamente più accattivante. Questa ultima fase dipenderà fortemente dalla tipologia di sistema di costruzione adottato.

2.1. Tecnologie disponibili

Differenti possono essere i processi con cui il sistema costruisce l'oggetto solido. Una delle più conosciute e semplici suddivisioni vede il materiale di costruzione quale elemento di distinzione. In particolar modo il materiale impiegato nella costruzione degli strati può essere allo stato solido oppure liquido. Per quanto riguarda poi il materiale allo stato solido, questo si può trovare in forma di filamenti oppure in forma di polvere. Per ognuno di questi materiali sarà poi possibile disporre di altrettante differenti soluzioni a seconda della strategia specifica di costruzione e delle peculiarità dello specifico materiale utilizzato.

2.1.1. L'utilizzo di materiali allo stato liquido

Quando si parla di materiale allo stato liquido il primo riferimento va alla Stereolitografia (SLA, abbreviazione di stereolithographic apparatus) che rappresenta anche storicamente il primo processo, venne presentato nel 1987, di costruzione di modelli fisici attraverso una logica additiva per strati. Questo sistema infatti, lavorando in presenza di una vasca di fotopolimero in forma liquida, materiale in grado di solidificare se colpito da una specifica radiazione luminosa, attraverso l'utilizzo di una sorgente laser allo stato solido (Nd:YVO4 con potenza fino a 216 mW), procede alla solidificazione degli strati facendo focalizzare selettivamente la radiazione luminosa attraverso una serie di ottiche, sullo strato in fase di costruzione (Fig.4).

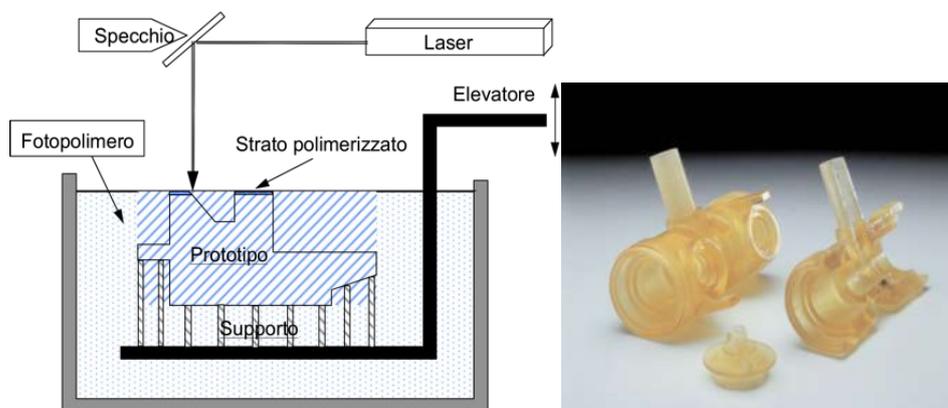


Figura 4
Stereolitografia (SLA): Il sistema ed un esempio di prototipo

L'interazione della radiazione laser con il fotopolimero innesca una reazione chimica che ha come effetto la solidificazione di quest'ultimo. Il movimento nel piano del fascio laser consente la realizzazione della prima sezione del pezzo su di una piastra di acciaio forata denominata elevatore. Una volta che la solidificazione delle pareti dell'oggetto nello specifico strato è completata la tavola di lavoro, quindi l'elevatore, viene abbassato, ed inizia la fase di ricoprimento al fine di avere un film di liquido sopra la sezione appena costruita, in modo da disporre di uno strato controllato di materiale liquido sul quale verrà nuovamente focalizzata la sorgente luminosa per la costruzione di un nuovo strato. Ottenuto il corretto livellamento, il processo riprende con la costruzione di

uno strato successivo che grazie alla presenza di materiale liquido tra i due strati, garantisce sufficiente coesione durante l'intero processo di costruzione.

Quando tutti gli strati saranno stati realizzati il modello fisico realizzato verrà estratto dalla vasca per mezzo dell'elevatore. E' chiaro che per ragioni di tempo il laser non può solidificare integralmente la sezione ma si limiterà al suo profilo e ad un certo numero di linee che congiungono il perimetro interno con quello esterno. Al termine di questa fase la parte prodotta (green part) contiene ancora del liquido intrappolato all'interno, inoltre le sue pareti non sono completamente polimerizzate: questi due fattori pregiudicano la consistenza meccanica del prototipo che dovrà essere migliorata con un ulteriore trattamento agli ultravioletti. Una volta completato anche il "post-processamento" potrà essere necessario eliminare gli elementi di supporto e effettuare qualche operazione di finitura manuale dell'oggetto.

Evoluzione della tradizionale SLA può essere considerato il sistema PoliJet che integra nel principio della foto polimerizzazione l'impiego del principio della stampa a getto (Fig.5).

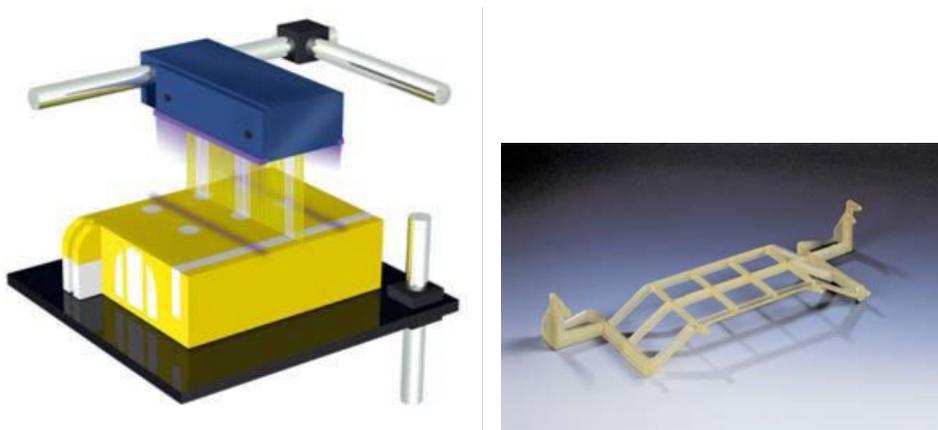


Figura 5

PoliJet (PJ): Il sistema di stampa ed un esempio di prototipo

Attraverso l'utilizzo di una testina con 1536 ugelli il sistema è in grado di depositare in modo selettivo il fotopolimero sulla tavola di lavoro al fine di costruire una singola sezione. Anziché dover disporre di una intera vasca di fotopolimero, che ovviamente può essere soggetta ad un processo di invecchiamento, questo processo sfrutta le potenzialità del materiale depositandolo solo selettivamente quando e dove serve. Il processo di solidificazione anziché richiedere un più dispendioso laser viene effettuato attraverso l'impiego di una lampada ad ultravioletti posizionata dietro la testa di stampa.

2.1.2. L'utilizzo di materiali allo stato solido

Partendo dalle tecnologie che utilizzano materiale in polvere l'attenzione si può focalizzare sulla Selective Laser Sintering (SLS). Questa tecnologia, che sebbene abbia all'interno del proprio nome il termine "Sintering" in realtà rappresenta un sistema che utilizza la fusione selettiva delle polveri, piuttosto che la sinterizzazione (Fig.6).

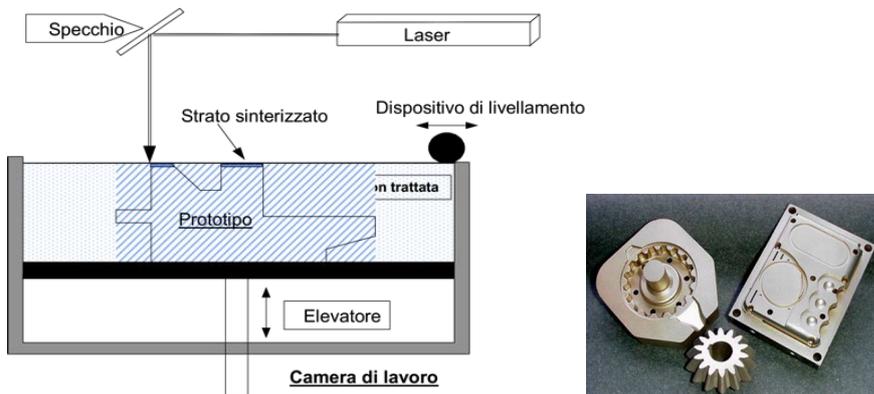


Figura 6
Sinterizzazione Selettiva (SLS): il sistema ed un esempio di prototipo metallico

La costruzione dei singoli strati inizia con la deposizione e livellamento di uno strato di polvere su cui viene focalizzato successivamente un fascio laser. In questo caso, trattandosi di una fusione selettiva delle polveri, la sorgente laser utilizzata deve avere una elevata potenza e per questo motivo si adotta un laser CO₂ con potenza variabile tra 25 e 200 W. Per supportare il processo di fusione ed evitare di un eccessivo dispendio nell'utilizzo della sorgente laser, la camera di lavoro è mantenuta ad una temperatura prossima a quella di fusione. Una volta realizzato lo strato, l'elevatore verrà abbassato e si riprenderà con la deposizione e pressatura di un nuovo strato di polvere sul quale fondere selettivamente le porzioni di sezione che corrispondono a parti piene dell'oggetto. Le porzioni di strato di polvere depositate che non interagiscono con la sorgente laser restando disaggregate avranno la funzione di supporto.

Per quanto riguarda le polveri metalliche, oltre ai sistemi che fanno parte della tecnologia Selective Laser Sintering, negli ultimi anni sono stati introdotti sul mercato alcune soluzioni in grado di aggregare attraverso fusione, sempre con l'utilizzo di una sorgente laser, le particelle di polvere metalliche senza aver bisogno di ricoprirle con materiale polimerico. Questi sistemi, noti sotto il nome di Selective Laser Melting, sono in grado di realizzare prototipi con caratteristiche strutturali superiori a quelli ottenibili con i sistemi SLS.

Spostandosi poi su sistemi che, benché utilizzino materiale in forma di polvere, non utilizzano più una sorgente laser, ma piuttosto una testina di stampa che deposita selettivamente collante, si arriva ai sistemi 3DPrinting (Fig.7). Attraverso lo stesso principio utilizzato dalle stampanti a getto di inchiostro il sistema, una volta che uno strato di polvere, normalmente gesso, viene depositato e livellato sulla tavola delle elevatore, la testina spruzza in modo selettivo il collante in quelle zone del piano di lavoro che corrispondono ad una porzione piena della sezione dell'oggetto in fase di costruzione.

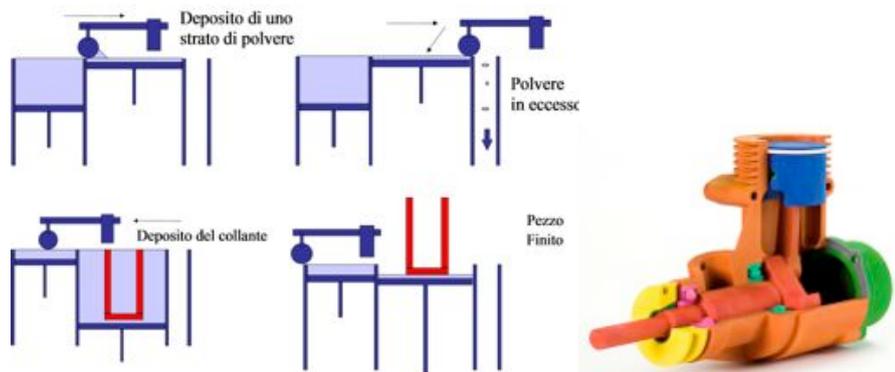


Figura 7
3D Printing (3DP): principio di stampa e fasi del processo e un esempio di prototipo

Sempre rimanendo nell'ambito del materiale allo stato solido, ma passando dalla polvere ai filamenti, si passa al processo FDM (Fused Deposition Modeling) (Fig.8).

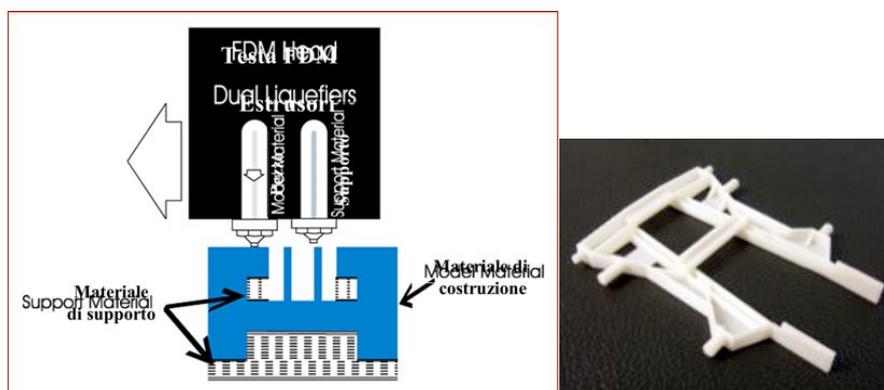


Figura 8
Fused Deposition Modeling (FDM): il sistema ed un esempio di prototipo

Attraverso una testa di estrusione, portata in temperatura, che si muove parallelamente alla tavola di lavoro il filo, portato alla temperatura di fusione, costruisce le pareti dell'oggetto. Allo stesso modo del processo di stereolitografia anche l'FDM deve prevedere la costruzione di un sistema di nervature che colleghino il profilo esterno a quello interno per migliorare la resistenza meccanica del prototipo. In presenza di parti sporgenti l'FDM deve prevedere la costruzione di supporti. Questi vengono realizzati attraverso la presenza di una seconda testina che a complemento di quella principale estrude un secondo materiale, che successivamente alla costruzione completa del prototipo verrà eliminato. Le temperature della testa di estrusione e della camera di lavoro sono parametri fondamentali: il materiale deve essere mantenuto ad una temperatura appena al di sopra del punto di plastificazione, pena la formazione di gocce che riducono la precisione del prototipo.

Per garantire elevate prestazioni, sia in termini di precisione che in termini di velocità la testa di estrusione viene fatta muovere attraverso l'utilizzo di un campo magnetico su di un cuscino d'aria in pressione.

2.1. Applicazioni

Per entrare nel merito delle possibili applicazioni in cui le stampanti 3D possono essere impiegate è necessario effettuare una prima classificazione dei prototipi che questi sistemi sono in grado di realizzare. Partendo dai prototipi estetici (Fig. 9), che vengono utilizzati per avere un primo riscontro visivo del progetto, si può poi passare ai prototipi funzionali (Fig.10) che prevedono l'utilizzo delle stampanti 3D per la realizzazione di modelli da destinare a test di collaudo, quali ad esempio prove fluidodinamiche, di resistenza, di corretto accoppiamento fra più parti e di assemblaggio.



Figura 9
Prototipo concettuale

Con riferimento poi a contesti più specifici del mondo della produzione industriale si parla non soltanto di prototipi "rapidi", ma anche di attrezzaggi rapidi ed anziché utilizzare il termine più conosciuto di *Rapid Prototyping* (Prototipazione Rapida) si utilizza quello di *Rapid Tooling*, in cui, come testè accennato, l'utilizzo delle stampati 3D ha lo scopo di realizzare attrezzature produttive con notevoli risparmi di tempo. Proprio nell'ambito del *Rapid Tooling* gli investimenti da parte dei costruttori di sistemi stanno crescendo sempre più. Questo scenario trova semplice giustificazione nel fatto che la realizzazione di una attrezzature di produzione, uno stampo per esempio, rappresenta per un'azienda un notevole investimento, sia in termini di tempo che di costo.



Figura 10
Prototipo funzionale: prove di assemblaggio

Si pensi che le attrezzature di produzione possono arrivare a coprire circa l'80% del costo di sviluppo del prodotto. I lunghi tempi di realizzazione, oltre ai costi davvero significativi, fanno capire come mai la possibilità di poter disporre di tecnologie in grado di realizzare attrezzature di prova, le così dette attrezzature di preserie, se non addirittura le attrezzature definitive, con maggiore flessibilità e con tempi notevolmente più ridotti, possa rappresentare un significativo vantaggio in termini di competitività.

Focalizzando l'attenzione sulle peculiarità dei singoli sistemi e partendo dalla SLA tra i materiali di costruzione più utilizzati è possibile citare le seguenti resine: acriliche, con bassa viscosità e destinate alla produzione di modelli estetici o repliche anatomiche dove il tempo di costruzione è da privilegiare rispetto alla precisione del pezzo, oppure epossidiche, dotate di elevata viscosità da impiegare laddove risulti fondamentale minimizzare gli errori geometrici per effettuare prove funzionali.

Spostandoci poi al processo SLS si può affermare che questo sistema è in grado di utilizzare differenti tipologie di materiali (termoplastici, cera, metalli, sabbia) e per questo motivo permette la realizzazione di prototipi adatti a prove di funzionamento poiché gli oggetti che vengono realizzati hanno caratteristiche simili a quelli realizzati con tecnologie convenzionali. In presenza di polveri di materiale polimerico il processo di "fusione selettiva" agisce direttamente sul polvere polimerica (nylon,...), mentre in presenza di materiali metallici o materiali ceramici il sistema opera con un polveri caratterizzate da un'anima metallica/ceramica ed un involucro in materiale termoplastico. In questo caso il processo di fusione selettiva intacca lo strato esterno di materiale polimerico portandolo a fusione e permettendo l'aggregazione delle particelle ceramiche/metalliche. In questo modo si dispone di un oggetto 20% materiale polimerico ed 80% metallico/ceramico. Alla fase di costruzione del prototipo segue una fase di post processamento nella quale il modello testè costruito viene posta all'interno di un forno ad alta temperatura ed attraverso un processo di osmosi il materiale termoplastico cede il posto al bronzo. In questo modo le caratteristiche strutturali del prototipo saranno più interessanti rispetto a quelle del prototipo prima del post trattamento.

I sistemi FDM poi impiegano per la costruzione dei prototipi ABS, Policarbonato e Polifenilsulfone. L'utilizzo di questi materiali polimerici fornisce sufficiente robustezza e consistenza agli oggetti realizzati permettendone l'impiego per prove funzionali. Per ciascun materiale impiegato nella costruzione del prototipo ne dovrà essere previsto

un secondo con proprietà termico meccaniche leggermente inferiori impiegato per la costruzione dei supporti.

Nel caso poi del processo 3DP, che utilizza il gesso come materiale di costruzione, le ridotte proprietà meccaniche del prototipo, unitamente ai costi di costruzione più contenuti ne permettono un utilizzo a pieno titolo per valutazioni concettuali ma non per prove funzionali (Tab.1).

Prototipi Concettuali	<i>3DP,SLA,SLS,FDM</i>
Prototipi Funzionali	<i>SLA, SLS, FDM</i>
Rapid Tooling	<i>SLA, SLS, FDM</i>

Tabella 1

Classificazione delle differenti tecnologie in funzione alla tipologia di prototipo realizzato

2.3. Tendenze

Sebbene l'utilizzo delle stampanti tridimensionali in ambito professionale o industriale abbia visto negli ultimi anni una marcata crescita, grazie all'utilizzo di materiali sempre più innovativi in grado di coprire un maggiore spettro di applicazioni, negli ultimi tempi qualcosa sta cambiando. Piuttosto che procedere unicamente nello sviluppo di soluzioni sempre più innovative e performanti per il solo mercato professionale, molti "vendors" si stanno spostando anche sullo sviluppo di sistemi a basso costo per applicazioni più "domestiche". Si pensi per esempio al progetto RepRap (*acronimo di Self REPLICating RAPid prototyping*) (Fig.11) che fornisce un sistema per la stampa tridimensionale ad un costo di circa 700 – 800 dollari, se ancora da assemblare, oppure di circa 1000 dollari, se già assemblato.

Oltre al costo ridotto questo sistema risulta essere molto interessante per la sua logica "Open". Attraverso l'utilizzo della piattaforma "Arduino" il progetto intende stimolare un processo di sviluppo e miglioramento continuo del sistema e delle sue possibili applicazioni in una logica di condivisione della conoscenza. Adottando il principio dell'estrusione di filamenti di materiale termoplastico (FDM), il sistema è in grado di raccogliere un significativo consenso anche nell'uso "domestico" grazie alla possibilità di sfruttare prodotti provenienti dalla raccolta differenziata. Lavorando infatti su sistemi che a partire dai prodotti della raccolta differenziata sono in grado di realizzare i filamenti necessari per il processo di stampa è possibile pensare ad un auto sostentamento dell'intero processo di stampa, quindi ad un ulteriore contenimento dei costi, oltre all'investimento iniziale.

Tutto questo si va a collocare nel nuovo scenario che ha come elemento fondante "la personal fabrication", ossia il processo di ideazione e stampa 3D dei prodotti su piccolissima scala. Se ora ci troviamo di fronte ad uno scenario nel quale si parla di "mass customisation" e quindi di produzione di massa, questa nuova tendenza richiama invece uno scenario con una dimensione molto più contenuta, quasi "artigianale".

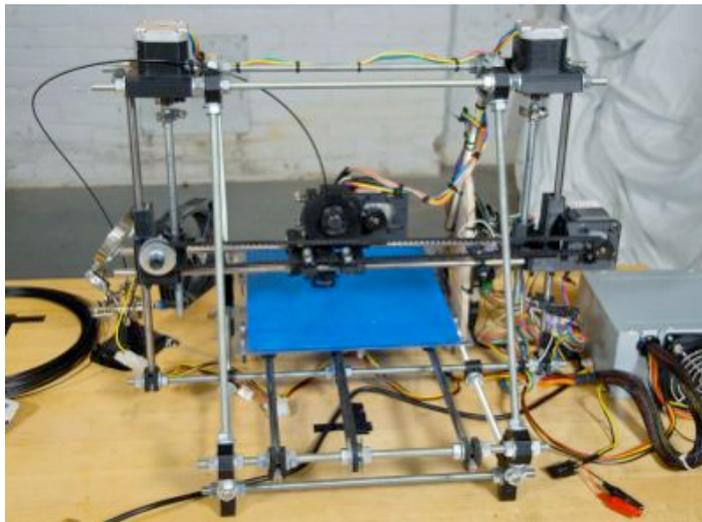


Figura 11
Il Sistema Reprap ed un esempio di prototipo

In un contesto di questo tipo le stampanti 3D possono essere utilizzate anche nella manutenzione domestica. Qualora per esempio fosse necessario sostituire il manico della macchinetta del caffè poiché rotto, sarebbe pensabile stamparselo autonomamente a casa. Qualora la persona non abbia competenze nell'ambito della modellazione 3D è possibile trovare in rete un numero sempre in crescita di modelli digitali 3D pronti per essere "materializzati" attraverso il processo di costruzione per strati.

3. Dal reale al virtuale: gli scanner 3D

Gli scanner 3D, meglio conosciuti in ambito industriale come sistemi di reverse engineering, od anche di "ingegneria inversa", lavorano rilevando la superficie esterna di un oggetto. Questa misurazione può normalmente essere realizzata attraverso l'utilizzo di specifiche teste di misura elettromeccaniche che a contatto con l'oggetto ne registrano la posizione rispetto ad una posizione di riferimento.

In alternativa si parla di sistemi senza contatto quando invece di una testa di scansione si prevede la presenza di una coppia sorgente/sensore, normalmente ottici, che attraverso la proiezione sulla superficie esterna dell'oggetto di una radiazione laser o di un "pattern luminoso" rilevano la superficie dell'oggetto analizzando il segnale riflesso.

In realtà, dovendo disporre, come risultato finale del processo di acquisizione, di un modello 3D consistente, l'acquisizione dello scanner 3D deve essere successivamente elaborata attraverso l'utilizzo di opportuni strumenti e strategie di elaborazione. Dallo Scanner 3D infatti si riesce solo ad ottenere una nuvola di punti (Fig.12), che contiene un set di coordinate tridimensionali (x,y,z) appartenenti alla superficie esterna dell'oggetto acquisito.

Normalmente, tra le altre cose, essendo la forma dell'oggetto complessa, non ci si trova di fronte ad una unica acquisizione, ma piuttosto ad un numero maggiore od uguale a due nuvole di punti, in funzione al livello di complessità dell'oggetto. Qualora infatti l'oggetto sia caratterizzato da una forma con presenza di un significativo numero di dettagli, distribuiti sull'intera sagoma dello stesso, sarà

necessario effettuare più prese. In pratica sarà necessario posizionare lo scanner 3D a più angolazioni differenti in modo da acquisire la totalità della sagoma dell'oggetto. Questo fa anche comprendere, che oltre ad effettuare una oculata scelta del sistema di acquisizione in funzione della specifica applicazione, è anche necessario effettuare una efficiente pianificazione delle acquisizioni per riuscire a catturare l'intera sagoma dell'oggetto. Se da un lato la presenza di più acquisizioni garantisce un più accurato rilievo di tutto l'oggetto, dall'altra comporta alcuni passaggi di elaborazione aggiuntivi prima di arrivare al modello digitale finale. Dovendo ottenere un unico modello tridimensionale dell'oggetto sarà necessario effettuare un opportuno allineamento delle differenti nuvole di punti ottenute nelle differenti fasi di acquisizione. Muovendo lo scanner 3D rispetto all'oggetto, ciascuna nuvola, sebbene sia una porzione dello stesso oggetto, non disponendo normalmente di una unica origine comune, deve subire un processo di "roto-traslazione", quindi una trasformazione geometrica nello spazio, per sovrapporre in una unica le differenti origini delle differenti nuvole di punti.



Figura 12
Il prototipo virtuale nello sviluppo prodotto

In questa fase, anche detta di "pre-processamento" oltre a questa attività di unione delle differenti acquisizioni si prevedono normalmente operazioni di filtraggio e pulitura della nuvola di punti da eventuali informazioni che possono essere dovute a fenomeni di acquisizione fuori controllo (spikes,...). Una volta in presenza di una nuvola di punti ripulita dal rumore, normalmente dovuto alle condizioni ambientali durante la fase di acquisizione, il modello può essere convertito in una "mesh", normalmente triangolare, attraverso la costruzione di poligoni quali unione dei differenti punti che caratterizzano la nuvola stessa. Questa fase oltre a supportare una più facile visualizzazione della sagoma dell'oggetto acquisito, quindi mostrare eventuali anomalie rispetto all'oggetto reale, grazie ad una serie di algoritmi di gestione dei poligoni, è in grado di supportare la chiusura di eventuali cavità anomale dando la possibilità di poter disporre di un primo modello, seppur approssimato da superficie poligonali piane, dell'oggetto acquisito (Fig.13).

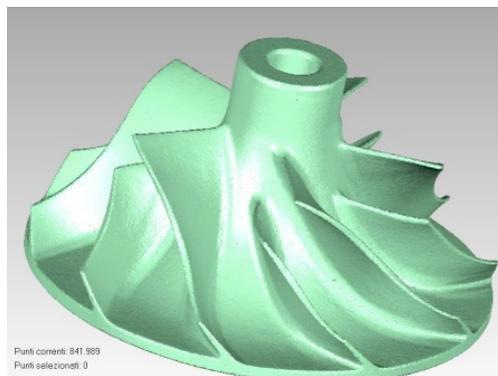


Figura 13
Un esempio di modello poligonale (mesh)

Sebbene il modello così ottenuto possa essere sufficiente per realizzare copie dello stesso oggetto attraverso l'utilizzo di soluzioni di stampa 3D, se è necessario effettuare integrazioni, modifiche sull'oggetto virtuale è necessario procedere nel processo di elaborazione.

Dovendo fruire, quale elemento finale, di superfici matematiche gestibili all'interno dei più comuni modellatori CAD 3D, è necessario implementare una prima fase di segmentazione, in cui una fitta maglia di curve guida vengono generate sul modello poligonale testè realizzato al fine di supportare la successiva costruzione delle superfici.

Una volta realizzata tutta l'intelaiatura delle curve guida nelle differenti regioni della superficie esterna dell'oggetto, attraverso opportuni algoritmi di "fitting" sarà poi possibile ottenere il modello tridimensionale completo e consistente (Fig.14)

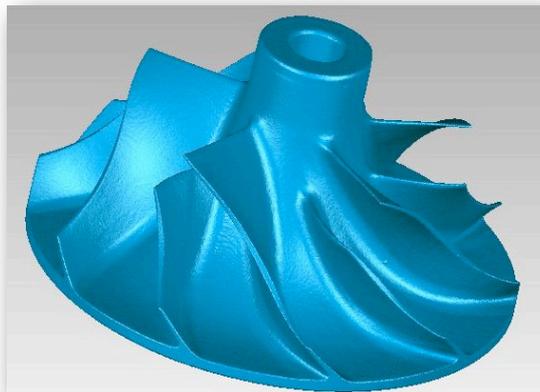


Figura 14
Un esempio di modello CAD completo

3.1. Tecnologie disponibili

Gli scanner 3D attualmente sul mercato hanno raggiunto negli ultimi anni un'elevata numerosità. Molti costruttori di sistemi di misura o di sistemi ottici hanno percepito l'opportunità di sfruttare le loro competenze metrologiche o ottiche per proporre sul mercato sistemi completi in grado di acquisire oggetti dalle forme complesse.

Sebbene le soluzioni disponibili sul mercato siano molteplici si può notare che, a parte qualche variante sulle metodiche di controllo o sulla strategia di acquisizione, tutti questi sistemi sfruttano principi fisici analoghi che ne permettono quindi una collocazione selettiva in due possibili macro famiglie: sistemi a contatto e sistemi senza contatto.

3.1.1. I sistemi a contatto

La famiglia dei sistemi a contatto raccoglie tutti quegli strumenti che utilizzano un contatto fisico di una testa di scansione con la superficie dell'oggetto. La prima famiglia di soluzioni che entrano a far parte di questa categoria sono le Macchine di Misura a Coordinate (CMM). Le Macchine di misura a coordinate sono caratterizzate da una struttura rigida in grado di fornire un'elevata precisione ed accuratezza durante l'acquisizione. Questi sistemi rilevano i loro dati toccando il pezzo mediante un tastatore elettromeccanico e rilevando i punti sulla superficie dell'oggetto grazie ad un tastatore che si flette fino a causare l'apertura del circuito elettrico di controllo durante il contatto con la superficie esterna dell'oggetto (Fig.15). Registrando le posizioni degli assi macchina negli istanti in cui il tastatore flette e chiude il circuito elettromeccanico, il sistema è in grado di acquisire la forma dell'oggetto.



Figura 15
Schema di della testa di scansione elettromeccanica

Oltre alle macchine di misura a coordinate la famiglia dei sistemi a contatto vede anche la presenza di sistemi che simulano il comportamento del braccio umano. Questi sistemi, noti con il termine di braccetti antropomorfi, operano secondo il principio del classico pantografo. In questi sistemi un tastatore è tenuto a contatto con la superficie da acquisire. Montato su di un braccetto antropomorfo il sistema rileva la sagoma dell'oggetto attraverso l'utilizzo di una serie di trasduttori, localizzati in corrispondenza dei giunti del braccio robotizzato (Fig.16). Secondo questo principio vengono valutati i movimenti della testa di scansione nello spazio cartesiano codificando le rilevazioni dei trasduttori di posizione collocati sui giunti. Il sistema ha un funzionamento manuale poiché la testa di scansione viene guidata dalla mano di un operatore sulla superficie del pezzo. L'operatore deve guidare la testa di acquisizione dell'oggetto, ma garantisce, rispetto agli altri strumenti di misura manuali, una maggiore uniformità e completezza della nuvola di punti grazie alla presenza di una logica di acquisizione pilotata via software che permette di impostare l'acquisizione secondo un passo di acquisizione costante, oppure secondo intervalli di tempo prefissati. La configurazione del numero di articolazioni determinerà la capacità di tastare della sonda.



Figura 16
Struttura di un braccetto antropomorfo

3.1.2. I sistemi senza contatto

La famiglia dei sistemi senza contatto contiene soluzioni in cui l'interazione con la superficie avviene impiegando una sorgente, luminosa, acustica o elettromagnetica. In ogni caso non si può parlare di un vero e proprio contatto, ma piuttosto di una riflessione della propagazione ottica, acustica od elettromagnetica in un mezzo e la conseguente riflessione a valle del contatto con la superficie dell'oggetto stesso. Proprio la differente tipologia di onda, elettromagnetica, acustica od ottica, porta ad una differente classificazione del sistema di acquisizione. Oltre a questa differenziazione, focalizzando in particolare l'attenzione sui sistemi ottici, il differente principio di misura adottato (tempo di volo, triangolazione,...) permette di fruire di sistemi differenti con differenti prestazioni a seconda dell'applicazione in cui vengono impiegati.

Partendo dalle sorgenti laser si può parlare di laser scanner (Fig.17), ossia sistemi ottici basati sul principio della "triangolazione ottica", oppure di laser radar quando la misurazione avviene attraverso la valutazione del tempo di volo di una sorgente laser o di olografia conoscopica quando si impiegano sistemi ottici basati sul principio dell'interferometria di una sorgente laser.

Sempre restando nei sistemi ottici, ma abbandonando le sorgenti laser per spostarsi sulla proiezione di "pattern luminosi" è possibile parlare dei sistemi a luce strutturata, ossia di sistemi che proiettano una griglia luminosa strutturata (Fig.18), e misura la distanza della sorgente dall'oggetto attraverso la valutazione della distorsione del pattern luminoso proiettato sull'oggetto.

Oltre ai sistemi ottici, come testè citato, si possono annoverare tra i sistemi di acquisizione senza contatto anche sistemi che sfruttando onde acustiche od anche sistemi che utilizzano un tracciamento elettromagnetico.

Oltre ai sistemi attivi, ossia sistemi che utilizzano la proiezione di un qualche fenomeno fisico (acustico, ottico, magnetico,...) ed attraverso l'analisi della sua riflessione rilevano ed estraggono una sagoma tridimensionale, esistono i sistemi passivi.



Figura 17

Principio della triangolazione laser ed un esempio di sistema

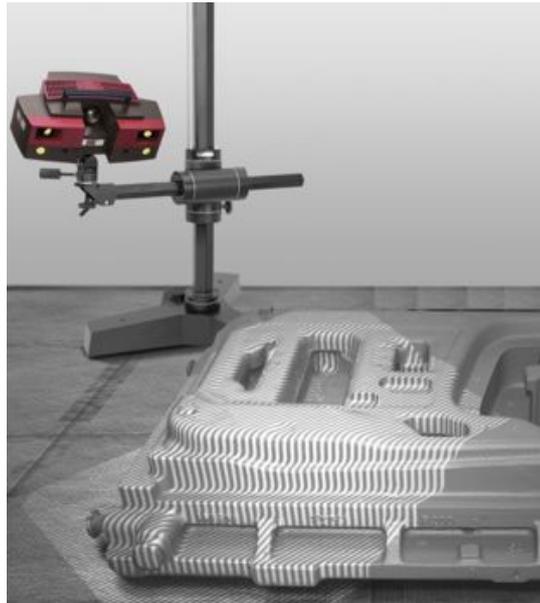


Figura 18

Un esempio di acquisizione con l'utilizzo di un sistema a luce strutturata

Senza l'utilizzo di alcuna proiezione, ma semplicemente attraverso l'acquisizione di più fotografie/immagini dello stesso oggetto, da punti di osservazione complementari, la fotogrammetria è in grado di costruire un modello tridimensionale (Fig.19).

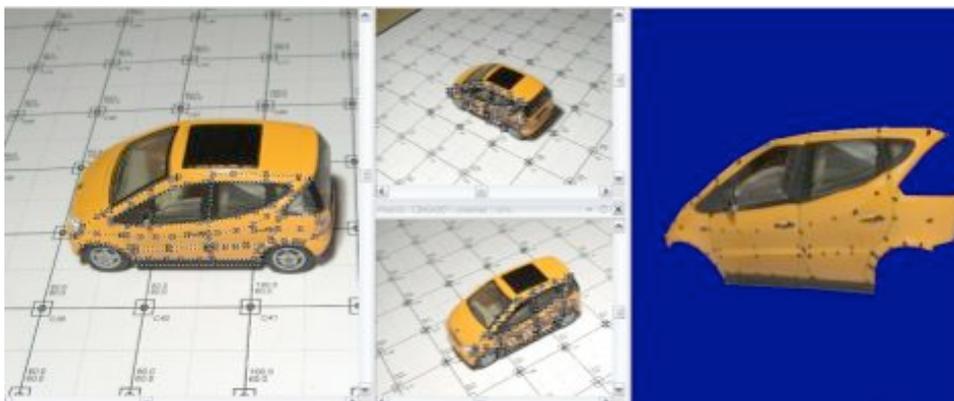


Figura 19

Un esempio di acquisizione con l'utilizzo di fotogrammetria

3.2. Applicazioni

L'applicazione di tecniche di acquisizione 3D si è estesa negli anni a molti settori. Lo sviluppo di soluzioni sempre più performanti e il ruolo sempre più importante dei modelli virtuali 3D, in differenti settori di applicazione, ha portato ad un crescente interesse verso questo processo da parte di differenti settori. In

particolar modo, negli ultimi anni alcune applicazioni hanno mostrato un maggiore sviluppo nei settori industriali anche se l'utilizzo di scanner 3d sta prendendo sempre più piede in settori nel dominio dell'archeologia, architettura e dell'entertainment.

Nell'ambito del settore industriale l'acquisizione 3D vede la sua presenza da maggior tempo e per questa ragione le sue applicazioni sono molte e diversificate. Partendo dalla realizzazione di protesi ortopediche, attrezzature sportive o militari l'attenzione alla forma rappresenta un elemento fondamentale per la riuscita del prodotto. L'ergonomia dello stesso deve essere curata accuratamente in modo da rispondere opportunamente alle esigenze del soggetto che deve utilizzarle. Per questa ragione lo sviluppo di questi prodotti può avvenire unicamente partendo dalla sagoma della parte del corpo umano con cui deve essere accoppiata. Questa informazione deve essere quanto più accurata possibile al fine di garantire che il prodotto realizzato sia in grado di garantire le prestazioni necessarie per cui è stato sviluppato (Fig.20)

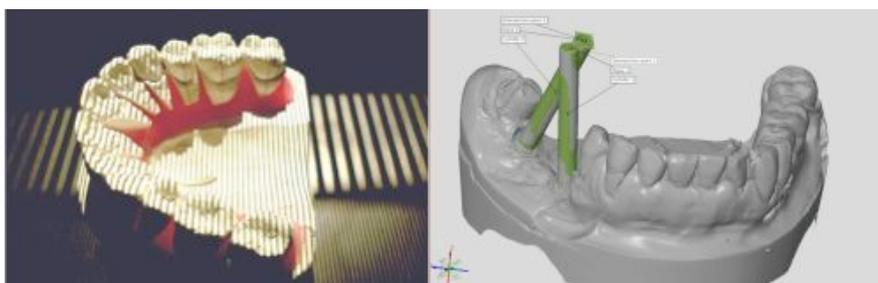


Figura 20

Un esempio di rilievo di protesi anatomica

L'acquisizione 3D gioca un ruolo altrettanto importante in presenza di particolari meccanici per i quali non sia disponibili modelli digitali tridimensionali. Questo accade in particolar modo per quanto riguarda oggetti storici, come per il caso delle auto storiche, in cui spesso e volentieri risulta necessario sostituire parti difettose e danneggiate. In questi casi non potendo disporre di ricambi, né di modelli digitali o altra documentazione progettuale, l'unica possibile soluzione è quella di poter effettuare un rilievo tridimensionale (Fig.21).



Figura 21

Un esempio di rilievo di particolare meccanico di auto storica

3.3. Tendenze

Negli ultimi anni si sta sviluppando una nuova tendenza nello sviluppo degli scanner 3D che vede il disaccoppiamento tra il sistema di acquisizione, o per meglio dire la testa di acquisizione, e il sistema di riferimento, o per meglio dire il sistema di tracciamento/misura. In pratica il sistema è costituito da due entità di misura separate, con tecnologie differenti, una delle quali ha lo scopo di acquisire la sagoma dell'oggetto, con tecnologia a contatto o laser, mentre la seconda unità, ossia il tracciatore ha il compito invece di tenere traccia della posizione della testa di scansione nello spazio. Questa tipologia di sistema, che nasce un po' dalla filosofia dei braccetti antropomorfi, ma che ha il vantaggio di non avere l'ingombro e le limitazioni del braccetto stesso, ha iniziato a prendere piede negli ultimi tempi grazie alla flessibilità e modularità che in grado di garantire.

La soluzione che è stata universalmente identificata dalla maggior parte dei costruttori è stata quella di prevedere due unità. La prima, ossia la testa di scansione che può prevedere un sistema a contatto oppure un sistema senza contatto, normalmente un sistema ottico che sfrutta la triangolazione laser, e poi un tracciatore, naturalmente senza contatto, in grado di mantenere traccia della posizione della testa di scansione nello spazio. Attraverso questo approccio è possibile per esempio utilizzare differenti tecnologie di scansione semplicemente utilizzando due differenti teste di scansione ma mantenendo lo stesso tracciatore. Mentre come si è detto le teste di scansione sfruttando normalmente una tecnologia di misura a contatto, la stessa dei braccetti antropomorfi, oppure sistemi a triangolazione laser, vista la vicinanza tra la testa di scansione e l'oggetto, il sistema di tracciamento può adottare più soluzioni di tracciamento.

La presenza del sistema di tracciamento ottico rende il sistema più flessibile ed in grado di operare su spazi maggiori. Inoltre il disaccoppiamento del sistema di tracciamento, per il posizionamento, dalla testa di scansione, permette di poter lavorare su ogni genere di prodotto, scavalcando le limitazioni di una o dell'altra tecnologia e fornendo al potenziale utente tutta la flessibilità che richiede, attraverso la sola sostituzione della testa di scansione e non dell'intero sistema (Fig.22)



Figura 22

*Testa di scansione con sistema di tracciamento ottico:
in alto, testa di scansione laser, in basso, con tastatore*

Una seconda soluzione anziché adottare il tracciamento ottico opera attraverso il tracciamento magnetico. Anche in questo caso i sensori magnetici vengono posizionati sulle teste di acquisizioni manuale forniscono posizione e orientamento dello scanner rispetto ad una unità base di rilevazione. Questi sistemi sono meno sensibili, rispetto a quelli ottici, delle condizioni ambientali, ma sono influenzati da qualsiasi fonte di campo magnetico, nonché da qualsiasi materiale ferromagnetico.

Un'ultima soluzione (Fig.23) che utilizza una testa di scansione, ma che ha abbandonato l'impiego del tracciatore prevede la presenza di una sola testa di scansione nella quale sono collocati un sistema di visione, che valuta la posizione della testa di scansione ed un sistema di acquisizione laser che effettua la scansione.

Attraverso la presenza di una serie di riflettori ottici posizionati direttamente sull'oggetto da acquisire, il sistema di visione è in grado di registrare la posizione della testa di scansione rispetto al sistema di riferimento pezzo, mentre il sistema laser ne acquisisce la sagoma esterna.

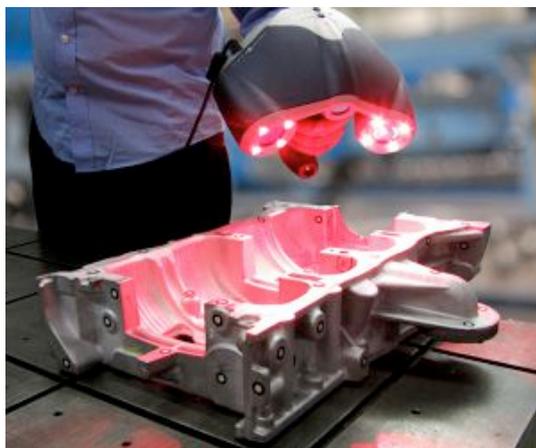


Figura 23

Un esempio di acquisizione senza braccio di misura nel sistema di tracciamento

Questa metodologia si presta bene a rilievi effettuati anche in ambienti caratterizzati da condizioni ambientali poco favorevoli (vibrazioni,...), visto che il sistema di riferimento è posizionato direttamente sull'oggetto da acquisire.

4. Sinergie tra stampanti 3D e scanner 3D

In presenza di oggetti di forma complessa, oggetti di stile, parti anatomiche,... l'integrazione tra gli scanner 3D e le stampanti 3D trova la sua massima efficacia. Si pensi ad esempio quando ci si trova in presenza di parti difettose o danneggiate di veicoli storici di cui non si dispone di alcun documento tecnico di progetto. In questo caso il processo di realizzazione del così detto "ricambio" diventa un processo lungo e laborioso che solo attraverso una serie di tentativi, ossia attraverso un processo di "aggiustaggio" può raggiungere la soluzione. Potendo invece in prima battuta acquisire la porzione di veicolo nel quale il componente danneggiato dovrà essere inserito, se non addirittura l'oggetto danneggiato, sarà possibile ottenere un modello tridimensionale virtuale dello

stesso, sul quale effettuare le opportune analisi e valutazioni progettuali al fine di realizzare la sagoma completa del componente funzionale al suo utilizzo. Una volta ottenuto il modello sarà quindi possibile, attraverso l'utilizzo di un sistema di stampa 3D per sinterizzazione, andare a realizzare in breve tempo l'oggetto fisico da inserire direttamente nel veicolo per il suo utilizzo. Il vantaggio di aver realizzato un'acquisizione tridimensionale, e poi la successiva costruzione di un modello tridimensionale virtuale, darà anche la possibilità di costruire una banca dati, eventualmente condivisibile in rete con logiche di utilizzo eventualmente selettive, che permetterà a tutti coloro che sono appassionati di auto storiche di eventualmente reperire i "ricambi" dei propri modelli.

Un discorso analogo può accadere in ambito medicale, quando per esempio possono essere necessarie protesi su misura. La disponibilità di sistemi di stampa 3D che lavorano con materiali biocompatibili ha fornito negli ultimi anni la possibilità di supportare quei soggetti, che a seguito di traumi od incidenti, hanno subito la perdita di porzioni di tessuto organico (porzioni di orecchio,...), attraverso la realizzazione di repliche. Per sfruttare queste tecnologie è però necessario disporre del modello virtuale del parte anatomica in oggetto, e questo può essere ottenuto attraverso l'impiego di scanner 3D ad elevata accuratezza in grado di rendere disponibile alla stampante 3D un modello il più fedele possibile ai lembi di tessuto che sono ancora presenti sul soggetto.

5. Bibliografia

1. Gatto A, Iuliano L. *Prototipazione rapida. La tecnologia per la competizione globale*, Tecniche Nuove, 1998,
2. Iuliano L. Vezzetti E., *An application for the thermoplastic injection moulding manufacturing*. International International Seminar on Progress in Innovative Manufacturing Engineering, 2001
3. Iuliano L. Vezzetti E., *Selective sampling in the reverse engineering of a free form surface*. International Conference on Advanced Research in Virtual and Rapid Prototyping, 2003
4. Vezzetti E., Moos S., Grivon D, Braul De Leon Rayant K., Fuentes Braco Gilmar E., Marino V. *Dispositivo e sistema di scansione tridimensionale, e relativo metodo* (TO2013A000202), 2013

Biografia

Luca Iuliano è Professore Ordinario di Tecnologie e Sistemi di Lavorazione presso il Dipartimento di Ingegneria Gestionale e della Produzione del Politecnico di Torino. E' autore e coautore di oltre 150 articoli scientifici nei settori delle lavorazioni meccaniche, della prototipazione rapida, del reverse engineering.

Enrico Vezzetti è Professore Associato di Disegno e Metodi per l'Ingegneria Industriale presso il Dipartimento di Ingegneria Gestionale e della Produzione del Politecnico di Torino. E' autore e coautore di oltre 100 articoli scientifici nei settori dello sviluppo prodotto, del reverse engineering, e della modellazione geometrica tridimensionale.

Anatomia del malware

M. Mezzalama, A. Liroy, H. Metwalley

L'esistenza di malware (virus, worm, cavalli di Troia) è uno degli aspetti negativi più rilevanti della rivoluzione digitale, con risvolti penali e economici. Il fenomeno coinvolge anche il settore dei dispositivi mobili (smartphone, tablet, ...) in cui si è passati da circa 2000 malware nel 2011, a più di 13000 nel 2012. Questo articolo analizza le tecniche per realizzare malware e permettere allo stesso di introdursi nei sistemi, rendersi residente, prenderne il controllo ed attivarsi a fronte di certi eventi, nascondendosi contemporaneamente ai programmi anti-virus. L'analisi considera sia i normali personal computer sia i più recenti dispositivi mobili.

Keywords: *Computer virus, Worm, Computer security for mobile devices, Antivirus*

1. Introduzione

L'informatica individuale e l'espansione delle comunicazioni Internet sono fenomeni di portata mondiale: circa il 35% della popolazione ha una connessione Internet, con picchi del 70-80% nelle nazioni più sviluppate [7]. Come ogni tecnologia pervasiva, anche l'informatica presenta "lati oscuri", e tra essi il "malware" è il fenomeno più rilevante, presente nel 90% dei cyber attacchi mondiali e che, solo nel 2004, ha provocato danni per oltre diciassette miliardi di dollari [3]. Il codice malevolo, però, è un fenomeno in continua crescita che ha ormai raggiunto numeri significativi: dal primo virus "Brain" (nel 1986) fino ad oggi sono stati creati oltre trenta milioni di programmi malevoli, con una crescita media di oltre il 100% negli ultimi anni.

Il termine *malware* deriva dalla contrazione di due parole inglesi, “malicious” e “software”, e indica programmi realizzati per danneggiare i sistemi su cui vengono eseguiti o per sottrarre informazioni sensibili. Nonostante spesso i programmi malevoli siano identificati genericamente come “virus informatico”, in realtà esistono tre grandi tipologie di malware: i *virus* sono programmi realizzati per creare danni e, successivamente, propagarsi inserendosi all’interno di altri file; i *worm*, invece, non hanno lo scopo di danneggiare direttamente il sistema ma cercano di auto replicarsi per saturare le risorse di sistema e di rete e, di conseguenza, creare danno; infine i *trojan horse* sono malware nascosti all’interno di legittime applicazioni software ma non possiedono codice dedicato alla propria replicazione e propagazione.

2. Struttura e organizzazione del malware

Nonostante ogni tipologia di codice virale usi strumenti, tecnologie e tattiche diverse, tutti possiedono un unico modello strutturale, basato sulle quattro fasi che il codice virale percorre (figura 1):

- **Infezione.** Il malware s’introduce all’interno del sistema, superando eventuali barriere di sicurezza, e s’installa al suo interno. In seguito il malware modifica le impostazioni del sistema adattandole alle proprie necessità, in primis quella di non essere rilevato.
- **Quiescenza.** Il codice virale rimane residente in memoria in attesa che si realizzi una determinata condizione, a seguito della quale si attiva (esecuzione delle azioni malevole e replicazione). Queste ultime possono essere ripetute più volte, quindi la fase si protrae fino a un’eventuale eliminazione da parte del codice stesso o di un software anti-malware.
- **Replicazione e propagazione (solo per virus e worm).** Al determinarsi di certi eventi o condizioni, il malware si replica e seleziona i bersagli verso cui propagarsi, infettando altri sistemi.
- **Azioni malevole.** Al verificarsi di certi eventi o condizioni, il codice virale esegue i propri compiti malevoli, come distruzione o furto dei dati del sistema. Se il sistema non viene compromesso definitivamente, il software ritorna nella fase di quiescenza.

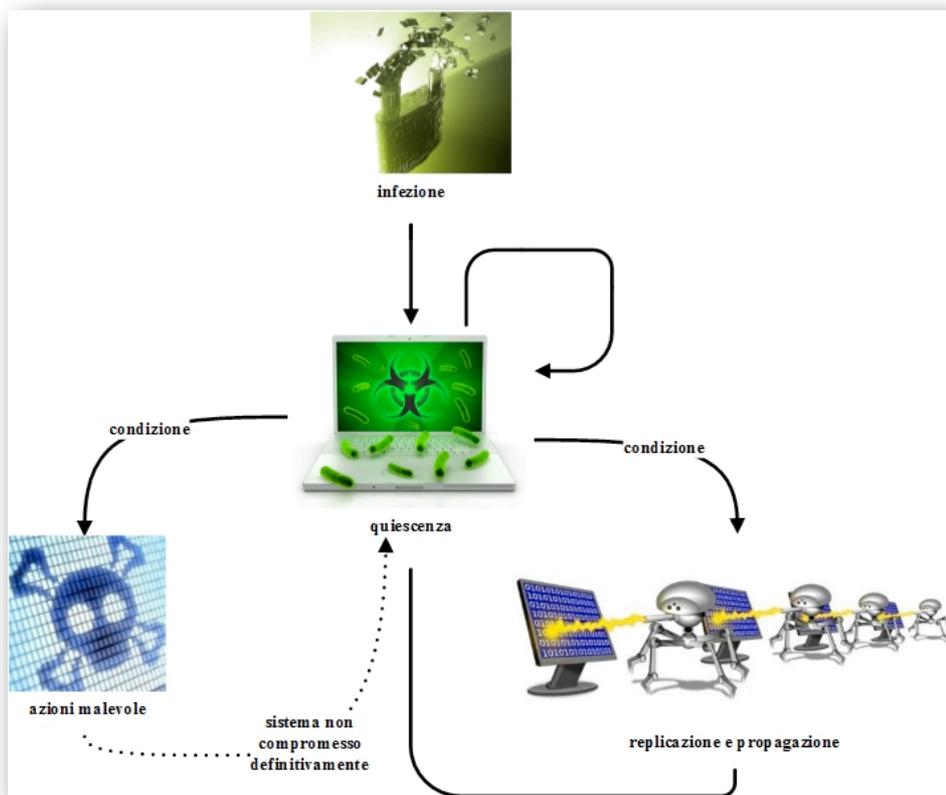


Figura 1
Ciclo di vita di un malware

2.1. Infezione

Il malware deve prima di tutto introdursi all'interno del sistema bersaglio e, se necessario, indurre la vittima a eseguirlo. La penetrazione passa in molti casi attraverso l'esecuzione di un file contenente, in modo diretto o indiretto, il codice virale. È ovvio che un codice malevolo non deve presentarsi in forma esplicita e, quindi, in questa fase spesso la parte malevola viene mascherata all'interno di programmi apparentemente legittimi attivati dall'utente in modo esplicito. Il primo compito di un malware è di inserirsi all'interno del sistema bersaglio e ciò avviene tramite tre principali canali:

1. **Trasferimento fisico.** Questa modalità, molto diffusa in passato ma ancora oggi frequente, prevede l'uso di un supporto di memorizzazione (come floppy disk, CD o unità USB) per il trasporto e l'inserimento del malware nei sistemi.

Le operazioni di trasporto e inserimento possono essere svolte dall'utente malevolo o, inconsapevolmente, dalla vittima stessa e prevedono un accesso fisico al PC. Stuxnet, uno dei malware più celebri e devastanti nella storia dei codici virali, usava questa tattica, inserendosi all'interno dei sistemi informatici tramite penne USB.

2. **Posta elettronica.** In questo caso il malware è allegato a messaggi di posta elettronica. L'utente viene così invitato in modo "suadente" ad aprire l'allegato, che può essere un file eseguibile o anche un documento elettronico (come un file Word o PDF contenente una macro pericolosa). "Melissa" è il primo malware che ha usato questo canale di diffusione ed ha infettato in breve tempo oltre ottantamila sistemi, causando danni a livello mondiale per più di un miliardo di dollari.
3. **Web.** Questo è attualmente il canale di diffusione più frequente e trasmette il codice malevolo attraverso un download da una pagina web. In particolare, in questo processo la vittima può svolgere un ruolo attivo o passivo. Nel primo caso, l'utente contrae il malware scaricando un file apparentemente innocente che, invece, si rivela essere malevolo. Nel secondo caso, invece, l'aggressore usa una serie di tecniche (note col nome di "drive-by download") che permettono la trasmissione di un malware con la semplice apertura di una pagina web. Quest'attacco sfrutta una o più vulnerabilità¹ presenti nei browser web ed evita, perciò, l'interazione con la vittima.

Una volta inseritosi nel sistema vittima, il malware deve attivarsi. Ciò richiede che l'utente debba forzatamente aprire il file contenente il codice virale. Questo contenuto, quindi, deve sembrare il più possibile inoffensivo e per questo motivo esistono tecniche di "mascheramento" per includere codice malevolo all'interno di contenuti leciti. Un caso tipico è quello dei virus informatici, che s'introducono autonomamente all'interno di altri file, solitamente eseguibili, modificandone la struttura interna. I file eseguibili possiedono una parte d'intestazione (in inglese, *header*) contenente informazioni destinate al sistema operativo, tra cui anche un puntatore che indica l'inizio del programma, cioè la prima istruzione macchina da eseguirsi all'attivazione del programma. Al momento dell'infezione, il virus si accoda al codice originale dilatando il file e, successivamente, sposta il puntatore originale al codice malevolo e ne crea un altro indirizzato verso l'istruzione di partenza originale, in modo che all'apertura del file il sistema operativo esegua prima il contenuto illecito e, in seguito, quello legittimo. Un esempio di questa tecnica è illustrata in figura 2, che mostra la differente struttura di un file normale e uno infettato [12].

¹ *Componente hardware o software, in corrispondenza alla quale le misure di sicurezza sono assenti, ridotte o compromesse. Esso rappresenta un punto debole e consente a un eventuale aggressore di compromettere il livello di sicurezza dell'intero sistema.*

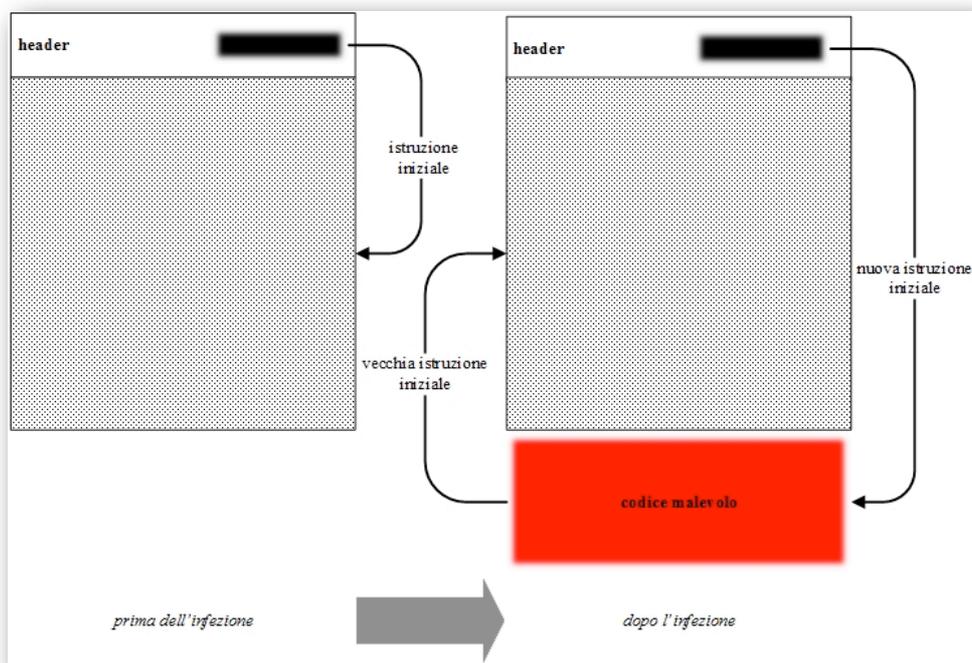


Figura 2
Mascheramento di un virus informatico

I worm, come i trojan, non possiedono la capacità di inserirsi autonomamente in altri file, ma vengono inclusi all'interno di programmi apparentemente legittimi. Per svolgere questa operazione l'utente malevolo usa di norma un *wrapper* (tra cui i più celebri sono "AFX File Lace", "Trojan Man" e "EliteWrap" [12]), uno strumento che consente di combinare due o più programmi all'interno di un solo file eseguibile. In questo modo, la vittima è convinta di aprire un programma apparentemente innocuo che, invece, contiene al proprio interno uno o più malware. Esempi sono il mascheramento in false applicazioni di sicurezza, aggiornamenti software o banner.

Per operare nel sistema, accedendo a tutte le sue risorse, il malware modifica le impostazioni del sistema per prenderne il controllo. Questa fase richiede lo sfruttamento di una vulnerabilità che permetta di ridurre o annullare le misure di sicurezza. Tecnicamente, quest'operazione si traduce nell'aumento dei privilegi di sistema attribuiti al programma malevolo: nell'ambito informatico, il privilegio classifica il livello di autorizzazioni fornite a un determinato programma e, conseguentemente, ad ogni sua azione. Di norma esistono due livelli, denominati "utente" e "amministratore" (o "root", in inglese), a seconda che un'operazione possa modificare componenti importanti del sistema, come le operazioni di I/O o le tabelle di sistema. Per acquisire elevati privilegi di sistema, un codice malevolo sfrutta una o più vulnerabilità. Le più diffuse sono di tipo "memory error", che

sfruttano criticità nella gestione di zone di memoria come stack o heap. La tipologia di attacco più conosciuta e usata di questa categoria è il “buffer overflow”. Questa tecnica, usata con successo contro numerosi programmi, come Adobe Reader, Microsoft Internet Explorer o Microsoft SQL Server, si basa sul fatto che, quando si attiva una routine, il suo indirizzo di ritorno viene memorizzato nella struttura dati denominata stack. L’utente malevolo cerca così di posizionare una serie di dati fuori dai limiti di un buffer². Alcuni linguaggi di programmazione, come Java, effettuano automaticamente controlli sul corretto uso del buffer, segnalando eventuali anomalie durante la fase di esecuzione del programma. Altri linguaggi, come il C, non operano alcuna verifica sulla corretta allocazione delle variabili, che quindi possono essere posizionate in indirizzi di memoria occupati da altri processi. Se, quindi, il programmatore non controlla scrupolosamente le dimensioni delle strutture dati prima dello spostamento in memoria, il programma diventa vulnerabile ad attacchi di questo tipo. Per sfruttare tale difetto, un utente malintenzionato (o direttamente un malware) invia all’applicazione una quantità di dati maggiore dello spazio assegnato in memoria, con l’obiettivo di provocarne un overflow della struttura e comprometterne zone critiche. Lavorando con attenzione i dati inviati al bersaglio, un attaccante può quindi eseguire qualsiasi tipo di codice malevolo e compiere facilmente azioni illecite senza l’autorizzazione della vittima [1].

La figura 3 illustra il medesimo stack: a sinistra correttamente allocato, a destra, invece, con un attacco di buffer overflow. Nell’uso appropriato, lo stack è usato per immagazzinare una stringa di dodici caratteri (in cui è inserita la parola “hello”). L’elemento più importante di questa struttura è il “Return Address”, che contiene l’indirizzo dell’istruzione da cui un programma deve riprendere la propria esecuzione dopo la chiamata di una eventuale funzione. Nel secondo caso, invece, è stata inserita la stringa “AAAAAAAAAAAAAAAAAAAAx08x35xC0x80” che, superando il limite dei dodici caratteri, provoca l’inserimento di dati illeciti nell’area di memoria dedicata all’elemento precedentemente illustrato. Per un utente malevolo è quindi possibile preparare una stringa per sovrascrivere il campo return address con l’indirizzo di un’area di memoria contenente un codice malevolo, che verrà eseguito dal sistema con i privilegi dell’applicazione chiamante. Per questo motivo, scoprire una vulnerabilità di questo tipo in un programma che possiede privilegi elevati permette ad un utente malevolo di compiere qualsiasi operazione desiderata [1].

² Un *buffer* è una sequenza di lunghezza predefinita di celle di memoria, concettualmente molto simile a un vettore.



Figura 3
Stack correttamente allocato (a sinistra) e con buffer overflow (a destra)

Ottenere i privilegi di sistema permette al malware di eseguire qualsiasi operazione sul sistema vittima, tra cui la più importante – nel caso di Windows – è la modifica del “registro di sistema”, una base dati che contiene al proprio interno le impostazioni del sistema operativo e ne regola il funzionamento. I campi di questa base dati (detti anche chiavi di registro), inaccessibili senza privilegi di amministratore, permettono di modificare il comportamento del sistema adattandolo ai bisogni del malware. Per un programma malevolo queste modifiche sono indispensabili perché gli permettono di accedere a risorse di sistema protette e di caricarsi in memoria ad ogni avvio del sistema operativo. Da recenti statistiche della nota azienda di sicurezza F-Secure, infatti, risulta che oltre il 40% dei malware modifica la chiave di registro “Run” che impone al sistema operativo di eseguire un determinato processo al proprio avvio [4]. In questo modo il programma malevolo si assicura la propria esecuzione. Gli esempi di malware che usavano questa chiave di registro sono numerosi: Sasser inseriva all’interno della chiave il valore “avserve.exe=%WinDir%\avserve.exe” (ove %WinDir% indica la cartella di installazione di Windows) mentre Sober inserisce copie di se stesso in tre file e crea tre chiavi di registro del tipo “<nome_casuale>=%System%\<nome_del_file_infetto>” [11].

2.2. Quiescenza

Completata la fase di penetrazione e infezione, il malware rimane residente in memoria, in attesa di una determinata condizione che lo avvii. Durante questa fase, è quiescente ma vigile per auto-protegersi da eventuali rilevamenti da parte della vittima o di un software di sicurezza.

La maggior parte dei malware rimangono residenti in memoria e quindi qualsiasi utente esperto può individuarlo tramite un'applicazione *task manager* che visualizza i processi in esecuzione nel sistema operativo, come "Gestione attività" in Windows. Per questo motivo, spesso gli attaccanti assegnano al processo malevolo il nome di un programma presente nel sistema. Per esser sicuri, inoltre, che essi siano realmente esistenti, vengono usati i nominativi di processi presenti di default nel sistema operativo, come "explorer" (processo di Microsoft Internet Explorer) o "notepad" (editor di testo di Microsoft Windows), o di applicativi che non desterebbero sospetto, come "win" (processo inesistente ma con un nome plausibile in un sistema Windows). Anche se la vittima scovasse la presenza del malware, non saprebbe quale tra i due processi (quello legittimo e quello illegittimo) eliminare. Per questo motivo, difficilmente l'utente elimina il processo malevolo, anche perché sbagliando la scelta verrebbe compromessa la stabilità del sistema operativo, in quanto spesso i malware assumono le sembianze di processi indispensabili per il corretto funzionamento [1].

Oltre a proteggere il malware dal rilevamento dell'utente, gli sviluppatori adottano anche tecniche di mascheramento per eludere i controlli di un eventuale anti-virus. Questi controlli possono sfruttare tecniche statiche o dinamiche.

2.3. Tecniche di rilevamento statiche

I metodi statici prevedono la rilevazione dei malware ricercando codice malevolo all'interno di file presenti su unità di memoria o flussi dati trasmessi in rete. L'idea è di attribuire ad ogni file una "firma" (basata su algoritmi di hash) che lo identifica univocamente e perde validità qualora il file venga modificato. La rilevazione di uno specifico malware si basa sulla rilevazione di sequenze univocamente associate al codice virale [1], anch'esse codificate con algoritmi di hash e note come *virus signature*. Il software di sicurezza scandisce il file e ricerca eventuali corrispondenze tra i dati e le sequenze dei codici virali noto. L'uso di questa tecnica implica la conoscenza e l'analisi di tutti i malware. Essa è quindi efficace contro gli attacchi noti ma è inutile contro i malware *zero-day*: nei giorni successivi alla creazione di un nuovo codice virale, esso non sarà ancora presente nel database degli anti-virus e sarà quindi in grado di superare i controlli di sicurezza. I produttori di anti-virus affrontano questo problema rilasciando periodicamente degli aggiornamenti del database per riconoscere i nuovi codici malevoli.

Per proteggere il sistema dal giorno del rilascio del malware fino all'aggiornamento del database vengono usate tecniche euristiche statiche, che permettono di riconoscere un codice malevolo con un alto livello di accuratezza senza, però, avvalersi delle signature. Questo tipo di tecnica prevede una prima fase di raccolta dati, durante la quale viene usato un certo numero di euristiche per esaminare il file in questione. Ogni singola euristica fa riferimento ad una particolare caratteristica virale, come la presenza di codice spazzatura o cifrato, l'uso di librerie insolite o la presenza di istruzioni non generate solitamente dai

compilatori o che modificano il vettore degli interrupt. Dopo aver raccolto questi dati, essi vengono analizzati, associando ad ogni euristica un grado di pericolosità e stilando un report che raccoglie le statistiche riguardanti il file scansionato. A seconda dei valori raggiunti dai parametri del report, il contenuto viene classificato come pulito, sospetto o infetto.

Per nascondersi dai controlli statici degli anti-virus, gli utenti malevoli adottano diverse tecniche di “offuscamento” del codice, per renderlo illeggibile o mutevole nel tempo. L’idea di base prevede che il corpo del malware (cioè la sequenza di istruzioni binarie che lo costituiscono) venga cifrato in modo che il codice virale non possa essere identificato mediante il rilevamento delle signature predefinite. La parte cifrata non può tuttavia essere eseguita se prima non viene decifrata; pertanto devono essere presenti anche istruzioni binarie e non cifrate eseguite all’avvio del malware. L’operazione di cifratura prevede l’uso di un’apposita chiave, nota come “*stub*”, per cifrare e decifrare il contenuto malevolo. Esistono diverse tecniche di cifratura ma quella più usata e semplice sfrutta una chiave segreta secondo i principi della crittografia simmetrica [3]. Ovviamente tale chiave non è statica, ma viene modificata di volta in volta per rendere il codice malevolo cifrato sempre diverso. Ad esempio, assumendo come chiave il contenuto di un’apposita area di memoria, è possibile mutare ad ogni infezione la chiave e, insieme ad essa, il corpo del codice virale.

Oltre alla cifratura, viene spesso usato un algoritmo di compressione per ridurre la dimensione del codice malevolo. Quest’operazione è spesso necessaria per evitare che la vittima venga insospettita da un’occupazione anomala di memoria. Questa tecnica, nota come “*pacchettizzazione*” (figura 4), permette di comprimere e cifrare il codice malevolo inserendogli in testa una sequenza di istruzioni per lo spaccettamento. Quest’approccio è ampiamente usato nei codici virali più recenti: secondo statistiche Symantec, oltre l’80% dei malware la usa.

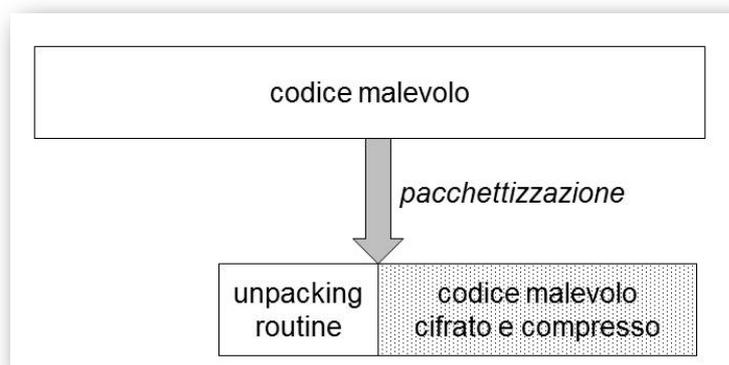


Figura 4
Pacchettizzazione di un codice malevolo

Il pregio di questa tecnica rappresenta però anche la sua principale debolezza: infatti, il codice binario che svolge le operazioni di cifratura e compressione (e viceversa) rimane identico ad ogni infezione e può perciò essere facilmente rilevato dagli anti-virus. Per risolvere questo problema, i malware più sofisticati e innovativi integrano al loro interno algoritmi in grado di variare le routine di cifratura e compressione. Ad esempio, per mutare una sequenza di istruzioni possono essere inserite istruzioni che non alterano il funzionamento del programma. Si consideri l'esempio in figura 5, dove le due serie di istruzioni hanno diverse sequenze binarie e quindi signature, ma svolgono le medesime operazioni. Infatti inserendo l'istruzione NOP (che non svolge alcuna operazione), il codice cambia ma il risultato è invariato. A seconda delle tecniche usate per la mutazione dei componenti di cifratura e compressione, si parla di malware "oligomorfi" (varie routine di cifratura ma in numero limitato), "polimorfi" (infinite varianti di cifratura) e "metamorfi" (infinite varianti del codice malevolo ma senza fare uso della cifratura). Collettivamente si parla di malware di tipo "stealth" (invisibile) per tutte queste tecniche di offuscamento [2].

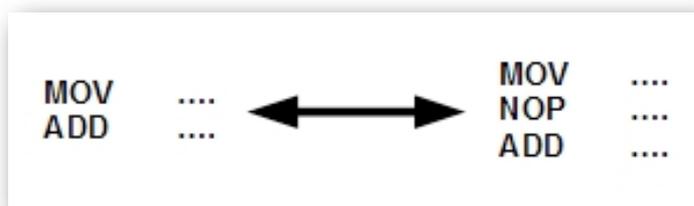


Figura 5
Esempio di istruzioni con uguale funzionamento ma signature differente

2.4. Tecniche di rilevamento dinamiche

Gli anti-virus usano anche tecniche dinamiche per rilevare il malware eseguendone il codice o controllandone il comportamento.

Il monitoraggio del comportamento (conosciuto anche come "behaviour blocker") è un componente dell'anti virus, solitamente residente in RAM, che controlla in tempo reale i programmi in esecuzione cercando comportamenti sospetti. In pratica, questi programmi analizzano tutte le operazioni svolte nel sistema, come quelle di lettura e scrittura su disco o gli accessi ad aree di memoria particolari. Sulla base di questi sintomi, il software di sicurezza riesce a rilevare programmi che eseguono azioni sospette e che, pertanto, sono potenzialmente dei malware. Se viene rilevata una azione allarmante, il software può bloccare l'operazione corrente, terminare il programma sospetto o chiedere all'utente di scegliere quale azione intraprendere [1]. Il pregio del monitor rappresenta però anche un suo punto debole, perché richiede l'esecuzione di un codice potenzialmente dannoso sul sistema reale. Per evitare questo problema, alcuni anti-virus emulano un

ambiente apposito per controllare il comportamento di un malware senza danneggiare il sistema reale: ad esempio, si analizza un codice cifrato eseguendolo su una macchina virtuale dedicata

Anche in questo caso gli scrittori di malware hanno messo in atto contromisure. Esistono infatti diversi strumenti, tra cui uno dei più famosi è *“red-pill”* [10], che rilevano se il programma è eseguito su una CPU reale o emulata. Integrando nel malware questi strumenti, è possibile eseguire il codice malevolo solo in presenza di un ambiente reale, evitando comportamenti illeciti in presenza di emulazione e superando i controlli imposti dal software di sicurezza.

I virus più recenti hanno iniziato a sfruttare anch’essi la virtualizzazione, eseguendosi ad un livello inferiore rispetto ai software di sicurezza. Questi programmi malevoli s’inseriscono all’interno del PC e convertono di nascosto un sistema fisico in uno virtuale, pienamente controllato dal malware che risulta così invisibile ai software di sicurezza che operano ad un livello superiore.

Oltre a nascondere il malware, durante questa fase il supervisor regola l’avvio delle azioni malevole e delle routine di replicazione e propagazione. Le fasi successive sono regolate da un componente aggiuntivo, il *“trigger”*, che le attiva al raggiungimento di una particolare condizione impostata dal realizzatore del malware. Gli eventi che scatenano il trigger possono essere molteplici, come ad esempio una data (ad esempio il virus *“Jerusalem”* si attivava ogni venerdì 13) o una particolare operazione eseguita dalla vittima.

Oltre ai malware elencati fino a questo punto, che operavano ad un livello strettamente locale, infettando un solo sistema per volta, esiste una particolare categoria di codice virale, conosciuta come *“malware distribuito”*, in cui ogni sistema infetto entra a far parte di una *“botnet”*, una rete formata da dispositivi collegati a Internet e controllati da un’unica entità, il *“botmaster”*. Una botnet può essere formata da milioni di sistemi infetti (una delle più grandi, chiamata *“Bredolab”*, era composta da oltre 30 milioni di PC), che vengono chiamati *“bot”* o *“zombie”*. I malware distribuiti possiedono un kernel che, oltre ad eseguire le normali operazioni, è in grado di contattare uno o più server remoti per scambiare informazioni. Per svolgere questa operazione, sfruttano i servizi di chat, connettendosi a un determinato canale protetto da una password. Per controllare e impartire ordini ai sistemi infetti vengono, inoltre, usati gli stessi canali di chat o reti peer-to-peer appositamente realizzate. Le botnet vengono usate per compiere attacchi *“distributed denial of service”* (DDoS) o campagne di spam: i DDoS mirano a compromettere il funzionamento di un sistema informatico inviando numerosi messaggi fino a esaurirne le risorse; le campagne di spam, invece, puntano a reclamizzare un determinato prodotto, inviando molte mail pubblicitarie.

2.5. Replicazione e propagazione

Parallelamente all'esecuzione delle azioni malevole, il malware deve anche replicarsi e selezionare nuovi obiettivi verso cui propagarsi. Questa fase è presente nei virus informatici e nei worm ma non nei trojan horse, che non si replicano automaticamente. Virus e worm possiedono, però, tattiche e comportamenti diversi: il primo punta a inserirsi all'interno di altri file e poi sfruttare la vittima per propagarsi, mentre il secondo si moltiplica e sfrutta la rete per propagarsi autonomamente [3].

I virus usano diverse tecniche per introdursi all'interno di altri file, solitamente eseguibili, e diffondersi. I virus di tipo "*parasitic*" inseriscono il proprio componente malevolo all'interno di porzioni di codice del file eseguibile da infettare. A seconda di dove venga inserito il codice virale, esistono diverse tipologie di questo tipo di malware: il *prepending*, che include il codice prima di quello originale, l'*appending*, che si salva dopo i contenuti originali, e il *cavity*, che si inserisce all'interno di eventuali spazi inusati nei settori di disco contenenti il file originale. I virus informatici, però, non compromettono solamente i file eseguibili ma si introducono anche in altri tipi di file, come boot sector o documenti. I boot sector virus infettano il settore di avvio presente nelle unità disco rimovibili (come unità USB o i vecchi floppy disk) o rigidi (hard disk) e ne assumono il controllo. Questo tipo di malware viene eseguito a ogni avvio del PC infetto prima del caricamento del sistema operativo e rimane in memoria fino allo spegnimento. Durante l'infezione, il virus si copia all'interno delle unità inserite nel lettore, che diventano un veicolo di diffusione se usati per avviare altri dispositivi, o danneggia i dati contenuti all'interno del sistema vittima.

Una particolare tipologia di virus informatici è rappresentata dai "macro virus", che non infettano file contenenti codice eseguibile ma file dati, inserendosi all'interno di documenti contenenti macro definizioni, sequenze di istruzioni scritte in linguaggio "Visual Basic for Application" e usate nei programmi della suite Microsoft Office per aumentare le potenzialità dei documenti di testo. Generalmente, per assicurarsi di essere eseguiti, essi infettano i modelli standard (nel caso di Word, il file "Normal.dot") che vengono presentati all'apertura di un programma di Office. In questo caso, ogni nuovo documento creato è automaticamente infetto. Un'altra tecnica usata da questi virus consiste nel modificare le macro associate alle voci di menu (ad esempio "Apri", "Salva" o "Salva con nome"): in questo modo, ogni volta che la vittima seleziona una di queste voci, il sistema esegue anche il codice malevolo, dando automaticamente inizio all'infezione di nuovi documenti [3].

I worm, al contrario dei virus informatici, non hanno capacità autonome di inserirsi in quanto, come detto in precedenza, sono gli stessi utenti malevoli ad inserirli in programmi leciti tramite i "wrapper". I worm, quindi, costituiscono un file a se stante e non hanno bisogno di infettarne altri. Per questo motivo, i worm possiedono tattiche di replicazione poco sofisticate, in quanto essi puntano

semplicemente a realizzare copie di se stessi mirando, quindi, ad infettare i PC e non i file. Questo tipo di malware adotta un sistema di propagazione particolarmente complesso costituito da due componenti, il “Target Selection Algorithm” e lo “Scanning Engine”. Il primo modulo seleziona gli indirizzi dei nodi collegati alla rete che costituiscono potenziali bersagli. Per realizzare questa operazione i worm usano diversi stratagemmi [3]:

- **Esplora risorse di rete.** Il malware esplora la rete simulando le richieste di un qualsiasi utente in cerca degli indirizzi dei sistemi vicini.
- **Richiesta DNS.** Il worm si collega al server DNS associato al sistema vittima e ottiene gli indirizzi di rete di altri nodi.
- **Contatti mail.** Il malware legge la lista dei contatti di posta elettronica dell’utente infetto. In questo modo, chiunque abbia scambiato messaggi di posta con la vittima diventa un potenziale bersaglio.

Lo scanning engine, invece, ha il compito di contattare gli indirizzi selezionati dal target selection algorithm e verificare se siano presenti le condizioni per una eventuale infezione. I worm, infatti, si propagano autonomamente sfruttando vulnerabilità presenti nei sistemi vittima. Lo scanning engine, quindi, inoltra verso il bersaglio una serie di pacchetti creati appositamente per verificare se il sistema sia vulnerabile agli exploit del malware e, se questa condizione si verifica, penetra al suo interno e lo infetta. L’efficacia e la velocità di questo componente è fondamentale per un worm, che ha come prerogativa la velocità di diffusione (Slammer infettò oltre 75000 server in soli dieci minuti) [3].

3. Il malware per dispositivi mobili

L’interesse degli utenti malevoli verso i dispositivi mobili nasce dagli anni 2000. Come si vede dalla figura 6, che illustra il numero totale di malware esistenti per dispositivi mobili secondo statistiche McAfee [9], in quel periodo i codici virali (consistenti in poche decine di unità) erano ancora poco presenti e non rappresentavano un problema. Le motivazioni di questa tendenza erano da ricercare, principalmente, nelle basse possibilità di guadagno (per gli utenti malevoli) e nelle limitate potenze di calcolo degli smartphone e della rete telefonica.

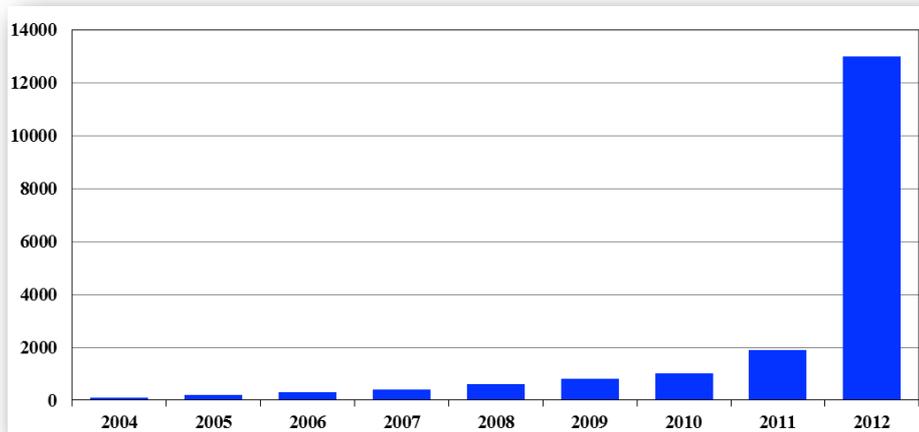


Figura 6
Tasso di crescita del malware per dispositivi mobili

La situazione cambia radicalmente a partire dal 2010, quando si realizzano una serie di condizioni che aumentano l'interesse degli utenti malevoli e quindi anche la creazione di malware. Infatti, attualmente esistono numerose congiunture che inducono gli aggressori a dedicarsi in maniera quasi esclusiva a questo settore:

- **Aumento delle vendite.** Le vendite dei soli smartphone hanno superato quelle di desktop e notebook nel 2011, con i tablet che effettueranno il sorpasso a breve. L'incremento vertiginoso della base di installazione di sistemi operativi mobili rappresenta uno dei principali fattori di interesse per gli attaccanti, attirati dalla possibilità di diffondere i propri malware su milioni di dispositivi.
- **Incremento delle prestazioni.** La potenza di calcolo degli smartphone è aumentata. Basti pensare, infatti, che le ultime generazioni di dispositivi possiedono processori quad core e elevati quantitativi di memoria RAM. Questa condizione permette di concepire malware potenti e devastanti quanto quelli esistenti su sistemi non mobili.
- **Aumento della velocità di rete.** Anche le reti dati hanno subito un incremento di prestazioni ragguardevole. Le ultime generazioni di collegamenti, infatti, possono raggiungere la velocità di un Gigabit per secondo. Le reti telefoniche, in accoppiata a quelle wireless, rappresentano un ambiente ideale per la diffusione e la realizzazione di botnet.
- **Sistemi operativi.** Le ultime generazioni di dispositivi mobili dispongono di sistemi operativi complessi ed avanzati come per PC, dove qualsiasi sviluppatore può creare liberamente applicazioni e, quindi, concepire codici virali senza alcuna restrizione.

- **Presenza di nuovi sensori.** Le ultime generazioni di smartphone e tablet possiedono sensori inesistenti su PC, come GPS, microfono o fotocamera, utili per ottenere preziose informazioni circa l'utente.

La prerogativa dei dispositivi mobili, quindi, è quella di avere sempre a portata di mano le informazioni e i contatti personali dell'utente, accessibili da qualsiasi luogo appoggiandosi a reti audio e dati senza fili. Gli smartphone e, recentemente, i tablet, contengono dati non solo personali ma anche di terzi, come quelli di amici e colleghi (i loro contatti, messaggi, appuntamenti, appunti condivisi e posizioni). Per questo motivo, gli obiettivi dei malware per dispositivi mobili sono mutati rispetto alla controparte per PC. Infatti, mentre i codici virali tradizionali miravano a realizzare una gamma molto variegata di attacchi, come denial of service, distributed denial of service, controllo del sistema da remoto oppure cancellazione o furto di dati sensibili, le applicazioni malevole per dispositivi mobili si sono focalizzate sull'ultima tipologia di attacchi.

Il mercato dei dispositivi mobili è attualmente dominato da due sistemi operativi: "Google Android" e "Apple iOS". Nonostante siano molto diversi tra loro, essi usano politiche di sicurezza simili:

1. **protezione contro accessi fisici non autorizzati** – l'utente può inserire meccanismi di protezione basati su codici PIN o segreti più complessi, che impediscono ad un utente malevolo di avere accesso al dispositivo senza l'autorizzazione del proprietario;
2. **sandbox** – ogni applicazione viene eseguita in una macchina virtuale dedicata (detta anche "sandbox"), che impedisce ad un codice malevolo di interagire con altri programmi in esecuzione o con lo stesso sistema operativo;
3. **cifratura dei contenuti** – vengono usati algoritmi per proteggere i contenuti personali dell'utente rendendoli illeggibili in caso di furto del dispositivo;
4. **firma delle applicazioni** – ogni applicazione è accompagnata da una firma digitale, per impedire che un malware possa inserirsi al suo interno.

Nonostante queste misure di sicurezza siano simili, la quasi totalità dei malware viene sviluppata per Android, interessando solo in minima parte iOS. Secondo le statistiche, il sistema operativo di Google raccoglie più dell'80% delle nuove minacce, mentre solo l'1% delle nuove famiglie di malware interessa iOS [5, 9]. Le motivazioni di questa tendenza sono da ricercare nelle diverse filosofie adottate da Google e Apple: Android, infatti, è un sistema operativo open-source in cui il proprietario del dispositivo ne ha un completo controllo, mentre iOS è una piattaforma chiusa che lascia molte meno possibilità di personalizzazione all'utente. Il diverso livello di apertura di un sistema operativo non si traduce solo in una diversa possibilità di personalizzazione, ma anche in un diverso modello di sicurezza. Apple, infatti, controlla meticolosamente ogni sorta di applicazione

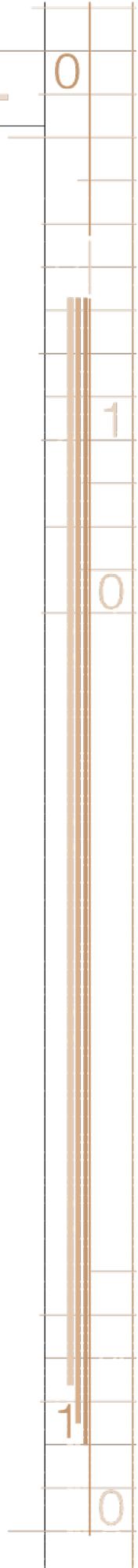
presente nel proprio canale di diffusione ufficiale, l'*App Store*, verificando l'identità degli sviluppatori e accertando la legittimità del codice da essi realizzato. Google, invece, effettua controlli più laschi sugli sviluppatori e sulle applicazioni disponibili sul proprio canale *Google Play*. Inoltre su qualsiasi dispositivo Android è possibile installare programmi esterni a Google Play e perciò un utente poco istruito circa le tematiche di sicurezza è esposto a tutti i pericoli provenienti dal malware. Queste motivazioni hanno portato allo sviluppo di numerosi codici virali per questa piattaforma, con un trend che, a partire dal 2010, ha subito un incremento continuo: ogni anno si assiste ad una crescita del malware per Android vicina al 100% [5].

I codici virali esistenti per Android possiedono un ciclo di vita del tutto identico a quelli presenti su PC, anche se la fase che ha subito maggiori cambiamenti è l'infezione. Anche in questo caso, quindi, la prima fase rappresenta un passaggio delicato per il malware, che deve adottare tecniche atte a non farsi trovare dall'utente, nel caso non ne sia richiesta l'interazione, o a convincerlo della propria legittimità, nel caso sia indispensabile che la vittima autorizzi l'installazione dell'applicazione. Nel passaggio da PC a mobile, i codici virali hanno però mutato i propri comportamenti e le proprie tattiche e, nel caso dell'inserimento iniziale nel dispositivo, questo cambiamento si è tradotto nella perdita di alcuni canali tradizionali, come il trasferimento fisico e la posta elettronica, a favore del web o dei "negozi" di applicazioni (detti anche comunemente "market"). Queste due tipologie di canali trasmissivi ben classificano le tattiche disponibili per i malware per dispositivi mobili: lo sfruttamento di vulnerabilità o l'inclusione di un codice virale in un'applicazione.

La prima possibilità prevede lo sfruttamento di una o più vulnerabilità per l'introduzione del malware e l'acquisizione dei privilegi a lui necessari. Solitamente vengono usate almeno due debolezze, una presente nel software di navigazione, per realizzare un attacco drive-by download e introdurre il programma malevolo, ed una seconda, presente direttamente nel sistema operativo, per acquisire i privilegi di amministratore e permettere al malware di installarsi silenziosamente senza richiedere l'autorizzazione della vittima. L'uso di vulnerabilità permette di realizzare codici virali efficaci e invasivi ma, di contro, poco compatibili. Infatti, le applicazioni malevole di questo tipo riescono ad installarsi solo sulla ristretta gamma di dispositivi mobili che ancora soffrono della debolezze scelte e che, quindi, non possiedono software del tutto aggiornato. La politica di rilascio degli aggiornamenti, inoltre, rappresenta un serio problema per Google: più del 45% dei dispositivi Android nel mondo possiede una versione obsoleta del sistema operativo ed ha quindi falle di sicurezza [6]. Le radici di questo problema sono da ricercarsi nell'elevata frequenza di rilascio di nuove versioni del sistema operativo e nella politica di aggiornamenti dei produttori di dispositivi. Ogni anno, infatti, viene rilasciata almeno una nuova versione di Android, ma essa non è subito installabile su qualsiasi dispositivo. Il sistema operativo di Google, infatti, è open-source ed è usabile da qualsiasi produttore

che, però, solitamente la adatta al proprio prodotto. Questo processo di personalizzazione, oltre ad essere lento, diventa progressivamente costoso e, quindi spesso smartphone o tablet relativamente recenti non vengono più aggiornati dal produttore, rimanendo fermi ad una versione vecchia di Android. Questa frammentazione permette di sfruttare vulnerabilità già largamente corrette da aggiornamenti che, in realtà, non vengono installati sulla maggior parte dei dispositivi. Inoltre, per aumentare la compatibilità delle proprie applicazioni, gli utenti malevoli integrano codice per sfruttare diverse debolezze. Un esempio è rappresentato da “DroidDream” e “DroidKungFu”, due celebri codici virali che sfruttano due diverse vulnerabilità per acquisire i privilegi di sistema in due diverse versioni di Android. La debolezza in questione, conosciuta come “Exploid” coinvolge il gestore di dispositivi per il kernel Linux. Fino alla versione 2.2 di Android, questo componente non controllava l’attendibilità della fonte dei messaggi di cambiamento di stato di una qualsiasi interfaccia fisica. In questo modo, se un’applicazione inviava un falso messaggio di connessione, il sistema operativo non controllava se esso provenisse dallo spazio kernel o utente, assegnando automaticamente i privilegi di root al programma che ha inviato il messaggio. Simulando l’inserimento di una periferica o attivando a ripetizione l’interfaccia wireless del dispositivo, quindi, era possibile ottenere facilmente i privilegi di amministratore.

Una seconda strategia di attacco è quella in cui l’utente malevolo tenta di realizzare un’applicazione che deve apparire il più possibile legittima, per convincere la vittima ad installarla e a dare inizio all’infezione. Questa tattica, quindi, non si basa sull’uso di debolezze software, ma sfrutta la relativa disinformazione degli utenti circa le tematiche di sicurezza. Le tecniche usate in questo caso sono molteplici ma si differenziano a seconda che si scelga di diffondere il malware attraverso Google Play o canali di diffusione alternativi. Questi ultimi rappresentano l’ambiente ideale per diffondere malware sotto forma di programmi piratati, ovvero versioni normalmente a pagamento diffuse, però, gratuitamente. Gli esempi di questo tipo di applicazione malevola sono numerosi e coinvolgono spesso giochi particolarmente celebri e desiderati, come “Angry Birds” o “Angry Birds Star Wars”. La tecnica più usata per realizzare un malware di questo tipo è il *repackaging*: tramite strumenti reperibili in rete (tra cui il celebre “apktool”) un utente malevolo può facilmente modificare e personalizzare l’applicazione originale inserendovi codice virale. Una volta che l’inclusione della componente malevola è completata, è possibile ricompilare i file ed ottenere una nuova applicazione identica all’originale, che però non possiede alcuna firma digitale, perché quella originale è stata compromessa durante il processo di decompilazione. Esistono però strumenti che permettono di firmare la nuova applicazione con un certificato “auto firmato” (noto anche come certificato *self-signed*), realizzabile da qualsiasi utente e che non necessita di alcuna validazione da parte di autorità di certificazione, cosa invece indispensabile per i certificati delle applicazioni ufficiali per Android. Infatti per la diffusione su Google Play è necessario che lo sviluppatore effettui una registrazione in cui deve



inserire i propri dati personali e, a seguito della quale, viene rilasciato un certificato da Google stessa. Se invece un utente desidera diffondere la propria applicazione attraverso canali non ufficiali, può usare un certificato auto firmato. In quest'ultimo caso, però, l'applicazione può solo distribuita fuori dallo store ufficiale, e quindi da una fonte sconosciuta: le impostazioni predefinite del sistema operativo non permettono l'installazione di questi contenuti ma, attraverso una semplice modifica, l'utente può consentirla. Completato il processo di firma, è possibile diffondere l'applicazione riconfezionata su qualsiasi canale di diffusione non ufficiale. Secondo statistiche redatte dall'azienda di sicurezza Lookout, questo tipo di tecnica coinvolge circa l'11% delle applicazioni presenti sui market alternativi [8].

4. Conclusioni

Nell'articolo si è evidenziato come il malware sia una minaccia quanto mai attuale per qualsiasi utente che possieda un dispositivo con accesso a Internet. I malintenzionati, infatti, sviluppano in continuazione nuove tattiche in grado di aggirare anche i più moderni sistemi di protezione, continuando la continua lotta tra "guardie e ladri" cominciata negli anni '90 con la diffusione di massa dei sistemi informatici. Inoltre, i dispositivi mobili, che fino a qualche anno fa erano solo in minima parte coinvolti da infezioni da parte di codici virali, stanno sempre di più diventando gli obiettivi preferiti degli utenti malevoli, grazie a caratteristiche introvabili sui PC tradizionali. Inoltre, la diffusione sempre più capillare dei dispositivi mobili ha fatto sì che sempre più utenti poco istruiti circa le tematiche di sicurezza compiano operazioni rischiose senza essere consapevoli dei pericoli. Prestando più attenzione nel compiere determinate azioni, infatti, sarebbe possibile diminuire notevolmente le possibilità di contrarre un malware. Nonostante che negli ultimi anni si siano fatti decisi passi avanti in tema di cultura della sicurezza informatica, questo campo risulta ancora poco noto. Recenti ricerche [13] evidenziano che quasi la metà degli utenti non effettua gli aggiornamenti software o addirittura non conosce l'esistenza di software di sicurezza appositi per dispositivi mobili. Risulta quindi indispensabile sensibilizzare gli utenti circa la sicurezza informatica, anche perché negli usi normali, soprattutto con smartphone e tablet, comportamenti adeguati sono più efficaci di molti software di sicurezza.

Riferimenti bibliografici

- [1] Aycock J., "Computer viruses and malware", Springer, 2006
- [2] Bashari B., Masrom M., Ibrahim S., "Camouflage in Malware: from Encryption to Metamorphism", International Journal of Computer Science and Network Security, Vol. 12 No. 8, Agosto 2012, pp. 74–83
- [3] Computer Economics, Inc, "2007 Malware Report: The Economic Impact of Viruses, Spyware, Adware, Botnets, and Other Malicious Code", Giugno 2007.
- [4] F-Secure Corporation, "Top10 malware registry launchpoints", <http://www.f-secure.com/weblog/archives/00001207.html>
- [5] F-Secure2 Corporation, "Mobile Threat report Q1 2013", 2013, http://www.f-secure.com/static/doc/labs_global/Research/Mobile_Threat_Report_Q1_2013.pdf
- [6] Google Android Developers, "Platform Versions", <http://developer.android.com/about/dashboards/index.html>
- [7] InternetWorld Stats, "World Internet Users and Population Stats", <http://www.internetworldstats.com/stats.htm>
- [8] Lookout Inc. "Alternative App Markets: Increase Choice of Apps, but Some Have a Higher Propensity for Security Risks", App Genome report, febbraio 2011, <https://www.lookout.com/appgenome/>
- [9] McAfee Labs, "McAfee threats report: Third quarter 2012", 2012, <http://www.mcafee.com/us/resources/reports/rp-quarterly-threat-q3-2012.pdf>
- [10] Paleari R., Martignoni L., Roglia G.F., Bruschi D., "A fistful of red-pills: How to automatically generate procedures to detect CPU emulators", 3rd USENIX Workshop on Offensive Technologies (WOOT), Agosto 2009
- [11] SecurityNET, *Virus informatici, internet worm e malware*. Edicred, Novembre 2005
- [12] Skoudis E., Zeltser L., "Malware: Fighting Malicious Code", Prentice Hall Ptr, 2004
- [13] Symantec Corp., "2012 NORTON CYBERCRIME REPORT", Luglio 2012

Biografia

Marco Mezzalama, Professore Ordinario di Sistemi di Elaborazione, ha svolto la propria attività didattica e scientifica presso il Dipartimento di Automatica e Informatica del Politecnico di Torino, del quale è stato Pro Rettore Vicario e Vice Rettore. Membro dell'Accademia delle Scienze, è stato revisore di progetti scientifici nazionali e internazionali. È autore di numerosi articoli scientifici e di alcuni libri. Ha interessi di ricerca nel settore della sicurezza informatica, delle architetture di elaborazione e dell'e-learning.

Email: marco.mezzalama@polito.it

Antonio Lioy, è Professore Ordinario di Sistemi di Elaborazione presso il Dipartimento di Automatica e Informatica del Politecnico di Torino. Attualmente la sua ricerca si concentra sulla sicurezza ICT, con particolare attenzione all'identità digitale (e-ID), ai sistemi informativi *trusted* ed ai sistemi di protezione basati su metodi formali (policy e ontologie). Agisce frequentemente come valutatore e revisore di progetti per la Commissione Europea. Autore di oltre 100 pubblicazioni, è stato Coordinatore del progetto POSITIF (IST-FP6) ed attualmente lo è del progetto SECURED (IST-FP7).

Email: antonio.lioy@polito.it

Hassan Metwalley, laureato in computer engineering presso il Politecnico di Torino nel 2013, ha svolto la propria tesi studiando i codici virali per i sistemi informatici. Attualmente lavora presso Syskoplan Reply occupandosi di applicazioni per sistemi operativi Android e iOS.

Email: hassan.metwalley@gmail.com

Information Fusion e sorveglianza

F. Flammini, A. Pappalardo, V. Vittorini

Uno degli ambiti maggiormente interessati dall'aumento del rapporto tra numero di elaboratori e numero di utenti è quello dei sistemi di monitoraggio distribuito per applicazioni di supervisione, diagnostica e sorveglianza. In tali sistemi, infatti, pochi operatori sono chiamati a controllare un elevato numero di sensori e sottosistemi, spesso distribuiti ed eterogenei. In questo articolo verranno illustrate tecniche automatiche di integrazione ed interpretazione dei dati per consentire di gestire una mole potenzialmente elevata di informazioni ed adottare decisioni efficaci in tempi rapidi.

Keywords: *Physical Security, Smart-sensing, Surveillance, Information Fusion*

1. Introduzione

Le reti di sensori costituite da dispositivi eterogenei sono adottate in molti ambiti al fine di rilevare eventi di diversa natura e supportare di conseguenza le decisioni degli operatori delle sale di controllo. Si pensi ad esempio ai seguenti contesti:

- monitoraggio ambientale per l'allerta precoce (*early warning*) di eventi potenzialmente disastrosi quali terremoti, alluvioni, smottamenti, incendi, etc.
- sistemi di allarme, sicurezza fisica¹ e sorveglianza per il rilevamento di intrusioni (*intrusion detection*) in aree protette, comportamenti anomali, individui sospetti, sostanze pericolose, etc.

¹ I sistemi di supervisione applicati alla sicurezza fisica sono stati recentemente battezzati come PSIM (Physical Security Information Management), per specializzarli rispetto ai più tradizionali SCADA (Supervisory Control and Data Acquisition).

- prognostica e diagnostica, per anticipare, rilevare e gestire guasti o incidenti in centrali, impianti industriali di produzione e in altre infrastrutture critiche.

In questi ed altri contesti emergenti, per ragioni di economicità e facilità di installazione, i sensori tradizionali sono sempre più affiancati o soppiantati da moderni *smart-sensors*, quali i cosiddetti *motest* delle Wireless Sensor Networks (WSN) [5], e *smart-cameras*, ovvero videocamere di rete "intelligenti", con un sistema di elaborazione dedicato a bordo. Tali dispositivi sono in grado di pre-elaborare le misure di parametri ambientali al fine di generare eventi anziché dati grezzi (*raw data*). Ad esempio, i sensori WSN possono notificare in automatico eventi quali temperatura, umidità o accelerazione oltre i limiti di normalità, mentre le telecamere intelligenti sono in grado di far funzionare sofisticati algoritmi di visione artificiale, per il rilevamento di presenza persone, attraversamento perimetri critici, oggetti incustoditi, elevato affollamento ed altro ancora [8].

La rapida crescita ed evoluzione dei sistemi di monitoraggio distribuito sopracitati fa sì che il numero di sensori, e quindi la mole di eventi potenzialmente generati (incluse le segnalazioni non attendibili, ovvero i cosiddetti "falsi positivi"), non sia compatibile con la capacità di gestione in tempo reale di un numero ridotto di operatori umani. Quello che deve essere realizzato, quindi, è un compromesso tra l'opportunità di raccogliere un numero di informazioni il più elevato possibile, per incrementare le capacità di riconoscimento e consapevolezza situazionale (brutta versione italiana dell'inglese *situation awareness*), ed il rispetto dei requisiti di usabilità ed ergonomia (i cosiddetti *human factors*), che impattano sulla correttezza e sulla tempistica della risposta agli eventi. E' quindi una delle molteplici applicazioni, in un settore più che mai interdisciplinare, dell'ambito di ricerca comunemente noto come *information fusion* per il supporto alle decisioni (si veda il Riquadro 1).

Da quanto detto, dovrebbe essere chiaro che i sistemi intelligenti di monitoraggio distribuito impiegati in attività di supervisione, diagnostica e sorveglianza, necessitano di un continuo processo di innovazione, da un lato per soddisfare requisiti applicativi in rapida evoluzione, dall'altro per affrontare le criticità ed i problemi aperti che si vengono a creare come naturale conseguenza di tale evoluzione. Gli aspetti più coinvolti da questo processo riguardano in particolare:

- La problematica generale della gestione di dati provenienti da molte sorgenti eterogenee di informazione da parte di un ridotto numero di utenti [7].
- La specializzazione al caso dei sistemi di supervisione, monitoraggio distribuito e sorveglianza, in cui un limitato numero di operatori deve prendere decisioni in funzione di segnalazioni provenienti da un numero elevato di sottosistemi e apparati sensoriali [13].

- La disponibilità di sistemi di integrazione e *reasoning* di tali informazioni allo scopo di supportare le decisioni degli operatori [12].

Questi sono gli aspetti che andremo ad affrontare in questo articolo, partendo dalle esigenze industriali che emergono da scenari applicativi reali (descritti nel paragrafo 2) ed illustrando i diversi ambiti di ricerca coinvolti, ivi comprese le tecnologie di rilevamento multi-modale e multi-mediale e le tecniche di aggregazione delle informazioni (paragrafi 3-6), escludendo dal nostro ambito di analisi gli aspetti più propriamente *human-related* relativi alla percezione e alla cognizione, a cui si daranno dei rapidi cenni e qualche puntatore alla letteratura di riferimento.

2. Information Fusion applicata alla sorveglianza

Il progresso delle tecnologie di sorveglianza registrato negli ultimi anni ha permesso di soddisfare requisiti applicativi in rapida evoluzione, ma ha anche creato aspettative irrealistiche da parte degli utenti, spesso poco consapevoli del rischio non trascurabile di falsi allarmi e dei loro effetti collaterali. Un'altra considerazione è che ha poco senso disporre di tecnologie eterogenee, come richiesto in ambienti complessi, senza sfruttarle adeguatamente. Pertanto, è evidente la necessità di comprendere come le tecnologie possano essere combinate ed integrate, per trovare il miglior compromesso tra costi, complessità di gestione e benefici attesi di un sistema di monitoraggio in un determinato ambito applicativo.

Tali aspetti sono strettamente connessi alle capacità di supporto alle decisioni che ci aspetta dai sistemi di supervisione, diagnostica e sorveglianza. In questo senso, supportare in maniera affidabile il processo di *decision-making* è fondamentale per evitare interpretazioni errate delle informazioni che potrebbero condurre all'attivazione di contromisure non appropriate.

Metodi, tecniche e strumenti disponibili a tal scopo dovrebbero essere inquadrati in un paradigma che tenga conto dei componenti principali del sistema e del ruolo fondamentale dell'operatore umano. Da una parte, c'è la necessità di inserire la prospettiva dell'utente nella progettazione di un sistema di *Information Fusion* (IF) [3]; dall'altra c'è bisogno di impiegare l'IF nella progettazione di un sistema di *Decision Support* (DS) per migliorare l'intero processo decisionale [12]. In altre parole, è necessaria la convergenza tra queste due aree di ricerca che, sebbene intimamente connesse, si sono sviluppate negli anni in maniera separata.

Più in dettaglio, scenari applicativi reali evidenziano esigenze che si pongono su livelli differenti: capacità funzionali, architetture di riferimento e grado di interazione con gli utenti.

2.1. Livelli di information fusion

Per prima cosa, bisogna definire il livello di astrazione della strategia di fusion [1]. In altre parole, è necessario comprendere se gli input del processo sono dati grezzi misurati dai sensori o dati pre-elaborati con l'intento di estrarre già un'informazione di più alto livello semantico. Analogamente, va definito il tipo di output del processo, che rappresenta il fine ultimo per il quale viene implementata la IF. Dal punto di vista dell'operatore, la strategia di fusion è quella di alto livello (*Decision Fusion*). La motivazione risiede nel fatto che in questo modo è possibile ragionare a partire da informazioni allo stesso livello semantico, ottenuto omogeneizzando molteplici *input*, ognuno caratterizzato da una diversa rappresentazione delle informazioni. Inoltre, tale strategia di fusion offre una scalabilità che è difficile da ottenere con una IF di basso livello, il che è un requisito importante laddove le aree da monitorare sono potenzialmente estese. Ciò non vieta di adottare strategie multi-livello: a basso livello si impiegano metodi di fusion mono o multi-modale (come sarà meglio spiegato in seguito), mentre ad alto livello si impiegano dati già elaborati. A livello di output, la strategia da perseguire nel contesto in esame dovrebbe consentire di valutare aspetti diversi: i) stati e attributi delle entità monitorate; ii) relazioni reciproche tra le entità e con l'ambiente circostante; iii) stati futuri e proiezioni delle situazioni interpretate, per valutare minacce, rischi e possibili impatti.

2.2. Gerarchia di centri di controllo e operatori

Riconoscere le situazioni in corso di svolgimento, la loro fase di evoluzione e possibili trend, è funzionale all'adozione delle misure più appropriate in risposta a quanto rilevato. Un processo congiunto di IF e DS consente dunque di potenziare in maniera significativa le capacità degli operatori umani, che dai centri di controllo supervisionano gli apparati installati in campo ed eseguono le procedure di gestione degli eventi. Le azioni incluse in queste procedure sono spesso orchestrate dagli stessi sistemi di monitoraggio e sorveglianza, la cui architettura può essere distribuita e gerarchica, con un numero ridotto di sale di controllo centrali e molte più sale locali, che raccolgono le segnalazioni in base a scopi e responsabilità differenti [2]. Attraverso apposite interfacce, un operatore può supervisionare, prendere in carico e confermare l'esecuzione di ciascuna fase procedurale che è nella sua responsabilità. Integrando queste funzionalità con quelle più avanzate che conseguono dall'IF, è possibile reagire in maniera più tempestiva e puntuale alle situazioni di emergenza. Ad esempio, in caso di sistemi di supervisione distribuiti e gerarchici, è possibile riconoscere fenomeni complessi, quali attacchi terroristici di tipo strategico, che prevedono l'esecuzione contemporanea di azioni dolose nei confronti di asset distanti dell'infrastruttura monitorata [9].

2.3. Information fusion e fattori umani

Molte infrastrutture sono geograficamente estese e richiedono la protezione di molteplici risorse e tipologie di utenti (operatori, manutentori, pubblico). Per garantire una copertura adeguata di tali infrastrutture è necessario impiegare un numero elevato di dispositivi (telecamere, sensori anti-intrusione, rilevatori di sostanze chimiche, etc.). Ciò rende impraticabile una sorveglianza unicamente *human-based*, dato che l'analisi manuale di numerosi flussi video ed altre segnalazioni di allarme è molto impegnativa, soggetta a cali di attenzione ed errori di gestione. Dal momento che il numero di operatori umani è tipicamente limitato, l'analisi del fattore umano assume una notevole importanza [12] [6]. Da una parte, infatti, vi è la tendenza dei singoli sottosistemi integrati a produrre una mole elevata di dati (eventi, allarmi, messaggi diagnostici, etc.), dall'altra le ridotte capacità umane nel gestirli in tempo reale, specialmente nel caso di eventi critici simultanei. Inoltre, attività fortemente non stimolanti e ripetitive rendono all'atto pratico poco efficace la sorveglianza assistita dall'uomo [13]. Tecnologie come l'analisi video intelligente attenuano considerevolmente i problemi dovuti ai fattori umani, che attualmente rappresentano uno dei principali handicap nell'analisi dei dati relativi ad attività di sorveglianza. Tuttavia, come anticipato, ciò non è ancora sufficiente, in quanto le tecnologie disponibili allo stato dell'arte non hanno livelli di affidabilità elevatissimi oppure generano eventi che devono essere comunque inquadrati nel contesto, interpretati e verificati prima di adottare eventuali contromisure. Pertanto, il grado di dipendenza dal livello di attenzione degli operatori è ancora molto elevato.

Relativamente alle azioni di risposta agli eventi, il controllo dei falsi allarmi è fondamentale. Da questo punto di vista, tutti i sistemi automatici di diagnostica (o similari) dovrebbero garantire una soglia di affidabilità minima, al di sotto della quale il sistema diventa all'atto pratico inefficace: nello specifico, in base a recenti studi sull'ergonomia, tale soglia viene superata quando oltre il 30% delle segnalazioni prodotte dal sistema sono da scartare perché false o irrilevanti [14]. La riduzione del numero di falsi allarmi riduce i costi di gestione ed ha un significativo impatto sulle prestazioni degli operatori, influenzandone l'efficienza in termini di rapidità di risposta e confidenza nell'accuratezza del sistema. L'implementazione di un processo di IF applicato alla sorveglianza è dunque funzionale anche al miglioramento dell'affidabilità delle segnalazioni, che dipende sostanzialmente da due parametri: la probabilità di rilevamento (*Probability Of Detection*, indicata con POD) ed il tasso di falsi allarmi (*False Alarm Rate*, indicato con FAR).

3. Smart-sensing

Per raggiungere il livello richiesto di conoscenza relativo ad un'area da monitorare è necessaria la "cattura" di flussi informativi appropriati. Un sistema di sorveglianza efficace utilizza sensori per monitorare parametri ambientali, determinare il verificarsi di condizioni anomale, rilevare l'occorrenza di eventi indesiderati. Essi rappresentano dunque le principali sorgenti di informazione, che concorrono a raggiungere una opportuna situation awareness.

I dispositivi sensoriali sono basati su tecnologie più o meno sofisticate a seconda delle caratteristiche richieste in termini di funzionalità, prestazioni, costi, adattabilità a particolari condizioni ambientali, etc. Uno degli ambiti applicativi in cui si richiede che i sensori esibiscano sofisticate funzionalità ed alte prestazioni è, ad esempio, la protezione delle infrastrutture critiche in contesti di homeland security.

Le attuali esigenze di sorveglianza richiedono che i moderni sistemi non rivestano più solo un ruolo di supporto, ma diventino il cuore del processo di inferenza. Poiché l'"intelligenza" del sistema risiede anche nella capacità di sensing, si è assistito negli ultimi anni allo sviluppo e all'utilizzo sempre crescente dei cosiddetti sensori smart.

Per capire cosa si intenda per smart-sensor, è opportuno specificare cosa intendiamo per sensore "standard". In generale, un sensore è un dispositivo realizzato con l'obiettivo di effettuare una misura di una grandezza fisica relativa ad un oggetto o all'ambiente, e trasformarla in un segnale elettrico analogico o digitale. In Figura 1 è mostrata l'architettura logica di un sensore, che comprende le seguenti componenti principali: sensing unit (per l'acquisizione del segnale), processing unit (ad esempio per l'amplificazione, la compensazione e il filtraggio del segnale), power unit (preposta all'alimentazione elettrica), sensor interface (per il collegamento fisico del dispositivo con altre componenti elettroniche).



Figura 1
Architettura logica di un sensore

La principale differenza tra sensori “smart” e sensori “standard” risiede nel fatto che i primi ospitano a bordo sistemi *embedded* completi, dotati di microprocessore, unità di memoria, trasmettitori radio, e relativi software di base per consentirne la programmazione. Di tali sensori “evoluti” si sono date negli ultimi anni diverse definizioni, tra le quali le due più comunemente accettate ne sottolineano rispettivamente gli aspetti tecnologici (sensori “smart”) oppure le caratteristiche funzionali (sensori “intelligenti”). La distinzione si riferisce principalmente al ruolo assunto dal microprocessore: se è utilizzato ad esempio per l’elaborazione digitale dei segnali, il calcolo e le funzioni di interfacciamento, o se viene anche impiegato a supporto di funzionalità avanzate (“intelligenti”) quali l’auto-diagnostica, l’auto-adattamento, l’auto-identificazione, l’auto-calibrazione, etc. In alcuni casi, il microprocessore può essere utilizzato anche per decidere quando memorizzare dati o per minimizzare il consumo di energia.

Traendo vantaggio dei recenti sviluppi della tecnologia nell’ambito della microelettronica e dei sistemi micro elettro-meccanici (MEMS), la dimensione di tali sensori è andata riducendosi nel tempo, mentre la produzione intensiva per una grande varietà di applicazioni basata su tecnologia VLSI (*Very Large Scale Integration*) ha portato ad una forte riduzione del loro costo. Questi fattori hanno contribuito a rendere la produzione dei sensori sempre più conveniente, a fronte di un’“intelligenza” sempre maggiore. Nel prossimo futuro, il progresso tecnologico porterà un miglioramento anche nelle altre componenti (memoria, trasmissione radio, batterie) consentendo autonomia elevata e costi di manutenzione ridotti.

Le prime applicazioni di sensori intelligenti sono state attuate nell’ambito dell’ingegneria civile, ad es. nel monitoraggio strutturale (*Structural Health Monitoring*, SHM). Oggi, l’adozione di sensori evoluti è sempre più frequente in diversi ambiti. In particolare, è in aumento il loro utilizzo nell’ambito della difesa del territorio al fine di implementare strategie di allerta precoce (*early warning*). In tale contesto, la sorveglianza di un’area da proteggere si è spostata dalla tradizionale TVCC (televisione a circuito chiuso) a sistemi che integrano dispositivi molto eterogenei e con differenti livelli di intelligenza, classificabili in determinate categorie funzionali [11]. Tra i sensori più noti rientrano quelli per: il rilevamento delle intrusioni in aree protette tramite radar ad infrarosso, microonde o ultrasuoni, il riconoscimento tramite *smart-cameras* di comportamenti anomali (es. bagaglio incustodito); il monitoraggio di parametri ambientali quali temperatura, luminosità, umidità, accelerazione, tramite *motes*; l’analisi audio intelligente per rilevare urla, spari e rottura vetri; il rilevamento di agenti nocivi (CBRNe: Chimico, Biologico, Radiologico, Nucleare, esplosivo); l’ispezione a distanza di una massa di individui per rilevare armi nascoste, tramite onde millimetriche a terahertz, etc.

4. Sorveglianza multi-modale e multi-mediale

Sebbene la video sorveglianza rappresenti la forma più popolare di monitoraggio, ne esistono diverse altre in via di diffusione. Come evidenziato nella precedente sezione, l'impiego di nuovi dispositivi è da un lato incoraggiato dagli avanzamenti tecnologici, dall'altro richiesto da requisiti in continua evoluzione. L'integrazione di tecnologie eterogenee conduce a sistemi di sorveglianza multi-modali e multi-mediali [14], in cui la diversità tecnologica rappresenta un efficace valore aggiunto. I sistemi di sorveglianza multi-mediali collezionano ed elaborano diversi flussi di informazioni: video, audio, e qualunque altro output provenga da sensori installati nella stessa area. Pertanto si può affermare che un sistema di sorveglianza multi-modale e multi-mediale (ovvero capace di fornire informazioni di diversa natura, catturate attraverso mezzi differenti) combina due caratteristiche principali:

- l'utilizzo di più sensori connessi in rete, con la possibilità di sovrapporre parzialmente le aree di monitoraggio;
- l'utilizzo di sensori eterogenei, in modo da rilevare più informazioni di interesse disponibili all'interno dell'area monitorata.

La combinazione può essere realizzata adottando diverse soluzioni. Ad esempio, una configurazione elementare può essere basata su videocamere intelligenti o su sensori acustici, appartenenti a classi diverse e quindi con diverse caratteristiche e potenzialità. La soluzione basata su videocamere potrebbe ad esempio comprendere dispositivi brandeggiabili (PTZ: *Pan-Tilt-Zoom*), con un campo visivo modificabile e regolabile da remoto, oltre che videocamere fisse, e con tecnologie ad infrarossi, termiche, etc. Questo tipo di soluzione è generalmente adottata per migliorare ed estendere la visione della scena sfruttando efficacemente *ridondanza* e *sovrapposizione* nell'installazione dei dispositivi. Da un punto di vista prettamente funzionale, una simile rete di video camere consente di catturare maggiori dettagli relativi agli eventi identificati e di coprire aree estese; ad esempio è possibile effettuare automaticamente lo zoom di un intruso e identificarlo, tracciare oggetti o persone sospette, ed effettuare l'analisi di una folla di persone. Inoltre, specifiche combinazioni di videocamere fisse e PTZ possono essere utilizzate, attraverso un approccio distribuito, regolando in automatico alcuni parametri dei dispositivi installati (come la risoluzione) a seconda della complessità della scena osservata (ad esempio in presenza di un numero più o meno alto di persone). In questo modo una PTZ può osservare una singola persona con alta risoluzione, o tracciare più persone allo stesso tempo ad una risoluzione inferiore.

Più in generale, differenti punti di vista di una stessa scena possono risolvere eventuali problemi dovuti ad occlusioni, guasti e sabotaggi. Questo è un punto

particolarmente importante nel caso di una videosorveglianza intelligente, ovvero realizzata con il supporto di una video-analisi via software dei flussi video catturati dalle telecamere. In tal caso infatti, le prestazioni di un algoritmo di *video analytics* dipendono anche dalle dimensioni dell'area sottoposta a sorveglianza, all'interno della quale il numero di elementi occlusivi (persone, pilastri, muri, o altri oggetti che ostruiscono il campo visivo) hanno evidentemente un impatto sul rilevamento ed il tracciamento di un obiettivo.

E' possibile pertanto migliorare la qualità del monitoraggio, e di conseguenza superare alcuni limiti degli algoritmi di video analisi anche in ambienti non favorevoli ed in condizioni di affollamento. Questi benefici hanno alcuni contro-altari, quali la difficoltà che si può incontrare nella corretta calibrazione dei dispositivi e nella corretta gestione delle videocamere disponibili. Quest'ultimo elemento è peraltro fondamentale per l'espletamento di funzionalità quali il tracciamento, essendo di grande importanza la corretta gestione dell'*hand-over* tra una video camera e l'altra, in modo da non perdere l'oggetto tracciato e la sua identità durante il suo spostamento.

Oltre alle occlusioni, l'efficacia della video-analisi può risentire anche di improvvisi cambiamenti di luminosità, della presenza di superfici riflettenti o di oggetti che si confondono con lo sfondo (*background*) della scena. In generale, la sorveglianza in situazioni complesse non può essere condotta soltanto attraverso video camere, ma è necessario – come già menzionato – acquisire ed integrare diverse fonti di informazione.

A questo va aggiunto che, per soddisfare le richieste sempre più sofisticate degli utenti finali, la complessità degli eventi che si desidera rilevare è sempre maggiore. Tipicamente, l'approccio alla base di queste configurazioni multimodali più avanzate può essere di tipo *bottom-up* o *top-down*.

L'approccio *bottom-up* è utilizzato per realizzare integrazioni ad-hoc, in cui i sensori sono considerati come *black-box* e non utilizzano nelle loro elaborazioni le informazioni provenienti da altri sensori. La logica di integrazione, quindi, è esterna e risiede ad un più alto livello di astrazione.

Nell'approccio *top-down*, invece, gli algoritmi eseguiti dai singoli sensori utilizzano le informazioni provenienti da altri dispositivi. In questo caso l'output di ogni sensore dipende dalla presenza di ulteriori sorgenti di informazione nell'area circostante. Ognuno di essi implementa dunque una parte della logica di integrazione, realizzata ad un più basso livello di astrazione. Ad es. mediante un approccio top-down applicato all'analisi video, si può comprendere quando un cambiamento nella scena è riconducibile ad una ordinaria evoluzione della stessa (ad es. a causa di fattori meteorologici) e aggiornare di conseguenza il background.

5. Processo di Information Fusion

Per garantire un adeguato livello di affidabilità, è fondamentale impiegare un approccio rigoroso, finalizzato a sfruttare i benefici offerti da tecnologie (smart-sensing), tecniche avanzate di sorveglianza (multi-modale e multi-mediale), e metodi di fusione delle informazioni. L'approccio al problema deve avere sia una base metodologica che una applicativa, al fine di trovare giusti compromessi in base alle dimensioni del flusso di informazioni da analizzare in tempo reale. Sono funzionali a tale scopo:

- La definizione di un'architettura per l'integrazione di tecnologie, tecniche e strumenti attualmente disponibili;
- Lo studio e l'applicazione di opportuni modelli e meccanismi di rilevamento.

Riguardo il primo punto, è necessario un paradigma finalizzato all'aumento delle capacità nella sorveglianza distribuita, a prescindere dal contesto applicativo. Esso risponde anche alla necessità di sfruttare la prospettiva dell'utente nella progettazione del sistema, e di migliorare l'intero processo decisionale.

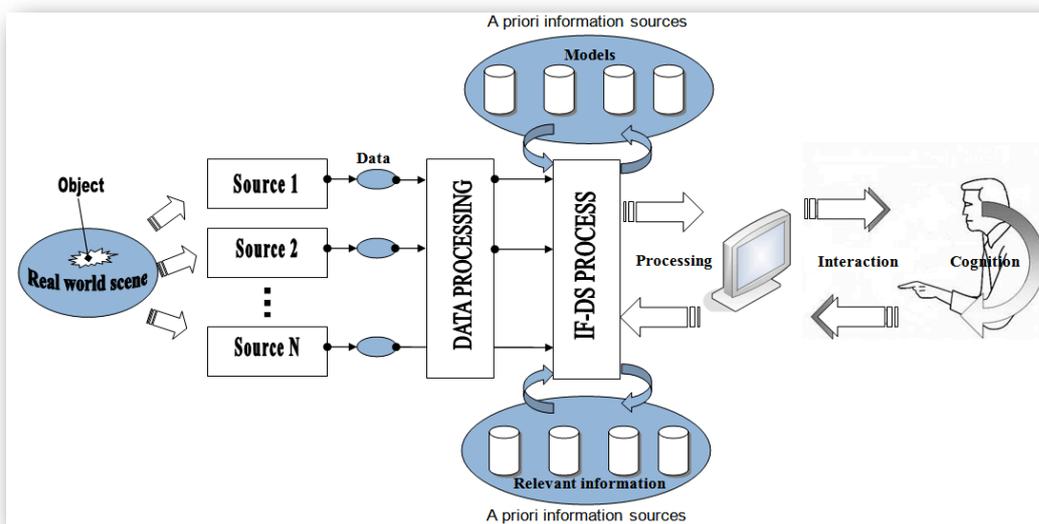


Figura 2
Un processo complessivo per il trattamento delle informazioni

La Figura 2 mostra il processo complessivo di trasformazione dell'informazione, a partire dai dati catturati dai sensori, fino ad arrivare ad indicazioni puntuali che un utente può usare per prendere decisioni. Esso si ispira ad uno dei più diffusi modelli di IF, denominato JDL (si veda il Riquadro 2) ed evidenzia tre importanti aspetti da tenere in considerazione nelle applicazioni di interesse:

- a. La necessità di pre-elaborare i dati grezzi provenienti dai sensori installati sul campo. Ciò è legato al livello di eterogeneità dei sottosistemi di monitoraggio, che possono variare dai sensori di temperatura alle telecamere intelligenti, e al livello semantico dell'informazione da essi fornita. A seconda dei casi, la pre-elaborazione può essere presa in carico dai sottosistemi stessi, dai sistemi di integrazione e gestione di più tecnologie (che nel dominio della sicurezza fisica prendono il nome di sistemi PSIM, *Physical Security Information Management*), o anche immediatamente prima della fusione delle informazioni.
- b. L'identificazione di differenti livelli di capacità (Figura 3). Infatti, la combinazione dei dati è finalizzata alla valutazione di: i) stati, attributi o identità delle singole entità monitorate; ii) le mutue relazioni tra le entità monitorate, anche rispetto all'ambiente circostante; iii) gli stati futuri e le proiezioni a partire dalla situazioni riconosciute, al fine di accertare il possibile accadimento di minacce e il relativo impatto.
- c. La realizzazione di un costante processo di affinamento, il quale può essere guidato dall'utente e/o automatico. In particolare, l'utente può contribuire al processo di IF, supportando l'aggiornamento dinamico delle informazioni definite a priori, ovvero dei modelli e parametri della fusion. Questo affinamento è essenziale per consentire, ad esempio, l'uso di informazioni sempre in linea con le indicazioni provenienti dal campo.

Nella Figura 2, ogni sorgente (*Source*) può rappresentare un singolo sensore (a prescindere dal proprio livello di "intelligenza"), un più complesso sottosistema di monitoraggio (ad esempio multimodale), o ancora un operatore umano. Successivamente, in base al tipo di informazione fornita in uscita, può essere richiesta un'elaborazione preliminare (*Data Processing*), oppure l'output delle sorgenti può entrare direttamente come input del processo *IF-DS*. Il processo combinato di IF e DS considera come ulteriori input una base di conoscenza, rappresentata da modelli di rilevamento, e uno o più database, impiegati per registrare informazioni rilevanti come i feedback degli utenti (si veda il Riquadro 3). L'interazione con l'utente consente di conciliare il processo cognitivo automatico con quello umano, permettendo la comprensione di eventi e situazioni, il riconoscimento di fenomeni emergenti, e l'adozione di più precise contromisure.

Dal punto di vista applicativo, il paradigma presentato può essere implementato da un sistema integrato per la sorveglianza e la gestione di situazioni critiche, le cui capacità dovrebbero includere:

- l'interfacciamento con diverse piattaforme di monitoraggio;
- il rilevamento di eventi di interesse nell'ambiente monitorato;
- l'invio di segnalazioni agli operatori dei centri di controllo, in modo da supportare procedure di emergenza o anche l'attivazione di risposte automatiche.

Dal momento che il sistema è potenzialmente in grado di "inondare" gli operatori con una mole di informazioni ingestibile, la capacità di correlazione degli eventi è estremamente importante, per consentire agli operatori di concentrarsi sulle notifiche significative, ovvero indicative di potenziali minacce. Un framework integrato comprensivo di tutte le capacità appena descritte sarebbe dunque in grado di rilevare fenomeni di una certa complessità con elevata affidabilità ed efficienza. Il problema della correlazione di eventi è stato ampiamente studiato nella letteratura scientifica ed un'ampia classe di soluzioni è stata definita. Tuttavia, i principali risultati hanno interessato applicazioni diverse dalla sorveglianza, come il rilevamento di intrusioni in reti informatiche o il riconoscimento di condizioni specifiche nelle cosiddette basi di dati attive. Nel contesto della sicurezza fisica, le capacità dei sistemi *legacy* sono molto limitate nell'analisi e nell'interpretazione dei dati in tempo reale. La scarsità di applicazioni in questo settore è motivata dalla mancanza di approcci che siano al tempo stesso efficaci e facili da implementare: il riconoscimento di situazioni in evoluzione, basato su modelli di minaccia definiti a priori e facili da aggiornare anche da parte degli operatori, è un obiettivo tutt'altro che semplice da raggiungere. E tuttavia questa capacità è necessaria quando un elevato numero di entità monitorate è coinvolto in complesse relazioni spazio-temporali. Assumendo che gli scenari di minaccia siano decomponibili, durante l'analisi del rischio, in fasi elementari che si susseguono con un ordine abbastanza predicibile, è possibile

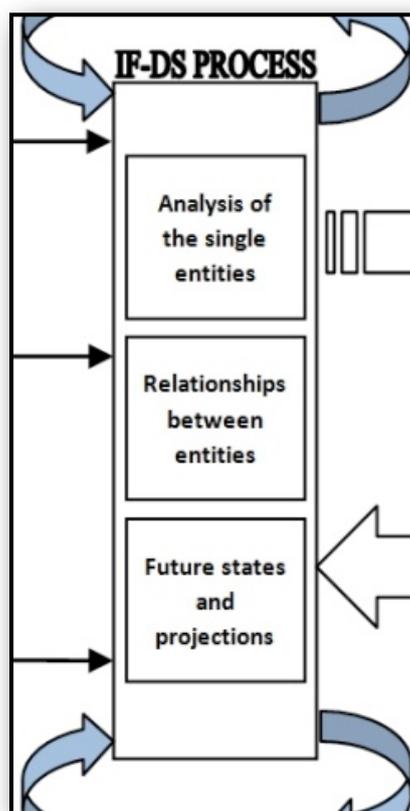


Figura 3
Dettaglio sul processo di IF e DS

pensare ad una correlazione logica, spaziale e temporale degli eventi. Al contempo, occorre prevedere tecniche capaci di garantire la “resilienza” rispetto a possibili errori di modellazione o mancati rilevamenti (es. dovuti a guasti dei sensori). Da questo punto di vista, un approccio promettente consiste nell’impiego di tecniche di riconoscimento euristico, basate su analisi di similarità degli scenari rilevati con i modelli di minaccia definiti a priori [10].

6. Conclusioni e problemi aperti

In questo articolo si sono passate in rassegna alcune tematiche che ruotano intorno alla problema della integrazione, sintesi e gestione delle informazioni provenienti dai sistemi di sorveglianza. Come si è visto, si tratta di approcci che hanno avuto uno sviluppo relativamente recente e che sono per loro natura spiccatamente multidisciplinari, combinando ricerche in ambiti di *signal processing* a livello sensoriale, *multi-sensor information fusion* per sistemi distribuiti, intelligenza artificiale (classificazione, *pattern recognition* e sistemi esperti), modellistica predittiva di affidabilità e prestazioni del rilevamento, studi ergonomici e di *human-factors* sugli operatori.

L’importanza di un meccanismo di fusione delle informazioni efficace, efficiente e flessibile risiede nel fatto che esso risulterebbe applicabile ai più svariati sistemi di supervisione e monitoraggio diagnostico, usati in molteplici ambiti applicativi, dal momento che questi condividono lo scopo abbastanza generico di fondere dati sensoriali eterogenei per supportare le decisioni degli operatori.

La ricerca in quest’ambito ha ricevuto una notevole spinta dalla consapevolezza delle difficoltà di gestione dei sistemi SCADA (e similari) in applicazioni reali, per la mole ingestibile di allarmi ricevuti mediamente ogni giorno da ciascun operatore, che può arrivare a molte centinaia, di cui solo una piccola percentuale realmente significativi. Un altro fattore fortemente motivante è legato alla necessità di rendere “intelligenti” i sistemi di sorveglianza per la *homeland security*, al fine di controllare i comportamenti degli individui (anche nell’interazione con Internet) per prevenire eventuali attacchi (eventualmente di tipo *cyber*), senza al contempo violarne la *privacy*.

Un grosso contributo allo sviluppo di tecniche di sorveglianza avanzata è dovuto al trasferimento di conoscenze, modelli e tecnologie da altri ambiti, quali quello della difesa, degli *Intrusion Detection Systems* (IDS) per le reti di calcolatori oltre che del *data mining* e delle basi di dati attive. Ad oggi, purtroppo, nonostante le apparenti similitudini, nessuno di questi ambiti è riuscito a fornire qualcosa di facilmente, direttamente ed esaustivamente utilizzabile per gli scopi descritti in questo articolo.

Diversi sono i problemi ancora da risolvere nell’ambito della fusione delle informazioni applicata alla sorveglianza. Tra questi vale la pena citare la possibilità di specifica degli scenari tramite modelli ontologici (si veda ad es. il

riferimento [10]) facili da mantenere, che consentano una trasformazione automatica in opportuni modelli di rilevamento *real-time* la cui complessità è nascosta agli utenti finali. Un altro punto che merita approfondimenti è quello della gestione dell'incertezza/incompletezza nella rappresentazione delle situazioni di interesse e della potenziale inaffidabilità nel rilevamento sensoriale, che può essere basata su logiche *fuzzy* ed approcci euristici, supportati da taluni formalismi di modellazione. Infine, molto interessante è la problematica di gestione dell'apprendimento della struttura e/o dei parametri dei modelli di rilevamento, con un desiderabile aggiornamento dinamico in base ai feedback degli operatori (una sorta di *supervised learning*, non necessariamente basata su reti neurali o bayesiane).

Riquadro 1 – L'Information Fusion

La fusione delle informazioni, o **information fusion**, è un'area di studio e di ricerca che ha ricadute estremamente significative in diversi domini applicativi, tra cui è possibile citare: difesa militare, previsioni meteorologiche e finanziarie, diagnostica medica, allerta precoce di calamità naturali, etc. L'importanza che i processi e le tecniche di information fusion assumono nella progettazione e implementazione di una sempre maggiore varietà di sistemi è dovuta alla crescente mole di informazioni generata da sorgenti eterogenee (siano esse sensori, satelliti, social networks, o osservatori umani) e alla sempre più pressante necessità di trasformare tali informazioni in *conoscenza*, al fine di supportare i processi decisionali con sempre maggiore efficacia ed efficienza. L'information fusion richiede pertanto *“l'integrazione e la sinergia di informazioni circa il comportamento di uno specifico sistema, provenienti da sorgenti differenti, allo scopo di supportare decisioni e azioni relative al sistema stesso. L'information fusion include nel suo ambito teorie, tecniche e strumenti il cui obiettivo è sfruttare la conoscenza implicita nelle informazioni acquisite da più sorgenti e nelle relazioni tra di esse”* (fonte: www.isif.org). Diversi sono i convegni e le riviste che trattano tematiche attinenti, tra i quali il riferimento probabilmente più significativo è rappresentato dalla rivista *Information Fusion* di Elsevier, a cui si rimanda per ulteriori approfondimenti sullo stato dell'arte in ambito sia teorico che applicativo.

Riquadro 2 – Il modello JDL

Uno dei modelli più diffusi per classificare il processo di fusione nasce dal U.S. Joint Directors of Laboratories (JDL), da cui prende il nome [12]. Nato negli anni '80 dal lavoro del Data Fusion Working Group, il modello JDL non descrive un flusso di attività, ma piuttosto individua categorie di funzioni che, a partire dai dati grezzi forniti dai sensori, consentono di pervenire ai diversi gradi di inferenza. Nella Figura 4, sono evidenziate le due principali categorie, cui appartengono funzioni divise in sottoclassi (o Livelli): la prima categoria è costituita dalle funzioni di *assessment* dei dati provenienti dalle sorgenti (Livelli da 0 a 3), la seconda da funzioni di *refinement*, alle quali è demandata la trasformazione delle informazioni in conoscenza (Livelli 4 e 5). Infine, il Livello 5 "apre" all'interazione tra il sistema di fusione e l'operatore umano, preposto a prendere le opportune decisioni in merito alle situazioni in essere (HCI: *Human-Computer Interaction*)

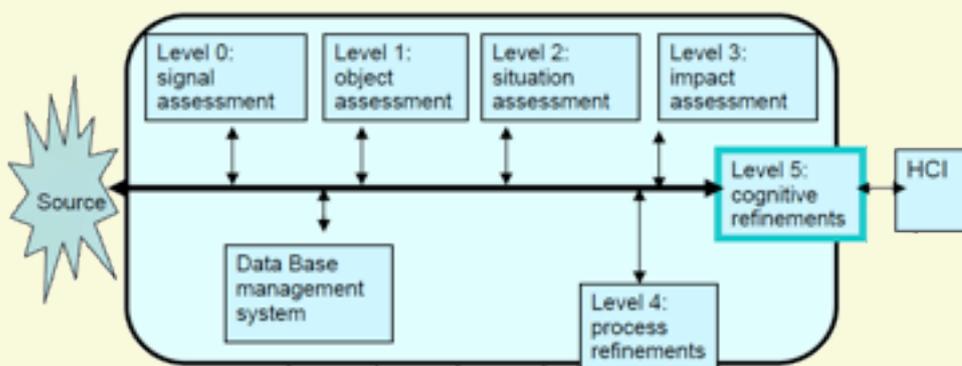


Figura 4
Il modello JDL

Riquadro 3 – Sorgenti e capacità nell'Information Fusion

Le informazioni relative al sistema reale che si sta osservando possono ricoprire un vasto orizzonte temporale. Come mostrato in Figura 5, in generale esse possono riferirsi al “presente” (i dati attuali provenienti, ad esempio, da sensori che stanno monitorando il sistema), al “passato” (attraverso la conoscenza pregressa, ad esempio serie storiche di eventi contenute in basi di dati), e al “futuro” (ad esempio, informazioni provenienti da simulazioni). Recentemente è stata riconosciuta l'importanza degli utenti nel contribuire efficacemente al processo di fusione delle informazioni.

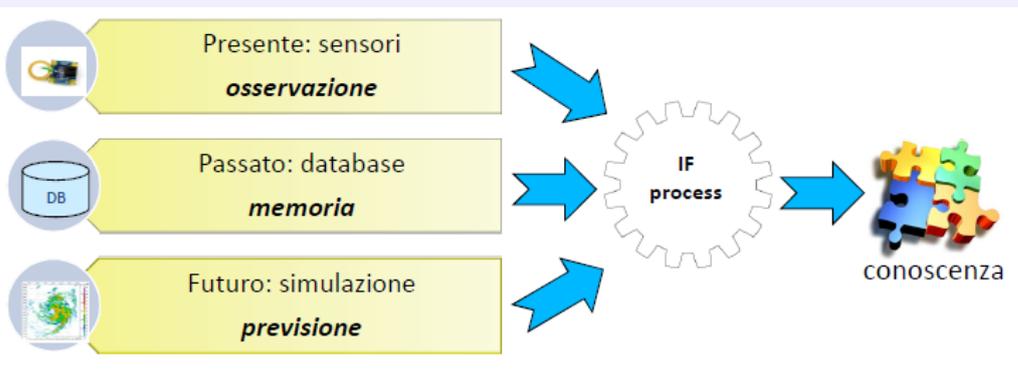


Figura 5
Schema di un sistema di Information Fusion

Il processo di fusione consiste in un progressivo meccanismo di inferenza che, a partire dai dati grezzi, permette di giungere: 1) alla determinazione dell'esistenza di entità di interesse; 2) alla loro localizzazione (posizione, velocità, etc.); 3) alla identificazione di tali entità; 4) alla definizione del loro comportamento; 5) alla comprensione della situazione in essere; 6) alla valutazione delle minacce che possono determinarsi dall'evoluzione futura (Figura 6).

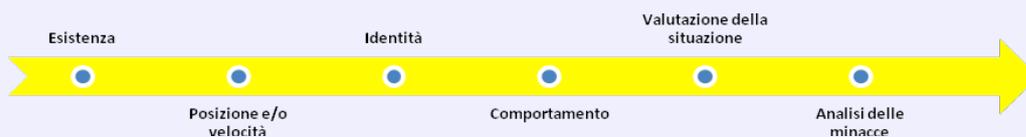


Figura 6
Livelli di inferenza

Riferimenti bibliografici

- [1] Atrey, P. K., Hossain, M. A., El-Saddik, A., Kankanhalli, M. S., *Multimodal fusion for multimedia analysis: a survey*, *Multimedia Syst.*, 16(6), 2010: pp. 345-379.
- [2] Bocchetti, G., Flammini, F., Pappalardo, A., Pragliola, C., *Dependable integrated surveillance systems for the physical security of metro railways*, in: Proc. of Third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC '09), Como (Italy), 2009: pp. 1-7.
- [3] Bossé, E., Guitouni, A., Valin, P., *An Essay to Characterise Information Fusion System*, in: Proceedings of the International conference of Information Fusion. Florence, Italy, 2006.
- [4] Castro, J.L., Delgado, M., Medina, J., Ruiz-Lozano, M.D., *Intelligent surveillance system with integration of heterogeneous information for intrusion detection*, in: *Expert Systems with Applications*, Vol. 38, No. 9, 2011: pp. 11182-11192.
- [5] Cesarini, M., Ghezzi, C., Tanca, L., Schreiber, F., *Reti di microdispositivi intelligenti*, in: *Mondo Digitale*, n. 17, 2006: pp. 47-55.
- [6] Cohen, J., Cohen, H. H., *Enhancing forensic human factors/ergonomics analyses using digital surveillance video*, in: Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting, 2007: pp. 1129-1132.
- [7] Cranor, L. F., Garfinkel, S., *Security and Usability. Designing Secure Systems that People Can Use*, O'Reilly Media, 2005
- [8] Cucchiara, R., *La visione artificiale per la videosorveglianza*, in: *Mondo Digitale*, n. 27, 2008: pp. 39-47.
- [9] Flammini, F., Mazzocca, N., Pappalardo, A., Pragliola, C., Vittorini, V., *Augmenting Surveillance System Capabilities by Exploiting Event Correlation and Distributed Attack Detection*, in: Proc. of Intl. Workshop on Security and Cognitive Informatics for Homeland Defence (SeCIHD'11), co-located with ARES'11, Springer LNCS 6908, 2011: pp. 191-204.
- [10] Flammini, F., Pappalardo, A., Pragliola, C., Vittorini, V., *A robust approach for on-line and off-line threat detection based on event tree similarity analysis*, in: Proc. of 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS'11), Klagenfurt, Austria, Aug. 30-Sept. 2, 2011: pp.414-419.
- [11] Flammini, F., Pappalardo, A., Vittorini V., *Challenges and emerging paradigms for augmented surveillance*, in: *Effective Surveillance for Homeland Security: Balancing Technology and Social Issues*, CRC Press, 2013: pp. 169-198
- [12] Hall, D.L., *Multisource information fusion for critical infrastructure situation awareness*, in: *Critical Infrastructure Security: Assessment, Prevention, Detection, Response*, WIT Press, 2012: pp. 267-277
- [13] St. John, M., Risser, M. R., *Sustaining vigilance by activating a secondary task when inattention is detected*, in: Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting, 2009.
- [14] Wickens, C., Dixon, S. *The benefits of imperfect diagnostic automation: a synthesis of the literature*, in: *Theoretical Issues in Ergonomics Science*, 8(3), 2007: pp. 201-212.
- [15] Zhu, Z., Huang, T.S., *Multimodal Surveillance: Sensors, Algorithms and Systems*, Artech House Publisher, 2007.

Biografia

Francesco Flammini ha ottenuto la laurea con lode (2003) ed il dottorato di ricerca (2006) in Ingegneria Informatica presso l'Università di Napoli Federico II. Dal 2003 lavora in Ansaldo STS (Finmeccanica) come progettista e ricercatore, occupandosi prevalentemente di verifica dei sistemi di controllo, protezione delle infrastrutture, e Innovation Network. Ha tenuto come professore a contratto corsi di informatica ed ingegneria del software. E' autore di numerosi articoli scientifici pubblicati su riviste, libri e atti di congressi internazionali. Svolge attività editoriali per libri e riviste sul tema dei sistemi sicuri ed affidabili ed è nel comitato di programma di diversi convegni internazionali, tra cui SAFECOMP. E' un ACM Distinguished Speaker ed un IEEE Senior Member.

E-mail: francesco.flammini@ieee.org

Alfio Pappalardo ha conseguito nel 2013 il dottorato di ricerca in Ingegneria Informatica ed Automatica presso l'Università di Napoli Federico II, con borsa di studio finanziata da Ansaldo STS. Nel Dicembre 2008 ha conseguito la Laurea in Ingegneria Informatica presso la stessa università, e nell'Ottobre 2009 il Master in Homeland Security presso l'Università di Bologna Alma Mater Studiorum. Dal 2009 si occupa di attività di ricerca nei settori della protezione delle infrastrutture critiche e della security nei trasporti ferroviari e metropolitani, presso Ansaldo STS. Attualmente è coinvolto in progetti internazionali per soluzioni integrate di security in sistemi metropolitani.

E-mail: alfio.pappalardo@unina.it

Valeria Vittorini è professore associato presso l'Università di Napoli Federico II, dove è docente di Fondamenti di Informatica e di Programmazione. È autrice di numerose pubblicazioni scientifiche sull'utilizzo di tecniche formali di modellazione di sistemi reali in diversi ambiti applicativi. Afferisce al Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione, presso cui svolge la propria attività di ricerca, prevalentemente incentrata sullo studio di sistemi distribuiti e sullo sviluppo di metodologie e strumenti per l'analisi di sistemi critici.

E-mail: valeria.vittorini@unina.it

Quando l'Informatica si veste di Robotica

Giuseppe Albano

L'impiego di "Apparati e Ambienti Mediatori" a matrice robotica sin dalla Scuola dell'Infanzia rappresenta una ineludibile opportunità di introdurre la disciplina informatica come efficace supporto al processo di apprendimento: vediamo come.

Keywords: *Educational Robotics, Constructionism, Computational Thinking, Procedural Thinking, Metacognition.*

1. Introduzione

"[...] Ed allora ecco che nella scuola i computer li troverete sì, ma non su ogni banco, nemmeno in ogni aula, ma ben chiusi dietro una porta blindata in aule (laboratori) il cui uso e la cui gestione non è sempre né facile né programmabile. Laboratori definiti, a piacere, "informatici" o "multimediali", in cui potete trovare le più svariate configurazioni (nel senso della disposizione fisica) e dotazioni (in rete e no, con una stampante a testa o senza nessuna periferica connessa, a volte anche con un masterizzatore su ogni pc (perfetto per una produzione di massa)), casse semplici o amplificate, mini, maxi o integrate, comunque sempre in grado di far risuonare per i corridoi attorne polifonie risultanti dal noto motivetto che accompagna l'accensione scandito in canone casuale.

E davanti ai computer chi? Se siete esterno alla scuola, e avete seguito il mio consiglio di intrufolarvi, provate ad andare nell'orario di servizio, e con la scusa di stare pensando ad iscrivere vostro figlio (la scusa funziona, anche se non avete figli ...) a quella scuola, chiedete di poter vedere i laboratori. In clima di concorrenza, di corsa all'accaparramento di alunni la forza del marketing supera ogni remora, e non sarà difficile convincere chi eventualmente vi chiederà di

tornare dopo le lezioni ("Sa, sono sempre di fretta e se non era per quest'appuntamento saltato ..."). Chi avete trovato nel laboratorio informatico (o multimediale)? Tre le ipotesi.

La prima: nessuno. In tal caso chi vi accompagna o si sente in imbarazzo e accenna a scuse varie tipo "Che combinazione, è proprio l'unica ora buca della settimana. Lei è proprio fortunato, così non disturbiamo la lezione", oppure non si rende neanche conto della stranezza della situazione e passa ad esaltarvi la potenza delle macchine e di quanto è costato quel ben di Dio. Mio consiglio: non guardate le macchine, ma cercate appeso da qualche parte – di solito sulla porta – il regolamento e l'orario d'aula. Leggete il regolamento e la lista di "divieti", è un documento molto esplicativo di come la si pensa in quella scuola sull'uso didattico dei computer. Se poi trovate anche l'orario, siete stati davvero fortunati. Controllate allora quante ore d'uso sono davvero programmate e chiedete quali materie si svolgono in quelle ore.

Seconda ipotesi: alle macchine trovate uno o più professori e/o tecnici. Osservate cosa stanno facendo, se conoscete ed usate i computer ed Internet sarete ben in grado di capire da soli. Fate finta di niente – voi quelle macchine non le conoscete per nulla – e fatevi spiegare da loro i progetti in corso che i ragazzi svolgono in quel laboratorio.

Terza ed ultima ipotesi: c'è una classe al lavoro. Bellissimo! Un ottimo campo di osservazione che analizzato con attenzione vi potrà portare a tre scenari possibili, rispetto alla normale attività didattica che si svolge nelle spoglie aule "normali". Potrete vedere qualcosa di uguale, di peggiore o – infine – di migliore rispetto all'oggetto ultimo di tutto questo parlare di tecnologie: l'insegnamento e l'apprendimento."

Il testo che introduce questo contributo è tratto da un articolo di Stefano Tommaso Donati del 2003; un articolo che non ha perso, anzi si può dire abbia consolidato la sua attualità in un panorama, quello della Scuola Italiana di base, oramai teatro di aperto confronto sul ruolo delle Tecnologie Digitali in ambito educativo.

L'articolo citato, in realtà, nell'intento di porre l'accento su comportamenti didattici inidonei allo sviluppo di competenze in materia digitale, per fortuna non rispecchia una realtà sicuramente non più attuale (l'articolo risale ad un decennio fa): molte sono al giorno d'oggi, infatti, le Scuole che adottano strategie didattico-educative tali da porre l'Allievo al centro del processo di apprendimento. In tali Scuole, l'impegno profuso dai Docenti è probante di una realtà ben diversa rispetto a quella che il Prof. Donati tratteggia nel contributo citato, e i laboratori multimediali sono una reale fucina di idee e attività al servizio di forme di didattica inclusiva e dagli evidenti tratti partecipativi.

Ancora una citazione, questa volta tratta da una tesi di laurea: quella del dott. Luca Leoni discussa presso l'Università di Bologna nell'anno accademico 2011-2012 ed avente per argomento "Competenze e competizioni di informatica:

valutazioni sperimentali". Nell'introduzione si legge: *"La scuola italiana rischia come il Titanic: naviga verso il futuro senza una chiara percezione dell'iceberg rappresentato dall'Informatica"*, ed ancora, l'Autore afferma che *"Nelle scuole sono presenti i computer, non l'Informatica."*, stigmatizzando in tal modo quale sia la percezione condivisa di quanto la Scuola sia ben lontana dall'affrontare e risolvere le problematiche annesse al ruolo e alle connotazioni che questa importante disciplina dovrebbe assumere nel contesto educativo.

Non meravigli la quasi totale assenza di note e riferimenti bibliografici di questo contributo: è intenzione di chi scrive disporsi ad un confronto di idee con il Lettore, cui è demandato il compito di valutare l'argomento descritto sulla base delle argomentazioni addotte, rapportate alla propria esperienza e, semmai, fornire attraverso una critica costruttiva, suggerimenti utili a migliorare la loro efficacia. Un confronto diretto, insomma, senza intermediari culturali che pure sono presenti fra le righe. Per crescere insieme, così come nello spirito delle stesse esperienze qui illustrate.

2. Il contesto

Quando parliamo di Scuole del primo ciclo dell'istruzione, parliamo di Fascia dell'Obbligo, oramai comprensiva dei primi due anni del biennio secondario superiore. Al suo interno, però, è determinante la presenza di una fascia d'età – quella dai quattro ai dieci anni – in cui la parola "insegnamento" dovrebbe essere sostituita dall'espressione "processo di apprendimento". Un'espressione ricca di suggestioni e spesso estesa, come peraltro è giusto che sia, all'intero percorso scolastico, ma che nei fatti in questo particolare periodo della vita dell'Individuo assume tinte forti più che in ogni altro.

Non ho parlato, si badi bene, di età scolare quanto di età cronologica, inserendo nel quadro dell'intervento educativo "intenzionale (quello, per capirci, in cui la Scuola interviene in modo intenzionale) anche il periodo prescolare affidato alla Scuola dell'Infanzia. Un segmento che sempre più assume legittima cittadinanza nel contesto scolastico per la sua forte valenza educativa.

È questo il periodo che intendiamo prendere in considerazione: dai quattro ai dieci anni il Bambino, lungi dall'essere considerato un "adulto in miniatura", condensa potenzialità che lo pongono in grado di forgiare atteggiamenti mentali e cognitivi che lo accompagneranno lungo tutto l'arco della propria vita. Ed è proprio in questo periodo che l'Informatica, quella "vera", può apportare il suo determinante contributo trasversale, sviluppando e consolidando atteggiamenti positivi e positivamente critici verso forme di apprendimento efficaci quanto permanenti.

In tal senso, l'informatica può essere considerata come un'inesauribile fonte di strumenti, attrezzi mentali a disposizione del Soggetto Educando; strumenti semplici nella loro essenza, tali da consentire di manipolare e modellizzare le

diverse situazioni per poterle padroneggiare e risolvere; Un vero e proprio ambiente di apprendimento, in cui l'errore è considerato come momento essenziale per progredire verso nuove conoscenze.

Perché ciò si realizzi, è necessario però valutare l'essenza di questa disciplina, riconducendo l'Informatica alla sua vera essenza, staccandola dal concetto prevalente per cui Informatica e Computer si identificano, così come pare sia per la Scuola, ignorando l'assunto secondo cui *“L'informatica non riguarda i computer più di quanto l'astronomia riguardi i telescopi (Edsger Dijkstra, citazione attribuita)”*.

3. informatica e metacognizione

Con l'espressione “metacognizione”, termine che può assumere significati diversi in relazione al contesto in cui viene utilizzato, possiamo sostanzialmente indicare l'attività di una mente che riflette su se stessa e sui processi che è in grado di innescare quando si trova ad affrontare una nuova conoscenza.

Gli atteggiamenti metacognitivi sono acquisiti in età infantile (3-4 anni), ed assumono nell'Individuo adulto un carattere solitamente inconsapevole, in altri termini, Egli riflette sulla conoscenza senza sapere di farlo, ma contestualmente questo lo aiuta a comprendere meglio ciò che conosce. Possiamo quindi affermare che la metacognizione è uno strumento che agevola ogni forma di apprendimento, indipendentemente dal contesto in cui l'argomento dell'apprendere può essere collocato. Ma cosa c'entra l'Informatica in tale processo? Qual è il suo ruolo nell'agevolare o sostenere i processi metacognitivi?

Una volta chiarito l'assunto iniziale – e cioè che l'Informatica non può e non deve essere confusa con il computer, che rispetto alla prima è da considerarsi come uno dei suoi prodotti, forse il più emblematico e rappresentativo – dobbiamo ricondurre il significato di questo termine al suo valore disciplinare.

Essa è sicuramente una scienza, il cui ambito d'indagine è l'informazione e i modelli che ne possono consentire la gestione in modalità automatica. Per fare questo, l'Informatica adotta strutture “elementari”, come gli algoritmi, i quali, riduttivamente, altro non sono se non rappresentazioni di processi scomposti in parti elementari. Attraverso architetture complesse di elementi algoritmici e codici (il codice binario è alla base dei sistemi informatici), l'Informatica perviene alla realizzazione di “modelli” che, applicati a sistemi automatici in grado di gestirli (l'hardware) consentono la gestione automatica di informazioni (i software), dando origine al quel sistema complesso che è un computer.

Un panorama, quello delineato, che appare ben lontano dalla Scuola dell'Infanzia e dalla Scuola Primaria come possibile ambiente in cui tali concetti, apparentemente alla portata solo di studi universitari, possano essere applicati a

sostegno di processi educativi e di apprendimento. La domanda è quindi: come è possibile introdurre l'Informatica in un tale contesto perché risulti utile ai processi di apprendimento propri di questa fascia d'età e, soprattutto, con quali obiettivi?

4. La memoria procedurale

Con questa espressione (la memoria procedurale viene anche definita "memoria implicita") si suole indicare una particolare tipologia di memoria, tipica dell'essere umano, in grado di "immagazzinare" le procedure tipiche di comportamenti di tipo complesso, sino a renderli automatici e spontanei. Un esempio significativo è quello riconducibile alla guida dell'auto, in cui si pongono in essere una serie di operazioni coordinate fra loro che ne consentono il controllo in ogni situazione. Inizialmente, il neo-patentato si pone alla guida della propria automobile con fare attento e meticoloso, impegnandosi a ricordare i consigli dell'Istruttore e rispettando rigidamente la sequenza di operazioni che ne consentono l'avvio, il movimento e la sosta. Gradualmente, egli diventa sempre più sicuro di sé, sino al punto in cui quasi non si accorge di quello che sta facendo: in questa fase, la procedura è stata completamente acquisita inducendo alla guida comportamenti spontanei quanto inconsapevoli.

Di tali esempi è piena la nostra vita di tutti i giorni; per accorgerci di quanto la memoria procedurale sia in grado di supportarla, basterà soffermarsi sui gesti e i comportamenti più usuali, quali il camminare, fare una telefonata, preparare il caffè.... tutto ciò che, in sintesi, richiede operazioni precise che inducano verso un obiettivo altrettanto preciso attraverso una sequenza di operazioni. È importante, però, non confondere i comportamenti procedurali con quelli riflessi o spontanei, frutto di reazioni all'ambiente che ci circonda.

Col tempo, la memoria procedurale potrebbe risultare compromessa a causa del sopraggiungere di deficit mentali o malattie; in tal caso, il Soggetto andrebbe incontro a gravi deficit, sino alla perdita di autonomia motoria, con conseguenze facilmente immaginabili.

5. Dalla memoria procedurale al pensiero procedurale

Analizzare e riprodurre i comportamenti procedurali, realizzando modelli applicabili in situazioni diverse, è compito dell'Informatica, che assume un valore metacognitivo quando contribuisce a sollecitare atteggiamenti mentali analitici nei confronti di situazioni cognitive diverse. In tal senso, è possibile parlare dello sviluppo di un pensiero di tipo procedurale, vale a dire di un vero e proprio atteggiamento autonomo e spontaneo del pensiero cognitivo rispetto ad una nuova conoscenza.

È questo un processo che si innesca naturalmente nel Soggetto sin dalla più tenera età (3 – 4 anni), tale da divenire del tutto inconsapevole e spontaneo col

passare del tempo. Se però, attraverso la mediazione di un intervento esterno e intenzionale, il Soggetto esercita tale forma di pensiero in situazioni e contesti esperienziali differenti, apprezzandone l'efficacia in termini di economia e produttività dei risultati, ecco che Egli lo applicherà in tutte quelle situazioni che richiedono di ricercare soluzioni ad un problema, quale che sia la natura di quest'ultimo: è quello che comunemente viene definito "Problem Solving", una strategia che travalica i confini del pensiero logico-matematico per "invadere" positivamente altri settori della conoscenza e dell'esperienza, definiti e definibili, per l'appunto, "problematici" laddove richiedano attività di analisi, di ricerca e di individuazione di risposte.

Pensiero procedurale e Problem Solving, da intendersi rispettivamente come strategia d'intervento e campo applicativo, si fondono in questo modo in un atteggiamento mentale che è alla base di quel processo di educazione permanente ("Imparare ad imparare") da più parti e universalmente ritenuto come "competenza chiave" per il corretto inserimento nell'attuale contesto economico, culturale e sociale del nostro Paese, in linea con gli orientamenti comunitari.

È quindi un problema non soltanto educativo quello che la Scuola si trova ad affrontare, recuperando quel ruolo di "maestra di vita" a suo tempo attribuitole e da troppo tempo, forse, usurpatole da altre agenzie educative (tecnologie e mass media); un compito arduo ed elevato, ma se vogliamo proprio dell'Istituzione scolastica, soprattutto in tempi come gli attuali, non sempre caratterizzati da messaggi positivi.

6. Apparati e ambienti mediatori

Il pensiero, nella sua attività cognitiva, è favorito dalla mediazione di Oggetti e Ambienti. È questo, in estrema sintesi, il pensiero pedagogico di Reuven Feuerstein e Seymour Papert, le cui teorie sono alla base del concetto di Informatica Metacognitiva o, meglio, delle strategie ad essa sottese per una sollecitazione efficace delle strutture mentali interessate. Le teorie dei due Pedagogisti sottolineano l'importanza di un approccio spontaneo alla conoscenza, mediato, per l'appunto, dall'impiego di oggetti in grado di favorire e supportare il processo di apprendimento.

Parliamo quindi di "Apparati e Ambienti Mediatori", costituiti da Robot programmabili, ambienti di programmazione iconici e testuali, ambienti di progettazione di documenti ipertestuali: il tutto a sua volta mediato da un Docente non più "trasmettitore" di contenuti ma disposto ad assumere il ruolo di "Regista" delle operazioni che possono essere compiute in quello che, complessivamente, può essere definito come veri e propri "Ambienti di Apprendimento", in cui il Bambino procede alla scoperta di occasioni di nuovi ambiti di conoscenza, favorito, e non ostacolato, dallo stesso errore, che in tale contesto assume un ruolo necessario e positivo.

7. Il ruolo dell'errore nel processo di apprendimento

Ed è proprio l'errore, anzi, il "Signor Errore", come amava definirlo Maria Montessori, che è posto alla base del processo di apprendimento: sbagliando s'impara, affermavano i nostri Genitori, e non avevano assolutamente torto.

Contrariamente al valore negativo e sanzionatorio che caratterizza un approccio di tipo trasmissivo, tipico di molte scuole e Docenti, le strategie basate sull'errore come momento fondamentale dell'apprendimento ne esaltano il ruolo e la valenza quale "innesco" per procedere in forma autonoma e significativa verso nuove conoscenze. Autonoma, poiché il Soggetto, in presenza dell'errore, assume spontaneamente un atteggiamento di riflessione e analisi della situazione in cui l'errore stesso, manifestandosi, ha impedito il raggiungimento di un determinato obiettivo; significativa perché l'apprendimento generato dall'errore viene dal Soggetto interiorizzato e correlato con le conoscenze pregresse, generando una vera e propria "competenza" trasferibile in altre situazioni e contesti operativi.

Un esempio, per quanto banale, può ulteriormente chiarire questo concetto. Capita spesso di operare al computer che, in quanto macchina esecutrice e non intelligente, incapace cioè di pensare e discernere fra ciò che è giusto o sbagliato fare, si limita ad eseguire l'istruzione impartita dall'operatore, ad esempio eliminare un testo selezionato in precedenza a seguito della pressione di un tasto.

Una situazione comune a quanti sono abituati ad usare il computer per elaborare un testo, e come tale oggetto di specifica attenzione nei vari corsi di formazione sul word processing.

Eppure, per quanto il Docente insista sulle corrette operazioni da effettuare in presenza di un documento di testo elaborato al computer, prima o poi capita di impartire un comando sbagliato, con conseguenze spesso traumatiche per chi commette un errore di questo tipo (per fortuna non irreversibili negli attuali software di word processing).

Prescindendo dagli effetti emotivi prodotti ("oddio, e ora come faccio.... Ho perso il mio lavoro!), riflettendo con fare analitico sulle operazioni effettuate, l'operatore non tarderà ad apprendere che la pressione di quel tasto, ad esempio la barra spaziatrice, è da evitare in presenza di un testo selezionato, pena il vederlo scomparire, sostituito da un malinconico spazio vuoto; difficilmente, Egli commetterà ulteriormente la stessa operazione.

Nell'esempio riportato, il Soggetto ha fatto i conti con se stesso, Allievo e Insegnante allo stesso tempo, partendo dall'errore commesso per avviare un processo di riflessione altrimenti assente, e acquisendo una nuova conoscenza.

L'errore, quindi, lungi dall'essere considerato come segnale di incapacità o incompetenza, rappresenta l'ineludibile punto di partenza per procedere verso nuove conoscenze, oltre a costituire un punto di riferimento per una nuova indagine, tesa a verificare i comportamenti posti in essere sino alla sua scomparsa nel processo avviato dal Soggetto. Tanto più esso assumerà un valore educativo se affrontato in maniera spontanea e autonoma, ancorché accettato dal Soggetto come fattore positivo.

Nell'Informatica Metacognitiva, l'errore è perciò un fattore fondamentale, così come lo è l'attività, definita di debugging, che attraverso una vera e propria "caccia all'errore" accresce sempre più le competenze messe in campo dal Soggetto, sino a generare, in concorso con le conoscenze e le abilità acquisite nel tempo, quella competenza che gli consente di applicare un tale atteggiamento in contesti differenti.

8. Apparatì e ambienti a matrice robotica

In tutto questo, la Robotica Educativa (espressione utilizzata per indicare una metodologia basata sull'impiego di hardware e software robotico in attività di tipo educativo) rappresenta uno dei possibili approcci all'Informatica Metacognitiva; un approccio "mediato" dall'impiego di apposito hardware e software che, in linea con il pensiero costruzionista di Papert, consentirebbe al Soggetto educando di "manipolare" i concetti, in altre parole di renderli concreti e in diversa misura percepibili. Un piccolo robot che si muove in uno spazio percettivo strutturato (ad esempio un pavimento opportunamente predisposto attraverso una griglia di movimento); un software i cui elementi grafici siano in grado di richiamare l'esperienza manipolativa attraverso la presenza di icone facilmente riconducibili ad azioni concrete; l'uso, infine, di linguaggi di programmazione testuali, in altre parole di simboli vicini alla comprensione da parte della macchina ma anche del Soggetto (perché riproducenti forme di linguaggio consuete); tutto questo è in grado di "mediare" la conquista graduale di abitudini mentali, sino a radicarle in veri e propri comportamenti autonomi. E quanto più precoce è l'approccio a una tale strategia, tanto più sarà possibile favorire la costruzione di architetture mentali di tipo trasversale rispetto alle diverse esperienze, architetture spendibili in qualunque situazione che preveda l'analisi di situazioni, la risoluzione di problemi, l'attivazione di comportamenti cognitivi.

9. Un simpatico "antenato"

1979. La MB, azienda inglese di giocattoli che prende il nome dalle iniziali di Milton Bradley, suo ideatore e fondatore, immette sul mercato un nuovo giocattolo robotizzato: il Big Track (Figura 1).

Dalla avveniristica forma ispirata ad un fantascientifico mezzo blindato, il Big Track è dotato di tre assi di cui quello centrale eroga la trazione. Gli assi possono

ruotare in modo inverso, permettendo al giocattolo di ruotare su se stesso come un carro armato.



Figura 1
Il Big Track della MB

Ciò che lo rende però unico nel suo genere, rispetto ai giocattoli dell'epoca, è la modalità con cui è possibile programmarne i movimenti, vale a dire attraverso una tastiera posta sul dorso dello scafo (Figura 2); una tastiera che consente di memorizzare all'interno di un microchip una serie di azioni che vanno dal procedere avanti e indietro, ruotare secondo un angolo di rotazione definito dall'operatore, inserire pause fra un movimento e l'altro, persino sparare con un suggestivo raggio di luce emesso dal suo "cannone laser", a colpo singolo o continuo. Oltre ai due servomotori

che ne assicurano i movimenti, infatti, il Big Track era anche dotato di un piccolo cervello elettronico e di una minuscola memoria RAM, utile a memorizzare i comandi impartiti per poi eseguirli in una singola procedura (per questo era dotato di un tasto di avvio della stessa).

Un oggetto che del giocattolo vero e proprio, in considerazione dell'epoca in cui venne immesso sul mercato, aveva ben poco: ci si trovava di fronte ad un vero e proprio robot di tipo "stand alone (funzionante in modo indipendente)" chiaramente ispirato all'ambiente di programmazione LOGO di Papert. Purtroppo, all'epoca, il Big Track non ebbe molta fortuna, almeno nel nostro Paese; paradossalmente, a tanti anni dalla sua prima presentazione oggi esso può vantare numerosi Fan, soprattutto fra un pubblico inaspettatamente adulto, che ancora può comprarlo a prezzi molto bassi sui diversi siti d'asta internazionali.

Chi ora è impegnato a scrivere questo contributo ha avviato le sue esperienze in materia di Robotica educativa proprio grazie a questo giocattolo, portato a scuola da uno dei suoi Allievi al rientro dalle vacanze natalizie. Il piccolo non riusciva a farlo funzionare, per cui chiedeva aiuto al suo insegnante.



Figura 2
La tastiera dei comandi del Big Track; da notare i tasti per il movimento, la rotazione, il tasto di avvio della procedura (GO)

Immediatamente, come era prevedibile, l'oggetto attrasse l'attenzione degli altri compagni, ancor più, lo confesso, quella del loro insegnante, il quale ne intuì le potenzialità come centro d'interesse per numerose attività (era una seconda classe elementare). I bambini, messi da parte i loro quaderni e abbandonati i banchi, fecero circolo intorno al loro compagno e a me, in quel momento nominato suo "tutore" proprio dai suoi occhi: con accanto qualcuno che, per il ruolo rivestito, doveva per forza "saperne più di lui" lo faceva sentire più sicuro e più spavaldo nei confronti dei compagni. Chissà cosa avrebbe pensato se solo avesse intuito che, forse, ne sapevo quanto lui su quel "coso" e che morivo dalla voglia di giocarci.

Le domande e la ricerca di risposte furono immediati, come immediata si ravvisò l'esigenza di procedere con maggiore ordine e disciplina: una disciplina, una volta tanto, scaturita non dalla imposizione quanto da un bisogno sentito del gruppo. Non ci volle molto a capire che il veicolo eseguiva movimenti in avanti e all'indietro che, misurati, corrispondevano alla sua lunghezza. Così, per capire quanti passi avrebbe dovuto compiere per raggiungere, ad esempio, la parete in fondo all'aula (liberata ovviamente dai banchi e dalle sedie), il gruppo fu costretto a rilevare un'unità di misura, moltiplicandola per tante volte quante corrispondevano ai passi di un compagno che, spontaneamente, si era offerto di misurare con i suoi piccoli passi tale lunghezza....

Le attività continuarono, fino a quando uno dei bambini, il padrone del Big Track, lanciò una sfida a me e ai compagni: egli sarebbe riuscito a programmarlo esattamente, una volta stabilito il suo obiettivo, senza guardare. La sfida fu raccolta con entusiasmo e scetticismo da tutti: il primo a non credere nella riuscita dell'impresa ero proprio io.

Meraviglia delle meraviglie.

Spontaneamente, senza suggerimento alcuno, il Bambino si procurò un foglio di carta e una matita su cui, procurano di non essere visto, aveva tracciato il percorso assegnato dai compagni, accompagnato da una serie di "blocchi" che indicavano passi, pause e rotazioni. In una parola, aveva realizzato un algoritmo di procedura. Quel foglio a quadretti, spiegazzato e dai tratti ingenui, aprì le porte della nostra aula alla robotica e ad un modo diverso di fare scuola.

10. Dal Big Track al Bee Bot

Dopo aver cercato di sottolineare, nel paragrafo precedente, come le origini del presente siano sempre da ricercare nell'esperienza pregressa, passiamo ora ad elencare alcuni fra questi dispositivi hardware (Apparati) e software (Ambienti), attualmente in uso in ambienti educativi. Essi sono stati individuati non in base alla loro più recente produzione rispetto ad altri, quanto sulla scorta della loro collaudata valenza ed efficacia in campo educativo, e riferiti a precise fasce d'età, secondo lo schema che segue:

Fascia di riferimento	Apparato	Ambiente
4/7 anni	Bee Bot	Focus on Bee Bot
8/10 anni	Pro-Bot	LOGO - Robomind
11/13 anni	Scribbler	Scribbler

Sono, sia chiaro, apparati e ambienti indicati come orientativi, selezionati come già riferito in base all'esperienza condotta da chi scrive in diversi contesti operativi e resi disponibili in percorsi di formazione e aggiornamento destinati ai Docenti. Va precisato che la loro efficacia non è intrinseca, ma può variare in base all'uso che ne viene fatto: sistematico, episodico in ambienti di apprendimento appositamente realizzati e progettati o nel chiuso di laboratori. Tutti rivelano comunque una matrice comune: l'attenzione allo sviluppo e al consolidamento di forme di pensiero procedurale a matrice informatica e a valenza trasversale.

Ci limiteremo, in questa sede, ad una descrizione delle risorse elencate, essendo la trattazione delle strategie connesse al loro impiego didattico-educativo non riducibile allo spazio riservatoci. Basterà affermare, in questa sede, come l'impiego di apparati e ambienti mediatori presupponga da parte del Docente la disponibilità ad assumere un ruolo apparentemente contrastante con quello solitamente attribuito a chi insegna: da semplice "trasmettitore" di contenuti, dovrà predisporre a rivestire egli stesso il ruolo di "Regista" e "Mediatore" nei confronti del Bambino, affiancandolo nel processo di apprendimento e procedendo con lui alla scoperta e alla costruzione di nuove conoscenze. È questo lo spirito che anima le stesse teorie poste alla base di un atteggiamento costruzionista, lo stesso teorizzato da Papert e da quelli che egli chiama "Micromondi".

11. Il Bee Bot

Questa simpatica "Ape Robot" (Figura 3), prodotta dall'inglese TTS e distribuita in Italia da Media Direct attraverso il sito Campustore (www.campustore.it) rappresenta una intelligente risorsa mediatrice in grado di favorire l'approccio alla Robotica sin dai tre – quattro anni di età: ergonomicamente ben concepito (può essere agevolmente afferrato con una sola mano), equilibrato nel peso e realizzato con materiali atossici, il piccolo Robot può diventare nel giro di poco tempo il compagno di giochi preferito dal Bambino, ben presto compreso nel suo ruolo di programmatore in erba.



Figura 3
Il Bee Bot dell'inglese
TTS

Il Bee Bot è un dispositivo “Stand Alone”, programmabile cioè indipendentemente dal computer con cui non è interfacciabile. I comandi sono disposti sul dorso della scocca, immediatamente e intuitivamente individuabili da parte del bambino grazie a precisi simboli. È inoltre dotato di due servomotori collegati alle ruote indipendenti, che gli consentono di ruotare verso destra o sinistra facendo perno su ciascuna ruota, a seconda dell’istruzione impartita (la rotazione è di 90 gradi e non può essere variata). Oltre ai comandi di movimento e rotazione, il Bee Bot dispone del tasto “Go (vai)”, che corrisponde al tasto “INVIO” del computer e consente l’avvio della procedura precedentemente programmata dal suo operatore. Il corredo dei comandi è completato dal tasto “Pause”, utile ad inserire intervalli nel percorso di diversa lunghezza, e del tasto “Clear” che provvede a ripulire la sua memoria dai comandi impartiti in precedenza:

L’immagine (Figura 4) evidenzia chiaramente i comandi descritti. Nella parte inferiore della scocca sono collocati il vano batterie e i pulsanti di accensione/spegnimento e di attivazione/disattivazione di luci e suoni (gli occhi del Bee Bot si illuminano per indicare l’inizio e la fine di una procedura, e il robot può emettere suoni allo stesso scopo e ad ogni movimento effettuato).



Figura 4
Il pannello comandi del Bee Bot

12. Il Software “Focus on Bee Bot”

Al Bee Bot è abbinato uno specifico software, denominato “Focus on Bee Bot (Figura 5)”, che può essere utilizzato dal Bambino per guidare il robot in un ambiente virtuale, costituito da percorsi diversi per tipologia e difficoltà di esecuzione. Questi ambienti, definiti “tappetini”, raffigurano situazioni e contesti diversi, dal circuito di gara da effettuare schivando ostacoli alle strade di una ipotetica cittadina, all’isola del tesoro. Non mancano percorsi simbolici, costituiti da lettere, numeri e forme, che propongono diverse attività per un corretto approccio al simbolo iconico e alfanumerico.

Come è possibile evincere dalla figura 5, l’ambiente operativo è ricco di opportunità: oltre al campo d’azione del Bee Bot, infatti, il piccolo operatore può disporre di una tastiera per i comandi del tutto simile a quella reale posta sul dorso del robot, e di una finestra che gli consente di visualizzare, modificare o cancellare le istruzioni impartite in caso di errore, facilitando una attività che non esitiamo a identificare con il processo di “debugging” in uso quando una procedura algoritmica presenta degli errori. In questo modo, il Bambino si abituerà presto a considerare l’errore come fattore che di per sé non invalida un intero processo, e che una volta individuato rende l’intera procedura eseguibile.



Figura 5
Il software "Focus on Bee Bot"

A completare il quadro delle opzioni, il software consentirà di variare l'angolo di visuale, di usufruire di procedure "passo passo", di scegliere fra varie attività e opzioni di personalizzazione del robot e dei percorsi e così via. È possibile, grazie ad un editor, realizzare ulteriori percorsi utilizzando immagini o disegni precedentemente realizzati con un semplice software disponibile nello stesso Sistema Operativo (ad esempio Microsoft Paint), favorendo in tal modo anche un approccio mirato al computer.

È facilmente intuibile come le attività proposte dal software possano essere trasferite in seno ad attività manipolative e di gruppo, attraverso la realizzazione di "tappetini", cartelloni, situazioni simulate e giochi di ruolo: è un discorso, come precedentemente affermato, affascinante quanto meritorio di ulteriore trattazione relativa alle strategie di intervento, discorso che in questa sede non ci è consentito affrontare.

13. Il Pro-Bot

Non ci si lasci ingannare dall'aspetto "giocattolaio" di questo dispositivo. Sotto la sua livrea di piccola automobile, il Pro-Bot (Figure 6 e 7) nasconde un'anima informatica davvero interessante e di tutto rispetto.

Con il Bee Bot, questo dispositivo ha in comune solo il suffisso del nome. Ci troviamo di fronte ad una macchina programmabile di sicuro interesse per alcune sue caratteristiche, che rendono manipolabile il robot virtuale di matrice papertiana protagonista di LOGO, il migliore e a mio avviso intramontabile ambiente di programmazione espressamente concepito per scopi educativi e didattici.



Figura 6
Il Pro-Bot

Oltre a replicare le possibilità di programmazione offerte dal suo fratellino minore, il Bee-Bot, Il Pro-Bot rivela numerosi altri spunti e opzioni, grazie alla sua dotazione di bordo. Il device dispone infatti di un display LCD che consente la visualizzazione delle istruzioni impartite e replica la sintassi propria del linguaggio LOGO. Mette inoltre a disposizione del programmatore comandi più complessi, come il comando RPT (in LOGO corrisponde al comando "Ripeti"), mediante il quale si possono realizzare procedure iterative e nidificate in grado di far disegnare al robot poligoni regolari e irregolari (a tal proposito è dotato di un apposito alloggiamento utile ad ospitare un pennarello con cui esso, procedendo nei suoi movimenti e rotazioni, descriva su un foglio il percorso tracciato).

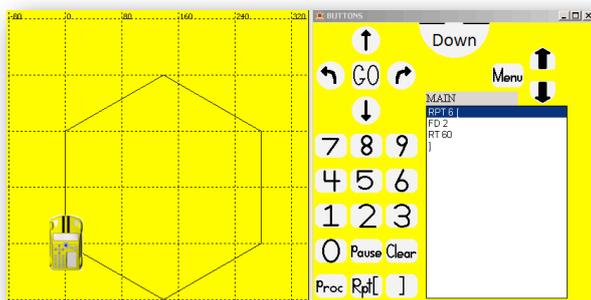


Figura 7
Il Pro-Bot in azione

Il pro-Bot è inoltre dotato di ben quattro sensori: due di contatto (anteriore e posteriore), un sensore di luce e uno di suono, tutti programmabili dall'operatore.

Inutile sottolineare che il Pro-Bot non è adatto ad essere utilizzato prima degli otto - nove anni, e che si propone senza soluzione di continuità con il Bee-Bot e relativo software, introducendo il

bambino in un ambito di programmazione più complesso, ricco di situazioni stimolanti a livello di analisi situazionale e Problem Solving e stimolando lo sviluppo di competenze di indubbio valore trasversale rispetto alle diverse situazioni cognitive e di apprendimento. Nonostante questo, il Pro-Bot, così come peraltro lo stesso Bee-Bot, non sono dispositivi di larga diffusione, coerentemente peraltro con la stessa Robotica impiegata in contesti metacognitivi. Ci si augura, col tempo, che la Scuola modifichi tale atteggiamento a favore di un'informatica maggiormente significativa, rispetto a quella diffusa, incentrata sull'uso del computer.

14. Lo Scribbler

Con lo Scribbler (Figura 8), illustrato nella figura che segue, chiudiamo la nostra veloce panoramica sugli Apparati e Ambienti Mediatori in grado di supportare percorsi di Informatica Metacognitiva sin dalla Scuola dell'Infanzia: una metodologia d'approccio indubbiamente attuale e che guarda all'Informatica non più come acquisizione di buone prassi produttive basate sull'impiego di applicativi, restituendole il valore di "Scienza del pensiero".



Figura 8
Lo Scribbler

Prodotto anche questa volta da un'azienda statunitense, la Parallax, lo Scribbler, letteralmente "Scarabocchiatore", è un robot programmabile per mezzo di un apposito software, questa volta non residente nella macchina, del tutto sprovvista di una interfaccia per comunicare con l'esterno (ad esempio un display). Niente paura: l'azienda produttrice mette a disposizione il software gratuitamente per quanti intendessero scaricarlo. Ovviamente, le due componenti sono interdipendenti, anche se lo Scribbler può comunque operare secondo una procedura autonoma, intrigante ma non suggestiva quanto la possibilità di programmarlo su un computer, scaricando poi il programma realizzato nella memoria del robot per vederlo all'opera.

La macchina porta a bordo una ricca dotazione, composta da ruote indipendenti (come sul Bee- Bot e sul Big Track, lo ricordate?), da un alloggiamento per il pennarello tracciante e da un robusto “cervello elettronico”. Ne completano la dotazione tre sensori di luce, due emettitori ad infrarossi accompagnati da un rilevatore di colori, apparati che gli consentono di interagire con l’ambiente. La comunicazione con il computer è assicurata dalla presenza di una porta seriale: ignoro se negli ultimi modelli è stata sostituita con una porta USB.

Quanto al software (Figura 9), è un programma molto ricco, basato sull’impiego di icone:

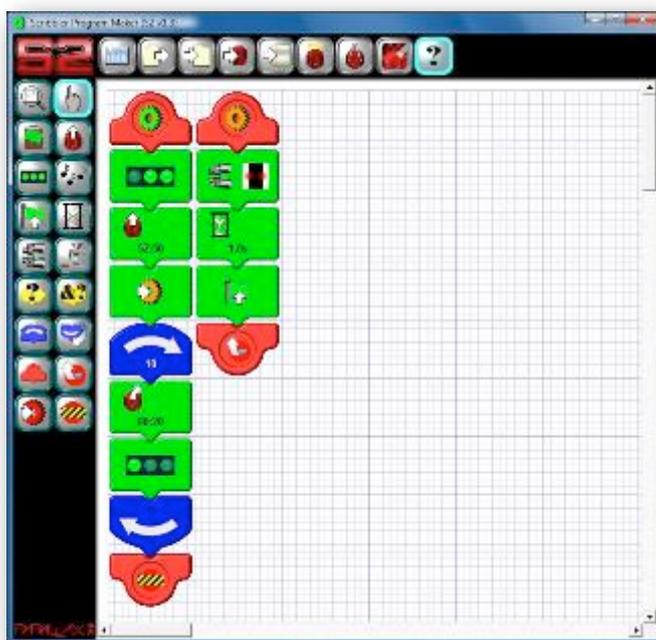


Figura 9
Il software di programmazione dello Scribblers

Per il suo tramite, l’operatore progetta una procedura complessa, costituita da movimenti e rilevamenti ambientali, può inserirvi cicli, richiamare altre procedure e così via. Un gioco “da duri” che, dopo le prime, iniziali quanto comprensibili incertezze, è in grado di coinvolgere Allievi e Docenti in un’attività di indubbio valore educativo. Il software è plurilingue, quindi anche in italiano.

Insieme al suo software, lo Scribblers può essere impiegato in attività laboratoriali sin dalla Scuola Primaria. Personalmente, lo vedo come risorsa corroborata dall’impiego preliminare degli altri due dispositivi precedentemente descritti (il Bee Bot e il Pro-Bot), ma non posso escluderne l’impiego anche in assenza di esperienze pregresse. Certo, sarà dura...

15. I software correlati

Una menzione a parte meritano i software che, pur non riferiti alle varie tipologie di hardware descritte e comunque disponibili nel panorama della “Robotica educativa (il virgolettato è d’obbligo in quanto trattasi di una definizione utile ad indicare questa tipologia di risorse; personalmente preferisco parlare di “Apparati mediatori”, non attribuendo a quella definizione un valore esaustivo e univoco rispetto all’argomento trattato in queste pagine)”, costituiscono veri e propri “Ambienti” in grado di suggerire attività diverse, tali da sollecitare forme di analisi e attenzione alle varie fasi di realizzazione di una procedura, per semplice o complessa che sia.

Dell’impiego di questi software molto si discute e si è discusso, soprattutto in merito alla loro efficacia e significatività educativa.

Possiamo dividerli sostanzialmente in due categorie: i software a interfaccia grafica e testuale. I primi risultano particolarmente diffusi, di facile impiego e di libera distribuzione. Questi software sono caratterizzati dalla presenza di icone che ne facilitano l’impiego, grazie alla possibilità di disporre di funzioni intuitive come il “drag and drop”. In pratica, l’operatore si trova di fronte ad una serie di “tessere” simili a quelle di un puzzle, ciascuna corrispondente ad una determinata funzione di un algoritmo. È il caso, peraltro, dello stesso software Scribbler, precedentemente descritto, comune però ad altri software come STARLOGO Robomind e lo stesso “Focus on Bee Bot”.

Se però in quest’ultimo l’interfaccia intuitiva diventa indispensabile, in considerazione della giovanissima età dei suoi fruitori, gli altri software sono frutto di una precisa scelta, quindi meritano un cenno a parte.

L’interfaccia grafica sicuramente favorisce, e di molto, la comprensione della struttura di un algoritmo; di contro, il fatto che l’operatore si limiti a spostare sullo schermo le tessere di un puzzle verificandone semplicemente la congruità con i movimenti e le funzioni da attribuire ad un robot virtuale, di certo non agevola la riflessione e quello che in ambito educativo si definisce “senso critico”, ovvero la capacità di analizzare e valutare l’efficacia di un processo, oltre a quella di applicare regole di tipo sintattico afferenti un linguaggio. Ciò non significa che si intenda attribuire a questi software una scala di valori commisurati alla loro efficacia educativa; semplicemente, se ne sottolineano limiti e vantaggi, affinché il Lettore possa orientarsi nelle sue scelte in forma consapevole e coerente con le sue necessità e obiettivi.

16. Il Robomind

Questo software (Figura 10) può essere definito la sintesi concettuale e operativa dei principi suesposti. Sviluppato in collaborazione con l'Università di Amsterdam, è un programma liberamente scaricabile e molto utile per introdurre il Soggetto in un ambiente di programmazione finalizzato a guidare i movimenti di un robot in un ambiente simulato.

Il software rende disponibili diverse opzioni: grazie all'operatore, il piccolo robot può muoversi avanti e indietro, disegnare strisce bianche o nere nello spazio operativo, ruotare verso destra o sinistra di 90 gradi. È persino in grado di afferrare e rilasciare oggetti, di guardarsi intorno e "prendere decisioni" sul da farsi al verificarsi o meno di determinate condizioni ambientali (ad esempio in presenza di ostacoli – vedi figura 11).

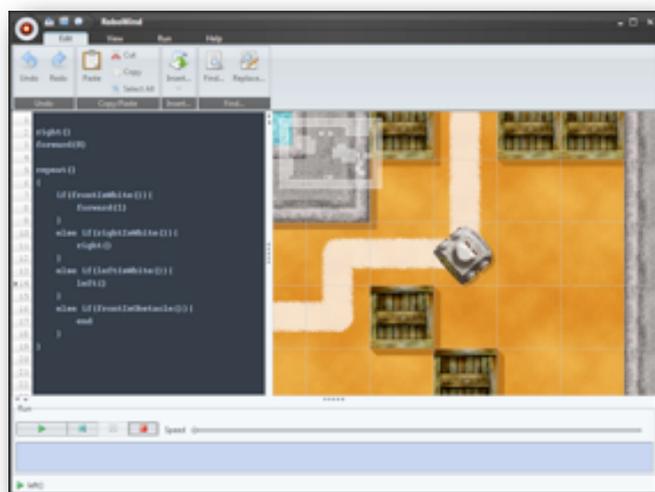
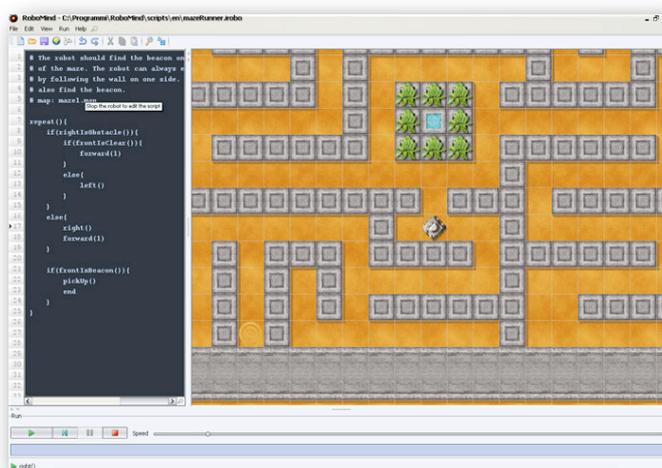


Figura 10
Il software Robomind

Figura 11
L'ambiente Robomind



Come è possibile intuire dalle immagini, l'ambiente Robomind integra la visuale grafica dello spazio in cui il robot si muove con una finestra in cui è possibile visualizzare le stringhe alfanumeriche corrispondenti a ciascuna istruzione impartita. Da qui, l'operatore può intervenire "criticamente" per operare modifiche procedurali o correzioni laddove decida di programmare il robot intervenendo direttamente sulla sintassi dei comandi da impartire.

Oltre alla modalità di programmazione, il software dispone anche della modalità "step by step": soprattutto nella fase iniziale, l'operatore può muovere il robot impartendo comandi singoli, in modo da verificare come questo risponda alle sue sollecitazioni passo dopo passo (vedi Figura 12).

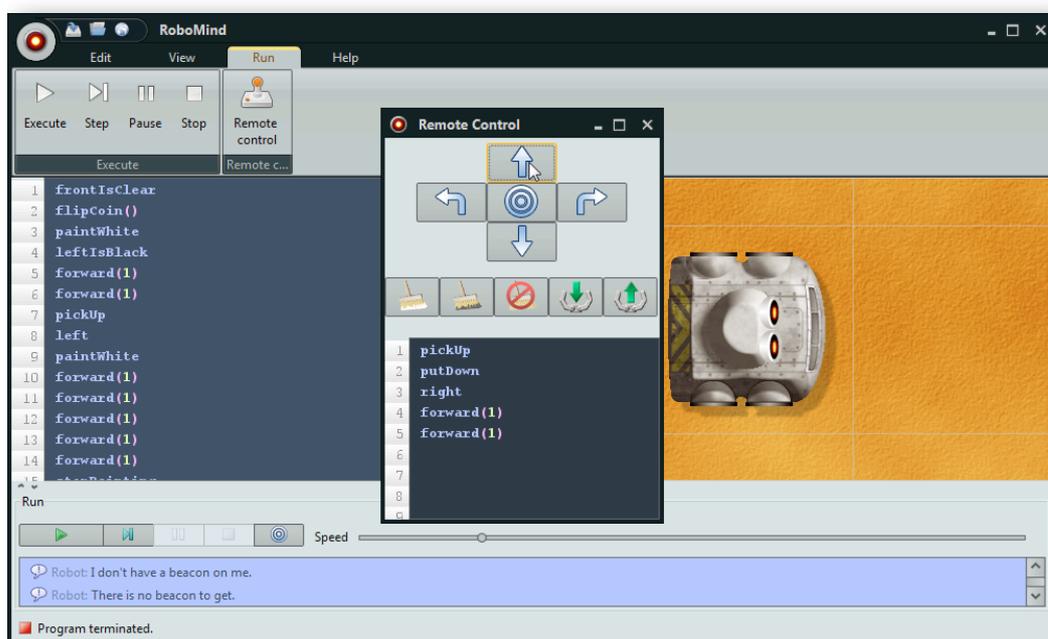


Figura 12
Il pannello comandi del Robomind

Il limite evidenziato dagli utilizzatori di questo software è costituito dalla scarsità dei comandi utilizzabili, pochi per la verità se si pensa al corredo di tipo grafico e iconico. In realtà, il Robomind è un software completo, che non indulge molto ad una programmazione intuitiva, sollecitando l'operatore a migrare ben presto verso un ambiente di programmazione vero e proprio, costituito da stringhe alfanumeriche. In tal modo si agevola sapientemente il passaggio ad un linguaggio di programmazione vero e proprio: il LOGO.

17. LOGO, Ali per la mente

È questo il titolo di una pubblicazione non proprio recente, autore lo stesso Seymour Papert, che accompagnava il lettore alla scoperta dell'ambiente LOGO e delle sue molteplici potenzialità educative. Denigrato da molti, da molti elogiato, il LOGO da molti anni sembra relegato alla preistoria dell'informatica, per la semplicità grafica e l'essenzialità che lo contraddistingue (ci si riferisce, naturalmente, alla versione papertiana). Eppure, così come il DOS, Sistema Operativo alfanumerico croce e delizia dei "vecchi" utenti del PC, a detta di molti oramai superato e obsoleto, costituisce in realtà la struttura portante degli attuali Sistemi Operativi a interfaccia grafica brandizzati Microsoft e targati Windows, così il LOGO è presente in tutti i software didattici riferibili all'ambiente robotico, non ultimo proprio il Robomind di cui abbiamo ampiamente dissertato.

Ma che cos'è?

Molti sapranno sicuramente rispondere a questa domanda; a quanti, fra coloro che leggono questo contributo, non avessero avuto occasione di "incrociarlo" in precedenti esperienze, risponderò con una semplice affermazione: LOGO è l'essenza della Robotica applicata ai processi di apprendimento, ma è anche il migliore approccio ai linguaggi di programmazione.

Basato sul linguaggio LISP, di antica memoria, LOGO vanta numerosi Fan, e non solo nel mondo della Scuola. Nasce nella mente di Papert in ossequio all'idea che non debbano essere i computer a governare i Bambini: semmai il contrario (e, badate, si era nel 1974: mai un'espressione di tanti anni fa potrebbe risultare più attuale).

La sua interfaccia grafica è molto semplice, come si può notare dall'immagine che segue:

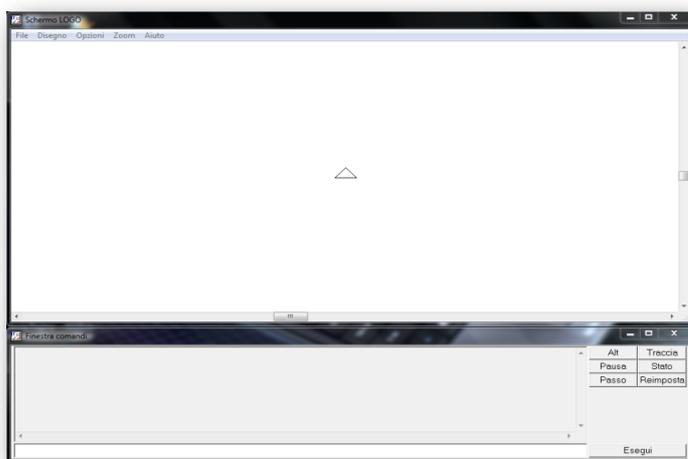


Figura 13
L'ambiente LOGO: da notare la semplicità dell'interfaccia grafica

Nella parte superiore, una finestra bianca sormontata da una barra menu; al centro, un piccolo triangolo: è il nostro robot, con il vertice superiore che indica la direzione in avanti. Sotto a quello che potremmo definire come “Ambiente di lavoro”, una finestra più piccola, utile a listare i comandi impartiti e ad elaborarli (nella piccola finestra inferiore) sotto forma di lettere e numeri. Non mancano alcuni pulsanti funzionali ad eseguire pause, reimpostare il programma, resettare la finestra comandi e così via.

Tutto qui. Ma è proprio in questa essenzialità che consiste la vera efficacia del LOGO. L’operatore, meglio se parliamo di un Allievo in età compresa fra dieci e tredici anni. Il programma riconduce l’azione alla correttezza sintattica delle istruzioni impartite, immediatamente eseguibili oppure inseribili in procedure anche complesse: in una parola, fa riflettere, e lo fa senza compromessi.

18. I campi di intervento

Al di là della sua efficacia intrinseca come ambiente di programmazione, LOGO ha altre caratteristiche, che gli consentono di interferire con alcune fra le discipline argomento consueto di insegnamento, conferendo loro un valore aggiunto in termini di interesse e affettività da parte dell’Allievo. Tipico è il campo della geometria, quella, per intenderci, solitamente insegnata per formule e teoremi, che attraverso LOGO non sono più e non tanto oggetto di semplice memorizzazione quanto di “costruzione” concettuale, una costruzione in linea con la teoria di riferimento (il costruzionismo di Papert, per l’appunto) di impatto educativo molto più efficace.

Un esempio valga per tutti, osservando l’immagine che segue (Figura 14):

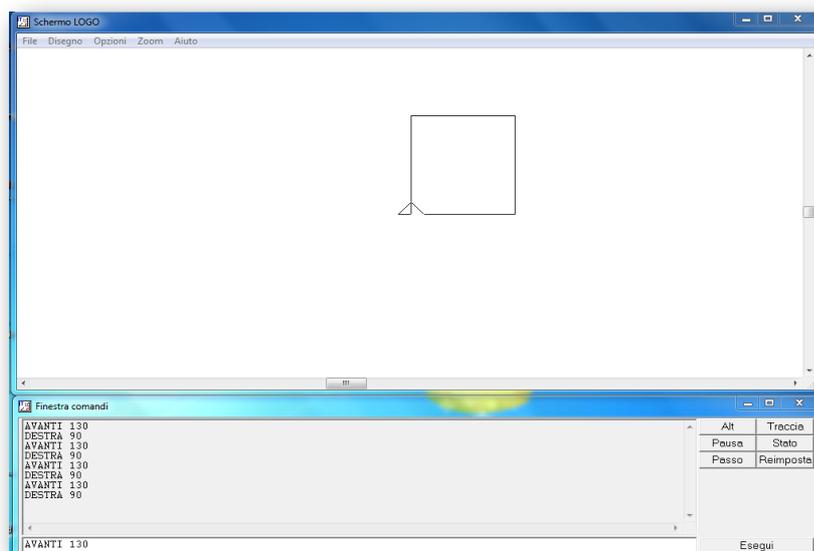


Figura 14
Il LOGO in azione

Nella stessa, è stato realizzato un semplice quadrato con un lato di 130 pixel (si badi bene, non centimetri). La figura è stata tracciata per mezzo di una procedura “passo passo” che ha impartito alla “Tartaruga” due semplici istruzioni, ripetute quattro volte, e cioè:

AVANTI 130
DESTRA 90

Indubbiamente, il fatto che la figura realizzata sia riconducibile al concetto di “Quadrato” non è una conoscenza alla portata della macchina: LOGO si è limitato ad eseguire le istruzioni ricevute, una volta verificata la loro correttezza sintattica.

Supponiamo che neanche l’operatore sappia che la figura disegnata sia un quadrato: per lui è semplicemente una “forma” realizzata con quattro lati uguali (AVANTI 130) e quattro rotazioni uguali (90 gradi cadauna). Per quanto ne sa, potrebbe chiamarsi anche “mattonella”.

Proviamo a costruirne un’altra di minori dimensioni, sfruttando però la stessa regola: quattro linee uguali e quattro rotazioni uguali (Figura 15:

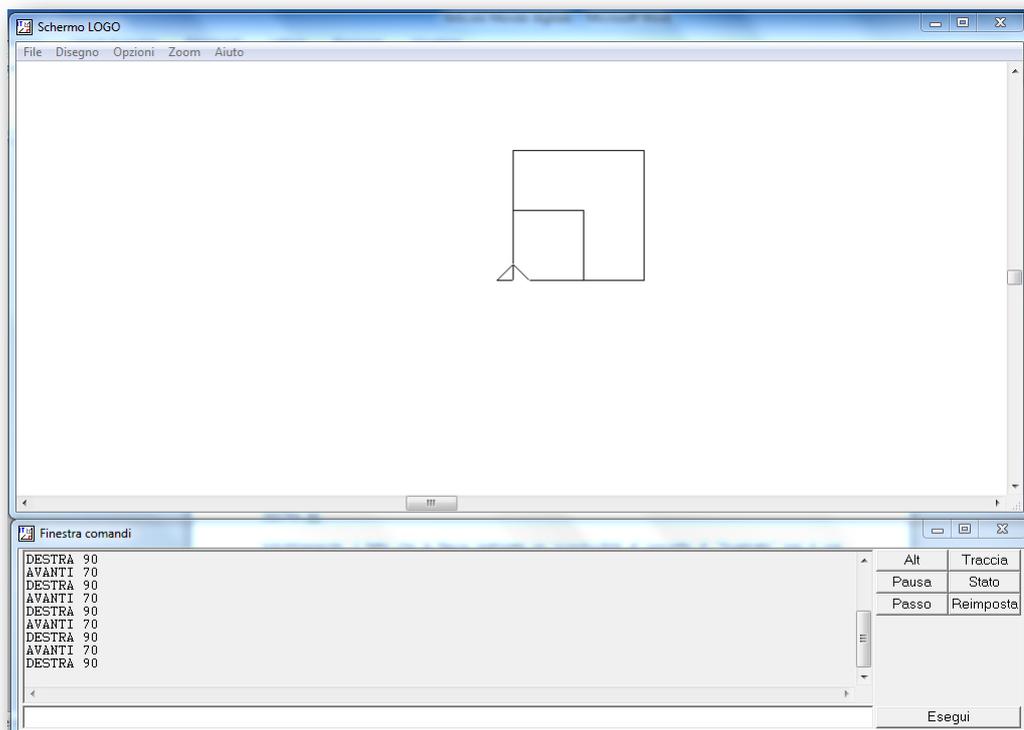


Figura 14
LOGO disegna quadrati concentrici

Volendo, possiamo continuare, senza mai cancellare la figura disegnata in precedenza. Otterremo così una serie di figure della stessa forma, tutte originate dallo stesso punto, e che si differenziano fra loro soltanto per la lunghezza in pixel attribuita con il comando "AVANTI".

Sarà facile, a questo punto, far scaturire la regola dalla stessa esperienza condotta, e chiamare per convenzione la figura realizzata in più dimensioni con il termine "Quadrato". Se provassimo a chiedere a questo punto al nostro Allievo cosa sia un quadrato, egli non potrà fare a meno di rispondere che è una figura "con quattro lati uguali e quattro rotazioni (angoli) uguali".

Lascio al Lettore ogni ulteriore considerazione. Gli esempi riportabili sono infiniti, e patrimonio comune di quanti, sia pure in passato, hanno avviato esperienze didattiche con LOGO. Qui lo spazio è però tiranno, e ci costringe a fermarci.

19. Conclusioni

Lo spazio che ci ospita, così come si presume l'indulgenza di quanti ci hanno seguito sin qui, sono alla fine, e Scuola ci inducono a trarre le nostre conclusioni.

Quella che in molti definiscono con l'espressione "Robotica Educativa (a mio avviso riduttiva rispetto alla vera essenza delle strategie descritte, vicina più ad un'ottica disciplinare che contestuale)" potrebbe rappresentare una nuova frontiera metodologica, nuova quanto radicata nel tempo, in considerazione della sua matrice papertiana (e non solo). Possiamo in realtà affermare che ci troviamo di fronte ad un vero e proprio "ritorno al futuro", in grado di influenzare e comunque interferire a 360° con le varie modalità che caratterizzano, a più livelli, l'intervento educativo.

Da più parti assistiamo al proliferare di esperienze in tal senso, diverse fra loro ma tutte riconducibili ad una matrice comune: la teoria costruzionista di Papert, il pensiero di Feuerstein (per quanto attiene il ruolo del "soggetto mediatore") e, non ultimo, proprio quel linguaggio, il LOGO, dai più dimenticato e relegato al ruolo di "linguaggio fuori moda". Tutte esperienze che meritano, da parte della Scuola e di quanti vi operano, la massima attenzione. Credo però che la stessa attenzione debba essere destinata all'argomento trattato da parte di chi si occupa di Informatica e, soprattutto, di approccio informatico a carattere educativo.

L'informatica non è Office. L'Informatica è una disciplina che, attraverso l'analisi della struttura di un'intelligenza artificiale, ci induce a riflettere sul suo funzionamento, generando abitudini mentali spendibili per "imparare a imparare", obiettivo ultimo della Scuola. Quello che auspichiamo in queste pagine è il confronto fra i vari approcci all'Informatica Metacognitiva e la costante attenzione degli Informatici a questo delicato settore, al fine di avviare efficaci sinergie destinate ad un Futuro che oggi ci guarda con gli occhi sbarrati e ansiosi di apprendere: il Bambino.

Biografia

Giuseppe Albano è Docente di Matematica e Informatica nelle Scuole del Primo Ciclo dell'istruzione dal 1973 al 2009, anno in cui lascia tale ruolo per dedicarsi alla ricerca e alla sperimentazione in campo educativo. Dal 1985, in concomitanza con l'avvio del Piano Nazionale di Formazione e Aggiornamento degli Insegnanti sulla scorta dei Nuovi Programmi per la Scuola Primaria (DP 104 del 12 febbraio 1985), è coinvolto in attività di formazione con l'IRRSAE di Puglia e Sardegna, attività che lo vede impegnato in numerose iniziative di formazione nelle Regioni citate; in tale periodo, perfeziona le proprie competenze nell'ambito dei criteri di programmazione dell'intervento educativo. Sin dal 1978 avvia le sue prime esperienze nell'impiego di Robot e Ambienti di programmazione, definiti poi dallo stesso come "Apparati e Ambienti mediatori" in materia di Informatica Metacognitiva, sviluppando percorsi che si rivelano efficaci per la sua attività di Docente e nell'ambito di interventi dedicati ad Allievi di Scuole primarie in zone ad alto rischio di dispersione scolastica. Nel 1992 fonda, insieme ad un gruppo di Colleghi, l'Associazione "TeleScuola" ed avvia le prime esperienze nazionali per la costituzione di una rete telematica fra le Scuole utilizzando la rete ITAPAC, riscuotendo ampi consensi da parte delle istituzioni coinvolte.

Attualmente, presta la propria consulenza in seno ad AICA per le azioni di certificazione nelle Regioni obiettivo Convergenza (Puglia, Calabria, Campania e Sicilia) ed è promotore di diverse iniziative per lo sviluppo, il consolidamento e la certificazione delle competenze digitali in Docenti e Allievi della Fascia Prescolare, Primaria e secondaria di I Grado; è altresì autore e promotore, in concorso con AICA, di iniziative progettuali e di certificazione destinate alle suddette istituzioni Scolastiche.

email: giuseppealbano@email.it

A telementoring system for supporting laparoscopic surgeries

Paolo Buono, Giuseppe Desolda, Rosa Lanzilotti
Università degli Studi di Bari Aldo Moro
Via Orabona 4, 70125 Bari
{paolo.buono, giuseppe.desolda, rosa.lanzilotti}@uniba.it

Abstract. *In medicine it is very important to perform training in the field without compromising patient's health and possibly saving costs. This is more important for surgery. This poster presents a telementoring system that allows expert surgeons to remotely assist and guide other surgeons during operations.*

Keywords: telementoring, on-line learning, ICT and health.

In un periodo di forti ristrettezze economiche, è doveroso per tutti ricercare soluzioni che garantiscano formazione di alta qualità a costi contenuti. Nel contesto medico è molto importante che la formazione in aula sia seguita da quella sul campo, in particolare per gli interventi chirurgici. È stato dimostrato il valore del telementoring (tutoraggio a distanza con l'uso delle moderne tecnologie) nella formazione di un chirurgo. Infatti, se mediamente un chirurgo deve compiere almeno 30 interventi per potersi considerare autonomo, con il telementoring può essere operativo da subito [1]. Il telementoring consente di ridurre la spesa sanitaria e di migliorare la qualità degli interventi (riduzione di complicanze e ricovero post-operatorio).

Il sistema LARE (Laparoscopia Assistita in REmoto), che stiamo sviluppando, ha l'obiettivo principale di consentire il telementoring durante un intervento chirurgico eseguito in laparoscopia, utilizzando tecniche avanzate di telestration (disegni sul video). Il chirurgo esperto (tutor) assiste e guida un chirurgo in sala operatoria (learner) da un'altra parte del mondo, interagendo via audio, osservando in tempo reale e in alta definizione le immagini che il learner sta vedendo in sala operatoria, ed indicando punti precisi su cui intervenire. Inoltre, LARE consente ad un numero elevato di persone di assistere all'intervento in modalità live, da diversi dispositivi. Chi assiste può partecipare scrivendo in una chat, visibile anche in sala operatoria. I video delle operazioni e dei tutoraggi sono criptati e memorizzati secondo un'architettura cloud ed è possibile rapidamente ritrovare parti di video mediante un software specificatamente progettato.

LARE è anche utile per l'aggiornamento di chirurghi già esperti e per fornire consulenze specialistiche. Il 9 febbraio 2013, LARE ha consentito a circa 300 chirurghi di assistere in live a due interventi in cui i tutor erano distanti più di 800 Km dalla sala operatoria in cui gli interventi si sono tenuti. Al congresso si illustrerà il sistema e si forniranno i dettagli relativi a tale evento.

[1] V. Poulakis, W. Dillenburger, M. Moeckel, R. de Vries, U. Witzsch, J. Zumbé, J. Rassweiler, and E. Becht, "Laparoscopic radical prostatectomy: prospective evaluation of the learning curve", *European urology*, vol. 47, no. 2, p. 167–75, 01-Feb-2005.

Self-BI per una Pubblica Amministrazione

Giovanni Bartolomeo, Anna Fino
Ministero della Giustizia
Largo L. Daga 2, 00164 ROMA
{giovanni.bartolomeo, anna.fino}@giustizia.it

Abstract. *This brief paper outlines how self-Business Intelligence techniques can improve data quality and reduce costs in a real-world public administration.*

Keywords: Business Intelligence, Data Reuse, Public Administration.

Le Pubbliche Amministrazioni (PA) europee sono sedute su una miniera d'oro: il corposo volume di informazioni raccolte¹. Uno dei principali ostacoli rimane tuttavia la difficoltà di ottenere dati completi, esaustivi e tempestivi². Per superare questa difficoltà sono state recentemente individuate due direzioni: (i) maggiore qualità del dato e (ii) canali di pubblicazione dati più diretti³. Un'ulteriore direzione per favorire questo processo virtuoso può consistere nell'elevare l'adozione della Business Intelligence (BI), allargandola anche al personale operativo⁴. Dunque dal rigido concetto di applicazione di BI per il personale direttivo occorre passare a quello di piattaforma flessibile in grado di permettere al personale operativo di creare le proprie applicazioni, infografiche e report tramite "self Business Intelligence" (self-BI).

Al Dipartimento dell'Amministrazione Penitenziaria stiamo riprogettando il sistema di *datawarehousing* in questi termini. Il prototipo implementato consiste in un *framework* di ETL (Pentaho Data Integration) ed una piattaforma di presentazione (Information Builders Webfocus 8.0). Attraverso questa soluzione sperimentale perseguiamo diversi obiettivi: (i) maggiore qualità del dato fin dalla nascita, con correzione di errori ed inconsistenze tramite feedback dallo stesso personale operativo, reso ora maggiormente consapevole degli utilizzi del dato. (ii) Riduzione dei costi a diversi livelli. A livello di sistema: l'implementazione di quelle funzionalità che altrimenti richiederebbero l'intervento dei programmatori è ora possibile tramite self-BI. A livello organizzativo: una maggiore ed oggettiva comprensione dei fenomeni all'interno dell'organizzazione (esempi sovraffollamento, eventi critici, dislocazione dipendenti) consente una gestione maggiormente efficiente. A livello periferico: tramite riutilizzo controllato del patrimonio informativo centrale anche a supporto dei nuovi "circuiti penitenziari regionali". La sperimentazione consente dunque una maggiore coerenza dei dati a costi più contenuti.

¹ Commissione Europea, comunicato stampa IP/11/1524 del 12 dicembre 2011.

² Pennini, A. Aprire i dati del bilancio dello Stato: prospettive e implicazioni per open data e PA.. Convegno sugli open data. Senigallia, 20 novembre 2010.

³ Palmirani, M. e S. Mazzini. Pubblicazione, uso e riuso degli Open Data. Open Data: un cantiere aperto. Forum PA 2013, Roma, 29 maggio 2013.

⁴ Kotorov, R. Talk a IB Summit Italia 2013, Milano, 18 giugno 2013.

Congresso Nazionale AICA 2013

Sperimentazione dei tablet a scuola per la promozione dello sviluppo in aree montane

Sergio Casiraghi, Michele Giugni¹, Graziano Murada²
Socio AICA 14729
Via Venusti 6, 23100 Sondrio
sergio.casiraghi@didasca.org

¹Assessore Cultura di Albosaggia (SO)
Piazza Dante Alighieri 1, 23010 Albosaggia (SO)
michelegiugni76@gmail.com

²Sindaco di Albosaggia (SO)
Piazza Dante Alighieri 1, 23010 Albosaggia (SO)
gmurada@fondazionefojanini.it

Abstract. *In Scuol@3.0, the combined use of tablet and cloud leads to significant changes in didactics and society.*

Keywords: scuol@3.0, school@3.0, education, tablet, cloud.

1. Introduzione dei tablet all'I.C. "Paesi Orobici" *Scuol@3.0*

Un piccolo comune può contribuire allo sviluppo sociale ed economico del territorio investendo in infrastrutture e cultura. Così il Comune di Albosaggia, in provincia di Sondrio, e la **Fondazione Albosaggia** hanno aperto connessioni wireless e dotato di *tablet* tutti gli alunni dalla classe terza di scuola primaria alla prima della secondaria di primo grado dell'I.C. "Paesi Orobici", attivando un'ampia sperimentazione, in corso da ben due anni, che ha già iniziato a dare buoni risultati di ricaduta sia sul piano scolastico che nell'ambito territoriale.

2. Percorsi sviluppati, metodi adottati e problemi affrontati

La sperimentazione ha garantito a medio termine promozione, sostegno ed assistenza all'uso dei *tablet* distribuiti in comodato d'uso ad alunni e famiglie, un fattivo contributo all'alfabetizzazione digitale nelle nostre aree montane in risposta a specifici bisogni emergenti nel rapporto tra istituzioni e territorio.

Lo sviluppo della sperimentazione ha permesso di definire nuovi *setting* operativi nelle classi coinvolte e in aree domestiche interessate dando spazio a forme di *connected learning* innovative attraverso la **Cloud di Google**.

Gli ostacoli tecnici quali *Digital Divide*, integrazione delle LIM o l'utilizzo responsabile degli *account* di rete da parte dei minorenni, sono stati superati.

La responsabilità dell'apprendimento condivisa da insegnanti, studenti e familiari ha capovolto l'insegnamento (*flipped classroom*) assumendo principi e concetti educativi contenuti nel metodo "*papillon*" in un quadro costruzionista.

Tra i tanti percorsi sviluppati, si segnalano quelli ambientali che hanno portato a diffondere l'uso dei **QR Code** in paese. Il processo intrapreso all'Istituto Comprensivo "Paesi Orobici", in associazione con DIDASCA, è stato documentato in rete attraverso il **blog**: <http://sperimentata.blogspot.it>

On-line il laboratorio didattico di economia aziendale per l'abilitazione all'insegnamento.

Virginia Dall'O'
UNIMIB

Dipartimento di Scienze Economico-Aziendali e Diritto per l'Economia
Piazza dell'Ateneo Nuovo 1, 20126 Milano
virginia.dallo@unimib.it

Abstract. *The didactic lab for the qualification of future teachers of business administration proposes contamination between formal and informal contexts in an extended cognitive WEB 2.0 environment. Experiential learning is based on the problem-solving method to allow the community of practice to balance the key competences (KC) and professional (PC) in view of the European Qualifications Framework (EQF), for an effective fallout on "net-generation".*

Keywords: WEB 2.0, experiential learning, business administration.

1. Per un profilo professionale sostenibile

La didattica laboratoriale per l'abilitazione all'insegnamento di economia aziendale (TFA ordinari 2013) è stata improntata su un approccio esperienziale "olistico" in rete [Dall'O' V.,2012], in modo da favorire il cambiamento di ruolo degli insegnanti da dispensatori di informazione a "generatori di conoscenza".

Per uno sviluppo armonico delle competenze chiave (KC) e professionali(PC) e in vista del Quadro Europeo delle qualifiche per l'apprendimento permanente (EQF) gli assets formativi hanno fatto leva simultaneamente sulla dimensione cognitiva, sociologica, tecnologica e metodologica del comportamento, integrate su attività di problem-solving.

Social e cloud hanno caratterizzato il percorso mirato allo sviluppo di una rete fisica e sociale incentrata:

- sulla co-progettazione di interventi didattici rivolti alle classi della secondaria superiore, diversificati in base ai risultati di apprendimento;
- sull'utilizzo delle ICT per la soluzione di casi didattico-disciplinari;
- su un ri-orientamento continuo verso le "buone prassi" quale variabile strategica dell'ambiente formativo.

L'operatività della comunità di pratica sul WEB ha permesso di prendere atto del gioco delle interdipendenze reciproche dei vari assets formativi e ha favorito l'acquisizione di competenze "adattative", oltreché di progettazione didattica, promuovendo un expertise meta-professionale in sintonia con le evoluzioni in corso nella società della conoscenza.

Bibliografia

Dall'O' V., Il laboratorio delle competenze, ARACNE Ed., Roma, 2012

Training professionals to cope with the opportunities of Very High Resolution satellite data for environmental analysis and planning at local scale

Luca Pugliese

*Istituto Internazionale per gli Alti Studi Scientifici - IIASS
Via Giuseppe Pellegrino 19, 84019, Vietri sul Mare, Salerno
l.pugliese@iiasvietri.it*

Abstract. *An overwhelming quantity of Very High Resolution (VHR, in the spatial sense) satellite data is already available, and more will come to support environmental analysis and planning tasks at local scale. This is an opportunity for Professionals to complete their technical reports with information extracted from this rich and economically viable source of territorial data. But, to fully exploit this source of geographical information, some focalized training activities should be taken by the Professionals, due to the specificity of VHR data in terms of algorithms and processing techniques applicable*

Keywords: Education, Remote Sensing, GIS.

During the last decade numerous satellite missions for Earth Observation (both Public initiatives and Commercial led) have been successfully launched. As a consequence an overwhelming quantity of data is already available, and more will come to support Environmental Analysis at different scales. Of particular interest for Professionals are the opportunities of analysis at local scale, offered by the Very High Resolution (in the spatial sense) data that are by now available at a lowering unit cost. To exploit the informative power of these data, both the traditional techniques of image processing should be revised and some "new" procedures should be learned. For example, Object Based Image Analysis has its roots in Pattern Recognition Theory and it is not a novelty in itself; but it was mainly applied to Image Analysis in Medicine or in Industrial Automation. Only the increment in the spatial resolution of the new satellite data, has pushed toward the massive application of this technique also for Earth Observation tasks.

The author is convinced that a course in Remote Sensing, oriented to VHR data analysis could be beneficial for experts and postgraduates willingly to widen their professional opportunities.

Pazienti online: il caso CIDP

M. Marra, R. Longo¹

CIDP Italia ONLUS

via Scipione Dal Ferro, 16, 73039 Tricase (Lecce)

info@cidp.it

Abstract. *Over the last few years, the importance of websites is increased regarding the search of peaces of information and treatments for several pathologies. The web site www.cidp.it has been created in order to become a reference point and to give answers about chronic inflammatory demyelinating disease.*

Keywords: cidp, e-health, rare disease.

Le neuropatie disimmuni sono un gruppo di malattie del sistema nervoso periferico la cui causa si ritiene essere legata ad una aggressione del sistema immunitario contro antigeni del nervo. Tale gruppo di malattie si è molto ampliato negli ultimi anni e la CIDP (polineuropatia cronica infiammatoria demielinizzante) è tra queste. La CIDP esordisce talvolta in modo acuto ma più frequentemente presenta dall'inizio un andamento cronico. In circa la metà dei casi (soprattutto nei giovani) la malattia ha un andamento cronico recidivante, mentre negli altri pazienti ha un andamento cronico progressivo sin dall'inizio. La prognosi a lungo termine è buona in circa la metà dei pazienti.

L'incidenza della patologia è di circa 3,5 casi per 100000 abitanti.

Il sito www.cidp.it nasce dopo l'esperienza familiare degli autori alla fine del 2009 con l'obiettivo di informare su una malattia sconosciuta ai più e ampiamente sotto-diagnosticata.

Attualmente il sito conta circa 600 visite al giorno ed è diventato il punto di riferimento per pazienti, specialisti e case farmaceutiche. Sul sito vi è una periodica pubblicazione di abstract tratti da NCBI spesso accompagnate da interviste ai maggiori specialisti italiani e internazionali.

Oltre all'informazione sulla malattia, il sito è un importante strumento di comunicazione per rivendicare il diritto alla cura di immunoglobuline; cosa invece che è spesso negata per mere questioni di budget sanitario.

Nell'anno solare 2012 il sito ha avuto 51034 visite con 42762 visitatori unici e 136710 pagine viste complessivamente. Le nuove visite sono l'83,08% e la frequenza di rimbalzo del 68,24%. Il 95% delle visite proviene dall'Italia con le città più attive nell'ordine: Roma (8234), Milano (7982), Torino (2642), Napoli (2633) e via via le altre.

Le sorgenti di traffico sono per il 91,2% i motori di ricerca, per il 3,1% i referral, e per il 5,7% traffico diretto. Riguardo ai motori di ricerca i termini più usati sono: "immunoglobuline" (15,08%), "polineuropatia" (6,45%), "cidp" e "prednisone" con il 3% e via via gli altri.

Le pagine più visitate sono quelle relative alle immunoglobuline (10,36%), dei sintomi (8,85%) e la cura (8,23%).

UNA CURVATURA DELL'ARTICOLAZIONE SISTEMI INFORMATIVI AZIENDALI DELL'INDIRIZZO AMMINISTRAZIONE, FINANZA E MARKETING DI UN ISTITUTO TECNICO ECONOMICO PER IL "RAGIONIERE DIGITALE"

Prof.ssa Paola Mastromatteo
Istituto di Istruzione Secondaria Superiore FRANCESCO SAVERIO NITTI
Via J.F. Kennedy, 140/142 – 80125 Napoli
nais022002@istruzione.it

Building knowledge-based economy is only possible if the information is readily available for everyone. The best way to convey information is provided by ICT. Information technology is essential in every field, it is no longer a privilege of a few but has become a necessity and a right for all. Every company, even the smallest, daily use computer tools to maintain high quality and competitiveness. A school curriculum supplemented by informatics training and certification of digital skills, may facilitate learning of informatics competencies, immediately useful in business administration and accounting work.

Keywords: digital skills, accountant, information technology.

1. Un progetto didattico per il "Ragioniere Digitale"

L'articolazione Sistemi Informativi Aziendali (SIA) dell'indirizzo Amministrazione, Finanza e Marketing (AMF) dell'Istituto Tecnico Economico (ITE) si caratterizza sia per le competenze generali nel campo delle attività economico-aziendali, sia per il riferimento all'ambito della gestione del sistema informativo aziendale, alla valutazione, scelta ed adattamento di software applicativi. Tali attività sono tese a migliorare l'efficienza aziendale attraverso la realizzazione di nuove procedure, con particolare riguardo al sistema di archiviazione, all'organizzazione della comunicazione in rete e alla sicurezza informatica. Gli ITE possono utilizzare la quota di autonomia del 20% dei curricoli, nell'ambito degli indirizzi in essere ed in coerenza con il profilo, sia per potenziare gli insegnamenti obbligatori, con particolare riferimento alle attività di laboratorio, sia per attivare ulteriori insegnamenti, finalizzati al raggiungimento degli obiettivi previsti dal piano dell'offerta formativa. L'I.I.S.S. F. S. NITTI intende utilizzare lo spazio di flessibilità, articolando l'opzione SIA per rispondere alle esigenze del territorio e ai fabbisogni formativi espressi dal mondo del lavoro e delle professioni. A tal fine, agli studenti che seguono il corso SIA con la sperimentazione della certificazione EU.C.I.P. (European Certification of Informatics Professionals) Core curricolare, sarà data l'opportunità di integrare e potenziare la propria preparazione seguendo un percorso formativo attivato in sinergia col Consorzio C.L.A.R.A. Il Consorzio C.L.A.R.A. (Computer Learning Applied Research Academy), partner del progetto, è una organizzazione no-profit che opera nel campo della formazione specialistica nel settore ICT finalizzata al conseguimento di certificazioni internazionali. Suo scopo è la promozione dello sviluppo della società dell'informazione attraverso il rafforzamento del potenziale umano, lo sviluppo dell'imprenditorialità ed il miglioramento della competitività delle imprese con particolare attenzione alle regioni del Sud Italia.

La tendenza negli ultimi anni (anno 2012 rispetto alla media 2009-2011, **dati Unioncamere – Ministero del Lavoro**) evidenzia un aumento della disponibilità delle imprese ad assumere **diplomati ad indirizzo informatico** anche senza esperienza lavorativa specifica; in un caso su cinque, le imprese hanno difficoltà a trovare candidati con le caratteristiche desiderate. Questi diplomati trovano opportunità di lavoro trasversalmente in tutti i settori economici, anche se c'è una certa prevalenza del settore dell'ICT, nel quale trovano sbocco professionale oltre 1/3 degli stessi.

I diplomi ad indirizzo amministrativo-commerciale, nelle articolazioni contabile o informatico, sono quelli più richiesti in assoluto dalle imprese italiane. Le opportunità di lavoro riguardano prevalentemente le piccole imprese; il 10% delle assunzioni è rivolta a giovani con meno di 25 anni di età e sale al 23% per i diplomati dell'indirizzo informatico. Una delle caratteristiche di rilievo che le imprese richiedono ai candidati in possesso di diploma - che concorrono a definire il "profilo ideale" del diplomato da assumere soprattutto nel settore amministrativo-commerciale - è il **possesso di conoscenze informatiche**.

L'obiettivo è di attuare una progettazione integrata sul tema: **Didattica delle competenze informatiche per il "Ragioniere Digitale"**, nell'ambito della sperimentazione EU.C.I.P. Core curricolare per il secondo biennio e quinto anno di un corso SIA. Il progetto prevede la realizzazione di tirocini formativi presso i laboratori del Consorzio C.L.A.R.A. per l'efficace trasferimento di **competenze digitali**, utili allo studente nel percorso curricolare finalizzato alla certificazione EU.C.I.P. Core e, come tali, più facilmente spendibili nel mondo del lavoro.

Lim “leggibile” Book

Gloria Cogliati, Giorgia Sanna¹, Daniela Garau²
IC Marco Polo Viani via Pistoia 68 55049 Viareggio Lucca
mitu.kiss@gmail.com

¹Pedagogista Consulente (Centro studi Eriksson.)
Via del Pioppeto 24Fraz. Gardolo - 38121 Trento
giorgia.sanna@eriksson.it

²IC Lerici P.zza Bacigalupi, 6 19032 Lerici (SP)
danielagarau1@gmail.com

Abstract. Learning objects are suitable for inclusion in classes of pupils with different abilities, certificated L.104/92, L.170/11 or, more generally included in the SEN. Anyone, involved in a learning process, has the same right to learn “not like everyone else”, but “*with the same opportunities*” which involves the use of methods, strategies, different times and ways.

Keywords: inclusion, same opportunities, readability, accessibility

Una proposta per lezioni digitali inclusive

Lo scopo del lavoro per la redazione dei contributi è quello di dare forma all’idea iniziale di costruire oggetti di studio adatti all’inclusione nelle classi di alunni con diverse abilità, certificati L.104/92 o L.170/10 o, ancora, inclusi più genericamente nei BES con la direttiva 27/12/12 e la CM n.8 del 6/3/13, perché le nostre scuole, oggi devono affrontare in ciascuna classe almeno una problematica, se non diverse in concomitanza. Tutti gli attori del processo di apprendimento hanno la stessa dignità di apprendere, “non come tutti gli altri”, ma “**con le stesse opportunità**” anche se ciò comporta l’uso di mezzi, strategie, tempi e modalità differenti. Lo studio di quanto già in letteratura ci ha portato a proporre e sperimentare, in diverse azioni di ricerca azione, come potrebbe migliorare il clima di classe, e l’apprendimento dei singoli proponendo una successione di lezioni, in diverse discipline, progettando gli oggetti con l’obiettivo specifico di rendere le lezioni, 1) ad alta leggibilità, 2) facilitate e accessibili 3) multimediali e in progress. Abbiamo fatto delle scelte precise sul colore dello sfondo e dei font, per ottenere il giusto contrasto e non affaticare la vista, infine, abbiamo inserito le registrazioni vocali, attivabili con un click, su tutto il testo scritto. Nella presentazione a schermo per la classe il testo compare grazie al click di attivazione ma nella stampa è sempre in primo piano senza che sovrapposizioni alle immagini diminuiscano la leggibilità. Il materiale prodotto può essere stampato per tutti in formato *.pdf grazie alle possibilità di esportazione offerte dal software della stessa lim, visualizzato pagina per pagina, per gli alunni con difficoltà di lettura o 3 immagini per pagina con uno spazio per le annotazioni laterali (appunti individuali) infine con 6 schermate per pagina. Il tutto consentirà nel passaggio dal cartaceo al digitale un ancoraggio possibile sia per gli alunni che per le famiglie.

Il vantaggio della facilitazione, con la stesura di testi basata su concetti chiave, potrà inoltre essere di aiuto a tutta la classe, e grazie ad un’attività di studio condiviso potrà essere ampliato a partire dagli spunti delineati.

Consensus conference http://www.aiditalia.org/upload/cc_disturbi_apprendimento_sito.pdf

Progetto Classe Turistica

Rita Giancotti, Marco Natali
Istituto Professionale di Stato Maffeo Pantaleoni
Via Brigida Postorino 27 – 00044 Frascati RM
RMRC03000T@istruzione.it

Abstract. *This paper describes the procedural phases referred to a project through which some students of a vocational public institute near Rome had realized a multimedial touristic product. The work, a web site contained in a CD, represents the result of a new approach of teaching thanks to the use of the New Technology.*

Keywords: project, multimedial touristic product, web site.

1. Il Progetto Classe Turistica

All'interno del POF dell'IPS Maffeo Pantaleoni di Frascati relativo all'anno scolastico 2012/2013 è stato inserito, per le classi IV del corso Tecnico dei Servizi Turistici, il Progetto Classe Turistica, allo scopo di realizzare un prodotto multimediale per la partecipazione all'omonimo concorso indetto annualmente dal Touring Club Italiano e rivolto a tutte le classi delle Scuole Secondarie di II grado. La partecipazione al concorso, nella sezione Vieni da noi, prevedeva la realizzazione di un prodotto multimediale di promozione del territorio di residenza degli alunni (nel nostro caso i Castelli Romani) mirato ad una particolare forma di turismo, il turismo scolastico. Per realizzare tale progetto, nella consapevolezza che non si può promuovere un territorio senza averne la conoscenza, l'Istituto si è avvalso della collaborazione di una Associazione che opera sul territorio per uno sviluppo turistico sostenibile, responsabile e sociale (Associazione Manacubba) e che ha curato l'individuazione e la realizzazione di tre visite guidate nel Comune di Monte Porzio Catone. Le classi si sono recate in visita a siti di interesse archeologico, culturale ed enogastronomico. Per la realizzazione del prodotto multimediale, sono stati individuati quattro gruppi di lavoro ad ognuno dei quali sono stati assegnati precisi compiti e funzioni in relazione alle seguenti forme di turismo: culturale, enogastronomico, del divertimento, turismo scolastico. Ogni gruppo di lavoro ha nominato un capogruppo, ha creato una propria cartella di lavoro contenente immagini, documenti, e materiali di lavoro sulla base dell'argomento assegnato, ha studiato sia su materiale fornito dall'Associazione che su materiale scaricato da Internet, ha realizzato filmati allo scopo di incentivare studenti d'Italia, e stranieri, a visitare il proprio territorio in chiave innovativa. Tutto il materiale prodotto dai vari gruppi è stato successivamente organizzato sotto forma di sito web, utilizzando Google Sites, che, trasformato in versione off line è stato registrato su CD assieme a cartelle di lavoro dei singoli gruppi contenenti immagini e documenti prodotti dagli alunni e materiali di studio. Il sito è visionabile all'indirizzo: <https://sites.google.com/site/magicromancastle/>.

Congresso Nazionale AICA 2013

Tecnologia ed Informatica nella Scuola Primaria

Liliana Dell'Isola, Giuseppina Pucciarelli
IV Circolo Didattico "Matteo Mari" Salerno
P.zza A. Trucillo 22 – 84100 Salerno
giuseppina.pucciarelli@istruzione.it

Abstract. *Computer Science assumes a significant role in the IV Didactic Circle of the Salerno's "Matteo Mari" Primary School. The activities of the technological project are structured in recreational-operational situations also through the use of the LIM (Lavagna Interattiva Multimediale – Multimedia Interactive Board). The obtained success has confirmed the role to confer to Computer Science in the Primary School.*

Keywords: Primary School Education, E-Learning, Digital Citizenship.

L'Informatica assume un ruolo di significativo spessore nella Scuola Primaria del IV Circolo Didattico "Matteo Mari" di Salerno.

A partire dall'anno scolastico 2012/2013 vengono organizzati percorsi di particolare valenza formativa soprattutto per gli alunni delle classi quarte e quinte, con sviluppo di attività riguardanti il sistema binario, il codice morse, l'uso di algoritmi, grafici e diagrammi di flusso.

Il progetto tecnologico – informatico, curato dalle docenti Giuseppina Pucciarelli, Anastasia Pastore, Anna Saccone, assume come finalità educativa di base l'acquisizione delle prime competenze del "cittadino digitale", in diretta corrispondenza con il quadro normativo del Programma Europeo E.T. 2020.

Le attività sono strutturate in momenti ludico-operativi, anche mediante l'utilizzo della LIM, creando un'azione educativo-didattica volta a favorire la maturazione di abilità interdisciplinari nel Soggetto Educando.

Al termine del primo anno di "sperimentazione", i risultati raggiunti sono stati pienamente soddisfacenti, facendo registrare significativi consensi da parte dell'utenza. Nel monitoraggio complessivo del Piano dell'Offerta Formativa 2012/2013, il successo ottenuto ha confermato il ruolo da attribuire all'Informatica nella Scuola Primaria, quale opportunità trasversale di crescita graduale degli alunni, attraverso l'utilizzo consapevole degli strumenti tecnologici, sostenendo lo sviluppo dei processi logici del pensiero e l'acquisizione di capacità di problem solving.

Abstraction and Lego: a K-12 experience

M.A. Bochicchio, A. Longo, M. Marra
DIDALab – Università del Salento
via per Monteroni, Complesso Ecotekne, 73100 Lecce (Le)
{mario.bochicchio, antonella.longo, massimo.marra}@unisalento.it

Abstract. *We present an experience with a group of K-12 students, using Lego bricks to learn abstraction and programming principles. The main evidence is that 7-11 years old children are ready to approach abstraction and programming.*

Keywords: Lego, programming principles, K-12.

L'utilizzo dei mattoncini Lego per l'insegnamento dei fondamenti della robotica nelle scuole elementari non è nuovo. In questo caso si utilizzano i kit di Lego MindStorm per insegnare l'astrazione e i fondamenti della programmazione a un gruppo di bambini di scuola elementare (dai 7 agli 11 anni). L'esperienza si è svolta presso il DIDA-Lab dell'Università del Salento nel settembre 2012 e ha visto il coinvolgimento di 16 ragazzi per 4 ore al giorno per 5 giorni consecutivi. Ai partecipanti non erano richiesti prerequisiti. Gli strumenti a disposizione erano 6 kit Lego Mindstorm con 6 portatili e 6 RCX (i controller programmabili della Lego). Il team di ricerca era formato da 4 unità e due studenti di supporto. Il primo giorno è servito per omogeneizzare e socializzare il gruppo di bambini; utilizzando i tradizionali mattoncini Lego i gruppi avendo avevano l'obiettivo di costruire oggetti diversi (automobili, giraffe, rane, etc.). Il secondo giorno sono stati illustrati gli ingranaggi e con i componenti Lego i bambini sono stati condotti verso la costruzione di carrucole e automobiline. A turno, i gruppi sono stati avvicinati all'astrazione e ai concetti base della programmazione (i comandi in sequenze, le condizioni, le opzioni), tramite del gioco della mosca-cieca. Il terzo giorno la classe è stata formata sui sensori, motori, luci e si è illustrato loro come combinarli in modo opportuno per potere fare svolgere, attraverso un telecomando, determinate funzioni. Il quarto giorno la classe è stata formata all'uso dell'applicativo grafico ROBOLAB, che graficamente consente di definire dei programmi che poi saranno eseguiti dall'assemblato Lego. Il quinto giorno la classe ha realizzato degli assemblati e piccoli programmi che sono stati poi illustrati ai genitori alla fine del corso.

L'esperienza è stata sicuramente positiva. La principale evidenza è legata al diverso interesse dei bambini in base all'età: i più piccoli erano interessati alle costruzioni in quanto tali, i più grandi guardavano gli ingranaggi e la programmazione. Gli ultimi due anni di scuola elementare sono apparsi il momento idoneo per avvicinare i ragazzi alle tecnologie e a stimolarli nello sviluppo di idee e a un approccio risolutore ai problemi. Accertato l'estremo interesse dei ragazzi verso questo tipo di attività, il prossimo passo sarà quello di avviare una sinergia con scuole e ludoteche presenti sul territorio per una replica in loco dell'iniziativa, avendo formalizzato il percorso e gli strumenti didattici più efficaci.

Congresso Nazionale AICA 2013

Macchine di Leonardo rianimano la fabbrica rinascimentale da Milano per il mondo

Mauro M. Langfelder
casateWood exhibiting Center
Via Casati 30 - 23880 Casatenovo (Lecco)
maurolangfelder@yahoo.it

1. Presentazione: come copertina di un probabile e-book

Marco d'Oggiono e **Cesare da Sesto** con il Caprotti (o **Salaino**) ed il **Boltraffio** appaiono intorno al loro maestro nel **monumento** che lo ricorda in una piazza al centro della città, che da un semestre sta ospitando **il Mondo di Leonardo**: la **Milano**, insubre e longobarda, protesa al rilancio dell'economia attraverso la cultura digitale.

La **mostra** è allestita nelle Sale del Re, affacciate sulla piazza accanto al monumento, e viene presentata a Fisciano, parallelamente dallo stesso autore, insieme ad un circostanziato contributo a parte, che ne descrive in *full paper* la valenza in piena **sintonia** con il tema del 50° Congresso di AICA.

Il monumento a Leonardo con i quattro discepoli campeggia nel poster fra le due mappe sfumate dell'Insubria (dei tempi dell'**Editto della Tolleranza**) e di una terra longobarda pervasa fino a percorrere tutta la penisola. E ricorda i tanti oltraggi subiti dalla città, che ha sempre saputo ricostruirsi, come nei mitici 40 giorni del **1906** quando, durante l'**Expo** del BiE di allora, furono distrutti da un incendio due padiglioni al Parco del Sempione ...

L'anelito alla ricostruzione sistematica dopo gli oltraggi potrebbe trovare oggi anche la capacità di reagire alle offese degli **writers** che un po' dappertutto ne offendono la tolleranza; mentre la città risale dalle sue brume storiche, e vuole testimoniare al mondo la sua ospitale accoglienza per l'Expo del **2015**.

I corsi di formazione permanente del **Dipartimento ABC** del Politecnico dedicati al **buildingSMART** del **BIM** (e dell'Interoperabilità) ripresenteranno il legno come materiale per le costruzioni prefabbricate, per valorizzare la modellazione eccezionale delle macchine di Leonardo, esposte nella rassegna milanese, miniate appunto in **legno**, a fare un **FabLab** di opere rinascimentali...

A cominciare dal banale ponte auto-costruito, che oggi è ammirato da tutte le generazioni (dai bimbi ai nonni) di visitatori, per arrivare alla complessa macchina del tempo ... quindi per pensare che la fruibilità in digitale dei suoi maggiori documenti, a partire dal suo **Codice Atlantico**, sia la testimonianza eccellente di quanto siano puntuali i temi che AICA propone; e che dalla Milano di Leonardo possono essere diffusi quasi in forma di **enciclopedia** virtuale ...

Un altro ebook, assai più ricco di quello delle Capitali in Italia, che accomuna Napoli e Milano, potrebbe proprio aprirsi, e scorrersi dopo questa **copertina** ... per una metropoli accogliente, nebbiosa, oltraggiata nella sua origine insubre e longobarda.

smartWash: sistema integrato a supporto di una lavanderia industriale

Alessandro Quarto ¹, Angelo Francesco Marangi ²

¹myHermes Srl

Corso Italia 63, 74121 – Taranto – Italia, alessandro.quarto@myhermessrl.com

²Finanza e Controllo – Consulenti Aziendali Associati Srl

Viale Magna Grecia 420/A, 7412 – Taranto – Italia, f.marangi@finanzaecontrollo.net

Abstract. *Aim of this work is to present the smartWash project. It is an integrated system for an industrial laundry. The system provides solutions for remote monitoring of linen, laundry fleets and resource consumption.*

Keywords: remote control, optimization, industrial laundry.

Il progetto smartWash, finanziato dalla Regione Puglia, (“Regolamento Regionale 20/2008, Avviso Piccole Imprese innovative di nuova costituzione”) definisce un ambiente innovativo a supporto delle attività di una lavanderia industriale. Il core del progetto è sviluppato su tre moduli:

- **smartTag:** sistema di ottimizzazione del check-in/out della biancheria. Permette di contrastare il furto e la perdita di capi; di monitorare il numero di lavaggi cui è sottoposto ciascun capo; di ottenere in real-time numero e tipo di capi ritirati. Il server consta di modulo di l’etichettatura con Tag RFID passivi lavabili e stirabili, e di un trasponder dedicato alla comunicazione con i client. I client, disposti nel vano posteriore dei veicoli di raccolta, constano di due blocchi: smartTag reader (tunnel schermato per la lettura di smartTag RFID passivi UHF EPC Gen2 Global a 96 bit) e modulo trasponder GPRS/UMTS.
- **smartSens:** sistema di pesatura remota dei carichi. Consente di eliminare le inevitabili differenze tra i valori dichiarati in consegna e quelli verificati in lavanderia. Conoscere in anticipo il quantitativo di biancheria in arrivo permette di ottimizzare risorse e cicli di lavaggio necessari. I client integrano sensori di forza su una base di pesatura.
- **smartTrack:** sistema di monitoraggio remoto delle rotte percorse dai veicoli di raccolta basato su tecnologia GPS-GPRS. Consente di valutare l’efficienza della logistica della lavanderia e di applicare tecniche di A.I. per pianificarne le attività (es. itinerari delle flotte, divisione ottimizzata dei carichi) [1]. smartTrack permette di ottimizzare i cicli di lavaggio da effettuare, prevedendo cicli di lavaggio non a pieno carico. L’elaborazione delle informazioni permette di determinare un impiego ridotto di materie prime (acqua, detersivi, additivi) ed un minore impatto ambientale ed economico di processo

[1] V. Di Lecce, A. Amato, F. De Luca, A. Giove, M. Calabrese, A Multi Agent System for Hazardous Material Transport Management, IASTED ASM, Corfu, Greece, June 23-25, 2008.

IBSE (Inquiry Based Science Education) in Virtual World

Annalisa Boniello, Marina Gallitelli¹

ISS Pitagora

Via Tiberio 1 Pozzuoli, Napoli

annalisaboniello@libero.it

1ISS Pitagora

Via Tiberio 1 Pozzuoli, Napoli

marina.gallitelli@gmail.com

Abstract. *The Project is based on IBSE in virtual world using simulation and a serious game on water and life. A Virtual Island, called Science Island, is built with a e-learning path where students, as little scientist, solve tests in every step.*

Keywords: IBSE, virtualworld, serious game.

1. IBSE (Inquiry Based Science Education) in Virtual World

Il progetto nasce dalla metodologia dell'Inquiry Based Science Education, applicata in ambienti virtuali 3D. È stata creata un'isola di scienze in 3D, chiamata [Science Island](#), in cui attraverso un gioco di ruolo gli alunni devono superare vari step attraverso 10 attività basate sull'inquiry, impersonando dei piccoli scienziati.

La tematica scelta per l'attività è "L'acqua e la vita", dal punto di vista chimico, fisico e biologico. Il livello di inquiry per il percorso è di tipo strutturato dove le prove e gli esperimenti sono stati strutturati dai docenti. Ogni step è diviso in fasi secondo le 5E, *engage – explore – explain – elaborate – evaluate*.

L'isola virtuale, è una isola costruita attraverso il software Opensim (http://opensimulator.org/wiki/Main_Page) nell'ambito del Progetto sulla didattica nei mondi virtuali di INDIRE chiamato EdMondo (<http://www.secondlearning.it/>). Un ambiente immersivo 3D interamente dedicato all'innovazione della didattica, all'interno di EdMondo esistono altri progetti didattici di sperimentazione in ambito matematico, archeologico e letterario. L'accesso a EdMondo è riservato a docenti di ogni ordine e grado ed ai loro studenti. Nell'isola ogni step parte da una domanda per poi passare ad attività sperimentali con determinati obiettivi di competenza per ritornare poi in virtuale dove, se l'obiettivo di competenza è stato raggiunto sarà possibile passare allo step successivo.

Gli alunni nell'isola seguono le sue domande, i suoi passi e i suoi esperimenti per poi alla fine arrivare dove si nasconde. È possibile visionare la sequenza narrativa iniziale al seguente link : <http://www.youtube.com/watch?v=-cF5une6MJw&feature=youtu.be> . Il target degli alunni è dai 13 ai 15 anni, quindi collocabile dalla terza media alla seconda superiore.



AICA
 Associazione Italiana
 per l'Informatica
 ed il Calcolo Automatico

50° Congresso Nazionale AICA 2013

Frontiere Digitali: dal Digital Divide alla Smart Society



Programma



UNIVERSITÀ DEGLI STUDI
 DI SALERNO



Dipartimento
 Studi e Ricerche Aziendali
 Management e Information Technology

**18-20
 settembre
 2013**

**Università di Salerno
 Campus di Fisciano**

patrocini



*Presidenza
 del Consiglio dei Ministri*



*Ministero
 dello Sviluppo Economico*

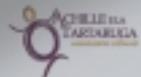


REGIONE CAMPANIA

sponsor



partnership



media partner



Mercoledì 18 settembre

10:30 **Inizio registrazione al Congresso AICA** (Aula Magna)

11:00 Sessioni scientifiche

(S0.1) **Supporto ICT al Patrimonio Culturale** (Aula Giurisprudenza Volterra)

(S0.2) **La Cittadinanza Digitale: Sicurezza, Fiducia e Tutela dei Diritti** (Aula Giurisprudenza De Capraris)

(S0.3) **E- Health: ICT per la Cura e l'Assistenza alla Persona** (Aula Giurisprudenza Castellano)

12:00 **Direttivo AICA** (Aula ex-Senato)

13.00 Pranzo libero

14:30 **Sessione di apertura** (Aula Magna)

Intervengono:

Genny Tortora, Chair Congresso AICA

Bruno Lamborghini, Presidente AICA

Saluti:

Raimondo Pasquino, Rettore dell'Università di Salerno

Aurelio Tommasetti, Rettore Eletto dell'Università di Salerno

Guido Trombetti, Vice Presidente della Regione Campania

15:00 **Sessione Plenaria Frontiere Digitali: dal Digital Divide alla Smart Society** (Aula Magna)

Introduce e coordina **Eduardo Scotti**, Giornalista La Repubblica

Keynote: **Francesco Sacco**, Università Bocconi e Agenda Digitale

Intervengono:

Annamaria Di Ruscio, Net Consulting - **Andrea Granelli**, Kanso -

Claudio Demartini, Politecnico di Torino - **Elio Catania**, Assinform

Incontro con i giovani vincitori di

Premi di Laurea ETIC 2013 AICA - Rotary

Olimpiadi Internazionali di Informatica 2013

17:30 Coffee Break

17:30 **Sessione poster (p.1)** (Aula Magna)

18:00 **Workshop Informatico dietro le quinte, dedicato agli studenti di informatica** (Aula Magna - Sala Stampa)

coordina **Stefano Levaldi**, IEEE Life Fellow, Sapienza, Università di Roma

intervengono:

Giuseppe Visaggio, Università di Bari - **Marcello Pappagallo**, Infobyte - **Roberto Bellini**, AICA

Michele Nappi, Università di Salerno

18:00 Sessioni scientifiche

(S1.1) **ICT e Formazione: Esperienze e Metodologie - I** (Aula Giurisprudenza Volterra)

(S1.2) **Usabilità: Applicazioni Utente-Centriche e Best Practices** (Aula Giurisprudenza De Capraris)

(S1.3) **Network Management e Cloud Computing - I** (Aula Giurisprudenza Castellano)

19:30 Partenza per il Welcome Party - Forte "La Carnale", Salerno

Giovedì 19 settembre

8:30 Sessioni scientifiche

(S2.1) **L'Informazione Geografica per l'Innovazione Sociale** (Aula Giurisprudenza Volterra)

(S2.2) **Biometria e Riconoscimento di Immagini** (Aula Giurisprudenza De Capraris)

(S2.3) **ICT e Formazione: Esperienze e Metodologie - II** (Aula Giurisprudenza Castellano)

8:30 **Workshop Salute: ICT per la cura e l'assistenza della persona** (Aula Consiglio)

coordina **Giuseppe Festa**, Università di Salerno

intervengono:

Alex Giordano, Ninja Marketing, Centro Studi Etnografia Digitale - **Bianca Barone**, Razorfish Healthcare

8:30 **Workshop ICT come elemento strategico di innovazione - esperienze di Spin Off e Start Up** (Aula Nicola Cilento)

coordinano **Filomena Ferrucci**, Università di Salerno e **Agostino Marengo**, Università di Bari

intervengono:

Fabrizio Cobis, MIUR - **Giuseppe Mastronardi**, Politecnico di Bari - **Giovanni De Caro**, Intesa San Paolo

Roberto Magliulo, Presidente Piccole Imprese, Confindustria Salerno - **Alfredo Lambiase**, Università di Salerno

- 10:30 **Coffee Break**
- 10:30 **Sessione poster (p.2)** (Aula Nicola Cilento)
- 11:00 **Sessione Plenaria Fab Lab: la nuova bottega rinascimentale** (Aula Nicola Cilento)
Introduce e coordina **JC De Martin**, Politecnico di Torino
Keynote: **Riccardo Luna**, autore di "Cambiamo tutto! la rivoluzione degli innovatori"
- A seguire panel discussion con:
Umberto Cugini, Politecnico di Milano - **Davide Gomba**, Officine Arduino, socio fondatore Fablab Torino
Antonio Grillo, FabLab Napoli - **Amleto Picerno**, Mediterranean Fablab
Stefano Micelli, Università Ca' Foscari e autore di "Futuro Artigiano" - **Andrea Abate**, Università di Salerno
Edoardo Calia, Istituto Superiore Mario Boella - **Michael Burnett**, Alpine Ventures
- 13:15 **Pranzo libero**
- 14:00 **Sessione Plenaria Sapere e Saper fare: il Sistema Nazionale di Certificazione delle Competenze e gli orientamenti Europei** (Aula Nicola Cilento)
Introduce e coordina **Claudio Demartini**, Politecnico di Torino
 Intervengono:
Fiona Fanning, CEPIS - **Anna Brancaccio**, MIUR - **Claudio Gentili**, Confindustria
Roberto Pettenello, Fondo BA - **Roberto Bellini**, AICA - **Rosamaria Barrese**, Agenzia per l'Italia Digitale
Franco Patini, Confindustria Digitale - **Rodolfo Zich**, Torino Wireless - **Angela Orabona**, USR Campania
- 15:45 **Coffee Break**
- 16:00 **Tavola rotonda Smart Mobility e Social Innovation** (Aula Nicola Cilento)
Introduce e modera **Antonio Pescapé**, DIETI, Università di Napoli Federico II
 Intervengono:
Salvatore Di Dio, TraffICO2 - **Antonio Cosma**, SMOB - **Luigi Mingrone**, Ci.Ro. - **Manuela Tufo**, S2-Move
Anna Donati, Delegata del Sindaco di Napoli per la mobilità sostenibile - **Fabrizio Cobis**, MIUR
- 16:00 **Workshop Sfide didattiche per l'informatica del 2020** (Aula Consiglio)
Introduce e coordina **Alberto Marchetti-Spaccamela**, Sapienza Università di Roma
Keynote: **Carlo Ghezzi**, Politecnico di Milano & presidente di Informatics Europe
 Intervengono:
Camil Demetrescu, Sapienza Università di Roma - **Claudio Demartini**, Politecnico di Torino
Alberto Montresor, Università di Trento - **Enrico Vicario**, Università di Firenze
- 16:00 **Workshop Digital Forensics: il metodo scientifico delle IT nelle indagini investigative** (Aula 8 Facoltà Lettere)
Introduce e coordina **Giuseppe Mastronardi**, Politecnico di Bari - *Respons.* AICA per la Digital Forensics
 Intervengono:
Vincenzo Giliberti, Ordine Ingegneri TA - *Commissione. Naz. Ingegneria Forense* - **Nanni Bassetti**, CAINE
Mario Savastano, CNR Napoli - **Michele Nappi**, Università di Salerno - **Pierpaolo Murrieri**, Selex ES
Roberto Cusani, Università Roma I "La Sapienza" - **Gianfranco De Fulvio**, RACIS Roma (Archivio Nazionale Impronte Digitali) - **Mirko Grimaldi**, Università del Salento - **Andrea Carmimeo**, Polizia Postale - Bari
Vincenzo Mastronardi, Università Roma I "La Sapienza" - **Antonio Piva**, AICA Nord-Est
- 16:00 **Workshop Robotica: il ruolo dell'impresa nella formazione delle competenze** (Aula 9 Facoltà Lettere)
Coordina **Enzo Marvaso**, Coordinatore Rete Robotica a Scuola
 Intervengono:
Stefano Suraniti, USR Piemonte - **Stefano Scarafiotti**, Business Unit Robotics Comau
Claudio Demartini, Politecnico di Torino - **Remo Praticò**, SEECI SUD
Ferdinando Solimeno, Scuole ed Istituzioni Comau
- 16:00 **Sessioni scientifiche**
 (S3.1.1) **E-Government: Strategie ed Esperienze** (Aula Giurisprudenza Volterra)
 (S3.1.2) **Network Management e Cloud Computing - II** (Aula Giurisprudenza Volterra)
 (S3.2.1) **Smart Society: Innovazione** (Aula Giurisprudenza De Capraris)
 (S3.2.2) **Cloud Computing: Services, Infrastructures** (Aula Giurisprudenza De Capraris)
- 19:00 **Partenza per Cena Congressuale**

Venerdì 20 settembre

8:30 **Assemblea GII** (Aula Consiglio)

8:30 Sessioni scientifiche

(S4.1.1) **Smart City e Green Economy** (Aula Giurisprudenza Volterra)

(S4.1.2) **Realtà Virtuale e Realtà Aumentata** (Aula Giurisprudenza Volterra)

(S4.2.1) **Sistemi di Supporto alle Decisioni** (Aula Giurisprudenza De Capraris)

(S4.2.2) **Digital Inclusion** (Aula Giurisprudenza De Capraris)

(S4.2.3) **ICT e Formazione: Esperienze e Metodologie - III** (Aula Giurisprudenza Castellano)

10:30 **Coffee Break**

10:30 **Sessione poster (p.3)** (Aula Nicola Cilento)

11:00 **Sessione Plenaria Big Data, Open Data, Analytics: strumenti e opportunità per la ripresa dello sviluppo** (Aula Nicola Cilento)

Introduce e coordina **Roberto Masiero**, *Innovation Group*

Keynote: **Robert Laurini**, *INSA Politecnico di Lione*

Relatori invitati:

Agostino Ragosa, *Agenda Digitale* - **Umberto Rapetto**, *Telecom Italia* - **Vicente Rojo Caballero**, *Indra, Spagna*

Domenico Donvito, *ISTAT* - **Alessandro Mantelero**, *Politecnico di Torino*

13.00 **Pranzo libero**

14:00 **Workshop Inspire: prepararsi all'atterraggio. Atto II** (Aula Consiglio)

coordinano **Franco Vico**, *AMFM GIS Italia* e **Monica Sebillio**, *Università di Salerno*

Intervengono:

Domenico Longhi, *Regione Abruzzo e CISIS - CPSG* - **Orlando Battipaglia**, **Vincenzo Parità**, *Regione*

Campania - **Domenico Modaffari**, *Regione Calabria* - **Angelo M. Gaccione**, *Provincia Cosenza*

Alessandra Vercillo, **Fabio Vinci**, *Epsilon- Italia* - **Carlo De Michele**, *Ariespace*

14:00 **Workshop L'Accessibilità digitale come misura di civiltà** (Aula Nicola Cilento)

coordinata **Ennio Paiella**, *Fondazione ASPHI onlus*

14:00 Sessioni scientifiche

(S5.1) **Nuove Strategie di E-Business** (Aula Giurisprudenza Volterra)

(S5.2) **Big Data** (Aula Giurisprudenza De Capraris)

(S5.3) **Software Design e Knowledge Management** (Aula Giurisprudenza Castellano)

Mercoledì 18 settembre ore 10.30 - 13.00

S0.1 Supporto ICT al Patrimonio Culturale (Aula Giurisprudenza Volterra)

- Open Data in Cultural Heritage Environment

Vincenza Ferrara¹, Sonia Sapia², Andrea Macchia³

¹Sapienza Università di Roma Digilab - ²Istituto Comprensivo "Val Maggia", Roma

³Italian Association of Conservation Scientists

- Un sistema di Visual Analytics a supporto dell'interpretazione di incisioni rupestri

Vincenzo Deufemia, Luca Paolino, Valentina Indelli Pisano, Genoveffa Tortora

Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

- Il laboratorio leonardesco come bottega rinascimentale (presentazione breve)

Mauro M. Langfelder Implementi BLM

S0.2 La Cittadinanza Digitale: Sicurezza, Fiducia e Tutela dei Diritti (Aula Giurisprudenza De Capraris)

-Promoting trust protecting citizens' rights in Digital citizenship

Angela Vozella - Centro Italiano Ricerche Aerospaziali

-Accesso alla rete: diritto sociale implicito nella costituzione per la fruizione di una cittadinanza attiva

Valentina Amenta - Istituto di Informatica e Telematica, Consiglio Nazionale Ricerche di Pisa

- La fiducia degli utenti nelle transazioni online tra sicurezza reale e percepita

Mario Petrone - Università del Molise

S.0.3 E- Health: ICT per la Cura e l'Assistenza alla Persona (Aula Giurisprudenza Castellano)

- A Tool for multimodal fusion of brain PET and MR

Giovanni Capobianco¹, Madalina Georgeta Ciobanu², Fausto Fasano², Matteo Merola¹

¹ University of Molise - X-Lab, DIBT - ² University of Salerno

- The multidimensional value of transparency in healthcare organizations computerization

Aurelio Tommasetti, Giuseppe Festa

Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

- Uno strumento di modellazione della scrittura a mano libera per l'analisi di patologie degenerative neuromuscolari

D. Impedovo^{1,2}, G. Pirlo^{1,2}, C. O'Reilly³, R. Plamondon³

¹Università di Bari "Aldo Moro", Dip. di Informatica, - ²Rete Puglia Centre, Bari - ³Ecole Polytechnique de Montreal, Canada

- Sistema Multi-Esperto Intelligente per la Diagnosi del Tumore al Seno

Sebastiano Impedovo, Donato Barbuzzi, Francesco Maurizio Mangini

Università di Bari "Aldo Moro", Dipartimento di Informatica

- UBI-CARE: social learning la gestione delle malattie croniche

Flora Berni, Pierpaolo Di Bitonto, Francesco Di Tria, Teresa Roselli, Veronica Rossano

Università di Bari "Aldo Moro", Dipartimento di Informatica

Mercoledì 18 settembre ore 18.00 - 19.00

S1.1 ICT e Formazione: Esperienze e Metodologie I (Aula Giurisprudenza Volterra)

- The PP&S100 Project: Process Control as an Information System Instance

Claudio G. Demartini¹, Marina Marchisio², Marco Mezzalama¹, Claudio Pardini³, Amelio Patrucco⁴

¹ Politecnico di Torino - ² Università di Torino - ³ IIS Carlo Anti, Villafranca di Verona - ⁴ Fondazione Torino Wireless

- From career education to vocational education and training in digital era: exploiting Communities of Practice

Maria De Marsico, Marco Temperini

Sapienza Università di Roma, Dipartimento di Informatica

- **A Social Game-Based approach supporting Environmental Learning**
Stefania Cuccurullo, Rita Francese, Ignazio Passero, Genoveffa Tortora
Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

S1.2 Usabilità: Applicazioni Utente-Centriche e Best Practices (Aula Giurisprudenza De Capraris)

- **Recommending touristic paths using context information**
Flora Amato, Francesco Gargiulo, Antonino Mazzeo, Vincenzo Moscato, Antonio Picariello, Carlo Sansone
Università di Napoli "Federico II", Dip. di Ingegneria Elettrica e delle Tecnologie dell'Informazione
- **Usability in software development company practices**
Carmelo Ardito, Paolo Buono, Maria Francesca Costabile, Rosa Lanzilotti, Antonio Piccinno
Università di Bari "Aldo Moro", IVU Lab, Dipartimento di Informatica
- **Using the USherlock tool to evaluate GUI usability: a comparison with canonical testing methods**
Rosanna Cassino, Maurizio Tucci
Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

S1.3 Network Management e Cloud Computing I (Aula Giurisprudenza Castellano)

- **Sicurezza nelle reti: utilizzo di architetture multi-core per il monitoraggio del traffico IP**
Marco Mezzalama, Gianluca Oglietti, Enrico Venuto
Politecnico di Torino - DAUIN
- **Cloud Forensics Challenges and Readiness**
Lucia De Marco^{1,2}, Filomena Ferrucci¹, M-Tahar Kechadi²
Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
University College Dublin, Ireland
- **CLOUD: Presente e Futuro**
Francesco De Angelis, Alberto Polzonetti, Samuele Sabbatini
Scuola di Scienze e Tecnologie UNICAM

Giovedì 19 settembre ore 8.30 - 10.30

S2.1 L'Informazione Geografica per l'Innovazione Sociale (Aula Giurisprudenza Volterra)

- **Smart Cities, Urban Sensing and Big Data: Mining Geo-location in Social Networks**
Daniele Sacco, Gianmario Motta, Linlin You, Nicola Bertolazzo, Chen Chen
Università di Pavia
- **Le nuove basi geografiche di supporto alla pianificazione ed alla gestione territoriale**
Gabriele Garnerò
Università e Politecnico di Torino DIST - Dip. Interateneo di Scienze, Progetto e Politiche del Territorio
- **Comprehending geographic scenarios through visual summaries**
D. De Chiara¹, V. Del Fatto², M. Sebillò¹, G. Vitiello¹
¹Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
²Libera Università di Bolzano
- **The use of Web service composition to pull down technology barriers**
Pasquale Di Giovanni^{1,2}, Monica Sebillò¹, Giuliana Vitiello¹, Michela Bertolotto²
¹Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
²University College Dublin, Ireland
- **A Mobile Visual Technique to Support Civil Protection in Risk Analysis**
Luca Paolino¹, Monica Sebillò¹, Genoveffa Tortora¹, Giuliana Vitiello¹, Marco Romano²
¹Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
²University III de Madrid
- **I servizi informativi integrati per la gestione del territorio della regione Campania**
Anna Martinoli, Orlando Battipaglia, Vincenzo Parità
Regione Campania, Governo del Territorio

S2.2 Biometria e Riconoscimento di Immagini (Aula Giurisprudenza De Capraris)

- Analisi regionale per il riconoscimento dell'iride

Donato Impedovo, Giuseppe Pirlo, Lorenzo Scianatico

Università di Bari "Aldo Moro", Dipartimento di Informatica

- Un sistema di autenticazione biometrica basato su tag RFID e riconoscimento del viso

Filippo Battaglia¹, Giancarlo Iannizzotto¹, Lucia Lo Bello²

¹Università di Messina, Dipartimento delle Scienze Cognitive, della Formazione e degli Studi Culturali

²Università di Catania, Dipartimento di Ingegneria Elettrica, Elettronica e Informatica

- Local Stability of Static Signatures using Optical Flow Analysis

G. Pirlo, D. Impedovo - Università di Bari "Aldo Moro", Dipartimento di Informatica

- Analysis of Stability and Complexity in Dynamic Signatures

G. Pirlo, D. Impedovo, F. Pomes - Università di Bari "Aldo Moro", Dipartimento di Informatica

- Caratterizzazione di malattie rare mediante analisi morfologica delle alterazioni del volto

Giuseppe Mastronardi, Cinzia Accettura, Domenico Daleno, Roberto F. Depasquale,

Massimiliano Dellisanti Fabiano Vilardi

Politecnico di Bari, DEI - Dipartimento di Ingegneria Elettrica e dell'Informazione

- A facial image based age estimation system: TAGE

E. Vitulano, D. Impedovo, G. Pirlo - Università di Bari "Aldo Moro", Dipartimento di Informatica

S2.3 ICT e Formazione: Esperienze e Metodologie II (Aula Giurisprudenza Castellano)

- The "CSCT Living Lab" for Computer Science and Computational Thinking

Claudio G. Demartini¹, Fabrizio Lamberti¹, Marina Marchisio², Claudio Pardini³, Amelio Patrucco⁴

¹Politecnico di Torino - ²Università di Torino - ³IIS Carlo Anti, Villafranca di Verona - ⁴Fondazione Torino Wireless

- La parola agli studenti: le ICT a scuola

Isabella Donato - Università di Genova, Scuola di Nuove Tecnologie per le Scienze Umane e Sociali

- e-Learning Platforms: A comparative analysis

G. De Serio, D. Impedovo, G. Pirlo, L. Scianatico

Università degli Studi di Bari "Aldo Moro", Dipartimento di Informatica

- Progettare e realizzare l'e-learning nella Pubblica Amministrazione: il caso della Provincia autonoma di Trento

Nicola Villa - tsm-Trentino School of Management

- Per un Museo Nazionale dell'Informatica

Giovanni A. Cignoni¹, Fabio Gadducci¹, Giuseppe Lettieri², Carlo Montangero¹

¹Università di Pisa, Dip. di Informatica - ²Università di Pisa, Dip. di Ingegneria dell'Informazione

- Formazione e Certificazione: possibili risparmi per il Sistema Italia, per le regioni, per i cittadini e la soddisfazione degli utenti

Antonio Piva¹, Michele Baggi² - ¹Università di Udine - AICA nord-est²

Giovedì 19 settembre ore 16.00 - 18.00

S3.1.1 E-Government: Strategie ed Esperienze (Aula Giurisprudenza Volterra)

- Fuzzy clustering in user segmentation

Leo Iaquinata¹, Anna Maria Fanelli², Giovanni Semeraro², M. Alessandra Torsello¹

¹University of Milano-Bicocca - Dep. of Informatics, Systems and Communication

²University of Bari - Dep. of Computer Science

- Envisioning Smart Disclosure in the Public Administration

Giuseppe Ciaccio, Antonio Pastorino and Marina Ribauda - DIBRIS, Università di Genova

- **Organizzazione e Governance dei Processi di IT Outsourcing**

Aurelio Tommasetti¹, Giovanni Vaia²

¹Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

²Università di Venezia Ca' Foscari

S3.1.2 Network Management e Cloud Computing II (Aula Giurisprudenza Volterra)

- **Il Cloud Computing nella strategia europea: l'armonizzazione delle regole tecniche e giuridiche**

Caterina Flick, Vincenzo Ambriola

Università di Pisa, Dipartimento di Informatica

- **Visualizing Internet-Measurements Data for Research Purposes: the NeuViz Data Visualization Tool**

Giuseppe Futia¹, Enrico Zimuel¹, Simone Basso¹, Juan Carlos De Martin¹

¹Politecnico di Torino - DAUIN, Nexa Center for Internet & Society

²Zend Technologies Inc, R&D Department

- **Una Panoramica delle Piattaforme Cloud e Cloud Pattern**

Beniamino Di Martino, Giuseppina Cretella, Antonio Esposito

Seconda Università di Napoli, Dipartimento di Ingegneria Industriale e dell'Informazione

S3.2.1 Smart Society: Innovazione (Aula Giurisprudenza De Capraris)

- **Digital Citizenship and Social Responsibility of Computer Professionals**

Norberto Patrignani¹, Marco De Marco² - ¹Catholic University of Milano - ²Università Guglielmo Marconi

- **L'“arcobaleno dei diritti della cittadinanza digitale” alla prova**

Andrea Trentini, Fiorella De Cindio - Università degli Studi di Milano, Dip. Informatica

- **Integrated Mobility for Individuals in Smarter Cities: a Crowd-sourcing approach**

Gianmario Motta, Daniele Sacco, Linlin You, Ruan Chonghe, Lu Zhang - Università di Pavia

- **Smart city: innovazione e occupabilità**

Dunia Pepe¹, Marta Palombi² - ¹Università Roma Tre, Isfol - ²Palombi Editore, Roma

- **The Role of the Innovation Poles in the Innovation System of the Region of Calabria**

C. Luciano Mallamaci¹, Domenico Saccà² - ¹Centro di Competenza ICT-SUD - ²Università della Calabria

S3.2.2 Cloud Computing: Services, Infrastructures (Aula Giurisprudenza De Capraris)

- **Cloud Agency: una Guida attraverso il Cloud**

Rocco Aversa, Luca Tasquier, Salvatore Venticinque

Seconda Università di Napoli, Dipartimento di Ingegneria Industriale e dell'Informazione

- **Building a Cloud Enabled Data Center**

Raffaele Stifani - IBM Italia

Venerdì 20 settembre ore 8.30 - 10.30

S4.1.1 Smart City e Green Economy (Aula Giurisprudenza Volterra)

- **Applicazioni Smart per Città Smart: un Contributo all'Innovazione**

Domenico Saccà¹, Simona Citrigno², Sabrina Graziano³

¹Università della Calabria - ²Centro di Competenza ICT-SUD - ³OKT Srl

- **Un metodo di valutazione delle emissioni di CO2 di un sito eCommerce B2C**

Valerio Morfino¹, Salvatore Rampone^{1,2}

¹Futuridea Innovazione Utile e Sostenibile, Benevento - ²Università del Sannio Dip. di Scienze e Tecnologie

- **Un sistema di supporto al posizionamento e al dimensionamento dei punti di accumulo dei rifiuti urbani domestici**

Paolino Di Felice, Leonardo D'Errico, Fabio Franchi

Università di L'Aquila, Dipartimento di Ingegneria Industriale e dell'Informazione e Economia

- Il contributo del settore fotovoltaico alla realizzazione di modelli energetici smart *(presentazione breve)*
Galdi Vincenzo, Testa Mario
Università degli Studi di Salerno, DIIN, Facoltà di Ingegneria

S4.1.2 Realtà Virtuale e Realtà Aumentata *(Aula Giurisprudenza Volterra)*

- Haptic Rendering of Deformable Surfaces in Medical Training
A. F. Abate¹, A. Casanova², M. Nappi¹, S. Ricciardi¹
¹Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
²Università degli Studi di Cagliari - Dipartimento di Scienze Mediche Internistiche "M. Aresu"

- In store augmented reality: retailing strategies for smart communities
Maria Teresa Cuomo, Debora Tortora, Gerardino Metallo
Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

S4.2.1 Sistemi di Supporto alle Decisioni *(Aula Giurisprudenza De Capraris)*

- A Multi-Criteria Decision Aid for Agent-Based Decision Makers
Alba Amato, Beniamino Di Martino, Salvatore Venticinque
Seconda Università di Napoli, Dipartimento di Ingegneria Industriale e dell'Informazione
- Optimizing a static greedy algorithm for influence maximization
Leonardo Capone, Nicola Di Mauro, Floriana Esposito - *Università di Bari "Aldo Moro"*
- Social and Smart Mobility for Future Cities: the S²-Move project *(presentazione breve)*
Davide De Pasquale, Pietro Marchetta, Eduard Natale, Alessandro Salvi, Antonio Tirri, Manuela Tufo
Intelligentia S.r.L. - University of Napoli Federico II . University of Sannio

S4.2.2 Digital Inclusion *(Aula Giurisprudenza De Capraris)*

- "Dis(ease)Ability Game" Quando applicare una teoria pedagogica innovativa è un gioco da ragazzi, anzi...una App.
Raffaella Conversano - *Istituto Comprensivo "A. Moro", Carosino (TA)*
- Il progetto "Sempre Giovani 2.0"
Domenico Consoli - *Istituto Tecnico Commerciale "C. Battisti, Fano"*
- Solidale - Web per ogni età *(presentazione breve)*
Claudio Tancini, Antonio Savarese - *Informatica Solidale*

S4.3 ICT e Formazione: Esperienze e Metodologie III *(Aula Giurisprudenza Castellano)*

- Confronto tra Metodologie Didattiche per l'Apprendimento delle Discipline Giuridiche
Teresa Consiglio, Elvira Vitiello - *Università Telematica "Leonardo da Vinci", Torrevicchia Teatina (CH)*
- A survey on the math test at the A-level scientific high school-leaving exam in Italy
Marco Scialdone, Salvatore Venticinque
Seconda Università degli Studi di Napoli, Dipartimento di Ingegneria Industriale e dell'Informazione
- Can Mathematics Education Take Advantage of Massively Multiplayers Online Competitions?
Mario A. Bochicchio, Antonella Longo, Antonio Bernardo - *Università del Salento, DIDALab*
- Competenza di lettura e didattica on line: ,tutor digitale
Piera Schiavone . I.I.S.S. "R.Canudo" (liceo scientifico), Gioia del Colle (Ba)
- Tecnologia ed Informatica nella Scuola Primaria *(presentazione breve)*
Liliana Dell'Isola, Giuseppina Pucciarelli - *Scuola Primaria "Matteo Mari", Salerno*
- Internet Social Network e DSA *(presentazione breve)*
N. Pastena, L. Nazzaro - *University of Salerno, Dep. of Human, Philosophical and Educational Sciences*
- Digital literacy nella società della conoscenza: alcune riflessioni educativo-didattiche per superare il digital divide *(presentazione breve)*
Alessandro Barca, Mariella Tripaldi - *XVI C.D. "Europa" di Taranto*

Venerdì 20 settembre ore 14.00 - 16.00

SS.1 Nuove Strategie di E-Business (Aula Giurisprudenza Volterra)

- Sampling issues and management solutions in internet-based market researches

Maria V. Ciasullo, Giuseppe Festa

Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

- Document Image Retrieval by LayoutAnalysis

G. Pirlo¹, M. Chimienti², M. Dassisti³, D. Impedovo⁴, A. Galiano⁴

¹ Università di Bari "A. Moro", Dipartimento di Informatica - ² Laboratorio Kad3, Monopoli (BA)

³ Politecnico di Bari, Dip. Meccanica, Management e Matematica - ⁴ Dyrecta Lab, Conversano (BA)

- Acquisitions, what will happen to my business? the acquired firm status

Bice Della Piana, Filomena Ferrucci, Gerardina Flammia, Carmine Gravino,

Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

- Il ruolo delle start-up ICT-based nel cambiamento del paradigma di business: il caso delle smart grids

Roberto Parente¹, Rosangela Feola¹, Walter Rassega²

¹ Università di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)

² Regione Basilicata, Ufficio Società dell'Informazione

- Investigating the Effect of Social Media on Trust Building in Customer-Supplier Relationships

Fabio Calefato, Filippo Lanubile, Nicole Novielli

Università degli Studi di Bari Aldo Moro, Dipartimento di Informatica

SS.2 Big Data (Aula Giurisprudenza De Capraris)

- Smart Cities, Urban Sensing and Big Data: Mining Geo-location in Social Networks

Daniele Sacco, Gianmario Motta, Linlin You, Nicola Bertolazzo, Chen Chen - Università di Pavia

- Thinking BigData: Motivation, Results and a Few Recipes for a Balanced Growth of High Performance Computing in Academia

P. Margara, N. Nepote, E. Piccolo, C. G. Demartini and P. Montuschi

Politecnico di Torino, Department of Control and Computer Engineering

- Integrating Statistical Data with the Semantic Web: The Istat Experience

Raffaella Aracri, Stefano De Francisci, Andrea Pagano, Monica Scannapieco, Laura Tosco, Luca Valentino

Istituto Nazionale di Statistica - Istat

- Big Data: limitazioni e opportunità geopolitiche e geoeconomiche

Carlo Muzzi - AICA

- Business Transformation with Big Data (presentazione breve)

Raffaele Stifani - IBM Italia

S.5.3 Software Design e Knowledge Management (Aula Giurisprudenza Castellano)

- Riconoscimento Automatico di Design Pattern da UML di Software Artifact

Beniamino Di Martino, Antonio Esposito

Seconda Università di Napoli, Dipartimento di Ingegneria Industriale e dell'Informazione

- The Knowledge Experience Base as a support for the global development of software

Pasquale Ardimento, Danilo Caivano, Nicola Convertini, Giuseppe Visaggio

¹ Università di Bari "Aldo Moro" Department of Informatics - ² Ser&Practices s.r.l. Department of Informatics

- Textual Analysis and Software Quality: Challenges and Opportunities

Gabriele Bavota¹, Andrea De Lucia², Rocco Oliveto³, Fabio Palomba², Annibale Panichella²

¹ University of Sannio, Department of Engineering

² University of Salerno, Department of Management and Information Technology

³ University of Molise, Department of Bioscience and Territory

Mercoledì 18 settembre ore 17.30 - 18.00

Sessione poster (*Aula Magna*)

(P.1) **A telementoring system for supporting laparoscopic surgeries**
 Paolo Buono, Giuseppe Desolda, Rosa Lanzilotti - *Università degli Studi di Bari Aldo Moro*

Self-BI per una Pubblica Amministrazione
 Giovanni Bartolomeo, Anna Fino - *Ministero della Giustizia - Dip. Amministrazione Penitenziaria*

Sperimentazione dei Tablet a scuola per la promozione dello sviluppo in aree montane
 Sergio Casiraghi¹, Michele Giugni², Graziano Murada³ - ¹AICA - ²Comune di Albosaggia (SO)

On-line il laboratorio didattico di economia aziendale per l'abilitazione all'insegnamento
 Virginia Dall'O¹ - *UNIMIB - Dipartimento di Scienze Economico-Aziendali e Diritto per l'Economia*

Giovedì 19 settembre ore 11.30 - 11.00

Sessione poster (*Aula Nicola Cilento*)

(P.2) **Training professionals to cope with the opportunities of Very High Resolution satellite data for environmental analysis and planning at local scale**
 Luca Pugliese - *Istituto Internazionale per gli Alti Studi Scientifici - IIASS*

Pazienti online: il caso CIDP
 M. Marra, R. Longo - *CIDP Italia ONLUS*

Una curvatura dell'articolazione sistemi informativi aziendali dell'indirizzo amministrazione, finanza e marketing di un istituto tecnico economico per il "ragioniere digitale"
 Paola Mastromatteo - *Istituto di Istruzione Secondaria Superiore FRANCESCO SAVERIO NITTI, Napoli*

Lim "leggibile" Book
 Gloria Cogliati¹, Giorgia Sanna², Daniela Garau³
¹IC Marco Polo Viani, Lucca - ²Centro studi Erikson - ³IC Lerici, Lerici (SP)

Progetto Classe Turistica
 Rita Giancotti, Marco Natali - *Istituto Professionale di Stato Maffeo Pantaleoni, Frascati (RM)*

Venerdì 20 settembre ore 10.30

Sessione poster

(P.3) **Esperienza di Didattica dell'Informatica nella Scuola Primaria** (*Aula Nicola Cilento*)
 Liliana Dell'Isola, Giuseppina Pucciarelli - *Scuola Primaria "Matteo Mari", Salerno*

Abstraction and Lego: a K-12 experience
 M.A. Bochicchio, A. Longo, M. Marra
Università del Salento - DIDALab

Macchine di Leonardo rianimano la fabbrica rinascimentale da Milano per il mondo
 Mauro M. Langfelder - *casateWood exhibiting Center*

smartWash: sistema integrato a supporto di una lavanderia industriale
 Alessandro Quarto¹, Angelo Francesco Marangi²
¹myHermes Srl - ²Finanza e Controllo - *Consulenti Aziendali Associati Srl*

IBSE (Inquiry Based Science Education) in Virtual World
 Annalisa Boniello, Marina Gallitelli - *ISS Pitagora, Pozzuoli (NA)*

Quadro riassuntivo

18 settembre						
10.30	Inizio Registrazioni					
11.00-11.00	S0.1	S0.2	S0.3	Riunione del Direttivo AICA (ore 12.00)		
13.00-14.30	Pranzo (libero)					
14.30-15.00	Sessione di apertura					
15.00-17.30	Plenaria					
17.30-18.00	Coffee break - Sessione poster					
18.00-19.00	S1.1	S1.2	S1.3	W1		
19 settembre						
8.30-10.30	S2.1	S2.2	S2.3	W2.1	W2.2	
10.30-11.00	Coffee break - Sessione poster					
11.00-11.00	Plenaria					
13.15-14.00	Pranzo (libero)					
14.00-15.45	Plenaria					
15.45-16.00	Coffee break					
16.00-18.00	S3.1.1	S3.2.1	Tavola Rotonda	W3.1	W3.2	W3.3
	S3.1.2	S3.2.2				
20 settembre						
8.30-10.30	S4.1.1	S4.2.1	S4.3	Assemblea GII		
	S4.1.2	S4.2.2				
10.30-11.00	Coffee break - Sessione poster					
11.00-11.00	Plenaria					
13.00-14.00	Pranzo (libero)					
14.00-16.00	S5.1	S5.2	S5.3	W4.1	W4.2	

design: davridescannapieco.it

Comitato Scientifico

Chair:

Genny Tortora, *Università di Salerno*

Co-Chairs:

Stefano Levialdi, *Sapienza Università di Roma*

Bruno Lamborghini, *AICA*

Comitato Organizzatore

Chair:

Genny Tortora, *Università di Salerno*

Co-Chairs:

Giuliana Vitiello, *Università di Salerno*

Monica Sebillo, *Università di Salerno*

Daniela Rovina, *AICA*

Per maggiori informazioni
aggiornamenti al programma:

aica2013@unisa.it - www.aica2013.it

Open Data in Cultural Heritage Environment

Vincenza Ferrara, Sonia Sapia¹, Andrea Macchia²

Digilab – Sapienza Università di Roma

Via dei Volsci, 122 – 00185 Rome

Vincenza.ferrara@uniroma1.it

¹*stituto Comprensivo “Val Maggia”*

Via Val Maggia, 21 00141 Rome

s.sapia@tiscali.it

²*Italian Association of Conservation Scientists*

Via Torquato Tasso, 108 00185 Rome

andrea-macchia@tiscali.it

Abstract. *Linked Open Data technology increases Cultural Heritage Repository. Many activities and studies are ongoing and provide new forms of access to culture, inspiring creativity and stimulating social and economic growth. Furthermore, current studies engage users in a new way of taking part in cultural heritage. In this context various tools have been developed where the museum objects becomes a new teaching aids. The tools developed share museum information with teachers who can re-use and re-contextualize digital museum resources in their multimedia lessons and create educational environments to improve engagement and student learning. This paper present the results of project to develop learning Repository and new goal to create annotation tool to increase the content of cultural heritage by teachers. This approach may be very interesting to look for a new way to promote cultural heritage in other context as lifelong learning and tourism.*

Keywords: Linked Open data, Cultural Heritage, Learning repository.

1. Introduzione

Negli ultimi anni le istituzioni culturali hanno sviluppato cataloghi online e Repository per gestire e promuovere il loro patrimonio. Dopo una fase massiva di digitalizzazione di oggetti culturali è diventato importante capire come meglio utilizzare le molteplici informazioni presenti sulla rete. A questo scopo l'applicazione di nuove tecnologie ha stimolato la realizzazione di un differente accesso ai contenuti e possibili usi e riusi delle informazioni da parte del pubblico come insegnanti, studenti, turisti o più in generale il cittadino per soddisfare diversi livelli di curiosità e di interesse.

La tecnologia Linked Open Data(LOD) [Berners-Lee, 2011] ha suggerito l'opportunità di organizzare e promuovere gli eterogenei contenuti di differenti cataloghi relativi al patrimonio culturale. Guardando a questa possibilità nel 2010, il Polo Museale Sapienza (PMS) ha lanciato un progetto per adottare la tecnologia LOD per promuovere e facilitare l'accesso al proprio patrimonio culturale e scientifico. Negli ultimi anni, infatti, i 20 musei della Sapienza Università di Roma hanno digitalizzato il loro patrimonio e adottato una differente tipologia di organizzazione dei dati in relazione alla particolare disciplina a cui sono collegati, implementando il contenuto online degli oggetti museali con immagini e contenuto informativo come descrizione, periodo storico di riferimento, tipologia del materiale e così via. Mentre i sistemi informativi sono stati prodotti per gestire le collezioni digitali dal punto di vista amministrativo o per l'accessibilità da parte di specialisti del settore e ricercatori, la stessa modalità non è risultata molto utile per insegnanti, studenti o più in generale per il pubblico. Inoltre non tutte le informazioni relative agli oggetti museali sono state messe a disposizione.

Per queste ragioni è nata la necessità di adottare una tecnologia che rendesse interoperabili i diversi sistemi e creasse un unico punto di accesso all'intero patrimonio del Polo Museale.

Una particolare attenzione è stata data alla funzione didattica dei musei e quindi allo studio per lo sviluppo di un Linked Data Repository per stimolare le attività degli insegnanti e degli studenti promuovendo la possibilità di integrare le risorse digitali del patrimonio culturale con i materiali educativi. Il PMS ha sperimentato l'applicazione di tale tecnologia nella produzione di un Repository di metadati nell'ambito di un progetto per l'inserimento del proprio patrimonio nella Europea Digital Library [Halshofer e Isaac, 2011]. Inoltre in diversi progetti il PMS sta sperimentando la condivisione dei propri contenuti rendendoli open per essere ri-utilizzati e ri-contestualizzati dagli insegnanti nell'ambito delle loro lezioni.

2. Riutilizzo del Patrimonio Culturale

L'Agenda Digitale Europea ritiene molto importante il riuso dei contenuti digitali. In questo contesto molte sono le indicazioni che intendono stimolare il riuso del patrimonio

Infatti la Raccomandazione della Commissione Europea dà le seguenti indicazioni: *“The online accessibility of cultural material will make it possible for citizens throughout Europe to access and use it for leisure, studies or work. Moreover, the digitised material can be reused – for both commercial and non-commercial purposes – for uses such as developing learning and educational content, documentaries, tourism applications, games, animations and design tools, provided that this is done with full respect for copyright and related rights”*[EU Commission Recommendation, 2011].

Il patrimonio culturale è stato, inoltre, un dominio privilegiato per applicazioni di accesso personalizzato ai contenuti online e recentemente la ricerca museale ritiene in questo modo di rispondere alle esigenze dell'utente costruendo un

supporto individuale [Ardissono et al, 2012]. In questo contesto i musei hanno sviluppato un modello di presentazione dei contenuti sia all'interno degli spazi espositivi che online per permettere di:

- Comunicare in modo efficiente con il loro pubblico.
- Raccogliere informazioni e feedbacks con il pubblico di riferimento
- Fornire modelli per gestire e risolvere problemi.

Inoltre queste istituzioni culturali hanno articolato le loro funzioni in relazione al territorio e ai visitatori [Falk, 2009]. L'obiettivo di queste istituzioni è lo sviluppo di nuove modalità per promuovere la cultura scientifica secondo la necessità di incrementare la richiesta di interdisciplinarietà nell'insegnamento basata su nuove forme di apprendimento e comunicazione con il supporto della tecnologia.

I musei come le biblioteche e gli archivi stanno esplorando, quindi, il valore aggiunto della condivisione delle loro risorse digitali secondo i principi dei Linked Open Data (LOD). Le risorse culturali possono essere condivise secondo due categorie del concetto "open": Open Data e Open Content. Open data si riferisce alle informazioni quali thesauri e metadata descrittivi mentre gli Open Content possono essere individuati, ad esempio, da video e foto. La combinazione di Open Data e Open Content è una delle tecnologie utili per la divulgazione del patrimonio culturale e può contribuire alla realizzazione della società della conoscenza [Oomen et al, 2012].

Diverse e molteplici sono le motivazioni per adottare una modalità open per le risorse culturali che renda disponibili ad esempio metadata, oggetti e collegamenti alle risorse digitali per incrementare un libero accesso alle collezioni e sviluppare, attraverso i contenuti online, nuove abilità e modelli di apprendimento. Infatti tale metodologia può rendere le collezioni più significative e rilevanti per gli utenti permettendo, inoltre, la produzione di nuovi servizi e supporti per un ri-uso creativo dei materiali con lo sviluppo di nuovi prodotti innovativi anche in termini di attività di tipo collaborativo [Lakhani e Panetta, 2007].

3.Ri-uso dei contenuti digitali nell'insegnamento

In classe, per lo studio della cultura e della storia così come per le scienze e la lingua, sono utilizzati testi e materiale didattico supplementare. Anche se tali materiali sono ben organizzati in base ai loro corrispondenti curricula, spesso, mancano informazioni sufficienti per gli alunni per una maggiore conoscenza dei beni culturali [Aihara et al, 2008].

Le risorse digitali possono fornire agli educatori strumenti utili per stimolare l'ambiente di apprendimento. [Burden e Atkinson, 2009] e diverse sono le teorie di apprendimento che sono state applicate per la costruzione di un sistema idoneo allo sviluppo di attività per didattica scolastica o più in generale per la formazione permanente come ad esempio la teoria di Kolb [Kolb, 1984], utilizzata in ambito museale [Marie, 2010].

Nel contesto museale i curatori hanno disegnato nuovi percorsi e creato laboratori all'interno dello spazio museale e hanno disegnato aree dedicate alla

didattica accessibili dai loro siti WEB. Queste aree virtuali forniscono giochi, itinerari personalizzati, oggetti per l'apprendimento (LO) e servizi educativi. Inoltre, recentemente, molte istituzioni culturali rendono accessibili e fruibili i loro materiali in modalità Open Data.

Nel panorama nazionale e internazionale diverse sono le modalità comunicative e le esperienze:

- Il Canada Museum Center fornisce l'accesso a un'area riservata per agli insegnanti per costruire lezioni online integrate da contenuti museali [Museum Center, 2013]
- Il Museo di Antropologia della Wake Forest University fornisce un programma didattico per la scuola primaria mettendo a disposizione online oggetti, fotografie, attività e materiale storico [Antropology Museum, 2013]
- Il Museo Galileo di Firenze ha inserito online giochi interattivi collegati alla discipline scientifiche [Museo Galileo, 2013].
- Il Rijksmuseum di Amsterdam mette a disposizione un Dimostratore che permette il download di interessanti immagini delle opera d'arte del museo [Aroyo et al, 2007] e recentemente ha reso disponibile una innovative applicazione che ha reso disponibili le immagini delle proprie opere ad alta risoluzione in forma gratuita per essere riutilizzate da qualunque utente in qualsiasi modalità.[Georgels, 2013].
- Il progetto The Natural Europe a partire da un network di risorse digitali ha sviluppato un Repository per l'educazione con lo scopo di condividere un ricco materiale scientifico con gli insegnanti come supporto alle attività didattiche [Schmidt et al, 2011].
- Per quanto riguarda gli Open Data un esempio è il Portale di Cultura Italia che mediante un OAI Provider rende disponibili metadati in formato XML e RDF strutturati secondo vari schemi [Portale Culturalitalia, 2013]

3. Il Progetto

Nel contesto citato il Polo Museale Sapienza con la scuola primaria "Val Maggia" di Roma ha lanciato un progetto per sperimentare l'utilizzo delle risorse digitali del proprio patrimonio culturale da parte degli insegnanti per essere integrate nelle lezioni multimediali. Il Polo Museale Sapienza comprende 20 musei e ha tra gli obiettivi quello di promuovere attività collegate alla didattica. [Polo Museale Sapienza, 2013]

Negli ultimi anni i Musei della Sapienza hanno prodotto cataloghi online per rendere accessibili il proprio patrimonio [Ferrara e Campanella, 2012] e sviluppato un proprio sistema informativo per la gestione dei dati e prodotto il proprio sito web [Ferrara, 2009].

Il PMS ha recentemente studiato la possibilità di creare un unico punto di accesso all'intero patrimonio culturale e scientifico e sviluppato un catalogo utilizzando i criteri indicati dall'ICCD e la tecnologia per la condivisione dei propri dati con altri Repository come ad esempio Europeana Digital Library partecipando al progetto europeo Linked Heritage [LinKed Heritage, 2013].

Per queste ragioni ha adottato la tecnologia LOD per rendere interoperabili i diversi database e costruire un unico punto di accesso alle risorse digitali. Il Repository è stato sviluppato strutturando i dati secondo lo schema LIDO per l'organizzazione e la condivisione dei metadati per essere utilizzati da diversi servizi online e costruire così un portale aggregatore delle risorse facilmente accessibili e condivisibili su WEB. [Tsalapati et al, 2012].

Il gruppo di ricerca, staff dei musei, insegnanti di scuola primaria e esperti di tecnologie hanno analizzato la possibilità di supportare le attività d'insegnamento con la tecnologia multimediale e le risorse digitali accessibili via WEB mediante percorsi personalizzati. Il gruppo di ricerca ha sviluppato un Repository per gestire i differenti ed eterogenei contenuti dei musei utilizzando la tecnologia LOD per promuovere il patrimonio, stimolare la visita ai musei e condividere i contenuti. Per la realizzazione del mapping dei dati è stato utilizzato un idoneo modello RDF per costruire un catalogo di metadati.

Dopo una iniziale ricognizione degli strumenti disponibili nella scuola come le LIM, computer e connessione ad internet è stata disegnata una applicazione per permettere agli insegnanti la costruzione di percorsi personalizzati accedendo via web ai cataloghi museali e la successiva attività di download di queste risorse sul proprio computer adottando un formato XML.

Queste informazioni, scelte e disponibili sul proprio computer, possono quindi essere utilizzate per la produzione di lezioni multimediali secondo una strategia di comunicazione basata su un continuo processo di collaborazione tra musei e insegnanti [Bowen e Filippini-Fantoni, 2004], [Yang et al, 2007].

Il file XML prodotto per la condivisione dei dati è costituito da metadati che permettono l'accesso ai contenuti relativi agli oggetti museali, al sito web del museo e alle immagini utili per la lezione:

```

<path>
  <name>erbario</name>
  <object>
    <id>RMSMUS12</id>
    <museum>Erbario - Museo di Botanica</museum>
    <linkmuseum>http://www.musei.uniroma1.it/erbario/index.html
    </linkmuseum>
    <inventory>1799</inventory>
    <linkform>
      <!
[CDATA[http://www.musei.uniroma1.it/erbario/catalogo/gestionedb/scheda.asp?
inventario=1799]]>
      </linkform>
      <name>Campanula persicifolia L.</name>
      <description />
      <image>1799.jpg</image>
      <linkimage>
        <![CDATA[http://www.musei.uniroma1.it/dbinfo/RMSMUS12/JPEG/1799.jpg
]]>
      </linkimage>

```

```
-</object>
</path>
```

Per permettere agli insegnanti di costruire le proprie lezioni multimediali è stato messo a disposizione un CMS per la produzione automatica di pagine web in modalità ipertesto sviluppato all'interno di un precedente progetto per la produzione automatica di siti WEB dei musei [Ferrara et al, 2012]. Il CMS ASDscuola (Accessible Site Developer) è un sistema open source portabile per gestire i contenuti e si basa su tecnologie JAVA e XML e produce pagine XHTML.

L'utente può scegliere le immagini degli oggetti museali correlate ai contenuti delle lezioni dall'itinerario scaricato. Il CMS ASD editor permette di accedere al file XML con le informazioni e le immagini degli oggetti all'interno della pagina dell'ipertesto. Successivamente l'insegnante potrà gli studenti al museo dove gli stessi potranno ritrovare gli oggetti osservati durante la lezione.

Gli studenti potranno utilizzare lo stesso sistema per costruire le loro relazione successivamente alla visita attivando il modello collaborativo. Fig. 1

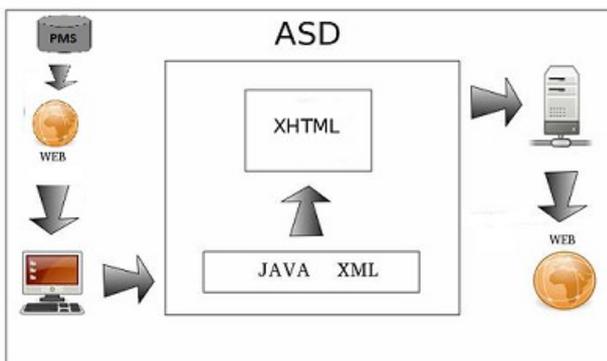


Fig. 1. Schema del progetto

Questa sperimentazione, i cui risultati e le lezioni multimediali prodotte dagli insegnanti sono disponibili sul sito web del progetto [Progetto MIUR, 2011], è stata vista con entusiasmo dalla classe docente ma ha fatto emergere delle criticità in merito ai contenuti relativi al patrimonio culturale resi disponibili che si provvederà a risolvere con un nuovo progetto finanziato dal MIUR.

E' stato rilevato, infatti, che spesso il linguaggio e la terminologia adottata dallo staff dei musei per descrivere il patrimonio non è facilmente accessibile e comprensibile per un pubblico non specialistico o per insegnanti e studenti.

Per risolvere questa questione la tecnologia può fornire l'opportunità di pensare a una modalità per l'implementazione e la rappresentazione dei contenuti. Il nuovo progetto avrà l'obiettivo di costruire un framework, basato sul web semantico, per sviluppare un tool di annotazione che incrementi il Repository con contenuti inseriti dagli insegnanti che possano così

soddisfare i diversi livelli di conoscenza e curiosità. [Dijkshoorn et al, 2012]. In questo modo sarà utile promuovere un ambiente virtuale per l'apprendimento basato su una attività di tipo collaborativo per condividere i contenuti e learning object (LO) con altre scuole.

4. Conclusioni e nuove prospettive

Una nuova strategia di comunicazione applicata alla teoria per l'apprendimenti e le tecnologie possono esse utiki per sviluppare relazioni migliori tra l'ambiente museale e quello didattico. I risultati del progetto presentato hanno dimostrato che le risorse digitali e gli strumenti tecnologici possono essere utilizzati come nuovo materiale didattico per costruire un ambiente di apprendimento. Un nuovo obiettivo sarà quello di valutare l'impatto delle tecnologie per verificare la qualità del coinvolgimento e dell'apprendimento da parte degli studenti. Il quadro proposto per l'insegnamento può essere considerato anche come una piattaforma utile per il riutilizzo delle risorse digitali nell'ambito della formazione permanente e del turismo culturale.

L'uso di Repository di metadati e gli ultimi sviluppi in Linked Open Data (LOD), per estendere la pubblicazione dei dati aperti, sta stimolando il disegno di nuovi RDF per creare nuovi modelli adatti per il collegamento di elementi di dati da fonti diverse. [Clough et al, 2011].

In conclusione, l'idea è di definire un nuovo modello di incremento e di utilizzo del patrimonio culturale, mediante la progettazione di tools che permettano l'annotazione da parte di utenti diversi dallo staff museali in modalità di tipo collaborativo adottando user-generated content tipica del WEB 2.0. Ogni oggetto culturale e scientifico può essere ri-contestualizzato in un LO e il suo riferimento reso accessibile mediante il contributo al disegno di un appropriato modello RDF per costruire un Repository dedicato al settore della formazione.

Questa modalità può stimolare la visita ai musei e migliorare le relazioni tra il contesto museale e quello didattico.

Bibliografia

[1] Aihara A., Yamada T., Kando N., Fujisawa S., Uehara Y., Baba T., Adachi J., Supporting creation and sharing of contents of cultural heritage objects for educational purposes in Purposes. Proc. of the Workshop on Personalized Access to Cultural Heritage (PATCH 2008), pp.5–14.

[2] Anthropology Museum Wake Forest University <http://moa.wfu.edu/>

[3] Ardissono L.; Kuflik T; Petrelli D., Personalization in cultural heritage: the road travelled and the one ahead in User Modeling and User-Adapted Interaction, 2012, pp. 73-99

[4] Aroyo L., Stash N., Wang Y., Gorgels P., Rutledge L., CHIP demonstrator: semantics-driven recommendations and museum tour generation in *The Semantic Web*, Springer Berlin Heidelberg, 2007, pp. 879-886

[5] Berners-Lee T., *Linked Data - Design Issues*. Retrieved April 30, 2011, <http://www.w3.org/DesignIssues/LinkedData.html>

[6] Bowen J., Filippini-Fantoni S., Personalization and the web from a museum perspective in *Proceedings of Museums on the Web Conference*, 2004, pp. 63-78.

[7] Burden K., Atkinson S., Personalising teaching and learning with digital resources: DiAL-e Framework case studies in Technology supported environment for personalised learning: Methods and case studies. Hershey, PA: IGI Global, 2009, pp.91-108

[8] Canadian Museum Center www.museevirtuel-virtualmuseum.ca/edu/Login.do?method=load&lang=en

[9] Clough P., Ford N., Stevenson M., Personalizing Access to Cultural Heritage Collections using Pathways, in *Proceedings of PATCH (2011)*, pp.12-19

[10] Dijkshoorn C., Oosterman J., Aroyo L., Houben G., Personalization in crowd-driven annotation for cultural heritage collections in E. Herder, K. Yacef, L. Chen, S. Weibelzahl (Eds.), *UMAP Workshops. CEUR Workshop Proceedings 872, CEUR-WS, 2012*. http://ceur-ws.org/Vol-872/patch2012_paper_3.pdf

[11] EU Commission Recommendation. "On the digitisation and online accessibility of cultural material and digital preservation", 2011, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:236:0028:0030:EN:PDF>

[12] Falk J., "Identity and the museum visitor experience", Walnut Creek: Left Coast Press, 2009.

[13] Ferrara V. ASD (Accessible Site Developer): un CMS per i Musei in: CONGRESSO NAZIONALE AICA 2009 Un nuovo "made in Italy" per lo sviluppo del Paese ICT per la valorizzazione dei beni e delle attività culturali. Roma, 4-6 novembre 2009

[14] Ferrara V., Sapia S., Macchia A., Risa E., Campanella L. La tecnologia per la promozione di una rete collaborativa tra scuola e musei. In: *Didamatica 2012 - Atti del Convegno*. Taranto, Italia, 2012

[15] Ferrara V., Campanella L. Polo Museale Sapienza: How promote University Museums in Arranging and rearranging: Planning university heritage for the future, PADOVA: Padova University Press, 2012, pp. 161-166.

[16] Gorgels P., *Rijksstudio: Make Your Own Masterpiece! Museum & Web* <http://mw2013.museumsandtheweb.com/paper/rijksstudio-make-your-own-masterpiece/>, 2013

[17] Haslhofer B., Isaac A., *data. europeana. eu: The Europeana Linked Open Data Pilot in International Conference on Dublin Core and Metadata Applications*, 2011, pp. 94-104.

[18] Kolb D. A., *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1984

[19] Lakhani K. R., Panetta J. A., *The Principles of Distributed Innovation in Innovations: Technology, Governance, Globalization*, 2007, pp. 97-112.

[20] *Linked Heritage Eu project* <http://www.linkedheritage.eu>

[21] Marie, J., The role of object-based learning in transferable skills development in University Museums and Collections Journal 3, 2010, pp. 187–190.

[22] Museo Galileo www.museogalileo.it/en/explore/onlinedidactic/scienceplay.html

[23] Oomen J., van Erp M., Baltussen L. B., Sharing cultural heritage the linked open data way: why you should sign up in Museums and the Web 2012, San Diego, CA, USA [www.museumsandtheweb.com/mw2012/papers/sharing_cultural_heritage_the_linked_o
pen_data](http://www.museumsandtheweb.com/mw2012/papers/sharing_cultural_heritage_the_linked_open_data)

[24] Polo Museale Sapienza <http://www.musei.uniroma1.it/eng>

[25] Portale Culturaitalia <http://dati.culturaitalia.it/sparql/home.jsp?locale=it>

[26] Progetto MIUR “L’oggetto museale come strumento per la didattica” www.musei.uniroma1.it/progettomiur/index.html

[27] Schmidt K., Schmitz H., Wolpers M., Developing a Network of Cultural Heritage Objects Repositories for Educational Purposes , in Proceedings of MTSR. 2011, pp. 337-348

[28] Tsalapati E., Simou N., Drosopoulos N., Stein R., Evolving LIDO based aggregations into Linked Data in CIDOC2012 - Enriching Cultural Heritage Helsinki, Finland <http://www.image.ece.ntua.gr/php/savepaper.php?id=767>, 2012

[29] Wang Y., ET A., Interactive User Modeling for Personalized Access to Museum Collections: The Rijksmuseum Case Study in Proceedings of the 11th international conference on User Modeling, Springer-Verlag Berlin, Heidelberg, 2007, pp.385-389.

Un sistema di Visual Analytics a supporto dell'interpretazione di incisioni rupestri

Vincenzo Deufemia, Luca Paolino, Valentina Indelli Pisano, Genoveffa Tortora
Department of Management & Information Technology (Università di Salerno)
Via Giovanni Paolo II, 132 (84084 Fisciano (SA))
{deufemia,lpaolino,tortora}@unisa.it

Abstract. *This paper presents a visual analytics system for supporting rock art archaeologists in exploring repositories of rock art scenes each consisting of hundreds of petroglyphs carved by ancient people on rocks. With their increasing complexity, analyzing these image repositories became a major task and challenge for rock art archaeologists. The proposed system combines visualization techniques with fuzzy-based analysis of rock art scenes to automatically reduce complexity in the study of petroglyph correlation. In this paper, we discuss how the proposed data views allow archaeologists to automatically infer information that is crucial for the correct interpretation of the scenes.*

Keywords: *Visual Analytics, fuzzy-based system, rock art.*

1. Introduzione

L'arte rupestre è un termine coniato in archeologia ed è riferito a qualsiasi tipo d'incisione fatta dall'uomo sulla pietra [1]. Tali incisioni, dette anche petroglifi o graffiti, sono state create rimuovendo la parte superficiale della roccia con tecniche di picchiettatura, intaglio e/o abrasione. Sebbene sia impossibile associargli interpretazioni certe, gli archeologi hanno proposto molte teorie che cercano di dare un significato ai petroglifi. Ad esempio l'immagine in Fig. 1(a) ha un significato religioso poiché è stata interpretata come un Cristo [2], mentre l'immagine in Fig. 1(b) ha un significato astronomico essendo stata interpretata come la costellazione delle Pleiadi [9].

Per preservare, studiare ed interpretare tali artefatti, gli archeologi hanno raccolto ed archiviato differenti tipi di informazioni come immagini, descrizioni testuali, coordinate GPS, rilievi in bianco e nero, ecc. La gestione e l'analisi di questi archivi, contenenti migliaia di dati a più dimensioni, è diventata un grosso problema ed un'importante sfida per gli archeologi rupestri. Seppure negli ultimi anni siano stati proposti molti metodi per automatizzare il riconoscimento dei petroglifi [6], pochi sforzi sono stati indirizzati a sviluppare strumenti che supportino gli archeologi nell'analisi di questi enormi archivi. Al fine di interpretare correttamente le scene contenenti simboli petroglifi, oltre all'analisi delle forme delle incisioni, è importante prendere in considerazione le relazioni

spaziali che intercorrono tra i simboli. Infatti, esiste un'alta correlazione tra la disposizione spaziale dei simboli nella scena e la loro interpretazione. Tuttavia, l'identificazione di tali relazioni è un compito molto impegnativo a causa delle dimensioni degli archivi e della complessità dei dati in essi contenuti.

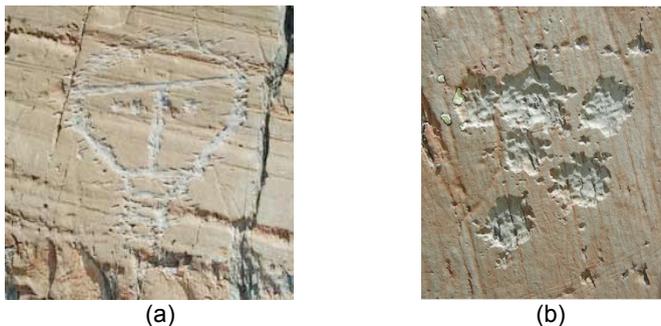


Fig. 1. (a) Foto di un petroglifo interpretato come un Cristo [2], (b) foto di un petroglifo interpretato come la costellazione delle Pleiadi [9].

Il nostro lavoro di ricerca ha avuto come obiettivo l'analisi di archivi di petroglifi utilizzando tecniche di visualizzazione che sfruttano le informazioni semantiche estratte dalle interpretazioni. Queste ultime fanno parte di un'ontologia che organizza i dati associati ai petroglifi, incluso le annotazioni semantiche e le possibili relazioni tra di stessi. Abbiamo quindi introdotto nuove tecniche di astrazione interattiva basate su ontologie per l'analisi visiva di grandi archivi di petroglifi. Tali tecniche d'astrazione si basano su informazioni semantiche e strutturali, e consentono agli utenti di individuare facilmente similitudini tra le forme e le relazioni dei petroglifi. Per ridurre la complessità nello studio della correlazione tra i petroglifi abbiamo introdotto una visualizzazione di dati che utilizza le relazioni spaziali fuzzy per determinare la struttura topologica delle scene e riassumere visivamente queste informazioni. Tali visualizzazioni possono essere usate per dedurre informazioni utili alla corretta interpretazione delle scene e alla validazione di nuove eventuali ipotesi.

Per validare le tecniche di visualizzazione proposte abbiamo sviluppato un sistema di *Visual Analytics* che supporta gli archeologi nelle loro attività di analisi considerando la struttura topologica delle immagini contenenti i petroglifi. Tale prototipo è stato, quindi, valutato utilizzando i dati relativi all'archivio dei petroglifi del Monte Bego.

L'articolo è organizzato come segue: nella Sezione 2 viene presentato lo stato dell'arte riguardante la visualizzazione delle informazioni archeologiche, la Sezione 3 introduce informazioni sull'arte rupestre e su una piattaforma per la conservazione digitale dei petroglifi, la Sezione 4 introduce il sistema proposto e il processo di analisi visiva interattiva, infine, le conclusioni sono discusse nella Sezione 5.

2. Stato dell'arte

Gli strumenti GIS sono stati ampiamente utilizzati nell'archeologia per le attività di analisi e ricerca [13]. Infatti, le grandi quantità di dati che devono essere interpretate dagli archeologi sono più facilmente comprensibili se visualizzati su mappe. Alcuni strumenti GIS sono stati potenziati ed utilizzati in applicazioni per il supporto decisionale, l'analisi e la modellazione [11][18].

L'ambiente di analisi geo-visivo presentato in [10] sfrutta il cubo spazio-tempo per analizzare le relazioni tra siti e reperti al fine di comprendere le interazioni tra le culture. Il cubo spazio-tempo permette di rappresentare le tre componenti principali dei dati spazio-temporali: quando, dove e cosa. Tale cubo visualizza lo spazio-tempo dei dati archeologici e fornisce funzioni interattive di filtro ed ordinamento che possono essere applicate per mostrare pattern e relazioni nascoste nei dati.

Il progetto *Cyber-Archaeology* si focalizza sullo sviluppo di tecnologie e strumenti per la documentazione dei beni culturali [12]. L'elemento centrale di *Cyber-Archaeology* è un sistema di analisi visiva che supporta l'analisi collaborativa di dati multispettrali riguardanti le grandi scale di tempo e di spazio. Questo sistema può anche essere utilizzato per guidare la scoperta di nuove informazioni, comprendere il significato di un nuovo manufatto, determinare se un manufatto deve essere monitorato o richiede un qualche tipo d'intervento.

In [3] la Visual Analytics è stata utilizzata per facilitare l'interpretazione delle immagini termografiche multi-temporali per fini di restauro dei beni culturali. L'ambiente visivo proposto consente di esplorare i dati termografici dal punto di vista spazio-temporale. Tale ambiente mira a rilevare pattern spaziali e spazio-temporali che potrebbero fornire informazioni sulla struttura e sul livello di degrado del manufatto.

ArchMatrix è un sistema visuale ed interattivo che supporta gli archeologi nell'archiviazione, nella gestione e nello studio dei reperti raccolti durante gli scavi archeologici [17]. *ArchMatrix* implementa il metodo *Matrix Harris*, utilizzato in archeologia per descrivere la posizione di unità stratigrafiche, e si basa su un archivio di grafi che consentono di applicare tecniche avanzate per il recupero dell'informazione.

Indiana Finder è un sistema di visualizzazione interattivo che supporta gli archeologi nell'analisi di grandi archivi di documenti e disegni [8]. Il sistema fornisce un supporto per attività investigative legate all'interpretazione di nuovi reperti, l'individuazione di anomalie nell'interpretazione e la scoperta di nuove informazioni.

Il sistema proposto in questo lavoro si differenzia dagli approcci precedenti sia in termini di tipologia di analisi che nel modo in cui le informazioni sono rappresentate. Nel nostro caso, infatti, le modalità di rappresentazione dei dati hanno lo scopo di fornire sinteticamente informazioni relative alla correlazione spaziale tra i simboli petroglifici presenti nelle scene e la loro interpretazione.

3.La preservazione digitale dell'arte

I petroglifi sono la più antica forma d'arte dell'uomo. Un petroglifo consiste in un unico simbolo scolpito in una roccia, come in Fig. 1(a), oppure può essere la composizione di due o più simboli, come nella Fig. 2(a). Solitamente gli archeologi che lavorano su un sito definiscono un dizionario di simboli considerando principalmente la forma dei petroglifi e cercano di dare un'interpretazione alle aggregazioni di più petroglifi, chiamate *scene*.

Le scene non solo rappresentano concetti reali della vita quotidiana, come gruppi di animali o armi, ma anche concetti astratti, come credenze religiose, idee sulla guerra e stili di vita dei popoli di quell'epoca. Anche se in alcuni casi i simboli che compongono una scena sembrano essere disposti senza alcun ordine, in molti altri casi esistono delle relazioni ricorrenti tra di essi e che sono utili alla corretta interpretazione della scena [2][4]. Ad esempio, Fig. 2(b) mostra un simbolo interpretato come *Dea madre* situato sopra un simbolo interpretato come *toro*. Dall'analisi dei petroglifi presenti nella zona del Monte Bego gli archeologi hanno osservato che la disposizione di questa coppia di simboli ricorre frequentemente ed è interpretata con il concetto *nascita* [4]. In questo senso tale combinazione di simboli può essere vista come un *pattern* che deve essere interpretato nella sua interezza piuttosto che nelle sue singole componenti.



(a)



(b)

Fig. 2. (a) una scena composta da diversi corniformi ed alabarde [2], (b) il concetto *nascita* rappresentato dal pattern *dea madre* che da alla luce il *toro* [4].

Per promuovere la conoscenza e la conservazione dell'arte rupestre e per supportare gli archeologi nella loro attività d'indagine la piattaforma *IndianaMAS* [14][15] mira ad integrare e completare le tecniche normalmente adottate per preservare i siti del patrimonio rupestre. L'obiettivo di *IndianaMAS* è la protezione e la conservazione digitale dei siti che sono patrimonio naturale e culturale di arte rupestre, memorizzando, organizzando ed esponendo informazioni riguardo questi siti in modo che venga incoraggiata la ricerca scientifica e che sia destato l'interesse e la sensibilità della comunità verso essi. La piattaforma *IndianaMAS* integra agenti intelligenti, ontologie, procedure per l'elaborazione del linguaggio naturale in differenti lingue, algoritmi per il riconoscimento delle immagini, algoritmi per l'integrazione dei dati [5]. Le

ontologie permettono di definire un vocabolario comune che può essere sfruttato efficacemente per organizzare i dati associati alle incisioni rupestri e per creare relazioni semantiche tra i dati. Gli algoritmi per il riconoscimento delle immagini vengono applicate per classificare singoli petroglifi.

Il database di petroglifi contiene le descrizioni testuali delle interpretazioni di scene, le informazioni ontologiche, le coordinate geografiche, le foto ed i disegni dei petroglifi. L'informazione ontologica è un elenco di metadati che riassume l'interpretazione che l'archeologo ha associato ad una scena. La Fig. 3 mostra una parte dell'ontologia che è stata definita per il dominio dell'arte rupestre.

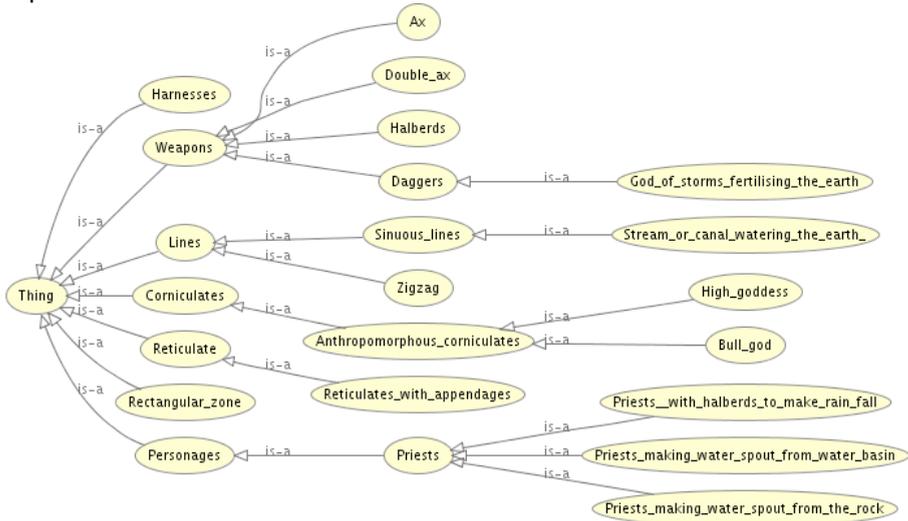


Fig. 3. Parte dell'ontologia definita per il dominio dell'arte rupestre.

4. Il sistema di Visual Analytics proposto

Come detto in precedenza, i pattern possono essere molto importanti per la loro corretta interpretazione delle scene in cui sono presenti. Finora gli archeologi hanno identificato questi pattern in maniera empirica. In questa sezione presentiamo un sistema di Visual Analytics il cui scopo è di supportare gli archeologi nell'individuazione e nell'analisi sistematica di tali pattern.

Il sistema proposto è uno strumento di analisi visiva progettato per sfruttare l'informazione semantica (ad esempio i metadati presenti nell'ontologia) associata ai petroglifi interpretati. Esso consente agli archeologi di isolare facilmente i petroglifi e visualizzare le loro relazioni più frequenti al fine di poter inferire e/o convalidare nuovi pattern. Questo strumento è in grado di elaborare un'astrazione strutturale, aggregando i petroglifi aventi forma simile, e un'astrazione semantica, sfruttando l'informazione ontologica associata alle scene, e di visualizzarne i risultati attraverso la vista Bubble. Inoltre, le relazioni topologiche e direzionali tra i petroglifi sono calcolate attraverso l'utilizzo della teoria fuzzy e visualizzati utilizzando la vista Ring.

4.1 Astrazione Semantica usando la Vista Bubble

Sia $S=\{s_1,\dots,s_n\}$ l'insieme dei simboli definito dagli archeologi per classificare un petroglifo e $O=\{o_1,\dots,o_m\}$ l'insieme dei concetti contenuti nell'ontologia. L'insieme P delle scene memorizzate nell'archivio può essere formalmente definito come $P=\{p_1,\dots,p_l\}$ dove $p_i=\{r_{i,1},\dots,r_{i,q}\}$ contiene i petroglifi della i -esima scena, $Shape(r_{ij})=s_k \in S$, e $Sem(p_i)\subseteq O$. La funzione $Shape$ determina il tipo di petroglifo utilizzando un algoritmo di *image matching*, mentre la funzione Sem determina l'insieme dei concetti ontologici associati ad un petroglifo.

A partire da queste informazioni, il sistema di Visual Analytics è in grado di costruire un grafo di tipi di simboli petroglifi a partire dall'insieme delle scene in P , includendo solo i nodi i cui tipi sono correlati con un concetto nell'ontologia. L'obiettivo è di visualizzare i tipi di simboli e le relazioni che co-occorrono con un determinato concetto ontologico. Per ottenere ciò è stata utilizzata l'astrazione per nascondere i tipi di simboli non rilevanti con i concetti ontologici selezionati ed il clustering per raggruppare i petroglifi aventi forma simile. Un'astrazione rispetto a un concetto ontologico m è definita come un grafo $G_m=(V, E)$ con $V = \{s_x: s_x \in S \text{ ed } \exists r_{j,k} \in p_j \text{ tale che } Shape(r_{j,k})=s_x \text{ e } m \in Sem(p_j)\}$ e $E = \{(s_x, s_y): s_x, s_y \in S \text{ e } \exists r_{j,k} \in p_j \text{ e } r_{j,n} \in p_l \text{ con } Shape(r_{j,k}) = s_x, Shape(r_{j,n}) = s_y, \text{ e } m \in Sem(p_j) \text{ e tali che esiste un rapporto spaziale tra } l_{j,k} \text{ e } l_{j,n}\}$. Quindi, il grafo costruito per il concetto ontologico m contiene i tipi di simboli co-occorrenti con m e le relazioni spaziali tra di essi. Nel sistema proposto la funzione $Shape$ è stata implementata utilizzando l'algoritmo proposto in [7], mentre il metodo per il calcolo delle relazioni spaziali tra i petroglifi nelle scene è descritto nella successiva sottosezione.

Il sistema visualizza l'astrazione semantica utilizzando la vista Bubble. In particolare, gli archeologi possono selezionare i concetti ontologici navigando l'ontologia attraverso la relazione di tassonomia (IS-A) o interagendo con una barra per selezionare il livello di granularità. La vista mostra una Bubble, rappresentata con un cerchio, per ogni concetto ontologico selezionato. L'intensità di colore della Bubble rappresenta la frequenza nelle scene del corrispondente concetto ontologico (colori chiari = poco frequente, colore scuro = molto frequente). La somiglianza semantica tra due concetti ontologici è calcolata come la distanza tra i corrispondenti nodi nell'albero ontologico ed è visualmente rappresentata utilizzando colori simili (colori simili = concetti ontologici semanticamente simili). La Bubble per il concetto ontologico m contiene l'informazione più importante per il grafo G_m . In particolare, la vista visualizza i simboli più frequenti che co-occorrono con m insieme alle loro relazioni. Anche qui, la dimensione di un simbolo indica quanto esso è frequente nelle scene.

L'astrazione basata sull'informazione ontologica fornisce una semplice vista per aiutare ad isolare i simboli chiave e le relazioni in un sottoinsieme di scene di petroglifi. Questa informazione, che è visualmente evidenziata nella vista Bubble, consente agli archeologi di scoprire pattern candidati. La Fig. 4 mostra la vista Bubble costruita per i concetti ontologici: *fertilizzazione*, *tempesta*, *nascita*, e *sacrificio*. Dal colore delle Bubble possiamo osservare che *tempesta* e *fertilizzazione* sono abbastanza vicini nell'albero ontologico. Inoltre la dimensione della Bubble indica che il concetto di *sacrificio* è meno frequente

nelle scene rispetto agli altri concetti. Ogni Bubble contiene il grafo ottenuto dal processo di astrazione. Il grafo all'interno della bolla *sacrificio* evidenzia un possibile pattern poiché i simboli *toro* e *alabarda* sono molto frequenti nelle scene che hanno il metadato *sacrificio*. Inoltre, questi due simboli sono spazialmente correlati in queste scene.

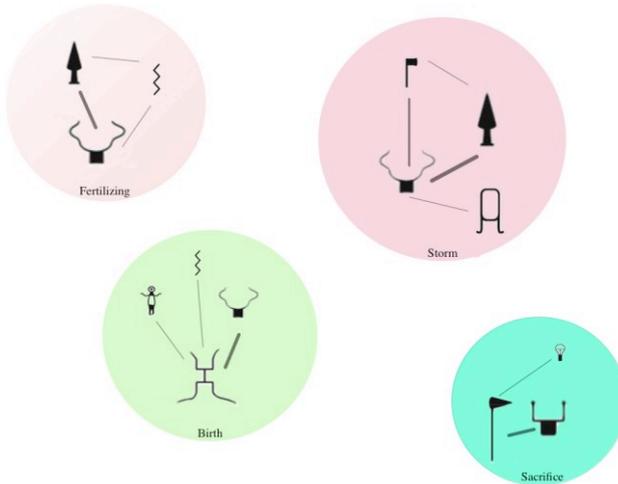


Fig. 4. La vista Bubble costruita per i concetti: *fertilizzazione, tempesta, nascita, e sacrificio.*

4.2 Analisi spaziale usando le relazioni Fuzzy

L'informazione visualizzata nella vista Bubble fornisce una panoramica della correlazione tra le relazioni tra i simboli e i concetti ontologici estratti dalle loro interpretazioni. Quando l'archeologo scopre un pattern candidato, egli ha bisogno di informazioni più dettagliate sulle relazioni dei simboli coinvolti nel pattern al fine di stabilire se esso è o meno un pattern. In particolare, l'archeologo deve analizzare il tipo di relazione spaziale esistente tra i simboli dei petroglifi e come tali simboli siano legati al metadato selezionato. Poiché i simboli sono stati incisi in maniera poco accurata, le relazioni spaziali tra questi simboli non possono essere calcolate utilizzando relazioni precise (crisp). Di conseguenza applichiamo il metodo proposto in [16] che combina relazioni topologiche fuzzy e relazioni direzionali nello spazio bidimensionale. In particolare, al fine di ottenere la struttura topologica di una scena usiamo il *Region Connection Calculus* (RCC), mentre le proprietà direzionali tra i simboli sono state elaborate usando il *Cardinal Direction Calculus* (CDC). Una matrice 3x8 è utilizzata per rappresentare le relazioni spaziali fuzzy. Le righe rappresentano le relazioni topologiche definite nella RCC3 (*disjoint, meet, partially overlap*) e le colonne rappresentano gli aspetti direzionali qualitativi della scena 2D come definito nel CDC8 (*est, nord-est, nord, nord-ovest, ovest, sud-ovest, sud e sud-est*). Le relazioni sono espresse in valori numerici, dove ogni valore rappresenta la percentuale di area tra due oggetti che corrisponde a una specifica relazione topologica in una certa direzione. Più precisamente, il

valore della cella per una relazione topologica di una data direzione è calcolato utilizzando il contributo fuzzy delle corrispondenti relazioni di Allen calcolate su specifici angoli derivanti dalla direzione.

Per valutare la relazione topologica e direzionale fuzzy, il sistema analizza la frequenza (quante volte la relazione è soddisfatta) e la forza (quanto è grande il valore fuzzy) delle relazioni e visualizza queste informazioni in una vista Ring. Per rappresentare visualmente questa informazione, la vista Ring mostra tre cerchi concentrici che corrispondono alle relazioni topologiche. Il cerchio più interno rappresenta la relazione *partially overlap*, il cerchio successivo rappresenta la relazione *meet* e infine il cerchio più esterno rappresenta la relazione di *disjoint*. Ogni cerchio è suddiviso in otto settori che corrispondono alle relazioni direzionali. I valori fuzzy ottenuto da una relazione topologica-direzionale tra due petroglifi sono mappati nella corrispondente cella del Ring. In particolare, abbiamo proposto due diverse visualizzazioni delle relazioni spaziali fuzzy. La prima colora ogni settore della vista Ring con la media dei valori fuzzy della corrispondente relazione fuzzy. Il colore rosso è usato per il settore *disjoint*, il colore viola per il settore *meet* e il colore verde per il settore *partially overlap* come mostrato in Fig. 5(a). Si noti che il nero indica l'assenza di relazioni, e la saturazione del colore mostra la forza della relazione, più il colore è saturo più la relazione è forte. Nella seconda visualizzazione i valori fuzzy sono mappati sui gradienti di colore, che evidenzia allo stesso tempo la frequenza e la forza delle relazioni. Come mostrato in Fig. 5(b), il colore grigio indica l'assenza di relazioni, mentre il colore viola ne indica la presenza. L'intensità del colore da un indice della forza della relazione.

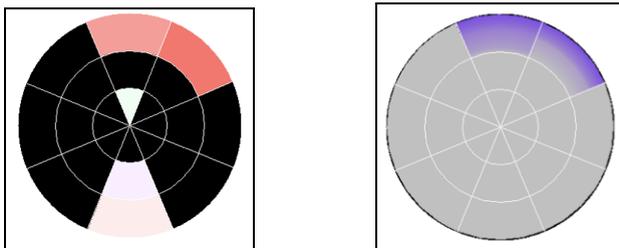


Fig. 5. Due diversi tipi di visualizzazione della vista Ring: (a) valori medi fuzzy e (b) tecnica del gradiente.

Gli archeologi possono anche interagire con i settori delle viste Ring per analizzare i metadati associati alle scene dove vale la relazione selezionata. La vista Metadata filtra l'insieme delle scene basata sulla relazione spaziale selezionata e mostra i metadati più frequenti utilizzando la visualizzazione *Tag Cloud*. Ad esempio, Fig. 6 mostra la vista Metadata generata selezionando il settore *Nord-Disjoint* del Ring. I termini più frequenti nelle scene che hanno i simboli toro e dea in relazione *Nord-Disjoint* sono: *Weapons*, *Double_Ax* e *Ax*. Utilizzando questa vista gli archeologi hanno una conferma del pattern scoperto.

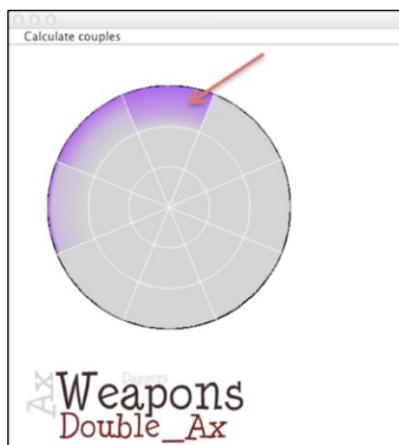


Fig. 6. Vista Ring e vista Metadata

5. Conclusioni

In quest'articolo abbiamo presentato uno strumento di Visual Analytics che consente di analizzare grandi archivi di petroglifi. Questo sistema supporta gli archeologi nell'individuazione di pattern ricorrenti tra i petroglifi, fornendo viste che evidenziano le correlazioni tra le relazioni spaziali tra i simboli petroglifi e le informazioni semantiche estratte dalle loro interpretazioni. Tali informazioni sono estratte da un'ontologia utilizzata per organizzare i dati associati ai petroglifi, incluse annotazioni semantiche e relazioni tra di essi. Le tecniche di astrazione utilizzate per la costruzione delle viste si basano sulle informazioni semantiche e strutturali dei petroglifi, mentre la teoria fuzzy è utilizzata per valutare le relazioni spaziali tra i simboli nelle scene.

Bibliografia

- [1] Bednarik R., *Rock Art Science: The Scientific Study of Palaeoart*, IFRAO-Brepols, 2001.
- [2] Bianchi N., *Mount Bego: Prehistoric Rock Carvings*, Adoranten, Scandinavian Society for Prehistoric Art, 70–80, 2010.
- [3] Danese, M., Demšar, U., Masini, N., Charlton, M., *Investigating Material Decay of Historic Buildings using Visual Analytics with Multi-Temporal Infrared Thermographic Data*. *Archaeometry*, 52, 3, 2010, 482–501.
- [4] de Lumley, H., Echassoux, A., *The Rock Carvings of the Chalcolithic and Ancient Bronze Age from the Mont Bego Area, The Cosmogonic Myths of the Early Metallurgic Settlers in the Southern Alps*. *L'Anthropologie*, 113, 5P2, 2009, 969–1004.
- [5] Deufemia V., Giordano M., Polese G., Tortora G., *A Visual Language based System for ETL Development, Software - Practice and Experience*, 2013, DOI: 10.1002/spe.2201.
- [6] Deufemia, V. Paolino, L., *Combining Unsupervised Clustering with a Non-Linear Deformation Model for Efficient Petroglyph Recognition*. In: *Advances in Visual*

Computing. Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York (2013).

[7] Deufemia V., Paolino L., de Lumley H., Petroglyph Recognition using Self-Organizing Maps and Fuzzy Visual Language Parsing, Proc. of International Conference on Tools with Artificial Intelligence (ICTAI'12), IEEE Press, 2012, 852–859.

[8] Deufemia V., Paolino L., Tortora G., Traverso A., Mascardi V., Ancona M., Martelli M., Bianchi N., de Lumley H., Investigative Analysis across Documents and Drawings: Visual Analytics for Archaeologists, in Proc. of the International Working Conference on Advanced Visual Interfaces (AVI'12), ACM Press, 2012, 539–546.

[9] Echassoux, A., de Lumley, H., Pecker, J.-C., Rocher, P., Les Gravures Rupestres des Pléiades de la Montagne Sacrée du Bego, Tende, Alpes-Maritimes, France. C. R. Palevol, 8, 2009, 461–469.

[10] Huisman, O., Feliciano Santiago, I., Kraak, M.-J., Retsios, B., Developing a Geovisual Analytics Environment for Investigating Archaeological Events: Extending the Space-Time Cube. Cartography and Geographic Inf. Science, 36, 3, 2009, 225–236.

[11] Kohler T. A., J. Kresl C. V. West, E. Carr, R. H. Wilshusen, A modeling Approach to Settlement Determinants and Spatial Efficiency Among Late Ancestral Pueblo Populations of the Mesa Verde Region, U.S. Southwest, in Kohler T. A. e Gumerman G. J. (eds) Dynamics in Human and Primate Societies, Oxford Univ. Press, 2000, 145–178.

[12] Levy T. E., Petrovic V., Wypych T., Gidding A., Knabb K., Hernandez D., Smith N. G., Schlulz J. P., Savage S. H., Kuester F., Ben-Yosef E., Buitenhuis C., Barrett C. J., Najjar M., DeFanti T., On-Site Digital Archaeology 3.0 and Cyber-Archaeology: Into the Future of the Past – New Developments, Delivery and the Creation of a Data Avalanche, in Forte M. (ed), Introduction to Cyber-Archaeology, Oxford, Archaeopress, 2010, 135–153.

[13] Lock G., e Stancic Z., Archaeology and GIS: A European Perspective, Taylor and Francis, London, 1995.

[14] Mascardi, V., Deufemia, V., Malafronte, D., Ricciarelli, A., Bianchi, N., and de Lumley, H.: Rock Art Interpretation within Indiana MAS. In: Jezic, G., Kusek, M., Nguyen, N.-T., Howlett, R., and Jain, L. (eds.): Agent and Multi-Agent Systems. Technologies and Applications. Lecture Notes in Computer Science, Vol. 7327. Springer-Verlag, Berlin Heidelberg New York (2012) 271–281.

[15] Mascardi, V., Briola, D., Locoro, A., Grignani, D., Deufemia, V., Paolino, L., Bianchi, N., de Lumley, H., Malafronte, D., and Ricciarelli, A.: A Holonic Multi-Agent System for Sketch, Image and Text Interpretation in the Rock Art Domain. International Journal of Innovative Computing Information and Control, 10, 1, 2014.

[16] Salamat, N., Zahzah, E., Two-Dimensional Fuzzy Spatial Relations: A New Way of Computing and Representation. Advances in Fuzzy Systems, 2012, 15pp.

[17] Valtolina S. e Barricelli B.R. ArchMatrix: a Visual Interactive System for Graph-based Knowledge Exploration in Archaeology, Proc. of the International Working Conference on Advanced Visual Interfaces (AVI'12), ACM Press, 2012, 681–684.

[18] van Leusen P., GIS and Archaeological Resource Management: a European Agenda, Lock G. Stancic, Z. (eds) Archaeology and Geographical Information Systems, A European perspective, London, Taylor and Francis, 1995.

Il laboratorio leonardesco come bottega rinascimentale

Mauro M. Langfelder
editore: *dispacci PDM - implementi BLM*
Via Casati 30 - 23880 Casatenovo (Lecco)
maurolangfelder@yahoo.it

Abstract. *The Introduction to Leonardo3, or The World of Leonardo, the show curring in Milan, deals with the wood models of his machines and the successful edutainment role it is getting on its visitors. Such two themes are central in the following two sections, related with the WoodLab sourcing close to Milan and the interaction of all ages visiting people and touching the digital small wood models, accompanied by the thousands of documents he wrote and drafted. The fourth section is trying to prove how the 4000 pages of Leonardo's documents, including his most famous and divulged books, and the 200 small wood models of his machines are forming a sort of digital eciclopaedia, enriched by lectures in various different languages. So that Leonardo might be defined as the precursor, if not the true founder, of the so called Interoperabilty, the new university discipline proposed by the IAI ABC department of Milan Polytechnic in terms of building and urban constructions. In so doing the conclusion informs that permanent education courses, under execution by IAI ABC will present the WoodLab as a tool for the launch of the restoring of abandoned industry villages sorrounding Milan as a proving drive to wood prefabrication.*

Keywords: edutainment, laboratory, renaissance

1. Introduzione. Leonardo3: legno ed edutainment

La mostra **Leonardo3** sul mondo di Leonardo da Vinci, corrente presso le Sale del Re in piazza Scala a Milano, presentando 200 sue macchine in modo modellato, centra uno dei temi del 50° congresso AICA. Lo fa attraverso l'esposizione di modellini in legno, accompagnati dall'illustrazione digitale delle pagine dei manoscritti cartacei originali, fruibili in modo selettivo in numero di oltre 4000..

La mostra riprende diverse rassegne, precedenti e parziali, tali di per sé da fare una ricca storia; ma a Leonardo3 è stato riconosciuto il Premio di Rappresentanza del Presidente della Repubblica Italiana.

Congresso Nazionale AICA 2013

La rassegna naturalmente comprende anche alcune opere artistiche che Leonardo ebbe modo di eseguire nei suoi trent'anni di vita a Milano. La presentazione interattiva dell'*Ultima Cena*, ricostruita anche nei suoi colori originari ed il monumento equestre a Francesco Sforza, opere contemporanee di Leonardo, completano la rassegna degli esempi di tutte le macchine coprenti di fatto vari percorsi scientifici dei dipartimenti del Politecnico di Milano, ma anche in particolare la Città Ideale di Leonardo, autentica antesignana delle *smart cities*.

Due sono soprattutto i motivi che ci inducono a circostanziare Leonardo3 in forma di *full paper*. Il rilievo educativo della rassegna, che si propone anche in forma di piacevole visita, come esempio probante di **edutainment**. Mentre la realizzazione dei modellini in **legno**, ci riconduce al materiale da noi scelto come base per il recupero di un opificio industriale, già dedicato ai tessuti elastici ed ora progettato nel suo riassetto edilizio.

Le possibilità offerte dalla prefabbricazione con un materiale emergente in forma di **WoodLab**, ci ha portato anche ad illustrarlo nelle slide di alcuni moduli di corsi di formazione permanente al Politecnico di Milano presso un dipartimento accomunante le facoltà di Ingegneria e di Architettura, che ha fondato il Capitolo Italiano dell'Alleanza Internazionale per l'Interoperabilità o **buildingSMART**.

Questo contributo in forma di *full paper* vuole enfatizzare l'interoperabilità come nuova materia di insegnamento universitario, partendo appunto dall'**ABC** (*Architecture, Built environment and Construction engineering*) del **Politecnico** milanese, per contribuire ad irradiarla verso l'intero sistema degli atenei italiani come fondante argomento e disciplina interdisciplinare; ma anche penetrare progressivamente con AICA nel mondo diffuso della scuola.

Peraltro un *poster* parallelo vuol dare significato espressivo al ruolo di Milano e di un Leonardo ricordato in monumento fra quattro dei suoi discepoli eletti per rinverdire un modo di fare scuola

2. Arte e scienza tornano in laboratorio

Il fatto che anche la maggior parte dei manoscritti di Leonardo siano presenti in forma digitale consente un'elaborazione integrata fra tutti i modelli meccanici ed i disegni e documenti pertinenti.

Il lavoro eccezionale ed eccellente che è stato fatto nella realizzazione dei modelli in legno di tutte le macchine si colloca in un momento in cui il legno viene nettamente rivalutato anche in Italia come materiale per costruzioni, soprattutto se considerato negli aspetti che caratterizzano la prefabbricazione e la conseguente netta riduzione dei tempi (e dei disturbi) di cantiere.

Un **WoodLab**, dunque una forma di FabLab, di particolare significato è quello cui si può pensare, guardando alla realtà della mostra; ma forse sarebbe più proprio, data la diversificazione di tutte le macchine, mirare ad una metropoli industriale estesa, come insieme di quartieri industriali dedicati.

Nei corsi che la formazione permanente del Politecnico di Milano dedica alla modellazione digitale del fabbricato (*building information modeling*) o **BIM** ed all'interoperabilità uno spazio adeguato potrà essere trovato per presentare

anche quanto qui descritto, quindi riutilizzato per l'educazione remota, estesa anche fuori dell'ambito degli atenei, particolarmente presso i professionisti di un po' tutte le categorie. Spesso infatti si ignora che Leonardo era figlio e nipote di notaio, in quel di Vinci ...

3. L'edutainment e le materie per godere dell'innovazione

La mostra Leonardo3, oltre alle lingue caratterizzanti la rassegna, verrà a breve completata da tedesco, francese, russo, cinese nell'accompagnamento dei visitatori. L'iniziativa tende a preparare l'accoglienza e l'ospitalità della città che protende verso il suo 2015 con un impegno appena crescente, mentre i preposti sono particolarmente concentrati sul dopo-EXPO.

La promozione dell'interoperabilità come disciplina universitaria e materia di insegnamento scolastico viene condotta in modo particolare, oltre che dal rettorato del Politecnico e dal ministero competente, anche dalla Confindustria che vi vede un'opportunità per il rilancio dell'economia, soprattutto sul piano internazionale. Ma anche vari ordini istituzionali, stimolati in modo particolare da ingegneri ed architetti, allargati a tutte le professioni intellettuali stanno muovendosi in questa stessa direzione.

Tuttavia l'aspetto che interessa di più in senso generale è quello di avviare attività educative in e-learning, partendo dalla formazione dei formatori, anche per affrontare l'edutainment nei suoi aspetti più diversificati ...

Il genio di Leonardo pare consolidarsi nella reciprocità del rapporto docente-discente divenuto simmetrico ed invitando tutti a riflettere (come qui si propone in modo esplicito) sul ruolo dell'interoperabilità praticata e divulgata nello spirito dell'avvicinamento delle arti e delle scienze, non a caso accomunate nella Firenze rinascimentale.

4. I libri che fanno un'enciclopedia digitale

Un pur breve cenno sulla storia milanese (poi globale) della presentazione (e della genesi) del materiale che sostanzia la mostra di piazza sella Scala, può essere opportuna. Già nel 2005 il Codice Atlantico fu presentato nella Biblioteca Ambrosiana con la collaborazione della Curia milanese e fu un anno di successi. Mentre a Vigevano iniziò la presentazione del Laboratorio di Leonardo nel 2009, dove ebbe 120 mila visitatori. Ma i **successi** lombardi si replicarono un po' in tutto il mondo negli anni seguenti e ci piace ricordare almeno: New York, Chicago, Toronto, Filadelfia, Città del Messico, San Paolo, Doha ... Anche Torino la ospitò; e così fu per Livorno, la città che i bocconiani ricordano come quella della genesi in Italia della famiglia di Angelo Sraffa, il primo preside e di fatto il fondatore dell'ateneo milanese

Il Mondo di Leonardo presenta vari libri, fra cui emerge il **Codice Atlantico**, ben noto al mondo della cultura, con le 1200 pagine descrittive delle invenzioni. Così per il **Manoscritto B**, e forse ancor più per il **Codice del Volo**, che impressiona per l'attualità delle specifiche invenzioni. La consultazione delle figure impaginate e l'abbinamento ai modelli rende la visita interessante

per tutti, dando forma ad una sorta di enciclopedia digitale, interessante tutte le età e culture, arricchita dal commento sonoro in più lingue, che pure qualifica gli **ebook**.

La possibilità di consultazione digitale delle parti delle macchine anima le invenzioni, rendendole accessibili a tutti, quindi diffuse. Tale piena espressione di raccolta di documenti, manoscritti carte o disegni, quindi fogli (di libri) digitali, può dare la spinta per favorire una diffusione prorompente degli **ebook**, proprio da Milano, che si configura anche come capitale dell'editoria.

I libri citati sono una vera **enciclopedia** digitale, e la visita alla mostra li rende fruibili per una diffusione compatibile e guidata, prolungata fino all'Expo del 2015, anche per far risaltare i valori scientifici che esprime, e magari protratta anche oltre..

5. Leonardo precursore dell'interoperabilità

Questo *full paper*, accompagnato forse anche dal poster correlato, almeno in modo da costituirne una figura, intende far cogliere anche il ruolo che si può svolgere per guidare i visitatori a penetrare nello spirito di Leonardo.

L'interoperabilità, votata ad essere materia di insegnamento universitario in un numero crescente di atenei, trova nella varietà delle macchine di Leonardo una base storica di rilievo, che ne fa un precursore se non proprio il maestro

Un libro sulla Milano di Leonardo potrebbe assistere i **giovani** che vorranno essere accompagnatori volontari dei visitatori, ed il suo ebook complementare potrebbe essere dispensato ai **visitatori** (ed i loro collegati) che potranno meglio conoscere anche i protagonisti della storia di Milano, attraverso i monumenti dedicati, cominciando da quello di piazza della Scala, arricchito dai discepoli ...

Sarebbe una conferma anche del fatto che la nuova disciplina o materia che si propone difficilmente potrebbe avere un Maestro migliore.

6. Conclusioni. Il WoodLab per ricostruire

Per la bibliografia si rimanda sia alla documentazione della mostra stessa, oltre che ai libri recenti dell'AICA, ricordati dal segnalibro di Didamatica 2013, ed ora in parte fruibili come ebook, come per quello titolato **Capitali in Italia**. Il prolungamento della mostra, da luglio 2013 a febbraio 2014 può essere un invito a visitarla per una presa di coscienza, magari in compagnia di ben più giovani amici. I corsi su modellazione e interoperabilità presso il Politecnico, schedati per il corrente autunno, faranno certamente riferimento all'opera di Leonardo ed alla rassegna presentata, che può presentare un tema di interesse per il mondo della scuola. La scienza dell'informazione può farci tornare a quell'artigianato, che fu arte ai tempi di Leonardo, per ritrovare basi per scoprire attraverso nuove iniziative il significato della bottega rinascimentale.

Promoting trust protecting citizens' rights in digital citizenship

Angela Vozella
CIRA
Via Maiorise, 81043, Capua (CE)
a.vozella@cira.it

Abstract. *This paper is inspired by the theme of citizen's rights (e.g. "Ensure privacy and freedom in the Internet...", a topic foreseen in Horizon 2020) [1] and their protection with regard to the pervasiveness of ICT. Not focused on a specific technology, it identifies issues to increase citizens' trust in new technologies based on software-intensive safety-critical applications.*

Keywords: safety critical systems, risk, liability.

1. Introduction

ICT widely spreads over the every daily life, from the management of Critical Infrastructures (C.I.) to home automation. Risks of such pervasiveness lie in vulnerability to hazards for systems and citizens; in fact, malicious attacks and system failures may result in catastrophic effects also considering the interdependence of C.I. [2]. This situation has lead society to be cautious in the adoption of new technology especially for safety critical applications. Clearly identified and perceived safety critical applications, as by instance the on board software for an aircraft, must comply to strict rules, imposed in this case by the Agency for Aviation Safety (EASA, ENAC,...). Producers of this type of safety critical applications have to guarantee to follow a predetermined process for the product life-cycle according to relevant safety categories [3].

Books and cinematography has widely diffused the idea that programmers could maliciously hide dangerous behaviours in the software code like being capable of autonomous malicious actions.

This paper provides a framework to identify possible concerns of the society raised by the use of new technology (commercial, corporate, governmental) in software intensive system and issues to overcome them specifically:

- in case of accident: liability (incl. issues like enforcement, impact of automation) and insurance;
- the protection against abusive use: privacy, data protection, security;
- public acceptance of new technology and ethics.

2. Liability

In the areas where it has been identified that existing regulations cannot support new technology adoption, a regulatory framework needs to be developed to determine which technologies or procedures are essential to reach the objective of a safe introduction. Although the matter of regulating the use of new technology is clearly relevant today and is being examined by many international authorities, the present legal framework is often inadequate.

By instance, considering the complexity of new systems, it is vitally important to make a clear distinction among people operating and interacting with the system at different stages in the system life cycle [4]. The liability for damages caused by a system failure should be attributed to the operator, that is, the person or entity that, ensures its functioning and makes known his or its status as operator. Determining the operator's liability is made without regard to personal responsibility (negligence or wilful misconduct). Therefore, it is strict liability based on the risk of a lawful activity. Under strict liability, no negligence of the operator needs to be proven, whereas with fault-based liability, an operator will only be found liable if some form of negligence is established.

Automated systems add further layers of complexity with respect to traditional software/hardware artefacts, since they may possess autonomous cognitive states and behaviours that are relevant from a legal perspective. When in complex organisations the automation has taken over more or less completely, humans become controllers of automated systems, rather than operators: in fact, automated systems directly operate to fulfil the task, exercising cognitive functions, acquiring information from the environment, processing it, and using the knowledge to achieve the goals assigned to them, as specified by their users, while humans monitor the work of automated systems.

In scenarios when one or several operators/controller and one or several automated support systems interact together for the fulfilment of a task, it would be better to describe humans and technology not as two interacting "components", but as constituting a joint (cognitive) system. Under this perspective, a relevant (and still open) question is that of how to deal with cases in which (as in some recent aviation accidents) conflicting information is provided to operators by humans (controllers) and automated systems, and more generally what kind of priorities should be given to different signals, and when humans may override automatic devices.

A shift from personal liability toward general enterprise liability (liability for creating a risk through the use of the technology) and product liability can be seen. Furthermore, speaking about the developer, the more automated the system becomes, the more organisations and individuals are involved in building, testing, and developing it: this is why the failures caused by highly automated tools will be managed distributing liability among the developers involved in building it. This would mean that experts will be called to establish what went wrong in the tool, and responsibilities for the developer of hardware, of software, the maintenance service provider, the integrator of the tool, etc. As liability may emerge from the introduction of a new inadequate technology,

liability may also emerge also from not adequately deploying an appropriate technology. The risk is that the current liability regime could result in unbearable costs for enterprises, so that to hamper the adoption of automated technologies, and in particular safety-enhancing technologies:

- concerning product liability, the issue is whether enterprises may always rely on the “state of the art” defence, so that they should not be held liable when automated technologies were developed according to available standards and when it was impossible to foresee malfunctioning of the technologies at the current state of the scientific and technological knowledge available in the field;
- concerning strict liability, we should consider whether current liability caps are appropriate also for higher level of automation, or complementary measures should be adopted to ensure that damages will be paid without creating a too high burden for enterprises (e.g. compulsory insurances for all stakeholders involved in the design/ development/ use of automated technologies, and the introduction of compensation funds).

A second question concerns how to properly analyse and manage the shift of liability due to automation, in order to achieve an optimal allocation of burdens. This will imply reconsidering the role of liability, also as a means to prevent accidents fostering the development of a safety culture within organizations.

In conclusion, in order to ensure a safe and responsible adoption of new technologies, an appropriate assessment of regulation and allocation of liability is crucial involving the following items:

- 1) To perform a rigorous risk assessment in order to identify the responsibilities of the different actors (operators, end users, controller, managers, manufacturers, designers). This will reduce level of uncertainty within insurance sector or any other risk-sensitive stakeholders by providing a first solid approach on assessing risk around new technology;
- 2) To analyse how existing laws and regulations (national and international, including standards and certification procedures) regulate the allocation of liabilities for the development and use of new systems, providing an adequate normative framework (Fig 1).

Outcomes must be shared among various stakeholders in order to contribute to achieve increased acceptance within the society of member states.

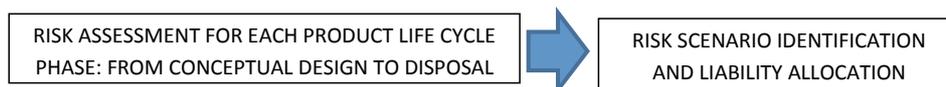


Fig 1 – Liability allocation for each risk scenario

3. Privacy and data protection

The introduction of new technology could raise many legal questions that include, inter alia, the issues of right to privacy and data protection.

In Europe, respect for private life was established in 1950 with the adoption of the European Convention of Human Rights (ECHR) in the framework of the Council of Europe, it may be described as a right which prevents public authorities from measures which are privacy invasive, unless certain conditions have been met. The right to data protection was introduced in the 1980s, as a consequence of technological developments. Data protection legislation aims at setting down conditions under which it is legitimate and lawful to process personal data. Data protection legislation obliges those entrusted with handling data to respect a set of rules and empowers the people concerned by granting them rights and guarantees. Finally, it provides for supervision by independent authorities. Protection of personal data is a right which is separate but closely linked to the right to privacy. Despite a consistent regulatory framework, privacy appears to be difficult to protect perhaps because there is no universal definition for privacy and its features vary according to the context, culture and legal tradition. The word 'privacy' stems from a Latin root, "privare" which means 'to separate' or to deprive, i.e. to separate from others what belongs to an individual only. Legal doctrine has identified four main facets of privacy:

1. Bodily privacy that relates to the protection of a person against invasive procedures such as genetic tests, drug tests and body searches,
2. Information privacy, also known as data protection, with reference to the rules governing the collection and handling of personal data such as medical records, Government records or financial records,
3. Privacy of communication that gives security and privacy of mail, telephones, electronic mail and other forms of communication,
4. Location privacy which sets limits on intrusion into domestic and other places such as at work or in public spaces; it includes searches, video surveillance and identity checks as well as one's geographical location and position in space.

This categorization suggests a continuum of zones of privacy and of personal, public and private space that can be imagined horizontal or vertical. It is difficult to draw a demarcation line between personal, private and public considering political, economic and even religious or cultural interests.

There are many opponents of an excessive defence of privacy claiming that privacy may prevent law enforcements or impede economic activities based on a free flow of information to be collected, stored and then sold, hindering law enforcement, national security operations and freedom of information.

So privacy must be balanced against other rights like freedom of information, national security and law enforcement. Several European provisions limit the right to privacy. The Convention for the protection of individuals with regard to automatic processing of personal data of 28 January 1981, "Recognizing that it is necessary to reconcile the fundamental values of the respect for privacy and the free flow of information between people", in Article 9 allows derogations for: a) protecting State security, public safety, the monetary interests of the State or the suppression of criminal offenses; b) protecting the data subject or the rights and freedoms of others. While Data Protection Regulation 45/2001/EC

containing provisions to protect personal data processed by European Union (EU) institutions and bodies states that private data must be: processed fairly and lawfully, collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes, accurate and, where necessary, kept up to date and in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data are collected. The Regulation also provides for the establishment of a “European Data Protection Authority”, an independent Community authority responsible for monitoring the correct application of the data protection rules by the EU institutions and bodies. Convention No. 108 “for the Protection of Individuals with regard to Automatic Processing of Personal data” aimed to strengthen data protection, i.e. the legal protection of individuals with regard to automatic processing of personal information relating to them. Compared with manual files, automated files have a superior storage capability and offer possibilities for a much wider variety of transactions, performed at high speed, while a further growth of automatic data processing in the administrative field was expected in the coming years as a result, inter alia, of the lowering of data processing costs, the availability of “intelligent” data processing devices and the establishment of new telecommunication facilities for data transmission.

The Council further stressed that “Information power” brings with it a corresponding social responsibility of the data users in the private and public sector. The proposals to review Directive 95/46/EC tabled by the Commission in January 2012, ask for tighter harmonization rules to prevent fragmentation in the way personal data protection is implemented across the Union and addressing the threats to privacy posed by recent technological developments. The General Regulation addresses several issues such as the need for special protection of health-related data, the need for clearer rules as regards data access and portability. [5] Finding a balance between the need to guarantee the right of individuals to “opt out” or delete their personal data (the so-called “right to be forgotten”) and the need to provide access to the same data to public officials is especially difficult. An even bigger concern is raised by the possible use which could be made of sensitive data by private subjects for commercial purposes.

The concept of “Privacy by Design” laid down in the first European directive on data protection, namely, in Art 17 of D-95/46/EC asks to embed “appropriate measures” in ICTs “both at the time of the design of the processing system and at the time of the processing itself, particularly in order to maintain security and thereby to prevent any unauthorized processing” of personal data (Fig. 2). This objective has been facilitated by the capacities of computers to draw upon the tools of artificial intelligence (AI), though it is not easy to balance performance requirements with legal and data protection issues.

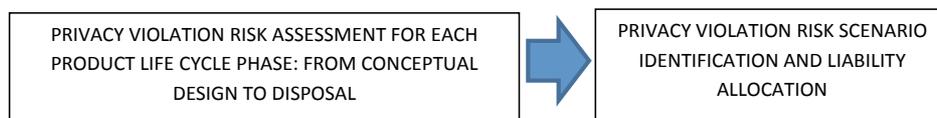


Fig 2 – Liability allocation for each risk scenario of privacy violation

4. New Technology Acceptable Risk

While there are detailed methods for the objective measure of the likelihood of a hazardous event based on a quantitative measure of historical safety performance, there are substantial differences in what is measured and what is perceived. Risk perception, and not the objective measure of risk, will be the driver behind the acceptance of any new technology whose applications can be considered safety critical. The perception of risk is driven by the magnitude of consequence more so than the associated likelihood of occurrence. Public perception of risk focuses on those hazards that have the potential to cause large consequences. People's risk perceptions are based on a combination of subjective judgment and limited knowledge of the true risks imposed by a new technology. According to a recent study there is a tendency by the public to overestimate small risks and to underestimate large risks, and that the public tends to focus on risk and how they can protect themselves from those risks. Conversely, experts tend to perceive risks within their competence area as much lower than the public. As a result, public trust seldom conforms to expert assessments of hazards associated with technologies, particularly when the technology is new to the public. That is why it is necessary to create credibility for the industry and not merely impose new products. In most cases, society has opposed any new technology that has associated risks. Distinctions must be made between those applications where the principal risk exposure is voluntary from those of involuntary risk exposure. This is because the public places a higher demand for protection from involuntary risks as opposed to voluntary. It is worth noting that the question relates to the public's acceptance of the risks associated with a new technology and not the public's acceptance of a new technology.

Other complex and often immeasurable factors such as morals and the economic and political climate are equally as important. The basic theory behind the acceptance of risk is the subjective assessment between:

- Society's perception of the level of exposure to the hazard;
- Society's perception of the benefits due to the hazardous activity.

On a general basis the media and public must be convinced that the perceived benefits (i.e., higher level for security, improved information, more services, lower costs) outweigh potential "costs" (i.e., increased noise, pollution, privacy concerns, safety risks, delays). The impact of the new technology on the environment can also influence society acceptance. New technology may be limited by noise, emissions, or other environmental constraints. Use of solar power, fuel cells, and other low emissions systems, could encourage their use among people caring for the Environment. The perceived benefit from a hazardous activity directly influences an individual's willingness to accept risk. It has been shown that the level of benefit awareness is directly proportional to the acceptable level of risk. It is likely that the public will make a distinction between those operations which provide a "greater good" (for example fire fighting or search and rescue) and those operations, which have only limited community benefits.

The benefits the new technology could ensure to the European Citizens represent one important cornerstone for their acceptance by the European people. Among these, citizens can see a number of them as having more direct benefits on their lives and we can classify them in three groups:

- Missions related to Civil Protection and/or social aid: technology used in monitoring, preventing and alert system for natural disasters as well as to inform/train people.
- Missions related to Security: for coastal surveillance or sensitive sites (ports, airports, power plants) monitoring.
- Mission related to Environment Protection / Preservation: monitoring and protecting natural environment.

Examining in more detail the above possible utilisations emerges a variety of specific tasks that are already perceived as important by EU Citizens, despite the possible different priorities among the Member States. Just recalling the results recently emerged from a number of statistical analysis and studies, it is possible to depict an almost exhaustive picture of what is actually recognised as “important benefits” by Citizens in the civil Protection, Security and Environment fields. A number of coordinated action of leveraging on these arguments could help in building the public awareness about new technology; as a consequence, will be facilitated the growth in familiarity for their technologies, also considering that some of them are already available in different field of application like automotive and mobile phones. Civil protection is involving thousands of workers in Europe and is typically perceived by citizens as an Institution able to intervene when a disaster occurs and sometime able to make prevention, alerting the people to abandon their places and/or organising active countermeasures to avoid or at least limit the effects of certain disasters.

European population has a personal perceiving about the seriousness of disasters depending on their specific culture and the particular geographic area they live in. Despite an ad-hoc analysis is always possible, it is recommended, to create a general common perception about disasters and consequently of the benefits that new technologies could give.

Gaining public trust in new technology needs time and specific actions. But any obtained trust could be easily damaged or lost in a high exposure accident.

To reduce the possibility of an adverse public reaction to new technology, a strategy for communicating with the public is needed based on the following actions:

- Make people perceive new technology as a natural part of future society-increasing familiarity and awareness about benefits by quickly and accurately reporting good and bad news concerning new technology perception versus objective evidence (e.g. creating a website where the public can get information and ask questions increasing familiarity)
- Select groups to be responsible for researching, analysing and matching the sentiment of the public
- Create a strategy to be used in case an accident occurs.

In line with these recommendations, the benefits of the new technology must be better explained to the public possibly in terms of: cost saving issues ,noise reduction, safety improvement. Furthermore the policy making process supporting the development of new technology applications needs to be transparent and involve the consultation of stakeholders, for example bodies like the European Group on Ethics, the LIBE Committee of the European Parliament or the European Union Agency for Fundamental Rights and European Data Protection Supervisor.

Last but not least, a certain range of permissible or forbidden uses of new technology with guidelines based on a 'privacy and data protection impact assessment' could be defined to increase the confidence of citizens.

Taking demographic factors into account, by this survey it will be possible to draw generalizations of different public groups comparing perceived risks associated with various uses with their perceived benefits.

The demographic groups could include; age, gender, level and type of education, voting status and other features. This assessment will include risks to different classes of people, and benefits to the users as well as society in general. This will also allow drawing correlations between certain types of uses in order to determine which applications are more easily accepted. It is possible that the public will have a significant level of discomfort with some applications. This discomfort will likely be more evident in older respondents, who are generally seen as more conservative.

As far as safety concerns are associated with different applications, the closer a person is exposed to the system where new technology might be in use, the more risk they will perceive. Correlations will be drawn between applications that receive similar risk perception responses, and a trial to analyse how or why these relationships exist. .

Summarizing the approach towards public acceptance could involve the following steps:

1. Definition of a "public acceptance facilitator group";
2. Implementation of a survey;
3. Evaluation of the survey results to plan facilitation strategy;
4. Implementation of the facilitation strategy;
5. Measurement of results and improvements.

5. Ethics

Ethics is commonly considered as a common sense definition of what is good for the humans and what is correct to do or not to do.. In a more formal way, Ethics is concerned with how we should conduct relations with others. When approved and codified by a State they become laws. Ethics could have different priorities depending on the traditional and cultural background of different Countries and peoples. Generally speaking, Citizens will expect new technology to have an ethical behaviour comparable with the human one, respecting some

commonly accepted rules. Then, they will expect to identify a legal or physical entity to blame and condemn in case ethical rules are broken.

If instead, they will be able to prove on a constant basis their capability in assisting human activities recognised as helpful like Civil Protection, Security and Environment protection, the Citizens will probably start to consider them as part of their life, like the smart phones. But sharing risk for benefits could not be enough as far as Ethics is concerned, because societal values are anyway at a higher moral level than benefits in most of the citizens

There is no doubt that some incorrect behaviour will have to be prosecuted. In this case defining the boundaries among “actors” responsibility for erroneous decisions taken during the design and operator responsibilities for the way they operate the system is crucial (Fig. 3)

As a resume ethical issues will include the definition and verification of: ethical principles for new technology, responsibilities for operators manufacturers and users considering the ethical principles and possible code of conduct, rules for operator selection and training, benefits and their compliance with ethical principles.[6]

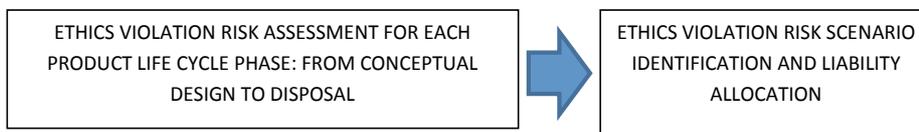


Fig 3 – Liability allocation for each risk scenario of ethics violation.

6. Conclusions

Usually for safety related technology a technical risk assessment is performed before the authority allows to use that technology. The risk assessment aims at demonstrating that risks are acceptable in terms of severity of consequences and occurrences. But a clear vision about Complementary Measures in terms of liability, protection of data, public acceptance is not performed. A scenario for each previous area of Complementary Measures is depicted, indicating the major elements that should be taken into account for a smooth integration of a new technology which is safety critical. Strict coordination among these measures has to be kept within a complementary measure impact assessment. Such analysis will drive the complementary measure implementation plan.

References

- [1] Report of the workshops on the Common Strategic Framework (CSF) for Research and Innovation: Inclusive, Innovative and Secure Societies Challenge
- [2] European Reference Network for Critical Infrastructure Protection Europe 2020

- [3] DO-178B, Software Considerations in Airborne Systems and Equipment Certification .
- [4] Sidney W. A. Dekker ,”Just culture: who gets to draw the line? “14 January 2008, Springer-Verlag London Limited.
- [5] EU Data Protection Directives <http://ec.europa.eu/justice/data-protection>.
- [6] European Group on Ethics in Science and New Technologies BEPA Bureau of European Policy Advisers.

ACCESSO ALLA RETE: DIRITTO SOCIALE IMPLICITO NELLA COSTITUZIONE PER LA FRUIZIONE DI UNA CITTADINANZA ATTIVA

Valentina Amenta

Istituto di Informatica e Telematica, Consiglio nazionale Ricerche di Pisa
Via Moruzzi, 1 56124 Pisa (PI)

Valentina.amenta@iit.cnr.it

Abstract. *In the article raises the configuration of the right of access to the Internet as a new law, which, although does not have an explicit regulatory approval, has a strong tone constitutional. The objective will be just what the legal status of the right of access to the Internet, which is proposed in its intrinsic structure and biphasic. From one side it is configured as a fundamental right and thus comprise all activities relating to the right to be informed and to inform. On the other hand, the Internet presents itself as social right instrumental to the exercise of other rights. This right can be claimed by members of a given community against their political authorities. But in national experiences such as the Italian, the formal equality, it seems inadequate to implement the principle of equality for the presence of subjective situations and unequal starting positions that benefit certain subsidiaries against others.*

Keywords: Access, broadband, Network.

1 INTRODUZIONE: Il diritto di accesso ad Internet

In un contesto in cui l'informazione, l'informatizzazione e la globalizzazione dei processi, delle regole e delle reti appaiono essere divenute chiavi insostituibili di accesso alla libertà, al progresso e alla democrazia, la sensazione è quella che stia nascendo e si stia sviluppando all'interno della nostra società una nuova categoria di diritti [Bovero, 2004].

Il nuovo diritto scaturito dall'evoluzione tecnologica è "il diritto di libertà informatica, che manifesta un nuovo aspetto dell'antica idea della libertà personale e costituisce l'avanzamento di una nuova frontiera della libertà umana verso la società futura e che si viene a collocare nel prisma del costituzionalismo contemporaneo [Frosini, 2008]".

Nella primigenia versione, risalente al 1981, la libertà informatica si configurava con una duplice accezione positiva e negativa.

La libertà informatica negativa esprime “il diritto di non rendere di dominio pubblico certe informazioni di carattere personale, privato, riservato”; la libertà informatica positiva, invece, esprime “la facoltà di esercitare un diritto di controllo sui dati concernenti la propria persona che sono fuoriusciti dalla cerchia della *privacy* per essere divenuti elementi di *input* di un programma elettronico; e dunque libertà informatica positiva, o diritto soggettivo riconosciuto, di conoscere, di correggere, di togliere o di aggiungere dati in una scheda personale elettronica”.

Una attenta dottrina afferma che “il diritto di libertà informatica assume una nuova forma del tradizionale diritto di libertà personale, come diritto di controllare le informazioni sulla propria persona, come diritto dello *habeas data* [Frosini, 1981]”. Il fine dello *habeas data costituzionale* è quello di garantire proprio “la libertà informatica, come garanzia personale a conoscere e accedere alle informazioni personali esistenti nelle banche dati, a controllare il loro contenuto e quindi a poterle modificare in caso di inesattezza o indebita archiviazione o trattamento, nonché a decidere sulla loro circolazione [[Rozo Acuna, 2001] ”. Con l'avvento della Rete, il diritto di libertà informatica “è diventato una pretesa di libertà in senso attivo, non libertà *da* ma libertà *di*, che è quella di valersi degli strumenti informatici di ogni genere. È il diritto di partecipazione alla società virtuale, che è stata generata dall'avvento degli elaboratori elettronici nella società tecnologica: è una società dai componenti mobili e dalle relazioni dinamiche, in cui ogni individuo partecipante è sovrano delle sue decisioni [Frosini, 2002]”. Ci troviamo dinanzi ad una nuova forma di libertà, ossia quella di comunicare con chi si vuole, divulgando le proprie idee, i propri pensieri e i propri materiali, e la libertà di ricevere. Si delinea, così, una libertà di comunicare come libertà di trasmettere e di ricevere.

In questo contesto, imperniato di una concezione tecnologica della libertà di comunicazione, fanno fatica e tardano ad emergere i contenuti delle tradizionali libertà costituzionali, in particolare quella di comunicazione e quella di manifestazione del pensiero.

In questo ambito, la materia dei diritti fondamentali si presenta come un punto di incontro tra temi di grande importanza, tra cui la definizione di diritto soggettivo, il concetto di costituzione e il senso della democrazia. La democrazia del XXI secolo si presenta in una forma diversa da quella che era nei secoli precedenti: “mutano i significati di rappresentanza e di sovranità, avanza una nuova democrazia di massa, che rompe le cerchie chiuse delle *élites* al potere, obbligando per così dire i rappresentanti della volontà popolare a scendere sulla piazza telematica e a confrontarsi direttamente con i rappresentati, nelle nuove forme assunte dalla tecnico politica [Rodotà, 1998]”. L'obiettivo della democrazia attuale, più di ogni altra cosa, nel campo dei diritti fondamentali resta quello di avviare una precisa commistione tra l'esperienza giuridica esterna, propria del cittadino cosiddetto normale e l'esperienza giuridica interna, propria degli operatori classici del diritto. Lo scopo si presenta

Accesso alla rete:Diritto sociale implicito nella Costituzione per la fruizione di una cittadinanza attiva

molto ambizioso, ossia restituire al cittadino la consapevolezza del suo ruolo primario anche in questa fase di cambiamento del diritto.

Quando si utilizza l'espressione "diritti fondamentali" si fa riferimento generalmente alla dicitura "diritti umani". Però, se da un lato, i diritti umani ambiscono all'universalizzazione, e sono protetti da regolazioni internazionali, dall'altro i diritti fondamentali non sono semplicemente diritti soggettivi, ma diritti soggettivi che svolgono un compito "funzionale" specifico di uno stato di diritto costituzionale.

All'interno della tematica in esame, questo lavoro intende approfondire proprio il problema della tutela dei nuovi diritti digitali, appartenenti ai cosiddetti diritti di quarta generazione, con particolare riguardo ad un diritto, il c.d. *diritto di accesso alla Rete*, che si configura attualmente come il diritto più innovativo, ma che sta tardando ad emergere come diritto fondamentale garantito.

Il contenuto del diritto di accesso telematico si prefigura non solo come più ampio rispetto alle forme tradizionali della libertà di comunicare, ma, anzi, sembra avere una funzione di supporto rispetto ad un ampio catalogo di diritti tradizionali. Ed è proprio il nuovo assetto delle telecomunicazioni che ha rafforzato l'idea e la necessità di ragionare sul diritto ad accedere alla Rete, da concepirsi "come una sorta di diritto-madre, rispetto a quello eventualmente da riconoscersi in ordine alla connessione ad *Internet* [Costanzo, 2000]".

È giusto credere, come un'autorevole dottrina [Zeno Zencovich, 2004] ha specificato, che non si può sottovalutare il ruolo sociale dell'accesso telematico perché ciò equivale a non garantire a chiunque quelle condizioni necessarie, al tempo d'oggi, per stabilire le relazioni con gli altri individui e per assicurare il pieno svolgimento della propria personalità.

Fra i coni d'ombra che i giuristi segnalano nell'emersione del diritto di accesso ad Internet come diritto, si delinea la necessità di qualificazione di tale diritto, ossia se esso possa essere collocato all'interno dei diritti cardine enunciati dalla nostra Costituzione, specificando se trattasi di diritto fondamentale o di diritto sociale, oppure se esso debba ricevere un'autonoma regolamentazione giuridica.

1.1 Diritto di accesso alla Rete: diritto fondamentale o diritto sociale implicito nella Costituzione?

L'interprete che voglia seguire la prima strada non può non muovere dalla considerazione degli artt. 15 e 21 del nostro testo costituzionale.

Il fondamento di questi articoli risiede nella consapevolezza che a prescindere dal sesso, dalla razza, dalla religione professata, tutti hanno diritto non solo alla libertà di manifestazione del proprio pensiero, ma anche alla libertà d'informarsi. Questi elementi sono imprescindibili per un concreto sviluppo della personalità di ciascuno e, quindi, per un'autentica uguaglianza fra i cittadini e per la possibilità per ciascuno di partecipare alla vita sociale. L'art. 15 Cost. riconosce il diritto del singolo e delle formazioni sociali di comunicare il proprio pensiero ad uno o più soggetti determinati e nel contempo ne garantisce la segretezza.

La tutela costituzionale ex art. 15 "si perde nella più ampia tutela della comunicazione in generale, quasi a comporre un ideale trittico con gli artt. 13 e

14 Cost. espressione della triplice dimensione (fisica, spaziale e spirituale) dell'inviolabilità della persona umana [Di Lello, 2007]”.

L'ambito oggettivo di tutela dell'art. 15 è proprio quello della comunicazione privata, che consiste sia nella corrispondenza in senso stretto sia in qualsiasi altra forma di comunicazione interpersonale.

La libertà d'informazione è intesa come libertà non tanto di cercare e conseguire determinate notizie, quanto, essendo volta all'apprensione di conoscenza in generale, di utilizzazione delle fonti d'informazione, indipendentemente dalle notizie in esse contenute [Loiodice, 1971].

Affinchè si possa parlare di comunicazione, ai sensi della nostra Carta Costituzionale, rilevano l'*animus* del soggetto di mantenere segreta la comunicazione e la *determinabilità* dei soggetti cui questa è indirizzata.

Bisogna, però, capire se il modello comunicativo creato da Internet possa rendere difficile la distinzione tra comunicazione privata (ambito dell'art. 15 Cost.) e manifestazione pubblica del pensiero, per la quale è prevista la tutela dell'art. 21 della Costituzione. La maggior parte della dottrina ritiene che il mezzo *Internet* “potrà beneficiare, a seconda dei casi, delle garanzie dei mezzi di manifestazione del pensiero *sic e simpliciter* o delle ulteriori garanzie accordate ai mezzi di comunicazione interpersonale. Verrà, quindi, a ricadere nelle garanzie dell'art. 15 il pensiero diretto a determinati destinatari attraverso *software* che ne garantiscano la segretezza, al contrario non ci si troverebbe di fronte a comunicazione riservata nel caso in cui l'utente utilizzi forme che mirino alla manifestazione pubblica del proprio pensiero. Ed infatti, quando, nella comunicazione diffusa in un *forum*, in una *Chat* o in un *Blog* si potranno ben applicare le garanzie tipiche della libertà di manifestazione del pensiero o il regime della libertà di riunione [Pubusa, 2006]”.

Qualora invece si volesse preferire un'interpretazione meramente letterale dell'art.21, la libertà di manifestazione del pensiero, la sua esistenza ed il suo esercizio “sarebbero rimesse a colui che esprime il proprio pensiero e, in particolare, al suo consenso, all'acquisizione di informazioni relative alla formazione e manifestazione del proprio pensiero. L'ambito d'azione della libertà d'informazione sarebbe dunque determinato da un soggetto diverso dal suo titolare, e la tutela sarebbe meramente occasionale, intervenendo solo qualora venga violata la libertà d'espressione”.

Chiedersi se sia possibile fare appello a un diritto che appare solo proclamato, ma non anche “protetto”, è chiedersi se un diritto “esiste”, anche nel caso che l'ordinamento non abbia predisposto per esso una precisa formazione.

Le disposizioni che codificano diritti fondamentali sono solitamente formulate in termini estremamente generici ed indeterminati. Di fronte a cataloghi di diritti così formulati, sorge il dubbio che è forse inevitabile che i diritti fondamentali siano proclamati in disposizioni formulate in maniera generica e indeterminata [Ferrajoli, 2010]. Così, l'individuazione di un nuovo diritto, come il diritto di accesso, può essere conseguita tramite una semplice lettura non restrittiva delle tipologie di base dei diritti fondamentali. Si potrebbe ammettere che i diritti fondamentali configurino un *genus* dotato di una capacità espansiva intrinseca,

che si duplica al suo interno nel rispetto delle direttrici implicite alla categoria. Infatti, i diritti fondamentali esistono prima dello Stato e per ciò non dipendono dalla legge, anzi costituiscono un limite alla libera produzione normativa. In ordinamenti giuridici dotati di costituzione lunga e pluralista, le disposizioni che codificano diritti fondamentali sono solitamente formulate in termini estremamente generici ed indeterminati. Una disciplina costituzionale che prevedesse diritti fondamentali in termini rigorosamente circostanziati risulterebbe irragionevole. Così, il diritto fondamentale implicito potrebbe essere considerato il presupposto di alcuni diritti fondamentali espliciti: questi ultimi non si spiegherebbero se non si affermasse il primo. Nell'ordinamento costituzionale italiano, peraltro, la creazione di diritti fondamentali impliciti è ampiamente agevolata dalla presenza dell'art.2 della Costituzione, solitamente interpretato come meta-norma che autorizza l'individuazione di tutti i diritti che si considerano essenziali allo sviluppo della persona umana [Pino, 2010].

Da questa impostazione potrebbe discendere un'interpretazione ampia e sistematica non solo dell'art. 21, ma di tutti i precetti costituzionali relativi alla problematica dell'informazione, in modo da rendere la relativa libertà autonoma e distinta dalla libertà di manifestazione del pensiero: ciò, però, non significa negare le naturali connessioni che uniscono l'espressione del pensiero all'acquisizione di conoscenza, ma individuare un'impostazione che, distinguendo i regimi giuridici, le valorizzi entrambe, cosicché esse possano rafforzarsi reciprocamente.

Alla luce di questa interpretazione, l'art.21 Costituzione potrebbe essere letto come se fosse scritto nel modo seguente: tutti hanno diritto di usare ogni fonte di informazione (ed *Internet* si presenta come un grande mezzo di diffusione delle informazioni) disponibile per diffondere il proprio pensiero [Loiodice, 1999]. Quanto finora detto, quindi, non si oppone all'estensione del regime costituzionale già esistente al mezzo *Internet*, posto che l'inciso di chiusura dell'art. 21, ossia <<ogni altro mezzo di diffusione>>, ha permesso alla norma costituzionale di sopravvivere all'avvento dei nuovi mezzi di comunicazione di massa. La libertà dell'art. 21 Cost. è garantita, secondo il tenore letterale dello stesso articolo, a <<chiunque>>, e ne deriva che l'accesso al mezzo deve essere riconosciuto a tutti, a prescindere dalla condizione sociale, dal luogo geografico di appartenenza. Potrebbe quindi non sussistere nessun ostacolo all'applicazione della norma costituzionale al mezzo telematico. Peraltro, se si considera che l'estensione della libertà di espressione è condizionata non solo dal contenuto del messaggio, ma anche dal mezzo con la quale è esercitata, forse il mezzo *Internet* realizza meglio di altri lo spirito della norma costituzionale [Di Lello, 2007].

La prospettiva delineata fin qui si riferisce alla manifestazione del pensiero.

Diversa è invece la prospettiva del <<navigatore>> del Web, di colui cioè che utilizza *Internet* per ricercare notizie. In questo caso l'attività svolta dall'utente non rientra in un'attività di manifestazione del pensiero in senso stretto, e tuttavia può farsi ricadere ugualmente sotto il regime dell'art.21 in virtù del corollario del principio del diritto di informazione.

Infatti, dalla libertà di informazione discendono, come corollari, il diritto ad informarsi, inteso come diritto a ricercare informazioni, ossia il diritto di accesso alle informazioni attraverso ogni mezzo telematico e non che le contiene, e il diritto ad essere informati, ossia il diritto alla disponibilità delle fonti informative. Questi due corollari, data la loro essenza, non possono esistere se non correlati nel senso che il primo non può esistere senza il secondo, e quest'ultimo non avrebbe ragion d'esistere senza l'altro.

Quindi, in riferimento al mezzo telematico, quando si parla di diritto ad informarsi non si intende fare riferimento ad un diritto attribuibile al soggetto a ricevere informazioni di un determinato contenuto, ma all'assenza di limiti all'accesso alla Rete per cercare notizie già presenti e disponibili sul *Web*. Non quindi, una pretesa, ma una semplice libertà del soggetto che trova come unica garanzia il non vedersi impedito l'accesso.

Sotto questo punto di vista il diritto a navigare in Internet sarebbe parimenti garantito come diritto a ricercare informazioni, in quanto "corollario necessario e imprescindibile della libertà d'informazione, ma la garanzia non coprirebbe la ricerca di informazioni che non siano disponibili".

La libertà informatica non si esaurisce però solo nella dimensione della comunicazione e dell'informazione. Essa abbraccia anche la libertà politica e l'organizzazione istituzionale, che ingloba nel suo alveo anche la fruizione dei servizi da parte degli utenti/cittadini. È manifesto, infatti, come il progresso tecnologico sia destinato a mutare sempre più gli apparati istituzionali sperimentati e come il processo democratico sia fortemente condizionato dal modo in cui circolano le informazioni, laddove cioè la possibilità di uso di queste da parte di tutti i cittadini si caratterizza per essere un prerequisito di quel processo. Così, la libertà informatica da slancio al principio democratico e a quello di imparzialità, che trovano la loro origine costituzionale innanzitutto nell'art. 3 della Costituzione, che dev'essere integrato dagli artt. 1 e 2 del medesimo testo costituzionale. Analizzando e seguendo il canale circolare che parte dall'art.1 e perviene all'art.3, si desumono i tre concetti chiave che indirizzano l'intero discorso: inclusione, partecipazione, politiche pubbliche. Queste locuzioni riflettono l'imperativo -principio vincolante- insito nell'art.3, 2° comma, che impone un collegamento costante fra fini e organizzazioni, tra valori fondanti costituzionalmente, e da cui consegue una struttura di poteri pubblici che a quei fini devono dare attuazione. L'art. 1, 2° comma stabilisce che la sovranità appartiene al popolo che la esercita nelle forme stabilite dalla Costituzione in forme plurime, per giungere all'art. 3, 2° comma che sancisce la rimozione degli ostacoli per consentire il pieno sviluppo della persona umana ai fini di una effettiva partecipazione alla vita politica, economica e sociale del Paese. È, quindi, un cerchio che si chiude: l'esercizio della sovranità, e in particolare dei diritti individuali, è in qualche modo finalizzato ad una relazionalità della persona che consenta una partecipazione piena, ossia una cittadinanza attiva e responsabile [Valastro, 2011].

Il primo riscontro normativo e la prima rappresentazione dell'individuo, visto in questa nuova veste di cittadino digitale, si rinviene nel Codice

dell'Amministrazione digitale (CAD), approvato con il d.lgs.82/2005, Questa normativa individua e disegna una *species* di cittadino digitale, che si adatta sia alle persone fisiche che alle persone giuridiche e che si erige su un doppio filo di diritti e obblighi, ossia il diritto di esigere dai pubblici uffici l'interazione in modalità digitale e l'obbligo per l'amministrazione di dotarsi di strutture adeguate per soddisfare la richiesta dell'utente. Tra le varie norme si sottolineano in tale sede quella sul diritto al procedimento amministrativo informatico, il diritto di effettuare pagamenti elettronici con le amministrazioni centrali, il diritto alla qualità dei servizi in termini di informazione e comunicazione tecnologiche, il diritto di comunicazione tramite posta elettronica. Questo nuovo rapporto nascente tra cittadino e amministrazione dev'essere inquadrato in un più ampio diritto sociale di cui lo Stato italiano si deve far portavoce per assicurare a tutti gli utenti la fruizione, presupponendo il diritto all'istruzione e all'accrescimento della cultura informatica. Questi ultimi diritti sono considerati dalla Corte costituzionale "corrispondenti a finalità di interesse generale, quale è lo sviluppo della cultura, nella specie attraverso l'uso dello strumento informatico, il cui perseguimento fa capo alla Repubblica in tutte le sue articolazioni (art.9 Cost.)".

Il diritto di accesso, in questo modo considerato, si qualifica dunque come un diritto sociale strumentale [Pietrangelo, 2007], che consente l'utilizzo di altri diritti sociali tradizionali. Questa nuova forma di partecipazione, più che mostrarsi come una fase nuova nello sviluppo dei diritti, si configura come una modalità alternativa di accesso agli stessi. L'innovazione tecnologica ha, infatti, aperto la strada ad un nuovo modo di intendere la cittadinanza. In sostanza, l'esercizio di molti diritti fondamentali è stato svincolato dal rapporto diretto tra il cittadino e gli apparati istituzionali, divenendo mediato da strumenti informatici che, per la loro stessa natura, si collocano al di fuori degli spazi fisici entro cui il rapporto tra il primo e i secondi ha comunemente luogo. La cittadinanza elettronica, in virtù della sua natura immateriale, sembrerebbe offrire inedite opportunità di esercitare diritti civili e politici fino ad ora rimasti sulla carta e nondimeno di rivendicare nuovi tipi di diritti. Essa, così, disegna una nuova possibilità di esercizio di diritti sociali, economici e politici. Com'è stato osservato, infatti, "la cittadinanza nell'era elettronica non solo presuppone un'alfabetizzazione informatica diffusa, ma richiede concrete possibilità di facile accesso alla Rete... Oggi la nuova cittadinanza è il diritto di non essere escluso dall'uso e dal beneficio delle strutture pubbliche e in particolare dall'uso delle risorse telematiche che queste strutture possono offrire. Inclusione e accesso sono momenti essenziali della nuova cittadinanza [Masucci, 2003]". Se, da un lato, vi sono coloro che hanno esaltato i vantaggi derivanti dalle implicazioni tecnologiche della Rete, affermando che essa sarà in grado, di per sé, di arricchire il proprio capitale umano e, in definitiva, di migliorare le proprie possibilità di vita [Costanzo, 1993]; dall'altro, vi è chi sottolinea [Pace, 2006] come la non omogenea diffusione di Internet tra la popolazione provocherà crescenti disuguaglianze, migliorando le prospettive di coloro che già si trovano in posizioni privilegiate e negando, al contempo, opportunità di avanzamento ai non privilegiati.

1.2. Conclusione: Diritto di accesso ad Internet, necessità di una sua esplicita regolamentazione?

In Italia, la questione della tutela di una cittadinanza digitale in funzione della creazione di un diritto di accesso ad Internet sotto il profilo considerato è stata recentemente affrontata, in dottrina, da un autorevole giurista, Stefano Rodotà. L'illustre studioso, all'Internet Governance Forum di Roma, nel mese di novembre 2010 ha elaborato una proposta per l'adozione dell'articolo 21-*bis*, per far sì che Internet sia riconosciuto come diritto fondamentale di tutti i cittadini italiani.

La formulazione dell'articolo in questione recita: "*Tutti hanno eguale diritto di accedere alla Rete Internet, in condizione di parità, con modalità tecnologicamente adeguate e che rimuovano ogni ostacolo di ordine economico e sociale*".

L'art. 21-*bis*, così come proposto, andrebbe a completare l'art. 21, che già esiste e che garantisce la libertà di stampa. L'illustre autore ha sottolineato che un ampliamento della Costituzione, pur se della prima parte, si rende ormai necessaria. Infatti, l'apertura verso questo diritto così prospettato rafforza in modo evidente, il principio di neutralità della Rete e di riflesso la considerazione della conoscenza in Rete come bene comune, fruibile da ogni cittadino. Analizzando il contenuto proposto per la nuova norma si possono cogliere i seguenti punti fondamentali:

- il nuovo articolo mette in luce la problematica del *digital divide* quando enuncia "con modalità tecnologicamente adeguate", e pone sullo Stato l'impegno al suo superamento.
- inoltre, tutti i cittadini, "in condizione di parità", devono avere accesso a Internet: devono poterlo fare "con modalità tecnologicamente adeguate", ovvero fornendo una linea ADSL quantomeno con una velocità minima garantita, senza che il costo di questa connessione ricada sui cittadini stessi ("che rimuovano ogni ostacolo di ordine economico e sociale").

Questo può voler significare che lo Stato dovrebbe intervenire per colmare eventuali carenze infrastrutturali, laddove gli operatori dovessero decidere di non ritenere economicamente conveniente investire, e che lo Stato dovrebbe decidere in anticipo i limiti di questo servizio e quali siano i soggetti coinvolti nel garantire la fornitura dello stesso agli utenti finali. Se l'articolo 21-*bis* diventasse una realtà, Internet dovrebbe essere riconosciuto come diritto costituzionale e social-fondamentale di tutti i cittadini.

Questo indirizzo non si configura come originale, ma è stato già introdotto in Finlandia dal luglio 2010. Tale Nazione ha consentito che le connessioni a banda larga diventassero un diritto per tutti i residenti. Per raggiungere tale obiettivo, il Governo ha stabilito che tutti i *provider* fossero obbligati a fornire a ogni residente una linea a banda larga con una velocità minima di 1 Mbps, che entro il 2015 dovrà raggiungere i 100 Mbps. I metodi per la crescita della banda larga adottate dalla Finlandia hanno fatto affidamento primariamente alle forze di mercato, incoraggiando l'intervento dello Stato solo dove esso fosse

strettamente necessario. Più nel dettaglio, anche nel caso in cui si fossero registrati fallimenti di mercato, i sussidi nazionali non potevano comunque eccedere un terzo della somma necessaria al progetto, con un tetto massimo di un altro terzo per i fondi europei e locali. Le forze private sarebbero, quindi, costrette a coprire almeno un terzo del costo del progetto.

Inoltre, questa notevolissima espansione del settore delle telecomunicazioni, dovrebbe essere supportata anche da un incremento di informatizzazione del settore pubblico attraverso una effettiva espansione dell'utilizzo delle nuove tecnologie. Compito dello Stato è, così, quello di eliminare gli ostacoli, non solo tecnici, ma anche cognitivi che ostacolano la vera uguaglianza digitale

Tra le politiche mirate a sviluppare il grado di accesso ad *Internet*, la maggior parte dei Paesi si è mossa promuovendo specifiche proposte di *skill training*, che coinvolgono luoghi di lavoro e scuole. Il più vasto programma di alfabetizzazione informatica è quello sviluppato dalla Corea che, con diversi programmi negli ultimi anni, ha impiegato diverse risorse. Il programma più determinante è stato quello di formazione dedicato alle casalinghe "*Cyber21*". Il piano consisteva in un corso di una settimana (in tutto 20 ore di lezione) che interessava oltre un migliaio di istituti di formazione ed era sovvenzionato interamente dal governo. Dev'essere ancora ricordata la Corea per l'ulteriore organizzazione di corsi di formazione gratuiti o a prezzi assai bassi per vaste porzioni della cittadinanza. Infatti, il programma "*10 Million People IT Education Project*", era rivolto ad aumentare la conoscenza delle nuove tecnologie, inglobando non solo le scuole, ma anche le persone anziane e i disabili, personale militare ed agricoltori. In Europa, invece, il processo di alfabetizzazione digitale affonda le sue radici nei programmi di *eLearning*, incentrati sul miglioramento delle competenze ICT di studenti e insegnanti. Questo metodo consente al fruitore di intraprendere un percorso di autoistruzione utilizzando e sfruttando i dati e i materiali presenti nella Rete.

Caratteristica peculiare di questa metodologia è la flessibilità, in quanto assicura contemporaneamente formazione permanente e inclusione sociale. In Inghilterra, è stato avviato il programma "*train to Gain*", che prevede una formazione meticolosa sul posto di lavoro. Esso ha coinvolto 127.000 datori di lavoro e fornito corsi di formazione ad oltre un milione di lavoratori. Inoltre, tutto ciò è stato incentivato da campagne pubblicitarie *ad hoc*, l'impiego di appositi *roadshow*, con bus che girando per le città promuovono e diffondono le potenzialità delle nuove tecnologie informatiche.

Pur ritenendo che l'esplicita previsione del diritto di accesso alla Rete come diritto fondamentale, tramite un nuovo articolo aggiunto alla Costituzione, eliminerebbe in radice ogni problema in ordine all'esistenza e alla qualificazione giuridica del diritto in parola, si può ritenere che la stessa non risulti strettamente necessaria. Ritenere che debba essere presente una specifica disposizione di rango costituzionale significa, in buona sostanza, dubitare, se non negare, che allo stato questo diritto, con siffatti connotati, non sia configurabile nel nostro ordinamento. Di questo, in realtà, si deve dubitare fortemente.

Bibliografia

Bovero M., La libertà e i diritti di libertà, in Quale libertà . Dizionario minimo contro i falsi liberali, a cura di Bovero, Roma-Bari, 2004.

Costanzo P., Informazione nel diritto costituzionale, in Digesto disc. Pubbl., VIII, Torino, 1993.

Costanzo P., Internet (Diritto pubblico) , Digesto (Disc. Pubbl.), Appendice, Torino, 2000.

Di Lello C., Internet e costituzione: garanzia del mezzo e suoi limiti, in Diritto dell'informazione e dell'informatica, 2007.

Ferrajoli L., Costituzionalismo principalista e costituzionalismo garantista, in Giurisprudenza costituzionale, 2010.

Frosini T.E., L'orizzonte giuridico dell'Internet, in Il diritto dell'informazione. e dell'informatica, 2002.

Frosini T.E., La protezione della riservatezza nella società informatica, in Privacy e banche dati, a cura di Matteucci, Bologna, 1981.

Frosini T.E., La libertà informatica: brevi note sull'attualità di una teoria giuridica, in Informatica e diritto, 2008.

Loiodice A., Informazione, in Enciclopedia del diritto, Milano, 1971, XXI.

Loiodice A., Libertà di comunicazione e principi costituzionali, in Trattato di diritto costituzionale, Padova, 1999, XXVIII.

Masucci A., Erogazione on line dei servizi pubblici e tele procedure amministrative, in Diritto pubblico, 2003.

Pace A.-Manetti M., Art. 21, in Comm. Cost., a cura di Branca-Pizzorusso, Bologna, 2006.

Pietrangelo M., La società dell'informazione tra realtà e norma, Milano, 2007.

Pino G., Diritti e interpretazione. Il ragionamento giuridico nello Stato costituzionale, Bologna, 2010.

Pubusa F., Diritto di Accesso ed Automazione, Torino, 2006.

Rodotà S., Libertà, opportunità, democrazia e informazione, in Internet e Privacy: quali regole?, Atti del convegno organizzato dal Garante per la protezione dei dati personali, Roma, 1998.

Rozo Acuna E., Habeas Data costituzionale: nuova garanzia giurisdizionale del diritto pubblico latinoamericano, in Diritto pubblico comparato ed europeo, 2002.

Valastro A., Le garanzie di effettività del diritto d'accesso ad Internet e la timidezza del legislatore italiano, in Il diritto d'accesso ad Internet. Atti della tavola rotonda svolta nell'ambito dell'IGF Italia 2010 , a cura di Pietrangelo, Napoli, 2011.

Zeno-Zencovich V., La libertà d'espressione. Media, mercato, potere nella società dell'informazione, Bologna, 2004.

La fiducia degli utenti nelle transazioni online tra sicurezza reale e percepita

Mario Petrone
Università degli Studi del Molise
Via De Sanctis s.n.c. 86100 Campobasso (CB)
petrone@unimol.it

Abstract. *The fundamental precondition for the acceptance and success of online transactions is in the trust and usability of the systems used. The usability and trust in online transactions are related to safety, real and perceived by the users. Consequently, the success of e-commerce, e-government or e-health is directly dependent on the relationship between the actual safety and perceived by the users. In this paper is presented the results of the analysis of that bond and of the relationship between security, trust and usability. Based on the above tests have been some measures to improve the acceptance and, consequently, the success of online transactions. A qualitative research approach was adopted, since the research seeks an understanding of human (i.e., customer and organizational employee) perceptions.*

Keywords: E-commerce, trust, , security.

1. Introduzione

Le tecnologie dell'informazione e della comunicazione (ICT) costituiscono uno strumento importante per il miglioramento della produttività del sistema economico; rappresentano un'opportunità per lo sviluppo del paese, per la competitività, per la crescita culturale e sociale; consentono di aumentare la conoscenza e la concorrenza nei mercati delle merci e dei servizi; richiedono un forte impegno da parte dei soggetti coinvolti: imprese, cittadini, istituzioni pubbliche.

Le indagini condotte mostrano che, nonostante i notevoli progressi compiuti negli ultimi anni, l'utilizzo delle transazioni online nei diversi settori dell'economia non ha ancora raggiunto livelli adeguati alla sviluppo della società dell'informazione: occorre far percepire maggiormente alle imprese (che usano la rete in prevalenza per presentare prodotti e servizi) i vantaggi delle transazioni telematiche in termini di miglioramento dell'efficienza aziendale, acquisizione di nuovi clienti e riduzione dei costi commerciali; occorre

accrescere la consapevolezza che i vantaggi aumentano se l'intera filiera (dalla produzione alla commercializzazione) è in rete.

Le applicazioni di e-commerce, e-government o e-health presentano ampi margini di miglioramento, anche se dati recenti segnalano un aumento della domanda. Per la sua ulteriore crescita occorre fugare il timore sui "fattori di rischio" (frodi, incertezza delle controparti, sicurezza della rete). Al riguardo si rileva, infatti, che ci sono diverse componenti che influenzano la relazione che si instaura nel corso di una transazione online; una tra le più importanti è la fiducia [RIF1] [Chiles and McMackin- 1996].

Con riferimento all'impiego dell'Information and Communication Technology (ICT), possiamo distinguere due tipi principali di relazioni in base a cui si individuano le tipologie di fiducia:

- tradizionale: relazioni in cui l'ICT gioca un ruolo secondario. In questo caso si individuano due tipologie principali di fiducia: istituzionale [RIF2] [McKnight, 1998] e sociale (spesso definita come "customer trust") [RIF3] [Granovetter, 1985].
- digitale (o online): relazioni totalmente basate sull'ICT. In questo contesto (e-business/ e-commerce/e-service/ e-marketplace) occorre considerare tre diverse tipologie di fiducia: istituzionale, sociale e tecnologica [RIF4] [Reeves and Nash 1996], [RIF5] [Misiolak et al. 2002], [RIF6] [Ratnasingam and Pavlou, 2002].

Nelle relazioni digitali la fiducia è considerata cruciale (Ba et al. 1999). Poiché le reti informatiche vengono considerate come un ambiente "non sicuro" (Ratnasingam 2002), l'ICT influenza fortemente il livello di fiducia [RIF5] (Misiolak et al. 2002) relativamente ai seguenti aspetti:

- sociale: fortemente legata alla percezione del rischio nello scambio di informazioni tra le parti [RIF7] [Koller 1988], influenzata principalmente da esperienze positive;
- organizzativo: relativa alle relazioni tra individui e organizzazioni supportate dall'information technology [RIF8] [Lewicki & Bunker, 1996], [RIF9] [Tyler & Degoey, 1996], [RIF10] [Pavlou et al. 2003], [RIF11] [Spagnoletti et al. 2007], [RIF2] [McKnight et al. 1998];
- tecnologico: relativo al rapporto con l'ICT usato a supporto per lo scambio delle informazioni (Reeves and Nash 1996), anche definita come la probabilità con cui le organizzazioni credono che l'infrastruttura tecnologica sottostante è in grado di facilitare le transazioni compatibilmente con le loro aspettative [RIF6] [Ratnasingam e Pavlou, 2002].

Con l'obiettivo di analizzare più approfonditamente il concetto di fiducia tecnologica, in questo lavoro non si considerano i fattori che influenzano direttamente la fiducia sociale o istituzionale (per esempio i meccanismi di feedback che mirano ad aumentare la reputazione del soggetti a cui fanno riferimento) nel contesto delle transazioni online; ma piuttosto si esaminano i meccanismi che agiscono direttamente sul livello di fiducia nella tecnologia utilizzata a supporto della relazione.

Le funzionalità tecnologiche che influenzano in maggior misura il livello di fiducia e l'usabilità dei sistemi sono quelle legate ai meccanismi di sicurezza associati alle seguenti fasi:

- accesso alle informazioni o risorse in cui viene presentata la richiesta e validata attraverso meccanismi di autenticazione, identificazione e autorizzazione;
- trasferimento delle informazioni in cui esse vengono trasmesse utilizzando una serie di mezzi e policy per garantirne la confidenzialità e l'integrità;
- gestione delle informazioni focalizzata sul garantire la loro disponibilità, necessaria per adempiere alle richieste valide, e per far fronte a potenziali perdite imprevedute dei dati (fault tolerance, backup, copie dei dati delocalizzate, etc.).

La complessità tecnica, la mancanza di trasparenza organizzativa e l'incertezza giuridica rendono difficile agli utenti valutare la sicurezza delle transazioni online e l'affidabilità delle parti interessate, riducendo di conseguenza la loro volontà di utilizzare i sistemi online. La verifica dell'affidabilità richiede notevoli sforzi. Le transazioni online pongono gli utenti di fronte a una serie di domande:

- come si può riconoscere l'affidabilità di un fornitore di servizi?
- come si può verificare la qualità di servizi o prodotti?
- qual è il grado di sicurezza e di affidabilità della tecnologia di trasmissione (cifratura, autenticazione)?
- cosa accade alle informazioni relative alla mia carta di credito?
- cosa accade in caso di mancata o errata consegna? Come funziona la gestione dei reclami?
- come si riconoscono le promesse di vincita fraudolente (lottery scams) o le truffe dell'anticipo (advanced fee frauds)?

Molteplici sondaggi rivelano che, con riferimento agli acquisti su Internet, la maggior parte degli utenti considera la protezione dei dati come la cosa più importante, seguita dalla sicurezza delle transazioni, dalla trasparenza dei processi e dal tempo necessario per effettuare l'acquisto online.

Relativamente all'analisi dei rischi, gli esperti sono unanimi: hacker, virus e worm continuano a mostrarci ogni giorno che i sistemi informatici non possono essere resi sicuri al 100%. Dove non si può fare affidamento sulla tecnica, il controllo dei rischi deve avvenire a livello sociale, anche se non è cosa facile, poiché le conoscenze in merito a pericoli e vantaggi dei sistemi complessi e al loro effetto sulla nostra vita sono obsolete già appena elaborate.

2.Fiducia e usabilità tra sicurezza reale e percepita

In considerazione del progressivo aumento dell'importanza economica, sociale e politica delle transazioni online, la fiducia che gli utenti attribuiscono ai metodi impiegati e l'usabilità dei sistemi implementati, rappresentano un presupposto fondamentale per la loro accettazione e il loro successo.

Venendo a mancare nelle transazioni online la possibilità di creare relazioni di fiducia basate su contatti interpersonali diretti tra le persone che occupano un ruolo di confine, è necessario individuare meccanismi di altro tipo, seppur finalizzati agli stessi obiettivi, di creazione delle condizioni di fiducia reciproca

basate su tre elementi essenziali [RIF2] [McKnight et al. 1998]; [RIF10] [Pavlou et al. 2003]):

- la percezione di interagire con un sistema tecnico sicuro;
- la presenza di meccanismi inter-organizzativi formali;
- la garanzia che le comunicazioni tra i sistemi ICT avvengano in maniera corretta grazie a standard tecnici e procedurali.

Per poter inquadrare correttamente i suddetti elementi è opportuno considerare che l'usabilità e la fiducia nelle transazioni online sono direttamente dipendenti dal rapporto tra la sicurezza reale e quella percepita dagli utenti. Conseguentemente il successo delle applicazioni di e-commerce, e-government o e-health è inevitabilmente connesso al suddetto rapporto. Come sinteticamente schematizzato nella Fig. 1, il raggiungimento di un alto livello di sicurezza reale consente di creare un ambiente sicuro mentre un basso livello di sicurezza reale determina la creazione di un ambiente insicuro e, quindi, rischioso. Parallelamente un innalzamento della sicurezza percepita origina sistemi affidabili e utilizzabili, mentre ad un basso livello di sicurezza percepita corrispondono sistemi non affidabili e poco utilizzabili.

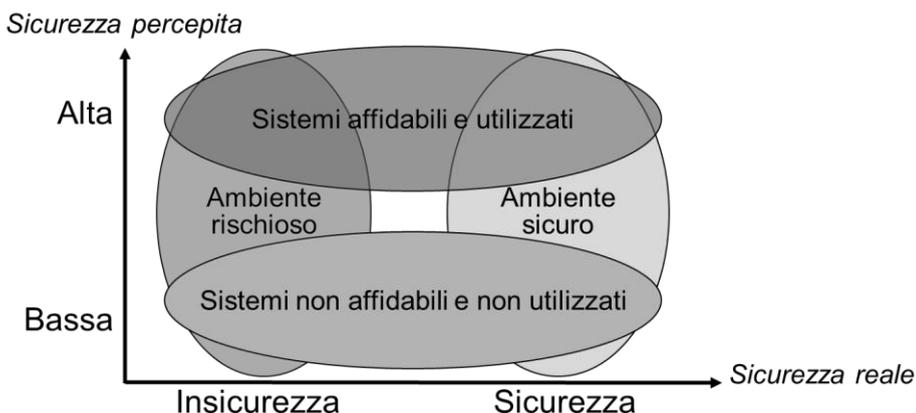


Fig.1 - Rapporto tra sicurezza “Sicurezza reale” e “Sicurezza percepita”

Considerato che ad un aumento della fiducia e dell'usabilità dei sistemi implementati corrisponde un maggiore successo delle transazioni online, è stata analizzata l'influenza che su di esse esercitano gli interventi finalizzati ad un incremento della sicurezza sia reale che percepita.

Nella Fig.2 viene sintetizzato il legame esistente tra la sicurezza reale e l'usabilità nonché l'impatto che detto legame comporta sul rapporto tra detta sicurezza e le transazioni online. In particolare, si rileva che ad un incremento della sicurezza reale corrisponde un incremento della complessità operativa e, quindi, un deterioramento dell'usabilità complessiva del sistema. Conseguentemente ad un incremento della sicurezza reale corrisponde una resistenza alla diffusione delle transazioni online.

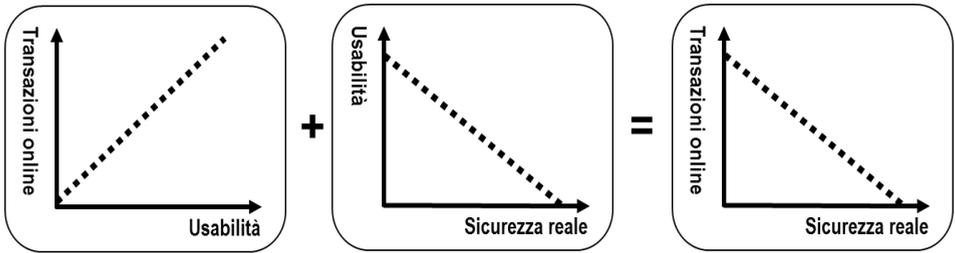


Fig.2 – Legame tra “Sicurezza reale”, “Transazioni online” e “Usabilità”

Nella Fig.3 viene sintetizzato il legame esistente tra la sicurezza percepita e la fiducia nonché l’impatto che detto legame comporta sul rapporto tra la suddetta sicurezza e le transazioni online.

In particolare, si rileva che ad un incremento della sicurezza percepita corrisponde un incremento della fiducia. Conseguentemente un incremento della sicurezza percepita concorre alla diffusione delle transazioni online.

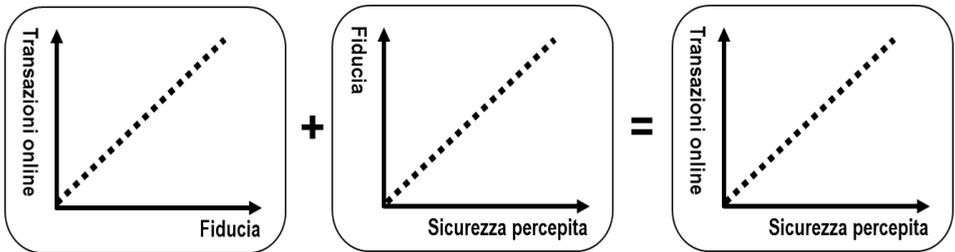


Fig.3 - Legame tra “Sicurezza percepita”, “Transazioni online” e “Fiducia”

Sulla base del confronto dei legami rappresentati nella Fig.2 e nella Fig.3, è stato possibile analizzare il rapporto tra la sicurezza, reale e percepita, e le transazioni online (Fig.4).

In particolare, da tale confronto si rileva che esiste uno spazio operativo entro il quale è opportuno far rientrare gli interventi nell’area della sicurezza in modo tale che gli stessi tengano conto delle problematiche connesse alla fiducia nonché all’usabilità.

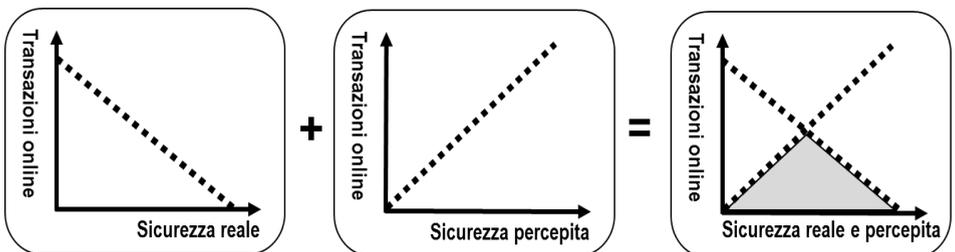


Fig.4 - Legame tra “Sicurezza reale”, “Sicurezza percepita” e “Transazioni online”

L'introduzione di nuovi meccanismi di sicurezza deve essere organizzata, mantenendo un equilibrato rapporto tra la sicurezza reale e quella percepita dagli utenti, in modo da creare fiducia ed allo stesso tempo garantire l'usabilità dei sistemi.

Si tratta di una fase in cui l'incertezza è forte e il senso di sicurezza debole [RIF12] [Siegrist, Earle & Gutscher, 2003]. Dal momento che non ci si può basare su esperienze pregresse, la confidenza non gioca alcun ruolo. In una situazione di questo genere la fiducia può essere creata promuovendo o garantendo visibilità a determinate persone o istituzioni personificate presso i gruppi target.

Si pensi, ad esempio, a persone che esprimono esigenze credibili e comparabili in materia di sicurezza, affidabilità e qualità di una soluzione tecnologica e che sviluppano, testano, consigliano o impiegano con successo una nuova tecnologia (similitudine di valori, esigenze, obiettivi comparabili).

Se sono disponibili persone o istituzioni o modelli di comportamento che rispondono a tali requisiti (pionieri), possono essere presi a riferimento anche da altre cerchie di persone.

Inoltre, tra i principali ostacoli alla diffusione delle transazioni online sono state individuate difficoltà legate ai limiti degli ambienti di supporto alla collaborazione ed all'inefficienza di alcune configurazioni organizzative.

In particolare, “non sono ancora state trovate modalità per favorire la nascita di un clima di fiducia e sicurezza all'interno della comunità” [RIF13] [Barbini 2004].

Un ulteriore elemento critico nei processi di implementazione dei sistemi inter-organizzativi è individuare la quantità e la qualità delle informazioni che le organizzazioni sono disposte a condividere.

Diversi autori hanno dimostrato l'importanza della presenza di una condizione di fiducia reciproca che consenta una forte integrazione anche senza dover costituire una relazione di controllo proprietario tra le controparti [RIF14] [Martinez 2004].

In questo scenario, l'intervento di una terza parte potrebbe agevolare la costruzione della fiducia, fornendo un supporto per superare differenze culturali ed eventuali barriere normative che intervengono ad esempio nella collaborazione tra utenti di diversa nazionalità.

Il ruolo di tali soggetti terzi dovrebbe infatti garantire, ad esempio creando un marketplace o un portale, che le transazioni avvengano in un ambiente certificato e “sicuro” in termini di soluzioni tecnologiche adottate per la comunicazione, di regolarità formale degli accordi e di fiducia istituzionale.

Ciononostante, la configurazione organizzativa incentrata sulla figura dell'integratore (broker), che rappresenta la più diffusa forma presente in letteratura [RIF15] [Bremer 1999], ha mostrato i suoi limiti sia in termini di capacità di identificare e gestire il business, sia in termini di capacità di gestire e controllare l'intera rete virtuale [RIF16] [D'Atri 2003]. Alcuni studi considerano infatti tra i principali fattori che ostacolano la diffusione di servizi online, elementi quali la fiducia, la percezione del rischio da parte dei clienti ed i problemi di

privacy. Inoltre questo tipo di fattori assume un peso diverso in contesti culturali differenti ed influenza la disponibilità degli utenti ad affidare i propri dati online a soggetti terzi [RIF17] [Dinev et al. 2006].

3. Metodo e risultati della ricerca

Per l'analisi della percezione della sicurezza tangibile ed intangibile da parte degli utenti è stato utilizzato un approccio qualitativo suggerito da A. Strauss, and J. Corbin in "Basics of qualitative research: grounded theory procedures and techniques" [RIF18]. Ai partecipanti alla rilevazione sono state poste domande aperte finalizzate ad acquisire la loro percezione della sicurezza relativamente ad un sito web. Sono stati somministrati 30 questionari a risposta aperta a tre gruppi di utenti dell'Università degli Studi del Molise: Docenti (10), Personale Tecnico-Amministrativo (10) e Studenti (10). Sulla base dei dati acquisiti è stata predisposta la Tabella 1 nella quale sono riportati gli indicatori tangibili ed intangibili di sicurezza rilevati dalle risposte fornite dagli utenti. Alcuni indicatori intangibili sembrano identici: per esempio notorietà, reputazione, rilevanza potrebbe essere considerati sinonimi. Tenuto conto della necessità di effettuare una ulteriore ricerca finalizzata alla verifica dei risultati che possono derivare dall'utilizzo degli indicatori rilevati, si è volutamente conservato quanto scritto dai partecipanti all'indagine.

Caratteristiche di sicurezza materiale ed immateriali	Tipologia caratteristica di sicurezza
Lucchetto	Tangibile
Certificato di sicurezza	Tangibile
Trasferimento tra interfacce del sito	Tangibile
Presentazione del sito web	Tangibile
Politiche di sicurezza	Tangibile
Riconoscimento tramite email	Tangibile
Simboli di terze parti	Tangibile
Indirizzo fisico , telefono ed e-mail	Tangibile
Identificazione dei problemi di sicurezza	Tangibile
Identità nota (azienda dispone di edificio fisico , es. sede Banca)	Intangibile
Sistema di gestione delle password	Tangibile
Notorietà del sistema di pagamento elettronico, come PayPal	Tangibile/Intangibile
Notorietà del marchio/azienda	Intangibile
Ambito operativo (nazionale, internazionale)	Intangibile
Riconoscibilità	Intangibile
Fiducia	Intangibile
Notorietà	Intangibile
Formalità del sito	Intangibile
Rispettabilità della società (nel caso di società di grandi dimensioni)	Intangibile
Reputazione	Intangibile
Rilevanza	Intangibile

Tabella 1 - Indicatori tangibili ed intangibili di sicurezza

Le risposte degli intervistati hanno consentito di individuare alcuni indicatori tangibili ed intangibili che possono essere utilizzati per migliorare la sicurezza, reale e percepita, di un sito web.

Dalla lettura della Tabella 1 si rileva che gli indicatori tangibili sono quelli di natura tecnologica come ad esempio HTTPS , lucchetti , certificati e simboli di sicurezza, mentre quelli intangibili sono direttamente derivati dalle conoscenze degli utenti.

Gli indicatori intangibili sono influenzati dalla società in termini di comunicazione e dal contesto in cui l'utente opera nonché dalle conoscenze che acquisisce da altri utenti e dalle esperienze maturate come nel caso della notorietà e della reputazione del sito.

La percezione delle caratteristiche intangibili è influenzata da comunicazioni informali e dal passaparola.

Gli indicatori tangibili devono essere compresi e verificati dagli utenti accedendo al sito web, piuttosto che acquisito attraverso l'interazione con altri utenti. Occorre che l'utente abbia alcune specifiche conoscenze ed esperienze che gli consentono di rilevare le informazioni necessarie per la valutazione della sicurezza come nel caso dei certificati di sicurezza relativamente al quale l'utente deve sapere che cosa significa averne uno, e come si può controllare se è scaduto o meno.

4. Conclusioni

Partendo dalla considerazione che la fiducia che gli utenti attribuiscono ai metodi impiegati e l'usabilità dei sistemi ICT implementati, rappresentano un presupposto fondamentale per l'accettazione ed il successo delle transazioni online, nel presente lavoro sono state analizzate le interdipendenze che la fiducia e l'usabilità hanno con la sicurezza reale dei sistemi e quella percepita dagli utenti.

Più in particolare è stato dimostrato che, mantenendo un equilibrato rapporto tra la sicurezza reale e quella percepita dagli utenti, è possibile creare fiducia ed allo stesso tempo garantire l'usabilità dei sistemi ottenendo come conseguenza una maggiore accettazione e diffusione delle transazioni online.

Al fine di individuare gli ambiti e le modalità di intervento è stata realizzata una indagine che ha consentito di individuare alcuni indicatori tangibili ed intangibili che possono essere utilizzati per migliorare la sicurezza, reale e percepita, di un sito web.

Sulla base dei risultati della ricerca effettuata coinvolgendo diverse tipologie di utenti, sono state definite alcune misure da adottare per migliorare l'accettazione e, conseguentemente, il successo delle transazioni online.

In particolare sono state individuate le seguenti linee direttive:

- la trasmissione diretta di conoscenze alla popolazione e alle imprese;
- lo sviluppo delle competenze dei mediatori di gruppi target (insegnanti, genitori, coetanei degli adolescenti, forze di polizia);
- un portale multicanale di informazione attraverso il quale garantire il reperimento delle informazioni e delle consulenze qualitativamente elevate e affidabili già esistenti;

– il collegamento e la coordinazione degli attori.

A sostegno degli obiettivi delle suddette linee direttive vanno aggiunti i seguenti azioni:

– la comunicazione deve consentire una divulgazione, adatta ai rispettivi gruppi target, delle attività previste dalle quattro linee direttive;

– la ricerca deve fornire una base per le decisioni politiche attraverso la realizzazione e l'ampliamento delle rilevazioni relative all'utilizzo delle soluzioni tecnologiche e la valutazione dell'efficacia delle iniziative.

Le fasi successive della ricerca saranno volte a verificare l'efficacia delle direttive proposte nonché ad analizzare l'utilizzo di soluzioni applicative combinate i sistemi biometrici di identificazione dell'utente e le relative ripercussioni sugli aspetti legati alla fiducia ed all'usabilità.

Bibliografia

[RIF1] Chiles, T.H. and McMackin, J. (1996). Integrating variable risk preferences, trust, and transaction cost economics, *Academy of Management Review* 21 73–99

[RIF2] McKnight, D.H., Cummings, L.L. E Chervany, N.L. (1998). Initial Trust Formation in New Organizational Relationships, in «*Academy of Management Re-view*», vol. 23, n. 3

[RIF3] Granovetter M. (1985), "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology*, 91(November): 481-510

[RIF4] Reeves, B., & Nass, C. (1996). *The media equation. How people treat computers, television, and new media like real people and places.* New York: Cambridge University Press.

[RIF5] Misiulek, N.I., N. Zakaria, and P. Zhang (2002). Trust in organizational acceptance of information technology: A conceptual model and preliminary evidence. In *Proc. Decision Sciences Institute 33rd Annual Meeting 2002*

[RIF6] Ratnasingam, P. and Pavlou, P. (2002), "Technology trust: the next value creator in B2B electronic commerce", *International Resources Management Association Conference - Washington, Seattle*

[RIF7] Koller, M. (1988). "Risk as a Determinant of Trust," *Basic and Applied Social Psychology* Volume 9, Issue 4, pp. 265-276

[RIF8] Lewicki, R.J., & Bunker, B.B. (1996). Developing and maintaining trust in work relationships. In R.M. Kramer & T.R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 114-139). Thousand Oaks, CA: Sage Publications

[RIF9] Tyler, T.R., & DeGoey, P. (1996). Trust in organizational authorities. The influence of motive attributions on willingness to accept decisions. In R.M. Kramer & T.R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 331-350). Thousand Oaks, CA: Sage Publications

[RIF10] Pavlou, P., Tan, Y.H. and Gefen, D. (2003), *Institutional Trust and Familiarity in Online Interorganizational Relationship*

[RIF11] Spagnoletti P., Za S., D'Atri A., (2007). "Institutional Trust and security, new boundaries for Virtual Enterprises", *Proc. of 2nd International Workshop on*

Interoperability Solutions to Trust, Security, Policies and QoS for Enhanced Enterprise Systems, IS-TSPQ2007, Funchal, Portugal

[RIF12] Siegrist, M., Earle, T. C. & Gutscher, H. (2003). Test of a trust and confidence model in the applied context of electromagnetic field (EMF) risks. *Risk Analysis*, 23, 705–716

[RIF13] BARBINI F.M. (2004), Il contributo dell'Impresa Virtuale all'Innovazione Organizzativa, in D'Atri A. (a cura di), *Innovazione Tecnologica e Tecnologie Innovative*, EtasLibri, Milano

[RIF14] MARTINEZ, M., (2004), *Organizzazione, informazioni e tecnologie*. Il Mulino

[RIF15] BREMER (1999), *Global Virtual Business: A Systematic Approach for Exploiting Business Opportunities in Dynamic Markets*, *International Journal of Agile Manufacturing*, vol.2, issue 1

[RIF16] D'ATRI, A. (2003). Organising and Managing Virtual Enterprises: the ECB Framework, in Camarinha-Matos, L.M. 2+and Afsarmanesh, H. (eds.). *Processes and Foundations for Virtual Organisations*. Kluwer Academic Publishers

[RIF17] DINEV, T., BELLOTTO, M., HART, P., RUSSO, V., SERRA, I., COLAUTTI, C. (2006). Privacy Calculus Model in E-commerce - a Study of Italy and the United States. *European Journal of Information Systems*, 15, 4, 389-402

[RIF18] A. Strauss, and J. Corbin, *Basics of qualitative research: grounded theory procedures and techniques*. SAGE Publication, London, 1990.

A Tool for Multimodal Fusion of Brain PET and MR

Giovanni Capobianco¹, Madalina Georgeta Ciobanu²,
Fausto Fasano¹, Matteo Merola¹

¹X-Lab, DiBT, University of Molise

C.da Fonte Lappone, 86090 Pesche (IS)

giovanni.capobianco@unimol.it, fausto.fasano@unimol.it, m.merola1@studenti.unimol.it

²University of Salerno

Via Giovanni Paolo II, 132 - 84084 - Fisciano (SA)

mciobanu@unisa.it

Abstract. *Neurological disorders diagnosis makes use of several diagnostic examinations, including cerebral medical imaging. In this field, two of the most widely used examinations are brain PET (Positron Emission Tomography) and MR (Magnetic Resonance).*

We present a technique and a tool for the integration of morphologic and functional information from brain MR and PET images into a single fused image. The registration of the image is based on the elliptic Fourier descriptors, while the different images have been merged using the multiresolution analysis based on the wavelet transformation. The Graphic User Interface is implemented in the MATLAB® environment. Two experienced neuroradiologists evaluated the accuracy of our approach through visual inspection of contours continuity and using some mathematical indices computed, such as universal image quality index, root mean square error, and mutual information metrics. The tool aims at supporting the doctor during the investigation of metabolic activity within the anatomically defined abnormality. This is usefull in the diagnosis of many neurodegenerative diseases and in the planning and follow-up of radiation treatment.

Keywords: medical image computing, neuroinformatic, computer supported diagnostics.

1. Introduction

Bioinformatics aims at providing methods and tools for storing, retrieving, organizing and analysing biological data and generate useful biological knowledge [5]. In recent years, the use of biomedical images has become more practical and easier thanks to the advances of technology. On the other hand,

the availability of adequate computational power has enabled the development of new disciplines, such as medical image computing and computer supported diagnostics. In the next future, we believe the computer will become a fundamental allied for the doctor during diagnostic and therapeutic decision.

The availability of information and biological-functional anatomy of the patient in a single examination is, for the insiders, a quality that enriches and completes the clinical picture of the subject and provides a set of information that allows a framework wider than the single case study.

In this work we define a fusion technique to merge medical images from different types of exams. In particular, PET (Positron Emission Tomography) and MRI (Magnetic Resonance Imaging), considered alone, are widely used in the diagnosis and treatment of pathologies affecting the central nervous system. Our goal is the creation of a new merged image that contains the structural information derived from the MR examination and functional information derived from the PET examination. To this aim, we propose a technique for the integration of morphologic and functional information from brain MR and PET images into a single fused image.

As the images are the output of two different exams, the first issue is to apply some transformation in order to make them aligned in a three-dimensional space. Multi-modal image registration technique [10] aims to establish a common reference between the two exams in order to align them for the fusion phase. The result of the fusion simplifies different scenarios. It helps doctors in localizing and staging cancer, in diagnosing of dementia, in identifying epileptic foci, in preparing surgical procedures in patients with critical issues in areas close to cancers and in decreasing the number of examinations to be applied to the patient. We propose to use the elliptic Fourier descriptors to solve this problem.

The second issue concerns the merging of the different images. In this case, we decided to use the multi-resolution analysis based on the wavelet transformation.

We also realised a tool that implements the proposed method leveraging the MATLAB^{®1} environment. The tool aims at supporting the doctor during the investigation of metabolic activity within the anatomically defined abnormality. This is useful in the diagnosis of many neurodegenerative diseases and in the planning and follow-up of radiation treatment.

We evaluated the effectiveness of both the registration and fusion phases. Two experienced neuroradiologists assessed the accuracy of the former phase through visual inspection of contours continuity. The latter phase has been assessed by using some mathematical indices, such as universal image quality index [14] root mean square error, and mutual information metrics [2].

The remainder of the paper is organized as follow. Section 2 presents related works. Sections 3 and 4 describe the proposed method and tool, respectively. Section 5 concludes the paper and gives some indication for future work.

¹ <http://www.mathworks.it/products/matlab/>

2. Related Work

The need for an integrated PET / MR is one of the ultimate goals in the evolution of positron emission tomography [3].

These techniques have recently provided an impetus to the discovery of methods for the early detection of diseases such as cervical cancer [12], for the staging of gliomata [7][1], for the preparation of surgical patients in critical areas next to tumours [13], for the diagnosis of dementia [4] or for the identification of fires epileptics [8]. Palubinskas G. and Reinartz [11] propose a general fusion framework consisting of main three steps: interpolation, fusion and histogram matching, even if they focused on multi-resolution and multi-sensor data fusion from satellite images.

Because of the high information intake that multi-modality imaging can provide, first with the attempts to fuse PET and CT (Computed Tomography) images [16], CT and MR images [15], and PET and MR images [6], software tools have been proposed.

Among these, the work by Deshmukh and Bhosale [2] is the most similar, as they present a PCA (Principal Component Analysis) based image fusion and focus on image fusion algorithm based on wavelet transform to improve resolution of the images. They also propose image fusion in Radon space. However, we also integrated the approach in an easy to use tool that easily supports the doctor during the analysis of the images and diagnosis of many neurodegenerative diseases.

3. The Proposed Approach

In this Section, we briefly describe each phases of our approach.

3.1. Input/Output

The first step consists of the definition of input/output data format. Among the several formats in which examinations can be represented, the most widely used are Analyze 7.5 and DICOM. We focused on the former because of its greater ease of use and the fact that the conversion from DICOM to Analyze 7.5 does not use any type of compression and, therefore, no information is lost during the conversion. Each individual DICOM slice is converted in a single file with the extension .img, while the header of the DICOM slice is contained in the .hdr file and contains the most important information contained in the original files.

The output of the program is a version DICOM examination containing the fusion of the two modes that can be saved for future reference and / or evaluations. Moreover, a fused image is also displayed in output to the end of the last stage of processing in order to be immediately available to the user.

3.2. Pre-processing

After the input phase of the program, the input images are pre-processed in order to make processing more efficient and effective. As an example, 16-bit greyscale images are converted is more suitable format for MATLAB®, i.e., a

16-bit integers ranging from -32768 to +32768 where the 0 is not black – as in greyscale images – but a medium grey.

3.3. Alignment and interpolation

After the pre-processing, to merge the two series of images into a single series, we perform an alignment phase to project them in a common space. Indeed, after the alignment and interpolation phases and after the registration phase the actual size of the individual voxels in both modes must be identical.

First of all, we analyse the header file for each series of images to determine the actual size - in millimetres - of the single unit. Then we proceed to lead the two modes to a common reference dimension.

Subsequently, we focus on the number of slices of the two tests and their actual depth size. In fact, the number of images that compose an examination varies according to the type of examination and the machine that performs the acquisition. We developed a procedure to uniform the number of slices of each test.

3.4. Registration

Once the two examinations are aligned in terms of resolution and actual size, the registration phase makes the information of the two examinations available to the fusion phase. To this aim, we have to spatially align them to a common reference. The recording technique developed in this work is based on the elliptical Fourier analysis. The elliptic Fourier descriptors are parameters that, given a closed contour, approximate it by means of trigonometric functions. These coefficients are calculated with the Discrete Fourier Transform using the procedure proposed by [9] applied to concatenated numbers that represent the borders of the image.

3.5. Fusion

When the acquisition and normalisation phases are complete, images can finally be merged. The fusion algorithm exploits the potential of multi-resolution time-frequency analysis typical of the wavelet transform. With this kind of transformation we can extract the most important details of the two series of images and merge them into a new image. An example of the fusion of a single slice is shown in Figure 1.

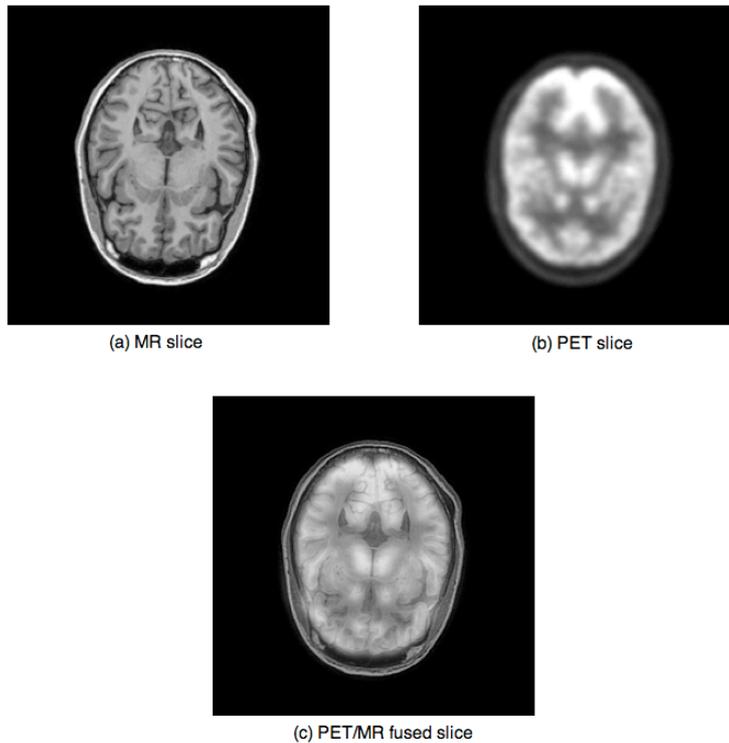


Figure 1. Fusion of PET and MR images

The fusion algorithm follows these steps:

1. decompose the image using a symlet wavelet transformation that takes the slice break as input and produces the following four matrices:
 - the matrix of the approximation of the image to decompose (low frequencies);
 - the matrix of horizontal details (horizontal high frequency);
 - the matrix of vertical details (vertical high frequency);
 - the matrix of diagonal details (diagonal high frequencies).
2. decompose the image using a symlet wavelet transformation that takes the above mentioned matrix of the approximation as input and returns four new matrices of level 2 for the approximation, horizontal details, vertical details, and diagonal details;
3. use the weighted average of coefficients of the approximation matrix of the previous step to produce a new fused matrix representing the low frequencies of the fused image;
4. merges all the level 2 matrices; for each pair of coefficients chooses the pixel that represents the most evident detail, e.g., the pixel with the highest value in the two modes;

5. merges the level 1 matrices using a similar approach as in step 4;
6. apply the inverse of the transformation of the step 2 to the level 2 matrices; this produces a fused image of level 1;
7. apply the inverse of the transformation of the step 1 to the level 1 matrices and the image produced in the step 6; this produces a final fused slice of the input images;
8. execute steps 1-7 for each slice of the series.

As shown in Fig. 2, the parameter used in the weighted average in step 3 determines how much information of the MR image slice must be merged with the information of the PET image.

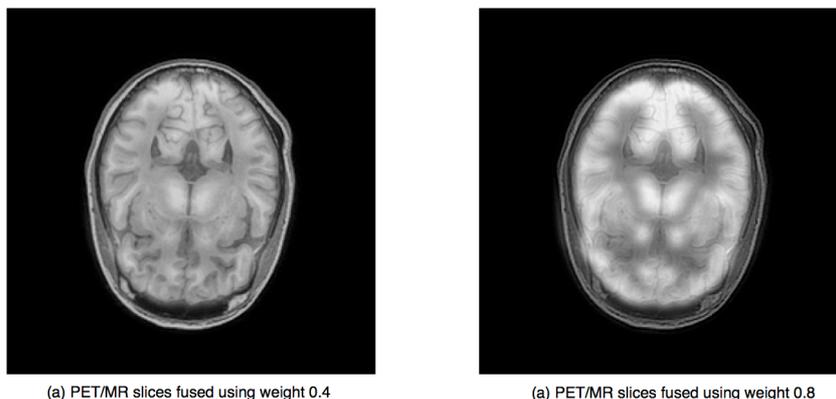


Figure 2. Effect of the weight coefficient during image fusion

4. The Tool for Multimodal Fusion of PET and MR images

We implemented the method described in Section 3 in a software tool using the MATLAB® environment. We also provided to the tool with a Graphical User Interface that guides the user through the steps of the method.

Figure 3 illustrates the interface to select the input images to merge. On the left-hand side the doctor has a preview of the selected image. He can also access the histogram of each image and perform some simple optimisation.

Once the images have been loaded into memory, the tool supports the user in the alignment phase necessary for the registration of the two exams output (see Figure 4). Furthermore, the number of slices of the two exams is matched by generating interpolated slides for the exams having a lower resolution. In Figure 3 you can note that the number of slices of the MR exam is 192, while the number of slices of the PET is 47. After the user selects the Align Slice Number functionality, the number of slices is 162 for both the exams. Finally, a multi-modal image registration is performed. The lower central section of the interface provides control to adjust the tool parameters and shows the values of

A Tool for Multimodal Fusion of Brain PET and MR TRE (Target Registration Error) average and variance. After the registration, the images can be fused in a single exam image (see Figure 5).

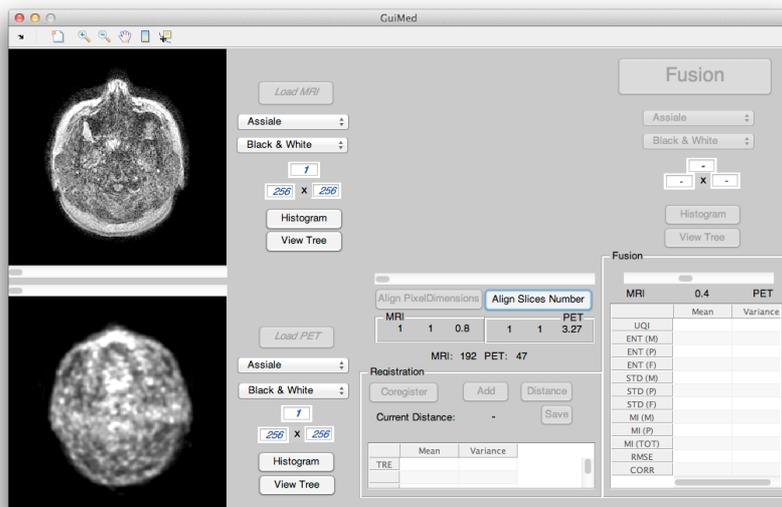


Figure 3. Screenshot of the GUI to select the input MR and PET images

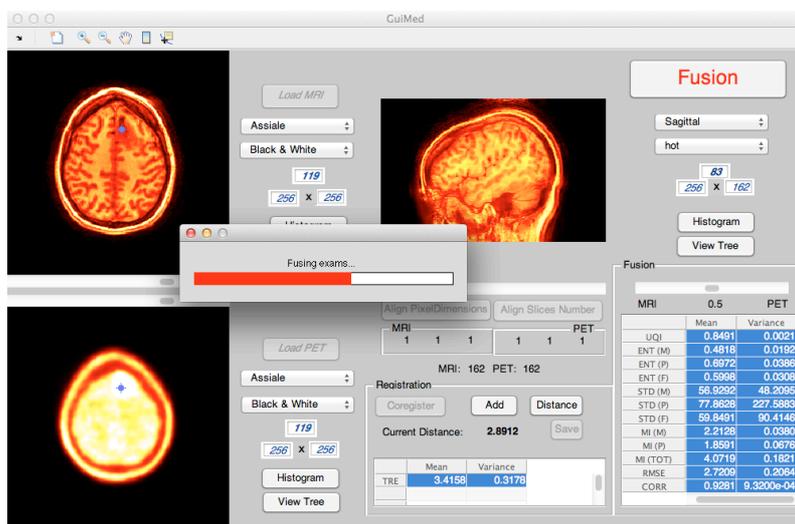


Figure 4. Screenshot of the GUI during the image fusion.

In Figure 5 the output of the fusion is shown. The user can access the histogram of the fused images and export it. He can also change the weight value discussed in the previous Section to determine how much information of

the MR image slice must be merged with the information of the PET image (a weight of 0.3 is selected in the example). Finally, the user the image quality metrics of the merged images are presented in the bottom-right corner.

In order to assess the accuracy of the registration, two experienced neuroradiologists visually inspected the contours continuity and verified that the tool correctly registered the images. On the other hands, the fusion accuracy has been assessed by using some mathematical indices, such as universal image quality index [14] root mean square error, and mutual information metrics [2].

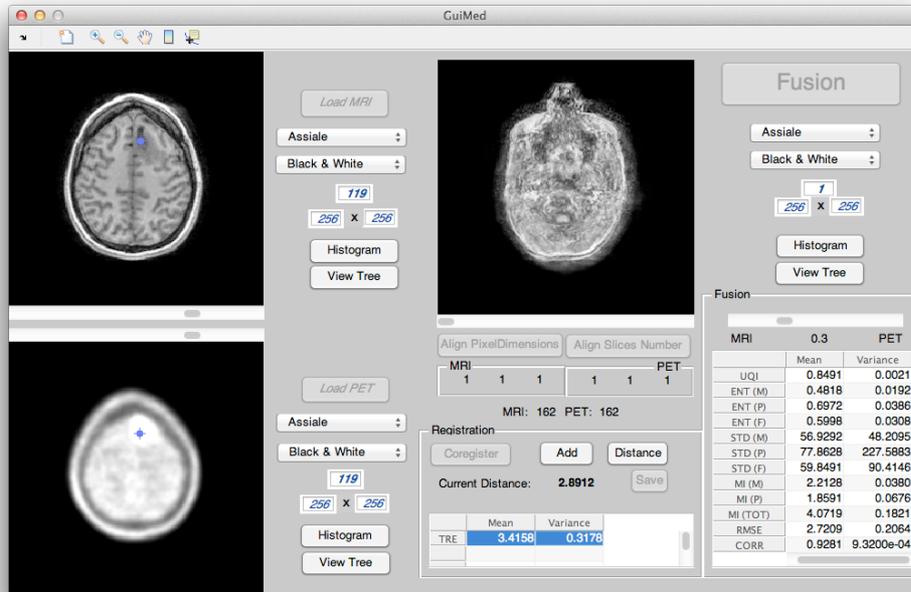


Figure 5. Screenshot of the GUI after the fusion

5. Conclusion and Future works

We have presented a fusion technique for medical images from different types of exams, i.e., PET and MRI. Indeed, these two types of exams provide the specialist with different kind of information: the MRI shows structural information while the PET exams contains functional information. Our goal is to integrate both morphologic and functional information in a single set of images generated starting from the brain MR and PET images to support the doctor during the diagnostic and therapeutic decision. We also presented a software tool that implements the proposed method enabling the user to interact with the image merge tool to achieve the best results according to his medical investigation.

As a future work we aim at integrating other types of images, such as X-Ray or CT. An evaluation of the effectiveness of the method and tool is also part of our future agenda.

A natural evolution of the presented research is the inclusion of a semi-automatic procedure for the extraction of pathological area. This would represent an important support for the medical personnel who could focus his attention on pathological areas also having anatomical and structural patient's references.

Acknowledgments

We thank the IRCCS INM Neuromed – Pozzilli (IS), Italy, for providing the sample MR and PET images. A special thanks is due to the Biomedical Engineering department and to Dr. Izzi and Dr. Pavone in particular. Finally, we are grateful to Dr. Sebastiano for his precious suggestions and support.

Bibliografia

- [1] Chen W. "Clinical applications of PET in brain tumors". *Journal Nuclear Medicine*, nr. 48, pp. 1468–1481, 2007.
- [2] Deshmukh M. and Bhosale U., "Image fusion and image quality assessment of fuse images". *International Journal of Image Processing*, vol. 4, no.5, pp. 484–508, 2010.
- [3] Heiss WD. "The potential of PET/MR for brain imaging". *Journal Nuclear Medicine*, no. 36 (1), pp. S105–S112.
- [4] Herholz K. "PET studies in dementia". *Annals of Nuclear Medicine*, vol. 17, no.2, pp. 79–89, 2003.
- [5] Hiwarkar T.A., Iyer S., "Soft computing methodologies in bioinformatics and its advance towards biological DNA", *International Journal of Engineering and Computer Science* vol. 2, no. 5, ISSN:2319-7242, pp.1506-1511, 2013.
- [6] Hwang S.C., Chen J.H., Lee J.J.S. and Yeh S.H., "The registration and presentation of MR and PET fusion imaging," *Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Bridging Disciplines for Biomedicine*, vol.5, pp.2305-2306, 1996
- [7] Jacobs A.H., Kracht L.W., Gossmann A., Ruger M.A., Thomas A.V., Thiel A., "Imaging in neurooncology", *NeuroRx* 2, pp. 333–347, 2005.
- [8] Koeppe M.J. and Woermann F.G., "Imaging structure and function in refractory focal epilepsy". *Lancet Neurology*, nr. 4, pp. 42–53, 2005.
- [9] Kuhl FP Giardina CR. "Elliptic Fourier Features of a Closed Contour". In: *Computer graphics and image processing* 18 (1982), pp. 236–258.
- [10] Maintz J.B.A., Viergever M.A., A Survey of Medical Image Registration, *Medical Image Analysis*, 1998, 2(1):1– 36.
- [11] Palubinskas G. and Reinartz P. "Multi-resolution, multi-sensor image fusion: general fusion framework", in Stilla U., Gamba P., Juergens C., Maktav D. (Eds)

- JURSE 2011, in Proc. of Joint Urban Remote Sensing Event, Munich, Germany, pp. 313-316, 2011
- [12] Seok-Ki K., Hyuck J.V., Sang-Yoon P., Ho-Young L., Sang-Soo S., Chong W.Y., Dae C.J., Sokbom K., Kyung-Sik C., "Additional value of MR/PET fusion compared with PET/CT in the detection of lymph node metastases in cervical cancer patients". *European Journal of cancer*, no. 45, pp. 2013–2109, 2009
- [13] Thiel A., Habedank B., Herholz K., Kessler J., Winhuisen L., Haupt W.F., Heiss W., "From the left to the right: How the brain compensates progressive loss of language function", *Brain and Language*, vol. 98, nr. 1, pp. 57-65, 2006, ISSN 0093-934X.
- [14] Wang Z. and Bovik, A.C., "A universal image quality index," *IEEE Signal Processing Letters*, vol.9, no.3, pp.81-84, 2002.
- [15] Wong A. and Bishop W., Efficient Least Squares Fusion of MRI and CT Images Using A Phase Congruency Model, *Pattern Recognition* vol. 29, pp.173–180, 2009.
- [16] Woo S.K., Kim K.M., Lee T.S., Jung J.H., Kim J.G., Kim J.S., Choi T.H., An G., and Cheon G.J., "Registration Method for the Detection of Tumors in Lung and Liver Using Multimodal Small Animal Imaging," *IEEE Transactions on Nuclear Science*, vol.56, no.3, pp.1454,1458, 2009.
- [17] Zou M.; Liu Y., "Multi-Sensor Image Fusion: Difficulties and Key Techniques," Proc. of *2nd International Congress on Image and Signal Processing, CISP '09*, pp.1,5, 2009

The multidimensional value of transparency in healthcare organizations computerization

Aurelio Tommasetti, Giuseppe Festa

Università degli Studi di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
Via Giovanni Paolo II – 132, 84084 Fisciano (SA)
tommasetti@unisa.it, g festa@unisa.it

Questo lavoro è frutto di un impegno comune. Tuttavia, i paragrafi 1 e 6 sono da attribuirsi ad Aurelio Tommasetti, mentre i paragrafi 2, 3, 4 e 5 sono da attribuirsi a Giuseppe Festa.

Abstract. *Healthcare informatics and the connected computerization represent a fundamental driving force for the evolution and development of healthcare organizations. Health information systems (in the widest sense, as proposed by AICA with the professional certification named ECDL Health) will become more and more important in the years to come, thanks to the benefits of business informatics in the health sector. In particular, health computerization may help to pursue physiologically a 'historic' instrumental value about health, that is transparency, not only as regards performances and their associated responsibilities, but also as regards organizational processes and related information flows. This study, as a conceptual paper, wants to develop a theoretical framework about the multidimensional value of health transparency, analyzing health information systems in the perspective of the transparency resulting by business computerization; in addition, by a specific content analysis, we try to contribute to a more robust arrangement of the ECDL Health Syllabus.*

Keywords: information systems, healthcare organizations, transparency.

1. Introduzione.

Gli attuali scenari socioeconomici stanno attraversando con intensità sempre maggiore fasi continue e discontinue di crescente complessità. Una delle concause più rilevanti di queste evoluzioni è certamente l'informatizzazione, che rappresenta un fenomeno da inquadrare non soltanto nei termini del trattamento hardware e software del dato (nel rispetto più rigoroso del tradizionale significato dell'informatica), ma anche (e, a parere di chi scrive, soprattutto) nei termini del cambiamento organizzativo che ne deriva, inevitabilmente caratterizzato da un maggiore ricorso alla riorganizzazione, alla re-ingegnerizzazione, alla proceduralizzazione e, in fin dei conti, alla standardizzazione (nel processo e/o nel risultato), facendo ampio uso, almeno nella progettazione della struttura organizzativa digitale, del criterio della discontinuità [De Witt, 2001; Tonti, 2002; Bracchi et al., 2010].

Infatti, in termini prima di tutto concettuali, progettare e sviluppare un sistema informativo (in generale) e informatico (in particolare) costringe il professionista contestualmente coinvolto a uno sforzo di comprensione, rappresentazione e ingegnerizzazione dello specifico flusso informativo e del relativo processo organizzativo, applicando necessariamente e continuamente approcci riduzionistici (frammentazione) e sistemici (integrazione). Soltanto a valle, pertanto, si arriverà (eventualmente) all'incapsulamento di flussi e processi all'interno di una soluzione informatica, sviluppandosi nel contempo una profonda conoscenza della fattispecie aziendale in oggetto.

Questo sforzo di comprensione (inevitabilmente assunto nella prospettiva dell'operatore, ma costantemente ispirato all'utente) potrebbe essere propedeutico anche alla generazione di trasparenza, che, in primo luogo, è possibile intendere, da un punto di vista insieme lessicale e semantico, come l'assenza di asimmetria informativa, ossia la non esistenza di fattispecie e/o soggetti in possesso di un diverso allineamento informativo. Tuttavia, una volta acquisito il contributo derivante dalla maggiore disponibilità d'informazioni, è possibile pensare anche a un ampliamento del concetto, perché altrimenti la trasparenza sarebbe "soltanto" simmetria informativa: nel concetto di trasparenza, invece, emergono anche altri valori, quali quelli dell'equi-finalità sociale (o almeno collettiva), dell'etica (o almeno della tensione all'etica) e così via, fino a inglobare le tradizionali ambizioni di efficacia, efficienza ed economicità, tipiche del contesto oggetto del presente studio, ossia la sanità, in un omnicomprensivo concetto di "sanità civile" (si specifica fin d'ora, per chiarezza, che la trasparenza di cui si discute in questo lavoro non si riferisce ai dati personali e sensibili del singolo cittadino / utente / paziente, per i quali è obbligatorio, doveroso e professionale il rispetto della privacy: cfr. d.lgs. 196/2003).

Inoltre, nell'ambito specifico dell'*information and communication technology*, "trasparente", almeno in lingua italiana, è anche quella fattispecie informatica, progettata e/o implementata a un qualsiasi livello, in cui un operatore (eventualmente anche utente) non sia tenuto a sapere come il sistema funzioni realmente per potersene servire nel suo processo d'uso. In questo caso, com'è agevole notare, la trasparenza si avvicina moltissimo all'usabilità, ossia alla minima distanza possibile tra *user model* e *design model*.

Usando quindi un parallelismo, per il beneficiario finale, vale a dire in senso lato il cittadino, il sistema sanitario è tanto più "usabile" quanto più è "trasparente", sia in termini operativi sia in termini sociali. Un importante contributo in tal senso è conferito proprio dall'informatica sanitaria o, meglio, dall'informatizzazione sanitaria, intesa come la

propensione mentale e culturale al governo e alla gestione dei dati, delle informazioni e delle conoscenze sanitarie in termini informatici, con relativi effetti su processi organizzativi e flussi informativi [Teti e Festa, 2009].

Questo studio, pertanto, si propone come un *conceptual paper*, finalizzato all'elaborazione di un *framework* teorico che consenta un'analisi integrata del valore multidimensionale della trasparenza sanitaria derivante dall'informatizzazione dell'organizzazione sanitaria, proponendo quindi proprio la trasparenza come un principio-guida per la progettazione del sistema informativo sanitario. Sulla scorta di tale rappresentazione concettuale, inoltre, si procede successivamente a un'indagine avente come oggetto il rinvenimento delle possibili dimensioni della trasparenza sanitaria all'interno del Syllabus dell'ECDL Health (la certificazione professionale, riconosciuta a livello internazionale e curata in Italia dall'AICA, destinata all'informatica sanitaria), allo scopo di contribuire a un possibile rafforzamento del Syllabus nelle parti riservate alla trasparenza informatica in sanità, soprattutto alla luce delle recenti normative italiane in materia.

2. Il valore multidimensionale della trasparenza in sanità.

L'economia neoclassica o marginalista, ossia quella che più ha influenzato le attuali interpretazioni delle fattispecie economiche, identifica il "valore" con l'apprezzamento di un'utilità (di tempo, luogo, possesso e forma), che, nel caso specifico della sanità, diventa "valore d'uso soggettivo" (questa la precisa locuzione neoclassica) delle capacità materiali, immateriali, finanziarie, professionali e tecnologiche (mutuando la tradizionale funzione di produzione) delle strutture sanitarie di soddisfare bisogni / desideri in termini fisici, psichici, sociali e ambientali (conformemente alla definizione di "salute" proposta dall'Organizzazione Mondiale della Sanità), considerando in tal senso il valore della sanità da un punto di vista "tecnico" [Tramarin, 2002].

A dire il vero, anche nell'economia classica si rinviene il principio di utilità, che tuttavia non è così centrale come in quella neoclassica, che enfatizza in particolare il concetto di utilità marginale quale perno della dinamica economica [Da Empoli, 2012]. Inoltre, nel termine "professionale", nel comparto specifico di studio e nella prospettiva di chi scrive, è compreso il contributo della comunicazione sanitaria, personale e pubblica, ormai sempre di più considerata, a pieno titolo e giustamente, una componente essenziale nell'erogazione del servizio e della prestazione sanitaria.

A legittimare il valore, inoltre, da un conseguente punto di vista "commerciale", è l'attitudine del bene ad avere "mercato" o, se si vuole, a rivestire interesse (e quindi valore) in quanto oggetto di scambio. In tal senso, è possibile che esisteranno beni diversi che avranno valori diversi (una cura oncologica ha per tantissimi aspetti un valore diverso rispetto a una cura estetica) oppure beni che avranno valore per alcuni e non per altri: per tali beni, in altre parole, potrebbe non esserci lo stesso interesse in capo ai medesimi soggetti di cui sopra (una cura oncologica è verticalmente fuori contesto e quindi senza valore per un soggetto bisognoso / desideroso di una cura estetica) o ancora beni simili che avranno valori contestualmente diversi (a seconda che siano richiesti da clienti con caratteristiche diverse nel tempo e/o nello spazio).

Inoltre, è pacifico come l'apprezzamento economico non si limiti alla sola valutazione monetaria, qualificabile nei termini del marketing mix come "prezzo", che costituisce la misura commerciale dello scambio, ma finisca inevitabilmente per comprendere anche gli ulteriori, rispetto alla privazione di denaro, elementi richiesti dalla funzione di scambio (il tempo necessario a raccogliere informazioni sulle opzioni di acquisto, la specificità contestuale dell'acquisto e del consumo, lo stress nella fase di decisione e di eventuale dissonanza cognitiva, ecc.). È evidente, di conseguenza, come la medesima utilità, per due acquirenti / consumatori diversi, possa avere lo stesso "prezzo", ma un diverso "valore", perché diversamente apprezzata in termini discrezionali e quindi soggettivi e quindi psicologici.

Infine, sempre in una prospettiva economico-aziendale, il valore rileva anche in termini etici e sociali: si pensa, in particolare, alla "catena mezzi-fini" e ai relativi valori strumentali e terminali. In tale modello (Reynolds e Gutman) è "strumentale" il valore che si ritiene necessario e/o sufficiente per il perseguimento di un'ulteriore finalità, mentre è "terminale" il valore che costituisce una specifica finalità [Lambin, 2008]. In sanità, per esempio, un valore strumentale è la competenza dell'organizzazione sanitaria (per il cittadino, in altre parole, è *eticamente* importante che un medico sia competente, perché così potrà erogare un servizio appropriato), mentre il valore terminale (il valore terminale per eccellenza) è la salute del paziente nello specifico e della collettività più in generale (che è *socialmente* importante in tutti i sistemi sanitari, ma soprattutto in quelli solidaristici, come quello italiano).

La "trasparenza" può essere attribuita all'una o all'altra categoria (strumentale o terminale). Nel contesto della Pubblica Amministrazione, per esempio, l'assenza della corruzione o anche del solo sospetto di corruzione è un fondamentale attributo del multidimensionale interesse pubblico (rilevando quindi la trasparenza come valore terminale a forte connotazione "sociale"), ma è chiaro che, soprattutto in una prospettiva organizzativa e particolarmente nel comparto dei servizi sanitari, rileva anche nella sua essenza di valore strumentale, ritenuto utile alla migliore gestione (più efficace, più efficiente, più economica e complessivamente più "legittima" e quindi più "etica") delle risorse attualmente e potenzialmente disponibili, in maniera trasversale e fluida.

Per esempio, considerando per semplicità di ragionamento un'unica utilità, un cittadino della Regione Campania potrebbe non avere un diretto interesse alla trasparenza della lista di attesa di una struttura sanitaria della Regione Lombardia, ma dovrebbe giustamente avere un interesse indiretto ("etico" e "sociale" insieme) a che ogni cittadino italiano goda di un simile beneficio, che è in realtà un diritto sancito dalla Costituzione (art. 3 come diritto all'eguaglianza e art. 32 come diritto alla salute). In realtà, l'integrazione del Servizio Sanitario Nazionale consente (e per certi versi quasi incentiva, in termini di competitività) la mobilità regionale, creando in capo al cittadino della

Regione Campania una pretesa attuale, in questo caso associabile a un interesse individuale e non più generale, alla gestione trasparente della lista di attesa anche nella Regione Lombardia.

L'esempio è puramente teorico e non presuppone ovviamente alcun pregiudizio, se non l'amara constatazione di un tasso di mobilità, nell'incrocio tra le due Regioni, positivo a Nord e negativo a Sud. Come si avrà modo di evidenziare nel corso del lavoro, a irrobustire ulteriormente la pretesa di trasparenza in sanità è proprio la fisiologica integrazione del Servizio Sanitario Nazionale, a livello strutturale (ambulatorio, presidio, azienda ospedaliera, policlinico universitario, istituto di ricovero e cura a carattere scientifico e così via), settoriale (nel senso di primo, secondo e terzo settore) e regionale (ossia tra regione e regione all'interno del territorio italiano, ma anche al di fuori dei confini nazionali, seguendo le procedure richieste dall'ASL di appartenenza).

Proprio nella prospettiva del cittadino / utente / paziente, infatti, si ricorda che, in termini normativi, «... la trasparenza è intesa come accessibilità totale, anche attraverso lo strumento della pubblicazione sui siti istituzionali delle amministrazioni pubbliche, delle informazioni concernenti ogni aspetto dell'organizzazione, degli indicatori relativi agli andamenti gestionali e all'utilizzo delle risorse per il perseguimento delle funzioni istituzionali, dei risultati dell'attività di misurazione e valutazione svolta dagli organi competenti, allo scopo di favorire forme diffuse di controllo del rispetto dei principi di buon andamento e imparzialità. Essa costituisce livello essenziale delle prestazioni erogate dalle amministrazioni pubbliche ai sensi dell'articolo 117, secondo comma, lettera m), della Costituzione» (art. 11 – “Trasparenza”, comma 1, d.lgs. 150/2009, “Attuazione della legge 4 marzo 2009, n. 15, in materia di ottimizzazione della produttività del lavoro pubblico e di efficienza e trasparenza delle pubbliche amministrazioni”).

Per certi versi, più in generale, è possibile “immaginare” un sistema d'informazione che si muova lungo un continuum con due estremi: la presenza di regolamentazione pubblica e l'assenza (mai totale, soprattutto in sanità) di regolamentazione pubblica. Nella sanità italiana, come si accennava in precedenza a proposito di livello “settoriale”, agiscono anche le strutture private, *for profit* e *non profit*, con diversi livelli di “pubblicità”: infatti, le strutture sanitarie private accreditate presso il Servizio Sanitario Nazionale hanno un funzionamento istituzionale praticamente identico a quelle del settore pubblico; nei casi non accreditati, invece, si fa riferimento in ogni caso a un'autorizzazione a svolgere un'attività clinico / sanitaria, che quindi soggiace a una certa “pubblicità” del funzionamento, ma è altrettanto chiaro che, essendo privata della “convenzione” in termini economici e venendo quindi meno il contrasto d'interessi a favore della collettività / società, l'asimmetria informativa potrebbe avere più spazio, a meno ovviamente delle positive pressioni derivanti dalla responsabilità, dalla deontologia, dall'etica, ecc.

In tal senso, sempre in termini normativi, si segnala che il d.lgs. 33/2013 (“Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni”) è «... predisposto in attuazione dei principi e criteri di delega previsti dall'art. 1, comma 35, della legge 6 novembre 2012, n. 190, recante “Disposizioni per la prevenzione e la repressione della corruzione e dell'illegalità nella pubblica amministrazione” [e] riordina, in un unico corpo normativo, le numerose disposizioni legislative in materia di obblighi di informazione, trasparenza e pubblicità da parte delle pubbliche amministrazioni. [...] Tra le novità più rilevanti [...] si segnalano: l'istituzione del diritto dell'accesso civico, l'obbligo di predisporre e pubblicare il Piano triennale per la trasparenza, l'obbligo di nominare il Responsabile della trasparenza in ogni amministrazione, la rivisitazione della disciplina in materia di trasparenza sullo stato patrimoniale di politici e amministratori pubblici e sulle loro nomine, l'obbligo di definire nella home page del sito istituzionale di ciascun Ente un'apposita sezione denominata “Amministrazione trasparente”» (fonte: AA.VV., Disposizioni in materia di trasparenza nelle pubbliche amministrazioni – Nota informativa sul d.lgs. n. 33/2013, ANCI, aprile 2013). L'art. 41 del decreto, in particolare, è riservato esclusivamente alla sanità (cfr. estratto dalla Gazzetta Ufficiale).

Trasparenza del servizio sanitario nazionale.

1. *Le amministrazioni e gli enti del servizio sanitario nazionale, dei servizi sanitari regionali, ivi comprese le aziende sanitarie territoriali ed ospedaliere, le agenzie e gli altri enti ed organismi pubblici che svolgono attività di programmazione e fornitura dei servizi sanitari, sono tenute all'adempimento di tutti gli obblighi di pubblicazione previsti dalla normativa vigente.*

2. *Le aziende sanitarie ed ospedaliere pubblicano tutte le informazioni e i dati concernenti le procedure di conferimento degli incarichi di direttore generale, direttore sanitario e direttore amministrativo, nonché degli incarichi di responsabile di dipartimento e di strutture semplici e complesse, ivi compresi i bandi e gli avvisi di selezione, lo svolgimento delle relative procedure, gli atti di conferimento.*

3. *Alla dirigenza sanitaria di cui al comma 2, fatta eccezione per i responsabili di strutture semplici, si applicano gli obblighi di pubblicazione di cui all'articolo 15. Per attività professionali, ai sensi del comma 1, lettera c) dell'articolo 15, si intendono anche le prestazioni professionali svolte in regime intramurario.*

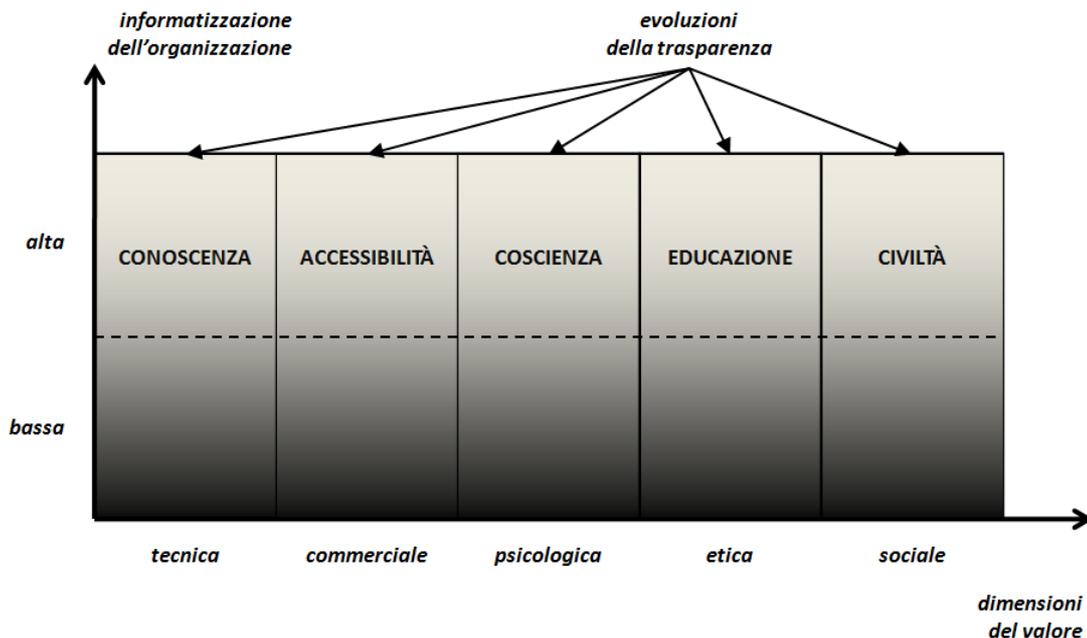
4. *È pubblicato e annualmente aggiornato l'elenco delle strutture sanitarie private accreditate. Sono altresì pubblicati gli accordi con esse intercorsi.*

5. *Le regioni includono il rispetto di obblighi di pubblicità previsti dalla normativa vigente fra i requisiti necessari all'accreditamento delle strutture sanitarie.*

6. *Gli enti, le aziende e le strutture pubbliche e private che erogano prestazioni per conto del servizio sanitario sono tenuti ad indicare nel proprio sito, in una apposita sezione denominata «Liste di attesa», i tempi di attesa previsti e i tempi medi effettivi di attesa per ciascuna tipologia di prestazione erogata.*

In generale, pertanto, sembra possibile articolare il concetto di “valore” in almeno cinque componenti fondamentali, di natura, come si è ragionato, tecnica, commerciale, psicologica, etica e sociale. Nella prospettiva del presente lavoro, inoltre, si vuole indagare nello specifico il valore della trasparenza in sanità, in particolare quando accompagnata dall'informatizzazione, contestualmente più o meno spinta (“bassa” o “alta”), dell'organizzazione sanitaria: da questa integrazione deriva un *framework* teorico, in altre parole una sorta di mappa concettuale, proposta nella Fig. 1, che sintetizza le modalità evolutive (che potrebbero essere considerate anche ulteriori “valori”) con cui si realizza maggiore o minore trasparenza nell'organizzazione sanitaria, grazie alla relativa informatizzazione (in

presenza di una bassa informatizzazione, infatti, dovrebbe esserci, come si rappresenta anche graficamente, maggiore opacità).



(Fig. 1) Evoluzioni / valori della trasparenza nell'informatizzazione sanitaria

La trasparenza, oltre che esigenza organizzativa, diventa in realtà una vera e propria aspettativa della società moderna, sempre più informata (ormai classico è l'esempio del nuovo edificio, sede a New York del New York Times, progettato dall'architetto Renzo Piano con largo uso di vetro trasparente, a simboleggiare l'assoluta onestà della testata rispetto ai propri lettori e al resto della società), dinamica, consapevole, vigile e bisognosa / desiderosa di qualità, investendo sempre più massicciamente, in termini economici e culturali, su comportamenti corretti (*etic*) e sussidiari (*sociali*). Puntando a valori ancora più "civili", una maggiore trasparenza dovrebbe comportare anche minore burocrazia negativa e, quindi, una potenziale maggiore produttività, ma dovrebbe soprattutto spingere le organizzazioni che se ne fanno promotrici a funzionare in maniera più sincera, aperta e sana.

Senza etica, peraltro, in una visione concretamente moderna degli scenari d'impresa non sembra nemmeno possibile ambire alla qualità, che, in assenza di etica "genericamente" d'impresa, si limiterebbe a un mero risultato tecnico, non realizzando le ormai normali ambizioni di sostenibilità imprenditoriale. Inoltre, è chiaro che la trasparenza trova un nemico non nella burocrazia (che, meccanica o professionale, rimane una delle configurazioni organizzative di base), ma nell'eccesso di burocrazia, ossia quando meccanismi inutilmente formali e dannosamente ridondanti appesantiscono l'efficacia e l'efficienza del funzionamento organizzativo.

3. Alcune riflessioni sulla valutazione multidimensionale della trasparenza in sanità.

Dati, informazioni e conoscenze costituiscono al giorno d'oggi veri e propri *asset*; anzi, si tratta probabilmente degli *asset* più rilevanti, che meritano un trattamento di assoluto rispetto, a cominciare, peraltro, dalla relativa protezione, sia in una fase interna (produzione e conservazione) sia in una fase esterna (distribuzione e fruizione). Essendo entità utili, possono evidentemente essere considerate beni, ossia, in termini più economico / aziendali, risorse, che normalmente sono scarse, necessitando quindi di un governo e di una gestione; in particolare, si tratta di risorse intangibili, che hanno un proprio valore (autonomo e di scambio), in quanto possono essere, come si anticipava in precedenza, prodotte, conservate, distribuite e fruite [Gambaro e Ricciardi, 2003].

È evidente che un soggetto che possieda una determinata informazione, ossia l'informazione utile in un determinato contesto, ha un vantaggio rispetto a un soggetto nelle stesse condizioni, ma privo dell'informazione in questione. Questa fattispecie, ossia l'asimmetria informativa, rappresenta in linea generale una situazione di ingiustizia economica, in molti contesti legittimamente e profittevolmente perseguibile [Schilling, 2009], ma, se collegata all'accesso al servizio sanitario, diventa intollerabile in generale e soprattutto in un sistema solidaristico come quello italiano, rigorosamente basato, almeno in teoria, sull'equità.

Non è un caso, infatti, che negli studi di economia generale la simmetria e l'asimmetria informativa costituiscano ipotesi capaci di proiettare scenari di studio completamente diversi. In verità, gli studi che si basano sulla perfetta simmetria informativa rappresentano evidenti artefatti della realtà, che tuttavia costituiscono delle basi fondamentali per l'elaborazione di modelli di funzionamento economico e comportamento sociale di indubbia rilevanza (si pensi, per esempio, agli studi sulla concorrenza perfetta, alle teorie di Modigliani-Miller, ecc.).

L'asimmetria informativa in sanità può manifestarsi a diversi livelli: quello più "sconco" si riferisce alla possibilità, per alcuni soggetti, di arrivare a godere di servizi o di accesso ai servizi secondo modalità più vantaggiose rispetto a un altro soggetto "concorrente" (è evidente, infatti, che prenotare per accedere a un esame radiografico

significa essere in concorrenza con un altro paziente / cittadino bisognoso dello stesso esame) o più in generale rispetto alla comunità. Altro nodo della questione è che nella sanità solidaristica non dovrebbero esserci meccanismi di vantaggio dell'uno o dell'altro (significativamente "privati"), ma di beneficio e benessere sociale: del resto, gli studi dell'economia sanitaria, anche se prevalentemente basati su meccanismi di negoziazione delle singole utilità (si pensi in particolare alla teoria dell'agenzia, con il medico nel ruolo di *agent* e il paziente in quello di *principal*), non possono trascurare la prospettiva dell'equità, soprattutto nei termini dell'efficienza allocativa e quindi nella prospettiva della programmazione sanitaria [Dirindin e Vineis, 2004].

Nella sanità, in particolare, il costo dell'informazione può generare una duplice degenerazione: l'azzardo morale (si pensi alla medicina difensiva) o la selezione avversa (pagano l'informazione soltanto i disagiati) [Ziliotti, 2001; Cappelli e Renzi, 2010]. In assenza di trasparenza, pertanto, sembrano valere le tradizionali riflessioni economiche in tema d'incertezza [Donia Sofio, 2001] sulla propensione – o no – al rischio (si pensi per esempio al triage) [Rumiati e Bonini, 2010].

La trasparenza, inoltre, potrebbe essere necessaria non soltanto per combattere le degenerazioni dell'asimmetria informativa, ma anche le negligenze relative all'informazione e alla comunicazione: molto spesso, infatti, i pazienti non conoscono il possibile spettro di prestazioni eseguibili da una determinata struttura, rinviando ad altri meccanismi di composizione della decisione l'approvvigionamento di queste delicatissime informazioni (medico di medicina generale, passaparola, mass media, ecc.). Può anche capitare, d'altra parte, che il soggetto meglio informato debba essere trasparente per poterne beneficiare (non implementandosi, in altre parole, un effetto "privato" dell'informazione), come anche che un soggetto meglio informato comunichi a molti per ottenere un filtro dalla conoscenza e/o dall'esperienza collettiva (si pensi al ruolo sempre più diffuso dei *social network* in materia di salute e sanità).

In termini economici, il contributo della trasparenza è analizzabile in primo luogo nella prospettiva dei due principali obiettivi di buon funzionamento di una qualsiasi organizzazione, vale a dire l'efficacia e l'efficienza, arricchite, nel comparto sanitario in generale e nel settore pubblico in particolare, dalla dimensione dell'economicità [Tommasetti e Cuccurullo, 2004; Marinò, 2008]. Anche soltanto per buon senso, infatti, sembra intuibile che un'organizzazione più trasparente (per legalità, deontologia, responsabilità, meritocrazia, competitività, ecc.) possa nutrire ambizioni di maggiore efficacia ed efficienza rispetto a un'altra più opaca; il riferimento all'economicità, inoltre, permette di allargare automaticamente la riflessione agli ulteriori concetti di universalità, equità, giustizia e così via, che sembrano francamente non implementabili in strutture e sistemi sanitari non trasparenti.

In termini più imprenditoriali, peraltro, la trasparenza diventa anche e soprattutto un fattore abilitante per la concorrenza, la flessibilità (sia tecnica sia produttiva), la focalizzazione sulla specializzazione e così via. In linea di massima, quindi, la trasparenza permette di alimentare nei mercati (sanitari) alcuni virtuosi conflitti d'interesse, consentendo anche una più profonda comprensione dei meccanismi di scelta del mix di prestazioni.

Si ricordi che nel comparto sanitario la concorrenza è (dovrebbe essere) sempre interna e non esterna, vale a dire impegnata a erogare prestazioni sempre migliori dopo aver assicurato un "livello essenziale di assistenza", auspicabilmente standard verso il benchmark e personalizzato verso l'assistito. L'opzione inevitabile della concorrenza interna comporta ovviamente che non ci possa (debba) essere concorrenza tra buona e cattiva sanità, ma soltanto tra livelli competitivi di qualità all'interno della buona sanità: la trasparenza, pertanto, diventa un passaggio obbligato in primo luogo per far emergere le fattispecie "buone" o "cattive" e, in secondo luogo, per poter classificare o almeno comparare le fattispecie "buone".

Nella prospettiva della concorrenza, sembra naturale che la trasparenza sia invocata o comunque indolore per l'operatore che si trovi in una posizione più competitiva (più performante, più etica, ecc.) ovvero per chi possieda una maggiore qualità verticale del proprio *output / outcome*. Questa maggiore *performance* in termini d'informazione e comunicazione (o almeno questa maggiore predisposizione / attitudine alla trasparenza) va a sua volta adeguatamente informata e comunicata, in una prospettiva di risonanza con gli *stakeholder* dell'ambiente socio-economico (interno ed esterno e, in quest'ultimo caso, micro e macro): in questo modo, inoltre, dovrebbe essere lecito anche aspettarsi un comportamento sia razionale sia emotivo che porti il mercato a un'attesa media in linea con l'offerente più trasparente, al cui standard (basso o alto che sia) dovrà allinearsi il resto dell'offerta.

Gli stessi "presidi" sanitari (ossia un qualsiasi strumento tecnico: dalla garza alla siringa, dal fonendoscopio all'elettrocardiografo, dall'apparecchiatura per la risonanza a quella per la tomografia assiale computerizzata e oltre) sono oggetti dotati di contenuto informativo: con il passar del tempo, diventano anche più precisi, analitici, discriminanti e così via grazie alla maggiore tecnologia incorporata, che è soprattutto conoscenza. In un certo senso, quindi, la tecnologia produce "migliore" informazione e l'health technology assessment (HTA, ossia la materia che si occupa della valutazione delle tecnologie sanitarie) finisce naturalmente per valutare non soltanto la tecnologia in sé, ma soprattutto la potenzialità dell'informazione producibile e del suo successivo impiego in un processo clinico [Arduini, 2010]. Pertanto, anche per la trasparenza, intesa come fattore capace di produrre maggiore efficacia, efficienza ed economicità in sanità, sembrano logicamente valere i criteri di valutazione economica dell'HTA, peraltro a loro volta sostanzialmente mutuati dalla farmaco-economia.

Nel prosieguo del nostro ragionamento, diventa quindi possibile anche tentare di determinare alcuni criteri per l'apprezzamento del valore (multidimensionale) della trasparenza sanitaria, in ragione della base di rappresentazione. Infatti, il valore individuale della trasparenza è in primo luogo ascrivibile all'assenza di asimmetria informativa: in termini economici, il criterio di valutazione di questa fattispecie dovrebbe coincidere con l'eliminazione del costo collettivo / sociale pari al *valore attuale netto allargato* che potrebbe acquisire il soggetto meglio informato; al singolo individuo, in particolare, si ritiene di poter ascrivere (almeno) le prime tre dimensioni rappresentate nel *framework*, ossia tecnica, commerciale e psicologica.

Un valore "allargato" potrebbe essere conseguito anche all'interno della stessa organizzazione sanitaria, perseguendo e conseguendo la semplificazione e la velocizzazione (a parità ovviamente di qualità e sicurezza) di diverse operazioni (si pensi alle prenotazioni interne rivolte dal reparto all'ambulatorio, al laboratorio, alla radiologia, alla sala operatoria, ecc.), ma anche una maggiore omogeneità (e quindi probabile maggiore efficacia, efficienza, economicità) in termini di procedure organizzative, sistemi di qualità, percorsi di accreditamento, ecc.

Nell'attuale società, caratterizzata da produzioni sempre più immateriali, il costo di diffusione (nei termini del supporto che contenga l'informazione) è inferiore al costo di produzione (si pensi alla trasmissione d'informazioni tramite internet): questa situazione è (o almeno dovrebbe essere) nota a entrambe le controparti, parlando genericamente di offerente (individuale) e acquirente (individuale / collettivo). In caso di asimmetria, invece, è perfettamente possibile che chi non possiede l'informazione (o non ne sappia fare il giusto uso, permettendoci di qualificare un'asimmetria di "teoria" e una di "pratica") sia destinato a subire un danno economico, pur essendo bassissimo l'eventuale costo della condivisione (il danno, d'altra parte, potrebbe derivare dal fatto che l'operatore in oggetto possiede l'informazione, sa come usarla, ma non ha i mezzi necessari – fisici, finanziari, ecc. – per usarla); del resto, l'informazione che diventa pubblica genera automaticamente trasparenza e perde immediatamente di valore per il singolo individuo (offerente), realizzandosi nello stesso tempo, pur se potenzialmente, un valore infinitamente maggiore per la società.

Infatti, chi ha prodotto una specifica informazione, sempre che sia valida e affidabile, molto probabilmente, anche se non certamente, ha sostenuto il migliore tra i costi possibili (dato che l'informazione ormai c'è, non ha senso riprodurla più economicamente, ma al massimo raffinarla per produrre altra informazione), spostando a questo punto la prospettiva di analisi dall'economia dell'informazione all'economia della conoscenza, in termini organizzativi e distributivi (ossia prevalentemente di coordinamento). Come in un qualsiasi altro contesto economico, dato che l'informazione ha un valore, avremo produttori, intermediari e utenti (che potrebbero anche coincidere): se l'utente volesse recuperare singolarmente l'informazione x, dovrebbe sostenere un certo costo: $C_x \times N$ sarebbe quindi il costo totale per tutti gli N utenti che volessero approvvigionarsi singolarmente dell'informazione x (per esempio, si pensi ai costi diretti e indiretti che sopportano singoli utenti, non tra loro comunicanti, per recarsi fisicamente al Centro Unificato Prenotazioni per avere cognizione della lista di attesa di un reparto di radiologia).

In realtà, in presenza di simmetria informativa, una volta realizzata l'infrastruttura di comunicazione, anche grazie a intermediari (con un certo costo di struttura da ammortizzare sul livello di uso), basterebbe produrre soltanto C_x , perché i costi di distribuzione dell'informazione verrebbero a essere a un certo punto quasi nulli (o in ogni caso asintoticamente prossimi allo zero), mentre in una struttura poliarchica, come a tutti gli effetti sono le organizzazioni che producono o consumano servizi e prestazioni in sanità, il numero di connessioni necessarie per coprire l'intero fabbisogno di comunicazione è pari a $N(N-1)$ in caso di flussi informativi unidirezionali ovvero $N(N-1) / 2$ in caso di flussi informativi bidirezionali. Si tratta, con evidenza, di un valore genericamente espresso in termini quantitativi (unità), che finirebbe per ricevere ulteriori contributi dai collegati aspetti sociali.

Inoltre, se cresce la complessità del fabbisogno sanitario, crescerà ovviamente anche la complessità dell'informazione clinica, ma non necessariamente la complessità dell'infrastruttura di comunicazione (si pensi ai dati sanitari, semplici o complessi, che viaggiano sulla stessa intranet). È ben evidente, pertanto, quanti risparmi interni ed esterni si potrebbero ricavare dalla trasparenza informativa, con conseguente aumento delle risorse disponibili per supportare la produzione di servizi sanitari a maggiore valore aggiunto.

Successivamente, è possibile arrivare a un valore "etico" della trasparenza, ascrivibile all'esigenza di godere di un funzionamento costantemente "normale" (legale, corretto, attento, ecc.) di un qualsivoglia servizio a favore di un individuo e/o della collettività: in termini economici, il criterio di valutazione di questa fattispecie dovrebbero essere i meccanismi dell'analisi decisionale, ossia *analisi costi / efficacia*, *analisi costi / benefici*, *analisi costi / utilità*. Grazie alla trasparenza, infatti, diventa possibile anche esaudire, chissà se completamente, l'ormai pressante richiesta di diminuzione della corruzione, del clientelismo, del favoritismo, del "nero" che allignano proprio dove non c'è trasparenza (e, quindi, al di là della correttezza della tecnica di misurazione, tale circostanza permette di incrementare anche il valore economico in generale per lo Stato e la società).

Infine, è possibile arrivare a un valore "sociale" della trasparenza, ascrivibile allo slancio di solidarietà verso un qualsiasi altro soggetto appartenente alla società, peraltro ignoto nella quasi totalità dei casi: in termini economici, il criterio di valutazione di questa fattispecie potrebbe essere il *value for money*. Il valore sociale della trasparenza, peraltro, è causa / ragione di ulteriori valori, a loro volta identificabili, sempre nella prospettiva della "catena mezzi-fini", come strumentali: si pensi all'importanza della parità di trattamento per gli utenti / pazienti oppure al contributo derivante dalla disponibilità dell'evidenza della pratica clinica per gli operatori sanitari.

La trasparenza, d'altro canto, è potenzialmente in grado di generare non soltanto vantaggi, ma anche svantaggi, almeno di termini di situazioni rischiose, come del resto qualsiasi fattispecie socio-economica: infatti, in termini sociali la trasparenza potrebbe perversamente portare anche all'insensibilità verso la singola situazione di disagio (per esempio, un rispetto impenetrabile della programmazione delle liste di attesa), mentre in termini economici si pensi soprattutto alla rigidità burocratica che dalla stessa trasparenza potrebbe derivare (si pensi a un ruolo eccessivo del concetto di "procedura"). È evidente che si tratta di palesi degenerazioni, ma si tratta purtroppo di effetti negativi assolutamente possibili in ambienti e contesti a elevata complessità di funzionamento, soprattutto da un punto di vista umano.

Altra ipotesi di rischio, al giorno d'oggi nient'affatto trascurabile, è anche quella dell'eccesso d'informazione derivante dalla trasparenza, che rappresenta una delle "sciagure" dell'attuale società dell'informazione. Nella sanità, in particolare, dall'eccesso dell'informazione è possibile che derivi anche una potenziale litigiosità giudiziaria o meglio un suo aumento (dato che si tratta di un fenomeno già notevole nel campo della salute).

4. La progettazione "trasparente" del sistema informativo sanitario.

Nella prospettiva del presente studio, una delle principali finalità perseguibili con un approccio multidimensionale alla trasparenza organizzativa (oltre al rispetto della normativa e all'eticità della scelta) è la maggiore propensione della struttura all'efficacia, all'efficienza e all'economicità, che dovrebbero correttamente significare, in una pro-

spettiva di eccellenza, produrre di più e meglio, con minori e quindi più sostenibili risorse. A questa ambizione economica, in un comparto complesso come quello sanitario, si affianca fisiologicamente anche e soprattutto l'apprezzamento sociale di un simile approccio, orientato in tal senso alla produzione e all'erogazione di salute [Bellucci e Cardoni, 2008].

Pertanto, diventa conseguentemente interessante che un sistema informativo aziendale, soprattutto per aziende particolari come quelle sanitarie, possa essere pensato anche, se non soprattutto, in funzione della trasparenza delle informazioni che tratterà, non soltanto come trasparenza verso l'esterno, ma anche come trasparenza verso l'interno. Finora, infatti, la trasparenza dell'organizzazione, soprattutto nel settore pubblico, è sempre stata vista come una sorta di "antidoto" al malfunzionamento (che nel caso della sanità diventa nello specifico "malasanità", non soltanto di "risultati", ma anche di "mezzi"), ma si è già evidenziato come sussistano importanti vantaggi anche per la stessa economicità organizzativa, in una sorta di esternalità positiva al contrario.

In questa direzione, assumendo quindi la prospettiva della trasparenza come *driver* della progettazione dell'organizzazione aziendale e soprattutto del sistema informativo aziendale, sembra possibile delineare con evidenza il percorso logico-metodologico che dovrebbe portare alla trasparenza informativa e, a valle, organizzativa, che rappresenta la principale finalità del nostro ragionamento. Infatti, seguendo un approccio "a gambero", molto caro al *project management* di stampo anglosassone, è agevole arguire che si arriva alla trasparenza dell'organizzazione (necessariamente) in presenza di trasparenza dell'informazione; a monte, questa è conseguenza della condivisione dell'informazione, che, a sua volta, dipende sempre di più dall'informatizzazione dell'organizzazione; ancora più a monte, questa è conseguenza della proceduralizzazione dell'organizzazione (ossia, nel linguaggio del *knowledge management*, operando trasformazioni dal continuo al discontinuo), che, a sua volta, dipende dalla formalizzazione dell'organizzazione, che, infine, discende, nel caso specifico della sanità e in una visione tutto sommato trasversale alle diverse aziende sanitarie, dal riferimento naturale alla burocrazia professionale (Mintzberg, 1996).

È chiaro, come osservato in precedenza, che la trasparenza non può costituire di per sé un valore terminale (se non nel contesto pubblico, verso il quale si è operato il richiamo al d.lgs. 150/2009) e per il quale "impiccarsi", ma più pragmaticamente un valore strumentale. Pertanto, nel funzionamento quotidiano dell'azienda sanitaria dovrà inevitabilmente esserci discrezionalità (si pensi al classico esempio dell'urgenza per le liste di attesa), perché ovviamente esiste il malato e non la malattia, o peggio la procedura in sé e per sé, ma è altrettanto ovvio che la discrezionalità non può e non deve diventare primato del singolo. Infatti, l'evidenza di regole interne e limiti di discrezionalità non impedisce l'opaco, ma almeno lo limita il più possibile: si pensi, solo per fare un per esempio, al "discernimento clinico", che costituisce un lampante esempio di discrezionalità correttamente esercitabile, se motivata e trasparente, in un rigido sistema di vincoli.

Su questo punto si ritiene necessaria una breve riflessione. Nella sanità moderna è chiaramente anacronistico pensare a un funzionamento organizzativo o a un comportamento professionale ispirati dalla sola procedura nel senso burocratico, ossia priva degli adeguati contributi di comunicazione, attenzione, empatia, ecc. Allo stesso tempo, tuttavia, non si può correre il rischio di essere soltanto dalla parte del paziente, se mai ci possa essere una simile provocazione, perché bisognerebbe essere anche dalla parte dell'operatore sanitario, sempre più esposto a montanti complessità. In tal senso, la procedura, chiaramente approfondita, ponderata, congegnata, sperimentata, validata e opportunamente e tempestivamente aggiornata, si ritiene costituisca un contributo fondamentale al buon esercizio della professione; anche da un punto di vista psicologico, perché rassicura l'operatore nel rapporto con il paziente. In tal senso, la trasparenza può evidentemente contribuire a migliorare il clima organizzativo e conseguentemente il profilo psicologico e prestazionale del singolo operatore sanitario.

Come ormai non competono tra loro le singole aziende, ma vere e proprie filiere integrate, così ormai non compete più il singolo medico, ma l'organizzazione sanitaria nel suo complesso. Dalla tensione informatica alla "trasparenza" può quindi derivare un contributo all'efficienza non soltanto della singola azienda, ma anche del complessivo network sanitario. Un sistema informativo sanitario trasparente, infatti, costituirebbe un esempio ipercompreso di struttura reticolare (in termini di risorse, prestazioni, output e outcome e così via), ma, in termini sociali, dovrebbe essere considerato anche come vettore di parità nell'accesso al servizio sanitario e alla prestazione sanitaria (in senso multidimensionale: quantità, qualità, rapidità, efficienza, costo, ecc.), che è circostanza valida per tutti i contesti, ma soprattutto per i sistemi sanitari solidaristici (parità di accesso può significare non opacità della composizione della lista di attesa, adeguata informazione sui servizi e sulle prestazioni, consapevolezza del mix tra offerta interna ed esterna – si pensi all'intramoenia – e così via).

Purtroppo, infatti, anche nel settore sanitario esistono notevoli livelli di asimmetria informativa (interni ed esterni rispetto alle singole organizzazioni) e non solo in scenari "storici" (come lavoro, crediti, assicurazioni, ecc.). Il settore dei servizi è certamente quello a maggiore contenuto informativo e quindi l'informazione (e ancor di più la sua eventuale assenza) è risorsa critica anche e soprattutto per i servizi sanitari, perché nella sanità moderna l'informazione viaggia a livello di reticolo, in una macroscopica *supply chain* di natura socio-economica (tramite, almeno in Italia, sussidiarietà verticale e orizzontale). Senza informazione e comunicazione, nel tempo e nello spazio, un'azienda sanitaria è quindi destinata a restare un'entità monolitica, non soltanto magari opaca nel funzionamento, ma anche avulsa per competitività dal mercato del sistema sanitario (infatti, è altamente probabile che, avendo possibilità di scelta, il paziente si rivolga alla struttura più trasparente).

La spinta alla trasparenza, peraltro, emerge anche dalla diffusione sempre più intensa degli strumenti del web 2.0, in cui si producono continuamente contenuti generati dagli utenti, spesso ritenuti persino più attendibili dell'informazione e della comunicazione ufficiale. Il perché di simili interpretazioni è presto detto: alle opinioni dei cittadini, soprattutto se organizzati in comunità numerose, viene automaticamente associato un valore superiore,

vuoi per innovativo passaparola vuoi per disinteresse allo sforzo della verifica vuoi per sfiducia derivante da una sorta di tradizionale assenza di trasparenza da parte degli organi ufficiali.

Alle porte del web, inoltre, sta bussando anche il web 3.0 (ossia il web semantico), che dovrebbe rappresentare una vera e propria rivoluzione culturale per l'internet tradizionalmente intesa, ma che in particolare per quanto riguarda la trasparenza dovrebbe fornire una spinta decisiva. Alla base della nuova filosofia del web, infatti, risiederà l'autenticazione del dato e del suo contenuto, che dovrebbe restituire maggiore affidabilità alle informazioni reperite su internet, con conseguente aumento della fiducia.

La trasparenza, inoltre, come accennato all'inizio del lavoro, deve essere usabile, nel senso che bisogna non soltanto distinguere gli ambienti / contesti a scarsa informazione / comunicazione da quelli a maggiore intensità informativa / comunicativa, ma anche prestare attenzione a quelli caratterizzati da eccesso d'informazione / comunicazione, che finisce con il determinare una sostanziale assenza di trasparenza. Si pensi, per esempio, a un particolare aspetto della Riforma Brunetta (l. 15/2009, d.lgs. 150/2009), che obbliga gli enti pubblici a comunicare tramite internet i compensi erogati nell'ambito degli incarichi conferiti: lo sforzo in sé è certamente rilevante, visto che si tratta di un vero e proprio cambiamento culturale, ma è doveroso sottolineare che si tratta di una pubblicità a volte farraginosa, perché implementata tramite link spesso "nascosti" oppure tramite dati pubblicati in formato "pdf" e quindi non immediatamente aggregabili / scomponibili per analisi integrate e / o incrociate, che, pertanto, generano una complessiva non usabilità del dato.

In questo percorso evolutivo, nell'informatizzazione sanitaria può svolgere una funzione molto importante l'ECDL Health, ossia la recente certificazione informatica per gli operatori sanitari (della quale per l'Italia è responsabile AICA), prevalentemente concentrata sui criteri di trattamento del dato informatico sanitario. Infatti, l'investimento in formazione finalizzata a diffondere il buon uso dell'informatica sanitaria, oltre a essere in passato prevista per legge (si fa riferimento alla prima versione dell'allegato B al d.lgs. 196/2003, modificato nel 2012 dal "decreto semplificazioni" del Governo Monti) risponde alla fondamentale attenzione da rivolgere alle risorse umane, che rivestono costantemente un ruolo irrinunciabile nell'organizzazione aziendale, perché, nel ragionamento in questione, il più sofisticato sistema informatico in realtà vale pochissimo se usato da un personale incapace, perché non adeguatamente formato e addestrato, alla sua più completa utilizzazione.

Di seguito, pertanto, si propone una contestualizzazione all'ECDL Health delle riflessioni finora sviluppate in tema di trasparenza, nella convinzione che tale valore, soprattutto in ragione della sua multidimensionalità, abbia una notevole potenzialità proprio nel comparto sanitario: in altre parole, considerare la trasparenza un importante principio per la progettazione dei sistemi informativi sanitari, per coerenza, genera a valle la necessità / opportunità di trasmettere agli operatori sanitari il senso, la motivazione. In ogni caso, oltre che nella prospettiva proposta in questo studio, l'adeguamento dell'orizzonte culturale dell'ECDL Health almeno alle recenti evoluzioni normative in materia di trasparenza, per il Servizio Sanitario Nazionale in generale, appare indispensabile.

5. Una content analysis sul Syllabus dell'ECDL Health.

Alla base della certificazione professionale in oggetto insiste un *corpus* di concetti, conoscenze e competenze in materia di sistemi informativi sanitari che è stato formalizzato all'interno del Syllabus, attualmente alla versione 1.1, che si articola in 4 sezioni (Concetti; Compiti assistenziali; Abilità dell'utente; Norme e procedure), per un totale di 52 item, suscettibili di verifica durante l'esame per il conseguimento della certificazione. Il Syllabus, ponendosi come perimetro tassonomico di riferimento per un'adeguata conoscenza dei sistemi informativi sanitari, peraltro in una prospettiva internazionale, è sensibile, come facilmente comprensibile, a revisioni nel tempo, dovute agli eventuali sviluppi delle competenze di riferimento (informatica, organizzazione, diritto, ecc.).

Nel corso del 2013 sarà rilasciata una nuova versione del Syllabus (1.5), attualmente ancora in fase di verifica interna al gruppo di lavoro AICA sull'ECDL Health. In termini complessivi, il Syllabus 1.5 conserva la suddivisione in 4 sezioni (ma da "Compiti assistenziali" si passa a "Compiti e responsabilità professionali") e incrementa leggermente il numero degli item, che passano da 52 a 54.

Anche in ragione di tale novità, si è ritenuto utile procedere a un'analisi di contenuto delle due versioni del Syllabus (l'attuale, 1.1 – e la prossima, 1.5) per verificare quanto e come il concetto di trasparenza sia preso in considerazione, soprattutto in ragione delle recenti evoluzioni normative avvenute in Italia (l. 15/2009, d.lgs. 150/2009, d.lgs. 33/2013). La metodologia dell'indagine è stata articolata nel seguente modo:

- a) definizione di un insieme (*set*) di *keyword*, collegate al valore multidimensionale della trasparenza studiato nel presente lavoro, da rintracciare all'interno dei testi;
- b) individuazione e conteggio della frequenza assoluta delle occorrenze;
- c) individuazione e conteggio della frequenza cumulata nell'area di riferimento (per prossimità lessicale);
- d) commento dell'emerso.

In ragione di tale metodologia, si è proceduto a considerare *keyword* le seguenti parole (associate in senso generale alla trasparenza informatica sanitaria): trasparenza, appropriatezza, tempestività, accessibilità, condivisione, riservatezza, identificazione, riconoscimento, integrazione (*integration*). Oltre al sostantivo, è stato considerato oggetto d'indagine anche il collegato aggettivo (declinato nel genere e nel numero) e avverbio (es.: appropriatezza, appropriato, appropriata, appropriati, appropriate, appropriatamente); inoltre, per giudizio, ossia dopo una prima analisi, si è deciso di comprendere anche le parole "accesso", "accessi", "identificare" e "riconoscere".

Si è quindi proceduto alla verifica delle occorrenze, usando, in ragione della limitatezza dell'universo oggetto dell'indagine, normali strumenti di *office automation*. I risultati dell'analisi sono riportati nella Tab. 1

Analisi Occorrenze	Syllabus		Aree		Analisi Occorrenze	Syllabus		Aree	
	1.1	1.5	1.1	1.5		1.1	1.5	1.1	1.5
Trasparenza	0	0	0	0	Riservatezza	5	6	6	8
Trasparente	0	0			Riservato	0	0		
Trasparenti	0	0	2	3	Riservata	1	1	2	6
Trasparentemente	0	0			Riservati	0	0		
Appropriatezza	0	0	1	1	Riservate	0	1	3	7
Appropriato	1	1			Riservatamente	0	0		
Appropriata	1	1	9	11	Identificazione	0	0	0	4
Appropriati	0	0			Identificato	0	0		
Appropriatezza	0	0	1	1	Identificata	0	0	0	4
Appropriatamente	0	1			Identificati	0	0		
Tempestività	0	0	1	1	Identificate	0	0	0	4
Tempestivo	0	0			Identificare	2	6		
Tempestiva	0	0	1	1	Riconoscimento	0	1	3	7
Tempestivi	0	0			Riconosciuto	0	0		
Tempestive	1	1	2	2	Riconosciuta	0	0	0	4
Tempestivamente	0	0			Riconosciuti	0	0		
Accessibilità	1	1	2	2	Riconosciute	0	0	0	4
Accessibile	0	0			Riconosciutamente	0	0		
Accessibili	0	0	2	2	Riconoscere	3	6	0	4
Accessibilmente	0	0			Integrazione	0	2		
Accesso	8	7	2	2	Integrazioni	0	1	0	4
Accessi	0	3			Integrato	0	0		
Condivisione	0	0	2	2	Integrata	0	0	0	4
Condiviso	0	0			Integrati	0	0		
Condivisa	0	0	2	2	Integrate	0	0	0	4
Condivisi	1	1			Integratamente	0	0		
Condivise	1	1	2	2	Integration	0	1	0	4
continua sull'altra colonna					Totale	25	42		

(Tab. 1) Tabella sinottica e comparativa dei risultati della content analysis

e permettono di commentare alcune evidenze:

- il concetto di *trasparenza* in senso stretto non è mai espressamente nominato in nessuna delle due versioni del Syllabus, dato che l'area di riferimento registra zero occorrenze, anche sommando tutte le keyword dell'area;
- il concetto che ha registrato il numero più elevato di occorrenze, sempre a livello di area e in entrambe le versioni, è quello relativo all'*accessibilità*, che per certi versi ovviamente richiama quello di trasparenza, sia nella prospettiva degli utenti interni (operatori sanitari) sia in quella degli utenti esterni (pazienti e cittadini);
- il concetto che ha registrato il maggior incremento percentuale di occorrenze, sempre a livello di area, è quello relativo all'*integrazione*, che, assieme all'*accessibilità*, sembra comunque implementare, almeno nella prospettiva interna dei sistemi informativi sanitari, l'orizzonte dell'organizzazione "trasparente";
- nella nuova versione del Syllabus tutti i concetti analizzati registrano un numero di occorrenze uguale o superiore rispetto alla precedente versione (lo stesso concetto di trasparenza, come si commentava in precedenza, risulta assente in entrambe le versioni).

Prima del suo definitivo rilascio, pertanto, si ritiene utile inserire nel Syllabus 1.5 alcuni riferimenti alla trasparenza, ormai invocata normativamente da diverse fonti, anche specificamente rivolte al comparto sanitario (cfr. d.lgs. 33/2013, art. 41). Tale adeguamento, inoltre, sembra dover essere orientato soprattutto verso l'apertura all'esterno del sistema informativo quale "sistema delle informazioni" [Carignani, 2004], dato che la *content analysis* ha restituito esiti positivi, a parere di chi scrive, nella prospettiva interna, ossia della collaborazione informativa tra flussi, processi e operatori.

6. Conclusione.

Con l'evoluzione dell'informatica e dei sistemi informativi aziendali, il trattamento delle informazioni, all'interno e all'esterno delle organizzazioni, è diventato sempre più facile (nel senso di *user friendly*), diffuso e aperto. In particolare, l'introduzione dell'internet (e delle collegate soluzioni intranet ed extranet), con il successivo sviluppo del web 2.0, ha concretamente modificato le modalità, le finalità e per certi versi gli "stili" di fruizione dei dati, delle informazioni e delle conoscenze, soprattutto nei contesti aziendali. Le innovazioni informatiche, inoltre, finiscono molto spesso per trasformarsi nel tempo in vere e proprie innovazioni sociali (si pensi alla domotica, agli *smartphone* e, naturalmente, ai *social network*), soprattutto in quei campi in cui la componente umana sia prevalente. In tal senso, quello a maggiore rilevanza è senza ombra di dubbio la sanità, i cui servizi sono per natura orientati a ero-

gare quanto di più umano ci possa essere, ossia la salute. Queste combinazioni, tecnologiche e sociologiche, devono oggi essere costantemente informate di “valori”, vuoi per responsabilità sociale, vuoi per previsione normativa, vuoi per prospettiva di miglioramento (ai fini dell’efficacia, dell’efficienza e dell’economicità). Tra questi valori, riveste oggi grande importanza la trasparenza, soprattutto nel comparto sanitario e anche in termini economici, allo scopo di assicurarne sostenibilità per le prossime generazioni [Kotler et al., 2010].

In questo lavoro si è considerata la trasparenza sanitaria nella sua multidimensionalità, indagando teoricamente una possibile ricostruzione dei suoi “valori”, fino a proporre un complessivo *framework*, allo scopo di evidenziarne le applicazioni da un punto di vista economico, aziendale e sociale. Questo studio, qualificandosi prevalentemente come *conceptual paper*, si è quindi basato sulle dinamiche generatrici di trasparenza provenienti dall’informatizzazione sanitaria, esaminando il funzionamento degli ambienti, dei mercati e delle aziende sanitarie nella prospettiva della gestione dell’informazione e della comunicazione in sanità, arrivando a identificare nella trasparenza un possibile principio per la progettazione dei sistemi informativi sanitari.

Inoltre, facendo riferimento all’ECDL Health, la certificazione professionale sull’informatica sanitaria della quale per l’Italia è responsabile AICA, è stata svolta anche una *content analysis* sulle versioni 1.1 e 1.5 del Syllabus, allo scopo di verificare l’attenzione riservata alla trasparenza dei sistemi informativi sanitari. Dall’indagine sono derivati risultati in chiaroscuro: un riferimento diretto alla trasparenza manca in entrambe le versioni, ma si registrano numerosi riferimenti a “proprietà” della trasparenza, soprattutto in un’accezione segnatamente informatica. Esistono quindi per il nuovo Syllabus possibili margini di sviluppo, dovuti alle recenti evoluzioni normative in Italia, alla necessità di una maggiore apertura dell’azienda sanitaria al confronto e alla valutazione e dovuti soprattutto alle sempre più pressanti richieste, provenienti da cittadini e operatori, d’informazione, partecipazione e responsabilizzazione.

In fin dei conti, infatti, la *content analysis* è stata svolta su un *corpus* “auspicato” di concetti, conoscenze e competenze in materia d’informatica sanitaria, chissà quanto possedute, praticate e approfondite nella realtà quotidiana delle strutture sanitarie. Un’organizzazione sanitaria più trasparente, basata su sistemi informativi sanitari più trasparenti, costantemente dotati di adeguati meccanismi di sicurezza informatica [Carignani e Rajola, 2010; Festa e Teti, 2010], rappresenta invece al giorno d’oggi una vera e propria “tensione” aziendale, sicuramente nella prospettiva del valore economico (grazie a probabile maggiore efficacia ed efficienza, ma soprattutto nella prospettiva del valore umano, sociale e civile connaturato nel bisogno di salute delle persone e delle popolazioni).

Bibliografia.

- Arduini R., Economia e gestione delle aziende sanitarie, Angeli, Milano, 2010.
- Bellucci A., Cardoni A., Elementi di economia delle aziende sanitarie, Giappichelli, Torino, 2008.
- Bracchi G., Francalanci C., Motta G., Sistemi informativi d’impresa, McGraw-Hill, 2010.
- Cappelli L., Renzi M.F., Management della qualità, Cedam, Padova, 2010.
- Carignani A. (a cura di), Tecnologie dell’informazione e della comunicazione per le aziende, McGraw-Hill, Milano, 2004).
- Carignani A., Rajola F., (a cura di) ICT e sistemi informativi aziendali, McGraw-Hill, Milano, 2010.
- De Witt G., Informatica moderna e produttività d’impresa, Angeli, Milano, 2001.
- Dirindin N., Vineis P., Elementi di economia sanitaria, il Mulino, Bologna, 2004.
- Donia Sofio A., Microeconomia sanitaria e politiche d’intervento, Aracne, 2001.
- Festa G., Teti A., Il contributo della certificazione ECDL Health alla sicurezza dell’informazione in sanità, in Di Resta F., Ferraris di Celle B., (a cura di) La sicurezza nei sistemi informativi sanitari, Edisef, Roma, 2010.
- Gambaro M., Ricciardi C.A., Economia dell’informazione e della comunicazione, Laterza, Bari, 2003.
- Kotler P., Shalowitz J., Stevens R.J., Turchetti G., Marketing per la sanità, McGraw-Hill, Milano, 2010.
- Lambin J.J., Market driven management – Marketing strategico e operativo, McGraw-Hill, Milano, 2008.
- Marinò L., La ricerca dell’economicità nelle aziende sanitarie locali, Giappichelli, Torino, 2008.
- Mintzberg H., La progettazione dell’organizzazione aziendale, il Mulino, Bologna, 1996.
- Rumiati R., Bonini N., Decisioni manageriali, il Mulino, Bologna, 2010.
- Schilling M.A., Gestione dell’innovazione, McGraw-Hill, Milano, 2009.
- Teti A., Festa G., Sistemi informativi per la sanità, Apogeo, Milano, 2009.
- Tommasetti A., Cuccurullo C., Mappe strategiche e relazioni sistemiche tra indicatori di gestione: il caso delle aziende sanitarie locali, in AA.VV., L’evoluzione del controllo di gestione. Modelli ed esperienze, Angeli, Milano, 2004.
- Tonti A., La semplificazione dei processi e delle procedure, Egea, Milano, 2002.
- Tramarin A., L’ospedale ammalato, Marsilio, 2002.
- Ziliotti M., L’economia dell’informazione, il Mulino, Bologna, 2001.

Uno strumento di modellazione della scrittura a mano libera per l'analisi di patologie degenerative neuromuscolari

D. Impedovo^(1,2), G. Pirlo^(1,2), C. O'Reilly⁽³⁾, R. Plamondon⁽³⁾

⁽¹⁾ Computer Science Department, *Bari University, Via Orabona 4, Bari, Italy*

⁽²⁾ Rete Puglia Centre, - *Via G. Petroni 15/F.1, Bari, Italy;*

⁽³⁾ Ecole Polytechnique de Montreal, Canada
giuseppe.pirlo@uniba.it

This paper presents a mathematical tool for the analysis of handwriting for research on neuromuscular diseases. In particular, the Delta-Lognormal and the Sigma-Lognormal models are considered to investigate handwriting generation processes. Successively a software system is presented for the analysis of handwriting and its potential is discussed for diagnosing and monitoring degenerative neurological diseases. The experimental results show the effectiveness of the approach and highlight some valuable directions for further research.

Neuromuscular Disease, Handwriting Generation Model, Handwriting Analysis, Delta-Lognormal, Sigma-Lognormal, disease diagnosis, disease monitoring.

1. Introduzione

Nelle società avanzate le malattie neuromuscolari, che spesso conducono a gravi forme di disabilità, sono in costante aumento. Le spese legate al trattamento dei pazienti affetti da queste patologie, che riguardano sia i costi di ospedalizzazione che di riabilitazione e cura a lungo-termine, sono enormi e la richiesta per lo sviluppo di nuovi strumenti efficaci di diagnosi precoce e di monitoraggio della malattia è sempre maggiore [Plamondon, 1995].

Questo lavoro presenta uno strumento matematico di modellazione dei processi di generazione della scrittura a mano libera e del suo uso per la diagnosi ed il monitoraggio delle malattie degenerative neuromuscolari. In particolare vengono presentati i modelli delta-lognormale e sigma-lognormale alla base dei processi di generazione della scrittura a mano libera e gli algoritmi per l'estrazione di parametri che descrivono i processi alla base dell'esecuzione di tratti manoscritti [Plamondon et al., 2012a]. Successivamente vengono descritte le principali problematiche legate alla fase di acquisizione dei dati, di

modellazione del processo di scrittura e di decisione, in ordine alla valutazione delle malattie neurodegenerative.

I risultati sono di grande interesse per lo sviluppo di un modello statistico a supporto di un protocollo diagnostico di valutazione delle patologie degenerative neuromuscolari.

2. Funzione di trasferimento del sistema neuromuscolare

Il modello proposto, basato sulla teoria cinematica di generazione dei movimenti rapidi, permette la descrizione dei processi di generazione di tracciati manoscritti attraverso l'equazione delta-lognormale. In particolare esso permette di predire il fondamentale rapporto velocità/accuratezza osservato nell'esecuzione di movimenti di scrittura diretti verso un punto *target*.

In letteratura diverse strategie sono state proposte per l'analisi e la sintesi dei processi neuromuscolari, in grado di controllare le azioni agoniste ed antagoniste. Queste strategie considerano i seguenti due criteri fondamentali:

1. Il modello deve essere sufficientemente realistico da riprodurre perfettamente specifiche traiettorie del movimento della punta della penna nei processi di scrittura;
2. Il suo potere descrittivo del modello deve essere consistente con le proprietà fondamentali dei singoli tratti di scrittura.

3. Il Modello Delta-Lognormale

Il metodo generale per studiare la risposta impulsiva di uno specifico controllore, ad esempio del modulo di controllo della velocità, è di considerare l'insieme complessivo di reti neurali e muscolari implicate nella produzione di un singolo tratto come un sistema lineare che produce il profilo di velocità curvilinea partendo da un impulso di comando di ampiezza D che si verifica all'istante t_0 . In questo modo il profilo di velocità curvilinea riflette direttamente la risposta impulsiva $H(t-t_0)$ del sistema neuromuscolare [Plamondon, 1995; Alimi e Plamondon, 1993].

La descrizione formale di questa risposta impulsiva può essere specificata considerando ciascun modulo di controllo come composto da due sistemi che rappresentano gli insiemi di reti neurali e muscolari implicati rispettivamente nella generazione delle attività agoniste ed antagoniste che producono uno specifico movimento. In tal modo la risposta impulsiva convergerà verso una curva lognormale e la risposta del modulo velocità potrà essere descritta come la differenza pesata di due curve lognormali, chiamata equazione delta-lognormale [Plamondon, 1994]. In questo contesto il controllo del modulo di velocità può essere visto come il risultato dalla simultanea attivazione (al tempo $t = t_0$) di due sistemi neuromuscolari tra loro antagonisti, con ampiezza di comando D_1 e D_2 , rispettivamente. Entrambi i sistemi reagiscono con una risposta impulsiva descritta mediante una funzione lognormale, i cui parametri μ_1, σ_1^2 e μ_2, σ_2^2 sono associati con il ritardo temporale e il tempo di risposta (su una scala temporale logaritmica) di ciascun processo, così come descritto dall'eq. (1):

$$v_{\sigma} = \frac{D_1}{\sigma_1 \sqrt{2\pi}(t-t_0)} \exp\left[-\frac{(\ln(t-t_0)-\mu_1)^2}{2\sigma_1^2}\right] - \frac{D_2}{\sigma_2 \sqrt{2\pi}(t-t_0)} \exp\left[-\frac{(\ln(t-t_0)-\mu_2)^2}{2\sigma_2^2}\right] \quad (1)$$

Anche per movimenti di rotazione, come la flessione e l'estensione del polso, la velocità angolare seguirà una specifica equazione delta-lognormale di ampiezza D_1 e D_2 . La Figura 1 mostra il profilo della delta-lognormale ottenuta usando i seguenti parametri: $D_1 = 2.35$ cm; $D_2 = 0.94$ cm; $t_0 = 0.348$ sec; $\mu_1 = -1.61$; $\mu_2 = -1.45$; $\sigma_1^2 = 0.0134$; $\sigma_2^2 = 0.00497$.

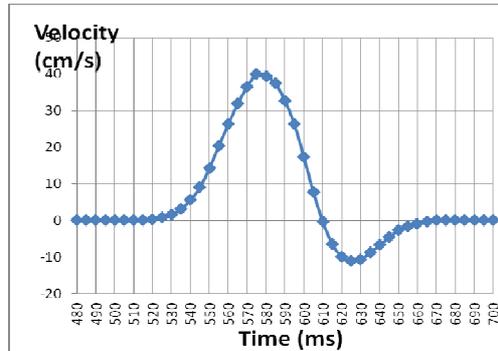


Figura 1: Profilo della funzione Delta lognormale

Riguardo ai due criteri alla base delle strategie di analisi dei processi di generazione della scrittura a mano libera, si evidenzia come il modello delta-lognormale è certamente tra i più promettenti in quanto:

- 1) l'equazione delta-lognormale è la più efficace nel ricostruire dati sperimentali legati ai processi di generazione del movimento [Plamondon, 1994].
- 2) i soggetti umani tendono a produrre tratti aventi circa la stessa durata temporale e la stessa precisione spaziale, al fine di assicurare la leggibilità degli stessi [Plamondon et al., 2012a].

4. Il modello Sigma-Lognormale

La teoria cinematica dei movimenti umani rapidi si basa sul modello sigma-lognormale per rappresentare l'informazione sia dei comandi motore che delle proprietà legate alla tempistica del sistema neuromuscolare implicato nella produzione dei movimenti complessi come la scrittura a mano libera di una parola [O'Reilly e Plamondon, 2009; Plamondon, 2012b].

Il modello Sigma-Lognormale considera la velocità di un singolo stroke j come avente forma lognormale Λ scalata secondo un parametro di comando (D) e shiftata nel tempo dal tempo di occorrenza del comando (t_0), così come riportato nell'equazione (2):

$$\left| \vec{v}_j(t; P_j) \right| = D_j \Lambda(t - t_{0j}; \mu_j, \sigma_j^2) = \frac{D_j}{\sigma(t - t_{0j})\sqrt{2\pi}} \exp \left\{ \frac{[\ln(t - t_{0j}) - \mu_j]^2}{-2\sigma_j^2} \right\} \quad (2)$$

dove $P_j = [D_j, t_{0j}, \mu_j, \sigma_j, \theta_{sj}, \theta_{ej}]$ rappresenta l'insieme di parametri della funzione Sigma-Lognormale [O'Reilly e Plamondon, 2009]:

- D_j : ampiezza dei comandi di input;
- t_{0j} : occorrenza nel tempo dei comandi di input;
- μ_j : ritardo temporale del sistema neuromuscolare espresso in scala temporale logaritmica;
- σ_j : risposta temporale del sistema neuromuscolare, espresso in scala temporale logaritmica;
- θ_{sj} : angolo di partenza della traiettoria circolare descritta dal modello lognormale lungo il pivot;
- θ_{ej} : angolo di arrivo della traiettoria circolare descritta dal modello lognormale lungo il pivot.

Inoltre, dall'ipotesi che ogni tratto lognormale rappresenti il movimento lungo il pivot, la posizione angolare può essere calcolata come [O'Reilly e Plamondon, 2009]:

$$\phi_j(t; P_j) = \theta_{sj} + \frac{\theta_{ej} - \theta_{sj}}{D_j} \int_0^t \left| \vec{v}(\tau; P_j) \right| d\tau \quad (3)$$

In questo contesto, una parola manoscritta può essere vista come l'output di un generatore che produce un insieme di tratti individuali sovrapposti nel tempo. Le complesse traiettorie risultanti possono essere modellate come somma vettoriale delle distribuzioni lognormali (essendo N_{LN} il numero totale di curve lognormali in cui il tracciato manoscritto è decomposto):

$$\vec{v}(t) = \sum \Lambda(t) = \sum_{j=1}^{N_{LN}} \vec{v}_j(t; P_j) \quad (4)$$

Le componenti della velocità nello spazio cartesiano possono essere ricavate dalla velocità tangenziale come segue [Plamondon e Djioua, 2006]:

$$\begin{aligned} \vec{v}_x(t) &= \sum_{j=1}^{N_{LN}} \left| \vec{v}_j(t; P_j) \right| \cos(\phi_j(t; P_j)) \\ \vec{v}_y(t) &= \sum_{j=1}^{N_{LN}} \left| \vec{v}_j(t; P_j) \right| \sin(\phi_j(t; P_j)) \end{aligned} \quad (5)$$

L'errore di ricostruzione del profilo di velocità usando i parametri della Sigma-Lognormale può essere valutato confrontando il rapporto segnale-rumore (SNR) tra il campione ricostruito e quello originale:

$$10 \log \left(\frac{\int_{t_s}^{t_e} [v_{xo}^2(t) + v_{yo}^2(t)] dt}{\int_{t_s}^{t_e} [(v_{xo}(t) - v_{xa}(t))^2 + (v_{yo}(t) - v_{ya}(t))^2] dt} \right) \quad (6)$$

dove t_s e t_e sono rispettivamente il tempo iniziale e finale della parola manoscritta.

5. Setup sperimentale

Al fine di investigare approfonditamente i processi degenerativi, è stato utilizzato un apposito software di acquisizione di tracciati manoscritti e di generazione dei parametri del modello sigma-lognormale. In particolare, dato in input un tracciato manoscritto, il software determina il numero di funzioni lognormali e l'insieme dei parametri di ciascuna funzione: $(t_0, D, \mu, \sigma, \theta_s, \theta_e)$.

L'insieme dei parametri fornisce diverse informazioni che sono usate in questa ricerca per l'analisi del sistema neuromuscolare al fine di rilevare l'avanzamento di patologie neurodegenerative.

In particolare, la dinamica della scrittura a mano libera può essere utilizzata sia per la diagnosi precoce delle malattie neuromuscolari che per il suo monitoraggio nel tempo, anche per valutare gli effetti dei trattamenti farmacologici.

Al fine di dare inizio alla fase sperimentale e di perfezionare il tuning del sistema, alcune valutazioni iniziali sono state realizzate usando il database ISUNIBA, sviluppato nel Dipartimento di Informatica dell'Università di Bari. Ventidue pazienti affetti da malattie neuromuscolari degenerative sono stati coinvolti nello sviluppo del database. Ciascun paziente ha fornito 10 parole manoscritte in due sessioni distinte (5 per sessione), approssimativamente avvenute ad un mese di distanza l'una dall'altra. Al fine di monitorare la progressione della malattia opportune ripetizioni delle acquisizioni sono programmate mensilmente. La Figura 2 mostra il dispositivo di acquisizione usato per gli esperimenti: la tavoletta WACOM INTUOS-5.



Figura 2: WACOM INTUOS-5 - Sessione di acquisizione

6. Risultati Sperimentali

Il sistema che è stato usato per l'estrazione dei parametri è "Script Studio", realizzato all'École Polytechnique di Montreal, Canada [O'Reilly e Plamondon, 2009].

Per gli obiettivi del presente lavoro sono stati utilizzati in particolare due moduli software di Script Studio:

- Il primo modulo ha l'obiettivo di acquisire le parole manoscritte;
- Il secondo modulo ha l'obiettivo di calcolare i parametri.

Ciascuna parola manoscritta è modellata in una matrice, in cui ciascuna colonna riporta le seguenti informazioni:

- Colonna 1: timestamp di ciascun punto;
- Colonne 2,3,4: posizione X,Y e Z di ciascun punto;
- Colonna 5: pressione;
- Colonna 6: azimuth;
- Colonna 7: elevazione;
- Colonna 8: velocità tangenziale.

Queste informazioni sono memorizzate in un file con estensione *.hws* che include anche informazioni sul formato, frequenza di campionamento e numero di elementi. Un esempio di file *.hws* è riportato nel seguito:

```
TypeFichier: HWS
NombrePoints: 1235
FrequenceEchantillonage(Hz): 200.00
1.250000 0.545690 0.000000 0.000000 0.000000 0.000000 0.000000 151.00402091
.....
5.587000 122.950404 -67.73745 3 0.000000 0.000000 0.000000 0.000000 10.876394
```

Ciascuna parola è memorizzata, tratto a tratto, in un file con estensione *.ana*. In particolare, ciascun elemento della matrice degli stroke include:

- Colonna 1: istante d'inizio del tratto (t_{0j});
- Colonna 2: ampiezza del tratto (D_j);
- Colonna 3,4: ritardo temporale logaritmico (μ_j) e tempo di risposta logaritmico (σ_j);
- Colonna 5,6: angolo iniziale (θ_{sj}) e finale (θ_{ej}) del tratto;

Le tre ultime colonne sono utilizzate per la generalizzazione in 3D e non sono di interesse per le finalità del presente lavoro.

Nel seguito si riporta un esempio di file *.ana*.

```
TypeFichier: ANA
Version: 3
NombreLognormales: 14
1.560254 32.21146 -0.8764 91 0.375699 -2.124533 -1.6 78667 0.000 0.000 1
.....
1.730000 11.151567 -1.599555 0.398123 -1.207766 -1.276859 0.000 0.000 1.
```

La stessa analisi è stata sviluppata per investigare modificazioni nelle parole manoscritte dovute a malattie neuromuscolari. A tale scopo sono stati considerati 22 pazienti affetti da malattie neuromuscolari degenerative e sono stati invitati, in specifiche sessioni di acquisizione, a scrivere la parola “mamma” sulla tavoletta elettronica.

Le Figure 3a, 4b e 5c mostrano la parola “mamma” scritta rispettivamente da un individuo sano, da un paziente nelle fasi iniziali della malattia e da un paziente in fase avanzata,

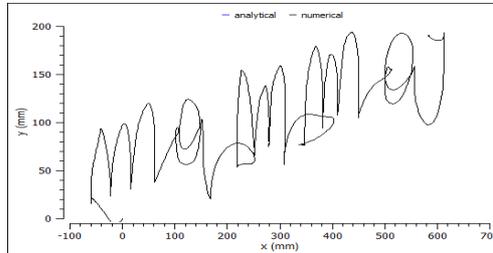


Figura 3a: Esempio di parola scritta da individuo sano

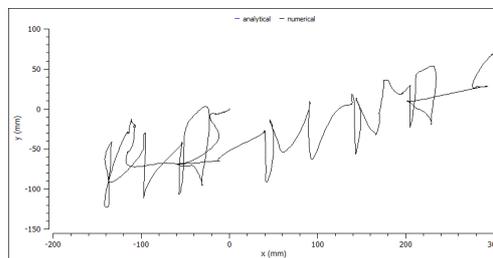


Figura 4a: Esempio di parola scritta da individuo nelle prime fasi della malattia

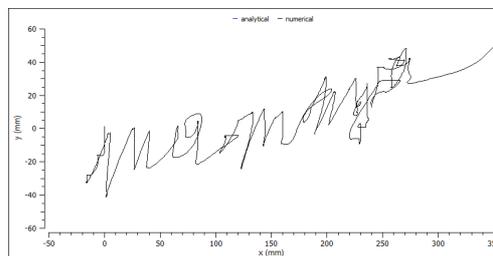


Figura 5a: Esempio di parola scritta da individuo nelle fasi avanzate della malattia

I file .ana rappresentati nelle Figure 3b, 4b e 5b mostrano rispettivamente i profili di velocità delle parole nelle Figure 3a, 4a e 5a.

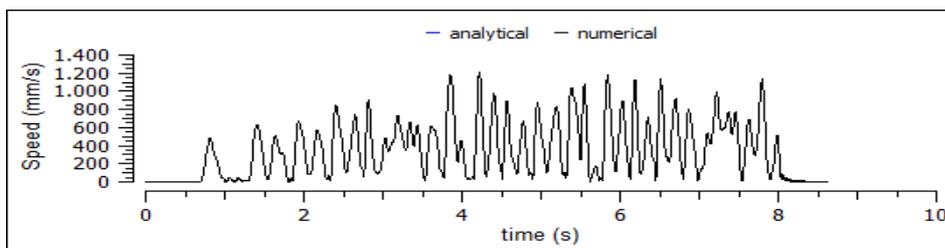


Figura 3b: profilo di velocità della parola in figura 3a

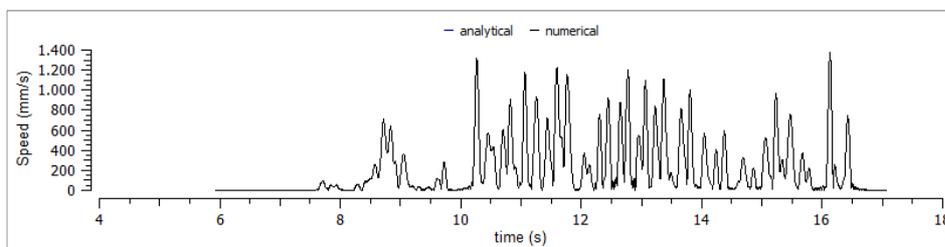


Figura 4b: profilo di velocità della parola in figura 4a

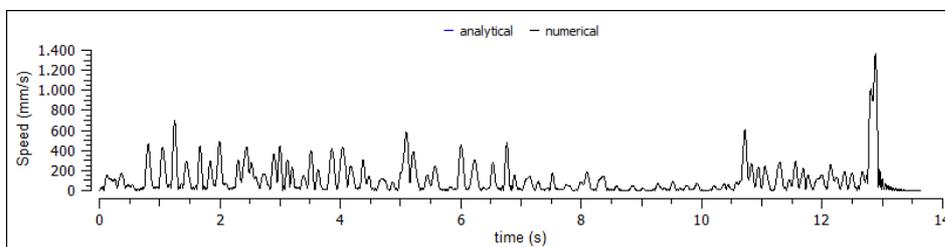


Figura 5b: profilo di velocità della parola in figura 5a

Dall'analisi dei profili di velocità si può osservare che negli individui sani i valori massimi di velocità sono piuttosto regolari in ampiezza (Figura 3b), mentre invece questa regolarità è fortemente ridotta negli individui negli stadi iniziali della malattia (figura 4b) e completamente persa negli individui in stadi avanzati della malattia stessa (figura 5b).

Inoltre si deve notare che la durata temporale passa da 3,5 sec. fino a 6 sec. per raggiungere la durata di 10 sec., come si vede confrontando le figure 3b, 4b e 5b. Anche il ritardo iniziale passa da 1 sec. a circa 2 sec., per raggiungere 4.5 sec. per persone in stadi avanzati della malattia.

Inoltre, guardando in maggior dettaglio i profili di velocità, è possibile verificare come individui sani difficilmente portano a zero il valore della velocità, mentre al contrario la velocità diventa sempre più spesso uguale a zero, e per periodi sempre più lunghi, allorchando si considerano individui con stadi di malattia sempre più avanzati.

7. Discussione e Conclusioni

In questo lavoro viene proposta l'analisi della scrittura a mano libera per individuare parametri utili alla diagnosi ed al monitoraggio di pazienti affetti da malattie degenerative neuromuscolari.

Alcuni risultati preliminari mostrano la diretta relazione tra il cambiamento delle caratteristiche dei processi di scrittura e i sintomi iniziali dell'insorgenza di patologie neurodegenerative. Ciò nonostante, ulteriori ricerche sono necessarie per migliorare la modellazione del sistema neuromuscolare e la sua applicazione allo studio delle malattie degenerative neuromuscolari [Woch et al., 2011; O'Reilly e Plamondon, 2012].

In tal senso ulteriori sviluppi della ricerca saranno orientati principalmente all'ampliamento del database dei pazienti, così come lo studio a lungo termine della correlazione tra l'evoluzione dei parametri del modello e la progressione della malattia.

Si evidenzia anche che tale ricerca, consentendo una maggiore comprensione dei processi alla base della scrittura a mano libera, certamente permetterà lo sviluppo di nuovi e più performanti sistemi di verifica automatica di firme manoscritte adattativi e personalizzabili [Impedovo e Pirlo, 2007; Di Lecce et al., 2000; Impedovo et al., 2009; Impedovo e Pirlo, 2012], in grado di fornire soluzioni evolute nei diversi contesti applicativi dell'e-commerce, e-health e e-government. nei quali è ormai fortemente richiesto l'uso efficace della firma biometrica [Vielhauer, 2005; Impedovo e Pirlo, 2011; Impedovo et al., 2012].

Acknowledgement

Gli autori ringraziano il Dott. Pietro Schino, Presidente dell'Associazione Alzheimer di Bari, e la Dott.ssa Caterina Pinto, Psicologa dell'Associazione Alzheimer di Bari, per il loro supporto allo sviluppo del database ISUNIBA.

Bibliografia

Alimi, A. M., Plamondon, R., Performance Analysis of Handwritten Strokes Generation Models, *Proc. of the 3-rd IWFHR*, Buffalo, USA, May 25-27, 1993, 272-283.

Di Lecce, V. , Dimauro, G. , Guerriero, A. , Impedovo, S. , Pirlo, G. , Salzo, A., A Multi-Expert System for Dynamic Signature Verification, *Proc. of the 1st MCS Conference*, LNCS, Springer-Verlag, J. Kittler and F. Roli (Eds.), Cagliari, Italy, 1857, 2000, 320-329.

Impedovo, D., Pirlo, G., Stasolla, E., Trullo, C.A., Learning Local Correspondences for Static Signature Verification, *Proc. of the 11th Int. Conf. of the Italian Association for Artificial Intelligence (AI*IA 2009)*, Reggio Emilia, Italy, 2009, 385-394.

Impedovo, S., Pirlo, G., *Verification of Handwritten Signatures: an Overview*, *Proc. of the 14th International Conference on Image Analysis and Processing - ICIAP 2007*, IEEE Computer Society Press, Modena, Italy, 2007, 191-196.

Impedovo, D. , Pirlo, G. , On the Measurement of Local Stability of Handwriting - An application to Static Signature Verification, *Proc. of the Biometric Measurements and*

Systems for Security and Medical Applications (BIOMS 2010), IEEE Computer Society Press, Taranto, Italy, 2010, 41-44.

Impedovo, D., Pirlo, G., Stability Analysis of Static Signatures for Automatic Signature Verification, *Proc. of the 16th International Conference of Image Analysis and Processing (ICIAP 2011)*, Lecture Notes in Computer Science, Springer Publ., 6979, Ravenna, Italy, 2011, 241-247.

Impedovo, D. , Pirlo, G. , Sarcinella, L. , Stasolla, E. , Trullo, C.A., Analysis of Stability in Static Signatures using Cosine Similarity, *Proc. of the 12th International Conference on Frontiers in Handwriting Recognition (ICFHR 2012)*, Monopoli, Bari, Italy, 2012, 231-235.

O'Reilly, C. , Plamondon, R., Development of a Sigma-Lognormal Representation for On-Line Signatures, *Pattern Recognition*, 42, 2009, 3324-3337.

O'Reilly, C., Plamondon, R., Design of a Neuromuscular Disorders Diagnostic System Using Human Movement Analysis, *Proc. of the 11th Int. Conf. on Information Sciences, Signal Processing and their Applications*, Montreal, Canada, 2012, 787-792.

Plamondon, R., Handwriting Generation: the Delta-Lognormal Theory, *Proc. of the 4th International Workshop on Frontiers in Handwriting Recognition*, 1994, 1-10.

Plamondon, R., A Kinematic Theory of Rapid Human Movements: Part I: Movement Representation and Generation, *Biological Cybernetics*, 72, 4, 1995, 295-307.

Plamondon, R., Djioua, M., A Multi-Level Representation Paradigm for Handwriting Stroke Generation, *Human Movement Science*, 5, 4-5, 2006, 586-607.

Plamondon, R. , O'Reilly, C., Galbally, J., Almaksour, A., Ancquetil, E., Recent Developments in the Study of Rapid Human Movements with the Kinematic Theory: Applications to Handwriting and Signature Synthesis, *Pattern Recognition Letters*, 2012, available on-line at: <http://dx.doi.org/10.1016/j.patrec.2012.06.004>.

R. Plamondon, Strokes against stroke- Stroke for strides, *Proc. of the 3-rd ICFHR*, Bari, Italy, 18-20 Sept. 2012, available online at: http://www.icfhr2012.uniba.it/ICFHR2012-Invited_I.pdf.

Vielhauer, C., A Behavioural Biometrics, *Public Service Review: European Union*, 9, 2005, 113-115.

Woch, A., Plamondon, R., O'Reilly, C., Kinematic Characteristics of Successful Movement Primitives in Young and Older Subjects: a Delta-Lognormal Comparison, *Human Movement Science*, 30, 2011, 1-17.

Sistema Multi-Esperto Intelligente per la Diagnosi del Tumore al Seno

Donato Barbuzzi¹, Francesco Maurizio Mangini¹, Donato Impedovo²,
Giuseppe Pirlo¹

¹Dipartimento di Informatica, Università degli Studi di Bari "Aldo Moro"
Via E. Orabona, 4 - 70125 Bari

{donato.barbuzzi, francescomaurizio.mangini, giuseppe.pirlo}@uniba.it

²Dipartimento di Ingegneria Elettrica e dell'Informazione, Politecnico di Bari
Via E. Orabona, 4 - 70125 Bari
impedovo@deemail.poliba.it

Abstract. Breast cancer (BC) is the most common cancer among women and the second leading cause of cancer deaths in women today, after lung cancer. The American Cancer Society estimates that more than 230000 women will be diagnosed with BC annually in the United States, and more than 39000 will die from the disease. However, according to an annual report published by several institutions, the BC death rate among women appears to be declining with an increased emphasis on early detection. Therefore, it is essential that BC is detected as early as possible. Although it is very hard, laborious and time-consuming to collect many labeled and unlabeled data, collecting unlabeled samples is easier than labeled ones. On the other hand, standard learning algorithms learn a diagnosis model from labeled samples only and, usually, they require many labeled samples. Therefore, it would be desirable to develop a predictive learning method to achieve high performance using both labeled samples and unlabeled samples. To address this problem, in this paper, we introduce a semi supervised learning algorithm able to use few labeled samples as representative set to build the diagnostic model which is used for classifying the unlabeled data. The experimental results, performed on the Wisconsin breast cancer dataset from the UCI machine learning repository datasets, suggest that our solution outperforms standard supervised learning algorithms, even when the number of training samples is very small.

Keywords: Breast Cancer diagnosis, Intelligent Multi-Expert System, Semi-Supervised Learning.

1. Introduzione

Nell'ambito dell'apprendimento automatico, la classificazione è generalmente indicata come apprendimento supervisionato. L'apprendimento supervisionato tradizionale necessita di un numero sufficiente di dati etichettati per costruire un training set tale da garantire una buona capacità di generalizzazione [Kittler et al, 1998] da parte del classificatore. Preparare un numero consistente di dati etichettati è più difficile e costoso rispetto a limitarsi a raccogliere dati non etichettati. Se si dispone solo di una piccola quantità di dati etichettati e la restante parte dei dati è priva di etichetta, le strategie di apprendimento semi supervisionato (SSL) [Chapelle et al, 2006] possono fornire prestazioni di classificazione soddisfacenti. L'apprendimento semi supervisionato combina i campioni etichettati e quelli privi di etichetta per implementare un efficiente sistema di classificazione.

I sistemi di apprendimento SSL sono stati applicati in diversi campi della medicina. In particolare, la diagnosi del tumore al seno costituisce un'area di interesse per questo tipo di sistemi.

Il tumore al seno è la seconda causa di morte tumorale in tutto il mondo. A questa patologia è stata rivolta una crescente attenzione in termini di prevenzione. In questi ultimi anni si è osservata una lieve contrazione di questo tipo di tumore ascrivibile alla diagnosi precoce della malattia, che può portare alla guarigione se efficacemente trattabile. Uno degli elementi fondamentali per la diagnosi del tumore al seno è la classificazione del tumore. I tumori, infatti, possono essere di vario tipo. In particolare, nel primo stadio diagnostico è essenziale la macro distinzione tra benigni e maligni.

La maggior parte dei tumori al seno vengono rilevati sotto forma di noduli al seno. Tali noduli possono anche essere benigni e quindi è compito del medico distinguere i noduli benigni da quelli maligni. Il tasso di sopravvivenza a lungo termine per le donne affette da tumore al seno è più alto se la malattia viene diagnosticata nel suo stadio iniziale [Brenner, 2002].

La diagnosi precoce necessita di una procedura diagnostica precisa e affidabile che permetta ai medici di operare la distinzione tra tumori della mammella benigni e maligni [Subashini et al, 2009]. Purtroppo, questo richiederebbe la disponibilità di molti medici esperti dato l'ingente numero di controlli senologici che vengono regolarmente eseguiti. E' evidente, quindi, la necessità di automatizzare il sistema di diagnostica.

Questo lavoro presenta una nuova strategia basata sull'uso del feedback, in un sistema multi-esperto con topologia in parallelo. L'idea alla base di questo articolo è che *il comportamento collettivo di un set di classificatori fornisce generalmente più informazione rispetto ad ogni singolo classificatore e questa maggiore informazione può essere utilizzata per conseguire una migliore classificazione* [Pirlo et al, 2009], [Impedovo e Pirlo, 2011], [Barbuzzi et al, 2012]. Più precisamente, come mostrato in figura (vedi Fig. 1), l'approccio proposto seleziona quei campioni che il sistema multi-esperto, nel suo complesso, classifica con elevata affidabilità per aggiornare il training set di quel singolo esperto che ha erroneamente etichettato gli stessi.

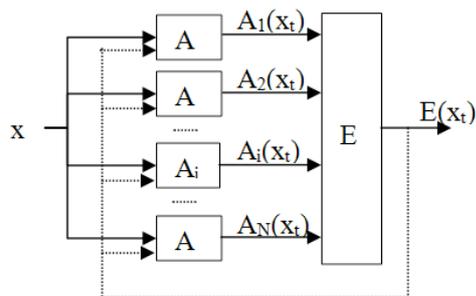


Fig. 1 - Feedback in un sistema multi-esperto con topologia in parallelo

I risultati sperimentali sono stati ottenuti considerando il dataset Wisconsin breast cancer dell'UCI machine learning repository. Essi dimostrano l'efficacia della strategia basata sull'uso del feedback anche rispetto agli algoritmi self-training e co-training reperibili in letteratura [Blum e Mitchell, 1998], [Chapelle et al, 2006], [Rattani et al, 2009], [Yuanyuan et al, 2011].

Questo documento è strutturato come segue. La sezione 2 presenta una breve panoramica delle tecniche di apprendimento semi supervisionato. La sezione 3 introduce l'algoritmo che si avvale dell'informazione fornita dal feedback. Nella sezione 4 sono illustrate le condizioni operative adottate. La sezione 5 presenta i risultati sperimentali ottenuti. Infine, nella sezione 6 sono riportate le osservazioni conclusive e altri temi di ricerca futuri.

2.Panoramica sulle tecniche di apprendimento Semi-Supervisionato

Supponiamo che nuovi dati *non etichettati* siano resi disponibili al trascorrere del tempo. La domanda che ci poniamo è: *come usare questi nuovi dati sconosciuti?*

L'etichettatura e la selezione dei dati rappresenta un processo laborioso e costoso che deve essere compiuto da esperti. Nuove tecniche di apprendimento semi-supervisionato possono essere usate per accrescere il training set di un classificatore e migliorare le performance dell'intero sistema multi-esperto. In letteratura, per raggiungere tale obiettivo, sono stati introdotti due algoritmi denominati: self-training e co-training [Blum e Mitchell, 1998], [Chapelle et al, 2006], [Rattani et al, 2009], [Yuanyuan et al, 2011].

Il self-training (o self-update) si basa sul concetto che un classificatore è riaddestrato considerando i campioni riconosciuti con più alta probabilità nel set dei dati non etichettati. Quindi, il classificatore è inizialmente addestrato adoperando il set dei dati etichettati e, poi, considerando quelle istanze incognite che il classificatore corrente classifica con maggiore sicurezza. Esse contribuiranno ad incrementare il training set (vedi Fig. 2). L'intero processo è ripetuto finché non si soddisfa un criterio di terminazione.

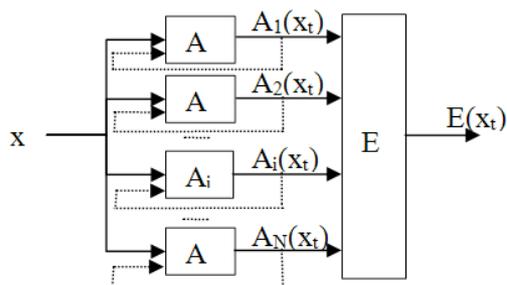


Fig. 2 - Self-Update in un sistema multi-esperto con topologia in parallelo

Il co-training (o co-update) si basa sul concetto che due classificatori possono migliorarsi l'un l'altro nel proprio operare. In particolare, il set delle caratteristiche è partizionato in due subset indipendenti, cosicché per una data classe le caratteristiche usate da un classificatore non possano essere correlate con le caratteristiche usate dall'altro classificatore. In questo modo, il primo esperto è aggiornato con gli elementi riconosciuti dall'altro con elevata affidabilità e viceversa (vedi Fig. 3). Il processo viene ripetuto finché non si soddisfa un criterio di terminazione.

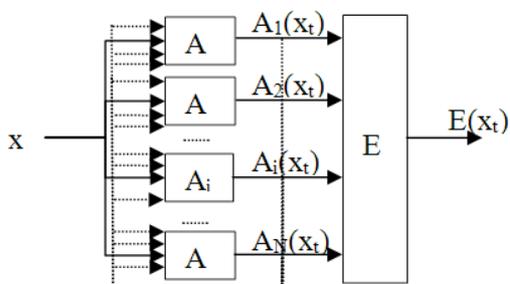


Fig. 3 - Co-Update in un sistema multi-esperto con topologia in parallelo

I criteri di stop sia per il Self-training che per il Co-training sono generalmente due: l'iterazione termina quando tutte le istanze sono state etichettate o quando si è raggiunto un numero massimo di iterazioni.

3.Strategia basata sull'uso del feedback

La strategia proposta costituita da un sistema multi-esperto, con topologia in parallelo, che sfrutta l'informazione fornita dal feedback è descritta di seguito.

3.1 Selezione delle istanze considerando il comportamento collettivo

Sia:

- C_j , for $j=1,2,\dots,M$: il set delle classi;

- $P = \{x_k \mid k = 1, 2, \dots, K\}$: il set dei campioni forniti al sistema multi-esperto (ME). In questo lavoro l'insieme P è partizionato in S sub-set $P_1, P_2, \dots, P_s, \dots, P_S$
In particolare P_1 è usato solo per l'apprendimento e i suoi campioni sono etichettati, mentre
 $P_2, P_3, \dots, P_s, \dots, P_S$
sono usati sia per la classificazione che per l'apprendimento (quando necessario) e i loro campioni sono non etichettati;
- $y_s \in \Omega$: l'etichetta del pattern x_s , $\Omega = \{C_1, C_2, \dots, C_M\}$;
- A_i : l' i -esimo classificatore, $i=1, 2, \dots, N$;
- $F_i(k) = (F_{i,1}(k), F_{i,2}(k), \dots, F_{i,r}(k), \dots, F_{i,R}(k))$: il vettore delle caratteristiche usato dal classificatore A_i per rappresentare il pattern $x_k \in P$;
- $KB_i(k)$: la base di conoscenza del classificatore A_i dopo l'esecuzione del subset P_k . In particolare
 $KB_i(k) = (KB_i^1(k), KB_i^2(k), \dots, KB_i^M(k))$;
- E : il sistema multi-esperto che combina le decisioni dei singoli classificatori per ottenere la decisione finale.

Nella prima fase ($s=1$), il classificatore A_i è addestrato usando i pattern $x_k \in P_i^* = P_1$. Perciò la base di conoscenza $KB_i(s)$ di A_i è inizialmente definita come:

$$KB_i(s) = (KB_i^1(s), KB_i^2(s), \dots, KB_i^1(s), \dots, KB_i^M(s)) \quad (1)$$

Successivamente, i subset $P_2, P_3, \dots, P_s, \dots, P_{S-1}$ dei campioni non etichettati sono forniti uno dopo l'altro al sistema multi-esperto sia per la classificazione che per l'apprendimento. P_S rappresenta il testing set.

Quando consideriamo nuovi dati (istanze di $P_2, P_3, \dots, P_s, \dots, P_{S-1}$), al fine di controllare e sfruttare il comportamento collettivo dei singoli classificatori, si adotta la seguente strategia:

$$\forall x_t \in P_s \ni \left(A_i(x_t) \neq E(x_t) \quad \text{and} \quad \left(\text{score}ME(x_t) \geq \text{threshold} \right) \right) : \text{update_}KB_i \quad (2)$$

In altre parole il training set di A_i è aggiornato con i campioni che sono erroneamente classificati dal singolo esperto e contemporaneamente etichettati dal sistema multi-esperto con un'alta affidabilità.

Vale la pena notare che questa strategia (vedi eq. 2) tiene conto sia delle prestazioni del classificatore individuale che delle prestazioni del sistema ME. Quindi è in grado di selezionare non solo i campioni da utilizzare per il processo di aggiornamento, ma anche i classificatori a cui tali campioni devono essere forniti, in modo da migliorare le prestazioni dell'intero sistema.

3.2 L'algoritmo

Nella procedura basata sull'uso del feedback, ogni esperto è addestrato su un training set, denominato P_1 . Un set di campioni non etichettati P_f è collezionato nel tempo. Tra tutti i campioni di P_f solo quelli classificati con un'alta affidabilità dal sistema multi esperto sono usati per accrescere la base di conoscenza di quegli esperti che hanno prodotto un errore di classificazione. Per calcolare la soglia di affidabilità, P_f è usato come validation set. La regola di riaddestramento rispetta sia la soglia di confidenza Thr^* per ogni esperto (scelta attraverso un processo di *trial-and-error*, pari al valore costante **0.85**) sia la soglia di affidabilità Thr_{ME} per il sistema multi-esperto (pari alla media delle soglie di tutti gli esperti del sistema). Successivamente tutti i campioni riconosciuti correttamente con un'affidabilità uguale o superiore al Thr_{ME} e classificati erroneamente da uno specifico esperto, sono selezionati. Ciò è fatto per ridurre la probabilità di introdurre impostori (falsa accettazione) nell'aggiornamento del training set. Questa procedura è spiegata nel seguente algoritmo (vedi Algorithm 1).

Algorithm 1: Feedback-based Process

1. Given:
 - $P_1 = \{p_1, \dots, p_M\}$: the Initial Training Set
 - $P_f = \bigcup_{u=2}^{S-1} P_u = \{p_{M+1}, \dots, p_N\}$: the new available (unknown) dataset
 - $KB_i(k)$: the knowledge base of the expert A_i , $i=1,2,\dots,N$
 2. For each expert:
 - (a) $A_i(u, P_1)$ is the classification of the sample $u \in P_f$ on P_1
 - (b) $thr^*=0.85$ is the threshold for A_i
 End for
 3. Apply the *combination rule*: $E(u, P_1)$ is the classification of the sample $u \in P_f$ on P_1 combining all experts.
 4. thr_{ME} is the estimated threshold for E
 5. Determine for each sample $u \in P_f$ a classification score of the Multi-Expert system $sME(E(u, P_1))$
-

```

6. For h=M+1,...,N
7. For each expert
  (a) If  $((A_i(u_h, P_1) \neq E(u_h, P_1))$  and
       $(sME(E(u, P_1)) \geq thrME)$  )
  (b) Then  $KB_i(k) = KB_i(k) \cup u_h$ 
End for
End for.
```

4. Condizioni operative

4.1 Il database

In questo articolo è usato il database *Wisconsin breast cancer* dell'UCI Machine Learning Repository. Nello specifico, il dataset è composto da 699 istanze ciascuna con due possibili classi di appartenenza *benigno* e *maligno* (rispettivamente 458 e 241).

Il database è stato inizialmente partizionato in 3 subsets:

- $P_1 = \{x_1, x_2, \dots, x_{15}\}$,
- $P_2 = \{x_{16}, \dots, x_{490}\}$,
- $P_3 = \{x_{491}, \dots, x_{699}\}$.

In particolare, P_1 rappresenta il training set, P_2 il set dei nuovi campioni (non etichettati) e P_3 il testing set.

Ciascun campione è stato partizionato in 3 parti, considerando caratteristiche indipendenti tra essi, che danno origine ai 3 differenti esperti del sistema. In particolare le seguenti caratteristiche sono state considerate in ciascun esperto:

- $F_1 = \{f_2, f_3, f_4\}$: Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape;
- $F_2 = \{f_5, f_6, f_7\}$: Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei;
- $F_3 = \{f_8, f_9, f_{10}\}$: Bland Chromatin, Normal Nucleoli, Mitoses.

f_1 rappresenta l'identificativo del campione e f_{11} rappresenta la classe di appartenenza del tumore (2 per benigno, 4 per maligno).

4.2 Classificatore e tecnica di combinazione

Il sistema usa tre differenti esperti, denominati A_1 , A_2 e A_3 . Ciascun esperto usa un differente set di caratteristiche: A_1 usa il set delle caratteristiche F_1 , A_2 usa il set delle caratteristiche F_2 e A_3 utilizza il set delle caratteristiche F_3 . Tutti i classificatori adottano la regola 3-nearest neighbor per eseguire la classificazione.

Nel sistema multi-esperto, considerando l'approccio proposto, la tecnica di combinazione delle informazioni gioca un ruolo cruciale nella selezione dei nuovi campioni. In particolare, in questo lavoro è stata presa in considerazione la tecnica Majority Vote (MV) per la combinazione delle singole decisioni. MV

considera le etichette fornite da ciascun classificatore, in modo che la decisione finale sia presa in base alla maggioranza delle etichette indicate. Questa tecnica di combinazione è adottata quando non si ha nessuna conoscenza sulle performance dei classificatori.

Per semplicità, consideriamo un sistema ME che adotta tre classificatori base combinati da un semplice MV (vedi Fig. 4). In questo caso, la base di conoscenza di A_1 con X_i sarà aggiornata, mentre il secondo e terzo esperto non saranno aggiornati. In questo caso le performance di A_1 dovrebbero essere migliorate sul training set e un pattern simile ad x_i , sarà correttamente riconosciuto migliorando così la performance dell'intero sistema.

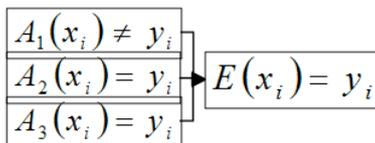


Fig. 4 - Esempio di una richiesta di aggiornamento

5. Risultati Sperimentali

I risultati sperimentali sono riportati in termini di errore percentuale (ER) di riconoscimento sul test set. Inoltre, per le differenti strategie di addestramento, è riportato il numero di campioni selezionati (SS) nel set dei nuovi dati non etichettati.

La tabella 1 mostra i risultati nel caso delle condizioni operative sopra illustrate.

La prima colonna “No-feed” riporta i risultati relativi all’uso di P_1 per il training e P_3 per il test, senza applicare alcuna strategia di feedback (0 sono i campioni selezionati tra i nuovi dati). “MV-feed” indica l’uso della modalità Majority Vote nel processo di feedback a livello di multi-esperto.

In quest’ultimo caso (vedi Tabella 1), l’approccio proposto apporta un miglioramento all’intero sistema di riconoscimento, rispettivamente, dello 0.95% e del 2.39% rispetto agli algoritmi di apprendimento semi-supervisionato proposti in letteratura: self-training e co-training.

Graficamente (vedi Fig. 5) è possibile osservare l’efficienza della strategia con feedback a livello di multi-esperto, che produce un errore di classificazione di sole 2 istanze su 209 del test set.

	No-Feed	Self-Training		Co-Training		MV-Feed	
	ER	ER	SS	ER	SS	ER	SS
A1	5.74	5.74	251	5.26	622	9.57	44
A2	9.57	2.87	304	6.22	569	4.31	44
A3	6.22	3.83	318	7.18	555	3.35	50
MV	5.26	1.91		3.35		0.96	

Tabella 1 - 3-NN: Training set P1, Feedback set P2 e Test set P3

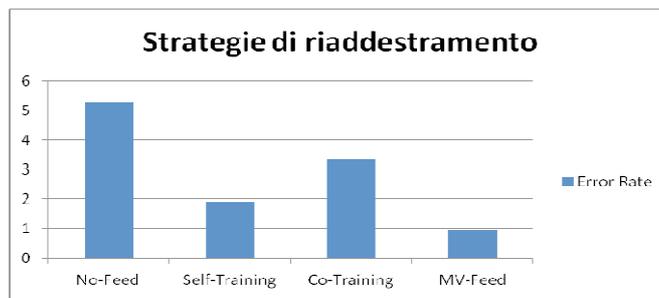


Fig. 5 – Errore di riconoscimento (in percentuale) tra le diverse strategie

Inoltre, in tabella, è mostrato il miglioramento conseguito anche in termine di istanze selezionate rispetto agli algoritmi proposti nello stato dell'arte, quando nuovi dati (non etichettati) siano resi disponibili nel tempo.

Ringraziamenti

Si ringrazia il Dr. William H. Wolberg dell'Università Wisconsin Hospitals, per aver fornito il database Breast Cancer Wisconsin (Original) disponibile online all'indirizzo:

<http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29>.

6. Conclusione e Sviluppi Futuri

In questo lavoro abbiamo proposto un sistema multi-esperto basato sull'uso del feedback che combina differenti insiemi di feature per diagnosticare il cancro al seno. La caratteristica distintiva del nostro sistema è l'impiego di una quantità ridotta di dati etichettati per costruire il modello diagnostico e l'utilizzo di un meccanismo di apprendimento incrementale per migliorare le prestazioni.

In particolare, quando un nuovo set di dati non etichettati è reso disponibile, la decisione finale, ottenuta combinando le decisioni dei singoli esperti, è usata per aggiornare la base di conoscenza di quei classificatori che lo necessitano, in accordo ad una strategia basata sull'uso del feedback.

I risultati sperimentali conseguiti dimostrano che un insieme di classificatori, configurato nel modo proposto, migliora le performance sia in termini di errore percentuale che di istanze estratte dal set dei dati non etichettati. Questo miglioramento ottenuto è superiore rispetto a quello raggiunto dagli altri algoritmi di apprendimento semi supervisionato proposti in letteratura.

Per migliorare ulteriormente il sistema presentato sarà necessario condurre ulteriori ricerche che prevedano non solo l'adozione di differenti tipi di classificatori, ma anche di diverse modalità di combinazione delle informazioni fornite dai classificatori stessi. Inoltre, si dovrà approfondire la possibilità di automatizzare il processo di individuazione delle soglie dei singoli esperti (Thr^*) e dell'intero sistema multi-esperto (Thr_{ME}) per la selezione di una nuove istanze.

Bibliografia

- [RIF 1] BarbuZZi D., Impedovo D. and Pirlo G. (2012). Benchmarking of Update Learning Strategies on Digit Classifier Systems . In: Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition. Capitolo, Bari, September 18-20, 2012, p. 35-40.
- [RIF 2] Blum A. and Mitchell T., "Combining Labeled and Unlabeled Data with Co-Training". ACM Proc. of the 11th annual Conference on Computational Learning Theory, pp. 92–100, 1998.
- [RIF 3] Brenner H., Long-term survival rates of cancer patients achieved by the end of the 20th century: a period analysis. *Lancet*. 360:1131–1135, 2002
- [RIF 4] Chapelle O., Scholkopf B. and Zien A., "Semi supervised Learning," The MIT Press, 2006
- [RIF 5] Impedovo D. and Pirlo G., "Updating Knowledge in Feedback-based Multi-Classifier Systems", in IEEE proc. of International Conference on Document Analysis and Recognition (ICDAR2011), pp. 227-231, 2011.
- [RIF 6] Kittler J., Hatef M., Duin R.P.W. and Matias J., "On combining classifiers", *IEEE Trans. on Pattern Analysis Machine Intelligence*, Vol.20, no.3, pp.226-239, 1998.
- [RIF 7] Pirlo G., Trullo C.A. and Impedovo D., "A Feedback-Based Multi-Classifier System", *IEEE proc. of International Conference on Document Analysis and Recognition (ICDAR 2009)*, pp. 713-717, 2009.
- [RIF 8] Rattani A., Freni B., Marcialis G.L. and Roli F., "Template update methods in adaptative biometric systems: a critical review", *proc. of International Conference on Biometrics, LNCS*, vol 5558, pp. 847-856, 2009.
- [RIF 9] Subashini T., Ramalingam V. and Palanivel S., "Breast mass classification based on cytological patterns using RBFNN and SVM" *Expert Systems with Applications*, 36(3): p. 5284-5290, 2009.
- [RIF 10] Yuanyuan G., Zhang H. and Xiaobo Liu: Instance Selection in Semi-supervised Learning. *Canadian Conference on AI 2011*: 158-169

UBI-CARE: social learning la gestione delle malattie croniche

Flora Berni, Pierpaolo Di Bitonto, Francesco Di Tria, Teresa Roselli, Veronica Rossano

Dipartimento di Informatica - Università degli Studi di Bari

Via E. Orabona, 4 - 70125 Bari

{flora.berni, dibitonto, francesco.ditria, teresa.roselli, veronica.rossano}@uniba.it

Ubicare is a social networking framework, based on semantic ontology and knowledge management techniques to support new healthcare models implemented through the Hub&Spoke paradigm. This paradigm provides for the concentration of high-complexity care in centres of excellence (Hub) supported by a network of local services (Spoke) which is responsible for sorting out patients, and for sending them to the reference centres only when a certain threshold of severity of clinical care is exceeded. The platform is characterized by ubiquity and pervasiveness with respect to the different “places” and “levels” of service.

Keywords: e-health, social learning, ubiquitous computing

1. Introduzione

L'esigenza di tenere sotto controllo la spesa sanitaria, la volontà di mantenere un servizio assistenziale efficiente con alti standard qualitativi e l'aumento delle malattie croniche hanno condotto negli ultimi anni alla sperimentazione di nuovi modelli di cura. Tali modelli spostano il focus dall'ospedale, più opportunamente deputato al trattamento delle patologie acute, verso la medicina generale e le strutture territoriali. Le esperienze più recenti per la cronicità riguardano principalmente la cardiologia, anche se si stanno diffondendo per pneumologia, nefrologia, diabete, oncologia. Il modello di servizio spesso si basa sulla costituzione di reti integrate di cura che consentono la gestione clinica in remoto dei pazienti grazie alla combinazione di tecnologie ICT: sistemi di monitoraggio a distanza, sistemi di condivisione dei dati sanitari e di protocolli di cura. Dal punto di vista più strettamente organizzativo questo nuovo paradigma dei servizi di assistenza sanitaria viene implementato attraverso modelli detti “Hub&Spoke” che prevedono la concentrazione dell'assistenza di elevata complessità in centri di eccellenza (Hub) supportati da una rete di servizi (Spoke) cui compete la selezione dei

pazienti e il loro invio a centri di riferimento, quando una determinata soglia di gravità clinico-assistenziale viene superata. Il modello prevede la creazione di network organizzativi che, per essere implementati, necessitano di collegamento e di una “condivisione di conoscenze” tra i diversi “luoghi” di assistenza. È evidente che lo sviluppo di applicazioni innovative di knowledge management e ubiquitous computing può, in tale contesto, essere il “fattore abilitante”.

In questo contesto nasce UBICARE con l'obiettivo di creare una rete di esperienze e conoscenze attraverso la quale

- condividere i dati clinici e favorire le sinergie fra i diversi livelli assistenziali (ambulatori ospedalieri, distretti socio-sanitari, studi di medicina generale);
- supportare il personale medico e paramedico nella diagnostica e nel monitoraggio del paziente (anche a distanza);
- educare il paziente ad uno stile di vita consono allo stato di salute in cui versa;
- formare il personale medico e paramedico sulle procedure di diagnosi, agli interventi terapeutici e al follow-up dei pazienti sia in modo “situato” ovvero durante lo svolgimento delle attività lavorative che in momenti appositamente dedicati alla formazione.

Il progetto mira a sviluppare un sistema pilota da sperimentare in due contesti reali: lo scompenso cardiaco cronico e la dialisi peritoneale.

Il presente contributo intende presentare un aggiornamento sullo stato avanzamento lavori del progetto già descritto in [1].

2. Scenario di riferimento

L'organizzazione del sistema sanitario regionale nel corso di questi anni, e sempre più nel corso dei prossimi sta virando verso il modello hub-spoke che prevede in caso di malattie croniche gravi o che comunque richiedono un monitoraggio continuo del malato (es. scompenso cardiaco cronico, trapianti, ecc.) pochi centri di eccellenza (hub) per la cura della sintomatologia acuta del problema e numerosi centri spoke (medici generici, personale infermieristico del distretto) per il follow-up del paziente. Tale organizzazione consente di contrarre i costi pubblici ottimizzando l'uso delle risorse demandando al medico di famiglia e/o al personale infermieristico nuovi compiti legati al trattamento della specifica patologia. Detti professionisti, tuttavia, non sempre sono in grado di adempiere a tali compiti a causa della scarsa competenza specifica. Tale gap di conoscenza, non può essere colmato solo per mezzo di attività formative tradizionali (comunque indispensabili), pertanto è necessario mettere a disposizione dei medici e del personale infermieristico strumenti ICT in grado di supportarli, di guidarli nelle decisioni, di facilitarne la collaborazione con le altre figure coinvolte nello scenario (pazienti, famiglie), di facilitarne l'accesso a risorse formative ed informative just in time.

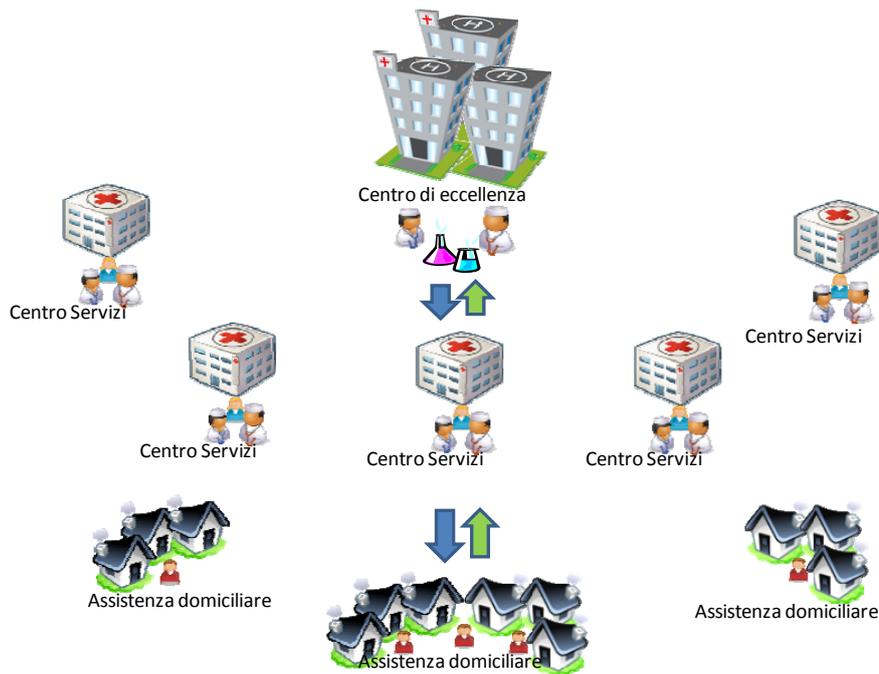


Fig.1 – Il modello hub & spoke

Dopo un'attenta analisi di questo scenario ed in particolare delle istanze provenienti dalle patologie croniche gravi è emerso che la soluzione tecnologica da realizzare deve:

- creare una rete di esperienze e conoscenze attraverso la quale condividere i dati clinici e favorire le sinergie fra i diversi livelli assistenziali (ambulatori ospedalieri, distretti socio-sanitari, studi di medicina generale);
- supportare il personale medico e paramedico nella diagnostica e nel monitoraggio del paziente;
- formare il personale medico e paramedico relativamente alle procedure di diagnosi, agli interventi terapeutici e al follow-up dei pazienti sia in modo "situato" ovvero durante lo svolgimento delle attività lavorative che in momenti appositamente dedicati alla formazione.

A tal fine, il sistema è in grado di supportare le decisioni in ambito medico per superare in tempo reale il gap di conoscenza esistente nei centri spoke fra la conoscenza posseduta (dal personale medico e paramedico) e la conoscenza necessaria per affrontare una specifica situazione clinica; supportare azioni di formazione per ridurre al minimo il gap delle conoscenze fra il personale del centro hub e quello del centro spoke; gestire e condividere la

conoscenza non solo relativa alle tematiche cliniche ma anche in termini di scambio di esperienze e condivisione di pratiche tra le famiglie dei pazienti.

3. La soluzione tecnologica proposta

Il modello architetturale scelto per il sistema integrato UBICARE è un'architettura ibrida. Il sistema è una social network che offre servizi e funzionalità finalizzati sia alla gestione remota di pazienti "cronici" che necessitano di una gestione ad elevata complessità, sia alla formazione a distanza delle varie figure che supportano e gestiscono il paziente. Per la realizzazione del sistema integrato UBICARE si è scelto, piuttosto che realizzare una social network ex-novo, di personalizzare un sistema open source. In questo modo si sfruttano le funzionalità già realizzate del social network engine e si focalizza il lavoro sulla personalizzazione e l'ampliamento delle funzionalità in linea con quanto richiesto dagli scenari applicativi. Il social network engine scelto per la realizzazione del sistema integrato UBICARE è ELGG una piattaforma open source facilmente personalizzabile e che ha una architettura a plug-in. In particolare, ELGG adotta come programming model lo standard MVC (Model-View-Controller) che consente la separazione dei compiti fra tre diversi livelli

- il *model* fornisce i metodi per accedere ai dati utili all'applicazione, estraendoli dal database e creando delle rappresentazioni ad oggetti
- il *view* visualizza i dati contenuti nel model e si occupa dell'interazione con utenti e agenti
- il *controller* riceve i comandi dell'utente (in genere attraverso il view) e li attua modificando lo stato degli altri due componenti

Questo schema, fra l'altro, implica anche la tradizionale separazione fra la logica applicativa a carico del *controller*, accesso ai dati a carico del *model*, e interfaccia utente a carico del *view*.

I moduli sono quindi sviluppati come plug-in di ELGG, in modo da comunicare fra di loro attraverso i differenti modelli di dati messi a disposizione. Tutti i plug-in, inoltre, possono accedere ad un database condiviso (il DB di ELGG stesso opportunamente esteso per contenere gli ulteriori dati necessari). In particolare, alcuni plug-in hanno il compito di popolare il database attraverso funzioni di acquisizione dati del paziente, altri invece hanno il compito di visualizzare tali dati a scopo di monitoraggio e di analisi.

Nell'ambito del sistema integrato UBICARE ci sono due componenti software esposte sotto forma di servizio web, usando quindi un'architettura SOA: il Sistema di Supporto alle Decisioni e il Data Service Gateway.

Il Sistema di Supporto alle Decisioni (DSS) offre servizi a tutti gli altri plug-in. A differenza delle altre componenti del sistema integrato, infatti, il sistema di supporto alle decisioni è un modulo indipendente il cui scopo è quello di fornire supporto alla diagnosi, ovvero fornire all'utente un suggerimento su uno specifico problema di diagnosi. Il sistema di supporto alle decisioni, inoltre, fornisce alla componente di gestione della formazione i profili clinici realistici utili per il training dell'utente.

Il Data Service Gateway (DSG) fornisce i dati clinici del paziente provenienti dalle basi di dati già esistenti a livello territoriale e nazionale da integrare nel database di ELGG. I dati clinici provenienti dai database esterni e i dati acquisiti giornalmente tramite il modulo di acquisizione dati sono conservati nella cartella clinica integrata del paziente che è oggetto di monitoraggio. Il DSG, ad un livello di astrazione molto elevato, è un database virtuale per l'integrazione dei dati.

3.1 Il sistema integrato UBICARE

L'analisi dettagliata dello scenario di riferimento e l'analisi dei pattern architetturali da utilizzare hanno dato come risultato l'architettura del sistema integrato UBICARE come rappresentata in figura 2.

L'Health care network è l'interfaccia multimodale dell'intero sistema. Grazie all'approccio a plug-in consente un'elevatissima scalabilità in termini di servizi erogabili. Il sistema integrato UBICARE utilizza l'Open Source Social Networking Engine (ELGG) per la sua capacità di fornire un framework solido su cui costruire un ambiente sociale. Inoltre, l'uso di un framework open source consente di riutilizzare le funzionalità tipiche di una piattaforma di sociale e di focalizzare gli effort di progettazione e sviluppo sulle componenti che risolvono specifici problemi dello scenario di riferimento e sulla personalizzazione di servizi già esistenti.

In particolare, le componenti già integrate di ELGG utili per il sistema integrato UBICARE sono i tool di comunicazione e la componente di profilazione che è personalizzata al fine di rispondere ai requisiti definiti per la componente di Context Aware & Personalization.

Il plug-in Context Aware & Personalization incapsula lo user model costruito dal framework ELGG e lo mette a disposizione di tutti i plug-in consentendo, in tal modo, la condivisione delle informazioni sull'utente. In particolare, la componente invia all'occorrenza ai differenti plug-in i dati inseriti manualmente dall'utente o acquisiti automaticamente dal framework (dispositivo utilizzato per l'interazione, posizione geografica, dati utente, attività svolte) consentendo la contestualizzazione e la personalizzazione del servizio.

In particolare, le componenti già integrate di ELGG utili per il sistema integrato UBICARE sono i tool di comunicazione e la componente di profilazione che sarà personalizzata al fine di rispondere ai requisiti definiti per la componente di Context Aware & Personalization.

Il plug-in Context Aware & Personalization incapsula lo user model costruito dal framework ELGG e lo mette a disposizione di tutti i plug-in consentendo, in tal modo, la condivisione delle informazioni sull'utente. In particolare, la componente invia all'occorrenza ai differenti plug-in i dati inseriti manualmente dall'utente o acquisiti automaticamente dal framework (dispositivo utilizzato per l'interazione, posizione geografica, dati utente, attività svolte) consentendo la contestualizzazione e la personalizzazione del servizio.

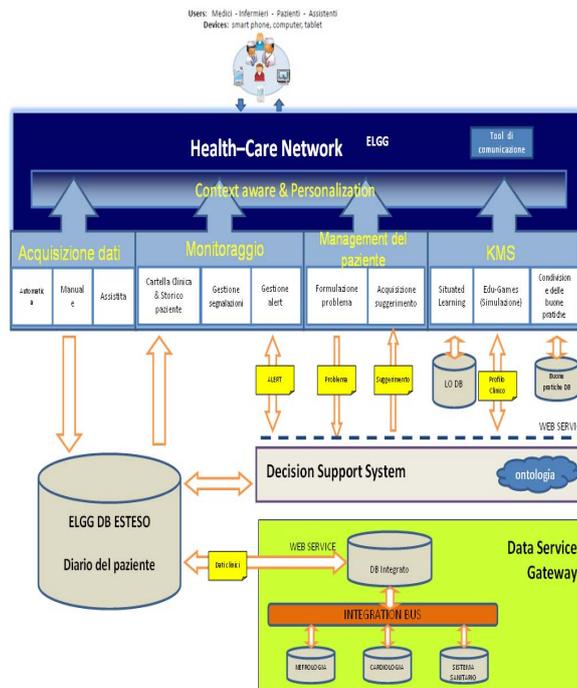


FIGURA 2: Architettura del sistema integrato UBICARE

Il plug-in Acquisizione dei dati paziente è un plug-in della piattaforma ELGG che consente l'acquisizione dei dati utili alla gestione e al monitoraggio del paziente. Le modalità di interazione per l'acquisizione dei dati saranno personalizzate al fine di rispondere ai seguenti requisiti:

- Acquisizione dati automatica: acquisizione giornaliera automatica dei parametri di peso, frequenza cardiaca, pressione arteriosa, diuresi ed ultrafiltrazione. Per i pazienti che al domicilio seguono uno schema di dialisi peritoneale automatizzata con l'ausilio di cycler capaci di effettuare automaticamente, durante le ore notturne, gli scambi dialitici, è possibile che tali apparecchiature inviino via modem i dati relativi alla seduta dialitica con allarmi o problemi tecnici rilevati durante la stessa (Scenario di Nefrologia)
- Acquisizione giornaliera automatica dei parametri di peso, frequenza cardiaca, pressione arteriosa, ECG mediante sensori in dotazione del paziente con sistema di allarme in caso di alterazioni di rilievo (Scenario di Cardiologia)
- Acquisizione mediante device impiantabili. Si basa sui dati acquisiti da dispositivi impiantabili chiamati cardiovertitori defibrillatori automatici. Questi sistemi sono utilizzati per prevenire la morte improvvisa dei pazienti; vengono impiantati sotto cute nella regione al disotto della clavicola e sono collegati al cuore da cateteri ovvero dei "fili" elettrici

capaci di stimolare l'attività elettrica del cuore e di erogare shock elettrici. Al tempo stesso tali sistemi sono in grado di acquisire rilevanti informazioni che vengono automaticamente inviate attraverso dei trasmettitori in possesso del paziente, collegati alla linea telefonica, al centro che segue clinicamente il paziente (Scenario di Cardiologia).

- Acquisizione dati assistita: acquisizione telefonica con questionario predefinito da parte dell'infermiere;
- Acquisizione telefonica con voce preregistrata e con segnalazione al centro scopenso o al centro ospedaliero di nefrologia in caso di alterazioni significative segnalate;
- Acquisizione dati via video chat che consentano al paziente di comunicare i dati all'infermiere o al medico che le inserisce contestualmente nel sistema;
- Acquisizione dati manuale: acquisizione parametri monitorati giornalmente da paziente e familiari. I dati sono rappresentati rispettivamente da peso, pressione arteriosa e frequenza cardiaca, nello scenario cardiologico, e da peso, frequenza cardiaca, pressione arteriosa, diuresi ed ultrafiltrazione, nello scenario nefrologico. In entrambi i casi deve essere possibile inserire le variazioni significative dei sintomi e dei segni clinici che indicano un peggioramento delle condizioni del paziente.
- Acquisizione parametri monitorati dal paziente via email inviate giornalmente che richiedono la compilazione di un form per la rilevazione rispettivamente del peso, della pressione arteriosa e della frequenza cardiaca, nello scenario cardiologico, e del peso, della frequenza cardiaca, della pressione arteriosa, della diuresi e dell'ultrafiltrazione, nello scenario nefrologico. In entrambi i casi deve essere possibile inserire le variazioni significative dei sintomi e dei segni clinici che indicano un peggioramento delle condizioni del paziente.

I dati raccolti con i processi di acquisizione sono conservati nel database di ELGG dove sono integrati con i dati provenienti dai sistemi informativi con cui il sistema integrato UBICARE interagisce.

Il plug-in Monitoraggio del paziente integra i dati del paziente memorizzati all'interno del database di ELGG con i dati provenienti dai sistemi informativi con cui il sistema UBICARE interagisce per creare una cartella clinica integrata che contenga i dati aggiornati della situazione clinica del paziente.

I dati della cartella clinica integrata, così creata, sono utilizzati dal Servizio di Supporto alle Decisioni (DSS) consentendo la Gestione degli alert, ovvero segnalazioni automatiche di alterazioni cliniche significative al personale medico e paramedico. La modalità di invio degli alert dovranno essere diverse e scelte dall'utente (sms, alert visivo sulla pagina dei pazienti, email).

Il monitoraggio del paziente deve anche prevedere la possibilità per il paziente, i familiari o il medico di medicina generale di inviare segnalazioni al

centro ospedaliero in caso le alterazioni cliniche significative non siano state rilevate automaticamente dalla cartella clinica integrata.

4. Conclusioni

Il progetto UBICARE intende creare attorno ai pazienti affetti da patologie croniche una rete assistenziale in grado di migliorare le conoscenze di coloro i quali gravitano attorno al paziente in modo da migliorare l'assistenza che è possibile offrire. I principali elementi di originalità e novità introdotti dal progetto riguardano le modalità con cui la diffusione della conoscenza in ambito medico è attuata nel territorio e le modalità con cui essa è gestita. Per quanto riguarda la diffusione della conoscenza in ambito medico, intesa come formazione e supporto alle attività lavorative, l'innovatività delle soluzioni proposte è documentata dalle evidenze sperimentali rintracciabili nei contesti nazionali ed internazionali. In [2] sono riportati i risultati di uno studio che valuta come estremamente efficace l'approccio della simulazione in situazioni cliniche complesse in quanto in grado di colmare il divario tra formazione teorica e pratica, consentendo l'acquisizione della capacità di prendere decisioni in tempi ridotti. La simulazione è stata ritenuta utile anche come strategia di training all'uso di strumentazione medica come già dimostrato dalla mySmartSimulations Company (<http://www.mysmartsim.com/index.asp>).

In Italia nello scorso anno è stata fondata a Firenze la Società Italiana di Simulazione in Medicina (SIMMED) <http://www.simmed.it/> che usa la piattaforma DrSim® per creare casi clinici in ambienti virtuali realistici. L'approccio dell'ubiquitous e context-aware learning è stato utilizzato in [3] nel reparto di Oncologia di University Hospital of North Norway per rendere più efficace la comunicazione tra i medici e per fornire ai medici un supporto in tempo reale all'uso delle diverse attrezzature sanitarie. Infine, l'uso delle social network per la condivisione di informazioni e la collaborazione per la soluzione di casi difficili è ormai una realtà molto diffusa e che ha dato ottimi risultati in termini di velocità nella formulazione delle diagnosi.

Per quanto riguarda la gestione della conoscenza, i principali elementi di novità sono la mancanza di una fase di training del sistema e l'utilizzo della modellazione trasparente ed automatica in logica descrittiva della storia clinica del paziente. Ciò permette l'utilizzo del sistema senza possedere alcuna conoscenza dei principi di intelligenza artificiale né di avere una preparazione informatica di alto livello. Al contempo, la qualità degli output del sistema permetteranno di trarre reale beneficio dal processo inferenziale specie nel caso di casistiche fortemente complesse oppure di notevoli moli di dati da esaminare.

Bibliografia

[1] F. Berni et al. UBI-CARE: ubiquitous learning per medici e pazienti. ISBN: 978-88-905406-7-7; (2012). In A. Andronico, A. Labella, F. Patini (Eds.): Didattica 2010 – Tecnologie Informatiche per la Didattica.

[2] Hourican, Susan and McGrath, Mary and Lyng, Colette and McMahon, Caron and Lehwaldt, Daniela (2008) Effectiveness of simulation on promoting student nurses management skills. *International Journal of Clinical Skills*, 2 (1). pp. 20-25. ISSN 1753-044X

[3] J. Ting, A. Tsang, W.H. Ip (2011) Coheals: A Context-aware Helthcare Learning System for Medical Equipments Usage Education. *INTED2011 Proceedings*, pp. 1928-1932, ISBN: 978-84-614-7423-3

Acknowledgement

Il progetto Ubicare è stato finanziato nell'ambito del programma di finanziamenti della Regione Puglia (Programma Operativo Regionale FESR 2007-2013 - Obiettivo Convergenza - ASSE I - Linea 1.2 Azione 1.2.4 - Aiuti a sostegno dei partenariati regionali per l'innovazione).

The PP&S100 Project: Process Control as an Information System Instance

Claudio G. Demartini¹, Marina Marchisio², Marco Mezzalama¹,
Claudio Pardini³, Amelio Patrucco⁴

¹Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino
claudio.demartini@polito.it, marco.mezzalama@polito.it

²Università di Torino, Via Carlo Alberto, 10, 10124 Torino,
marina.marchisio@unito.it

³IIS Carlo Anti, Via Magenta, 7 – 37069 Villafranca di Verona,
dirigente@carloanti.it

⁴Fondazione Torino Wireless, Via Francesco Morosini, 19 – 10128 Torino
amelio.patrucco@torinowireless.it

Abstract *The Project PP&S100, Problem Posing & Solving, is part of a set of initiatives promoted by the General Directorate of the Italian Ministry of Education, Research and University, for supporting the many innovations that have recently affected curricula at the upper secondary level. Main goals of the project include strengthening computer science culture and enhancing its role as a scientific discipline; founding education processes on logics, mathematics and computer science weaved together to pursue an interdisciplinary scenario, building a culture of "problem posing and solving" by investing across a broad disciplinary group of subjects, ensuring growth of computer science based training of trainers and practicing activities within specific social networks and virtual learning environment to share learning materials, teaching supplies, mentoring and self-evaluation. This work presents a hierarchical conceptual model for "Computational Thinking" assumed by PP&S100, formalized according to the outcomes shown by information systems courses held at the Politecnico di Torino and validated using Process Control as an information system instance.*

Keywords: Problem Posing&Solving, Computer Science, Computational Thinking, Process Control, Collaborative Learning.

1. Introduction

The purpose of the project, which gathers among the primary proponents also the Italian Association for Automatic Computing (AICA), the National Research Center (CNR), the National Industrial Unions, the University of Turin

Congresso Nazionale AICA 2013

and the Politecnico di Torino, is to sustain changes proposed at the regulatory level transforming "programs specified by the Ministry of Education" into national Guidelines for high schools, technical schools and vocational organizations. This change relies mainly on teachers for a clearer knowledge management and independent design of educational routes most suitable for achieving learning outcomes that guidelines declare and set for the entire national territory. The Project PP&S100 [Palumbo and Zich, 2012], Problem Posing & Solving for 100 schools, now grown to 150, is part of a set of initiatives promoted by the General Directorate of the Italian Ministry of Education, Research and University, for supporting the many innovations that have recently affected curricula at the upper secondary level.

It has an immediate impact in the scientific application domains (physics, chemistry, natural sciences and so forth) but also brings a relevant potential within all subject areas, even those placed in the social and humanistic domains, thanks to the basis provided by trans-disciplinary languages and computer science representations.

The project culturally focused on problem posing and solving, aims to tap the invaluable potential of information technology as an enabler of innovation. This already takes place when mathematics and computer science should strengthen one another, being delivered jointly in the same classes, as it is the case in the first two years at high school, where they should work both enforcing mutual outcomes - even if, for truth's sake and most of the times - this happens only on a formal base.

Main goals of the project include:

- strengthening computer science culture, enhancing its role as a scientific discipline, by introducing specific courses in the first two years where they are not now given
- set up education processes integrating logic, mathematics and computer science
- building a culture of "Problem posing & solving", by investing across the broad disciplinary group, within the various curricula, adopting a systematic activity based on the use of the logical-mathematical computer science for the formalization, quantification, simulation and analysis of problems of various complexity.
- ensuring growth of computer science-based training of trainers called to go with the transformation process.
- adopting a relevant amount of social network activity and virtual learning environment, by sharing learning materials, teaching supplies, mentoring and self-evaluation.

The project, lasting three years, has planned to focus the goals looking at the first two couples of years of high school. Its working organization has planned to involve at least 110 institutes of secondary education, which should act as "poles" of territorial networks. At the end of the first year, 150 classes of about 4500 students had already been engaged.

Since the beginning the Project addressed Mathematics and new technologies, also dealing with learning management with an advanced computer environment (ACE). The latter played a basic role to build new

teaching plans in Maths with respect to the framework drawn by problem posing and solving. The Science Faculty of the University of Turin made a great experience in this field and, thanks to it, put in place a Moodle asset and a Maple suite for the project itself. Both of them, had been built tightly interlaced and were successfully used for several years in teaching Mathematics for all the students coming from the same Faculty, to whom a part of those attending other courses joined soon.

The whole infrastructure started first with a Moodle virtual learning environment, always available at the address <http://minerva.i-learn.unito.it/>. Then that Moodle platform grew integrating a Maple Suite, consisting of an advanced symbolic computing environment and MapleNet, supporting file distribution in such a way that interaction remains active even if Maple is not running locally on the computer. MapleTA added to soon, being a tool apt to develop tests and assignments, including definition of open mathematical questions and self-evaluations, and, in a while, MapleSIM followed, being useful to carry out virtual labs in Physics, Electronics, Mechanics, Thermodynamics and so forth. The whole system was able to sustain more than 150 teachers, which gave birth soon to a Community, taking advantage of the Moodle platform, which made a strong collaborative learning scenario possible, fully and constantly practiced.

An effective tutoring and coaching service was also established addressing teacher learning to pursue the constructivist approach [Brooks, & Brooks, 1999] in the classroom and to support products development to be shared on the Moodle platform. Since February 2013 more than 150 students communities grew in the project, where they could learn according to a multifaceted perspective: at school during their classes and in the labs, but also at home, by trying to solve problems together, understanding them first, according to the posing perspective, and choosing a suitable algorithm to carry out a possible solution, most of the times working inside the learning management platform as reported in [Zich et al].

A few classes – belonging to 10 high schools and 10 technical institutes, selected on a voluntary basis – are planned to start in the second project year (2013/2014). They are to provide for the breeding ground to put in place a proposal for a new discipline intended to deepen information and computer science: a first step towards the creation of a process able to deploy the same experience after the project ending. The discipline is being organized according to a framework articulated into two lab modules spread over two years of high school curricula.

Section 2 introduces emerging scenarios for computer science in the school, focusing on the layer abstract model assumed as a reference in the project. Section 3 describes the Computational Thinking Living Lab arguing on Process Control as an information system instance to validate the abstraction model, while Section 4 reports on concluding remarks and future work.

2. Learning Computer Science in Education: Emerging Scenarios

Inquiry based Science Education and Learning [Olson and Loucks-Horsley, 2011] can be considered as the result of different contributions given by several domains such as, among others, "constructivism" [Brooks and Brooks, 1999] [Marlowe and Page, 2005], "Bloom's taxonomy of learning" [Forehand, 2005] and "multiple intelligences", without forgetting Paulo Freire's "Pedagogy of the Oppressed" [Freire, 1993], where the terms "problem posing" occurred for the first time.

Problem posing and solving fits well the computer science discipline, which is able to give proper means, tools, methodologies and theories to describe real and conceptual phenomena. This is mainly because logics, formal languages and physical elementary devices are deeply rooted in the founding discipline principles.

A glance at the rest of the world [Peyton-Jones et al, 2011] shows that many other countries have expressed the same creative impetus that already hit England some years ago [Livingstone and Hope, 2011]. In fact, excitement around computing led UK to introduce this subject into schools curricula, making of that event an effective representative experience for all other countries. In particular, European countries [Gander et al, 2013], together with many others around the world, are now trying to collect information from UK, because they came to share what U.S.A. anticipated in their White Paper since 2005:

- a clear distinction between computer science as a discipline and rigorous computer applications or digital literacy should prevail
- recognize that all children and young students should learn information science in the same way they learn science or mathematics.
- be aware that growing demand for ICT professionals from the labor market reduces the potential supply of qualified teachers.

2.1 Computational Thinking

Computational Thinking (CT) is often elicited as a best practice of inquiry based science approaches, as defined by J. Wing: "*Computational thinking is the reasoning processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent*" [Wing, 2008]. This definition can be assumed as a reference point with respect to the objectives pursued by the PP&S100 project, as the terms "formulating problems" and "formulating solutions" suggest.

The most important complex mechanism in CT is the abstraction process, which is used, well deepened and represented in Unified Modeling Language UML or Entity Relationship (ER) model, to define patterns, specializing entities or generalizing others, leaving from objects or other simpler entities ("generalization" abstraction and its converse, the so-called "specialization"). It is also used to capture essential common properties of a set of objects applying an abstraction procedure named "classification". Also an algorithm is usually

given as an abstraction of a process having a determined objective, which operates executing a sequence of steps, working on a set of inputs to produce a corresponding set of outputs. A process itself can be considered an abstract data type representing a set of values and operations designed to manipulate those values, which are not visible from outside of that abstract data type representation. From that drawing comes out that efficient algorithms cannot be designed without using appropriate abstract data types to handle input data or internal temporary ones to be provided to the same process. Well depicted by both UML and ER models at least, abstraction gives analysts the power to deal with complexity scaling the representation of reality, leaving from the bit representation, as the lowest description level, up to the largest system you can find in reality. The hierarchical abstraction design follows a hierarchical sequence of layers, each one featuring its own specific abstraction level, thus recalling the communication domain where a layered architecture allows segmentation of services assigning specific functionalities to any layer to keep it disjointed from the adjacent ones. This enhances modularity, in fact when implementing or even changing functionalities, a specific layer is chosen to focus on, without affecting others, since adjacent interactions take place through specific inter-layer services. The loosely structured architecture of the Internet shows how effective and robust may be a well-formed abstraction. An example is provided by the TCP/IP layer, which, placed over the network and under the session layer has enabled a multitude of applications to proliferate at layers above, and a multitude of different platforms, media, and devices to operate at lower layers.

Fig. 1 depicts the model drawn to introduce the rationale of the computer science discipline to be developed in the project, fully adopting the “Inquiry based Science Education model” and its specialization represented by the “Computational Thinking” scenario. The most abstract modeling phase is placed at layer 4 (I4), the Problem Posing Layer, where the model is shaped through an inquiry-based process focusing on main questions linked to the problem to be understood. To give a flavor of that layer functionality, the Logical Framework Analysis [EU, 2004] has been recalled, to emphasize the fact that a collaborative interaction should also be taken into account and put in



Fig. 1 – CT and the Layer Abstraction Model

place among all the stakeholders involved in the problem.

The next layer (I3) deals with the specification of the solution to be proposed for the problem posed at the previous one. Algorithms and data representation should be given there according to a formal abstract description, being they better shaped using specific environments, such as MAPLE/MAPLE SIM or

MATLAB/SIMULINK [Chonacky and Winch, 2005], LABVIEW [National Instruments, 28/8/2013] and UML [Wikipedia, 28/8/2013].

MAPLE, MATLAB and LABVIEW, together with their linked applications, may be considered as appropriate examples of Advanced Computation Environment (ACE) to specify mathematical models for complex phenomena. In parallel, for non-mathematical domains, in a complementary perspective with respect to ACE, the Unified Modeling Language [Unified Modeling Language, 28/8/2013] and IDEF (Icam DEFinition for Function Modeling, where 'ICAM' is an acronym for Integrated Computer Aided Manufacturing) [IDEF0, 28/8/2013], are also introduced. Thus a multi-faceted model description can be carried out, representing the same applications/phenomena by means of use-cases and activity diagrams to describe functionalities, class diagrams for static objects description, sequence diagrams for tracing object dynamic interactions and so forth. Both scenarios, mathematical and non-mathematical, are also able to sustain high-level simulations. In the most cases both environments can also translate given formal models descriptions into a first draft of the programming language description for the same systems/phenomena. These descriptions are handled at layer two (I2), where a more detailed picture can take care of the various aspects also linked to the hardware platform to be used subsequently at layer one (I1). At level two more precise simulations can also be planned, being given the infrastructure already traced in terms of operating systems, network features, file system and so forth, relying also on compilers (C, C++) or interpreters (Python, Scratch) to correctly implement the semantics of the language. To cite an example for all in terms of CASE tools, Visual Paradigm [Visual Paradigm, 28/8/2013], supporting also ER-Model based environments, and Eclipse [Eclipse Project, 28/8/2013] could be used as a specific reference. Hence, as shown by the experience already made at the Information Systems Course at Politecnico, conceptually showed in Fig. 1, Computational Thinking stems from both mathematical and engineering thinking. Unlike mathematics, where constraints take the shape of theorems, corollaries, assumptions, information systems are constrained by the physics of the underlying information-processing units and their related operating environment. This means that boundary conditions, failures, malicious agents, and unpredictability of the real world must also be taken into account, having in mind that, unlike other engineering disciplines like Mechanics or Constructions, Information Science — based on intangible artifacts like software, allows anyone to develop virtual worlds totally unconstrained by physical issues, in fact in cyberspace the only constraint to creativity may be given by imagination.

3. “Computational Thinking” Living Lab and the TestBed

This work deals with the development of a complete test-bed that aims at stating the feasibility of the learning process, based on previously exposed methodologies, when applied to a very simple but relevant context. The label assigned to this experience is “Living Lab”, focused on enhancing the relevance of computer science in the school as a means to represent reality with respect to both static and dynamic perspectives.

The Lab rationale has the following main points:

- check how learners can leverage learning approaches based on Problem Solving and Computational Thinking - being the latter a learning domain emerged recently and often considered in literature as a specialization of Problem Solving - when applied to scientific subjects in the secondary school to create a continuum from mathematics, to computer science, up to other "mathematics-based sciences".
- check how Simulation and Prototyping can act as "accelerators of understanding" of complex systems and physical phenomena.
- prepare a draft syllabus for the first module of Computer Science for the secondary schools, taking care of the features of the different specialization branches
- build an environment for trial studies around a "living-lab" framework, where basic elements are combined to develop prototypes, which span from simple process control systems together with their adjustment parameters, up to more complex systems like robots and organizations management networks.

The multi-disciplinary nature of the whole process is relevant for the objectives pursued by the project, since it encompasses perspectives linked to the "hardware" platform planning, together with a description of the system as a whole, traced at various abstraction levels. To validate what has been first modeled, several simulation phases are carried out, giving space to the following realization, which has to be made available at the final stage in terms of product or service.

To validate the abstraction model introduced in the project, a first example has been drawn using a heating system as an elementary process to which a regulation device should have been adopted. The reason for this choice relies on the easiness with which the system can be modeled and the wide number of applications available in different domains, ranging from home automation, industrial systems, chemical and rural plants, energy savings, and so forth.

3.1 The Driver

In order to make Computational Thinking well understood, and validate the abstraction model drafted above, a tangible and well-known ordinary living experience has been taken as an inspiring background. Since ancient times mankind has always tried to overcome any constraints imposed by the physical properties of being. Any natural phenomena has been analyzed and deepened with the aim of controlling its effects whether to get advantages, when positive, or to reduce its impact, in case of damage.

Hence, Controls and Regulators can be considered as ubiquitous and pervasive mechanisms, used to constraints economic systems behavior, when large-scale systems are considered, but also to control robots, or other smaller devices, when small-scale systems, even though complex, are examined. Controls and Regulators are also immediately perceived within people daily life, several examples do exist like heating and air conditioning systems, showers, and so forth. It is worth not forgetting that man lives because some physical

features in its body, like blood pressure, heart beat rate, are regulated naturally, so that they are constrained within certain thresholds. Even when those features overcome those constraints, artificial devices, controllers or regulators, can be provided to force a correct behavior. An insulin pump device is a typical example where insulin levels corrections are executed to regulate glucose level in blood, when pancreas natural control system is not behaving normally. A picture showing what may happen without these controls is given by the consequences of a possible control failure: when blood glucose level is beneath its lower threshold, a loss of consciousness may occur, in contrast, when the level increases above the higher threshold, damage to liver, eyes, and kidneys may happen.

All the previous comments concern the definition of the possible main questions, related to the impact of a specific control device, to be placed mainly at the higher level (I4) of the layered model drawn in Fig. 1, corresponding to the "Problem-Posing" domain. The following layer (I3) is assigned to the "Problem Solving" domain, where a formal specification, representing the solution model, can grow, . While the third layer (I2) offers a translation of the formal specification into a descriptive programming language, layer I1 hosts an effective implementation on the basis of that language and a suitably chosen hardware platform. Hence, if the question "Why are we developing a control system" concerns the analysis of a specific problem, taking care of

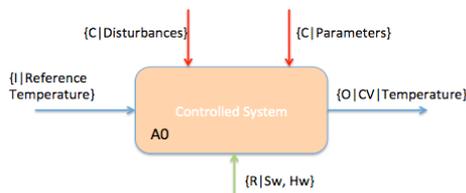


Fig. 2 - IDEF0 block to describe a Controlled System

stakeholders, the impact of caused effects they should sustain, their influence on the domain, their interest to overcome the problem, and so forth, the question "What is a Control System?" should be placed at the second layer, since it concerns the solution to be realized to contrast the problem. Following the inquiry-based approach, it is expected that the next question should be "What is a System?". Hence moving from it, the IDEF0 methodology is often suggested as a means to provide analysts with a suitable way to shape a graphical abstraction of what in nature, but also within organizations, can be perceived as a system, often understood simply through the analysis of its input and outputs. IDEF0 allows an extended representation of what is ordinarily named as a "block schema", having the added value offered by three different sets of inputs, the first labeled as {constraints}, entering the box from above and representing the constraints the process has to respect, the other, labeled {resources}, from below, representing all the means used by the system/process when running, while {inputs} represent the set of information and/or raw material entering the process from the left side of the box. The picture is completed by an arrow coming out from the box, placed at the right side, labeled {outputs}, representing material and information processed and thus produced by the process. Always following an abstraction schema, a real

controlled system can effectively be represented taking advantage of the same graphic representation, focusing on the conceptual model of a particular device, such as that one able to regulate the room temperature, which may be chosen as a first simple example to be dealt with the students.

Future developments will focus on the description of a MapleSim Diagram, choosing appropriate devices. UML diagrams will also be drawn to give an object paradigm perspective. Finally an Arduino platform will be used as a prototyping environment for the physical development.

4. Conclusions

PP&S100 aims at supporting the many innovations put in place by recent regulations adopted by the Ministry and affecting curricula at the upper secondary level of the national education system. One of the main objectives of the project deals with strengthening computer science culture and also enhancing its role as a scientific discipline. To this purpose a model is drawn in this work to shape the rationale of this discipline to be introduced at the secondary school, first cycle, fully adopting the “Inquiry based science education model”, rooted into the constructivist scenario and its specialization, represented also - among others - by the “Computational Thinking” approach. A hierarchical framework is defined using abstraction models shaped as an inquiry-based process and focusing on main questions linked to the problem to be posed and then described according to a “solution” perspective. A four-layer based hierarchy has been considered taking advantage of the project/product life cycle, which is usually founded on a process, which naturally evolves from the highest abstraction level (inception) up to a more practical (implementation and maintenance) phase, involving real world impact in terms of applied actions and sensed physical quantities. To enhance conceptual reading of functionalities provided by each hierarchical level, references to commercial available tools are also taken into account. To make “computational thinking” interpretation easier for large scale education and training systems, some catalyzers have been identified, which can suggest how enhancing and enabling interleaved disciplines domains, through ordinary problems faced by students and teachers in their daily experience.

References

[Palumbo and Zich, 2012] Palumbo C., Zich R. *Matematica ed informatica: costruire le basi di una nuova didattica*, Bricks, Anno 2, n. 4, ISSN 2239-6187, 2012, pp. 10-19.

[Olson and Loucks-Horsley, 2011] Olson S. and Loucks-Horsley S. “Inquiry and the National Science Education Standards: A Guide for Teaching and Learning, Ed., Comm. on the Devel. of an Add to the Nat. Science Education Stds on Scientific Inquiry, NRC, National Academy Press, Washington, D.C, 2011.

[Freire, 1993] Freire. P., *Pedagogy of the Oppressed*. Continuum International Publishing Group, 1970,1993.

[Marlowe and Page, 2005] Marlowe B., Page M., Creating and sustaining the constructivist classroom. 2nd Ed. Thousand Oaks, California: Corwin Press Inc., 2005.

[Zich et al] Zich R., Pardini C., Marchisio M. (2012). Moodle&Maple: una struttura integrata al servizio del Progetto MIUR su Problem Posing and Solving. Atti di MoodleMoot Italia, 2012, pp. 1-10.

[Forehand, 2005] Forehand, M., Bloom's taxonomy: Original and revised. In Ed. M. Orey, Emerging perspectives on learning, teaching, and technology, 2005.

[Brooks and Brooks, 1999] M. G. Brooks and Brooks J. G., "The Constructivist Classroom, The Courage to Be Constructivist", November 1999, Vol. 57, Num. 3.

[Wing, 2008] Wing J. M., Computational thinking and thinking about computing, Phil. Trans., R. Soc., 2008, 366, 3717-3725.

[Peyton-Jones et al, 2011] Simon Peyton-Jones et al, "Computing at School, International comparisons", Microsoft Research UK, 2011

[Livingstone and Hope, 2011] I. Livingstone and A. Hope, "Next Gen. Transforming the UK into the world's leading talent hub for the video games and visual effects industries", A Review by Ian Livingstone and Alex Hope, February, 2011, NESTA.

[Gander, 2013] Gander W., Informatics education: Europe cannot afford to miss the boat, Report of the joint Informatics Europe & ACM Europe Working Group on Informatics Education, April, 2013

[Wikipedia, 28/8/2013] Wikipedia, UML General Information, 12/10/2007, http://en.wikipedia.org/wiki/Unified_Modeling_Language, accessed: 28 August 2013.

[Chonacky and Winch, 2005] Chonacky N., Winch D., 3MS for Instruction: MAPLE, MATHEMATICA, and MATLAB, Computing in Science & Engineering , Vol.: 7, n. 4, July/August, 2005.

[National Instruments, 28/8/2013] National Instruments, LABVIEW, <http://www.ni.com/labview>, accessed: 28 August 2013.

[IDEF0, 28/8/2013] IDEF0, <http://www.idef.com/IDEF0.htm>, accessed: 28 August 2013.

[Unified Modeling Language, 28/8/2013] Unified Modeling Language, <http://www.uml.org/>, accessed: 28 August 2013.

[Visual Paradigm, 28/8/2013] Visual Paradigm, <http://www.visual-paradigm.com/product/vpuml/>, accessed: 28 August 2013.

[Eclipse Project, 28/8/2013] Eclipse Project, <http://www.eclipse.org/>, accessed: 28 August 2013.

[EU, 2004] European Commission-EuropeAid Co-operation Office, Guidelines on Aid Delivery Methods, Volume 1: Project Cycle Management, March 2004, Brussels.

From career education to vocational education and training in digital era: exploiting Communities of Practice.

Maria De Marsico¹, Marco Temperini²

¹Sapienza University of Rome
Dept. Of Computer Science
Via Salaria 113, 00198 - Rome
demarsico@di.uniroma1.it

² Sapienza University of Rome
Dept. of Computer, Control and Management Engineering
Via Ariosto 25, 00185 - Rome
marte@dis.uniroma1.it

Abstract. *The aim of this contribution is to sketch the possible role of the Communities of Practice (CoPs) in projecting in the digital dimension both the activities aimed at orienting towards a future professional life, and those aimed at Vocational Education and Training (VET). While the former are especially targeted at students just exiting from high school, the latter rather span an overall lifetime, including life-long learning, continuous in-service training and professional outplacement. Though the stakeholders participating the two processes might be different, the core underlying processes have much in common.*

Keywords: career education, vocational education and training, community of practice.

1. Introduction

The present economic and financial crisis that Europe and the whole world are experiencing has further underlined the need to devise new ways to fight unemployment and under-employment. In general, it is more and more necessary to avoid the waste of human skills and energy which very often characterizes the present labor market. This complex phenomenon must be approached from two opposite but somehow complementary sides. From the first one, youngsters approaching either university or their first employment should be guided to become aware of their full potential and real skills and abilities, and to choose the most appropriate career for their talent (career education). From the other side, Vocational Education and Training (VET) should efficiently support both continuous refresh of professional skills, and

possible quick and satisfactory (for both the employer and the employee) outplacement. Traditional training courses, either organized in-house or by third-party, often reveal unsuccessful in effectively satisfying the anticipated needs. Most of all, a significant part of protocols, rules, strategies which are integral part of everyday tasks remains often hidden or badly underlined, so that the "official" training often fails in providing a ready-to-exploit expertise. To overcome this problem, the spread of digital resources may inspire new models and strategies. In particular, the flexible model of the Community of Practice (CoP) can be easily adapted to support both career education and VET. Lave and Wenger [Lave and Wenger, 1991] define a CoP as a group of people "who share an interest, a craft, and/or a profession". The members share a common interest in a particular domain, and exchange practical experiences to increase knowledge and skills related to their field. In this way the members learn from each other, and also the group develops new potentialities and dynamics. In fact, a CoP can be realized in everyday life through direct contacts in a physical place, but also, thanks to the enabling technologies, at a distance through online discussion boards and newsgroups. Either a traditional forum may be configured as a basic example of CoP, or dedicated frameworks may be set up.

This paper aims to sketch a solution to make choices and provide services that promote entry into the labor market, acting as a first step in the design of a more comprehensive system that can trigger a permanent cycle of training and outplacement to allow all citizens to keep their skills up to date and adapt to rapidly changing work environments. The consequences of such a virtuous cycle in the long-term concern not only recent graduates looking for first occupation, but also people wishing to reposition themselves in the labor world and, ultimately, the companies that could encounter candidates more motivated and aware of the requirements for a successful working relationship.

2. CoPs history at a glance

The concept of Community of Practice was already around in social and educational sciences (e.g., see [Bourdieu, 1977] and [Vygotsky, 1978]), however its success and spreading of concrete implementations were spurred by the work of Lave and Wenger ([Lave and Wenger, 1991] [Lave, 1991]).

The core idea which aimed at changing the way people work and learn was just the discarding of the separation between work practice, mostly conservative and stable in time, and learning, which was in this way necessarily focused on the most static aspects. A large part of experiential knowledge, the so-called implicit knowledge gathered within an organization and while performing concrete tasks, was left out from training due to such separation. However, people actually work according to patterns and rules which are often different from manuals descriptions and from the "theoretical" descriptions of job tasks. The actual practice relies on often complementary as well as completely alternative strategies and rules. Nonetheless, organizations tend(ed) to rely on the "static" form of training to transmit work practice. The roots for this kind of behavior are mainly philosophical. A special value is assigned to "abstract

knowledge", while the concrete practice details are considered as contingent, easily derived after the relevant abstractions, and therefore less essential.

It is interesting to notice that, actually, in the past of most cultures, learning and training were firmly based on the role of the artisan workshops, that were the privileged places for transmitting and preserving arts and crafts. Common practice and storytelling were the main instruments for the knowledge/skill transmission. Strangely enough, it was the industrial age that stated the prevalence of abstract knowledge over actual practice and established a separation between learning and working, as well as between *learners* and *workers*. The new theories claimed a return to the "ancient wisdom": appropriate practice is central to understanding roles and concrete practices related to a specific job, while abstractions detached from practice might obscure less trivial details of practice. Further, they assigned a central role to social exchange of experience and information. It is interesting to notice that such topic formerly focused on organizational learning [Brown and Duguid, 1991], at present deserves an increasingly significant role in learning in general. Theories by Vygotsky [Vygotsky, 1978] and by other researchers paved the way towards projecting the same ideas on education and learning, in the line of socio-constructivism. Learning is a social process happening within the zone of possible achievements of the student [De Marsico, Sterbini, and Temperini, 2013]. The core concept in the practice-based theory of learning devised by Lave and Wenger's [Lave and Wenger, 1991] is the "legitimate peripheral participation" (LPP) in "communities-of-practice." (CoP). In a CoP external observers may be allowed to watch without actively participating. The novices firstly access the community from the periphery; as they get experienced and gain reputation, they move towards full participation and membership. Therefore we can usually distinguish 1) a group of core experts, who achieved high reputation and reliability due to their experience and according to the assessment of their peers, 2) the major group of active participants, who participate to the exchange of information and experiences, and 3) the peripheral participants, who start as observes of the dynamics and contents of the CoP, and eventually and possibly level-up. Personal growth is also reflected by community growth: the community itself develops through different levels of interaction among the members, as well as different kinds of activities. Self-development originates by active participation to the community. As a matter of fact, the two central elements of the CoP approach are *situated learning* [Hummel, 1993] and *community reflection on practice*. Knowledge is acquired from and applied to everyday settings, and is discussed with peers and experts; this makes up a true social system [Wenger, 1998]. These concepts focus on learning, rather than teaching [Lave and Wenger, 1991], though not building an educational theory. Internet, with early forms of forums, and the modern Social Networks (SN), built upon Web2.0 dynamics, has technologically supported the exchange of not only personal data but also of working experience and learning difficulties on virtual communities. Dedicated software frameworks (e.g., ELGG - <http://elgg.org>) allow to set up CoPs, which have become one of the most appreciated forms of social learning [Blackmore, 2010]. Recent documents from projects in New Zealand also report the use of CoPs to implement career

education networks. Their aim is to support an aware transition from school to work and/or further education [Vaughan and O'Neil, 2010].

We can say that learning is both the main activity and the core topic of the CoP strategies. Buysse, Sparkman and Wesley [Buysse, Sparkman, & Wesley, 2003] underline "the growing need to integrate educational research and practice" to connect what we know with what we do. However, shared lists of recommended practices often fail to promote the personal responsibility and exploration ability, pertaining to both educational researchers and teachers, and to parents and students. CoP practices should encourage and motivate every member of the educational community to analyze, constructively criticize, and effectively complement each other's experiences. Teachers should be encouraged in taking active part in research activities. To these aims, the authors consider CoPs a very promising tool, even compared with other proposed approaches to join research and practice, e.g., action research [Calhoun, 1994], or professional development schools. The core idea is breaking "the linear relationship through which information is handed down from those who discover professional knowledge to those who provide and receive educational services" [Buysse, Sparkman, & Wesley, 2003].

3. CoPs for career education

Currently the orientation towards a new job starts, at best, at the time of entering into the new organization ("staff orientation"). The orientation, which should begin in any case before the actual activity of the newly-hired, should provide the new employee with basic information about the organization, the position occupied and expectations linked to it by the company, and the overall community in which operations do occur. However, a significant part of such information and related interactions should be anticipated at earlier phases prior to joining the labor world. In collaboration with universities and training centers, a first level of information should be provided before of the choice of university course or of applying for a future work, so to minimize the phenomenon of early school leavers, that often depends on incorrect assessment of commitment and perspectives. Moreover, additional information services should be provided in a later stage, before of the conclusion of the previous studies, so to support a more informed and convinced choice of both the completion of the educational/training path, and of the new career path.

Our proposal relies on a CoP-oriented approach to career education, which should include the development of an integrated IT platform and innovative services, able to support students during their academic career, from the decision about the most suitable course to enroll up, to the first contact with the labour world after graduation. In particular, it is important to model the processes that take the students from the moment they come out of high school and must choose the path best suited to their training and personal inclinations, and to follow them in their tracks involving either work or university or both. In fact, many universities are somehow equipped to respond to these needs, especially in promoting periodic meetings (possibly supported by technology) between students and companies and in presenting courses and routes.

These experiences provide obvious benefits, and also limitations. The meetings with the companies are particularly effective in connecting aspiring trainees with industrial stakeholders mainly interested in a temporary collaboration (the *stage*), which is part of the university curriculum, with a goal generally limited to short-term needs from both parties. In this perspective such meetings are mainly addressed to students at the end of their degree course. In addition, the cadence is mostly biannual, leaving many students run out of information when it would be needed. All this means that the target is mainly centered on the needs of the companies. On the other hand, existing career portals supports business contacts and the matching of supply and demand in the longer term, but the student is still a passive subject, and above all has very little information about the company's activity as a whole, which is often crucial to a satisfactory choice for both stakeholders.

In both cases, therefore, the companies' attention is particularly directed to illustrate the objectives of short or medium term. While this may be relatively satisfactory for a student looking for an internship, it shows serious limitations when the same student is looking for tools to get a better view on the development prospects of the companies (s)he made contact with, and on the fields in which the stage activity can possibly convey his/her professionalism, as well as on the possibility of conversion of skills acquired in different fields. Especially the current organization of the information is essentially static, and does not provide any form of customization. Moreover, as already pointed out above, students are not provided with any really interactive and informative communication tool with the labor world, nor with a support tool in the selection of training courses. Going even backwards, there is little or no concrete systematic information about careers and perspectives related to a given university course, so that the choice is often driven by friends, relatives, or even by vague as well as unfounded expectations. The result is a huge premature university death rate, with the related waste of resources and energies from both the student and the surrounding subjects.

This situation leads to one of the major ills recognized in the Italian education system: the distance between the world of production and the world of education. However, the current socio-economic situation highlights the need, even more than in the past, to support more effectively the match between supply and demand in the labor market. The solutions proposed to overcome the mentioned limitations include, for example, the provision of examples of company's work processes, so to allow application and testing of capabilities and inclinations of job applicants. For graduates, additional information and examples of the activities may be of help to focus talents and abilities, so to access career paths best suited to their preparation and most satisfactory against their expectations. Particularly valuable are the testimony of company employees who share their experiences, difficulties and satisfactions obtained in order to allow a comparison of experiences and a more conscious approach to the company. Question / answer forums and other means of communication between graduate and company should contribute to an open mechanism for the exchange of information and experiences, which can bring both the company and the graduate to take informed and effective decisions. In

practice, we propose a first step in transforming the career education in a transversal tool, characterized by an ever-changing series of personalized services to facilitate the achievement of the objectives of education, training, empowerment, social inclusion and well-being of citizens.

To achieve the above objectives, guidance services should be seamlessly accessible, starting at as early a stage as possible in the university study career. This could avoid many failure-stories due to lack of awareness of the requirements and objectives of a study/training program. In this broader perspective, it is important also to provide a correct and clear information on the part of educational organizations, in synergy and collaboration with those alleged to counterparty possible future professional relationships.

In practice, we propose to implement a CoP, extended to both academy and companies, where the students enter in a "peripheral" observing role, but proceed along the path towards active participation, when they will be able in turn to give suggestions to younger newcomers.

This scenario requires a significant amount of energy and resources, and an effort to redesign at least part of the information architecture linked to the career education. On the other hand, the necessary actions might be framed in the context of the EU Horizon 2020 Strategy, with particular reference to the most recent document (June 2013) of the European Community that calls for action to solve the problems of youth unemployment (http://ec.europa.eu/europe2020/pdf/youth_en.pdf). These actions would also help to strengthen the links within the so-called "knowledge triangle" of research, education and innovation, which are the three main hubs around which the modern knowledge society should develop. This can only bring benefits to the same companies involved in terms of improved productivity and competitiveness. Good examples of the proposed approach come from abroad, e.g., the experience portal IGGY "Connecting and challenging the world's brightest young minds" developed by the University of Warwick under the auspices of the World Council for Gifted & Talented Children. Through the portal, the members access to a range of services including participation in discussions and debates on topics of interest with experts from worldwide, interaction with senior students from different universities, input from experts and professors, participation in projects and competitions, access to contents.

4. CoPs for VET: a concrete example

Digital support for VET is fundamental for workers, but it is of paramount importance for trainers. A main motivation for the European project Understand IT (<http://aitel.hist.no/understandit/>) was the documented need for supporting VET teachers, trainers and tutors in including ICT in their teaching activities, and in training their trainees in using similar tools for life-long learning. Despite the more and more advanced learning technologies, many teachers, most of all elder ones, are frightened or even disturbed by technology. The preceding project Vitae, whose results were exploited and refined in Understand IT, developed a set of net-based collaborative learning activities, using modern Web 2.0 tools. The resulting strategy aims at helping VET teachers to discover

how to use these tools in their own practice. Communities of practice were encouraged in Understand IT, as a crucial mean to peer-assisted learning, together with mentoring, coaching and peer-to-peer activities.

The first concrete step of the project was the review of the actual situation of the national educational systems, national educational standards and curricula of participating countries (Denmark, Germany, Lithuania, Norway, Italy and Portugal). In practice, all countries are using virtual classes and learning platforms, but they still lack a systematic integration of ICT into the pedagogical activities, especially as for the use of Web 2.0 tools. There is a growing awareness of the crucial role of using Web in learning activities, but who, or what institutions in each country, should care about the training of involved teachers and trainers? In many settings resistance to change must be addressed, as the main challenge for educators is to give up control, and award learners new degrees of trust, exploiting peer-to-peer activities.

The results of the project were tried out by running courses in 4 languages in the participating countries. To this aim, it was not possible to overlook national situations and educational settings before proceeding to a joint planning towards the introduction of ICT in VET. In particular, before trying the introduction of a trans-national CoP. The project results especially emphasize the need for more independent learning, global and intellectual awareness, and proof of progress through the compilation of appropriate portfolios. In particular, it is often the case that learners are required to be able to use communication forums and assess the trustworthiness of online information in a critical way. On the other end, educators are more and more often required to train their students to use e-portfolios: these support evidence both of the learning advancement and of the teaching practice.

The powerful message sent by the learner centric model [Quintana, Carra Krajcik, and Elliot, 2001] can be enhanced by enriching the role of the teacher to include greater individual support and guidance for the learners. In addition, peer-mentoring and peer-coaching seem to be winning strategies, most of all when learners are teachers in their everyday life, as it was the case for the project experimental activity. However, coaching is an art, that implies carefully tuned presence/absence of the coach, to spur coachee's abilities regarding critical planning, and accurate evaluation of obstacles and of related overcoming activities. On the other hand, this model is somehow implicit in how CoP work [Wenger, 1998]. This was the main motivation spurring the establishing of a Learning Community of Practice (LCoP), to support participants' exchanges about their collaborative activities.

In addition to considerations regarding the search for effective *training-the-trainers* strategies, we also took into account the need for projecting national VET programs onto an European/international perspective. For the project this meant to cope with the language issues related to both the management and the interaction of several "local learning communities", based on the different languages of involved countries. Solutions were devised to allow interface localization, support to content translation and tagging (actually, these aspects are seldom addressed in existing CoPs), and channeling of the

communication among local learning communities. Local peers could so have the possibility to add an international flavor to their achievements

We first run a pilot version of the VITAE course, and then experimented it with "real" teachers. The pilot version of the course was attended by members of the project (with other members running as observers). The course comprised several tasks, ranging from practice with Web 2.0 tools for doing and analyzing one's own teaching, to analysis of existing CoPs and social networks, to coaching and mutual coaching sessions with the course coordinator and with peers, to the final composition of a plan to experiment one's learning outcomes in her present teaching activity. Main outcomes of such trial were basically in the deeper appreciation of the amount of work the students are asked to perform, in order to timely deal with the assigned tasks, and in the knowledge about what parts of the course need localization and/or translation (and what of such parts are hardly tractable in such a sense). So we resolved to shorten the course (time is a premium for our prospective - working - students) and planned how to deal with localization/translation.

5. A possible implementation

Service Oriented Architecture (SOA) defines a logical model with distributed software modules offering services in a standardized manner. The application of this paradigm brings a series of advantages. Software as a Service allows the encapsulation of software components, isolating them so that only the layer relative to the actual service is exposed to the outside. This produces independence from the implementation and security of the internal system. The encapsulated application logic is completely decentralized and accessible over the Internet from different platforms, devices and programming languages, for full interoperability. Developing a set of components around a layer of existing software is simple and should not require changes in the original application code. The adoption of SOA paradigm allows for a gradual implementation, through the use of existing infrastructure and the simultaneous application of the latest technologies. Moreover, it allows to reduce the costs of training because it is based on a common set of shared services and activities.

The conceptual SOA framework is generally characterized by three specialized layers: a front-end layer (interface), a back-end layer (services logic), and a data persistence layer. Platform services would include: self-assessment of skills, knowledge and flaws; an information system available to companies to expose, in addition to research positions, the main strategic lines of development; a layered CoP where students and employees are able to obtain/offer training and guidance pathway identified. The different levels of information and interactive services require specialized modules, whose composition and structure, offered by a SOA framework, will be completely transparent to end users. First of all it will be necessary to define the different roles involved (student, educator, trainer, recruiter, etc.) as well as related services and levels of information relating to each. In addition to these roles, somehow "traditional", we propose to promote a concrete "coaching" activity by former students, eventually coming from the same educational setting, already

currently employed by the companies involved, which can establish an interaction geared more to a peer-to-peer relationship, and then predictably more frank and open, with new aspirants. In addition, senior experts or teachers can also be consulted at all stages of the student's participation in the community. The set of planned activities takes the form of a CoP organized by areas, for roles and levels of detail, with extreme flexibility in switching between different informative areas. In addition, both the information and the actions of the parties may be subject to a process of "reputation gain" supported by a mechanism of peer-evaluation available to the community, so the information deemed most useful will be classified and stored for future research in a knowledge base. The latter is expected to grow in size over time, both in terms of depth (level of detail) and amplitude (variety of areas and topics) and representativeness. This may be achieved through appropriate participatory tools for tagging and annotation, so that the users themselves can actively contribute to information organization and to improve the community protocols. In addition, planning to exchange expertise and information within European Community, the platform might be opened to support automated or peer-mediated translation of contributions considered interesting. The overall architecture is sketched in Fig.1. For instance, company services may include publishing of various contents, e.g., job offers, as well as mentoring tools, e.g. editors to create guided tours of enterprise activities. Student services may include profile publishing, peer-coaching, and communication and cooperative working facilities. Expert services may essentially include facilities to mediate among stakeholders, or provide expert advice about specific questions.

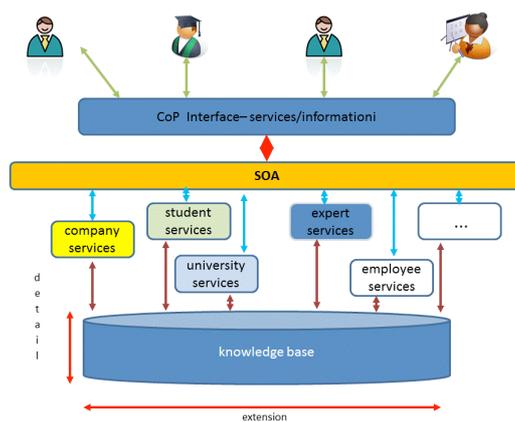


Fig. 1 – Proposed architecture

6. Conclusions

We sketched a possible architecture including the use of layered CoPs to handle both student orienting (career education) and VET for long-life learning

and outplacement. The approach involves the active and stimulating participation of all stakeholders involved in the transition from school to labor world, and exploits the new possibilities offered by CoPs in the digital world.

References

Blackmore, C.(2010). *Social Learning Systems and Communities of Practice*. The Open University, 2010

Bourdieu, P. (1977) *Outline of a Theory of Practice*. Cambridge: Cambridge University Press

Brown, J.S., Duguid, P. (1991). Organizational Learning and Communities-of-Practice: toward a unified view of working, learning and innovating. *Organization Science*, 2, 1, 40-57.

Buysse, V., Sparkman, K. L., & Wesley, P. W. (2003). Communities of practice: Connecting what we know with what we do. *Exceptional Children*, 69, 3, 263-278

Calhoun, E.F. (1994). *How to use action research in the self-renewing school*. Alexandria, VA: Association for Supervision and Curriculum Development.

De Marsico, M., Sterbini, A., Temperini, M. A Framework to Support Social-Collaborative Personalized e-Learning. M. Kurosu (Ed.): *Human-Computer Interaction, Part II, HCII 2013, LNCS 8005*, pp. 351–360.

Hummel, H.G.K. (1993). Distance education and Situated Learning: Paradox or Partnership? *Educational Technology*, 11-22.

Lave, J., & Wenger, E. (1991). *Situated learning. Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.

Lave, J. (1991). Situating learning in Communities of Practice. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (Eds.) *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, DC.

Quintana C, Carra A, Krajcik J, Elliot S. Learner-centred design: reflections and new directions. In: Carrol (ed.) *Human-computer interaction in the new millennium*. ACM Press, 2001, 605-626

Vaughan, K., O'Neil, P. (2010). *Career Education Networks and Communities of Practice - A report from the School-Communities strand of the Education Employment Linkages project*. EEL Research Report N. 6. Available online at http://dspace.lincoln.ac.nz/dspace/bitstream/10182/4370/1/eel_rr_06.pdf

Vygotsky, L. (1978) *Mind in Society: Development of Higher Psychological Processes*. Cambridge MA: Harvard University Press.

Wenger, E. (1998). *Communities of Practice: Communities of practice: Learning, meaning, and identity*. Cambridge Un. Press

Wenger, E. (2010). Communities of Practice and Social Learning Systems: the career of a concept. In: C. Blackmore (ed.) *Social Learning Systems and Communities of Practice*. The Open University, 2010, 179-198

A Social Game-Based approach supporting Environmental Learning

Stefania Cuccurullo, Rita Francese, Ignazio Passero, Genoveffa Tortora

Dipartimento di Studi e Ricerche Aziendali

Via Giovanni Paolo II, 132

84084 Fisciano (SA) - ITALY

{scuccurullo, francese, ipassero, tortora}@unisa.it

Abstract. *Environmental education is a priority at worldwide level. Thus, it is indispensable to change people ecological, economic and social perceptions by integrating the concept of sustainable development in their learning processes. In this paper we present and evaluate a game-based learning approach based on reputation, collaboration and competition aiming at promoting the waste disposal practice among the primary school children. The proposed game, named Pappi World, is a 3D Virtual World designed according to pedagogical theories concerning behavior and attitude changes.*

Keywords: game-based learning, social learning, virtual world.

1. Introduzione

Global environmental problems become increasingly pressing. To achieve the sustainable development it is necessary to change the ecological, economic and social attitudes and perceptions of people by integrating the concept of sustainable development in their educational process. Environmental education is linked to the acquisition of a responsible environmental behavior that goes beyond the mere knowledge. In fact, in addition to knowledge people has to develop attitudes and motivations towards correct behaviors. These activities are successful if they involve citizens from their early childhood.

In this paper we discuss the social aspects of a 3D game to support the learning of a correct behavior for the practice of recycling waste. The proposed game, named Pappi World, aims at supporting elementary school children in the learning of how to recycle waste. It proposes different learning approaches, such as: individual learning, in which students learn how to properly differentiate waste; social learning, based on the reputation of the participants. The game also supports collaborative and competitive learning mode. A preliminary evaluation of the students attitude and of their social interaction has also been performed.

This paper is organized as follows: Section 2 discusses related work, while Section 3 briefly introduces the game features, while Section 4 details the proposed learning approach. Section 5 describes a preliminary evaluation of a learning experience using Pappi World. Section 6 finally concludes by outlining the lines of future research.

2.Related Work

VR and 3D graphics have been largely adopted for e-learning (Monahan et al., 2008; De Lucia et al., 2009b). Barab et al. (2005) describe a learning and teaching project adopting a multiuser, virtual environment to propose to 9-12 children educational tasks. It combines strategies used in commercial gaming environments with lessons from educational research on learning and motivation.

An example of simulation-based learning game is Math-City, aiming at supporting the students in learning mathematical concepts (Polycarpou, Krausea, Rader, Kembel, Poupore & Chiu, 2010). The progress in mathematics is adopted by the students to grow their city. No social aspect is considered.

Wrzesien & Raya (2010) adopt a serious Virtual World for teaching children natural science and ecology. They follow both a collaborative and a competitive approaches, enabling the students to discover their potential and skills during the game and compare them with other players while competing, it also allows them to share and learn during the collaborative play.

3.The game

The game has been designed according to the guidelines specified by the Italian Environmental Education and Instruction Ministries, which include the understanding the concept of waste as a resource and the value of the differentiated collection, according to the practice of the three R's (Reuse, Recovery and Recycling). In addition, the educative process should include laboratory activities, also in a edutainment modality. The laboratory, through experimentation, play and research activities provides new knowledge, creating the basis for reflection and exteriorization.

The aim of the game, funded by Campania Region (Italy), is to contribute to the sustainable development of the ideal sustainable world of the alien Pappi. This world is structured as a set of archipelagos, each one representing a school. Each island of the archipelago, representing a class, can be seen in detail. On an island there are several cities, representing the students' land. Finally, a city is an ad-hoc developed 3D world, shown in Fig. 1, which the students explore and where they learn the waste disposal practice. The city also offers a game room where the students can play individual game oriented toward the waste disposal learning.

4.The learning approach

According to the Tbilisi Declaration of 1977, environmental education was defined as formal and informal education, aiming at raising the awareness of people for ecological dynamics, their natural surroundings, environmental problems, and interdependencies with economic, social and cultural development. As a consequence, it is not enough to interact with learning contents to change the people behavior. People will behave correctly when they will change their attitudes, motivation, and values. (Smith & Ragan, 1999; Miller, 2005).



Fig.1 - Waste Collection

In the past, government organizations insisted on the knowledge of the problem (Dietz & Paul C. Stern, 2002), distributing informational materials. Concerning a correct practice of recycling, the subject of this article, they provided information on where, how and when to recycle (*Procedural knowledge*). Another type of knowledge is the knowledge of the Impact (*Impact Knowledge*), i.e. on the benefits of recycling of waste in terms of pollution reduction and energy conservation. A third type of knowledge is the *Normative Knowledge*. It represents the knowledge about the behavior of others. It has a strong impact on the attitude to properly recycle waste, basing on the awareness that relatives, friends and the whole community is dedicated with interest to this practice, and that each individual is judged by others and appreciated for its correct behavior. As shown by (Dietz & Paul C. Stern, 2002),

this kind of knowledge has a strong impact on the changes in behavior and attitudes..

The game offers different learning approach, as described in the following.

3.1 Individual Learning

The individual learning activities proposed by Pappi World aim at promoting Procedural Knowledge (how to collect and differentiate waste). Each student owns a village, represented by the 3D world shown in Fig.1, and he/she has to collect the garbage, putting it in the appropriate bin. Garbage is randomly generated. The student can collect other waste by playing an individual game available in a game room, useful to collect extra waste recycling scores. These games are simply to use. The player can play with them until the needed quantity of material has been collected. They cannot freely be accessed, to avoid to distract the user from the game purposes.

3.2 Collaborative Learning

The students learn the relevance of a recycle policy by exploiting the collected garbage as a mechanism for growing their village. Indeed, they can convert the collected material in new houses and furniture for progressing their village. If they do not own the right quantity of a given material (plastic, paper, etc.), they can exchange it with other players. The material exchange interface is shown in Fig. 2. This practice enforces the idea that differentiating garbage is a useful practice for improving their village. Pappi World also add the "waste miles" concept, derived from "food miles" . Food miles is a term which refers to the distance food is transported from the time of its production until it reaches the consumer. Food miles are one factor used when assessing the environmental impact of food, including the impact on global warming. We introduce this concept also in the waste case, given a high value to exchanges of material between students that are closer.

3.3 Competitive Learning

A way to collect more garbage can be to go in the village of other students and take theirs. Competitive learning exists when one student achieves his goal and all other students fail to reach that goal (Johnson & Johnson,1991). Competitive learning have many criticisms, but if it is well designed can be very stimulating. In this game it is possible to take material from students of other classes, stimulating *intergroup competition* that can be seen as an appropriate competitive strategy since it maximizes the number of winners.

3.4 Reputation

The concept of reputation has a great relevance in the social dimension. In general, a reputation system makes evident the contribution of each member of a community, such as a class or a school. It is a mean for motivating

A Social Game-Based approach supporting Environmental Learning participants, influencing their attitude and beliefs. To get the best results it is important that an individual can directly compare his or her behavior with the provided information (Dietz & Paul C. Stern, 2002). It is also worth making comparisons considering groups near to that person, such as the neighbors or the classmates, instead of the country or the city.

Pappi World adopts the following reputation mechanisms:

- *promoting class/school leader*, i.e., a student particularly motivated, that recycles diligently and encourages his/her classmate to recycle;
- disseminating data on community (class/school) recycling rates, such as the percentage of people that regularly recycle for each community;
- *city growing*, the representation of the village in the school island visually show the village level. In this way a visual feedback on the recycle level of each student of a class is provided, highlighting the most virtuous member of the class community.



Fig.2 - Material Exchange

3.5 The teacher role

The teacher acts as a facilitator of the game. He supervises the actions of each student and controls the class progresses. He reinforces the role of the

class leader and stimulates the participation by public commenting on the game web site.

5. Preliminary evaluation

The participants in this study were 24 children between 9 to 10 years randomly selected from the primary school of the Town of Baronissi (Italy). They all attended the 5th grade and had the same learning objectives related to natural science and ecology.

Method

The aim of the evaluation was to assess the learner's attitude toward experiential learning with the Pappi World . In the experiment we considered the factors which promote the use of the system

Evaluation Organization

The study has been performed in one laboratory session (i.e., three supervisors for all the 24 participants). In the first step, a presentation of 10 minutes introduced to all the subjects the main features of Pappi World. The students played Pappi World for an hour. After playing, they answered the Student Post-Game Perception Questionnaire, reported in Table 1. A 7-points Likert scale anchored at 1 by Strongly Agree and at 7 by Strongly Disagree has been adopted.

Table 1. The Student Post-Game Perception Questionnaire.

ID	Question
Q1	The game environment was stimulating. (Environment)
Q2	When I made a mistake while playing, I was able to solve the problem easy and quickly. (Behavioral Control)
Q3	I quickly learnt to play the game (Learnability)
Q4	I'm satisfied with this game. (Satisfaction)
Q5	I will recommend the game to a friend.
Q6	I enjoyed the game. (Perceived Enjoyment)
Q7	I would like to use the game in the future. (Intention to use)
Q8	I was able to control the avatar and the events. (Control)
Q9	The avatar control distracted me from what I had to do. (Distraction)
Q10	It was easy to collect garbage. (Easy to use Perception)
Q11	I played without effort. (Mental effort)
Q12	I think I have learnt to make the waste separation better. (Perceived usefulness)

Q13	Material exchange was stimulating. (Collaboration,)
Q14	It was stimulating to collect the garbage in the city of the others. (Competition)
Q15	It is important for me to appear in the list of the most virtuous students that are an example of correct behavior. (Reputation)

The students involved were accustomed to answer questionnaires. Considering their young age open questions were avoided and the questionnaires have been designed to keep focused and straight forward, trying to avoid any confusion. In particular, we took into account the following factors derived from the TAM model (Davis, 1989; Davis, Bagozzi, & Warshaw, 1989):

Perceived usefulness (question Q12), defined as “the degree to which a person believes that using a particular system would enhance his or her job performance”. In the learning context, the user believes that a system would yield positive benefits for learning.

Perceived ease of use (question Q10), “the degree to which a person believes that using a particular system would be free of effort”

Perceived enjoyment (question Q6), “the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated” (Davis, Bagozzi, & Warshaw, 1992).

We also included question Q1, which may have a significant influence on the determinants of attitude toward using the game.

When the student easily interacts with the environment and is able to manage problems without the need of an external help his avatar behaves as expected and moves in the virtual world according to his commands (question Q2). Learnability is an important quality of a game since low learnability gets the student to abandon the game (Q3). Question Q4 measures the overall satisfaction, while question Q5 considers if the user is satisfied with the game so as to recommend it to friends. Also Intention to use is a relevant aspect (Q7).

Since the game has been developed using a 3D Virtual World environment, aspects concerning the presence perception should be verified (Witmer and Singer, 1998; De Lucia et al., 2009b).

A high control on the environment and on the interaction contributes to get a high experience of presence (Q8). Distraction (Q9) is another factor that contributes to increase immersion and participation when is reduced. If the user does not control his avatar in a natural way the presence perception is reduced. Question Q10 is specific to assess the ease of the garbage collection feature of the game. It could happen that the game is easy to use in general, but the user could have some problems with its most relevant feature. Question Q11 collects the perception related to the mental effort of the user while playing. An excessive effort could discourage the use of the game.

Social aspects are considered in the following three questions: Question Q13 is related to the satisfaction concerning the material exchange feature (collaboration), while Question Q14 collects the perception related to the

possibility of growing his own city by collecting the garbage of the other participants, competing with them to take their garbage. Finally, question Q15 collects the perception of the user of being appreciated by the community for his success.

Results

Fig.1 shows the boxplots representing the results of the user perception questionnaire in Table 1. All the participants considered PW stimulating (Q1). The games was perceived easy enough to be controlled (Q2): the positive scores are more than 70%. It was also perceived to be easy to learn (Q3), with 100% of positive answers. Participants were very satisfied (Q4), with 100% of positive answers. 100% of them would recommend the game to a friend (Q5); PW is considered very enjoyable (Q6) and 88% of participants would like to use it in the future (Q7). The avatar control was not considered a distracting factor (Q9) and the garbage collection was perceived easy to perform (Q10). No particular mental effort was required (Q11), which were perceived instructive (Q12). The student perceptions related to the collaborative/competitive activities proposed by Pappi World are reported in Table 3. The judgments expressed were very positive.

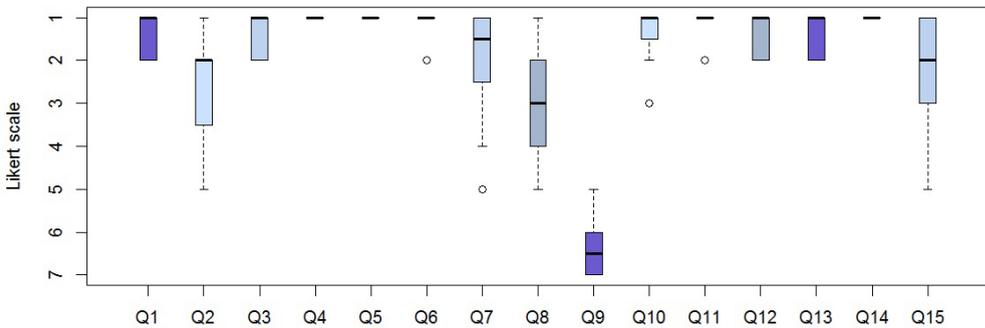


Fig.1 - Perception Results

Table 2. Social activity related results.

Question	% positive answer	Median
Q13	100	1
Q14	100	1
Q15	92	2

We also collected the data reported in Table 3 concerning the city level reached by the students.

Table 3. City evolution during the game..

μ	σ	\bar{x}	min	max
2,37	0,76	2	1	4

This data confirm that the game is appropriate for three laboratory sessions of one hour. Indeed, after one hour, half of the students reached the second level of the game. Only one remained at level 1, while two of them reached level 4. Thus, one hour session is enough for exploiting the reputation approach, being the differences among the students visible. We also evaluate if the students really collected the garbage of the others. Indeed, they subjectively appreciate this aspect, but only if this had de facto happened during the game, it can be considered as an important mechanism. Table 4 reports the results concerning the competitive approach (take waste from the others) and the collaborative one (exchange of material). A greater number of visits has been performed w.r.t. exchanges. Thus, the first approach seems to be preferred.

Table 4. The collaborative approach data.

Collaborative approach	μ	σ	\hat{x}	min	max
Neighbor visit	6	1,83	6	3	8
Waste exchange	3,70	1,4	4	2	6

6. Conclusion

This paper describes the social approach proposed by the game Pappi World and discusses a preliminary evaluation. The game follows pedagogical guidelines concerning the use of social features such as collaboration, competition, public scoring and reputation to induce the belief that waste can be a relevant resource. Each student becomes a member of a community where the recycling practice is diligently followed. The teacher has the role of an active mediator. We also collected the written comments of three teachers participating to the experience that expressed positive judgment towards the adoption of this game in a laboratory activity. In particular, they suggested to investigate how to present the game also to special needs children.

We plan to perform a deeper evaluation adopting the game in a longer education experience.

Acknowledgements

This research has been funded by Campania Region. The authors of this paper would like to thank all the participants to the evaluation of the game.

References

- Barab, Sasha, et al. (2005) Making learning fun: Quest Atlantis, a game without guns." *Educational Technology Research and Development* 53.1: 86-107.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology*, 22(14), 1111–1132.

De Lucia, A., Francese, R., Passero, I., & Tortora, G. (2008a). Supporting jigsaw-based collaborative learning in second life. *Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies, ICALT 2008* (pp. 806–808).

De Lucia, A., Francese, R., Passero, I., & Tortora, G. (2009b). Development and evaluation of a virtual campus on second life: The case of SecondDMI. *Computers & Education*, 52, 220–233.

De Lucia, A., Francese, R., Passero, I., & Tortora, G. (2009c). Development and evaluation of a system enhancing second life to support synchronous role-based collaborative learning. *Softw., Pract. Exper.*, 39, 1025–1054.

Dietz, T., & Paul C. Stern, C., Editors (2002). *New Tools for Environmental Protection: Education, Information, and Voluntary Measures*. The National Academies Press.

Johnson, DW & Johnson, RT (1994). *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning*, 4th ed, Allyn and Bacon, Boston. UNSW library

Kebritchi, M., & Hirumi, A. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers and Education*, 51, 1729 – 1743.

Law, E., KickmeierRust, M., Albert, D., & Holzinger, A. (2008). Challenges in the development and evaluation of immersive digital educational games. In A. Holzinger (Ed.), *HCI and Usability for Education and Work* (pp. 19–30). Springer Berlin / Heidelberg volume 5298 of *Lecture Notes in Computer Science*.

Miller, M. (2005). Teaching and Learning in Affective Domain. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Retrieved 08/12/2013, from <http://projects.coe.uga.edu/epltt/>

Ministero dell'Ambiente e della tutela del territorio e del mare, *Linee guida per l'educazione ambientale e allo sviluppo sostenibile* http://www.minambiente.it/export/sites/default/archivio/notizie/Linee_guida_ScuolaxAmbiente_e_Legalitx_aggiornato.pdf

Monahan, T., McArdle, G., Bertolotto, M. (2008) Virtual reality for collaborative e-learning, *Computers & Education*, Volume 50, Issue 4, May 2008, Pages 1339-1353.

Smith, P. & Ragan, T.J. (1999). *Instructional design*. New York: John Wiley & Sons.

Bob G. Witmer, Michael J. Singer: Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence* 7(3): 225-240 (1998).

Wrzesien, M., & Raya, M. A. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the ejunior project. *Computers & Education*, 55, 178 – 187.

Recommending Touristic Paths using context information

Flora Amanto, Francesco Gargiulo, Vincenzo Moscato,
Antonio Picariello, Carlo Sansone
Università degli Studi di Napoli "Federico II"
Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione
Via Claudio 21, 80125, Napoli
{flora.amato, francesco.grg, vmoscato, picus, carlosan}@unina.it

Abstract. *In the developing of applications for touristic paths planning, context-aware recommendation services can profitably exploited to suggest automatically useful information for users depending on the context. In fact, Recommenders' technology can accommodate location's dependent information with user's needs in a mobile environment, related to the touristic domain. In this paper, we describe a system providing context-aware recommendation services for the generation of touristic paths driven by the context and based on both the experience of previous users and personal preferences of tourists. The information are gathered by several heterogeneous sources (sensors, web portals, repositories related to touristic events and locations) and are stored and analyzed in a cloud architecture that is particularly suitable to process and manage the huge amount of extracted data.*

Keywords: Recommender Systems, Context-Awareness, Big Data.

1. Introduction

Nowadays, travel and tourism industry is one of the most important and dynamic sectors, especially for B2C e-commerce. The tourism industry is regarded as one of the biggest sectors in the world generating an estimated 11% of the global gross domestic product and employing 200 million people and serving 700 million tourists worldwide, a figure, which is expected to double by the year 2020.

The great advances of the information technologies are continuously changing the way in which touristic services are performed, enabling people to access accurate and "on-line" information – known by now as *big data* - as well as to undertake reservations and plans in real time, thus reducing costs and satisfactions with respect to conventional methods.

Just to make an example, taking a quick look at what is happening in the last few years, people search for information before traveling: usually they make online air-ticket bookings, and hotel and room reservations, decide the kind of restaurants to go and the particular events to attend.

As a matter of fact, itinerary planning is often a difficult and time-consuming task for people, especially if they visit a destination for the first time. It involves substantial research to identify what are the main locations to visit, the time spending at each location, what should be next locations, and the time it will take to get from one place to another and of course how to reach the next places (public transportation, taxi, and so on).

Without any prior knowledge, at the moment a lot of people still rely on travel books or on recommendation travel blogs. However, these options have a lot of problems: travel books do not cover all cities/locations and, perhaps more importantly, are not free; travel blogs reflect a single person's view, with no guarantees that the information provided are reliable.

In this framework, *recommender systems* are becoming more and more important as applications that can be exploited to suggest products and provide people with useful information to facilitate their decision-making processes. In other words, such kind of system implicitly assumes that it can map user needs and constraints, through appropriate recommendation algorithms, and convert them into a number of locations-events-products selections using appropriate knowledge that is retrieved and analyzed by the system [Ricci et al, 2011].

In this paper, we propose a touristic context-aware recommendation system based on both the experience of previous users and on personal preferences of tourists and dependent on the context. The information is gathered by several heterogeneous sources (sensors, web portals, repositories related to touristic events and locations). All the information are eventually stored and analyzed in a cloud architecture that is particularly suitable to process and manage the huge amount of data extracted.

The paper is organized as in the following. Section 2 reports a brief literature overview on the evolution of recommender systems. Section 3 describes the proposed recommendation strategy for generation touristic paths. Section 4 outlines the system overview, while some preliminary experimental results are presented and discussed in Section 5. Finally, Section 6 reports some conclusion and the future work.

2. Literature Overview

As described in the previous Section, the focus of the paper lies in the introduction of context-aware recommendation services for touristic paths planning.

Within this framework, recommendation is generally known as the problem of estimating *ratings* - sometimes called also *utilities* - for the set of items that has not yet been seen by a given user [Ricci et al, 2011].

In *Content-Based* recommender systems, the rating r_j^i of item o_j is estimated using the utilities $r(u_i, o_k)$ assigned by the user u_i to items o_k that are in some way "similar" to item o_j .

One of the main drawbacks of these techniques is that the system can only recommend items that are similar to those already rated by the user itself (*overspecialization*).

Collaborative Filtering is, in the opposite, the process of filtering or evaluating items using the opinions of other people. Collaborative systems predict the rank of items r_j for a particular user u_i based on the utility $r(u_h, o_k)$ of items o_k previously rated by other users u_h “similar” to u_i . It takes its root from something human beings have been doing for centuries: sharing opinions with others.

Collaborative systems have their own limitations, which mainly relate to the *cold-start problem*, that describes situations in which a recommender is unable to make meaningful recommendations due to an initial lack of ratings. This problem can occur under three scenarios, namely, *new user*, *new item* and *new community*.

Content-based filtering and collaborative filtering are then usually combined by means of the *hybrid* approach that helps to avoid certain limitations of each method. Different ways to combine collaborative and content-based methods into a hybrid recommender system can be devised, namely: (a) implementing collaborative and content-based methods separately and combining their predictions; (b) incorporating some content-based characteristics into a collaborative approach; (c) incorporating some collaborative characteristics into a content-based approach; and (d) constructing a general unifying model that incorporates both content-based and collaborative characteristics.

In the area of recommendation systems, in the last few years, the use of additional contextual information has recently brought to the introduction of the *Context-aware Recommender Systems* (CARS) [Adomavicius et al, 2005]. In the *Contextual Pre-filtering* techniques context information are used to initially select the set of relevant items, while a classic recommender is used to predict ratings. In the *Contextual Post-filtering* approaches context are essentially used in the last step of the recommending process to “contextualize”, for each user, the output of a traditional recommender.

As it will be evident later in the paper, with respect to the described state of the art our approach can be classified as a hybrid strategy that incorporates some content-based characteristics into a collaborative strategy. It exploits system logs to implicitly derive information about individual users and the community of users as a whole, considering their past browsing sessions as a sort of *unary* ratings. Similarly to several collaborative filtering techniques, it is a kind of *active filtering* strategy in which past browsing sessions, modeled as a directed graph, determine the most suitable items to be recommended. Similarly, to information retrieval and filtering approaches, our approach gives high importance to the characteristics of the object a user is currently looking for, in order to effectively compute the utility of other items.

Eventually, as in a classical CARS, context information is opportunely exploited in a pre-filtering stage to determine the most suitable candidates for recommendation considering user preferences, and, in post-filtering stage to arrange the computed recommendations on the base of some environmental conditions.

3. A Recommendation strategy for building personalized touristic paths

The problem of building effective context-aware recommendation services, able to support an intelligent planning of touristic paths, implies the identification of “items” that are most likely to satisfy the interests of a user at any given point of his/her exploration depending on the context conditions.

In other terms, here we need to address some fundamental research questions:

1. How can we *model the context*?
2. How can we select a set of objects that are good candidates for a recommendation (*pre-filtering strategy*)?
3. How can we rank the set of candidates? In other words, which kind of *recommendation strategy* can we adopt?
4. How can we organize the recommended objects in visiting paths (*post-filtering strategy*)?

In our idea, to give an answer to each introduced question we assume that each interesting “touristic item” (e.g., hotel, restaurant, museum, etc.), composing a touristic path, is characterized by a set of metadata, corresponding to specific values of taxonomic attributes of an available *a-priori knowledge* useful for touristic applications. Moreover, each item is equipped with several context information, describing the “situation” for the place that is captured by apposite sensors.

2.1 Modeling the Context

The context is represented by means of the well-known *key-value model* [Adomavicius et al, 2005] that uses some pre-defined variables to describe several conditions of a visiting place.

In particular, we define four main classes of context parameters (*dimensions* of the context):

1. *time* dimension of the considered place (the relative time requested to user to reach the place, the opening/closing time, etc.);
2. *location* dimension for the considered place (the address, the actual position in terms of GPS coordinates, the relative position of user respect to the place, etc.);
3. *environmental* dimension for the considered place in terms of current weather and environmental conditions (e.g. temperature, humidity, rainfall degree, wind, season, moment of the day, etc.);
4. *social* dimension for the considered place in terms of number of users (e.g. visitors, guests, etc.) close to the considered place and the number of positive/negative feedbacks/comments.

A *context instance* is then given by a set of finite values related to the context dimensions.

In addition to the context data, we use a *Knowledge Base* containing a set of rules able to provide, in each moment and for each recommended, item a *comfort degree* on the base of some context parameters values.

For example, the following part of rule:

low_degree←-(*env_cond.temp*>30 AND *env_cond.humidity*>80%) OR...

indicates a low comfort degree for an outdoor place, such as an archeological site.

The comfort degree is eventually used in the post-filtering recommendation activity to arrange the order of items of the same type (if there is more than one alternative in the path). The context parameters values can be periodically computed using apposite web services or exploiting the sensors available on the place.

Once captured the context status of a place, the basic idea is that when a user is interested in the suggestion of a touristic path, the system:

1. receives the request and collects the information about user preferences/needs, for example the list of interesting items (e.g. museums, restaurants, hotel) and of the related constraints in terms of metadata values (e.g. he/she would like to visit museums with baroque pictures, then to eat in a cheap restaurant offering pizza, and finally to accommodate in a comfortable hotel near the sea);
2. selects a set of candidate objects for each type of objects that satisfy the user needs (pre-filtering strategy);
3. ranks these objects using a proper recommendation strategy;
4. arranges such objects in apposite touristic visiting paths considering the comfort degree and the available cartography (post-filtering strategy).

3.2 Pre-filtering

Each object subject to recommendation may be represented in different and heterogeneous metadata/*feature* spaces.

For instance, a hotel may be described by annotations concerning its kind (e.g. 4 star accommodation), position (near the sea), accommodation fee (180\$ for a night) and so on.

Each of these sets of features contributes to the characterization of the objects to different extents. Hence, it is important to consider congruently each type of descriptor during the recommendation process.

The first step consists in clustering together "similar" objects, where the similarity should consider all (or subsets of) the different spaces of features. To this purpose, we employ *high-order star-structured co-clustering* techniques [Ienco et al, 2013] to address the problem of heterogeneous data pre-filtering.

In this context, the same set of objects is represented in different feature spaces. Such data represent objects of a certain type, connected to other types of data, the features, so that the overall data schema forms a star structure of inter-relationships.

The co-clustering task consists in clustering simultaneously the set of objects and the set of values in the different feature spaces. In this way we obtain a partition of the objects influenced by each of the feature spaces and at the same time a partition of each feature space.

The pre-filtering stage leverages the clustering results to select a set of candidate objects by using the user's preferences, which are modeled as sets of descriptors in the same spaces as the objects' descriptors.

3.3 Objects' ranking

We use as recommendation strategy an importance ranking method that some of the authors have proposed in [Albanese et al, 2011].

Such a method combines metadata information of objects (items), past behavior of individual users and overall behavior of the whole community of users and context information.

Our basic idea is to assume that when an object o_i is chosen after an object o_j in the same browsing session, this event means that o_j "is voting" for o_i . Similarly, the fact that an object o_i is very similar to o_j can also be interpreted as o_j "recommending" o_i (and vice-versa).

Thus, our idea is to model a browsing system for a set of object O as a labeled graph (G, l) , where $G=(O, E)$ is a directed graph and $l: E \rightarrow \{pattern, sim\} \times R^+$ is a function that associates each edge in $E \subseteq O \times O$ with a pair (t, w) , where t is the type of the edge which can assume two enumerative values (*pattern* and *similarity*) and w is the weight of the edge.

According to this model, we can list two different cases:

- a *pattern label* for an edge (o_j, o_i) denotes the fact that an object o_i was accessed immediately after an object o_j and, in this case, the weight w_j^i is the number of times o_i was accessed immediately after o_j ;
- the *similarity label* for an edge (o_j, o_i) denotes the fact that an object o_i is similar to o_j and, in this case, the weight w_j^i is the similarity between o_j and o_i .

Thus, a link from o_j to o_i indicates that part of the importance of o_j is transferred to o_i .

Given a labeled graph (G, l) , we can formulate the definition of recommendation grade of a multimedia object more formally as follows.

Definition 3.1: (Recommendation Grade ρ)

$$\forall o_i \in O \quad \rho(o_i) = \sum_{o_j \in P_G(o_i)} w_j^i \rho(o_j) \quad (1)$$

where $P_G = \{o_j \in O \mid (o_j, o_i) \in E\}$ is the set of predecessors of o_i in G , and w_j^i is the normalized weight of the edge from o_j to o_i . For each $o_j \in O$, $\sum_{o_i \in S_G(o_j)} w_j^i = 1$ must hold, where $S_G = \{o_i \in O \mid (o_j, o_i) \in E\}$ is the set of successors of o_j in G .

It is easy to see that the vector $R = [\rho(o_1) \dots \rho(o_n)]^T$ can be computed as the solution to the following equation:

$$R = C \cdot R \quad (2)$$

where $C = \{w_{ij}\}$ is an ad-hoc matrix that defines how the importance of each object is transferred to other objects and can be seen as a linear combination of the following elements [Albanese et al, 2011]:

- A *local browsing matrix* $A_i = \{a'_{ij}\}$ for each user $u_i \in U$. Its generic element a'_{ij} is defined as the ratio of the number of times object o_i has been accessed by user u_i immediately after o_j to the number of times any object in O has been accessed by u_i immediately after o_j .
- A *global browsing matrix* $A = \{a_{ij}\}$. Its generic element a_{ij} is defined as the ratio of the number of times object o_i has been accessed by any user immediately after o_j to the number of times any object in O has been accessed immediately after o_j .
- A *similarity matrix* $B = \{b_{ij}\}$ such that $b_{ij} = \sigma(o_i, o_j) / \Gamma$ if $\sigma(o_i, o_j) \geq \tau \quad \forall i \neq j, 0$ otherwise. σ is any similarity function defined over O which calculates for each couple of objects their relatedness in terms of metadata descriptors; τ is a threshold and Γ is a normalization factors which guarantees that $\sum_i b_{ij} = 1$. In particular, we exploit a *semantic similarity* function based on the object metadata that has been computed used the *Li-Bandar-McLean* metric for semantic relatedness of concepts based on a vocabulary [Budanitsky and Hirst, 2001].

So far, we have a suitable manner to represent object features and to compare the related similarity also considering semantics in terms of object metadata [Amato et al, 2009]; now, our main goal is to compute customized rankings for each individual user.

In this case, we can then rewrite equation 2 considering the ranking for each user as follows:

$$R_i = C \cdot R_i \quad (3)$$

where $R_i = [\rho(o_1) \dots \rho(o_n)]^T$ is the vector of recommendation grades, customized for a user u_i .

We note that solving equation 3 corresponds to find the stationary vector of C , i.e., the eigenvector with eigenvalue 1. We demonstrated in [Albanese et al, 2010] - here all the details for computation of recommendation grades are also reported - that C , under certain assumptions and transformations, is a real square matrix having positive elements, with a unique largest real eigenvalue and the corresponding eigenvector has strictly positive components. In such conditions, equation 3 can be solved using the *Power Method* algorithm.

It is important to note that C does not have to be computed for all the database objects, but it needs to be computed only for those objects that are good candidates, i.e. the objects that are closest to users and in which a user is really interested in (pre-filtering strategy).

3.4 Post-Filtering

Finally, the list of suggested items is organized in apposite visiting paths considering the available cartography: they are not fixed and are arranged on the base of environmental situations and comfort degrees (post-filtering strategy).

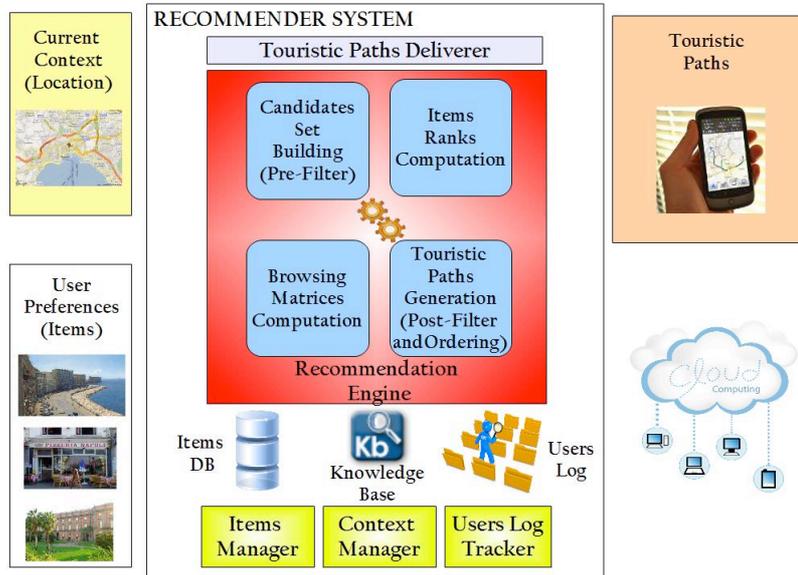


Fig.1 – The System Overview

4. The System Overview

Figure 1 shows an overview of our recommender system, which takes as input the current context in terms of user location and preferences (items of interest with the related metadata values) and generates a touristic path.

We can distinguish the following main components.

- **Items Manager** - A repository manager stores the items to be suggested with the related descriptions.
- **Users Log Tracker** - It is a module devoted to capture and store - in an appropriate format - all the users' browsing sessions in terms of accessed items during their explorations.

- **Context Manager** - It is a module devoted to gather from sensors, web portals and other kind of repository the context information exploiting proper *web services*.
- **Recommendation Engine** - It is the system core that for each user and on the base of current context dynamically proposes a set of recommended objects ordered on the base of their utility; in particular, it is composed by: (i) a *Browsing Matrices Computation Module* - able to transform the collected browsing sessions into three matrices (a global matrix which takes into account the overall browsing behavior of the users, a local matrix which considers the behavior of a single user and a similarity matrix storing the objects' semantic similarity); (ii) a *Candidate Set Building Module* - computes the subset of items that fit with users' needs; (iii) a *Items Ranks Computation Module* - performs the ranking of the selected candidates for recommendation; (iv) a *Touristic Paths Generation Module* - capable of arranging the suggested items on the base of environmental situations and the comfort degrees.
- **Items Deliverer** - It aims at delivering recommended paths to each user in a format that will depend on the user profile and device.

All the modules have been realized using *JAVA* technologies, *PostgreSQL* as relational DBMS and *OWLIM/Sesame* for storing objects' metadata as set of RDF triples. The Items DB is constituted by a collection of about 5,000 touristic places related to several famous Italian cities (e.g., Naples, Florence and Rome) with the relative metadata description and some multimedia information.

5. Preliminary Experiments

Previous research work on recommender system evaluation has mainly focused on algorithm *accuracy*, especially objective prediction accuracy. More recently, researchers began examining issues related to users subjective opinions and developing additional criteria to evaluate recommender systems. In particular, they suggest that user's satisfaction does not always (or, at least, not only) correlate with the overall recommender's accuracy and evaluation frameworks for measuring the perceived qualities of a recommender and for predicting user's behavioural intentions as a result of these qualities should be taken into account.

We carried out several preliminary experiments to investigate how helpful the recommendations offered by our system are, demonstrating that the introduction of such techniques can improve the tourists' experience and satisfy their needs.

In the training phase, we exploited several thousand of users' *logs* to capture their browsing sessions in terms of visited places and to build a consistent matrix *A*. Successively, we asked a group of about 50 people to express several preferences/needs for touristic paths in a given city.

Finally, we measured the user satisfaction degree with respect to the related top 10 paths generated by the systems in a scale ranging from 1 to 5 (as a sort of rating in a recommender) and observed the average values in Table 1.

Item Rank	User Satisfaction
1	4,8
2	4,2
3	4
4-10	2,78

Table 1 – Average User Satisfaction Values

6. Conclusions

Our proposal represents a “context-aware” extension (pre-filtering and post-filtering activities) of a recommender strategy defined by some of the authors in previous works able to support touristic paths planning. We have shown that the customized recommendations may be computed combining several features of objects, past behavior of individual users and overall behavior of the entire community of users and using user preferences and context information as pre-filtering and post-filtering strategies, respectively.

Experimental results showed that our approach is quite promising and encourages further research.

Bibliografia

[Adomavicius et al, 2005] Adomavicius, G., Sankaranarayanan, R., Sen, S., Tuzhilin, A., Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS)*, 23, 1, 2005, 103–145.

[Albanese et al, 2010] Albanese, M., d’Acierno, A., Moscato, V., Persia ,F., Picariello, A.: Modeling recommendation as a social choice problem. *Proceedings of ACM Conference on Recommender Systems (RecSys 21010)*, 329–332, 2010

[Albanese et al, 2011] Albanese, M., d’Acierno, A., Moscato, V., Persia ,F., Picariello, A.: A multimedia semantic recommender system for cultural heritage applications. *Proceedings of IEEE Conference on Semantic Computing (ICSC 2011)* , 403–410, 2011

[Amato et al, 2009] Amato, F., Mazzeo, A., Moscato, V., Picariello, A. (2009). A system for semantic retrieval and long-term preservation of multimedia documents in the e-government domain. *International Journal of Web and Grid Services*, 5(4), 323-338.

[Budanitsky and Hirst, 2001] Budanitsky, A., Hirst, G., Semantic distance in wordnet: An experimental, application oriented evaluation of five measures. *Proceedings of the Workshop on WordNet and other Lexical Resources*, 2001.

[Ienco et al, 2013] Ienco, D., Robardet, C., Pensa, R.G., Meo, R., Parameter-less co-clustering for star-structured heterogeneous data. *Data Mining and Knowledge Discovery*, 26, 2, 2013, 217-254.

[Ricci et al, 2011] Ricci, F., Rokach, L., Shapira, B., Kantor, P.B., *Recommender Systems Handbook*. Springer, 2011.

Usability in software development company practices

Carmelo Ardito, Paolo Buono, Maria Francesca Costabile,
Rosa Lanzilotti, Antonio Piccinno
IVU Lab, Dipartimento di Informatica, Università di Bari "Aldo Moro"
via Orabona 4, 70125 Bari
{name.surname}@uniba.it

Abstract. *The importance of adopting usability engineering methods in software development is eventually recognized by an increasing number of companies. However, several studies show that, in spite of the declared awareness of usability as an important software quality and the benefits reported in the literature, many companies still neglect in their development practices those activities that are essential to ensure that their products are usable and capable of generating a satisfying user experience. This paper describes the case of the project K-People, funded by "Regione Puglia", in which usability engineering methods were employed in the development of a company web portal. It can be seen that the designers' attitude completely changes once they directly experience how effective such methods really are in the design and development of quality software.*

Keywords: software life cycle, human-centered design, usability methods.

1. Introduzione

L'usabilità di un sistema software è la misura in cui i suoi utenti sono in grado di utilizzarlo per compiere specifiche attività [ISO/IEC 1998]. Un basso livello di usabilità implica che gli utenti hanno difficoltà ad utilizzare un sistema, indipendentemente da quanto siano complesse le sue funzionalità [Nielsen 1993]. Questo si ripercuote negativamente sull'esperienza d'uso del sistema. Oggi, infatti, l'attenzione dei ricercatori si sta spostando sulla esperienza d'uso (User eXperience, UX), cioè si studia come motivare, attrarre e coinvolgere gli utenti nell'interazione con il sistema. Progettare per la UX richiede la comprensione delle esigenze degli utenti, sia da un punto di vista pragmatico (p.e. funzionalità del sistema e l'interazione) sia edonistico [Väänänen-Vainio-Mattila et al. 2008]. La UX estende il concetto più tradizionale di usabilità, focalizzata principalmente sulla facilità d'uso, ed enfatizza gli attributi soggettivi come l'estetica, le emozioni e il coinvolgimento sociale. L'usabilità rimane un attributo importante per una buona UX, anche se talvolta

le persone sono disposte a rinunciare in parte all'usabilità a fronte di una migliore UX. L'esempio più eclatante è dato dal successo che hanno avuto, appena immessi sul mercato, strumenti come iPod e iPhone che, pur presentando alcune funzionalità oggettivamente meno usabili rispetto a prodotti concorrenti, hanno avuto un grosso successo per l'*appeal* che suscitavano nei loro utenti.

La metodologia di progettazione human-centered (HCD) è il punto di partenza per prendere in considerazione usabilità e UX nel ciclo di vita del software [ISO/IEC 1998]. Tale metodologia enfatizza l'importanza di concentrarsi sugli utenti, i loro compiti e il contesto in cui essi operano e di effettuare una progettazione iterativa mediante la creazione e la valutazione di prototipi di complessità crescente.

Fin dai primi anni '90, sono stati proposti vari metodi per l'usabilità e, più recentemente per la UX, che rientrano in quella che viene definita l'*ingegneria dell'usabilità* [Nielsen 1993]. Una rassegna dettagliata è in [Scapin e Law 2007]. Molta ricerca è stata condotta per integrare efficacemente tali metodi nelle pratiche di sviluppo del software più diffuse. È stato inoltre dimostrato come la valutazione di usabilità nei sistemi software porti ad un aumento delle vendite [Dray e Karat 1994], aumento della produttività [Karat 1997], riduzione dei costi di addestramento degli utenti [Dray e Karat 1994] e riduzione dei costi di assistenza [Reed 1992]. Inoltre, nonostante gli sviluppatori dichiarino che l'usabilità è più importante delle funzionalità del sistema stesso [Frese e Hesse 1995], la letteratura fornisce studi che mostrano come nelle aziende di software le attività di ingegneria dell'usabilità siano limitate o, addirittura, assenti. Sembra quindi che, nella pratica, i benefici dell'usabilità abbiano un impatto minimo. Alcuni ricercatori affermano che, al fine di ottimizzare l'impatto dell'usabilità e della UX sullo sviluppo del software, è fondamentale analizzare le attuali pratiche di sviluppo aziendali, andando a verificare direttamente come gli sviluppatori lavorano all'interno dei loro processi di sviluppo [Boivie et al. 2003; Lethbridge et al 2005; Robinson et al. 2007].

In linea con questo approccio, questo articolo riporta l'esperienza del progetto *K-People*, che mostra come la collaborazione con esperti di ingegneria dell'usabilità è stata importante per far acquisire all'azienda, coordinatrice del progetto, la consapevolezza che vari metodi possono essere incorporati nei processi di progettazione e sviluppo di software senza richiedere costi eccessivi e dando, invece, un contributo rilevante alla qualità del software prodotto. Tale esperienza dimostra che per cambiare la mentalità attuale non basta leggere articoli o seguire corsi di formazione che illustrino il valore dei metodi di ingegneria dell'usabilità, ma bisogna indurre progettisti e sviluppatori ad applicare e sperimentare tali metodi nella pratica per consentire di "toccare con mano" non solo i vantaggi che ne derivano, ma comprendere che taluni metodi effettivamente richiedono poche risorse pur producendo risultati di valore.

L'articolo è organizzato come segue. La sezione 2 riporta brevemente i risultati di due studi che avevano lo scopo di indagare sull'adozione di pratiche di ingegneria dell'usabilità in aziende che sviluppano software. La sezione 3 riassume il progetto *K-People*. La sezione 4 descrive la progettazione "human-centred" del portale web sviluppato durante il progetto. La sezione 5 commenta

i risultati di interviste a personale dell'azienda che ha lavorato al progetto *K-People*. La sezione 6 conclude l'articolo.

2. L'usabilità nella pratica dello sviluppo di software

Lo studio riportato in [Bak et al. 2008] aveva l'obiettivo di analizzare le pratiche aziendali relative all'utilizzo di metodi di ingegneria dell'usabilità nella progettazione e sviluppo di sistemi software. Ad aziende che sviluppano software, è stato sottoposto un questionario sui metodi da esse utilizzati per la valutazione di usabilità dei loro prodotti. Il questionario chiedeva inoltre di evidenziare sia i vantaggi che derivano dall'utilizzo di tali metodi, sia gli eventuali ostacoli. Lo studio ha coinvolto 39 aziende che operano in una specifica regione nel nord della Danimarca.

Tale studio ha ispirato un'analoga indagine effettuata in un'area completamente differente, un'area mediterranea e precisamente la Regione Puglia [Ardito et al. 2011]. L'indagine ha replicato quella della Danimarca nel senso che è stato utilizzato il medesimo questionario, somministrato ad aziende con caratteristiche simili a quelle danesi. I risultati sono facilmente comparabili. Da un lato evidenziano delle differenze nella comprensione dell'usabilità; in particolare tale concetto nello studio italiano è meglio compreso anche se molti fanno ancora molta confusione. Questo dimostra che con il passare del tempo, gli sviluppatori stanno diventando coscienti di ciò che rappresenta l'usabilità e l'importanza delle valutazioni di usabilità. Dall'altro lato, il numero di aziende che utilizzano metodi di ingegneria dell'usabilità nei loro processi di progettazione e sviluppo di software è simile e, purtroppo, è ancora una percentuale piuttosto bassa. I principali vantaggi riportati dai rispondenti di entrambe le indagini sono: il miglioramento della qualità del prodotto; l'aumento della soddisfazione degli utenti; la maggiore competitività del prodotto.

Uno dei principali problemi evidenziati dallo studio è che ancora troppi sviluppatori professionisti non sanno bene che cosa è l'usabilità e sanno ancora meno di UX.. Molti di loro sono interessati principalmente alle funzionalità del sistema e all'efficienza del codice, cioè a qualità del software che sono di grande interesse per chi sviluppa il sistema ma hanno poco impatto sugli utenti finali.

Effettivamente, nella formazione di progettisti e sviluppatori che oggi sono attivi sul mercato c'è stata carenza rispetto a principi e teorie per la creazione di sistemi interattivi usabili; la rassegna in [Ardito et al. 2011] ha evidenziato che il termine "usabilità" da molti non è ancora ben compreso; talvolta è confuso con "accessibilità", forse sulla base della legge n.4 del 9 gennaio 2004, comunemente chiamata "Legge Stanca". Solo negli ultimi anni la disciplina di Interazione Uomo-Macchina è stata inserita nel curriculum dei corsi di laurea in Informatica, anche se in molte sedi è ancora considerata una disciplina non obbligatoria. Di conseguenza, come riporta Roberto Polillo nella prefazione del suo libro [Polillo 2010], la formazione dei progettisti è inadeguata per il contesto dei nuovi sistemi interattivi "perché prescinde totalmente dallo studio dell'uso dei sistemi. Ai progettisti viene insegnato a trovare soluzioni tecniche a problemi tecnici. Non viene mai insegnato a sollevare – sia pure per un momento – lo

sguardo dal codice dei programmi o dagli schemi tecnici, per riflettere sul senso di ciò che stanno facendo, e di esaminare l'effetto delle loro scelte progettuali sull'attività degli utenti".

L'indagine ha anche evidenziato che non c'è una corretta conoscenza di quali siano i metodi appropriati per costruire sistemi usabili senza spendere risorse eccessive. Infatti, proprio le risorse necessarie in termini di costo, tempo e persone coinvolte viene avvertito come uno dei principali ostacoli. C'è poi, secondo le aziende interpellate, una mancanza di metodi che possano essere effettivamente integrati nelle pratiche di progettazione e sviluppo da loro adottate. Infine, le aziende dichiarano di avere molte difficoltà nel reperire utenti da coinvolgere nelle valutazioni di usabilità e questo indica anche che non sono consapevoli che ci sono metodi di valutazione di usabilità che non ne richiedono la presenza o che anche un test con paio di utenti dà risultati molto utili.

In questa situazione è stata particolarmente stimolante la richiesta che ci è pervenuta da parte della società Webscience per collaborare nel progetto *K-People* [K-People 2009], finanziato dalla Regione Puglia. Il contributo richiesto al nostro gruppo di ricerca in qualità di esperti di ingegneria dell'usabilità consisteva nell'identificare metodologie e tecniche da utilizzare nelle varie fasi di costruzione del portale web relativo al progetto *K-People* affinché risultasse usabile e capace di produrre un'efficace esperienza dell'utente [Hassenzahl e Tractinsky 2006]. La richiesta è stata particolarmente stimolante: ci dava la possibilità di dimostrare nella pratica cosa significa progettare per l'usabilità e come applicare i metodi appropriati.

3. Il progetto *K-People*

Il progetto *K-People*, finanziato dalla Regione Puglia e coordinato dall'azienda Webscience S.r.l., ha portato allo sviluppo di un sistema che supporta gli impiegati di un'azienda nella gestione di informazioni relative a processi destrutturati in cui, cioè, non è predefinibile la sequenza delle azioni, gli input necessari e gli attori coinvolti (ad esempio i processi di tipo decisionale, creativo, collaborativo, ...). Tali processi si differenziano da quelli formali e strutturati per diversi aspetti: dipendono dal contesto, e quindi sono difficilmente standardizzabili; prevedono intense relazioni collaborative; sono fortemente dipendenti dalla discrezionalità e dal giudizio umano; si basano su dinamiche comportamentali non prevedibili; sono difficilmente controllabili e misurabili. In questi processi, l'enfasi va spostata sugli utenti coinvolti, supportandoli con strumenti che consentano di utilizzare nel modo migliore le risorse che l'azienda mette a disposizione, assistendoli in tutte le condizioni e assecondando le loro abitudini di utilizzo di diversi strumenti di comunicazione, anche in mobilità.

Il sistema *K-People* è composto di un insieme di moduli Open Source in grado di integrarsi con i sistemi informativi aziendali tradizionali e con i sistemi di produttività individuale. L'obiettivo è supportare l'utente, denominato "knowledge worker", nel reperimento delle informazioni e nella gestione delle proprie attività nei processi destrutturati. L'utente interagisce con il sistema attraverso un portale web che gli consente di gestire processi strutturati e destrutturati, attivando pattern comportamentali (una serie di attività predefinite

che gli utenti svolgono usualmente all'interno di un processo) e monitorando flussi operativi complessi, quali la richiesta di contributi per l'esecuzione di specifiche attività, l'invio di solleciti a persone coinvolte in attività collaborative, la definizione di decisioni collaborative, la pianificazione di una riunione, ecc.

Per abilitare questa visione innovativa dei sistemi informativi aziendali, il sistema *K-People* consente:

- la gestione integrata di informazioni strutturate e destrutturate, mantenendo una visione orientata ai processi;
- l'integrazione di soluzioni per il supporto di processi aziendali destrutturati, in particolare soluzioni integrate per la comunicazione interna ed esterna all'impresa, per la collaborazione, per la gestione di documenti, per la business intelligence, per la creazione e il supporto di social networking;
- la creazione di uno spazio di lavoro personalizzato sulle esigenze del singolo lavoratore, che risponde alle nuove esigenze del knowledge worker, rivolgendo l'attenzione alla possibilità di effettuare lavoro multicanale, attraverso cellulare, mediante ambienti desktop web configurabili e attraverso applicativi per l'office automation.

Rispetto allo stato dell'arte, *K-People*, possiede un elevato potenziale innovativo, in quanto consente di rivoluzionare il tradizionale approccio alla gestione delle informazioni aziendali, permettendo all'utente di modellare le funzionalità e i canali di accesso del sistema informativo intorno alle proprie esigenze e in funzione del contesto di utilizzo.

4. Progettazione Human-Centred del portale *K-People*

Se in passato i bisogni delle persone che avrebbero utilizzato un prodotto software erano trascurati, perché gli sviluppatori creavano prodotti per persone molto simili a loro, oggi non può essere più così: si pensi al numero sempre crescente di persone che utilizzano i sistemi informatici per supporto alle loro attività quotidiane di lavoro, di intrattenimento o di mera ricerca di informazioni. I sistemi che oggi si sviluppano devono permettere agli utenti di concentrarsi sui loro compiti e non sulle modalità per eseguire tali compiti. Occorrono, quindi, tecniche che aiutino a cambiare il modo con cui si progettano i sistemi; metodi che siano centrati sulle persone che utilizzeranno tali sistemi e che tengano conto delle loro capacità e necessità [Costabile 2001].

Nell'HCD è di primaria importanza la fase iniziale di analisi dei requisiti, in cui l'enfasi non deve essere data solo ai requisiti hardware o alle specifiche funzionalità software, ma bisogna analizzare in dettaglio chi utilizzerà il sistema, i compiti che dovrà eseguire e il contesto d'uso. Solo a questo punto il prodotto può iniziare a essere progettato e la fase di progettazione dovrà concentrarsi sulla possibile interazione tra utente e sistema, sviluppando prototipi dell'interfaccia utente a vari livelli, da quelli iniziali su carta a quelli più avanzati. È fondamentale la valutazione formativa di questi prototipi per definire un prodotto che sia adeguato ai requisiti degli utenti finali.

Prima di procedere con la progettazione human-centred del sistema *K-People*, ci siamo resi conto che era necessario creare consapevolezza nei dipendenti dell'azienda Webscience su cosa significa l'usabilità e come poter

progettare per l'usabilità, visto che dalle indagini in [Bak et al. 2008] e [Ardito et al. 2011] era emersa quanta confusione c'è ancora su questi concetti.

A questo scopo abbiamo tenuto un corso breve di una giornata presso l'azienda, rivolto ai dipendenti e al management, in cui sono stati illustrati modelli, metodologie e tecniche dell'ingegneria dell'usabilità. In particolare, è stato illustrato l'HCD, enfatizzando anche l'importanza del coinvolgimento degli utenti finali nella progettazione, in qualità di esperti del dominio applicativo.

Seguendo l'approccio di "participatory design", secondo cui è opportuno costituire un team di progetto che comprende vari stakeholders del progetto con competenze in discipline diverse [Schuler e Namioka 1993], per il progetto *K-People* è stato creato un team multidisciplinare composto da esperti di Interazione Uomo-Macchina (ricercatori dell'IVU Lab dell'Università di Bari), ingegneri del software di Webscience S.r.l. e rappresentanti degli utenti finali scelti tra il management e tra gli altri dipendenti dell'azienda addetti al settore tecnico e a quello amministrativo. La scelta di questi tipi di utenti è dovuta allo scopo del sistema (supportare processi aziendali destrutturati) e alla facilità di reperirli, essendo l'azienda Webscience coordinatrice del progetto. Quindi, trovare utenti disponibili è spesso molto più facile di quanto si pensi. Se si considera, ad esempio, quante applicazioni di e-government, e-commerce, e-health si rivolgono a persone di vario tipo, si comprende che si possono fare test veloci coinvolgendo amici e/o persone di famiglia, variando caratteristiche come età, sesso, livello di cultura, esperienza nell'uso delle tecnologie.

Nelle fasi iniziali di analisi, tramite riunioni, brainstorming, interviste e visite sul posto, sono stati creati i profili di utente e, per ciascuno, analizzati i compiti di interesse. Sono stati, inoltre, disegnati scenari di uso del sistema. Successivamente, sono state create soluzioni progettuali consistenti in modelli e schemi di navigazione, prototipi a bassa fedeltà sotto forma di schizzi su carta delle schermate principali che sono stati discussi con il team di progettazione.

Gli utenti finali sono stati coinvolti non solo per la valutazione ma anche per la progettazione del sistema *K-People*. Jacob Nielsen, nell'intervista pubblicata in [Preece et al. 2002], suggerisce un modello "a sandwich", che prevede l'alternanza di attività di progettazione e valutazione. Nielsen consiglia di partire con un test con anche solo 4-5 utenti (user test) di due o tre proposte progettuali preliminari e, in base alle indicazioni che ne derivano, sviluppare e migliorare quelle più promettenti. Successivamente, eseguire una valutazione euristica. Poi evolvere ulteriormente il progetto, fare qualche test rapido con utenti, evolvere e fare valutazione euristica, e così via. La valutazione euristica è molto utile soprattutto per un progetto molto grande e, come è noto, è un metodo che ha un ottimo rapporto costi/benefici e non richiede utenti, ma solo l'esperto di usabilità che esegue l'ispezione.

Poiché il costo delle valutazioni di usabilità è sentito come uno dei maggiori problemi da parte delle aziende, il nostro approccio è di utilizzare metodi, come la valutazione euristica e i test condotti con un numero ridotto di utenti, che Nielsen definisce "discounted usability", cioè metodi che sono efficienti in rapporto al loro costo molto limitato.

Gli sviluppatori di *K-People* non avevano mai preso in considerazione la possibilità di fare un semplice test anche con un solo utente. Sono rimasti

particolarmente colpiti nel constatare quanti problemi si possono evidenziare con una sola persona che interagisce con un prototipo a bassa fedeltà. Si sono, pertanto, resi conto di quanto sia vantaggioso l'utilizzo di tali prototipi e quanto poco costoso sia farlo valutare da un utente finale il cui modello mentale del compito da eseguire è tanto diverso da quello del progettista [Norman 2000].

La valutazione euristica non coinvolge direttamente gli utenti finali, ma solo gli esperti di usabilità, i quali ispezionano i prototipi o i sistemi e identificano, in accordo con i principi di usabilità (euristiche), gli elementi che potrebbero causare problemi agli utenti [Nielsen e Mack 1994]. I prototipi del portale di *K-People* sono stati ispezionati da cinque esperti di Interazione Uomo-Macchina che hanno usato le euristiche di Nielsen per valutarne la conformità rispetto a: *Apprendibilità, Efficienza, Memorabilità, Basso numero di errori, Soddisfazione nell'uso*. Ogni ispettore ha riportato i problemi individuati in un proprio report di ispezione. Gli ispettori si sono poi riuniti per discutere i problemi riscontrati individualmente e hanno prodotto un report di valutazione complessivo contenente i problemi di usabilità, distribuiti rispetto alla gravità in un intervallo da 1 (puramente estetico) a 4 (catastrofico). Il report è stato quindi sottoposto all'attenzione del team multidisciplinare per discutere i problemi riscontrati, trovare le soluzioni migliori e procedere alla creazione di un nuovo prototipo.

Il processo iterativo di progettazione di un prototipo e sua valutazione utilizzando alternativamente valutazione euristica e user testing con un paio di utenti è continuato fino alla generazione di un prototipo a più alta fedeltà che soddisfacesse i requisiti identificati. Diversamente da quanto possa apparire, questo processo iterativo non è particolarmente costoso perché si lavora con prototipi dell'interfaccia utente che si possono realizzare facilmente e che sono valutati con metodi che richiedono poche risorse. Il prototipo ad alta fedeltà è stato valutato con un numero più significativo di utenti. Dieci utenti sono stati osservati, uno per volta e in sessioni singole, da un ricercatore esperto di usabilità, mentre un dipendente dell'azienda Webscience ha svolto il ruolo di facilitatore, intervenendo per sostenere l'utente quando era in difficoltà (senza suggerire nulla) o invitandolo a passare al compito successivo se si bloccava per molto tempo. Ogni utente ha eseguito sette compiti su un laptop collegato ad Internet, tramite cui ha potuto accedere al portale del sistema *K-People*.

L'analisi dei dati raccolti durante questo test ha evidenziato ulteriori problemi, diversi da quelli scoperti nelle precedenti valutazioni, che talvolta hanno impedito ad un utente di portare correttamente a termine un compito. Il team multidisciplinare si è riunito nuovamente per discutere tali risultati e risolvere i problemi che avevano messo in difficoltà gli utenti. Una schermata significativa della versione finale del portale *K-People* è mostrata in Fig. 1.

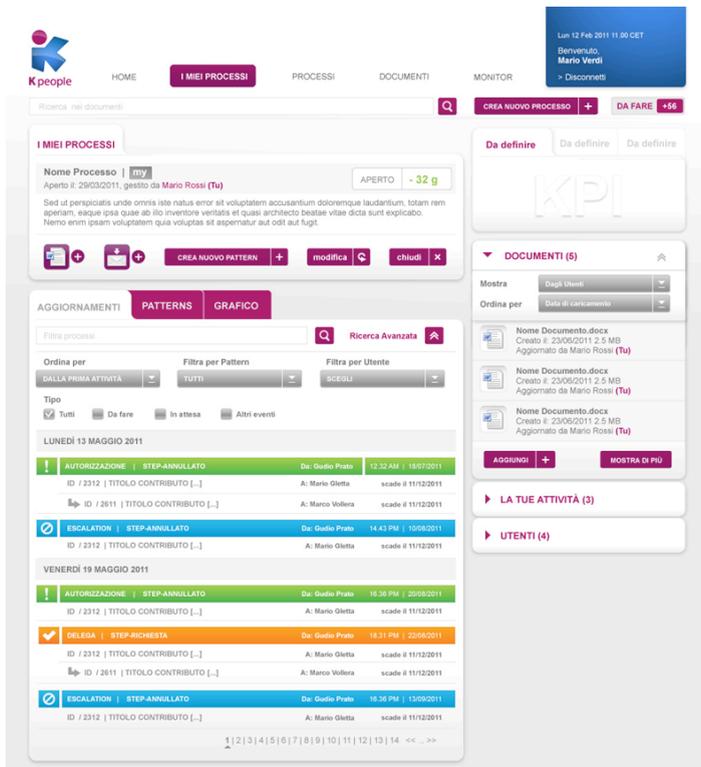


Fig. 1. Una schermata del prototipo finale del portale *K-People*

5. Interviste finali

Alla fine del progetto, sono stati intervistati tre dipendenti di Webscience che avevano lavorato nel progetto *K-People*, che per la prima volta avevano avuto un impatto sulle problematiche di usabilità nella pratica del loro lavoro. Scopo principale era capire se gli intervistati avevano apprezzato il nuovo approccio allo sviluppo dei sistemi e se intendono adottarlo per il futuro. Sono stati intervistati: un responsabile della progettazione con mansioni manageriali nell'azienda, un analista e uno sviluppatore. Le interviste, di tipo semi-strutturato, sono state svolte singolarmente utilizzando Skype™, sono state registrate e sono durate ognuna circa 45 minuti.

Le opinioni più significative per i nostri scopi sono state espresse dal primo intervistato. Infatti, lo sviluppatore ha avuto un impatto minore rispetto al nuovo approccio alla progettazione, fase in cui non è direttamente coinvolto. L'analista ha apprezzato molto l'approccio HCD, lo sviluppo mediante prototipi a bassa fedeltà e gli altri metodi utilizzati, ma ha affermato esplicitamente che il suo ruolo non gli consente di intervenire nella decisione dell'azienda di applicare l'HCD anche in futuro. Il responsabile della progettazione ha colto pienamente le potenzialità dell'approccio e ha sottolineato con entusiasmo quanto sia lui che altri suoi colleghi in Webscience si siano resi conto che metodi veloci di

valutazione di usabilità, utilizzati già nella prime fasi di progettazione di un sistema su prototipi a basso costo, riescono ad identificare problemi che sarebbe eccessivamente costoso risolvere nelle fasi avanzate dello sviluppo. Tuttavia, anche lui, rispetto alla domanda critica se adottare anche in futuro l'approccio HCD, ha detto chiaramente che bisogna discuterlo con i responsabili dell'azienda. Dunque, è evidente che, anche in aziende aperte all'integrazione di tecniche di ingegneria di usabilità nelle proprio pratiche di sviluppo, come Webscience, bisogna ancora lavorare per cambiare la mentalità, non solo di progettisti e sviluppatori, ma anche di coloro che sono ai vertici dell'azienda.

6. Conclusioni

Studi effettuati con piccole e medie aziende software hanno evidenziato che gli sviluppatori stanno diventando più consapevoli di ciò che è l'usabilità del software e dell'importanza della valutazione dell'usabilità [Bak et al. 2008], [Ardito et al. 2011], [Ardito et al. 2013]. Tuttavia, il numero di aziende che eseguono valutazioni di usabilità in maniera sistematica e appropriata è ancora ridotto. Nonostante le aziende affermino che ci sono vantaggi nel progettare sistemi usabili, quali maggiore soddisfazione dell'utente, aumento della competitività del prodotto e dell'azienda, esse percepiscono degli ostacoli significativi che le scoraggiano a utilizzare metodi di ingegneria dell'usabilità.

Per cercare di modificare tale atteggiamento, abbiamo colto con entusiasmo la proposta di Webscience di collaborare nel progetto *K-People*. Il nostro intento era mostrare in modo operativo che esistono metodi di ingegneria dell'usabilità che, pur richiedendo poche risorse in termini di costi, persone e tempo, sono molto efficaci. Abbiamo applicato tali metodi nello svolgimento del progetto e non solo abbiamo ottenuto un prodotto che meglio soddisfa le esigenze dei suoi utenti, ma i progettisti che hanno collaborato con noi hanno verificato personalmente quanto ingiustificati siano i loro preconcetti sull'applicabilità dei metodi dell'ingegneria dell'usabilità alle pratiche di sviluppo del software. Va sottolineato che si deve ancora lavorare molto perché questa consapevolezza sia acquisita anche dai vertici delle aziende, perché sono loro che hanno il potere decisionale di modificare i processi di sviluppo per creare sistemi usabili.

Ringraziamenti

Si ringraziano EU e Regione Puglia che hanno finanziato questo lavoro attraverso il progetto di ricerca *K-People*, programma F.E.S.R., P.O. Regione Puglia 2007-2013 - Asse I Linea 1.1 – Azione 1.1.2. Aiuti agli Investimenti in Ricerca per le PMI. Si ringrazia, inoltre, Webscience S.r.l.

Bibliografia

- [Ardito et al. 2011] Ardito, C., Buono, P., Caivano, D., Costabile, M.F., Lanzilotti, R., Bruun, A., Stage, J., Usability evaluation: a survey of software development organizations, in Proc. of *SEKE 2011* (Knowledge Systems Institute), 282-287.
- [Ardito et al. 2013] Ardito, C., Buono, P., Caivano, D., Costabile, M.F., Lanzilotti, R. Investigating and promoting UX practice in industry: an experimental study. Submitted for publication.

- [Bak et al. 2008] Bak, J.O., Nguyen, K., Risgaard, P., Stage, J., Obstacles to usability evaluation in practice: a survey of software development organizations, in Proc. of *NordiCHI 2008 (ACM)*, 23-32.
- [Boivie et al. 2003] Boivie, I., Aaborg, C., Persson, J., Lofberg, M., Why usability gets lost or usability in in-house software development, *Interacting with Computers* 15(4), 2003, 623-639.
- [Costabile 2001] Costabile, M.F., Usability in the software life cycle, in *Handbook of Software Engineering and Knowledge Engineering*, ed. Chang, S. K., (World Scientific Publishing, 2001), 179-192.
- [Dray e Karat 1994] Dray, S.M., Karat, C.M., *Human factors cost justification for an internal development project*, (Academic Press, Inc., 1994), 111-122.
- [Frese e Hesse 1995] Frese, M. Hesse, W., The work situation in software-development. *Methods & Tools* (1995).
- [Hassenzahl e Tractinsky 2006] Hassenzahl, M., Tractinsky, N., User experience - a research agenda, *Behaviour & Information Technology* 25(2), 2006, 91-97.
- [ISO/IEC 1998] ISO/IEC, *9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability*, 1998.
- [K-People 2009] K-People, "Programma F.E.S.R., P.O. Regione Puglia 2007-2013 - Asse I Linea 1.1 – Azione 1.1.2. Aiuti agli Investimenti in Ricerca per le PMI," 2009; <http://kpeople.webscience.it/>.
- [Karat 1997] Karat, C.M., Cost-justifying usability engineering in the software life cycle., in *Handbook of Human-Computer Interaction*, eds. Helander, M., Landauer, T. K. and Prabhu, P., (Elsevier, 1997), 767-778.
- [Lethbridge et al. 2005] Lethbridge, T.C., Sim, S., Singer, J., Studying Software Engineers: Data Collection Techniques for Software Field Studies. *Empir Software Eng*, 10(3), 311-341.
- [Nielsen 1993] Nielsen, J., *Usability Engineering* (Morgan Kaufmann, 1993).
- [Nielsen e Mack 1994] Nielsen, J., Mack, R.L., *Usability inspection methods* (John Wiley & Sons, Inc., 1994).
- [Norman 2000] Norman, D., *The Design of Everyday Things* (MIT Press, 2000).
- [Polillo 2010] Polillo, R., *Facile da Usare* (Apogeo, 2010).
- [Preece et al. 2002] Preece, J., Rogers, Y., Sharp, H., *Interaction Design* (John Wiley & Sons, Inc., 2002).
- [Reed 1992] Reed, S., Who defines usability? You do!, *PC Comp.Dec.* (1992), 220-232.
- [Robinson et al. 2007] Robinson, H., Segal, J., Sharp, H., Informed empirical studies of software practice. *Information and Software Technology*, 49(6), 540-551.
- [Scapin e Law 2007] Scapin, D., Law, E., Refine Usability Evaluation Methods (R3-UEM) Proc. of Mause International Workshop on Review, COST 294. 2007.
- [Shneiderman 2002] Shneiderman, B., *Leonardo's laptop: human needs and the new computing technologies* (MIT Press, 2002).
- [Schuler e Namioka 1993] Schuler, D., Namioka, A., *Participatory design: principles and practices*. (L. Erlbaum Associates, 1993).
- [Väänänen-Vainio-Mattila et al. 2008] Väänänen-Vainio-Mattila, K., Roto, V., Hassenzahl, M., Now let's do it in practice: user experience evaluation methods in product development, in Proc. of *CHI '08 (ACM, 2008)*, 3961-3964.

Using the USherlock tool to evaluate GUI usability: a comparison with canonical testing methods

Rosanna Cassino and Maurizio Tucci
Department of Management and Information Technology,
University of Salerno, Fisciano, ITALY, 84084
{rcassino, mtucci}@unisa.it

Abstract. *The usability of web sites and software applications is a key issue for their success. Nowadays, with the development of the digital society, end-users can easily access to an extraordinary amount of software applications and websites, so usability validation tools can greatly help them to cope with this extremely wide available choice. This work presents the usability evaluation results of some graphical user interfaces performed by the automatic tool USherlock, in comparison with standard testing. Unlike other existing tools, which usually perform usability analysis based on the examination of source code, USherlock only works on the front-end of a GUI, whose source code is unknown or unavailable. The tool derives the structure of the interface in terms of its static and interactive elements, in a deductive manner. Then, the system automatically simulates the activation of each interactive element of the GUI to perform a dynamic usability evaluation according to a number of well-known usability metrics. The present experimental analysis was carried out on several web pages, and the results are compared to the testing made by canonical methods: user questionnaires are developed on a sample of students tester of IUMUS course of University of Salerno. The outcoming of the experiment shows that similar results are obtained, motivating the development of an automatic tool like USherlock to reduce the costs and times of classical tests.*

Keywords: GUI Evaluation, Usability test, Usability Evaluation Tools.

1. Introduction

Nowadays the use of desktop pc and of mobile devices at the most different categories of users was allowed thanks to the presence of applications that support sophisticated interfaces often characterized by an ever increasing complexity in terms of functionality, look and feel and behaviour. Since much of the judgment of a user about an application is determined by the quality of its interface, it follows that the design of the same can be considered one of the most important steps in the development of an entire software system. The GUI

allows the user to interact directly with the device by manipulating graphical objects. In other words, it is the level of a software application that takes care of the dialogue between the user and the system using a graphical environment. In this perspective, the usability of a software system and especially of its user interface now has become crucial and decisive for the achievement of the applications and of the devices where they are executed.

It is possible to measure how a GUI is “usable” by several evaluation methods and tools of the same. The most common methodologies for interactive graphical interfaces evaluation can be classified in the following categories: analytical evaluation; evaluation by observation; evaluation by interviews; experimental evaluation; automated evaluation.

The first four methods require interaction from end-users or system specialists, and involve manual and repeated steps to refine the assessment. These approaches detect usability deficiencies of the graphic environment developed by running test cases and verification of results and/or by completing questionnaires. In addition to being very expensive and laborious, they often produce results that are considerably dependent on measures of acquisition, on the precise definition of usability, on the type and number of tasks, on the data processed and on evaluation standards.

On the contrary, the automatic evaluation techniques are designed to avoid these problems both in terms of cost (zero in the case of a free application) both in terms of running time: an automatic tool is able to locate in minutes (if not seconds) many critical issues; to get the same results with heuristic methods would take many hours of interviews and simulations of use cases.

In addition, comparative testing carried out in the field shows that automated tools for analytical evaluation are very efficient in terms of execution time, objectivity and reliability of the results obtained [6].

In this paper we present the results of the GUI usability evaluation performed by a fully automated tool that uses an alternative approach for the computation of some metrics of usability of the interface, independent of the source code and based only on the analysis of visual components extracted using image processing algorithms. Several static properties of a GUI are evaluated by the tool (named USherlock), which interacts with each interface element to determine its nature according to the changes produced on the interface itself. Depending on the type of interaction (input/output; pause; double click; click and double click; insertion of a character) and in case a visual feedback is found, the tool identifies some typologies of “dynamic” elements: button, link, text area, etc..., and adds a new node to a tree which represents the hierarchical structure of each frame of the interface. For each element or set of elements classified the evaluation process runs all the usability controls. At the end of the evaluation process, each node is assigned a list of the inconsistencies identified and a score (rating between 1 and 10), which indicates the “quality” of the node.

Using USherlock we have examined some web interfaces. Several usability issues were detected, such as: quite poor visual contents, colors of the elements that have an excessive contrast compared to the background color, lack of feedback on some elements, frames too full of elements, elements that are poorly aligned, opening pop-up windows and non-reversible actions.

The results are compared to the testing made by canonical methods: user questionnaires are developed on a sample of students tester of the Computer Science course of University of Salerno. The outcoming of the experiment shows that similar results are obtained, motivating the development of an automatic tool like USherlock to reduce the costs and times of classical tests.

The rest of the paper is organized as follows. Section 2 presents any related work. The section 3 describes a brief overview of the architecture of the implemented tool. Section 4 presents the results of the evaluation performed on several pages of web sites and the comparison with the testing carried out by evaluation questionnaires developed from user tester. Section 5 contains some conclusions and further researches.

2. Related works

The judgment of a user about an application is determined by the quality of its interface, than the design of the related interface can be considered one of the most important steps in the development of an entire software system.

It is possible to measure how a GUI is “usable” by several evaluation methods and tools. [1] describes a formal technique that explores those features of a specific design that fail to satisfy a set of properties.

In [3] we have presented a methodology to specify and evaluate interactive visual environments, in particular web interfaces, based on the SR-Action Grammars formalism and we present a bottom – up approach to aid the designer to develop graphical applications that automatically respect a significant number of usability rules, such as consistency, completeness and user control, before the software is released and tested by standard methods.

Most of the existing automated tools use a back-side approach, which derives from the analysis of the “hidden” side (source code, logs, etc ...) of the system: if the source code is not available, this kind of evaluation cannot be performed. On the other hand, reverse engineering techniques are implemented in order to extract static properties of the graphical interfaces, semi automatic tool are delegate, after, to analyze the properties of the several GUI components [8]. To evaluate the efficiency of the interactive aspects, most of the systems extract test cases of the typical activities and tester users are assigned to the evaluation. None of the tools and techniques we have examined provides a completely automatic and integrated platform to fully extract and evaluate GUI components and interaction mechanisms.

GUITAR [4] and DART [7] are tools that extract the GUI structure of an application using GUI Ripper, automatically generate test cases for the application based on the extracted information, create expected output (oracle) for the test cases, execute the test case on the application and determine if the tests ran successfully.

The tools described in [5] evaluate visual and textual properties of user interfaces. The set of the presented systems provide graphical analysis tools such as a dialog box summary table that shows an overview of visual properties of all dialog boxes. SHERLOCK provides terminology analysis tools including an Interface Concordance, an Interface Spellchecker, and Terminology Baskets

to check for inconsistent use of familiar groups of terms. Button analysis tools include a Button Concordance and a Button Layout Table to detect variant capitalization, distinct typefaces, distinct colours, variant button sizes and inconsistent button placements.

In [2] we introduced a new front-end process to automatically extract a model of the GUI by dynamically “traversing” all its windows and extracting all the widgets, properties, and values. The usefulness of this process is demonstrated by recovering a structural model called a GUI forest and dynamic models called event-flow graphs and integration trees. Results of case studies show that GUI ripping is effective and requires very little human intervention.

In this paper we present the results of the tests performed by a fully automated system that uses this new approach for the evaluation of some metrics of usability of GUIs, independent of the source code and based only on the analysis of visual components identified and classified using image processing algorithms. The next section presents a brief overview of the implemented tool.

3. The proposed approach

Most of the presented approaches of evaluation require long and expensive testing phases. For this reason, in recent years there has been a rapid spread of automatic tools of evaluation, which allow achieving more objectives in a simple, rapid and economical manner. In this perspective, we refer to an approach to evaluate the usability of a GUI analyzing only its graphical output.

Our front-end approach for GUI evaluation is based on the automatic extraction of a structure of the interface starting from what the user sees on the screen. We use an automatic mechanism to analyze the interactive points of the GUI and then we perform an automatic evaluation of the static and dynamic aspects of the interface. The system is fully integrated and does not require any manual intervention either during the identification of graphical components or during the evaluation process.

As a matter of fact, the static properties of a visual interface, such as aspect ratio, widget nesting, relationship between background and widgets, widgets density, widgets deployment, margins between widgets, widgets alignment, color widgets, background color, easily recognizable edges in widget, widgets shape and size, clear and recognizable buttons, are evaluated using an image processing algorithm to analyze the graphical aspect of the interface. In particular, the evaluation process runs on each frame of the GUI to classify them according to the checkpoints listed above. The nature and behavior of each interface element is determined according to the changes produced on the interface by interacting with it.

At the end of the evaluation process, each frame is assigned a list of the inconsistencies identified and a score (rating between 1 and 10), indicating its “quality”.

In this perspective, for each element is necessary to know:

- RGB component of each point
- position and size

Using USherlock tool to compare GUI usability evaluation results with canonical testing methods

- classification (what kind of widget is: button, label, text area, etc ...);
- input events "heard" and output generated.

To effectively perform the analysis of the previous properties of an interface, it is necessary to:

- Capture all the frames of the interface
- Identify all the relevant elements
- Create a hierarchy of elements in an appropriate data structure
- Classify the elements according to their interactive features
- Evaluate each of the properties listed above
- Presentation a report with the obtained results.

The results of evaluation can be shown to the user in a simplified or detailed manner. The first form of presentation is an aggregated report that shows a numerical score between 5 (maximum score) and 1 (minimum score) for each element. For the root node, the score is computed using a recursive relationship that links the inconsistencies of the node to the score of its children-nodes. The second form of presentation is a list of the inconsistencies identified for each node of the hierarchy. For the i-th element, the list presents the following information: the upper-left pixel coordinates, size (width and height), surface, medium color, dynamic nature, the list of listened events, inconsistencies, evaluation degree.

Based on the described approach we have implemented the USherlock tool. The implementation details of the system are presented in [2].

In the following section we present the results of the usability evaluation of some web sites and the comparison with the canonical test developed by evaluation questionnaires submitted to the evaluation of a group of testers.

4. Experimental results

To demonstrate the usefulness of the implemented approach we have analyzed several interfaces of four typologies of web sites:

1. a popular search engine;
2. an institutional site;
3. a site dedicated to the issues of usability and accessibility of websites;
4. a business site of information technology services.

For each site we have analyzed eight web pages.

For the same set of pages, we conducted a canonical test performed by a group of tester users (students of the Computer Science course of the University of Salerno). In particular, for each site, an evaluation questionnaire was proposed based on the outcome of a set of predefined tasks. Table 1 shows the user questionnaire common to the analysis of all web pages.

USABILITY PROPERTIES	QUESTIONS	VALUE
Widgets density	How would you rate the amount of elements within the page? (1: too much or too little, 2: confused, 3: acceptable, 4: reasonably good, 5: good)?	

Widgets deployment	How would you rate the distribution of elements on the page (1: out of proportion - too many elements above or below, 2: good - elements equally distributed)?	
Widgets color	What do you think of the color of the elements on the page (1: too strong relative to the background; 2: almost unrecognizable from the background; 3: insufficient; 4: sufficient, 5: good)?	
Visibility of the system status	Do you think it is clear the outcome of the execution of an action on the page (e.g. carrying out research) (1: very poor, 2: poor, 3: just enough, 4: sufficient, 5: good)?	
Background color	How would you rate the background color of the page? (1: confused, 2: hardly recognizable compared to the other elements of the interface. 3: insufficient, 4: sufficient. 5: good)	
Aspect ratio	The size of the windows are well-proportioned? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Widget nesting	Do you think that the elements, text boxes and labels on the page are correctly nested? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Relationship between background and widgets	You think that the relationship between the number of elements on the page and the space reserved for them are uniform? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Margins between widgets	The edges of the elements are distinct and equidistant from each other? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Widgets alignment	How do you rate the alignment of elements inserted within another element of the page? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Easily recognizable edges in widgets	The edges of the elements are clearly distinguishable from the background color of the page? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Clear and recognizable buttons	Do you think that buttons are clear and easily recognizable by their color? (1: very little. 2: poor, 3: just enough, 4: sufficient, 5: good)	
Clear and recognizable buttons to mouseover	How do you rate the recognition of the buttons when they are crossed by the mouse arrow? 1: you do not notice at all, 2: you can tell by chance, 3: middle, 4: enough, 5: very relevant)	
Permitted actions and immediate feedback	What is the percentage of non-active links found? (1: very low, 2: low, 3: middle, 4: middle high, 5: high)	

Using USherlock tool to compare GUI usability evaluation results with canonical testing methods

Reversibility	In case of wrong action, it was possible to return to the previous state in a simple and effective manner? (1: absolutely no, 2: no, 3: enough, 4: yes, 5: absolutely)	
Enter Text	When inserting text inside the appropriate sections, the contrast between text and background is evident? (1: very little, 2: poor, 3: just enough, 4: sufficient, 5: good)	

TABLE. 1 - User questionnaire.

For short, we illustrate only the evaluation results of the popular search engine. Figure 1 show the first page of the web site examined.

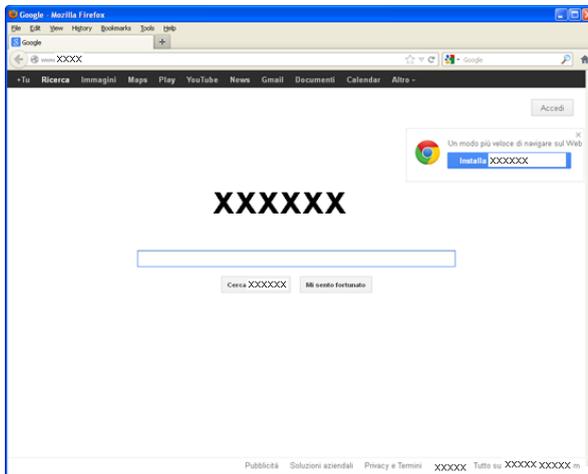


Fig. 1 – A web page of a popular search engine.

The web pages analyzed are:

1. <http://www.google.it/webhp?hl=it&tab=ww>
2. <http://www.google.it/imghp?hl=it&tab=wi>
3. <http://www.google.com/doodles/finder/2013/All%https://accounts.google.com/ServiceLogin?hl=it&continue=http://www.google.it/>
4. <http://translate.google.it/?hl=it&tab=wT>
5. <http://www.google.it/intl/it/about/products/>
6. <https://accounts.google.com/SignUp?service=mail&continue=https%3A%2F%2Fmail.google.com%2Fmail%2F%3Ftab%3Dnm<mpl=default>
7. <https://accounts.google.com/ServiceLogin?service=blogger&passive=1209600&continue=http://www.blogger.com/home&followup=http://www.blogger.com/home<mpl=start#s01>
8. <http://books.google.it/bkshp?hl=it&tab=np>

Automatic evaluation

Overall, the evaluation performed by USherlock has highlighted a high aspect ratio (value: 1.71), a rather simple interface (Widgets density - noWidgetRatio: 2.048E7), several not aligned elements (Widgets alignment = 2) and a certain difficulty to return to the main page (reversibility = 2). Figure 2 shows the summary, for each property, of the average of the values obtained by the tool automatic analysis performed on the pages listed above.

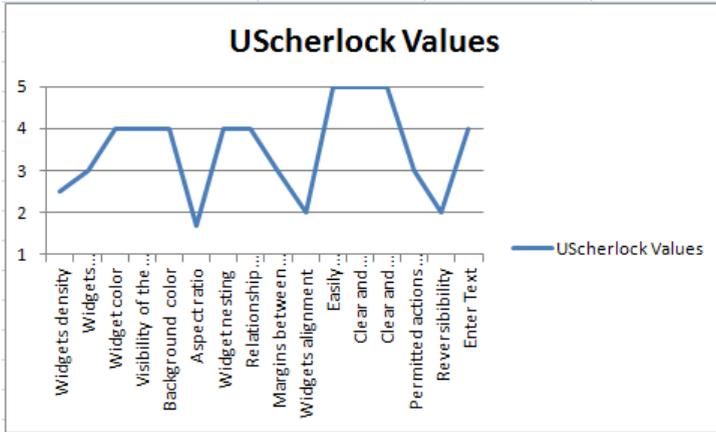


Fig. 2 – Automatic evaluation values of the search engine web site.

Manual evaluation

Similarly, figure 3 shows the evaluation values obtained by taking, for each property, expressed in terms of the average of values in the evaluation questionnaires.

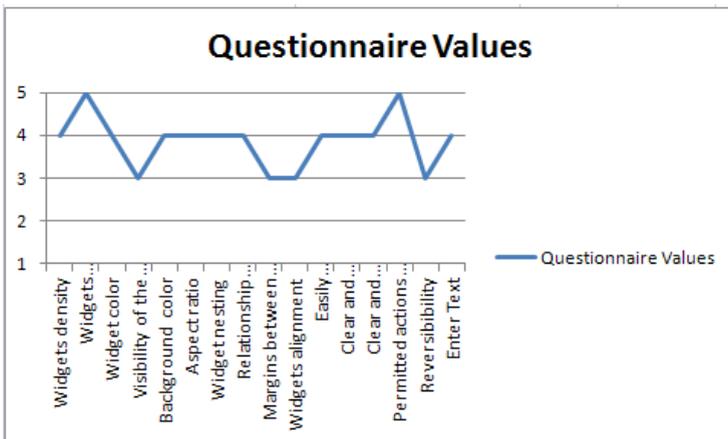


Fig. 3 – Manual evaluation values of the search engine web site.

Using USherlock tool to compare GUI usability evaluation results with canonical testing methods

The results of the evaluation questionnaire of the web page visibly show a substantial similarity between the indications of the USherlock analyzer and the tester users votes. Slight differences appear due to the subjective nature of user evaluation, as the answers to questionnaires are marginally influenced by how the questions are formulated.

Similar results are obtained for the evaluation of home page of the institutional site and of the other seven pages linked:

1. <http://www.istruzione.it/>
2. <http://hubmiur.pubblica.istruzione.it/web/universita/>
3. http://cercauniversita.cineca.it/index.php?module=strutture&page=StruttureSearchParams&advanced_serch=1
4. <http://www.afam.miur.it/>
5. <http://archivio.pubblica.istruzione.it/istanzeonline/index.shtml>
6. <https://www.researchitaly.it/conoscere/>
7. http://archivio.pubblica.istruzione.it/webmail/posta_amministrativi.shtml
8. <http://hubmiur.pubblica.istruzione.it/web/istruzione/famiglie/ordinamenti>

Again, analogous results are obtained for the evaluation of the home page of the site dedicated to the issues of usability and accessibility of websites and of the other six pages linked:

1. <http://www.w3.org/>
2. <http://www.w3.org/standards/xml/>
3. <http://www.w3.org/participate/>
4. <http://open-stand.org/>
5. http://docs.webplatform.org/wiki/Main_Page
6. <http://www.w3.org/community/>
7. https://www.w3.org/community/wp-login.php?redirect_to=/community/groups/propose_cg

Finally, comparable evaluation results are obtained for the evaluation of home page of the business site of information technology services and of the other seven pages linked:

1. <http://latorracainformatica.it/>
2. <http://latorracainformatica.sistemi.net/>
3. <http://latorracainformatica.sistemi.net/pagina.asp?idlv=11&idDoc=3&idpa>
4. http://latorracainformatica.it/area_riservata_11.html
5. http://latorracainformatica.it/dove_siamo_10.html
6. http://latorracainformatica.it/informazioni_12.html
7. http://wsb.register.it/foman/form_mailing_action.jsp
8. http://latorracainformatica.it/site_map.html

5. Conclusions

In this paper we have presented the evaluation results of the usability analysis of several graphical user interfaces performed in automatic manner by the USherlock tool. In the implemented system the static properties of a GUI, such as aspect ratio, widget nesting, relationship between background and

widgets, widgets density, widgets deployment, margins between widgets, widgets alignment, colour widgets, background colour, easily recognizable edges in widget, widgets shape and size, clear and recognizable buttons, are evaluated using an image processing algorithm to analyze the graphical aspect of the interface.

We have shown that similar values are obtained by manual tests conducted by the use of evaluation questionnaire submitted to a group of tester users. The use of an automatic system allows to save a considerable amount of time and evaluation costs at least for those usability metrics that are quantitatively measurable. Subjective aspects of usability as the sense of satisfaction are characteristics that cannot be measured by any automated method. However, the metrics evaluated using the proposed methodology can be a useful complement to standard techniques of evaluation.

Further developments will focus on to measure the efficiency in terms of execution times for the various types of graphical user interfaces examined and on the improvement of the system through the use of OCR systems and inference engines.

Bibliography

[1] José C. Campos, Michael D. Harrison “Interaction Engineering Using the IVY Tool”, - EICS '09 Proceedings of the 1st ACM SIGCHI symposium on Engineering interactive computing systems.

[2] R. Cassino, M. Tucci: “Automatic Usability Evaluation of GUI: A Front-side Approach Using No Source Code Information”, Organizational Change and Information Systems, Lecture Notes in Information Systems and Organization Volume 2, 2013, pp 439-447

[3] R. Cassino, M. Tucci: “Developing usable web interfaces with the aid of automatic verification of their formal specification”, Journal of Visual Languages & Computing, Volume 22, Issue 2, April 2011, Pages 140–149.

[4] Hackner DR, Memon AM, “Test case generator for GUITAR”, Companion of the 30th International Conference on Software Engineering:959-960, 2008.

[5] Rohit Mahajan and Ben Shneiderman, “Visual and textual consistency checking tools for graphical user interfaces”, IEEE Transactions on Software Engineering, Vol. 23 Issue 11, Page(s): 722 – 735, 1997.

[6] Melody Y., Ivory And Marti A. Hearst, “The State of the Art in Automating Usability Evaluation of User Interfaces”, Journal ACM Computing Surveys (CSUR) Surveys Homepage archive Volume 33 Issue 4, December 2001.

[7] Atif Memon, Adithya Nagarajan and Qing Xie, “Automating regression testing for evolving GUI software”, Journal of Software Maintenance and Evolution: Research and Practice, Volume 17, Issue 1, pages 27–64, January/February 2005.

[8] Atif M. Memon, Bao N. Nguyen, “Advances in Automated Model-Based System Testing of Software Applications with a GUI Front-End”, Advances in Computers Volume 80, 2010, Pages 121–162.

Sicurezza nelle reti: utilizzo di architetture multi-core per il monitoraggio del traffico IP

Marco Mezzalama, Gianluca Oglietti¹, Enrico Venuto²

Politecnico di Torino – DAUIN

Corso Duca degli Abruzzi, 24 - 10129 Torino (TO)

marco.mezzalama@polito.it

¹Politecnico di Torino

Corso Duca degli Abruzzi, 24 - 10129 Torino (TO)

gianluca.oglietti@polito.it

²Politecnico di Torino

Corso Duca degli Abruzzi, 24 - 10129 Torino (TO)

enrico.venuto@polito.it

Abstract. *The IP packet capture activity has always assumed great importance in the computer networks security. It's daily used in fact to monitor and analyze the IP traffic passing through a computer networks with the purpose to identify anomalous behaviors that could be associated with security problems. The new infrastructures for high throughput networks, also used in small or medium sized local networks, have made this activity more and more difficult showing some limits of the most recent multi-core capture systems used today. This paper has the purpose to describe the main technologies used in a generic capture system, to identify its possible limits, to diagnose its causes and to discover the possible solutions that must be adopted.*

Keywords: security, networks, multi-core, packet capture, high throughput.

1. Introduzione

L'attività di acquisizione dei pacchetti IP ha assunto una sempre maggiore importanza nella **sicurezza delle reti informatiche** in quanto permette di individuare eventuali comportamenti anomali spesso legati ad infezioni virali, tentativi di hacking, SPAM o truffe informatiche. Tale attività però è divenuta sempre più difficoltosa per la crescente quantità di traffico (throughput) da acquisire. A causa del costante aumento della banda necessaria agli utenti per accedere ai servizi disponibili su Internet si è infatti assistito nell'ultimo decennio ad un sostanziale incremento della capacità dei canali di comunicazione impiegati nelle moderne reti informatiche: nelle reti locali (LAN) si è passati dai 100 Mbit/s dei cavi in rame ai 10 Gbit/s delle fibre ottiche mentre nelle reti metropolitane/geografiche (MAN/WAN) è possibile osservare collegamenti punto-punto in fibra con capacità di 1Tbit/s [1] o, in laboratorio, di 100 Tbit/s [2].

La realizzazione di infrastrutture di rete ad alte capacità anche in reti LAN di piccole o medie dimensioni ha quindi reso sempre più difficoltosa l'attività di acquisizione del traffico IP mostrando alcuni limiti dei sistemi finora utilizzati. Effettuando alcuni test è infatti possibile riscontrare un notevole **aumento del numero di pacchetti persi** già in presenza di throughput dell'ordine di 1Gbit/s utilizzando un calcolatore multi-core e una scheda di rete di ultima generazione.

Nei paragrafi seguenti questo problema verrà analizzato facendo riferimento ad architetture hardware Intel ma i risultati ottenuti possono essere estesi anche ad altre famiglie di CPU.

2. Background

I calcolatori e le schede di rete di ultima generazione presentano al loro interno molte tecnologie che è necessario conoscere per realizzare un sistema di acquisizione per il monitoraggio del traffico IP in reti ad elevato throughput.

2.1. Il modello architetturale NUMA

I calcolatori di ultima generazione sono caratterizzati da architetture hardware progettate per impiegare sulla stessa scheda madre **un elevato numero di unità di calcolo** (o core) a volte distribuite anche su più processori. Una delle architetture multi processore oggi più utilizzate risulta essere l'**architettura NUMA** (Non Uniform Memory Access).

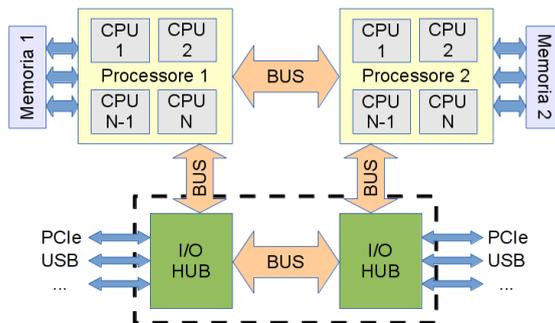


Fig. 1 - Schema a blocchi di un'architettura NUMA generica

La caratteristica principale dell'architettura NUMA sta nel fatto che ad ogni processore multi-core viene direttamente connessa solo una parte dell'intera memoria di sistema. L'insieme formato dal processore multi-core e dalla memoria direttamente connessa ad esso viene chiamato **nodo NUMA**. La possibilità di poter disporre sullo stesso sistema di due o più nodi NUMA permette di incrementare notevolmente le prestazioni del calcolatore in quanto ogni processore multi-core può accedere in parallelo alla porzione di memoria ad esso collegata in modo completamente indipendente. Un altro sostanziale incremento delle prestazioni viene dal fatto che utilizzando questa architettura è possibile ridurre al minimo sia il numero dei fallimenti di accesso alla memoria cache del processore (cache misses) che il numero totale di accessi alla memoria di sistema. Questi ultimi benefici però, in generale, possono essere

ottenuti solamente schedulando i processi sui core dei nodi in cui si trovano anche i dati a cui devono accedere. La schedulazione ottimale dei processi sui core di un sistema NUMA, nella realtà, risulta essere molto più complessa [3] in quanto dipende fortemente da come i processi utilizzano la memoria e dal numero di processi in esecuzione contemporaneamente sullo stesso nodo.

La Fig. 1 mostra lo schema a blocchi generico di un'architettura NUMA. Come è possibile osservare essa è composta da due nodi interconnessi alle periferiche di sistema tramite due **I/O Hub**; ogni nodo risulta poi essere collegato all'altro e ad uno solo dei due I/O hub tramite un **bus dedicato** (chiamato nei sistemi Intel Quick-Path Interconnect o QPI). Un'architettura di questo tipo quindi **non è simmetrica**: senza modificare le interconnessioni fra periferiche e I/O hub le prestazioni possono variare in base al nodo scelto per eseguire il codice dei programmi. Le prestazioni migliori si ottengono solitamente schedulando i processi che devono accedere ad una particolare periferica sullo stesso nodo su cui quella particolare periferica risulta essere interconnessa e, ovviamente, utilizzando un numero di processi non superiore al numero di core presenti all'interno del nodo utilizzato.

2.2 Schede di rete

All'interno delle schede di rete, negli ultimi anni, sono state introdotte alcune tecnologie che via via hanno permesso di accedere ai sempre più veloci mezzi trasmissivi. Le prime di queste tecnologie avevano come obiettivo la diminuzione del numero di operazioni a carico della CPU di sistema (tecnologie di tipo Traffic Offload Engine [4]) e la diminuzione del numero di interrupt ad essa inviati (tecnologia Interrupt Moderation [5]). All'aumentare del throughput però queste nuove tecnologie, da sole, si sono rivelate insufficienti.

Nelle schede di rete di ultima generazione, come quelle basate ad esempio sul controller Intel 82599 [6], sono state quindi implementate altre tecnologie che permettono di ottimizzare le operazioni di lettura e scrittura dei dati in memoria (tecnologia Direct Cache Access [7]) e, tramite l'introduzione di più code in ricezione/trasmisione, di utilizzare tutti i core a disposizione nei recenti calcolatori: **Extended Message Signaled Interrupt** (o MSI-X che permette di assegnare in modo univoco l'interrupt generato da una coda sempre ad uno stesso core utilizzando la tecnica SMP affinity [8]) e **Receive Side Scaling** (o RSS che fornisce al controller della scheda un algoritmo che, in hardware, è in grado di suddividere tutti i pacchetti acquisiti fra le code abilitate). Queste ultime due tecnologie risultano particolarmente importanti ai fini di questo contributo.

3. Realizzazione di un sistema di acquisizione

Nei paragrafi seguenti sono state raccolte alcune informazioni relative al sistema utilizzato nelle prove di acquisizione del traffico di rete.

3.1 Software

Sul server è stato installato il sistema operativo GNU/Linux **Debian** 6.0.6 (64 bit) con kernel aggiornato alla versione 3.2.22. La scheda di rete da 10Gbit/s è

stata configurata usando il driver **ixgbe** aggiornato alla versione 3.7.17-NAPI.

3.2 Hardware

Il sistema impiegato durante i test di acquisizione del traffico è un server basato sulla scheda madre X8DTU-6TF+ [9] la cui struttura interna (northbridge) è mostrata in Fig. 2. Come è possibile osservare si tratta di un'architettura NUMA composta da due nodi collegati ad un I/O hub (IOH 36D – Intel 5520 Chipset [10]) tramite bus QPI. La scheda madre mette a disposizione due bus PCIe versione 2.0 (su slot x16 e x4) e una scheda di rete integrata da 10 Gbit/s basata sul controller Intel 82599EB. Ogni nodo è costituito da 6 GB di RAM e da un processore Xeon X5560 @ 2.80GHz (8 core totali: 4+4HT).

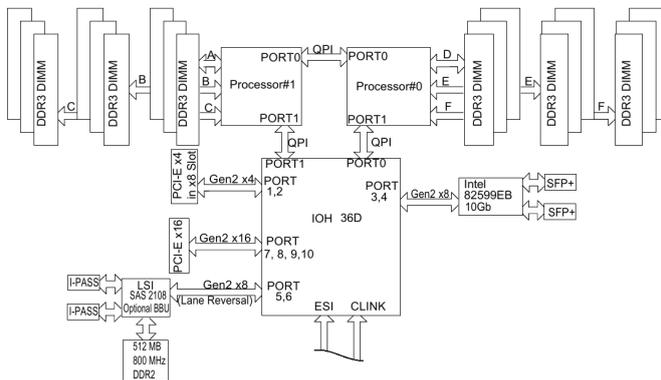


Fig. 2 – Parte della struttura interna della scheda madre utilizzata

3.3 Configurazione di sistema

Per configurare il driver **ixgbe** è necessario estrarre alcune informazioni dal sistema. Tramite il comando **numactl** è possibile ricavare le informazioni sull'architettura NUMA: numero di nodi presenti, suddivisione dei core fra i vari nodi e quantitativo di memoria installato. Fatto ciò è possibile identificare le interfacce di rete, estraendo i **PCI_ID** con il comando **lspci**, e verificare a quale nodo risultino essere interconnesse andando a controllare i relativi file **/sys/bus/pci/devices/PCI_ID/local_cpulist**. Le informazioni ottenute indicano che sul sistema multi-core sono presenti due interfacce di rete 10 Gbit/s direttamente interconnesse al nodo 0 (core 0,1,2,3,8,9,10 e 11).

3.4 Tipologia delle prove

Le prove sono state eseguite inviando al sistema di acquisizione un blocco di 25.000.000 di pacchetti IP a **throughput differenti**: 1, 2, 3, 4, 5, 6, 7, 8, 9 e 10 Gbit/s. Per ognuno di questi throughput sono stati inviati sette **flussi di pacchetti di dimensioni differenti** come consigliato dalla RFC 2544 [11]: 64, 128, 256, 512, 1024, 1280 e 1518 byte. Il software di acquisizione aveva il solo compito di contare i pacchetti ricevuti per ogni differente flusso. Tutte le prove sono state effettuate tre volte e la Tab. 1 contiene la media di queste ripetizioni.

4 Descrizione delle prove effettuate

Per analizzare il comportamento di un'architettura multi processore di tipo NUMA durante l'acquisizione di traffico IP in reti LAN ad elevato throughput sono state effettuate 4 prove differenti descritte nei seguenti paragrafi.

4.1 Prima prova

L'obiettivo di questa prova era di acquisire il traffico IP configurando il driver ixgbe in modo da **utilizzare al meglio tutte le tecnologie a disposizione**. Per questo motivo su ogni interfaccia sono state inizializzate 8 code di ricezione (una per ogni core disponibile in un nodo) mentre il driver è stato configurato in modo da utilizzare la memoria messa a disposizione dal nodo 0, nodo al quale la scheda risulta essere interconnessa. Le interfacce sono identificate da Linux con i nomi eth2 ed eth3 ma in tutte le prove è stata utilizzata la sola interfaccia eth2. Nel seguito l'elenco di tutte le operazioni eseguite:

1) Preconfigurazione del sistema: prima di modificare l'assegnazione degli interrupt (SMP affinity) è necessario disattivare il demone irqbalance (ha il compito di ridistribuire dinamicamente gli interrupt generati dalle periferiche fra tutti i core disponibili con l'obiettivo di ottimizzare le prestazioni). È necessario inoltre rimuovere il driver ixgbe dalla memoria prima di poterlo riconfigurare.

```
~# killall irqbalance  
~# rmmod ixgbe
```

2) Caricamento del modulo ixgbe in memoria: è necessario configurare il modulo ixgbe in modo da utilizzare solamente la memoria presente sul nodo 0 (Node=0,0) e inizializzare 8 code in ricezione su ogni interfaccia (RSS=8,8).

```
~# insmod /PATH/ixgbe.ko Node=0,0 RSS=8,8
```

3) Configurazione dell'interfaccia di rete: per acquisire correttamente il traffico da un'interfaccia di rete è consigliabile disabilitare l'auto negoziazione della velocità (impostandola manualmente alla velocità desiderata, in questo caso 10 Gbit/s), disabilitare il controllo di flusso del protocollo Ethernet e massimizzare la dimensione dei buffer circolari di ricezione.

```
~# ethtool -A eth2 autoneg off rx off tx off  
~# ethtool -s eth2 speed 10000  
~# ifconfig eth2 up  
~# ethtool -G eth2 rx 4096
```

4) Assegnazione degli interrupt delle code ai rispettivi core del nodo 0: per assegnare gli interrupt generati dalle code ai core del nodo 0 è sufficiente estrarre il loro identificativo (IRQ_ID) dal file /proc/interrupt e modificare di conseguenza i relativi file /proc/irq/IRQ_ID/smp_affinity.

Il software utilizzato per contare il numero di pacchetti acquisiti è tshark:

```
~# tshark -qi eth2
```

4.2 Seconda prova

L'obiettivo di questa prova era di acquisire il traffico IP configurando il driver

ixgbe in modo **non ottimale**, andando cioè ad utilizzare il nodo 1 su cui non risulta essere direttamente connessa la scheda di rete. Per poter effettuare un confronto con i dati ottenuti nel caso precedente sono state quindi nuovamente inizializzate 8 code di ricezione. Nel seguito l'elenco delle operazioni eseguite:

1) Preconfigurazione del sistema (come prima prova).

2) Caricamento del modulo ixgbe in memoria:

```
~# insmod /PATH/ixgbe.ko Node=1,1 RSS=8,8
```

3) Configurazione dell'interfaccia di rete (come prima prova).

4) Assegnazione degli interrupt delle code ai rispettivi core del nodo 1.

Per contare i pacchetti acquisiti è stato utilizzato ancora il software tshark.

4.3 Terza prova

L'obiettivo di questa prova era quello di verificare cosa accade utilizzando **una singola coda** all'interno dell'interfaccia di rete (RSS=0,0). L'unico interrupt generato è stato assegnato ad un core del nodo 0 (core 2) e, per **ottimizzare l'uso della cache**, l'esecuzione del programma di acquisizione è stata schedulata sullo stesso core. Nel seguito l'elenco delle operazioni eseguite:

1) Preconfigurazione del sistema (come prima prova).

2) Caricamento del modulo ixgbe in memoria:

```
~# insmod /PATH/ixgbe.ko Node=0,0 RSS=0,0
```

3) Configurazione dell'interfaccia di rete (come prima prova).

4) Assegnazione dell'interrupt dell'unica coda al core 2 del nodo 0.

Per contare i pacchetti acquisiti è stato utilizzato ancora il software tshark. Per garantire però che lo scheduler esegua il comando tshark sul core 2 del nodo 0 è necessario utilizzare anche il comando taskset:

```
~# taskset -pc 2 PID_TSHARK
```

```
~# taskset -pc 2 PID_TSHARK_CHILD
```

Dove PID_TSHARK è il pid del processo tshark mentre PID_TSHARK_CHILD è il pid del suo processo figlio (dumpcap).

4.4 Quarta prova

L'obiettivo di questa prova era quello di verificare cosa accade quando **non viene garantita l'ottimizzazione della cache** del processore. In questo caso il programma di acquisizione del traffico IP è stato schedulato su un core del nodo 0 differente (core 3) rispetto a quello utilizzato per la gestione degli interrupt generati dall'unica coda abilitata sull'interfaccia (RSS=0,0). Nel seguito l'elenco delle operazioni eseguite:

1) Preconfigurazione del sistema (come prima prova).

2) Caricamento del modulo ixgbe in memoria:

```
~# insmod /PATH/ixgbe.ko Node=0,0 RSS=0,0
```

3) Configurazione dell'interfaccia di rete (come prima prova).

4) Assegnazione dell'interrupt dell'unica coda al core 2 del nodo 0.

Per contare i pacchetti sono stati utilizzati i comandi tshark e taskset:

```
~# taskset -pc 3 PID_TSHARK
~# taskset -pc 3 PID_TSHARK_CHILD
```

5 Discussione dei risultati ottenuti

I dati ottenuti in queste prove, come è possibile osservare in Tab. 1, sono abbastanza sconcertanti.

Pacchetti Inviati (byte @ Gbit/s)	% Pacchetti acquisiti nella prova				Pacchetti Inviati (byte @ Gbit/s)	% Pacchetti acquisiti nella prova			
	1	2	3	4		1	2	3	4
64 @ 1	15,1	15,0	0,3	58,2	64 @ 6	0,4	0,6	0	10,2
128 @ 1	12,0	11,9	11,9	76,5	128 @ 6	0,5	0,6	0	14,7
256 @ 1	14,5	12,3	11,8	65,5	256 @ 6	9,9	9,7	0	22,3
512 @ 1	22,5	22,7	11,8	54,4	512 @ 6	14,8	14,6	0,7	22,2
1024 @ 1	29,6	29,2	17,7	40,2	1024 @ 6	12,8	12,7	13,6	15,7
1280 @ 1	33,4	32,9	22,2	48,9	1280 @ 6	12,2	12,5	13,8	14,5
1518 @ 1	37,9	37,4	21,9	48,6	1518 @ 6	12,0	12,0	13,7	13,8
64 @ 2	11,7	11,6	0	29,9	64 @ 7	0,3	0,6	0	8,7
128 @ 2	25,1	24,9	0	44,0	128 @ 7	0,3	0,6	0	13,1
256 @ 2	12,9	12,8	11,6	49,5	256 @ 7	6,8	6,5	0	19,4
512 @ 2	22,4	21,7	11,8	39,6	512 @ 7	13,6	13,3	0	20,6
1024 @ 2	29,5	29,0	17,3	34,9	1024 @ 7	11,6	11,4	11,8	14,6
1280 @ 2	32,1	32,1	22,2	34,5	1280 @ 7	10,9	10,9	12,7	13,2
1518 @ 2	31,4	31,8	30,2	34,0	1518 @ 7	10,6	10,6	12,0	12,6
64 @ 3	2,2	1,4	0	19,6	64 @ 8	0,3	0,4	0	7,6
128 @ 3	13,9	13,9	0	30,3	128 @ 8	0,3	0,4	0	11,4
256 @ 3	25,7	26,1	1,4	39,2	256 @ 8	4,8	4,6	0	16,9
512 @ 3	22,5	22,5	11,8	31,5	512 @ 8	12,8	12,0	0	19,3
1024 @ 3	22,1	22,1	17,2	24,9	1024 @ 8	10,2	10,2	8,4	13,5
1280 @ 3	22,0	22,0	22,1	24,3	1280 @ 8	9,9	9,9	11,3	12,2
1518 @ 3	21,9	21,9	23,0	23,0	1518 @ 8	9,4	9,4	11,1	11,4
64 @ 4	0,9	0,9	0	15,2	64 @ 9	0,3	0,4	0	6,7
128 @ 4	7,4	7,0	0	23,2	128 @ 9	0,3	0,5	0	10,3
256 @ 4	19,3	18,4	0	32,3	256 @ 9	2,9	2,7	0	14,8
512 @ 4	19,3	19,1	8,4	26,6	512 @ 9	10,7	11,1	0	18,3
1024 @ 4	17,5	17,5	17,2	20,4	1024 @ 9	9,4	9,4	5,9	12,2
1280 @ 4	17,0	17,1	18,2	19,9	1280 @ 9	9,0	9,0	10,1	11,5
1518 @ 4	17,0	17,0	17,9	19,2	1518 @ 9	8,6	8,6	9,9	10,7
64 @ 5	0,7	0,7	0	12,1	64 @ 10	0,2	0,2	0	6,1
128 @ 5	3,3	2,7	0	18,2	128 @ 10	0,2	0,2	0	9,3
256 @ 5	13,4	12,5	0	27,0	256 @ 10	1,1	0,3	0	13,6
512 @ 5	16,8	16,4	3,9	24,5	512 @ 10	9,3	8,5	0	17,7
1024 @ 5	14,6	14,5	16,1	18,2	1024 @ 10	8,8	8,6	3,7	12,0
1280 @ 5	14,4	14,4	15,7	17,0	1280 @ 10	8,3	8,2	8,1	11,0
1518 @ 5	14,0	14,0	15,5	16,3	1518 @ 10	8,0	8,0	9,3	10,1

Tab. 1 – Risultati ottenuti

Nonostante sia stato utilizzato un server in grado di supportare tutte le tecnologie hardware disponibili sulla scheda di rete impiegata, infatti, non si è mai riusciti ad acquisire più del 76,5% dei pacchetti anche a throughput relativamente bassi (1Gbit/s). In media poi, durante queste prove, il sistema di acquisizione del traffico non è stato mai in grado di acquisire più del 25% dei pacchetti inviati: 12,5% nella prima prova, 12,3% durante la seconda, 7,6% nella terza e 23,1% nella quarta.

5.1 Risultati ottenuti durante la prima prova

Impostare i driver della scheda in modo da inizializzare 8 code in ricezione non si è rivelata una scelta ottimale in quanto **lo stack di rete di Linux** non è in

grado di gestirle correttamente. Come è possibile osservare dai dati riportati nella Tab 1 questo test ha ottenuto i risultati peggiori in presenza di pacchetti di piccole dimensioni (64, 128 e 256 byte) e cioè quando il numero di pacchetti acquisiti dalla scheda nell'unità di tempo era massimo. Il sistema operativo infatti, in queste condizioni, non è stato in grado di acquisire il traffico IP perché lo stack di rete ha dovuto impiegare la quasi totalità delle risorse per riunire i pacchetti provenienti dalle varie code e presentarli al programma a livello utente tramite l'unica interfaccia disponibile (eth2) come è possibile osservare in Fig. 3.

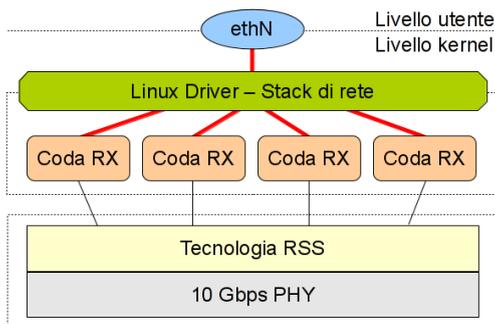


Fig. 3 - Unica interfaccia di uscita dello stack di rete di Linux

5.2 Risultati ottenuti durante la seconda prova

I risultati ottenuti durante la seconda prova permettono di valutare se sia effettivamente necessario configurare il driver ixgbe in modo da utilizzare le risorse del nodo corretto (nodo su cui la scheda di rete risulta essere direttamente interconnessa). Osservando i dati ottenuti si ottiene che in media sono stati acquisiti lo 0,2% di pacchetti in meno rispetto al caso precedente (12,3% contro un comunque sconcertante 12,5% della prima prova). In prima approssimazione sembrerebbe quindi che utilizzare il nodo corretto non sia così importante. È necessario però sottolineare il fatto che, durante queste prove, il processo di elaborazione del traffico a livello utente (tshark) si limitava a contare i pacchetti acquisiti. Durante queste prove, inoltre, il bus di interconnessione QPI risultava essere utilizzato solo marginalmente in quanto sul sistema non erano in esecuzione ulteriori processi che lo potevano saturare. Alla luce di queste considerazioni e visti i risultati ottenuti (degrado dello 0,2%), in un caso reale, è assolutamente necessario configurare il driver ixgbe in modo da **utilizzare correttamente l'architettura NUMA** impiegata.

5.3 Risultati ottenuti durante la terza prova

I risultati ottenuti durante la terza prova sono i **peggiori in assoluto**. Tali risultati sono dovuti al fatto che il programma di acquisizione è stato schedato per essere eseguito sullo stesso core impiegato per la gestione degli interrupt della coda. Se da un lato infatti questa configurazione ottimizzava l'utilizzo della cache del core dall'altro ne **occupava completamente le risorse**. L'unico core utilizzato nel test non aveva infatti risorse di calcolo sufficienti per eseguire sia il

codice del programma a livello utente che tutte le operazioni necessarie per acquisire i pacchetti dalla scheda di rete. Molti dei pacchetti, in questo caso, venivano eliminati direttamente dal Packet Throttling [12] a livello del driver della scheda di rete.

5.4 Risultati ottenuti durante la quarta prova

I risultati ottenuti durante questa prova **sono nettamente i migliori** in quanto, rispetto alla terza prova, sono stati utilizzati due core differenti: uno per la gestione degli interrupt provenienti dalla scheda ed uno per l'esecuzione del programma a livello utente. Questa configurazione, anche se non garantiva di ottimizzare l'utilizzo della cache, permetteva di avere a disposizione una maggiore quantità di risorse di calcolo minimizzando inoltre il numero di context switch su entrambe i core utilizzati. La scelta di utilizzare comunque due core appartenenti allo stesso nodo garantiva, se non l'ottimizzazione della cache di primo livello, di limitare il numero di errori di cache miss totali. Rispetto al primo test poi, la scelta di disabilitare la tecnologia RSS permetteva al sistema di non sprecare risorse per riunire i pacchetti provenienti da diverse code.

Il sistema però, durante questo test, ha acquisito in media solo il 23% dei pacchetti inviati: un risultato comunque **non sufficiente** per essere utilizzato per monitorare la sicurezza in reti ad elevato throughput.

6 Conclusioni

L'attività di acquisizione dei pacchetti IP transitanti all'interno di una rete di calcolatori risulta essere ormai fondamentale in molti campi dell'IT in quanto viene quotidianamente utilizzata per il monitoraggio delle reti informatiche. Tale attività assume una particolare importanza nella **Sicurezza Informatica**. Poter acquisire **tutti i pacchetti IP** in transito in una rete permette infatti di individuare eventuali comportamenti anomali spesso legati a problemi di sicurezza più o meno gravi: avere a disposizione un **sistema di acquisizione perfettamente funzionante** può consentire agli operatori che lavorano in questo campo di individuare molte minacce informatiche quali, ad esempio, infezioni virali, tentativi di hacking, SPAM o tentativi di truffa.

Negli ultimi anni l'attività di monitoraggio del traffico IP in transito all'interno delle reti LAN è diventata però **sempre più complessa** a causa del crescente throughput da acquisire. Fino ad oggi infatti, con throughput dell'ordine di qualche centinaio di Mbit/s, era sufficiente utilizzare un normale calcolatore mono/multi-core per riuscire ad acquisire tutto il traffico in transito su di una rete LAN senza dover conoscere a fondo l'architettura hardware utilizzata. Con throughput che invece si avvicinano ormai ai **10 Gbit/s** risulta essere fondamentale non solo conoscere la struttura hardware impiegata ma anche conoscere a fondo molte delle tecnologie implementate dai produttori sia sulle nuove schede di rete che sulle schede madri assemblate sui server di ultima generazione. Uno degli obiettivi di questo contributo era infatti quello di mostrare come, impiegando un sistema di acquisizione realizzato con un moderno server multi processore basato su di una architettura di tipo NUMA, i

risultati ottenuti potessero variare in modo consistente agendo su alcuni parametri mai presi in considerazione sui sistemi di acquisizione finora utilizzati come il numero di code utilizzate sulla scheda di rete, l'affinity scheda – processore, l'affinity coda – core e l'affinity processo – core.

Alla luce degli sconcertanti risultati ottenuti durante la migliore delle prove effettuate (dove è stato acquisito meno di $\frac{1}{4}$ del traffico inviato alla scheda) è inoltre possibile affermare che se si intende monitorare il traffico in transito su reti LAN ad elevato throughput, ad esempio con l'obiettivo di identificare eventuali problemi di sicurezza informatica, **non è sufficiente utilizzare potenti calcolatori o schede di rete di ultima generazione** ma è assolutamente necessario apportare profondi cambiamenti anche sul software utilizzato: driver della scheda, kernel del sistema operativo e software a livello utente. Gli scarsi risultati ottenuti sono infatti da imputare ai driver standard delle schede di rete, all'interfaccia NAPI e allo stack di rete del kernel che sono stati creati per garantire il normale funzionamento di una interfaccia di rete e **non risultano essere quindi ottimizzati per l'acquisizione dei pacchetti**. Per poter realizzare un sistema d'acquisizione del traffico di rete è quindi assolutamente necessario introdurre un nuovo layer software [13] che si vada ad affiancare, quando necessario, allo stack di rete e all'interfaccia NAPI.

Bibliografia

- [1] "Ericsson and Telstra successfully trial 1Tbps optical link", Ericsson Communications, <http://goo.gl/0azN5>, 2013.
- [2] "100 Tbit/s Free-Space Data Link using Orbital Angular Momentum Mode Division Multiplexing Combined with Wavelength Division Multiplexing", Hao Huang et al, 2013.
- [3] "Memory Management in NUMA Multicore Systems: Trapped between Cache Contention and Interconnect Overhead", Zoltan Majo et al, 2011.
- [4] "TCP offload engine", <http://goo.gl/7Qpl3>.
- [5] "Interrupt Moderation Using Intel GbE Controllers", Intel.
- [6] "Intel 82599 10 GbE Controller Datasheet", Intel.
- [7] "Direct cache access for high bandwidth network I/O", Huggahalli R., Iyer R., Tetrick S., Computer Architecture, 2005.
- [8] "SMP IRQ affinity", Ingo Molnar, Max Krasnyansky, <http://goo.gl/fmoHb>.
- [9] "X8DTU-6TF+", Supermicro, <http://goo.gl/ibWhb>.
- [10] "Intel® 5520 Chipset and Intel 5500 Chipset Datasheet", Intel.
- [11] "Benchmarking Methodology for Network Interconnect Devices", S. Bradner, J. McQuaid, <http://www6.ietf.org/rfc/rfc2544>, 1999.
- [12] "napi", Linux Foundation, <http://goo.gl/H38Fu>, 2009.
- [13] "Improving Passive Packet Capture: Beyond Device Polling"; Luca Deri; 2004.

Cloud Forensic Challenges and Readiness

Lucia De Marco^{1,2}, Filomena Ferrucci¹, M-Tahar Kechadi²

¹Department of Management and Information Technology

University of Salerno, Via Giovanni Paolo II, 132 - 84084 Fisciano (Salerno) – Italy

²School of Computer Science and Informatics, University College Dublin - Ireland

ldemarco@unisa.it, fferrucci@unisa.it, tahar.kechadi@ucd.ie

Abstract. *Cloud Forensics is an emerging topic in the Digital Forensics context. It concerns with the management of digital crimes in the Cloud Computing environment. The specific characteristics of Cloud Computing determine new challenges for practitioners and researchers dealing with Cloud crimes. In this paper, we discuss the impact of Cloud Computing into the Digital Forensics field and the relation between the specific Cloud Computing features with the challenges they provide from the forensics point of view taking into account technical, organizational, and legal dimensions. An approach for conducting proactive investigations and at the same time improving Cloud security is based on the development of a Cloud Forensic Readiness System, meant to collect and store data coming from the Cloud infrastructure to be then exploited for conducting digital crimes investigations more effectively. A reference architecture for such a system is described.*

Keywords: *Cloud Computing, Digital Forensics, Cloud Forensics, Forensic Readiness*

1. Introduction

Technological progresses have been so attractive to positively influence all the aspects of our lives, from personal to business. On the other side, the progress has been attracting for misuses, originating more and more sophisticated digital crimes. From this perspective, the Digital Forensics (DF) discipline [Palmer, 2001] is becoming more and more significant in order to facilitate and solve those crimes; indeed, it manages with scientifically derived and proven methods the evidences coming from digital devices for reconstructing a correct events timeline that is fundamental for facilitating cases resolution.

Since '80s, DF has been progressing at the same speed of ICT. Several tools and procedures have been established [Ambhire and Meshram, 2012],

[Casey, 2002] [Casey, 2011] and adaptations to the technological innovations have been developing. Cloud Forensics (CF) is a new frontier in DF [Ruan et. al, 2011], since the Cloud Computing (CC) technology is the current ICT evolution [Mell and Grance, 2011]. The Cloud has been adopting very rapidly for several purposes [Gartner, 2013] and Public Cloud services spending is expected to record an annual growth rate of 17.7% from 2011 through 2016, i.e., 210 billion dollars (see Fig. 1). Such a success derives from some specific appealing characteristics of CC; nevertheless, these Cloud characteristics represent issues for performing DF investigations. An approach for leveraging some of these emerging issues is based on the implementation of Digital Forensic Readiness (DFR) [Rowlingson, 2004] into the Cloud. It focuses on rendering existing Cloud architectures capable of collecting sensitive and critical information related to digital crimes before they happen, leading to save time and money for the investigations [De Marco et. al, 2013].

The paper is structured as follows: in Section 2 the main challenges characterizing Cloud Forensics are illustrated; in Section 3 Digital Forensic readiness is discussed in the context of the Cloud; finally, conclusions and future work are presented in Section 4.



Fig. 1 - Public Cloud Services Market, 2010-2016 [Gartner, 2013]

2. Cloud Forensics

DF has been following the evolution of the computing architectures [Pollitt, 2010] [Garfinkel, 2010]. Traditional DF involved the collection, the analysis and the reporting of digital evidence coming from single machines that were seized by law enforcements and controlled by them since their arriving on the crime scene until the case resolution. The adopted procedures and tools were capable of performing investigations on single hard drives and external storage devices. With the spreading of computer networks, the forensics practitioners assisted to the evolution of Network Forensics (NF) [Palmer, 2001], which applies the DF science to computers' networks. The devices and the digital evidence to be investigated are located along the network paths, e.g., routers,

access points, switches and server machines, together with the data found into the single machines, either client or server. The giant steps of technological progress are posing stronger and stronger limits to DF. For instance, the smartphone technology is escalating a sub-branch called mobile forensics [Casey, 2011], which has been requiring efforts for adapting existing tools to the architecture of the devices. More recently, we are assisting to the increasing adoption of Cloud Computing services [Gartner, 2013], meant to be a revolutionary event for DF. In fact, forensics people, i.e., government, lawyers, and police, must deal with the issues deriving from the Cloud Computing infrastructures, in order to use the Cloud evidence for court cases respecting their admissibility and reliability principles [Casey, 2011].

Cloud Forensics (CF) is essentially defined as the application of forensic science to the Cloud Computing technology. It is conceived as a subset of NF [Ruan et. al, 2011], but some differences exist between the manners for managing investigations in the cloud and in computer networks. The Cloud is based on broad network access and the devices belonging to the network paths are spread all over the world without precise routing information; hence, the data they store potentially belong to different jurisdictions. Thus, the NF tools and procedures can be used for the Cloud but they must be adjusted for it, because of some Cloud specific characteristics.

The Cloud has not to be intended only through its technical aspects, because it also includes other features that determined a CF structure modeled by three-dimensions, which are technical, organizational, and legal [Ruan et. al, 2011]. The technical dimension deals with tools and procedures for performing forensic investigations; the organizational one concerns with the manner of establishing a forensic capability, e.g., what are the roles and the responsibilities to assign into a Cloud organization; finally, the legal dimension covers issues about multi-jurisdiction, multi-tenancy, and Service Level Agreements (SLAs) policies. The most discussed CF challenges can be structured through the same three-dimensional model with the aim of providing a manner for understanding how to relate these challenges to the specific features from which they derive [Zargari and Benford, 2012] [Reilly et. al, 2011] [Ruan et. al, 2011] [Ruan et. al, 2013] [Birk and Wegener, 2011].

Technical Challenges	CC Features				
	Elasticity	Location	VM	Broad Access	Network
Reduced data access	X	X	X		
Lack of physical control	X	X	X		
Lack of standard	X	X	X		
Multiple log formats	X		X		
No timestamps synchronization		X			X
No routing information		X	X		X

Table 1 – CF Technical Dimension Challenges

Into the technical context several aspects must be discussed (see Table 1). Cloud Services are rapidly elastic, meaning that they are provisioned and

released for responding to scaling demands. Elasticity is guaranteed involving multiple servers, located all over the world, and Virtual Machines (VMs) in order to simulate infinite resources availability [Mell and Grance, 2011]. These features determine a reduced access to data, because the Cloud Services Providers intentionally hide its location to facilitate movements and replications. Furthermore, physical control over the architecture components is lacking; it varies for the 3 Services Models (SaaS, PaaS, IaaS), being bigger under the Providers in SaaS, and bigger under the Customers in IaaS. In the Cloud there is also lack of standards and multiple log files formats; again, among the several data centers and servers there is no timestamps synchronization. Finally, Cloud Services are accessed over the network through heterogeneous client platforms; but, unlike Network Forensics, in the Cloud there is no routing information to derive the digital attacks sources.

CF Organizational Challenges	CC Features				
	Third Services	Party	Multiple Locations	Cross-Providers	SLA
Lack of expertise			X	X	
Legal measures	X		X	X	X

Table 2 – CF Organizational Dimension Challenges

From the organizational perspective other features must be considered (see Table 2). Conducting a Cloud forensic investigation involves data and services belonging to both Providers and Customers. There are cases where the Cloud Services Providers obtain services from third parties; hence the scope of the investigation becomes wider. The lack of legal experience specific for the Cloud features determined uncertainty about the measures to undertake in case of cross-providers or third parties resources supplying; also the SLAs might be tailored for including proper rules.

CF Legal Challenges	CC Features			
	Elasticity	Multiple locations	SLA	Broad Network Access
Multi – tenancy	X		X	
Multiple jurisdiction		X	X	X

Table 3 – CF Legal Dimension Challenges

CC features determined challenges also in the legal domain (see Table 3). The resources can be used by multiple consumers via a multi-tenant model, and they are dynamically assigned according to the demand. The principal issue pertains to manage the multi-tenancies in order to preserve others' privacy. Moreover, the multiple jurisdictions issue is specific for the Cloud, because it derives from the presence of data centers all over the world; and also from the users' capability of accessing documents and services from

everywhere, being no more physically constrained. Several legal principles are taken into account in order to address this matter [ENISA, 2009] [Orton et. al, 2012] [Ward and Sipior, 2010].

3. Cloud Forensic Readiness

An approach for addressing some issues of DF in order to render a system ready for forensic purposes is represented by Digital Forensic Readiness (DFR). It was defined as "the ability of an organization to maximize its potential to collect digital evidence and minimizing the costs of an investigation" [Tan, 2001]. In order to achieve these goals, a DFR capability must be conceived as the development of a system with the principal aim of identifying, collecting, and storing critical data coming from the underlying computing infrastructure. Rowlingson proposed a ten steps process that should allow an organization to implement DFR [Rowlingson, 2004]; it includes the key activities to perform for gathering potential digital evidence complying with admissibility and reliability principles for court cases. An interesting approach for managing DFR is instead discussed in [Reddy and Venter, 2013]; a novel architecture is provided for assisting large organizations who want to realize an optimal level for managing DFR. That architecture is composed of detailed functional requirements, which are determined by a literature survey, and it is also supported by an early proof-of concept prototype system to demonstrate that it is feasible.

Other works stressed the importance of a DFR capability in order to enhance the internal security of an organization. In [Grobler and Louwrens, 2007] the overlap between the DF and Information Security best practices was examined; these best practices exclude requirements for the preservation of digital evidence, necessary for investigations. Indeed, into the organizations adopting the actual best practices, the events and information related to the security controls are not prosecutable. DFR is the solution for implementing the respect of the legal admissibility guidelines on all the processes and evidence. In [Danielsson and Tjostheim, 2004] the importance of the availability of digital evidence is discussed; because they are necessary for investigating on criminal cases, they must be collected in a proper and pro-active manner, in order to render the investigations effective and successful; thus, for this purpose, a DFR capability must be implemented into an organization, and it must follow a structured approach; it includes several features regarding national and international legislation, together with their constraints and requirements about data collection and preservation; it aims to identify the sources of digital evidences and how they must be preserved by tools; finally, it provides guidance for providing a report of the incidents to law enforcements, including the content, the format, the criteria for the report itself and how the interaction between the law enforcements and the affected parties is regulated.

A DFR capability can be added to Cloud Computing [Ruan et. al, 2013] through a dedicated system, which will perform pro-active forensic investigations and will offer some positive side effects. To the best of our knowledge few attempts have been performed in order to provide a forensic data collection capability to the Cloud. In [Dykstra and Sherman, 2012] some

existing forensic tools like EnCase were analyzed in the CF context; the result affirmed that the Cloud data collected by those tools are unreliable, because proper Cloud features require more efforts for performing forensics than simply tailoring the existing tools and procedures; this is due to the fact that new technical requirements must be managed for complying with the legal principles required for the digital evidences. A proper remote forensic acquisition suite of tools in an open-source cloud environment has been proposed in [Dykstra and Sherman, 2013]; this suite, named FROST, provides a forensic capability into the IaaS level of OpenStack, an open-source Cloud Computing platform. FROST performs data collection from the CSPs and from the host operating system, and renders the data available to the users, because it is assumed that the customers are cooperative during the investigations. The data collected in FROST includes virtual machines images, logs coming from the API requests, and the OpenStack firewall logs. This suite is considered by the authors also a way for enhancing forensic readiness into the Cloud.

A reference architecture for a Cloud Forensic Readiness System (CFRS) was proposed in [De Marco et. al, 2013]; the approach undertaken to conceive such architecture was based on an examination of the most common Cloud technical features, in order to understand the potential sources of digital evidences among artifacts and tools. Subsequently, the CFRS principal sub-systems were designed, together with the operations they must perform; finally, the necessary communication channels were introduced.

Such reference architecture was proposed motivated by the necessity of providing a general approach and understanding the essential operations that a CFRS must perform. Its main advantage resides in its design, which is not constrained by any specific and technical configuration; rather, it is flexible and customizable, and it can be considered as a template for all the organizations and Cloud Service Providers who want to implement a DFR capability.

The CFRS is designed over an existing Cloud infrastructure without altering the original components (see Fig. 2). It includes two main sub-systems; the former is the Forensics Data Base, which is dedicated to the collection of the Cloud Services information, suitable for becoming potential evidence, and to their classification depending on the type; the latter is the Readiness Core Module, which performs different activities on the collected data, executed by dedicated subsystems, and described in the following.

Into the Forensics Data Base, the Forensics Artifacts sub-system is dedicated to the storage of VMs Images, Single Sign-On logs, system states and applications logs, running system memory. Cloud Auditors logs and error logs coming from hypervisors are instead collected by the Forensics Log. The Monitored Data sub-system refers to information coming from Cloud facilities dedicated to data monitoring and control [CSA, 2011], e.g., Database and File Activity Monitoring, URL Filtering, Data Loss Prevention, Digital Rights Management System and Content Discovery System.

Into the Readiness Core Module, the previously collected data are encrypted and stored by the dedicated sub-components, i.e., Data Encryption and Data storage, respectively. The Data Management sub-system performs forensic analysis and knowledge extraction with the purpose of reconstructing a correct

and reliable events' timeline; finally, the chain of custody report [NIJ, 2008], necessary for cases resolution, is performed by the Chain of Custody sub-system.

A communication and data exchange channel is necessary between the Cloud and the CFRS, and also among the several sub-systems. For this purpose, the OVF standard language [OVF] is suitable both for its design and distribution. This language is capable of creating and distributing Software applications to be executed on different VMs, independently from hypervisors and from CPUs architectures. Moreover, it can be extended for future VM hypervisors developments.

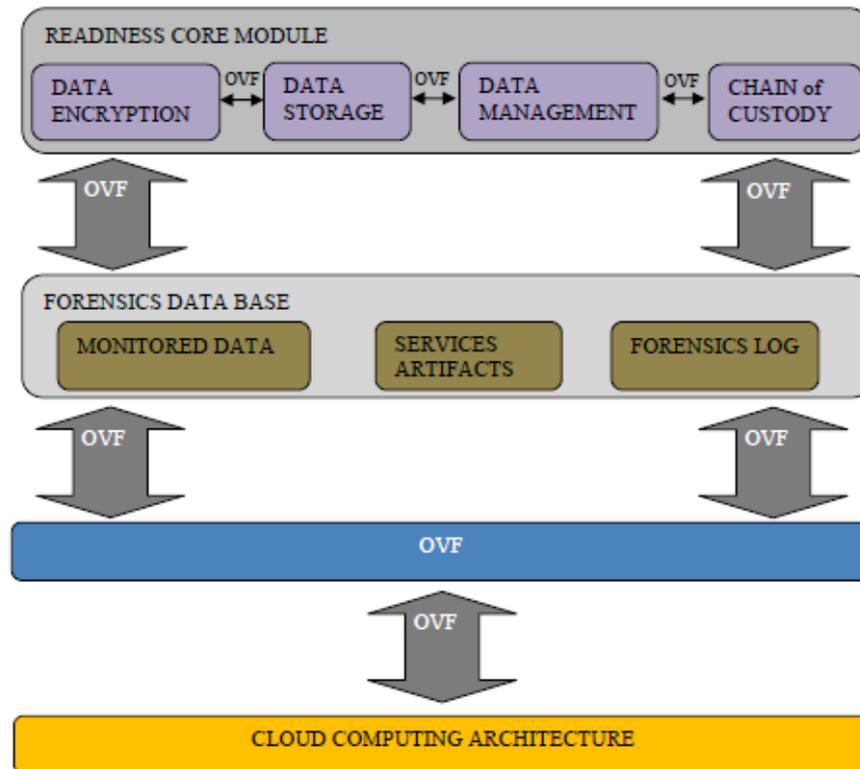


Fig. 2 – CFRS Reference Architecture

The implementation and the usage of such a CFRS are very important for multiple purposes; indeed, a CC environment can exploit it to improve customers' data privacy and its internal security, because a major control and monitoring will be performed for protecting critical information by the forensic readiness system itself [De Marco et. al, 2013]. A Cloud organization or a CSP can be prepared for managing possible incidents, and the manner of gathering enough digital evidence minimizes the effects on the organizations routine operations, which would be considerable in case of a simple re-active

investigation. The CFRS can also furnish a better way for demonstrating the respect of the Service Level Agreements clauses that guarantee the level of the services, because, for instance, the sub-system dedicated to the event reconstruction is able of determining the sources of the attacks or the exact time where a data violation happens. Moreover, a CFRS adoption can be helpful in order to address some CF challenges described in section 2; first of all, from a technical perspective, a manner for aligning the multiple log files formats and for synchronizing the several machines timestamps can be implemented into the Data Management sub-system, otherwise reconstructing a correct and reliable events' timeline cannot be feasible. From the organizational perspective, the CFRS can be considered as a valid approach for assigning the necessary roles and responsibilities for managing the cloud incidents, such as Investigators and Incident Handlers, because the people trained for that can be manager of the Readiness Core Module that performs activities related to cases resolution. Finally, from a legal point of view, the CFRS can highlight the main issues regarding the jurisdictions borders; it can become the instrument for alerting the proper governments' offices, helping to address a more general problem.

4. Conclusions and Future Work

Cloud Forensics is an emerging topic with the goal of performing digital investigations into the Cloud architecture. CF is facing with several challenges that come from the specific CC features, suggesting that adaptations for investigation procedures and tools are necessary for the Cloud. In this paper, an examination of the CF dimensions and open challenges was provided. Moreover, the digital forensic readiness was examined; it can be an approach for providing more security to the users and facilitating the criminal investigations. The illustrated CFRS reference architecture was meant for detecting the potential sources of digital evidence that can be proactively collected from the Cloud and for being subsequently analyzed.

Addressing the identified CF challenges should allow us performing both re-active and pro-active forensics investigations into the Cloud; in fact, our research direction deals with two main topics. With the former, a structured process model for conducting re-active forensics investigations will be provided: it requires a systematic review of the existing investigation process models, in order to understand their common features and phases, and produce a process model for the Cloud; the design of such a model will include a set of phases and possible sub-phases, a set of involved CF actors, and a collection of scenarios for the Cloud crimes, such as the ones mentioned in Section 2. With the second, which concerns pro-active CF, we will design and prototype a CFRS starting from the presented reference architecture; we will investigate in which part of a Cloud architecture a CFRS can be implemented; we will attempt to customize such a reference architecture, probably commencing with an open source cloud platform; we will also understand if the cloud artifacts that we included in our proposal are actually suitable for being manipulated by the OVF data exchange module; and we will finally decide what are the most important

knowledge extraction procedures to adopt for forensics purposes, i.e., data mining techniques. Finally, we will merge the two topics by simulating Cloud attacks into a system that is furnished with a Forensic Readiness capability in order to highlight the challenges and the advantages of the implemented CFRS.

References

- Ambhire, V. R., Meshram, B. B., Digital Forensic Tools. IOSR Journal of Engineering, Vol. 2(3), Mar. 2012, pp. 392-398
- Birk, D., Wegener, C., Technical Issues of Forensic Investigations in Cloud Computing Environments, in IEEE Sixth International Workshop on Systematic Approaches to DF Engineering, Oakland, California, USA IEEE, May 2011, pp. 1–10
- Casey, E., Handbook of Computer Crime Investigation, Academic Press, New York, 2002
- Casey, E., Digital Evidence and Computer Crime, 3rd Edition, Academic Press, New York, 2011
- Choo, K.K.R., Cloud computing: challenges and future directions. Trends & issues in crime and criminal justice no. 400, 2010, pp. 1-6
- Cloud Security Alliance; Security Guidance for Critical Areas of Focus in Cloud Computing v 3.0, 2011 <https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf>
- Danielsson, J., Tjostheim, I., The Need for a Structured Approach to Digital Forensic readiness in Digital forensic readiness and e-commerce (IADIS) 2004, pp. 417 - 421
- De Marco, L., Kechadi, M-T., Ferrucci, F., Cloud Forensic readiness: Foundations, in Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (LNICST) series, Proceedings of the 5th International Conference on DF & Cyber Crime (ICDF2C), Moscow, Russia, September, 2013
- Dykstra, J., Sherman, A.T., Acquiring Forensic Evidence from Infrastructure-as-a-Service Cloud Computing: Exploring and Evaluating Tools, Trust, and Techniques, in Proceedings of the 12th Annual DF Research Conference (DFRWS'12), Washington, DC, USA, Digital Investigation, vol. 9, August 2012, pp. 90–98
- Dykstra, J., Sherman, A.T., Design and Implementation of FROST: Digital Forensic Tools for the OpenStack Cloud Computing Platform, preprint submitted to the 13th Annual DFRWS Conference held in Monterey, CA from August 4-7, 2013
- European Network and Information Security Agency (ENISA), Cloud Computing: Benefits, risks and recommendations for information security, 2009 <http://www.enisa.europa.eu/activities/risk-management/files/deliverables/cloud-computing-risk-assessment>
- Garfinkel, S.L., Digital Forensics research: The next 10 years. Digital Investigation, Volume 7, Supplement, August 2010, pp. S64-S73
- Gartner, Forecast Overview: Public Cloud Services, Worldwide, 2011-2016, 2013 <http://www.forbes.com/sites/louiscolombus/2013/02/19/gartner-predicts-infrastructure-services-will-accelerate-cloud-computing-growth/>
- Grobler, T., Louwrens, B., Digital forensic readiness as a component of information security best practice, in Proceedings of New Approaches for Security, Privacy and Trust

in Complex Environments, IFIP TC-11 22nd International Information Security Conference, vol. 232, 2007, pp. 13-24

Lau, F., Rubin, S.H., Smith, M.H., Trajkovic, L., Distributed denial of service attacks, in IEEE International Conference on Systems, Man, and Cybernetics, Nashville, TN, USA, vol.3, 2000, , pp. 2275-2280

Mell, P., Grance, T., Final Version of NIST Cloud Computing Definition Published, 2011, <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>

National Institute of Justice, Electronic Crime Scene Investigation: A Guide for First Responders, 2008, <https://www.ncjrs.gov/pdffiles1/nij/219941.pdf>

Orton, I., Aaron, A., Endicott-Popovsky, B., Legal Process and Requirements for Cloud Forensic Investigations. Cybercrime and Cloud Forensics: Applications for Investigation Processes, ed. Keyun Ruan, IGI Global, Forthcoming, 2012

Open Virtualization Format: OVF Standard. <http://www.dmtf.org/standards/ovf>

Palmer, G., A Road Map for DF Research, Report from the First DF Research Workshop (DFRWS), Utica, New York, 2001

Pollitt, M., A History of DF, IFIP Advances in Information and Communication Technology Volume 337, 2010, pp. 3-15

Reddy, K., Venter, H.S., The architecture of a digital forensic readiness management system. Computers & Security, Volume 32, February 2013, pp. 73-89

Reilly, D., Wren, C., Berry, T., Cloud Computing: Pros and Cons for Computer Forensics Investigations. International Journal of Multimedia and Image Processing (IJMIP), vol. 1, 2011, pp. 26-34

Rowlingson, R., A ten step process for forensic readiness. International Journal of Digital Evidence, 2 (3), 2004, pp. 1 – 28

Ruan, K., Carthy, J., Kechadi, T., Crosbie, M., Cloud forensics: an overview. Advances in DF 7, vol. 361, 2011, pp. 35–49

Ruan, K., Carthy, J., Kechadi, T., Baggili, I., Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results. Digital Investigation, Volume 10, Issue 1, June 2013, pp.-34-43

Specht, S.M., Lee, R.B., Distributed Denial of Service: Taxonomies of Attacks, Tools and Countermeasures, in Proceedings of the 17th International Conf. Parallel and Distributed Computing Systems (PDCS '04), International Workshop on Security in Parallel and Distributed Systems, September 2004, pp. 543-550

Tan, J., Forensic Readiness, Technical report, @Stake Organization, 2001 http://isis.poly.edu/kulesh/forensics/forensic_readiness.pdf

Ward, B.T., Sipior, J.C., The Internet Jurisdiction Risk of Cloud Computing. Information Systems Management, Volume 27, Issue 4, 2010, pages 334-339

Windows, Windows Rootkit Overview. Symantec, 2006 <http://www.symantec.com/avcenter/reference/windows.rootkit.overview.pdf>

Zargari, S., Benford, D., Cloud Forensics: Concepts, Issues, and Challenges, in Third International Conference on Emerging Intelligent Data and Web Technologies (EIDWT), IEEE, 2012, pp. 236-243

CLOUD: Presente e Futuro

De Angelis Francesco, Polzonetti Alberto¹, Sabbatini Samuele²

Scuola di Scienze e Tecnologie UNICAM
Via del Bastione 1, 62032 Camerino (MC)
francesco.deangelis@unicam.it

¹Scuola di Scienze e Tecnologie UNICAM
Via del Bastione 1, 62032 Camerino (MC)
alberto.polzonetti@unicam.it

²Scuola di Scienze e Tecnologie UNICAM
Via del Bastione 1, 62032 Camerino (MC)
samuele.sabbatini@unicam.it

Cloud is an emerging technology that allows making the increasing demand of advanced technology at a lower cost, in an easier and more effective manner. This article will contain some information regarding this new technology. The emphasis will be both on state of the art and objectives of future research about cloud. In particular, Section 1 provides introductory information about cloud. Section 2 provides a set of common definitions and deepens cloud's features from different perspectives. Section 3 analyzes the usage scenarios for cloud and in particular, what features are considered most important. Section 4 compares the strengths and characteristics of cloud against its current usage. Section 5 explores the major research topics and assesses their complexity and relevance. Section 6 concludes the report with a series of general considerations.

Keywords: cloud, research, future.

1. Introduzione

Il Concetto di Cloud permette un ottimo rapporto costo/prestazione consentendo agli utenti ed ai fornitori un facile accesso alle risorse in modalità self-service riducendo così il costo per l'amministrazione del sistema e migliorando l'utilizzo delle risorse. Negli ultimi anni, il cloud ha generato un impatto importante sull' IT globale, dando origine a nuovi mercati e nuove comunità di utenti. Il concetto di cloud non è, quindi, una novità in sé, in principio nacque da un bisogno industriale diretto a migliorare l'utilizzo delle risorse senza incidere nelle esigenze dei consumatori utilizzando le risorse disponibili in modo più efficiente.

I sistemi cloud sono emersi più da una necessità commerciale che da progetti di ricerca e sviluppo, in questo modo è stato preseguito un percorso evolutivo a breve termine per soddisfare le esigenze dei clienti. Per creare una sostenibilità a lungo termine, è necessario concentrarsi nella ricerca e negli aspetti di sviluppo del cloud computing che sono di rilevanza a lungo

termine, piuttosto che partecipare alle fluttuazioni a breve termine che creano una grande competizione, e possono inoltre essere di scarso impatto sullo sviluppo globale. Il marketing, oltre che una maggiore diversificazione dei servizi, ha creato una grande campagna pubblicitaria intorno al cloud che ha portato ad elevate aspettative da parte dell'utente, che potrebbero non essere realisticamente soddisfatte. Il marketing tende a promettere caratteristiche che sono facilmente confuse con caratteristiche dal significato diverso in altri domini, creando potenzialmente idee sbagliate quali: i cloud (a) sono sempre disponibili ("sempre e ovunque"), (b) offrono infinite risorse con più elasticità e senza soluzione di continuità, (c) hanno prestazioni da super computer a costi minimi, (d) forniscono notevoli benefici ambientali. Anche se il cloud può offrire un notevole sostegno in tutte queste direzioni, esse devono essere limitate dai vincoli di sistema come il mobile computing ed "Internet of Things". Questi limiti non sono però chiari a tutti i clienti, a volte ai fornitori stessi, ed è quindi uno dei principali compiti di ricerca l'identificazione della posizione esplicita del cloud all'interno di questi domini.

Molti utenti si sentono costretti a migrare al cloud in modo da rimanere competitivi e di beneficiare dei miglioramenti presunti offerti da tali infrastrutture. Senza una chiara conoscenza di quali casi d'uso siano particolarmente adatti per ambienti cloud, vi è un elevato rischio che un utente inesperto stimi le capacità e quindi la fruibilità del cloud erroneamente, rimanendo insoddisfatto con le capacità effettive che ottiene dalla nube. Dato che gli utenti ottengono maggiore esperienza nell'utilizzo del Cloud, le loro capacità, i punti di forza e le carenze diventano sempre più evidenti. I fornitori sono quindi sotto crescente pressione per soddisfare queste aspettative, in parte derivanti dalle loro promesse, per soddisfare le esigenze degli utenti, affinché l'interesse per il Cloud non possa calare di nuovo in futuro. Secondo l'analisi Gartner [Fenn e LeHong, 2011], il cloud computing ha raggiunto un picco di aspettative gonfiate nel 2009/2010 ed è ora in fase di crescente disillusione dove gli utenti diventano sempre più consapevoli delle carenze del sistema e iniziano a valutare e considerare la scelta del passare al cloud o meno.

Questo lavoro cerca di individuare gli aspetti chiave di sistemi cloud e la loro rilevanza per il futuro dell'ecosistema IT. Si cerca di individuare le lacune ancora esistenti e le questioni di ricerca da affrontare al fine di correlare le tecnologie esistenti alla logica del cloud. Lungo questa linea, tenta di stabilire una serie di definizioni comuni che identificano chiaramente gli aspetti intrinseci del cloud e li separa dalle altre aree correlate.

2. Definizione di Cloud

In questo momento esistono molte definizioni, simili tra loro, di cloud ma la più usata è quella fornita dal National Institute of Standards and Technology (NIST) [Mell e Grance, 2011] che descrive il cloud in questo modo:

“CLOUD computing is a model for enabling ubiquitous, convenient, ondemand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or

service provider interaction. This CLOUD model is composed of five essential characteristics, three service models (Software / Platform / Infrastructure as a Service), and four deployment models, whereas the five characteristics are: on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. The deployment models include private, community, public and hybrid CLOUD”.

Nonostante questa definizione sia molto usata, si può notare come essa non definisca il concetto di cloud ma ne descriva solamente lo stato attuale, senza spiegare il concetto che realmente risiede dietro tale modello tecnologico. In questo contesto è difficile sfruttare queste definizioni come linee guida per la ricerca e lo sviluppo in tale settore, inoltre queste definizioni saranno costantemente aggiornate a seconda dello stato attuale delle cose. Per questo, nel report della Comunità europea “Future of CLOUD Computing” [Schubert et al., 2010], si è cercato di dare una definizione di cloud più a lungo termine:

“A 'CLOUD' is an elastic execution environment of resources involving multiple stakeholders and providing a metered service at multiple granularities for a specified level of quality (of service)“

Questa definizione, però, non prende in considerazione il fatto che con il termine CLOUD ogni soggetto in causa ha una diversa prospettiva (e di conseguenza anche il suo concetto, il suo obiettivo e la sua intenzione). Nelle righe seguenti saranno spiegati i punti di vista dei tre soggetti principali coinvolti nel CLOUD: l'utente finale, lo Sviluppatore e il Provider.

Dal punto di vista dell' **utente finale** il cloud deve estendere i servizi generali internet-based con particolare interesse verso la alta disponibilità di servizio, costi ridotti (attraverso un migliorato uso delle risorse), facile da usare ma soprattutto sicuro. L'utente finale tipo non è interessato ai dettagli tecnici che permettono il funzionamento di tali applicazioni. Per quanto riguarda il **Provider** è di fondamentale importanza gestire e fornire le risorse necessarie all'utente finale rendendole economicamente accessibili. Per garantire questa possibilità il provider deve avere applicazioni, servizi e framework che facciano un ottimo uso delle risorse in hosting in modo da ridurre i costi amministrativi e allo stesso tempo soddisfare i requisiti dell'utente. Dal punto di vista dello **sviluppatore** (e del ricercatore) è di particolare interesse cosa può offrire tecnicamente un ambiente CLOUD, rispetto a cosa lo sviluppatore ha bisogno nel suo codice per implementare le funzionalità richieste. Il CLOUD, per sua natura, non prescrive uno specifico approccio tecnico per risolvere questi problemi ma una serie di problemi tecnici derivano implicitamente dal cercare di affrontarli.

3. Scenari di utilizzo e caratteristiche

Come detto nella sezione precedente il CLOUD non ha una reale definizione ma è formato dalla descrizione dello stato attuale della tecnologia, una descrizione quindi in continuo evolversi. Questa situazione nasce dal fatto che non è stata la ricerca (con un'idea a lungo termine) a

segnare la direzione, ma il marketing, le cui politiche ed idee sono a breve (brevissimo) termine. In questo contesto molto interessante è la vista dello stato attuale dei sistemi cloud presentato nel report “Future of Cloud Computing” come riportato in Figura 1.

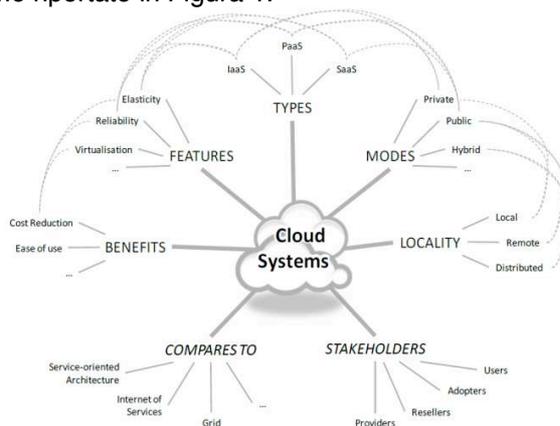


Fig.1 – Visione generale degli aspetti relativi al cloud [Schubert et al., 2010]

In questa sezione si parlerà dell’attuale utilizzo della tecnologia cloud, quindi analizzeremo dettagliatamente i vari tipi di cloud attualmente disponibili, e in che modalità sono sviluppati. Per quanto riguarda i vari modelli di cloud possiamo dividerli in tre grandi macro-aree: Infrastructure as a Service(IaaS), Platform as a Service (PaaS), Software as a Service(SaaS). Con **Infrastructure as a Service (IaaS)** si intende la capacità di fornire al consumatore l’elaborazione, lo storage, le reti, o altre risorse fondamentali di calcolo in cui il consumatore è in grado di distribuire ed eseguire software arbitrario, che può includere sia sistemi operativi che applicazioni. Il consumatore non deve gestire o controllare l’infrastruttura cloud sottostante, ma ha il controllo su sistemi operativi, storage e applicazioni distribuite e controllo, eventualmente limitato, di componenti di rete selezionati (ad esempio firewall host). Con **Platform as a Service (PaaS)** si intende la capacità di fornire al consumatore la possibilità di sviluppare sull’infrastruttura cloud applicazioni create dall’utente o acquisite utilizzando linguaggi di programmazione, librerie, servizi e tutti gli strumenti supportati dal provider. Il consumatore non dovrà gestire o controllare l’infrastruttura cloud di base tra cui la rete, i server, i sistemi operativi, o di archiviazione, ma ha il controllo sulle applicazioni distribuite ed eventualmente le impostazioni di configurazione per l’ambiente di applicazione-hosting. Con **Software as a Service (SaaS)** si intende la capacità di fornire al consumatore la possibilità di utilizzare le applicazioni del provider in esecuzione su una infrastruttura cloud. Le applicazioni sono accessibili da vari dispositivi client attraverso un’interfaccia thin client, come ad esempio un browser Web (ad esempio, la posta elettronica web-based), o attraverso un programma di interfaccia. Il consumatore non gestisce o controlla l’infrastruttura di base compresi la rete, i server, i sistemi operativi, gli storage e le funzionalità delle applicazioni, anche individuali, con la

possibile eccezione delle impostazioni di configurazione dell'applicazione specifiche per l'utente limitato.

A causa della rapida esplosione del cloud anche altre terminologie sono diventate di uso comune, ma sono comunque tutte riconducibili a delle sfaccettature più dettagliate delle 3 macroaree sopra definite. Nella Figura 2 è possibile vedere uno schema di tutti i modelli di cloud attualmente utilizzati.

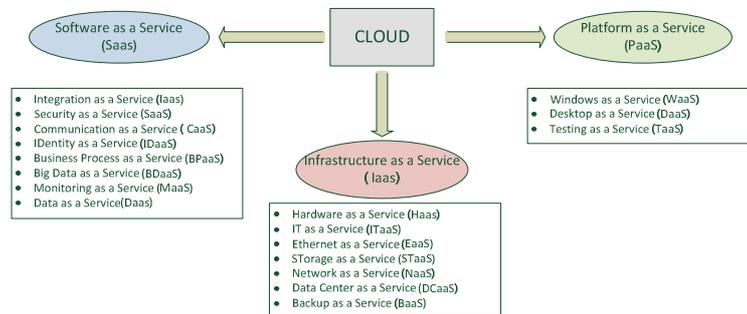


Fig.2 – Elenco dei principali modelli del cloud

Dopo aver parlato dei vari tipi di cloud analizzeremo i vari tipi di implementazione soffermandoci sui modelli più diffusi: Cloud Privato, Cloud Community, Cloud Pubblico, Cloud ibrido (Figura 3).

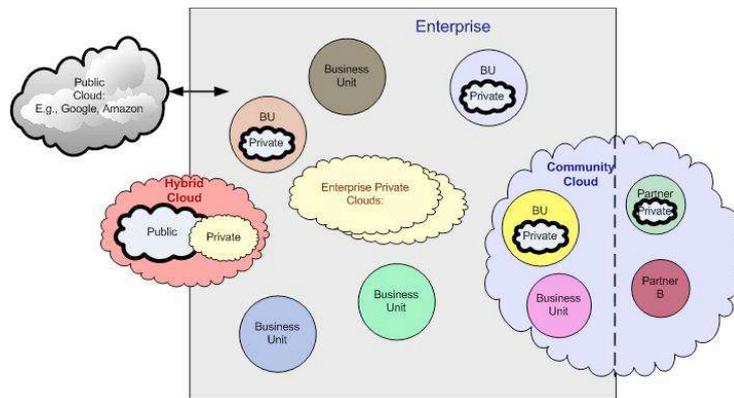


Fig.3 – Esempio delle varie tipologie di cloud dal punto di vista aziendale

Nel **Cloud Privato**, l'infrastruttura cloud è utilizzata esclusivamente da parte di un'unica organizzazione che comprende più i consumatori (ad esempio, le unità di business). Esso può essere di proprietà, o gestita da una azienda terza parte, o una combinazione di essi. L'infrastruttura può risiedere sia in sede dell'azienda consumatrice sia direttamente nell'azienda provider ma con componenti ad uso esclusivo. Nel **Cloud Community** l'infrastruttura cloud è utilizzata non più per la singola organizzazione, ma è condiviso da una specifica comunità di consumatori che hanno stessi obiettivi e/o richieste (ad esempio mission, requisiti di sicurezza, politica e considerazioni di conformità). Anche in questo caso l'infrastruttura può

risiedere ed essere gestita in una o più organizzazioni della comunità. Nel **Cloud pubblico** l'infrastruttura cloud risiede direttamente nei locali del provider e non nei consumatori, questo perché essendo un servizio pubblico l'utilizzo è aperto a tutti e non ad una cerchia ristretta di organizzazioni. Nel **Cloud ibrido** l'infrastruttura cloud è una composizione di due o più infrastrutture cloud distinte (privato, community o pubblico) che restano soggetti unici, ma legati insieme da una tecnologia standardizzata o proprietaria che permette ai dati la portabilità delle applicazioni (ad esempio, il cloud bursting per il bilanciamento del carico tra nuvole).

4. Problematiche attuali

Come già detto in precedenza la tecnologia cloud non è nata attraverso studi o ricerche ma attraverso delle esigenze di mercato manifestatesi in questi ultimi anni, ciò ha comportato anche ad una serie di potenzialità ed opportunità mancate che tale modello tecnologico può apportare, tenendo bene in mente che anche queste potenzialità non sono tutt'ora state inquadrate correttamente. Questa sezione vuole approfondire questi temi soffermandosi su cosa potrebbe fare questa tecnologia se sfruttata al massimo del suo potenziale.

Uno degli aspetti di maggior rilievo è la scalabilità, sia intesa come scalabilità verticale che orizzontale. Attualmente la scalabilità verticale (parallelismo) è molto difficile da raggiungere e questo comporta una riduzione significativa di performance. D'altra parte la scalabilità orizzontale (replica di istanze) può essere ancora ottimizzata, infatti mentre attualmente gli oggetti vengono replicati completamente in caso di richieste multiple, spesso i dati che sono necessari al consumatore sono solo una parte dell'istanza e non l'istanza completa riducendone anche in questo caso le performance. Una delle più importanti conseguenze di questo problema è lo spreco delle risorse utilizzate, nonostante il cloud si ponga come una tecnologia in grado di ridurre le risorse necessarie per determinate azioni ancora ci sono molti margini di miglioramento, il rischio infatti è che con l'aumentare delle risorse anche il cloud attuale non sia in grado di soddisfare le richieste, considerando che molte tecnologie (come la virtualizzazione) aumentano il normale uso di risorse.

Uno dei problemi più sottostimati è il multi-tenancy, attualmente il QoS è associato direttamente ad una risorsa, in caso di picchi di richieste su questa risorsa esiste la possibilità che alcuni utenti non riescano ad accedervi. Teoricamente questo limite potrebbe essere aggirato assegnando ad ogni singolo utente uno specifico QoS, in questo modo la risorsa necessaria all'utente può essere raggiunta senza alcuna restrizione [Koh et al., 2007] [Pu et al., 2010].

Un altro importante aspetto non preso in considerazione fino a poco tempo fa è il fatto che cloud diversi hanno interfacce diverse. I primi passi per ovviare a questa problematica è la creazione di piattaforme standard per generare interfacce standard. Piattaforme come OpenStack ed OpenNebula sono progetti open source nati per questo scopo in cui rispettivamente grandi aziende e comunità europea hanno deciso di puntare per avere una soluzione standard per le proprie applicazioni. Le attuali applicazioni sono

un altro limite in quanto non sono predisposte per essere trasferite su infrastrutture cloud. Questa predisposizione non è dovuta solo al fatto che le applicazioni non sono programmate per soddisfare le caratteristiche della tecnologia ma anche dal fatto che lo stesso cloud, inteso quello attuale, non è in grado di soddisfare i requisiti impliciti delle applicazioni.

La gestione dei dati è un altro settore in cui il cloud non sta esprimendo tutte le potenzialità, infatti il software cresce ad una velocità maggiore dell'hardware, che quindi è costretto a gestire sempre più attività e che le applicazioni stanno aumentando di interattività client-server. Un futuro sistema potrebbe avvalersi di informazioni dislocate differentemente nella rete senza alcuno specifico indirizzamento all'istanza primaria, ma comunque in grado di "replicarla" e renderla disponibile. In questo contesto rendere i dati frammentati nella rete rendono necessari alcuni supporti per rendere disponibili le risorse necessarie senza l'utilizzo di complicati e pesanti software. Questa problematica è strettamente correlata al problema della gestione dei big data dove i dati eccedono la gestibilità dei software e dei sistemi [Longbottom, 2011][[The Economist, 2010]. L'ultimo aspetto da considerare è la sicurezza e la privacy. Portare dati sensibili nella rete significa doverne aumentare la sicurezza e quindi realizzare nuovi protocolli e policies in grado di rendere affidabile questo passaggio. Ancora infatti molti aspetti legali sulla gestione e protezione dei dati non sono stati definiti, e questo mettendo incertezza al consumatore fa sì che sia ancora restio ad usare il cloud per alcune operazioni più sensibili di altre.

Oltre ai temi citati precedentemente che solo adesso si iniziano ad affrontare, ce ne sono altri che ancora sono completamente ignorati, ma che possono portare ulteriore valore al Cloud. Ad esempio attualmente le infrastrutture di ogni singolo provider sono isolate rispetto alle altre, avere un mondo in cui tutte queste infrastrutture possano diventare interoperabili andrebbe ad abbattere tantissimi costi, ed aumentare la qualità del servizio.

5.Proposte di ricerca

In questa sezione sono analizzati i problemi di ricerca da risolvere per migliorare la tecnologia cloud soffermandosi sulla complessità e il tempo necessario di risoluzione. Il seguito di questa sezione analizzerà i caratteri tecnologici da affrontare per aiutare l'evoluzione del cloud.

Il primo aspetto da prendere in considerazione è la gestione dei dati il cui volume sta crescendo a dismisura, e i quali inoltre diventano sempre più complessi e dispersi nella rete. Spesso i dati infatti sono generati da sensori sparsi in tutta la rete, mentre le strutture cloud sono generalmente centralizzate. La gestione dei dati è forse la problematica per eccellenza e che a causa della sua complessità non ha obiettivi di risoluzione a breve termine. Altro aspetto correlato ai dati è il bisogno di avere reti sempre più performanti, i dispositivi mobili in questi ultimi anni stanno esplodendo e questo fa sì che sempre più dispositivi si allaccino alla rete e quindi pretendano una banda sufficiente per soddisfare ogni singola richiesta. In questo caso le tempistiche sono più brevi rispetto alla gestione dei dati grazie anche la minore complessità del problema.

La descrizione e l'utilizzo di risorse è un altro tema molto importante (Metadata). Il bisogno di descrivere le risorse è nato dal fatto di avere software automatici in grado di leggere queste descrizioni e prendere le corrette decisioni. I sistemi distribuiti (GRID) hanno già fatto in modo di iniziare le ricerche in questa direzione e quindi i requisiti base possono essere raggiunti in medio termine, mentre una soluzione completamente autonoma aumenta notevolmente la complessità e di conseguenza anche i tempi della ricerca. Oltre a descrivere la risorsa è molto importante come essa venga utilizzata, per ottimizzarne l'utilizzo è necessario conoscere cosa intende fare l'utente in modo da non creare alcun conflitto modificando il sistema in modo elastico. Uno dei problemi più grandi è che i dati derivano da una grande vastità di fonti rendendoli diversi l'uno dall'altro, quindi il continuo proliferare di nuove fonti aumenterà considerevolmente la complessità. Come già accennato in precedenza, il multi-tenancy è molto importante nella tecnologia cloud, e quindi è molto importante studiare tecniche che permettano di gestire questa feature raggiungendo allo stesso tempo un adeguato livello di gestione delle risorse, della sicurezza e degli altri aspetti richiesti. Attualmente le procedure di controllo non sono adeguate e la risoluzione di questo problema di altissima complessità non è da considerarsi solubile in tempi brevi.

Tra le aree di ricerca c'è senza dubbio la creazione di federazioni di provider, con un numero crescente di fornitori di servizi cloud IaaS al PaaS, una federazione (da fornitori sempre più legati) in grado di fornire maggiore resistenza ed elasticità diventa una strategia vincente. Tuttavia, i fornitori attualmente leader cloud hanno sempre rifiutato di accettare un'interfaccia comune (API), inoltre ci sono problemi di portabilità e le questioni complesse della legalità dei diritti di accesso. Il raggiungimento di questo obiettivo è senza dubbio a lungo termine, ma può essere raggiunto attraverso piccoli passi effettuabili in più breve tempo, come ad esempio può essere la creazione di piattaforme standard su cui implementare i servizi cloud.

Con l'arrivo del cloud molte tecniche attualmente in uso sono diventate automaticamente vecchie in quanto non sono in grado di sfruttare a pieno il potenziale della nuova struttura hardware, è quindi necessario lo sviluppo di nuovi sistemi e/o metodi di programmazione in cui i metadati siano al centro del codice dando la possibilità di fornire parametri relativi al ridimensionamento, all'elasticità e alla sicurezza. In questo contesto è difficile definire i termini di ricerca in quanto è richiesto un elevato livello per modificare e/o creare molti dei paradigmi di programmazione esistenti.

L'ultimo aspetto tecnologico, ma non il meno importante, è la sicurezza. Quando i dati vanno "out of house" ovviamente vengono a manifestarsi tutti i problemi relativi alla sicurezza, sia legislativi (ogni organizzazione ha una propria legge) sia tecnologici. Queste problematiche derivano da aspetti di multi-tenancy, da dati internazionali, dalla distribuzione di codice e dai dati associati delle risorse umane. Inoltre un'operazione di ICT complessa che coinvolge molte persone in molti paesi con standard di professionalità ed etica variabili fanno sì che si abbiano rischi più elevati di sicurezza. A livello temporale probabilmente questa è la peggior area, la sensazione attuale è che problemi di sicurezza saranno sempre presenti e bisognerà quindi

affrontare minacce sempre più evolute e di conseguenza aggiornare e rimodellare in continuazione nuovi sistemi di sicurezza.

La Tabella 1 mostra gli argomenti (tecnologici e non) di ricerca più rilevanti per far evolvere e migliorare la tecnologia del cloud computing.

Gestione Dati	Comunicazione e Rete
<ol style="list-style-type: none"> 1. Gestione big data su reti a larga scala; 2. Soddisfare il requisito di real time per determinati dati multimediali; 3. Distribuire una gigantesca quantità di dati dai sensori verso i centri Cloud; 4. Integrità e struttura dei dati; 5. Sicurezza dei dati. 	<ol style="list-style-type: none"> 1. Banda garantita (non rispetto alla risorsa ma all'utente); 2. Mantenimento della connessione mobile.
Metadata	Interoperabilità-Portabilità
<ol style="list-style-type: none"> 1. Modalità per descrivere metadata in maniera generica; 2. Capire come e quali software possono interagire con i metadata è la chiave di molti fattori (elasticità, interoperabilità, ..); 3. Virtualizzazione, tramite descrizione dei metadata, dei devices, della rete, dei dati e delle infrastrutture distribuite. 	<ol style="list-style-type: none"> 1. Metodi di sviluppo di sistemi semiautomatici che permettono di avere compormamenti dinamici nel tempo; 2. Unire clouds of different provider; 3. Gestione di un' applicazione in un sistema federato; 3. Gestione dello SLA tra 2 provider; 4. Interfacce e formato dati standard.
Multi-Tenancy	Sicurezza
<ol style="list-style-type: none"> 1. Isolare le prestazioni e le sezioni di rete (necessario per la gestione delle risorse e la sicurezza); 2. Metodo di sviluppo di nuovi sistemi, probabilmente utilizzando tecniche orientate ai servizi. 	<ol style="list-style-type: none"> 1. Trattare le richieste degli utenti senza richiedere informazioni sensibili; 2. Gestione delle policies di sicurezza; 3. Individuazione di errori ed attacchi; 4. Isolamento di processi di carico ad alta sicurezza.
Programmabilità e Usabilità	Gestione delle risorse
<ol style="list-style-type: none"> 1. Ristrutturare gli algoritmi; 2. Nuovi modelli di design; 3. Capire l' interdipendenza tra le varie applicazioni; 4. Monitoraggio delle informazioni; 5. Multi-Tenancy; 6. Ottimizzare gli algoritmi per l'efficienza energetica; 7. Elasticità, orizzontale e verticale per consentire l'utilizzo delle risorse ottimizzato mantenendo SLA / QoS. 	<ol style="list-style-type: none"> 1. Gestire in maniera efficiente l' eterogeneità dei dati; 2. Schedulamento multidimensionale, dinamico e a larga scala per rispettare il timing e QoS; 3. Upgrade e & Downgrade efficiente; 4. Esplorare tutte le specifiche a tutti i livelli; 5. Gestione del consumo di risorse; 6. Gestione energetica efficiente delle risorse.
Politica e Legislazione	Costo del Modello e Business
<ol style="list-style-type: none"> 1. Problemi di privacy nei trasferimenti internazionali di dati da utente a cloud; 	<ol style="list-style-type: none"> 1. Appropriati modelli di pagamento; 2. Controllo pagamenti in caso di multi tenancy.

Tab.1 – Principali tematiche di ricerca del cloud computing

6. Conclusioni

Il cloud attualmente non è una sofisticata ed affermata tecnologia allo stesso tempo però i sistemi cloud in genere soffrono seri (a lungo termine) inconvenienti. Questi inconvenienti sono meno conosciuti a causa della

mancanza di esperienza nel dominio cloud e ad una chiara definizione delle aspettative degli utenti. Vi è tuttavia un forte potenziale e la necessità di una base stabile a lungo termine, su cui la tecnologia cloud e in generale il futuro di internet possono continuamente evolversi [Zimmermann, 2011].

In queste circostanze la ricerca e lo sviluppo non sono fermi, in particolare nelle aree ad alta densità industriale nuove soluzioni e nuove tecnologie sono sviluppate rapidamente in base alle esigenze degli utenti e all'evoluzione del mercato. In opposizione a questo mondo esiste un altro mondo, quello accademico, nel quale la ricerca si muove più lentamente, anche se la qualità dei risultati e la loro rilevanza a lungo termine possono essere più elevati. In aggiunta a questo, i risultati della ricerca sono generalmente accompagnati da grande scetticismo, in quanto gli utenti non hanno fiducia in questi risultati poiché spesso non sono pronti per essere convertiti in vantaggi tecnologici del mondo reale e per questo motivo il loro assorbimento richiede un lavoro iniziale di adattamento.

Il mercato IT e l'industria non hanno ancora raggiunto il loro pieno potenziale sia per campo di applicazione (numero di utenti e dispositivi) così come per l'eterogeneità di risorse. Per fronteggiare questa crescita sarà quindi estremamente necessario spingere al limite le prestazioni, il networking, la comunicazione ecc. Il concetto di Cloud offre un enorme potenziale per affrontare questa crescita, sia in dimensioni che complessità, ma non deve essere confuso con la "soluzione finale" del problema. Mentre il mercato si espanderà ulteriormente, sorgono altri problemi che prima o poi richiederanno funzionalità aggiuntive che il concetto cloud non riuscirà a soddisfare, dando così origine ad un nuovo modello.

Bibliografia

Fenn J., LeHong H., Hype Cycle for Emerging Technologies 2011, Gartner. Disponibile online: http://www.gartner.com/DisplayDocument?doc_cd=215650

Koh Y., Knauerhase R., Brett P., Bowman M., Wen Z., Pu C., An Analysis of Performance Interference Effects in Virtual Environments, IEEE International Symposium on Performance Analysis of Systems & Software, 2007.

Longbottom, C., Big data: large problems or huge opportunities? ComputerWeekly, 2011. Disponibile online: <http://www.computerweekly.com/news/2240105424/Big-data-large-problems-or-huge-opportunities>

Mell P., Grance, T., The NIST Definition of Cloud Computing, National Institute of Standards and Technology, 2011.

Pu X., Liu L., Mei Y., Sivathanu S., Koh Y., Pu, C., Understanding Performance Interference of I/O Workload in Virtualized Cloud Environments, IEEE 3rd International Conference on Cloud Computing, 2010.

Schubert L., Jeffery K., Neidecker-Lutz B., The Future of Cloud Computing – Report from the first Cloud Computing Expert Working Group Meeting, 2010.

The Economist, Data, data everywhere, 2010. Available online: http://www.economist.com/node/15557443?story_id=15557443

Zimmermann R., Cloud Computing. FIA Budapest plenary session, 2011.

Smart Cities, Urban Sensing and Big Data: Mining Geo-location in Social Networks

Daniele Sacco¹, Gianmario Motta², Linlin You³, Nicola Bertolazzo⁴, Chen Chen⁵

Università di Pavia

Via Ferrata 1, 27100, Pavia (PV)

¹ daniele.sacco01@ateneopv.it

² motta05@unipv.it

³ linlin.you01@ateneopv.it

⁴ nicola.bertolazzo01@ateneopv.it

⁵ chen.chen01@ateneopv.it

Location based social networks offer spatio-temporal information which can be accessed through public Application Programming Interfaces (APIs) and drew the interest of researchers with diverse scientific backgrounds. This availability of data enables a potential use of geo-located content as an additional, low cost and infrastructure-less source of information for urban sensing in Smart Cities. All these aspects bounded with the need of real-time analytics for urban sensing takes to Big Data management and its related issues. A systematic literature review outlines related works and gaps in current research. We propose a reference model to exploit Big Data and Open Data for urban sensing and we validate it by a case study. Finally, we give recommendations for future research about location and mobility mining of social network data.

Keywords: big data, urban sensing, smart city, location-based social network, twitter, data mining, open data

1. Introduction

Last years have seen a paradigm shift of the role of the user on the web, from consumer to producer of information. Due to last technological developments and their world-wide dissemination, Web 2.0 changed users' approach to information creation and exploitation. The paradigm shift has been also a social shift. Web has become an essential need for people, a sort of commodity, that is used to communicate, interact, share information and even maintain relationships thanks to social networks.

The last advent of smartphones, equipped with GPS sensors allowing users to geo-locate themselves, can take to the next shift, from a social and collaborative Web 2.0 to a local and mobile Web 3.0. One among the first achievements has been the integration of Geographic Information Systems

(GISs) and social networks resulting in new location-based capabilities. Social networks that include location information into their contents are called Location Based Social Networks (LBSNs). So, LBSNs offer spatio-temporal information which can be accessed through public Application Programming Interfaces (APIs) and draw the interest of researchers with diverse scientific backgrounds. This availability of data enables a potential use of geo-located content as an additional, low cost and infrastructure-less source of information for urban sensing in Smart Cities.

All these aspects bounded with the need of real-time analytics for urban sensing takes to Big Data management and its related issues. Real-time urban sensing use citizens as active and passive sensors and can reveal important insights of human behavior in the city. Diverse use scenarios can enable new perspectives in society level, e.g. community healthcare, public safety, city resource management and transportation management.

2. Systematic Literature Review

Systematic Literature Review (SLR) can be regarded as “explicitly formulated, reproducible and up-to-date summary” [Egger et al, 2008] that includes and extends the statistical results of a meta-analysis methodology. As opposed to narrative reviews, it is based on a structured method that is always explicitly specified at the beginning of the review.

Our objective is to identify initiatives, experiences and viewpoints on location and mobility mining of social network data. So, our research question is “How to exploit geo-located data from social networks and what level of maturity has reached its application?”. So, the expected outcomes of our SLR are: (a) a complete overview of the “state of the art”, (b) the identification of gaps in current research, solutions, trends and future research and suggestions to the community of researchers and practitioners, and (c) recommendations about best practices for location and mobility mining of social network data.

Extracted information contains techniques, issues, models and any other kind of topic useful to provide an accurate snapshot of the current state of location and mobility mining of social network data. 87 out of 109 articles have been selected and classified by year of publication, geographical area, research method and publication channel.

There are no significant articles before 2009 because Location Based Social Networks (LBSNs) raised in the same year. Afterwards researchers could start to consider data provided by LBSNs. The articles considered in our research range from 2009 to second quarter of 2013. Considering that publications in 2013 refer only to the first months of the year, the number of publications tends to have an exponential growth (2^x): 2 publications in 2009, 4 in 2010, 14 in 2011, 38 in 2012. Case studies represent the majority of the study types. The number of case studies and instrument development publications (87.3% in total) reflects the experimental approach to the issue. It is also motivated by the low number of theoretical papers (5.7%) and by the lack of position papers.

3. Discussion

We have identified 3 main domains: (a) data sources, (b) technologies, (c) use scenarios. Let us consider the main contributions to each domain.

4.1 Data Sources

The data sources that can be inferred for urban sensing are heterogeneous. Three innovative data sources exist: (a) mobile sensor data about the individual devices, (b) infrastructure sensor data about the context, and (c) social data from social network and other internet services [Zhang et al, 2011]. Data sources can be used independently but the combination of the three kinds can provide a comprehensive understanding of human behavior and its context. Here we focus only on data sources for geo-location from the social network tier and how they have been used in literature.

Twitter is a widely-used platform for the real-time social sharing of short text-based messages called tweets. Twitter and smartphone usage reflected same growth, indeed Twitter users interact frequently on mobile devices. As Twitter is easy to use and interactions are short, many users post tweets despite they are engaged in other activities. This gives Twitter data good spatial and temporal coverage because tweets can be automatically geo-located [Mai et Hranac, 2013]. Twitter provides a free real-time streaming API through which a sample of all tweets can be retrieved. The API provides filters that can be set on these data streams to capture tweets within a geographic area or only those containing certain terms. However, data stream is limited to 1% of total tweet volume. So, only a subset of the total tweets can be used.

Foursquare is a location-based social network where users can “check in” to different locations and share them with friends both on Foursquare itself and other social networks. Users can upload pictures at a venue or leave ‘tips’ on the venue page (e.g. a user may ‘check-in’ to a hotel and leave a ‘tip’ about how bad the service is) [Cheng et al, 2011]. Foursquare check-in data is not directly accessible: however, users typically decide to share their check-ins publicly on Twitter, so they can be retrieved via Twitter streaming API.

Several other papers use data sets published for research purposes, because no API is available or the social network stopped its service. For example, Gowalla was a location-based social network created in 2009. The concept behind the service was to advertise your exact location to all your friends in real-time [Scellato et al, 2010]. Also BrightKite was a social networking website where users could share their location, to post notes and to upload photos. By making “check-ins” at places, users could see people who is nearby. Now, only already collected data sets are available [Li et Chen, 2009].

Other available services are Momo and Flickr. However, a question could rise: why not to use Facebook, the most popular service? The main reason is that Facebook API can be used to retrieve data from those users who accepted to publish their posts to your application or system, so it is not publicly available.

4.2 Technologies

We discuss here the main technologies used for urban sensing. The SLR shows that related works focus on machine learning techniques.

K-means can be used to reveal clusters of common behaviors across land segments. The land use of each cluster can be derived by analyzing the activity vectors of the regions within the cluster. K-means depends on the initial random selected seeds and it needs to specify the number of clusters k (land uses) to identify [Frias-Martinex et al, 2012]. [Lee et al, 2011] used a similar approach, but they also formed a Voronoi diagram using the center points (latitude, longitude) of the K-means results and regarded the formed regions as a set of region of interests, to identify the occurrence of local events.

A Self-Organizing Map (SOM) is an unsupervised neural network that reduces the input data dimensionality to be able to represent its distribution as a map. So, SOM forms a map where similar samples are mapped close together. [Frias-Martinex et al, 2012] used SOM to build a map that segments the urban land into geographical areas with different concentrations of tweets in the time period under study.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) has specific characteristics: (a) it is based on the concept of density reachability, producing satisfying results identifying arbitrarily shaped clusters, (b) the number of clusters is not given a priori, and (c) the algorithm tolerates noise, allowing for some data points not to be assigned to any cluster [Villatoro et al, 2013]. The advantage of density based clustering algorithms is that clusters are defined by the density of data points and not by spatial size and form of cluster.

Spectral methods for data clustering are popular because of the quality of the clusters that are produced and the simplicity of implementation. Spectral clustering is able to find arbitrarily shaped clusters and does not pose any constraints on them (in contrast to the k -means, for example, which assumes cluster to be convex). It requires anyway the parameter k to define the number of desired clusters [Roslrer et Liebig, 2013].

Mean-shift is a non-parametric clustering technique that detects the modes of an underlying probability distribution from a set of discrete samples. So, mean-shift can be used both as an algorithm to detect local maxima (modes) as well as a clustering technique (areas associated to the modes). [Frias-Martinex et al, 2012] assume that there exists an unobservable underlying probability distribution of where people tweet from. The modes of that distribution are determined to represent urban landmarks or points of interest in the city.

4.3 Use scenarios

The identification of use scenarios aims to enable new perspectives in society level, e.g. community healthcare, public safety, city resource management and transportation management..

We here propose a classification for visualization objectives in urban sensing, representing an evolution of existing Business Intelligence (BI) solutions towards Geographic Information Systems (GISs) and Big Data visualization [Stodder, 2013], as shown in Tab. 1.

Class	Description	BI similarities	Objectives
Urban characterization	The results can be stored for users as a “snapshot” of a certain point in time. Users examine snapshots to identify changes in data over time, so they must be provisioned and presented consistently so that trends and comparisons are valid	It recalls dashboards. The view is static and it is previously defined by data analysts.	To visualize and predict social ties and urban structure
Spatial discovery	It enables users to interact with data through analytical processes. Visual functionality for filtering, comparing, slicing and dicing, drilling down, and correlating data can then be integrated with the users' analytical application functions for forecasting, modeling, and statistical, what-if, and predictive analytics.	It recalls On-Line Analytical Processing (OLAP). The view is dynamic and it allows users to navigate data.	To analyze behaviors “on-line”, in space and time
Exception alerting	It notifies users of particularly important changes in the data or when situations arise that demand immediate attention. Alerts mean that something important in real-time data or event streams is happening.	It recalls event processing in modern Business Activity Monitoring (BAM) solutions, that detect and warn of problems or exceptions in real-time.	To detect events or exceptions to standard behaviors, e.g. disasters, diseases, unexpected crowds

Tab. 1 – Visualization objectives classification

The three visualization classes reflect papers analyzed by our SLR. Respectively, 30 papers deal with urban characterization, 18 with spatial discovery, 12 with exception alerting. The rest of papers use social networks geo-located data mainly to build recommender systems or user profiling. These papers are single user oriented, so we do not take them into account in following paragraphs because they do not give a broad view of the city.

4.3.1 Urban characterization

[Rosler et Liebig, 2013] provide insights on the activity profiles in urban environments. Clusters identified by Foursquare check-ins help to describe the socio-dynamics of urban districts in different times of the day. [Ferrari et al, 2011] extend the work on activity profiles by providing also mobility patterns that occur in an urban environment and understanding of social commonalities between people. Traditional municipal organizational units such as neighborhoods are studied by Livehoods project from [Cranshaw et al, 2012], who shows that their boundaries do not always reflect the character of life in these areas. [Joseph et al, 2012] mine check-ins to identify groups of people which are of different types (e.g. tourists), communities (e.g. users tightly clustered in space) and interests, and how they use urban space.

By processing Twitter data, [Wakamiya et al, 2011] are able to examine the relation between regions of common crowd activity patterns and major categories of local facilities. [Wakamiya et al, 2012] extend their work and correlate psychological and geospatial proximity of urban areas by borrowing crowd's experiences from geo-tagged tweets, in order to demonstrate that people often rely on geospatial cognition to the real space than the exact physical distance in the real world.

4.3.2 Spatial discovery

[Silva et al, 2012] study social behaviors by monitoring check-ins in Foursquare, [Sagl et al, 2012] by Twitter and Flickr data. They are able to analyze city dynamics spatially and temporally and to identify seasonality in human behaviors. [Cheng et al, 2011] investigate 22 million check-ins across 220,000 users and report a quantitative assessment of human mobility patterns by analyzing the spatial, temporal, social, and textual aspects associated with these footprints.

4.3.3 Exception alerting

[Boettcher et Lee, 2012] introduce EventRadar, a detection system that finds local events such as release parties, musicians in a park, or art exhibitions. [Watanabe et al, 2011] provide a similar system to detect events. [Baldwin et al, 2012] also try to predict events impact. [Mai et Hranac, 2013] assert that tweets can be matched to traffic incidents by examining the content of the tweets for key words and comparing locations of the tweets and incidents. Results are confirmed for areas with sufficient density of Twitter usage.

4. A Big Data-based approach

Big data is a popular term used to refer to the exponential growth, availability and use of information. However, as Gartner states (<http://www.gartner.com/it-glossary/big-data/>), Big Data doesn't focus only on the high volumes of information, but also on "data velocity", that involves streams of data that are produced and processed in real-time, and on "data variety", that refers to more types of information to analyze, e.g. social media, context aware data, documents, images, videos, audio, etc.

All papers in our review obviously consider "data variety" because location-based social networks' data is geo-located and unstructured (typically JSON files), however the challenge results from the expansion of all three properties given by Gartner, rather than just "data volume" or "data velocity" alone.

We here propose a big data-based approach to mining of social network geo-located data. Our approach intends to consider the 3 properties by applying technologies that fit Big Data management. Our literature review found out no papers take into account big data-based system architectures to process "variety", "velocity" and "volume" of social network data. Here we propose a simple architecture to process this kind of data and possible extensions to support more analytics.

In order to validate our approach, we report a case study about public transport monitoring. First we developed a simulator that transforms timetables from open transport data in Torino (Italy) to a visualization tool that shows planned position of each bus in a specific time (each moving point in Fig. 1). We decided to enrich this visualization and leverage urban sensing. Our specific question was: “how to correlate transport planning and urban activity areas?”. Exploitation of location-based data from social network is a viable way.

To define activity areas in a city we built a tool for spatial discovery that could allow us to model crowds by density clusters within the city and drill down them according to different scales. In order to keep the case study simple, our approach has been to compute density on a fixed grid applied over the map (rectangles with same dimension). However, what kind of data could be used? We exploited Twitter streaming API, collected geo-located data in Torino area in real-time and clustered them on the map for each specific time range (1 hour).

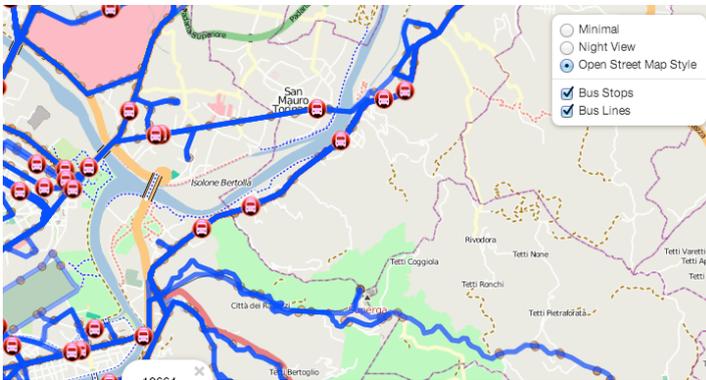


Fig. 1 – Public transport simulator

To retrieve Twitter data we implemented FluenTD agents by Node.js on a virtual machine in Amazon Elastic BeanStalk. FluenTD (<http://fluentd.org/>) is an open-source log collector that enables to have a logging architecture with more than an hundred types of systems, by treating logs as JSON. Node.js (<http://nodejs.org/>) is a platform built on JavaScript for easily building fast, scalable network applications. It uses an event-driven, non-blocking I/O model that makes it lightweight, efficient, and oriented to data-intensive real-time applications that run across distributed devices. FluenTD and Node.js integration allows to build real-time collection and filtering of geo-located tweets, and their storage in TreasureData (<http://www.treasure-data.com/>). TreasureData is a “Big Data as a Service” cloud solution that offers a time series, columnar, Hadoop-based, schema-free data warehouse stored on Amazon S3. It allows you to access data using Hive query language (<http://hive.apache.org/>) by JDBC. An Extraction-Transformation-Loading (ETL) process access data in TreasureData, processes it and uploads it in CartoDB every 5 minutes. CartoDB is a “database as a service” cloud solution, based on a PostgreSQL database with GIS extension. As PostgreSQL databases can use

extensions to run data mining algorithms, we decided to apply density-based clustering online, whenever the user changes the zoom scale in the map. Fig. 2 shows the final architecture for our prototype.

As this solution intends to implement Service Oriented Architecture (SOA), it has two main advantages: (a) it can be easily extended, and (b) a component can be replaced by others because of decoupling. For example, the data stored in the data mart can be used to perform different analytics or to build an alerting system. As SAP Hana has a Predictive Analysis Library that offers native support to DBSCAN, we could use SAP Hana to implement it as a supplementary layer between TreasureData and CartoDB. If we want to easily implement Latent Dirichlet Allocation we could replace TreasureData with Mahout on a Hadoop cluster. The ETL process, that may represent a bottleneck in terms of performances, can be replaced by ad hoc solutions, e.g. interfaces developed in Node.js to provide data filtering and storage in the data mart.

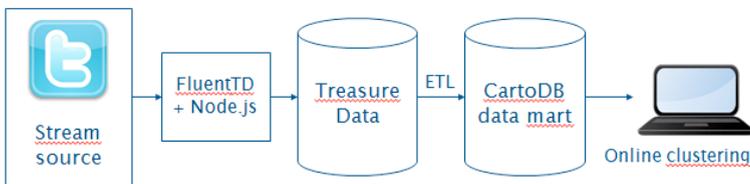


Fig. 2 – A Service Oriented Architecture for Big Data management

Fig. 4 shows early results of our prototype. It considers tweets collected from 8 AM to 9 AM on July 10th, 2013. Color density of rectangles change according to tweet density in the same area.

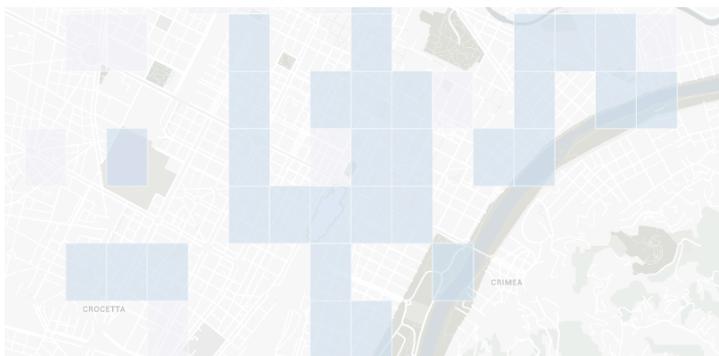


Fig. 4 - Tweet density in Torino

Final integration between the crowd modeling tool and the public transport simulator allows spatial discovery of urban areas not covered by transport service during peaks of presence of people who may need to move from one place to another within the city.

5. Conclusion

Our SLR demonstrated the exponential growth in number of papers about this trending topic. The high number of case studies and instrument development publications reflects the experimental approach to the issue and it may also lead developers and freelancers to provide smart solutions to contextual problems in the city.

Most of related works focus on validation of data mining techniques, rather than their application in real use scenarios. Our prototype demonstrated that the integration of urban sensing and open data can help municipalities to reveal enhanced insights about their services to citizens.

So, future research should move from validation of techniques to their real application in Smart Cities by exploiting Big Data technologies, thus providing real time analytics to municipalities and end users.

Our future works intend to integrate different data sources, such as sensors, and more social media, in order to provide deeper insights for urban sensing. Furthermore, our system needs to scale up and validate performances against user needs. Next step will be the development of a service orchestration layer to provide complete validation of our Service Oriented Architecture.

References

- [Baldwin et al, 2012] Baldwin, T., Cook, P., Han, B., Harwood, A., Karunasekera, S., & Moshtaghi, M. (2012, April). A support platform for event detection using social intelligence. 13th Conference of the European Chapter of the Association for Computational Linguistics (pp. 69-72).
- [Boettcher et Lee, 2012] Boettcher, A., & Lee, D. (2012, November). EventRadar: A Real-Time Local Event Detection Scheme Using Twitter Stream. In Green Computing and Communications (GreenCom), IEEE International Conference on (pp. 358-367).
- [Chen et al, 2013] Chen, T., Kaafar, M. A., & Boreli, R. (2013). The Where and When of Finding New Friends: Analysis of a Location-based Social Discovery Network. International Conference On Weblogs And Social Media
- [Cheng et al, 2011] Cheng, Z., Caverlee, J., Lee, K., & Sui, D. Z. (2011). Exploring Millions of Footprints in Location Sharing Services. ICWSM, 2011, 81-88.
- [Cranshaw et al, 2012] Cranshaw, J., Schwartz, R., Hong, J. I., & Sadeh, N. M. (2012, June). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. In ICWSM.
- [Egger et al, 2008] Egger, M., Smith, G. D., & Altman, D. (Eds.). (2008). Systematic reviews in health care: meta-analysis in context. Wiley.
- [Ferrari et al, 2011] Ferrari, L., Rosi, A., Mamei, M., & Zambonelli, F. (2011, November). Extracting urban patterns from location-based social networks. 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks (pp. 9-16)
- [Frias-Martinez et al, 2012] Frias-Martinez, V., Soto, V., Hohwald, H., & Frias-Martinez, E. (2012, September). Characterizing Urban Landscapes using Geolocated Tweets. 2012 IEEE International Conference on Social Computing (pp. 239-248).

[Joseph et al, 2012] Joseph, K., Tan, C. H., & Carley, K. M. (2012, September). Beyond local, categories and friends: clustering foursquare users with latent topics. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 919-926)

[Lee, 2012] Lee, C. H. (2012). Mining spatio-temporal information on microblogging streams using a density-based online clustering method. Expert Systems with Applications, 39(10), 9623-964

[Lee et al, 2011] Lee, R., Wakamiya, S., & Sumiya, K. (2011). Discovery of unusual regional social activities using geo-tagged microblogs. World Wide Web, 14(4), 321-349.

[Li et Chen, 2009] Li, N., & Chen, G. (2009, August). Analysis of a location-based social network. IEEE International Conference on CSE'09 (Vol. 4, pp. 263-270)..

[Mai et Hranac, 2013] Mai, E., & Hranac, R. (2013, January). Twitter Interactions as a Data Source for Transportation Incidents. In Transportation Research Board 92nd Annual Meeting (No. 13-1636).

[Rosler et Liebig, 2013] Rösler, R., & Liebig, T. (2013). Using Data from Location Based Social Networks for Urban Activity Clustering. In Geographic Information Science at the Heart of Europe (pp. 55-72). Springer International Publishing.

[Sagl et al, 2012] Sagl, G., Resch, B., Hawelka, B., & Beinat, E. (2012). From Social Sensor Data to Collective Human Behaviour Patterns: Analysing Spatio-Temporal Dynamics in Urban Environments. GI-Forum: Geovisualization, Society and Learning.

[Scellato et al, 2010] Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010, June). Distance matters: geo-social metrics for online social networks. In Proceedings of the 3rd conference on Online social networks (pp. 8-8). USENIX Association.

[Silva et al, 2012] Silva, T. H., Melo, P. O., Almeida, J. M., Salles, J., & Loureiro, A. A. (2012, November). Visualizing the invisible image of cities. In Green Computing and Communications (GreenCom), 2012 IEEE International Conference on (pp. 382-389)

[Stodder, 2013] Stodder, D. (2013). Data Visualization and Discovery for Better Business Decisions. TDWI Best Practices Report. TDWI Research

[Villatoro et al, 2013] Villatoro, D., Serna, J., Rodríguez, V., & Torrent-Moreno, M. (2013). The TweetBeat of the City: Microblogging for Discovering Behavioural Patterns during the MWC2012. Citizen in Sensor Networks (pp. 43-56). Springer Berlin

[Wakamiya et al, 2011] Wakamiya, S., Lee, R., & Sumiya, K. (2011, November). Crowd-based urban characterization: extracting crowd behavioral patterns in urban areas from twitter. ACM SIGSPATIAL Workshop on LBSNs (pp. 77-84)

[Wakamiya et al, 2012] Wakamiya, S., Lee, R., & Sumiya, K. Measuring Crowd-sourced Cognitive Distance between Urban Clusters with Twitter for Socio-cognitive Map Generation. International Conference On Emerging Databases

[Watanabe et al, 2011] Watanabe, K., Ochi, M., Okabe, M., & Onai, R. (2011, October). Jasmine: a real-time local-event detection system based on geolocation information of microblogs. ACM conference on Information and knowledge management

[Zhang et al, 2011] Zhang, D., Guo, B., & Yu, Z. (2011). The emergence of social and community intelligence. Computer, 44(7), 21-28.

Le nuove basi geografiche di supporto alla pianificazione ed alla gestione territoriale

Gabriele GARNERO

Università degli Studi e Politecnico di Torino

DIST - Dipartimento Interateneo di Scienze, Progetto e Politiche del Territorio

V. Leonardo da Vinci, 44 - 10095 GRUGLIASCO (TO) - ITALY

gabriele.garnero@unito.it

ORCID ID: 0000-0003-3964-8798

Abstract. *The techniques and modes through which territorial data bases are currently produced and, potentially, updated prove to be reliable and accurate tools for the operators as much as they allow to set up diverse operating modes, which help develop territorial planning/management processes.*

For the purposes of this paper, the Author will analyze the peculiar aspects in both territorial data production procedures and data validation rules. Finally, some application experiences are provided showing how the available supports can be employed in new scenarios, which allows to gain interesting perspectives within the framework of territorial planning/management processes

Keywords: GIS, Viewshed, Landscape.

1. Introduzione

Le tecniche e le modalità attraverso le quali vengono oggi prodotte e, potenzialmente, aggiornate le basi dati territoriali, mettono a disposizione degli operatori non solo strumenti di conoscenza affidabili ed accurati, ma consentono di impostare differenti modalità operative, attraverso le quali fondare i processi di pianificazione e gestione del territorio.

Nella presente memoria, l'Autore analizza sia gli aspetti di produzione del dato territoriale sia le regole attraverso le quali il dato stesso è validato.

Si forniscono infine esperienze applicative, nelle quali i supporti disponibili vengono utilizzati in nuovi scenari, che consentono di prevedere interessanti prospettive nell'ambito dei processi di pianificazione e gestione del territorio.

2. I dati cartografici di base

Le produzioni cartografiche oggi in atto da parte di soggetti differenti (amministrazioni centrali, enti locali,) sono caratterizzate da una

impostazione basata su una struttura di dati spaziali multi-scala costituita da un insieme di oggetti ciascuno dei quali è caratterizzato da:

- codice identificativo univoco;
- geometria 3D (punto, polilinea, area) georeferenziata in un dato sistema di riferimento cartografico (UTM/WGS84);
- attributi alfanumerici (tabelle).

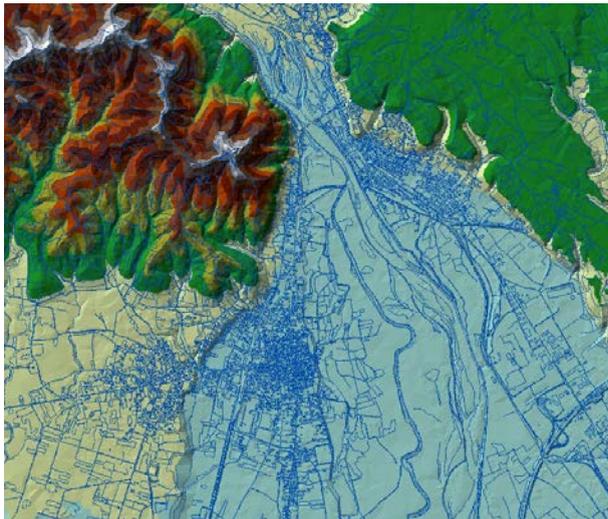
L'introduzione dei DB Topografici non è dunque solo il concretizzarsi di una codifica condivisa per la Cartografia Numerica e di una serie di regole per la sua strutturazione, ma rappresenta l'impostazione di un solido concetto di dato spaziale.

L'utilizzo dei DB Territoriali diventa pertanto uno strumento il cui uso parte dalla pianificazione urbanistica per costituire una base dati geografica aggiornabile per la costruzione dei SIT per la gestione del territorio e delle risorse ambientali.

2.1 Il contesto normativo

A livello nazionale, le attività di coordinamento delle produzioni cartografiche sono state portate avanti nell'ambito dell'Intesa tra Stato, Regioni ed Enti Locali sui Sistemi Informativi Geografici (*IntesaGIS*), stipulata nel 1996: questo progetto rappresentava all'epoca il tentativo più organico di modificare in termini positivi la situazione dell'Informazione Geografica in Italia, per creare uno stimolo verso una partecipazione più ampia delle istituzioni, delle imprese e del mondo scientifico.

Obiettivo dell'Intesa era lo sviluppo di interventi coordinati per realizzare basi informative territoriali informatizzate a copertura dell'intero territorio nazionale necessarie per l'esercizio delle funzioni di interesse locale, regionale e nazionale.



**Fig. 1: produzione prototipale con Specifiche IntesaGIS – Regione Piemonte (2005)
Sezione 094130 (Gattinara)**

Le nuove basi geografiche di supporto alla pianificazione ed alla gestione territoriale

Per la realizzazione dell'Intesa e degli Accordi ad essa collegati era stato costituito un Comitato Tecnico di Coordinamento per l'elaborazione degli indirizzi, delle specifiche comuni e per la definizione, nelle diverse aree del Paese, di accordi di programma tra gli Enti interessati alla realizzazione delle basi informative territoriali, a partire dai Database topografici e dal collegamento di questi con archivi catastali aggiornati.

Le attività dell'Intesa e, in seguito, del Comitato nazionale GIS-CNIPA, portarono alla produzione delle Specifiche comuni, adottate in seguito da varie produzioni prototipali e di base:

- specifiche per il DB Topografico (alle grandi scale);
- specifiche per il DBPrior10k;
- specifiche per le reti di raffittimento dell'IGM95;
- specifiche per i DTM;
- specifiche per l'utilizzo dei dati catastali.

Il successivo passo a livello nazionale fu rappresentato dall'adozione del Codice dell'Amministrazione Digitale (CAD), con D. L. 7 marzo 2005, n. 82, a mezzo del quale:

- lo Stato, le Regioni e le autonomie locali assicurano la disponibilità, la gestione, l'accesso, la trasmissione, la conservazione e la fruibilità dell'informazione in modalità digitale (Art. 2);
- i dati delle pubbliche amministrazioni sono formati, raccolti, conservati, resi disponibili e accessibili con l'uso delle tecnologie dell'informazione e della comunicazione che ne consentano la fruizione e riutilizzazione, alle condizioni fissate dall'ordinamento, da parte delle altre pubbliche amministrazioni e dai privati (Art. 50);
- qualunque dato trattato da una pubblica amministrazione è utilizzabile da un'altra pubblica amministrazione nei limiti dell'esercizio delle proprie funzioni (Art. 50);
- al fine di rendere possibile l'utilizzo in via telematica dei dati di una pubblica amministrazione da parte dei sistemi informatici di altre amministrazioni, l'amministrazione titolare dei dati costruisce, gestisce ed eroga i servizi informatici allo scopo necessari, secondo le regole tecniche del sistema pubblico di connettività (Art. 50).

Lo stesso CAD, all'Art. 59, definisce il concetto di dato territoriale come qualunque informazione geograficamente localizzata; definisce altresì il concetto di base dati di interesse nazionale (Art. 60) come l'insieme delle informazioni raccolte e gestite digitalmente dalle pubbliche amministrazioni, omogenee per tipologia e contenuto e la cui conoscenza è utilizzabile dalle pubbliche amministrazioni per l'esercizio delle proprie funzioni.

A livello europeo opera invece la Direttiva INSPIRE (*Infrastructure for SPatial InfoRmation in Europe*), che istituisce un'infrastruttura per l'informazione territoriale nella Comunità europea.

La Direttiva è entrata in vigore il 15 maggio 2007 ed intende creare, grazie a norme comuni di attuazione integrate da misure comunitarie, una struttura comune che renda l'informazione territoriale dei vari Stati compatibile e utilizzabile in un contesto transfrontaliero, in modo da superare i problemi

riguardo alla disponibilità, alla qualità, all'organizzazione e all'accessibilità dei dati.

Questi, in sintesi, gli aspetti più importanti della Direttiva:

- INSPIRE si basa sulle infrastrutture per l'informazione territoriale create dagli Stati membri: a tal fine l'infrastruttura deve essere stabilita e resa operativa dai singoli Stati, che devono garantire che i dati territoriali siano archiviati, resi disponibili e conservati al livello più idoneo, al fine di evitare duplicazioni di dati: questi vanno raccolti una sola volta e gestiti laddove ciò può essere fatto in maniera più efficiente. Non è richiesta la raccolta di nuovi dati spaziali, ma qualsiasi dato territoriale dovrà adeguarsi alle indicazioni della Direttiva;
- l'interesse principale della Direttiva è rivolto alle politiche ambientali comunitarie e alle politiche o alle attività che possono avere ripercussioni sull'ambiente. Quando sarà pienamente operativa permetterà di combinare dati transfrontalieri da uno Stato membro all'altro con continuità e condividerli con le applicazioni e tra gli utilizzatori;
- la Direttiva mira ad agevolare la ricerca dei dati spaziali attraverso il *web*, tramite servizi di rete che ne permettano l'utilizzo in molteplici modi, dalla visualizzazione, al *downloading*, alle varie trasformazioni. I dati devono essere facilmente individuabili e adatti ad un uso specifico, facili da comprendere ed interpretare.

Al fine di sviluppare e rendere attinenti i dati e i servizi, la Direttiva deve tenere in conto alcuni temi ad essa correlati come la salvaguardia della *privacy* e i diritti di proprietà intellettuale, ma allo stesso tempo lascia libertà ai singoli Stati membri di stabilire un compenso per i propri set di dati o servizi offerti.

2.2 Gli Open Data

A partire dalle considerazioni e dai disposti normativi cui si è fatto cenno, oggi il contesto e il dibattito culturale si sono spostati sul concetto di *Open Data*.

Gli americani per primi hanno posto regole esplicite e trasparenti in materia: nel dicembre 2009 Barack Obama, al suo primo mandato alla Casa Bianca, emana la direttiva sull'*Open Government* che definisce i dati e i formati aperti, i servizi di *download* dal *web*, di riuso . Il Governo inaugura anche il portale data.gov dedicato all'*Open Data*, sul quale risulta subito evidente che i dati più richiesti e scaricati sono quelli geografici.

In Italia alcune Regioni rendono disponibili già da qualche anno dati geografici attraverso servizi di consultazione, interoperabilità e *download*: la prima Regione ad adottare una licenza realmente *open* è stata il Piemonte, che dal maggio 2010 offre dati geografici e non sul portale www.dati.piemonte.it.

Ma è con la L. 221/2012 (conversione del Decreto Crescita 2) che il Governo Italiano fa chiarezza sugli *Open Data*, istituendo contestualmente l'Agenzia per l'Italia Digitale con le finalità di coordinare e monitorare l'operato della Pubblica Amministrazione.

Un altro importante aspetto della Legge definisce il cosiddetto *Open by Default*: “ i dati e i documenti che le amministrazioni titolari pubblicano, con

Le nuove basi geografiche di supporto alla pianificazione ed alla gestione territoriale *qualsiasi modalità, senza l'espressa adozione di una licenza di cui all'articolo 2, comma 1, lettera h), del decreto legislativo 24 gennaio 2006, n. 36, si intendono rilasciati come dati di tipo aperto ai sensi all'articolo 68, comma 3, del presente Codice.* Questa definizione inverte di fatto il principio in uso nel diritto d'autore italiano che affermava che, salvo diversa indicazione, tutti i diritti sono riservati.

3. Le Specifiche cartografiche

Non è nei limiti della presente memoria l'entrare nel dettaglio dei contenuti tecnici delle specifiche di produzione previste alle varie scale e per i diversi prodotti cartografici.

Sono ora vigenti, a livello normativo, i quattro decreti del 10 novembre 2011 emanati dal Ministro per la Pubblica amministrazione e l'innovazione di concerto con il Ministro dell'Ambiente e della tutela del territorio e del mare, pubblicati sulla Gazzetta Ufficiale n. 48 del 27 febbraio 2012 (Suppl.Ord. n. 37).

Attraverso queste norme, a completamento dell'iter previsto dall'articolo 59 comma 5 del CAD, sono state adottate le prime specifiche definite dal Comitato per le regole tecniche sui dati territoriali delle pubbliche amministrazioni. In particolare, i quattro provvedimenti in questione riguardano, rispettivamente:

- Adozione del Sistema di riferimento geodetico nazionale;
- Regole tecniche per la definizione delle specifiche di contenuto dei database geotopografici;
- Regole tecniche per la definizione del contenuto del Repertorio nazionale dei dati territoriali, nonché delle modalità di prima costituzione e di aggiornamento dello stesso;
- Regole tecniche per la formazione, la documentazione e lo scambio di ortofoto digitali alla scala nominale 1:10000.

Vengono nel seguito unicamente descritti alcuni aspetti peculiari che caratterizzano le moderne produzioni.

3.1 Le Specifiche per i DB Territoriali

Come in tutte le carte numeriche, il concetto di scala è diverso da quello delle carte tradizionali, nelle quali ha un significato fisico (rapporto medio tra le distanze sul foglio di carta e quelle misurate nella realtà); in cartografia numerica gli oggetti sono rappresentati nel sistema cartografico di rilievo, non in scala. La precisione e la risoluzione dei dati dipendono tuttavia dalla cosiddetta scala nominale, cioè sono mutuati da quelli di una corrispondente cartografia cartacea.

Nei DB Territoriali ciascuna zona viene rilevata una sola volta alla scala più opportuna (anche in funzione del *budget* disponibile): il risultato è un DB multiscale.

La scala nominale determina la precisione richiesta nel rilievo, e quindi la tolleranza geometrica che gli oggetti devono garantire nel DB, formalizzando un concetto di Tolleranza inteso come massima discrepanza tra il valore "vero" di una grandezza e quello riportato sulla carta (nel DB). La Tolleranza dev'essere rispettata statisticamente dal 95% degli oggetti.

Tolleranza 2D/H	Tipo di Tolleranza	DBT 1k	DBT 2k	DBT 5k	DBT 10k
2D	su tutti i punti	0.60	1.20	3.00	6.00
H	su tutti i punti	0.60	0.80	2.00	4.00
2D/H	sulle medie	± 0.10 $H=\pm 0.15$	± 0.20	± 0.50	± 1.00
	sugli sqm delle medie	± 0.20 $H=\pm 0.30$	± 0.40	± 1.00	± 2.00

Tab. 1: valori delle Tolleranze previste (da produzioni Regione Lombardia) [m]

Viene inoltre definito un limite di cattura, che esprime la risoluzione degli oggetti del DB, formalizzando quali elementi devono essere rappresentati ad una certa scala nominale.

Il valore del limite di cattura è posto pari a $\frac{1}{2}$ della Tolleranza planimetrica alla scala.

Se un oggetto è inferiore al limite di cattura va rilevato soltanto se è ritenuto importante, ed in questo caso viene rappresentato in modalità “collassata”, con una dimensione fittizia pari a $\frac{1}{4}$ della Tolleranza.

Il Catalogo individua i Dati Territoriali organizzati in Strati, Temi e Classi: la struttura di riferimento è la Classe, che definisce la tipologia degli oggetti territoriali:

- proprietà;
- struttura del dato;
- regole di acquisizione, strutturazione e relazione tra gli oggetti.

Gli Strati e i Temi non rappresentano una classificazione, ma raccolgono le Classi in sottoinsiemi morfologicamente o funzionalmente omogenei.

Per ogni Classe è definito l'insieme degli attributi, e per ogni attributo è definito il tipo di dato e i limiti del valore.

Il principale elemento di novità, che ha profondamente inciso e sta modificando il modo di intendere la produzione cartografica dell'intero comparto fotogrammetrico nazionale, è rappresentato dalla formalizzazione dei vincoli d'integrità, cioè delle proprietà intrinseche che gli elementi rappresentati devono soddisfare (s'intendono le proprietà verificabili sugli elementi rappresentati stessi, senza osservazioni dirette del mondo reale).

In sostanza viene introdotta una sorta di “*intelligenza nel dato*”, capace di validare se stesso e le possibili interazioni con gli elementi circostanti.

I parametri di qualità presi in esame sono indicati nel documento 1n1014:

- **Congruenza logica:** riguarda la valutazione delle caratteristiche di contenuto del DB sia per il rispetto dello schema fisico di formato sia per la strutturazione concettuale della componente spaziale:
 - di formato
 - di dominio

Le nuove basi geografiche di supporto alla pianificazione ed alla gestione territoriale

- di geometria
- di topologia
- **Accuratezza posizionale:** si valuta lo scostamento delle coordinate dalla reale posizione sul terreno rispetto alla tolleranza indicata;
- **Completezza:** fornisce l'attendibilità della sola presenza/assenza di un determinato oggetto topografico nel DBT:
 - eccesso
 - omissione
- **Accuratezza tematica:** valuta la correttezza della codifica degli oggetti del DB e la completa informazione di attribuzione:
 - di classificazione
 - toponomastica

La conformità di un insieme di dati alle specifiche indica come e cosa deve esserci nel dato, nelle basi geografiche questa conformità è declinata secondo due aspetti:

- **Conformità reale:** la conformità reale riguarda la corrispondenza tra il contenuto informativo e la porzione di mondo reale alla quale il dato si riferisce;
- **Conformità intrinseca:** la conformità intrinseca riguarda la consistenza dell'informazione contenuta nel dato secondo le specifiche.

3.2 Le Specifiche per i DTM

Relativamente ai DTM (*Digital Terrain Model*), le principali indicazioni formalizzate nelle Specifiche sono relative ai seguenti aspetti:

- il cambiamento di tendenza è rappresentato dal fatto che il principale prodotto relativamente all'altimetria è ora rappresentato dal DTM, mentre le curve di livello assumono unicamente una funzione di rappresentazione cartografica, derivata dal modello digitale stesso e finalizzato all'osservazione da parte dell'utente: per gli aspetti legati alle elaborazioni si privilegia l'utilizzo dei DTM;
- le Specifiche definiscono una serie di requisiti qualitativi dal punto di vista della precisione cui devono soddisfare i DTM, in particolare istituendo una serie di differenti Livelli, caratterizzati ciascuno dal punto di vista della precisione e della risoluzione di griglia;
- vengono definite le specifiche per la produzione, tra le quali:
 - ordinariamente è prevista la produzione di un TIN (*Triangulated Irregular Network*) da cui ottenere il grigliato regolare del DTM per interpolazione;
 - per la produzione dei modelli digitali è necessario impiegare tutte le informazioni disponibili riconducibili al suolo, quindi tutti gli elementi che costituiscono la planimetria delle rappresentazioni cartografiche, ristretta ai soli elementi la cui quota è riferita al suolo;
 - per la generazione del modello digitale è necessario integrare con punti (*mass points*) e linee di discontinuità (*breaklines*) rilevati unicamente per la produzione del DTM (senza valenza cartografica). Per la misura dei punti isolati è auspicabile

utilizzare le metodologie della fotogrammetria digitale che prevedono l'uso dell'autocorrelazione ovvero le tecniche LiDAR, a seconda del livello che ci si propone di ottenere;

Dall'ultima versione del documento CISIS "Ortoimmagini e modelli altimetrici a grande scala - Linee Guida", in Tab 2 si riportano i valori dei Livelli di maggior diffusione (valori in metri):

	Livello			
	1	2	3	4
Tipologia	DEM o DSM	DEM o DSM	DEM o DSM	DEM o DSM
Accuratezza in quota: in campo aperto <i>PH(a)</i>	5	2	1	0.30
Accuratezza in quota: con copertura arborea > 70% <i>PH(b)</i> (nel caso di DEM)	10	1/4 altezza media alberi	1/4 altezza media alberi	0.60
Accuratezza in quota: edifici (nel caso di DSM) <i>PH(c)</i>	5	2.50	1.50	0.40
Tolleranza in quota: in campo aperto <i>TH(a)</i>	10	4	2	0.60
Tolleranza in quota: con copertura arborea > 70% <i>TH(b)</i> (nel caso di DEM)	20	1/2 altezza media alberi	1/2 altezza media alberi	1.20
Tolleranza in quota: edifici (nel caso di DSM) <i>TH(c)</i>	10	5	3	0.80
Accuratezza planimetrica: <i>PEN</i>	5	2	1	0.30
Tolleranza planimetrica: <i>TEN</i>	10	4	2	0.60
Passo di griglia:	20	20	10	5

Tab. 2: Valori caratteristici dei principali Livelli per DTM e DSM [m]

4. Nuove rappresentazioni: analisi della sensibilità visiva del paesaggio

Le basi dati rese disponibili consentono di contribuire alla costruzione di nuove rappresentazioni, che possono fornire valori aggiunti sia alle più tradizionali attività di gestione (SIT comunali, gestione delle utility, gestione della strumentazione urbanistica, incrocio con le basi catastali,), sia a moderne modellazioni territoriali non altrimenti possibili.

Tra le varie esemplificazioni possibili, si ritiene opportuno affrontare una problematica relativamente originale, oggetto di recenti sperimentazioni da parte del gruppo di ricerca cui l'Autore appartiene, focalizzata al controllo della qualità estetico-percettiva del paesaggio attraverso un approccio "quantitativo" basato sull'uso dei SIT.

Le aree maggiormente visibili del territorio possono essere individuate in modo automatico ed informatizzato: le analisi dei bacini visuali (*Viewshed Analysis*) consentono di ottenere una simulazione complessa delle relazioni tra morfologia del paesaggio e punti di osservazione: è infatti una tecnica di analisi spaziale che utilizza gli algoritmi delle *lines of sight* per determinare la visibilità di aree da un determinato punto di osservazione del territorio.

La tecnica consiste nel calcolare il campo di osservazione (*bacino visuale*) rispetto alla posizione e all'orizzonte visivo di un osservatore: sulla base di un

Le nuove basi geografiche di supporto alla pianificazione ed alla gestione territoriale DTM è infatti possibile determinare la visibilità relativa da punti di vista predeterminati (oppure da una successione di punti, come per i percorsi) per ogni cella in cui è discretizzata l'area di studio. Il prodotto risultante di tale analisi è un'immagine raster il cui contenuto informativo dipende dal particolare modello di visibilità adottato (*binary viewshed*, *cumulative viewshed*, *identifying viewshed*, ecc.).

Le indagini cui fa riferimento la figura che segue sono state condotte relativamente allo studio per la progettazione paesistica della “*Strada del Papa*”, percorso devozionale incentrato sulla spiritualità salesiana che collega la casa natale di Don Bosco (il Santo dei giovani) con quella di Domenico Savio (il primo giovane Santo), nel territorio di Castelnuovo Don Bosco (AT).

Attraverso la funzione *Viewshed* nell'ambiente ArcGIS 10 è possibile ottenere un'immagine raster che rappresenta la visibilità a partire da un determinato punto di osservazione: la *viewshed analysis* consente di ottenere un'immagine raster in cui il valore di ogni cella può essere 0 (non visibile) o 1 (visibile) e che rappresenta il bacino visivo dal punto prescelto.

Ottenuta un'immagine per ogni punto di osservazione o, come nel caso in oggetto, a partire da un percorso, è possibile effettuare un *overlay* tra i diversi risultati ed ottenere una nuova elaborazione raster, che mette in risalto la “visibilità assoluta” del paesaggio dall'insieme dei punti di vista.

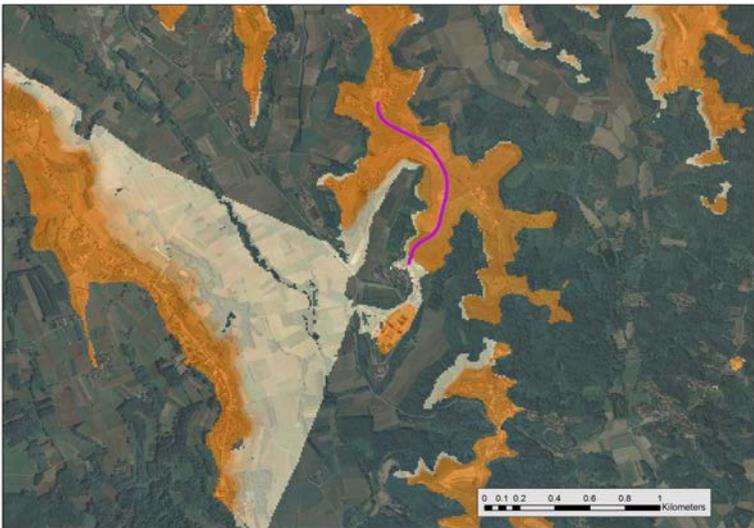


Fig. 2: carta della sensibilità paesistica della “Strada del Papa”

Per costruire una carta della sensibilità visiva del paesaggio bisogna suddividere in classi il numero di sovrapposizioni, per sapere quali aree sono più visibili rispetto alle altre: tecnicamente, è necessario operare sulla tabella degli attributi dell'ultima elaborazione, effettuando un *summarize*, che incorporerà il numero delle sovrapposizioni tra *viewshed* diverse per ogni cella.

Tali aree sono quindi interpretate come quelle che presentano maggiore sensibilità visiva rispetto alle trasformazioni territoriali, ossia l'impatto della trasformazione è potenzialmente maggiore in tali aree, perché visibili da più

punti: l'analisi è dunque utile per ragionare sui criteri localizzativi di nuovi interventi, soprattutto di quelli potenzialmente impattanti, diventando così strumento di estremo interesse per il progetto urbanistico.

5. Conclusioni

L'esempio riportato non è che uno degli ambiti di utilizzo delle moderne basi di rappresentazione, valido per comprendere come sia possibile ipotizzare analisi ed interrogazioni che solo i moderni supporti possono consentire.

Bibliografia

Biallo, G.: (2013). Dati Geografici Aperti: istruzioni per l'uso, e-book disponibile al link http://www.opengeodata.it/index.php?option=com_jdownloads&Itemid=0&view=finish&cid=4&catid=3&m=0

Cassatella C. (2011): Analisi scenico-percettiva del paesaggio. In: Fare paesaggio : dalla pianificazione di area vasta all'operatività locale / In A. Peano (cur.). Alinea, Firenze, pp. 225-228. ISBN 9788860556127

Cassatella C. (2012): Lo sguardo sulla città / Townscapes. In: Paesaggio e bellezza / Enjoy the Landscape / Cassatella C., Bagliani F. CELID, Torino, pp. 92-105. ISBN 9788876619557

Chiabrando, R., Fabrizio, E., Garnero, G. (2011): On the applicability of the visual impact assessment OAISPP tool to photovoltaic plants, Renewable and Sustainable Energy Reviews, 15 (1), pp. 845-850, DOI: 10.1016/j.rser.2010.09.030

Garnero, G.; Corrias, A.; Manigas, L.; Zedda, S.V. (2013): VGI, Augmented Reality and Smart Web Application: Projects of Development in the Territory of the Sardinia Region, in B. Murgante et al. (Eds.): ICCSA 2013, Part IV, LNCS 7974, pp. 77-92. Springer, Heidelberg

Fabrizio, E.; Garnero, G. (2013/IN PRESS): The assessment of the visual perception in viewshed analysis for the landscape settings, Atti 10th AIIA Conference, Viterbo

Ferrante, F.; Garnero, G. (2013): Tecnologie e dati spaziali per una moderna governance del territorio. Strumenti a supporto della riforma del Catasto (Technologies and spatial data for modern land governance. Tools to support the cadastral reform.), Territorio Italia - Governo del Territorio, Catasto, Mercato immobiliare, n. 01/2013, p. 9-26 (Italian and English version)

Godone, D., Garnero, G. (2013): The role of morphometric parameters in Digital Terrain Models interpolation accuracy: A case study, European Journal of Remote Sensing, 46 (1), pp. 198-214, DOI: 10.5721/EuJRS20134611

Minucciani, V.; Garnero, G. (2013): Available and Implementable Technologies for Virtual Tourism: A Prototypal Station Project, in B. Murgante et al. (Eds.): ICCSA 2013, Part IV, LNCS 7974, pp. 193-204. Springer, Heidelberg

Normativa INSPIRE: disponibile on-line al link http://eur-lex.europa.eu/LexUriServ/site/it/oj/2007/l_108/l_10820070425it00010014.pdf

Comprehending geographic scenarios through visual summaries

D. De Chiara, V. Del Fatto¹, M. Sebillo, G. Vitiello
University of Salerno
via Giovanni Paolo II 132, Fisciano (Salerno), Italy
ddechiera@unisa.it, msebillo@unisa.it, gvitiello@unisa.it
¹ Free University of Bolzano
Piazza Domenicani 3, Bolzano-Bozen
vincenzo.delfatto@unibz.it

Abstract *In the last decade, the need to support decision makers in solving problems related to a territory and its phenomena has stimulated Geographic Information Visualization (GeoVis) researchers to propose highly interactive visualization tools able to both synthesize information from large datasets and perform complex analytical tasks. The goal of the present paper is to propose two GeoVis methods based on a recent InfoVis technique, known as Chorem and Tag Cloud, which combines advanced GeoVis techniques for visualizing and analyzing geographic data and related spatio-temporal phenomena. The methods elaborate a simplified maps in order to help to identify significant geospatial information, data, and knowledge by supporting analytical processes that join innate human abilities of vision and cognition with computer-based visual interfaces. As initial result a system prototypes have been realized in order to obtain an overview of large dataset. They are focused on data extraction and aggregation, and output visualization, and adopt various techniques to allow users to select data to visualize starting from a geographic dataset.*

Keywords: Geographic Information Visualization, GeoVisual Analytics, GeoVisualization Techniques, Interactive Visual Environments for Accessing and Analyzing Geospatial Information.

1.Introduction

Information Visualization, also known as InfoVis, is recognised as the scientific area where interactive visual representations of abstract data aim to amplify user's cognition of information [1]. In particular, recent InfoVis research has been targeted to support humans in solving problems through highly

Congresso Nazionale AICA 2013

interactive visualization tools able to both synthesize information from large datasets and perform complex analytical tasks [2]. A similar need has also stimulated researchers from the Geographic Information (GI) field, who are proposing advanced tools able to semantically integrate quantitative, qualitative and cognitive aspects of a domain of interest, and generate visual information useful to provide an immediate overview of it [3]. GeoVisualization (GeoVis, for short) represents the research field where the efforts from both these two disciplines are integrated in order to define techniques and metaphors, such as maps and diagrams, capable to represent geographic data and improve their abstraction and interactivity. Due to pervasiveness of geographic component, applications from several domains can be labelled as GeoVis applications and strong relationships among them can be established, which allow designers to export advanced solutions from apparently different domains and extend them by properly taking into account the geographic component.

The goal of the present paper is to propose two GeoVis methods based on a recent InfoVis techniques, known as Chorems and Tag Cloud, which combine maps with advanced GeoVis techniques for visualizing large geographic data and related spatio-temporal phenomena.

Chorems represent a schematic territory representation which eliminates details not useful to the map comprehension [9]. An example of such a visual metaphor is given in Figure 1, where three maps containing chorems referring to the island of Corse are depicted. In particular, Figure 1(a) shows the transhumance between summer and winter pastures, Figure 1(b) shows how the wind influences ecosystems and Figure 1(c) shows agricultural specialities.

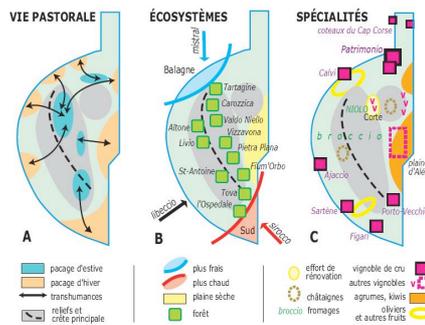


Fig. 1. Map containing chorems about the island of Corse [9].

A Tag Cloud is a visual representation typical for text data, used to depict keyword metadata (tags) whose importance is expressed by font size or colour [4]. The proposed methods adopt the Chorem and Tag Cloud rationale and extends it by exploiting techniques for summarizing datasets and simplifying their geographic representation. Starting from a geographic dataset, the proposed methods extract relevant information about a geographic area by counting and/or summarizing data, and generate simplified maps that provide useful support in locating facts and new patterns. The goal of our future work is to realize interactive maps by which users may perform visual analytics tasks.

The paper is organized as follows. Section "Background" recalls the preliminaries about Visual Summaries and Visualization methods.

Section 3 introduces visual analysis through maps of chorems. Section 4 introduces the visualization of data summaries by georeferenced Clouds of Tags. Conclusions and future work are drawn in Section "Final Remarks".

2. Background

This section introduces the context of the paper, which is based on several topics related to map visualization. In particular, two cartographic solutions for visual summaries are briefly recalled useful to form the basis for the present proposals.

Cartograms [5] are maps where a particular distortion is introduced to express a specific map attribute. According to the definition given in [7], a cartogram is a small diagram which shows quantitative information or an "abstracted and simplified map, the base of which is not true to scale". They are targeted to show the relative importance of a variable, and make easy visual comparisons, as illustrated in Figure 2, where the Cartogram of Mexico's population distribution is represented.



Fig. 2. A Cartogram showing Mexico's population distribution.

While cartograms show values of a single variable at a time, chorems allow designers to assemble into a single map more than one thematic layer, thus representing the relative importance of a set of objects and phenomena related each others. According to the definition by Roger Brunet [10], a chorem is a schematic territory representation, which eliminates any detail not necessary to the map comprehension. Figure 1(b) shows a set of chorems related to the island of Corse, representing the transhumance between summer and winter pastures. This capability of chorems may be exploited to provide decision makers with a means to acquire syntactic information, as well as semantic aspects, useful to support human activity to model, interpret and analyse the reality of interest [11, 12].

Tags are usually single words whose importance is visually expressed by font size or colour. A tag cloud is a visual representation typical for text data, used to depict keyword metadata (tags) on websites, or to visualize free form

Comprehending geographic scenarios through visual summaries GeoVis techniques. In particular the techniques exploit the GeoVis capability of summarizing datasets and simplifying their geographic representation also by altering their original shape. Chorems technique represent a schematic territory representation which eliminates details not useful to the map comprehension. The potential of chorems has been initially discussed by Laurini et al. in [6], who listed different roles that chorems may play in supporting expert users in daily activities. The authors argued that chorems can be used to represent geographic knowledge, to visually summarize database contents and finally to underlie the creation of a novel entry system for geographic databases. In spite of their wide diffusion, no standard structure was available at the time the international project started.

In order to bridge this gap, a definition and a classification of chorems as visual syntheses of geographic database content were first provided in [23], to homogenize chorem construction and usage and to provide a usable framework for computer systems. In particular, the prototype for chorem map generation and results obtained from it were encouraging and represented the starting point for further processing tasks aiming at deriving spatial analysis data, and supporting expert users in decision-making, thus helping them face the complexity of the adopted applications and get rapid and exhaustive responses in critical situations. In fact, as pointed out by Andrienko et al., in many activities, it is essential to support the analyst with means which summarize the outcomes of his work and help him present them in a way that the ultimate users may easily understand the derived knowledge. With this observation in mind, our approach is then meant to support those people, who may need to take rapid decisions even with little expertise in spatial analysis. In particular, the goal of the current research has been to enhance the role that a chorem map may play in geographic domains, by exploiting its synthesis capability also for querying and accessing data of interest when they are visually summarized as a set of chorems on a map. A preliminary idea has been introduced in [20] where we have proposed to exploit the salient aspects of chorems also to understand the causes of a given phenomenon. Indeed, the strong simplification and the expressive enough summary may be used either to catch a thematic global view of a territory and its phenomena or to investigate and analyze such phenomena by accessing data characterizing them.

To develop this idea, we have first extended the semantics associated with the chorem concept by including information about its construction process in terms of initial data source and intermediate steps. This allows us to investigate phenomena of interest by visually manipulating chorems and applying a set of given spatial operators on them. In particular, we have adopted the Visual Information-Seeking mantra stated by Ben Shneiderman, namely Overview, Zoom and Filter, Details on Demand [21] and the Visual Analytics Mantra presented in [22] Analyze First - Show the Important - Zoom, Filter and Analyze Further - Details on Demand, and assigned these operations a specific behavior. This approach then combines GeoVisualization and interaction techniques to represent and examine well identified spatio-temporal phenomena in an innovative manner. Users may acquire information from the underlying database by interacting with maps of chorems which visually summarize its content. Each interaction task assumes a context-sensitive

meaning and invokes a proper functionality among the ones specified in agreement with the Visual Information-Seeking mantra.

A point of distinction of the proposed GeoVisualization system against existing solutions is that the adopted visualization technique relies on the joint representation of a simplified map and an expressive synthesis of spatio-temporal phenomena of interest.

Finally, we have also introduced an extension of the ChorML language, initially introduced to store information about chorems. The enhanced version we propose is the basis for the prototype where the operators have been experimented. In particular, they have been used to exemplify some visual analytics tasks useful to illustrate how the chorem-based approach guarantees simple and immediate solutions for analysts and ultimate users. The idea is based on the “Overview, zoom and filter, details on demand” approach, i.e. macroscopic versus microscopic. In agreement with this approach, we assign each interaction task with a specific meaning, aiming at satisfying each type of user’s requirement.

In figure 5 we illustrate the image of the operations we propose, meant to allow users to interact with a chorematic map in order to perform visual analytics tasks.

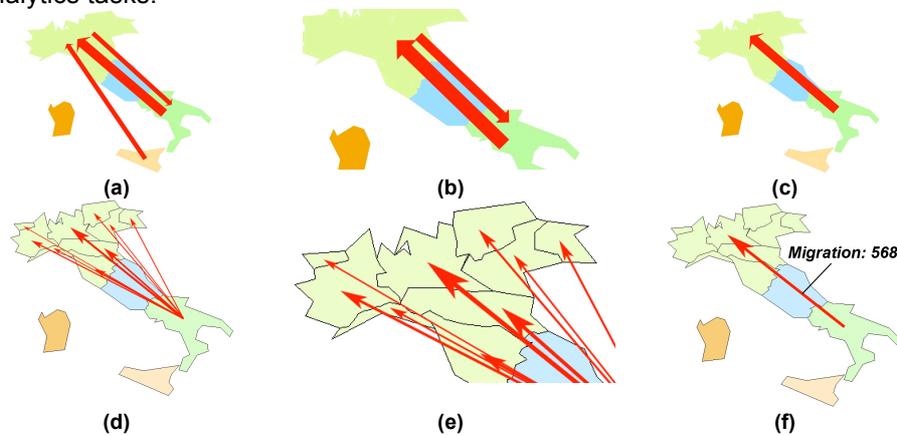


Figure 5 (a) A chorem map under investigation (b) The application of a geographic filter (c) The application of a semantic filter (d) The application of a semantic zoom (e) A geographic zoom (f) The output of the example

4. Analyzing geographic information through tag clouds

Tag cloud technique adopts tags to express concepts extracted from data sources, and finally, it adapts the georeferencing concept underlying the Tag Maps, by associating it with the whole cloud. The method application results in a simplified map containing a georeferenced cloud of tags [24], each referring to a different map characteristics. In particular, it counts and summarizes data from a geographic dataset and expresses them through tag clouds within the

boundary and the alphanumeric data to visualize. The *Data Aggregation and Generalization* module has been implemented to simplify the map boundary, as shown in Figure 7(a), by using the well-known Ramer-Douglas-Peucker (RDP) algorithm [18, 19]. It produces the simplified map in XML and SVG format. Moreover, it also allows the user to select the weight function useful to establish the tag dimension. The *Cloud Creation* module has been implemented as a PHP application to call a remote API (Tagul API [15]) via XML-RPC, which generates the tag cloud. The API takes an XML file as input parameter, containing the derived simplified map boundary and the calculated tag weight function, and returns an SVG file. As shown in Figure 7(b), the cloud assumes the shape of the map boundary. Finally, as shown in Figure 7(c), the *Data Representation* module has been implemented to merge the SVG version of the simplified map boundary and the tag cloud in order to derive the resulting map.

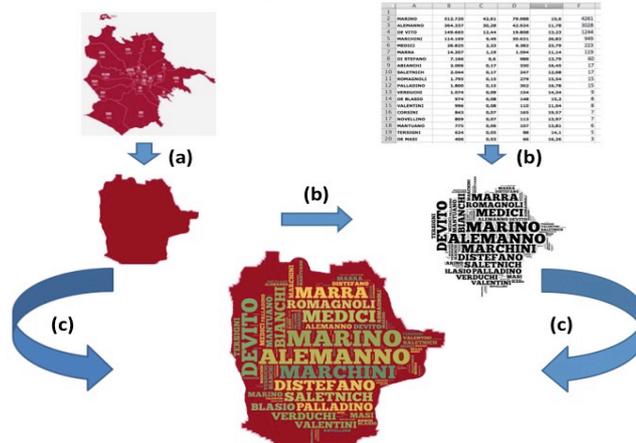


Fig. 7. The generation process of a tagged map representing the results of the elections for mayor of the city of Rome (Italy) 2013.

5.Final Remarks

The main goal of this dissertation has been to contribute to the GeoVisualization and GeoVisual Analytics research field by defining and realizing methods and techniques to improve human-(geo)information discourse. General speaking, contributing to the achievement of this goal means first to investigate potentialities of recent advances in diverse disciplines which have the territory as a common element. Indeed, as they analyze phenomena related to it by different point of views, GeoVisual Analytics researchers should understand how findings may be shared and integrated. A second relevant aspect concerns the combination of alternative graphic representations and geocomputational approach which can be exploited to stimulate the visual thought process, enabling new map forms used as visual thinking and decision-support.

These observations along with the twenty-year experience matured by computer science researchers from the University of Salerno within the visual language and multimedia database field has focused attention on the changing and expanding role of maps in science also useful to face general-purpose issues related to territory. In particular, recent literature demonstrates that studies conducted within the GeoVisualization domain have produced proficient working systems developed with a twofold goal, namely providing users with environments where users can easily view geographic data in many alternative views through visual geospatial displays, and hiding the inner complexity of geospatial data thus overcoming difficulties by unskilled users.

References

1. Card, S., Mackinlay, J., Shneiderman, B. (1999) *Readings in Information Visualization: Using Vision to Think*, Morgan Kaufmann, San Francisco, CA, USA.
2. Shneiderman, B., and Plaisant, C., (2010) *Designing the User Interface: Strategies for Effective Human-Computer Interaction (Fifth Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA .
3. Andrienko, G., Jern, M., Dykes, J., Fabrikant, S. I., Weaver, C. (2007) *Geovisualization and synergies from InfoVis and Visual Analytics in Proceedings of the 11th International Conference Information Visualization IV 2007*, IEEE Computer Society Press, pp. 485-488.
4. Viegas, F., and Wattenberg., M., (July 2008) *TIMELINES: Tag clouds and the Case for Vernacular Visualization*. *Interactions* 15, 4 , 49-52.
5. Dorling, D., (1994) *Cartograms for Human Geography*. *Visualization in Geographical Information Systems*, pages 85-102.
6. K. Lopez, R. Laurini, V. Del Fatto, D. Sol, R. Loreto, M. Sebillio, G. Vitiello "Visualizing Geographical Analysis Results From Spatial Databases Based on the Chorems", *Proceedings of The 2009 International Conference on Modeling, Simulation and Visualization Methods MSV'09*, (joint conference in *WORLDCOMP'09 - The 2009 World Congress in Computer Science, Computer Engineering, and Applied Computing*). Las Vegas, USA, July 13-16, 2009. H. R. Arabnia, L. Deligiannidis (Eds.) Pages 29-35 CSREA Press 2009, ISBN 1-60132-120-17. ASCE. (1994) *The Glossary of the Mapping Sciences*. New York, ASCE.
8. Worldmapper. Worldmapper website, <http://www.worldmapper.org/index.html>.
9. Brunet, R. (2004) *La Corse, Région d'Europe*. *Mappemonde*. 4: 1-16.
10. Brunet., R. (1986) *La Carte-modele et les Choremes*. *Mappemonde*, 4:2-6.
11. V. Del Fatto, R. Laurini, K. Lopez, R. Loreto, F. Milleret-Raffort, M. Sebillio, D. Sol-Martinez, G. Vitiello "Potentialities of Chorems as visual summaries of spatial databases contents", *Proceedings of the 9th International Conference on Visual Information and Information Systems VISUAL2007*, Shanghai (CINA), 28 – 29 May 2007. G. Qiu et al. (Eds.): *VISUAL 2007*, LNCS 4781, pp. 537–548, 2007. © Springer-Verlag Berlin Heidelberg 2007.
12. De Chiara, D., Del Fatto, V., Laurini, R., Sebillio, M., Vitiello, G. (2011). *A Chorem-based Approach for Visually Analyzing Spatial Data*, in *Journal of Visual Languages & Computing (JVLC)*, Elsevier. Volume 22, Issue 3, Pages 173-193.

13. Viegas, F., B., Wattenberg, M., Feinberg, J. (2009). Participatory Visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics* 15, 6, 1137-1144.
14. Tagul website: <http://tagul.com/>
15. Tagul API website: <http://tagul.com/api>
16. Tagxedo website: <http://tagxedo.com>
17. Jaffe, A., Naaman, M., Tassa, T., Davis, M. (2006). Generating Summaries and Visualization for Large Collections of Geo-referenced Photographs. In *Proceedings of the 8th ACM international workshop on Multimedia information retrieval (MIR '06)*. ACM, New York, NY, USA, 89-98.
18. Douglas, D., H., and Peucker, T., K., (1973) Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. *Canadian Cartographer*, 10:112-122.
19. Ramer, U. (1972) An Iterative Procedure for the Polygonal Approximation of Plane Curves. *Computer Graphics and Image Processing*, 1(3):244-56.
20. De Chiara D., Del Fatto V., Laurini R., Sebillo M., Vitiello G., "Visual Analysis of Spatial Data through Maps of Chorems", in *Proceedings of the 15th International Conference on Distributed Multimedia Systems – Workshop on Visual Languages and Computing VLC2009*, San Francisco (USA), 10-12 September 2009. KSI Skokie Illinois, pp. 295-300.
21. Shneiderman B., The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pages 336-343 (1996).
22. Keim D.A., Mansmann F., Schneidewind J., Ziegler H., Challenges in Visual Data Analysis, *Proceedings of Information Visualization (IV 2006)*, IEEE, p. 9-16, (2006).
23. Del Fatto V., Laurini R., Lopez K., Sebillo M., Vitiello G., A chorem-based approach for visually synthesizing complex phenomena, in *Journal of Information Visualization*, vol. 7, n.3, pp. 253-264 (2008). G. Andrienko, N. Andrienko, J. Dykes, S. Fabrikant and M. Wachowicz (Eds). Palgrave Macmillan Journals.
24. D. De Chiara, V. Del Fatto, M. Sebillo "Visualizing Geographical Information Through Tag Clouds", in *Information Systems: Crossroads for Organization, Management, Accounting and Engineering*. M. De Marco, D. Te'eni, V. Albano, S. Za (Editors), pp. 209-216. © Springer Physica-Verlag Berlin Heidelberg 2012. ISBN: 978-3-7908-2788-0.
25. D. De Chiara, V. Del Fatto, M. Sebillo, G. Tortora, and G. Vitiello "Tag@Map: A Web-Based Application for Visually Analyzing Geographic Information through Georeferenced Tag Clouds" in *Proceedings of the 11th International Symposium on Web and Wireless Geographical Information Systems*, April 12-13, 2012 – Naples, Italy. *Lecture Notes in Computer Science*, S. Di Martino, A. Peron, and T. Tezuka (Eds.): W2GIS 2012, LNCS 7236, pp. 72–81, 2012. © Springer-Verlag Berlin Heidelberg 2012. ISBN 978-3-642-29246-0.

The use of Web service composition to pull down technology barriers

Michela Bertolotto¹, Pasquale Di Giovanni^{1,2}, Monica Sebillo², Giuliana Vitiello²

¹ School of Computer Science and Informatics
Belfield, Dublin 4, Ireland
michela.bertolotto@ucd.ie

² Department of Management and Information Technology (DISTR)
Via Giovanni Paolo II, 132-84084 Fisciano (SA), Italy
{pdigiovanni, msebillo, gvitiello}@unisa.it

Abstract. During the last years, the transition from stand alone and centralized systems to distributed ones has had a profound impact on the way a growing number of software solutions are designed and developed. Among the various proposals aimed at supporting the development of distributed systems, the Service Oriented Computing (SOC) has emerged as one of the leading approaches for designing and implementing distributed applications. The key concepts of this paradigm are the idea of service, an independent software module that performs a well-defined set of operations and the possibility to seamlessly combine such modules in order to offer more sophisticated functionalities. In this paper, after analyzing the main characteristics that have made the SOC paradigm a so widespread solution we discuss, as a concrete use case, the development of a software system aimed at helping farmers in Sri Lanka to improve the quality of their cultivations and related earnings.

Keywords: Service oriented computing, Services composition, Interoperability.

1. Introduction

In recent years, the development of the network infrastructures and the growing importance of the Web and its related technologies have, on one side, changed the way users access information and, on the other side, had a profound impact on the design of computer applications.

One of the most obvious results of this change is the possibility for software components located in different geographic locations to communicate in order to satisfy the requests of final users. To facilitate the design and development of such systems various solutions have been proposed and, among them, the Service Oriented Computing (SOC) paradigm has emerged as one of the leading approaches to develop distributed applications.

The fundamental idea behind this paradigm is the concept of service, an independent software module able to perform a defined set of operations. A service exposes its capabilities through its public interface that is usually described using public standards and technologies accessible by anyone in a platform independent way. Therefore its functionalities can be invoked by any type of software system: traditional desktop application, mobile applications and even other services. Such characteristic, in addition to encouraging the development of loosely coupled solutions, facilitates the reuse of functionalities since the capabilities offered by each service can be reused in more contexts and not only in the specific context for which the service has been originally developed. Moreover the availability of a technology agnostic public interface hides the implementation details to the potential clients letting, for example, an application developed using Microsoft .Net to seamlessly communicate with a service written in Java.

The ability to see services as independent building blocks and the need to only know their public interface to take advantage of their functionalities constitute the basis for the other key idea behind the success of the SOC paradigm namely services composition i.e., the ability to compose different services, developed also by different organizations to provide with complex functionalities. However, although the idea behind the services composition is quite simple, as we will see, there are some issues that must be carefully taken into account in order to guarantee a seamless exchange of information.

The remainder of this paper is organized as follows. In Section 2 we briefly describe the main features of the current standards for the development of reusable services. Section 3 discusses the fundamental characteristic of services composition as a mean to realize complex solutions based on the principles of the SOC paradigm. A description of a concrete example of a services based system is provided in Section 4 and, finally, some conclusions are drawn in Section 5.

2. Current standards for service based development

As outlined in the introduction, the functionalities of a service are exposed through its public interface. A complete description of such an interface is the only thing that a service client needs to know in order to invoke and use the service features. Once the potential client knows the operations supported by the service, all the communication between the two entities is based on various messages exchange mechanisms such as the traditional Request-Response pattern.

Such a communication must be supported by a specific middleware that must guarantee some basic activities [Tsalgatidou and Pilioura, 2002]:

- Service creation,
- Service description,
- Service publishing for potential users to locate it,
- Service discovery by potential users,
- Service invocation and binding,

The use of Web service composition to pull down technology barriers

- Service un-publishing in case it is no longer available or needed.

However since the actual implementation of services and clients might be realized using different platforms and programming languages, the use of proprietary formats for the information exchange is simply unfeasible. Therefore there is the need to describe the public interface and to provide a messaging framework in a neutral manner using globally accepted standards that must be available on each platform.

In this context, the World Wide Web Consortium (W3C) has defined a series of universally accepted standards based on the use of the Extensible Markup Language (XML) in order to guarantee their independence from a specific platform or technology. In particular, in the following subsections we will briefly describe the SOAP protocol for the exchange of messages and the Web Services Description Language (WSDL) for the description of the service interface. Moreover, in the subsection 2.3, we will compare the W3C services design choices with the service standards used by the Geographic Information Systems (GIS) community.

2.1 SOAP

SOAP is a "lightweight protocol for exchange of information in a decentralized, distributed environment" [Gudgin et al, 2001] and constitutes the backbone of the messaging framework of most service oriented solutions.

It is entirely based on XML and heavily relies on XML technologies such as XML Schema and XML Namespaces.

The structure of a SOAP message is made up of three elements: The Envelope, the Header and the Body. The Envelope can be seen as the container of the SOAP message. The Header element is used to carry additional information and although optional represents, in practice, a fundamental part for the implementation of a series of additional important protocols such as those related to the information security and reliability. Finally, the Body element contains the real payload of the exchanged message [Tsalgatidou and Pilioura, 2002] [Erl, 2005]. Such a payload is usually represented by plain XML although binary data (images or PDF documents, for example) can be embedded in a SOAP message using particular techniques such as the Message Transmission Optimization Mechanism (MTOM) or the Base64 encoding [Powell, 2004].

2.2 The Web Services Description Language

The WSDL [Christensen et al., 2001] is an XML based language for describing W3C services and how to access them. A WSDL document separates the abstract aspects of a service description from the concrete aspects such as the binding with a certain network protocol preserving, this way, the public interface of the service by changes in the underlying technology. Seven elements constitute the typical structure of a WSDL document namely Types, Message, Operation, Port Type, Binding, Port, Service. The former four constitute the Abstract Description of a service while the latter three allow the connection of the abstract interface of a service to a real technology and

transport protocol such as the Hypertext Transfer Protocol (HTTP). In the remaining of this subsection we describe in greater detail the role of each element.

The Types element can be seen as the container of the data type definitions used inside the Web Service. The Message element represents the data being communicated. In a WSDL document, it comes after the Type element. Each WSDL document can have one or more Message elements. Each Message has a univocal name and contains one or more children referred to as "Parts". The Parts can be compared to the parameters of a function in a traditional programming language. The data type of a part element can be a simple type or a type defined in the Types element. The Operation element describes, instead, the features that the Web Service will expose. Each Operation element is composed of Input and Output Elements, which refer to Messages exchanged during the communication. The Port Type element is a named set of abstract Operations and the abstract Messages involved. Every Port Type has a unique name and is made of several Operations.

The Binding, Port and Service elements describe how the Abstract Description is mapped into a concrete format. The Binding element, in particular, defines message format and protocols details for the Operations and Messages defined by a particular Port Type. The Port element represents an instance of an abstract port (Port Type) obtained as a combination of a binding and a network address and, finally Service, the higher-level element in a WSDL document declares a W3C service as a collection of related Ports.

2.3 Geographic services standards

Despite the wide acceptance of the W3C proposals the GIS community, also for historical reasons, has developed, during time, its own set of standards for the fulfillment of geospatial data oriented services.

In particular the proposals of the Open Geospatial Consortium (OGC) have become the de facto standard for developing distributed geographic applications but although based on the use of XML for the exchange of data and HTTP as the transport protocol, they are incompatible with the W3C services.

The first important difference between OGC and W3C services is represented by the strong standardization imposed by the OGC about the public interface of a geographic service. Moreover, unlike W3C services each type of OGC service represents a different standard with a fixed public interface. However, to facilitate the development process, the aspects that are common to all types of OGC services are described in the OGC Common Standard specification [Whiteside and Greenwood, 2010]. The Common Standard describes another fundamental difference: the only way for a client to know the capabilities of an OGC service is to parse the XML document returned by the standardized GetCapabilities operation. In the W3C realm the service capabilities are publicly exposed in its WSDL document.

The last difference we mention here concerns the binding type and the binding time of operations. In the OGC services the type of a response message can dynamically vary based on the client requests while in W3C

The use of Web service composition to pull down technology barriers services the message payload is completely defined at design time [Schaffer, 2008].

Two of the currently most used OGC standards are the Web Map Service standard and the Web Feature Service standard. The former provides an HTTP based interface for requesting georeferenced map images from one or more distributed geospatial databases while the latter provides interface to access and manipulate geographic features (an abstraction of real world phenomena associated with an Earth location).

3. Services Composition

The general idea behind services composition is quite simple: two or more services are combined together in order to use their functionalities to "build networks of collaborating applications distributed within and across organizations" [Nano and Zisman, 2007]. A composed service can be the result of the composition of elementary services, previously composed services or a combination of the above.

Among the various advantages of the composition we can surely mention a better reuse of the service functionalities since the same service can be used in different contexts and a better flexibility since the internal representation of a service can be modified (e.g., for refactoring or optimization reasons) without affecting the behavior of the whole system (provided that the service public interface remains unaltered). Therefore instead of developing a whole application from scratch, an organization can realize its system by composing different types of services renting, for example, third-party services for certain functionalities (e.g., the functionalities offered by a financial entity for the electronic payments) and focusing on the development of those functionalities that represent its core business. [Di Nitto et al, 2008].

Nevertheless, despite the undeniable advantages, services composition is a very challenging task still involving research both in academia and industry even though, due to space constraints, we'll briefly summarize only the most important issues.

First of all, before starting to actually compose the various services, factors such as quality of service, security, response time and cost of the involved services must be carefully evaluated since they can have a huge impact on the final result of the computation especially when the involved components are developed by different organizations. Therefore, as a common practice, the service providers and the consumers agree upon a Service Level Agreement (SLA), a contract that disciplines the rights and the obligations of the involved participants [Di Nitto et al, 2008].

For what concerns the real composition the fact that, now, all the operations and algorithms might be the result of the composition of different services represents the first important issue to deal with. Such a peculiarity implies, as a direct consequence, the need to coordinate the sequence of operations to ensure the correctness of the computation and avoid inconsistencies [Dustdar and Schreiner, 2005] and constitutes an important difference when the service based applications are compared with the traditional monolithic systems.

Another important issue is related to the choice of the composition strategy. In fact, the software architect might choose to compose services during the design of the system (Static composition) or during its execution (Dynamic composition). Static composition is a suitable option only when there is the certainty that the involved services will rarely change [Dustdar and Schreiner, 2005].

The change in the structure of a service represents another problem that must be taken into account. Two main types of changes can be identified [Di Nitto et al, 2008]:

- Shallow service changes when the changes are localized to a single service
- Deep service changes when the changes may affect the entire service chain.

The last fundamental issue we mention here concerns the need to have a commonly accepted “composition model and a language to specify the services involved in the composition” [Dustdar and Schreiner, 2005] that describes in a formal manner and without ambiguity the sequence of the operations to be performed.

The most widely accepted approaches are the services Choreography and Orchestration.

In the Choreography approach each participant is provided with some rules that must follow. The global behavior of the system is the result of the interaction of the various parties each following its own rules.

In the Orchestration approach, instead, the composition process is represented by a sequence of steps, conditions and exceptions coordinated by a central controller [Rosen, 2008]. In the Orchestration context, the Web Services Business Process Execution Language (WS-BPEL) represents the de facto standard to create applications by composing existing services [Alves et al, 2007]. WS-BPEL is expressed by using XML and heavily relies on several W3C specifications such as the previously described WSDL.

Several elements contribute to define the definition of a WS-BPEL process. The most important of them are the <PARTNERLINK> elements that identify the various services involved in the process, the <VARIABLES> elements that contain the variables holding the messages exchanged during the composition and the elements used to arrange the execution order such as the <SEQUENCE> element or the <IF> element for the conditional execution.

As a concrete example of services composition, in the next section we briefly describe our choices for what concerns the design of a software system aimed at helping Sri Lankan farmers to improve the quality of their cultivations and related earnings.

4.The Social Life Networks for the Middle of the Pyramid project

The Social Life Networks for the Middle of the Pyramid (SLN4MoP) is an international collaborative research program that aims to provide real-time information to support activities related to livelihood targeted to meet the needs

The use of Web service composition to pull down technology barriers of people living in developing countries. In particular, agriculture in most of these countries represents a major economic sector, employing the largest share of the workforce, while suffering from low productivity. The reasons for this low productivity may include land fragmentation, lack of postharvest infrastructure, weak market linkages, information and knowledge asymmetries (or lack thereof), and, most important, low technology utilization. Therefore, as a pilot study, we have designed and are developing a software system aimed at helping Sri Lankan farmers to improve their rural activities and gain higher incomes.

From a high level point of view, the proposed system presents a traditional client-server architecture: a client sends a request to a remote backend which processes it and sends back the desired results. For what concerns the actual structure of the system and the subdivision of the available functionalities a three-tier subdivision, namely a client tier, a middle tier and a data tier has been chosen.

For the time being, the client tier mainly consists of applications running on modern mobile devices, although other potential stakeholders such as government agencies could use different type of applications (e.g., traditional desktop application) for their needs. Using such mobile applications farmers can get detailed information about the best crops to grow according to the type of soil or the expected income, detailed information about the best fertilizers to use, information about the selling price of the selected products and can send data about their choices. Such data, aggregated in an anonymous form, will be processed by the middle tier to update, for example, the average cultivated quantity of a certain crop into a specific area. Detailed information about the design choices concerning the development of the mobile application used by the farmers can be found in [Di Giovanni et al, 2012] [Di Giovanni et al, 2013] [Sebillo et al, 2013].

Different and potentially distributed data sources constitute the data tier such as information about the farmers, the various types of cultivable crops, the land extension of a farm or the different soil types. Moreover, such information could be stored using different type formats.

For what concerns the middle tier, the technical choices behind its design have been influenced by several factors. First of all the system must be able to provide not only with different types of information to the various types of potential stakeholders involved in the project but must also be able to provide the same information with different levels of granularity according, for example, to the specific user's preferences or security policies. Therefore flexibility is one of the requirements that influenced our choices.

Other important requirements to satisfy were the ability to add new features without affecting the existing components, the independence of the system functionalities from the specific format of the various data sources and the possibility to seamlessly add new data sources or replace existing one without modifying the behavior of existing implementations. Finally, in order to accelerate the development process, the opportunity to use different development tools from different vendors was a desirable addition.

To satisfy such requirements, instead of developing a single, monolithic application we chose to develop independent and reusable modules for each

functionality. We also decided to use a common protocol for the exchange of information among such modules. With these premises, organizing the design of the middle tier around the principles of the SOC paradigm has been a natural step.

After the initial partitioning of the functionalities into a set of services, we further specialized the derived services into three layers of abstraction [Erl, 2005]:

- The application service layer
- The business service layer
- The orchestration service layer

Such a subdivision presents several advantages. Among them we can mention a better reuse of the functionalities, a better maintenance and most important an improved flexibility of the whole system since new modules can be added or existing modules can be substituted e.g., with a more optimized version without affecting the behavior of the remaining services. Moreover third party solutions can easily interact with the functionalities exposed by our services.

The application service layer represents the lower level of the abstraction and includes all those services that expose solution agnostic and reusable functionalities. Such functionalities are used in the upper level, the business service layer and, finally, the services in the orchestration layer organize and compose the services of the lower layers in order to offer the desired functionalities to the final users.

The actual technological platforms we chose for the development of our service oriented system deserve further considerations. The majority of our platform is being developed in compliance with the above mentioned services standards proposed by the W3C. However, although W3C service are one of the best alternatives for the development of enterprise class distributed applications, as we have seen, they are not the best choice for the management of geospatial data.

Geospatial data are of utmost importance for the purposes of our system (for example, the suggestions for a specific user about the best crops to grow are based not only on the current market trends but also on the specific characteristics of the soil where his/her farm is located). Moreover since the geospatial information could be useful also for third party entities (e.g., a government agency might be interested to visualize the various soil types available in a specific region using a traditional desktop GIS application such as uDIG) we chose to develop the services that deal with geospatial information following the standards proposed by the OGC (in particular, our implementation will be compliant with the Web Feature Service standard).

Unfortunately, as we have seen, despite sharing some common characteristics such as the use of XML for the exchange of information, the W3C and OGC services are incompatible.

Therefore, to guarantee the highest level of interoperability among our various software modules, according to the solutions discussed in the literature, we are developing a service wrapper whose purpose is to query the proper

The use of Web service composition to pull down technology barriers OGC service and return the geospatial information in a W3C compliant way keeping the structure of the original W3C or OGC services unchanged.

In particular, the wrapper carries out the following tasks:

1. It receives a SOAP message from a generic client containing a request for a specific geospatial data
2. It translates the SOAP message in a format suitable for the underlying OGC service and sends the request
3. It translates the received response in a SOAP message and sends it to the client.

A detailed analysis of the challenges we are facing during the development of the wrapper can be found in [Iouf et al, 2008].

We briefly mention that such issues can be divided into three main categories namely Data handling, Functionalities mapping and Metadata management.

The first type of issue concerns the different data types that can be returned by an OGC service, the second issue deals with the difficulties to correctly expose the interface of an OGC service using a WSDL document and, finally, the third issue relates with the need to deal with the spatial metadata required by each OGC service.

5. Conclusions

The SOC paradigm has radically changed the way distributed applications are designed and developed. The key idea behind this approach is the concept of service, an autonomous software module that combined together with other services can be used to create complex solutions. Unfortunately, services composition is a non-trivial task and several aspects must be taken into account. In this paper we have provided an overview of the current main standards used for the actual services development and a description of the most important challenges that arise during the composition task. We have also described some design choices we made for the development of a services based system aimed at helping Sri Lankan farmers to improve their farming activities.

Bibliography

[1] Alves A., Arkin A., Askary S., Barreto C., Bloch B., Curbera F., Ford M., Golland Y., Guizar A., Kartha N., Liu C. K., Khalaf R., König D., Marin M., Mehta V., Thatte S., van der Rijn D., Yendluri P., Yiu A., Web Services Business Process Execution Language Version 2.0, OASIS, 2007, <http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.pdf>

[2] Christensen E., Curbera F., Meredith G., Weerawarana S., Web Services Description Language (WSDL) 1.1, World Wide Web Consortium, 2001, <http://www.w3.org/TR/wSDL>

[3] Di Giovanni P., Romano M., Sebillo M., Tortora G., Vitiello G., De Silva L., Goonethilaka J., Wikramanayake G., Ginige T., Ginige A., User Centered Scenario Based Approach for Developing Mobile Interfaces for Social Life Networks, Proc. of the First International Workshop on Usability and Accessibility Focused Requirements

Engineering (UsARE 2012), June 4, 2012, Zurich, Switzerland ISBN 978-1-4673-1846-4, IEEE, 2012 18-24.

[4] Di Giovanni P., Romano M., Sebillio M., Tortora G., Vitiello G., De Silva L., Goonethilaka J., Wikramanayake G., Ginige T., Ginige A., Building Social Life Networks through Mobile Interfaces - the Case Study of Sri Lanka Farmers, in Spagnoletti P. (ed.) Organizational Change and Information Systems, Lecture Notes in Information Systems and Organisation 2, Springer-Verlag Berlin Heidelberg 2013.

[5] Di Nitto, E., Ghezzi, C., Metzger, A., Papazoglou, M., Pohl K., A journey to highly dynamic, self-adaptive service-based applications, Autom Softw Eng (2008) 15: 313–341.

[6] Dustdar, S., Schreiner, W., A survey on web services composition. Int. J. Web Grid Serv. 2005 1(1): 1-30.

[7] Erl T., Service-oriented architecture: concepts, technology, and design, Prentice Hall PTR, Upper Saddle River, New Jersey, 2005.

[8] Gudgin M., Hadley M., Moreau JJ., Nielsen H. F., SOAP Version 1.2. World Wide Web Consortium, 2001, <http://www.w3.org/TR/2001/WD-soap12-20010709/>

[9] Ioup E., Lin B., Sample J., Shaw K., Rabemanantsoa A., Reimbold J., Geospatial Web Services : Bridging the Gap between OGC and Web Services, in Sample J.T., Shaw K., Tu S. and Abdelguerfi M. (eds) Geospatial Services and Applications for the Internet. Springer, New York, 2008, 73-93..

[10] Nano E., Zisman A., Realizing Service-Centric Software Systems, Software, IEEE 24,6, 2007, 28-30.

[11] Powell M., Web Services, Opaque Data, and the Attachments Problem, Microsoft Corporation, 2004. <http://msdn.microsoft.com/en-us/library/ms996462.aspx>

[12] Rosen M., Orchestration or choreography?. <http://www.bptrends.com/> 2008

[13] Schffe o mmon nginee ing ep o t pen Geospatial Consortium Inc, 2008 <http://www.opengeospatial.org/standards/dp>

[14] Sebillio M., Tortora G., Vitiello G., Di Giovanni P., Romano M., A Framework for Community-Oriented Mobile Interaction Design in Emerging Regions, Kurosu M. (Ed.) Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, Springer-Verlag Berlin Heidelberg, 2013, 342–351.

[15] Tsalgatidou A., Pilioura T., An overview of standards and related technology in web services. Distrib. Parallel Databases, 12,2-3, 2002, 135-162.

[16] Whiteside A., Greenwood J., OGC Web Services Common Standard. Open Geospatial Consortium Inc., 2010, <http://www.opengeospatial.org/standards/common>

A Mobile Visual Technique to Support Civil Protection in Risk Analysis

Luca Paolino¹, Monica Sebillo¹, Genoveffa Tortora¹, Giuliana Vitiello¹, Marco Romano²

¹*Dipartimento di Studio e Ricerca Aziendale (Management & Information Technology)
Via Giovanni Paolo II, Fisciano(SA) - Italy
{lpaolino, msebillo, tortora, gvitiello}@unisa.it*

²*Departamento de Informática (University Carlos III de Madrid)
Avenida de la Universidad 30, Leganés (Madrid) – Spain
mromano@inf.uc3m.es*

Abstract. *The importance of territory defense is becoming an activity of national civil defense agencies more and more important. In our research we describe a participatory design process involving real stakeholders from the emergency management field, which has led to the development of a mobile application supporting on-site operators. The application was based on the requirements we extracted by interviewing operators on a real scenario. The utility of the application is successively described on three typical scenarios.*

Keywords: mobility, digital society, risk analysis.

1. Introduction

Vulnerable sites enlarge upon one or more dangerous or vulnerable geographic features inside a specific area. Experts consider a geographic feature as “dangerous” if, due to human or natural causes, it can induce the partial or total failure or loss of functionality of other features in its neighborhood [10]. Examples of dangerous sites are portions of land subject to landslides, chemical plants, dumps, etc. Differently, the presence of dangerous features, which can be subject to partial or total failure or loss of functionality, makes a geographic feature “vulnerable”. Some examples of vulnerable features are civil infrastructures and human settlements. Anyway, the same geographic feature can be considered both as dangerous and vulnerable.

In this context, it is very important that nations are able to locate and evaluate the vulnerability and dangerousness of the own territories also considering human or natural factors. For this reason, every nation has organized special departments and agencies whose aim is territory control. The diversity of emergency situations which originated from natural disasters occurred on the Earth in recent years, have raised several ICT research issues covering different areas, including data visualization, (geo)visual analytics,

Congresso Nazionale AICA 2013

advanced (mobile) interfaces, communication technology and collaborative environments. The research efforts put to address those issues have come to highlight that reliable information and good communications are crucial requirements to achieve effectiveness in emergency response preparedness and management.

In this paper we describe a participatory design process involving real stakeholders from the emergency management field, which has led to the development of a mobile application supporting on-site operators for dynamic, advanced analysis of heterogeneous data. The visual analytics power of a technique named Framy [6] is exploited to facilitate the analysis of the risk levels of the area surrounding the operator.

The application requirements came out of a contextual inquiry we performed in the territory of Salerno, Italy, which involved the Civil Defense Department of forest rangers.

We were able to interview subjects on vulnerable areas, on the kind of activities they commonly perform when dealing an emergency and on the extent to which mobile devices (including cell phones and personal digital assistants, PDAs) are adopted for continuous information exchange between on-site operators and the emergency operating center. The goal of the study was to analyze current practices and gain insight on the activities requiring the use of computers, as they really occur during the risk analysis phase and on the activities that could benefit from the adoption of (mobile) computer applications. The derived requirements guided the subsequent design of a mobile application that could address the need to dynamically analyse spatial data on small sized screen.

The paper is organized as follows. In Section 2, we describe the current practices of the rangers appointed to patrol, identify and analyze vulnerable areas. Then we present the requirements that an information system should present to support this kind of users during their operations. Section 3 presents the mobile application we have designed on the basis of the requirements and three typical scenarios of usage of the application which better illustrate the advantages coming from the adoption of Framy in this field of civil defense. Some final remarks conclude the paper.

2. Practices and Requirements of Emergency Management by Civil Defense Departments

Operators of the Civil Defense Department of Salerno are in charge to evaluate the risk level of vulnerable areas of the above-mentioned province. Some evaluations can be performed directly by the command center by using advanced software that analyzes data stored in a database. On the other hand, some evaluations are performed directly in situ because the presence of an operator is essential. Moreover, the majority of data stored in the central database can be collected only in situ. Therefore task forces of operator explore the vulnerable areas to gather information and evaluate the risks. These task forces are continuously in contact with both command center and the other rangers by means of radios, pocket radios and mobile phones. They exploit professional GPS modules to annotate positions with accuracy and to orient

themselves. The kinds of areas that they can explore are different. They can be in charge of patrol urban areas, forests or vulnerable technological parks such as dams. During the exploration they need data coming from the other groups of operators to make decisions about the risk level of an area. For example, a team notices that the water level of a river that flows near a urban area is too high; another team can retrieve information about a violent storm coming from a near region towards the city. These two pieces of information, if combined together, can produce a high-risk level for the urban area.

Another reason to combine data is to allow operators to make decisions about the route to be followed during the exploration. For example, a team can alert about obstacles in their area such as fires or landslides that can make the route hard. Decision makers, on the basis of such information, are able to choose the most convenient route.

The proposed scenario of existing working practices gave us the opportunity to reason about what were the major requirements that emerged from the context. At this stage, we were primarily interested in deriving a list of requirements that we could take into account throughout the design process. Summing up all the considerations and discussions we had upon the fieldwork completion, and reasoning on the derived scenario, we were therefore able to elicit an initial set of requirements divided into five categories according to the classification described by Preece et al. [9]. The initial set of requirements and the corresponding rationale for the design of an application supporting rangers patrolling rural areas or estimating risk levels of the same.

Functional requirements: 1) The application allows the exchange of heterogeneous data between operators working both directly on vulnerable sites in mobility and in command centers. 2) The application analyzes in runtime data to establish the risk level. Rationale (R): By sharing valuable information about natural phenomena and morphological data, users would be able to make better decisions during the risk analysis and the exploration of vulnerable site.

Environmental and contextual Requirements(ECR): The application is used on the go. R: users are mobile, need light instruments to move from a place to another nimbly. ECR: The application may be used in off-line mode. R: Some environment could not be reached by communication networks. ECR: Users can receive a training provided by experts. R: Users are members of the staff of a professional organizations that can provides an adequate training .

Data requirements(DR): 1)The application has the access to data related to the distribution of geographical elements such as the extension of mountain areas, rivers and forest. 2) Moreover it has the access to elements that present both geographical and temporal components such as storms, fires, natural disasters, in other words elements that are only momentarily present in the area. R: Operators are interested to get information about geographical elements all around them but even about natural phenomena that have temporally a geographical position. DR: Data must be accurate and updated frequently. R: Users make crucial decisions on the basis of provided data.

User profiles: 1) The range of the instruction level varies between Ordinary Level and Master Degree. 2) Users already use some technological instruments

during the exploration of a vulnerable area, such as GPS module, phone, and radio. R: All of the information is carried out from the initial survey conducted directly in situ.

Usability requirements(UR): The application should be easy to learn and should re-quire a little training effort. R: The application is used frequently during by operators working directly on vulnerable sites. UR: The user interface should be effective: it should provide a simple management of users mistakes. R: The application provides support to a critical task on the basis of data updated directly by users. Therefore, it is paramount to reduce the number of possible unintentional user mistakes.

3. Application design and Usage Scenarios

Starting from the previous set of requirements we decided to exploit the visual analytics power of the technique named Framy [6] to provide users with a usable mobile applications aimed to analyze the risk levels of the area surrounding the operators. Framy has proven a useful technique to address different problems related to the management of spatial data on small sized screens, including augmented reality applications [3] and multimodal interfaces [2]. Also, in the context of emergencies, it has been successfully adopted [4] to support decision makers working directly on the disaster sites during the critical phases of the crisis management. This inspired us in reusing the technique in a different context where the mobility and the advanced analysis of heterogeneous data play the main roles.

Framy is explicitly targeted to enhance geographic information visualization on small-sized displays by providing information clues about features of either the on-screen or the off-screen space. The rationale behind it is to display semi-transparent coloured frames along the border of the device screen. Each frame is then partitioned and coloured with a saturation index featuring the Hue-Saturation-Value model (Smith 1978). The intensity of such an index is proportional to the result of an aggregate function which either aggregates a property of objects located in the corresponding map sector either inside or outside the screen such as sum and count, or calculates the distance between the map focus and a point of interest (POI) in that portion.

Thus, for instance, a frame may indicate the amount of facilities located within some specific sectors or the sum of values of a given attribute. If several aggregates are displayed on the map, nested frames may be visualised along the borders, each one corresponding to a different aggregate, with a different colour.

Formally, the colour intensity of a frame portion C_i can be expressed as

$$\text{Intensity}(C_i) = f(g(S_i)) \text{ for each } i \text{ in } (1, \dots, n)$$

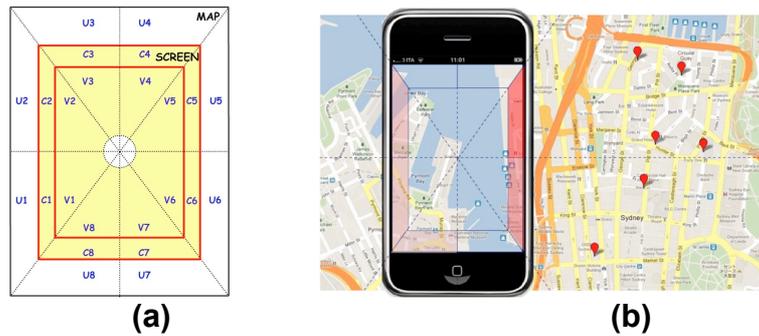


Fig. 1. (a) An example of (on/off-)screen subdivision accomplished by Framy, (b) Application running Framy

where f is a monotonic function; g is a function which aggregates a set of spatial data; and given U_i (invisible) which represents the map portion of a sector which lies out-side the screen and V_i (visible) which represents the map portion which is inside screen, S_i is one of either U_i or V_i or $U_i \cup V_i$ (Figure 1a).

Figure 1b illustrates an application running the Framy technique. The frame is partitioned into eight coloured portions. The intensity of each portion is proportional to the number of POIs located around the map focus.

Details about the aggregate values associated to each sector, such as the number of POIs, the POI names or their distances from specified locations, may be required by tapping on the corresponding sector.

In order to allow users to easily distinguish colours and correctly interpret their meanings, the Framy visualisation technique has been supplied [6] with colour rules de-rived from Itten's [8] theory which provides a way to select highly contrasted colours. In particular, the rules we have adopted concern the colour combination of a frame and the map background, the colour combination of a frame and the corresponding features and the colour combination of nested frames.

As for the territory analysis and the evaluation of vulnerability and dangerousness, the application can be used directly during the exploration activity by exploiting the integrated localization sensors and data gather on the go, or can even be used to analyze data being in other locations.

As Figure 2a shows, the applied technique is used, first of all, to analyze data surrounding the operators, such as the presence of forests, rivers, urban settlements and natural phenomena as storms and snowfalls, and then to visual aggregate this data by using the visual metaphor of the frames. Combined elements generate a risk level: as higher the level is more intense will be the color of the corresponding frame portion. Therefore, just by looking the frame, the operators immediately understand in which directions are the vulnerable areas and their estimated risk levels.

The criteria to combine elements can be set directly by the operators. For example they can decide that an area is dangerous if at least three elements are present, for example a storm, a river and an urban area. Finally for each kind of combination the operator can set a different risk level.

The evaluation of the risk is performed directly by the mobile application and not by a remote server. This choice is necessary because the operators may work in areas out of the communication networks and can work with data in off-line.

The application allows users to send updated information to the server of the command center. Then, it updates the current situation to the rangers in the interested area.

In the example depicted in Figure 2a, icons represent the elements that can create a risk. Circle shapes indicate the territorial extension of those elements. In northwestern areas snowfalls can cause serious damages to the routes through the mountains and forests, while in northeastern areas the presence of a river close to a urban area during the storm causes a risk level even higher.

The usage of the frame can be fully personalized. For example, the user can set different nested frames; each frame has a different hue that is related to different risk evaluations (Figure 3a). Another possibility is to exploit the analytical power given by applying two aggregate functions to the same frame. The first function modifies the saturation while the second the brightness. It is formally expressed in the following formulas:

- saturation (Ci)= $f_1(g_1(S_i, L_j))$ for each i in (1,...,n) and j in (1..m)
- brightness (Ci)= $f_2(g_2(S_i, L_j))$ for each i in (1,...,n) and j in (1..m)

where f_1, f_2 are monotonic functions; g_1, g_2 are functions which aggregate a set of spatial data of specific layer.

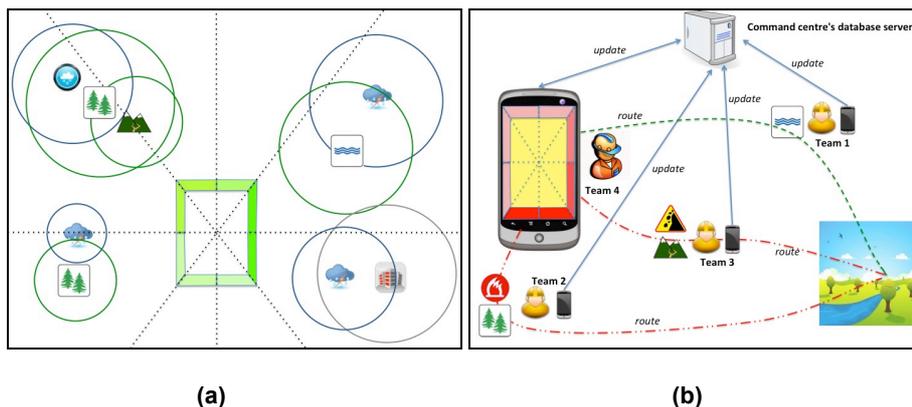


Fig. 2. (a) The frames display the analysis of the natural phenomena combined with the human structures, (b) Identification of the easiest path to follow

Scenario 1: identification of a safe path in a rural area. In Figure 2b, it is depicted a typical scenario where the application may be exploited. Four teams are working in patrolling a quite huge vulnerable area. The area is a natural

park that during the summer often is at risk of fires. Moreover, the whole mountain range of the park is at risk of landslide.

Team 1 is in charge to patrol an area beside the river. This area has been often flooded during the last winter; therefore it is necessary to check the status of the ground. Team 2 is exploring a forest area where notice a beginning of fire made by a pyromaniac. Team 3 is assigned to a mountain area at risk of landslide. Team 4 is the last team to start the mission and is in charge to monitor the most distant area. In order to get that area they can follow three different routes through the mountains, the river or the forest. The first three teams send the information gathered during the monitoring activities to the command center's server that informs all the teams about the current situation. The application in possession of team 4 calculates the risk level for the different areas all around the team. By analyzing the frame intensities, the decision maker takes in account the high difficulty level related to the routes passing through the forest effected by the fire or the mountain that present some landslides. While, because team 1 reported that the area of the river is in a good status, the frame portion related to that area presents a low intensity level. Therefore, the decision maker determines to follow the route through the river monitored by team 1.

Scenario 2: Arsons Prevention. In this section we show a typical scenario where the application can be used to support the evaluation of arson risks. A decision maker of a civil defense task force is planning the patrol of a huge area in order to prevent the possibility of arsons. He sets the application with two personalized frames. As Figure 3a shows, the first frame (the green one) represents the distribution of the areas vulnerable to the arsons. The risk is calculated on the basis of the vulnerable elements present in those areas. More in detail, the areas in the South, Southwest and Southeast are considered more vulnerable than the others, because of the high presence of forest. In this case, the intensities of the corresponding frame portions are very high. The urban areas, represented by the motorway and the city, are considered less vulnerable and the intensities of the corresponding frame portions are significantly lower. The areas on the west part are considered not particularly vulnerable because the river makes those areas not easily accessible.

The second frame, the red one, represents the distribution of the arsons of the last two years. In particular the portion indicates that the areas with the majority of arsons in the past are the ones near the motorway and the mountain. In the southern area the number of arsons is quite small while they are totally absent near the river and the city. Therefore the red frame, on one hand, confirms that the difficult accessibility of the eastern areas protects them from the arsons, on the other hand, it highlights that the area around the motorway, even if it is not considered pretty vulnerable judging from the green frame, is the area most exposed to the risk of arsons. This is probably due to the easy access of pyromaniacs to that area.

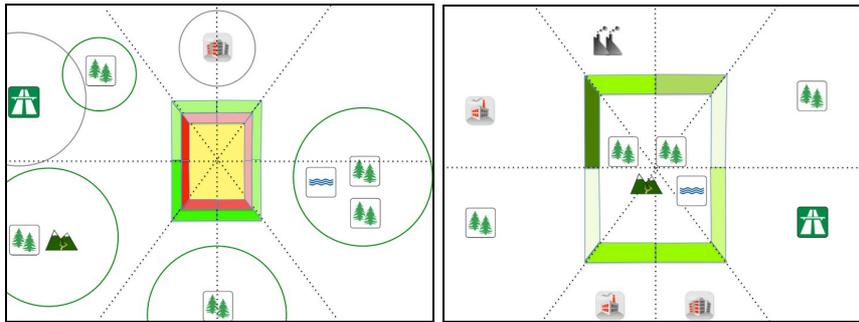


Fig. 3. (a) The usage of the technique to analyze the risks of fires, (b) The usage of Framy to evaluate the industrial pollution

Scenario 3: Pollution Control. The last scenario shows as one single frame can be exploited to analyze different kinds of information.

The civil defense is evaluating the risks of contamination coming from the industrial and urban areas situated all around a natural park. As it is shown in Figure 3b, the single frame displays the level of produced pollution through the saturation, while the brightness indicates the distance between the origin of the pollution and the park.

More specifically, even if the motorway does not produce a high level of pollution, its distance from the park is really short, so that this increases the level of contamination risk. The saturation of the corresponding portion is low but its brightness is really high. Another example is given by the industrial park in the North; the pollution level is very high and even the distance from the natural park is pretty short. In this case both the brightness and the intensity levels are quite high.

3. Conclusions

In the paper we have briefly described the design of a mobile application which allows to visualize information and aggregation of results on small screen devices. The scenario-based method we adopted was especially useful for participatory design. The participation of potential future users in the development process started from the requirement analysis and continued throughout the design and prototyping phases, relying on their expertise in the domain and their feedback on all the crucial choices. Interaction scenarios were also used to provide evidence of the appropriateness of the proposed system for on-site emergency response preparedness and management. At present we are planning a usability evaluation study for the developed application. Of course, in order to assess the validity of the study we will involve emergency operators other than those who collaborated during design.

5. References

1. Ginige, A., Romano, M., Sebillio, M., Vitiello, G., Di Giovanni, P.; "Spatial data and mobile applications: general solutions for interface design", Proceedings of AVI '12 Proceedings of the International Working Conference on Advanced Visual Interfaces, pages 189-196 ACM New York, NY, USA ©2012
2. Di Chiara, G.; Paolino, L.; Romano, M.; Sebillio, M.; Tortora, G.; Vitiello, G.; Ginige, A., "The Framy user interface for visually-impaired users," Digital Information Management (ICDIM), 2011 Sixth International Conference on , vol., no., pp.36,41, 26-28 Sept. 2011
3. De Chiara D., Paolino L., Romano M., Sebillio M., Vitiello G.; "Link2u: Connecting Social Network Users Through Mobile Interfaces". Proceedings of IEEE Pacific-Rim Conference on Multimedia (Shanghai) , Springer Verlag 2010
4. Paolino, L., Romano, M., Sebillio, M., Vitiello, G., "Supporting the on-site emergency management through a visualisation technique for mobile devices". Journal JLBC 4(3&4): 222-239, Elsevier, 2010.
5. De Chiara D., Paolino L., Romano M., Sebillio M., Tortora G., Vitiello G., "Augmented Map navigation through Customizable Mobile Interfaces". Proceedings of The 16th International Conference on Distributed Multimedia Systems (Chicago), Knowledge Systems Institute 2010
6. Paolino, L., Vitiello, G., and Tortora, G., "Framy: visualizing geographic data on mobile interfaces. Journal of Location Based Services", 2 (3), 236–252. 2008
7. Paolino L., Romano M., Sebillio M., Tortora G., Vitiello G.; "Audio Visual Information Clues about Geographic Data on Mobile Interfaces". Proceedings of The IEEE Pacific-Rim Conference on Multimedia (Bangkok), Springer Verlag 2009
8. Itten, J.; The Art of Color, New York, John Wiley & Sons, Inc. 1961.
9. Preece J., Rogers Y. and Sharp H., "Interaction Design", Apogeo, 2004
10. Della Rocca A. B., Fattoruso G., Lo Curzio S., Pasanisi F., Pica R., Peloso A., Pollino M., Tebano C., Trocciola A., De Chiara D., Tortora G., "SISI Project: Developing GIS-Based Tools for Vulnerability Assessment". VISUAL 2008: 327-330

I servizi informativi integrati per la gestione del territorio della regione Campania

Anna Martinoli, Orlando Battipaglia, Vincenzo Parità
AGC 16 Governo del territorio - Regione Campania
{o.battipaglia, a.martinoli@maildip.regione.campania.it}; v.parita@regione.campania.it.

According to the Legge Regionale n.16/04, D.P.C.M. 11/11 2011 and art. 59 from D.lgs. n. 82/2005 “Codice dell’Amministrazione digitale”, the Campania Region has realized a virtual community of users – producers and consumers – based on the usage of a Regional Geographic Information System. Services, functionality and data, provided through the infrastructure of the regional network and the integrated geographic database, are meant to support an efficient planning of Campania territory resources. The GeoPortale (<http://sit.regione.campania.it>) represents the entry point to these services. This paper describes such a functionality and the employed technologies.

Keywords: global positioning system, cartography, digital orthophoto

1. Introduzione

Il Sistema Informativo Territoriale della regione Campania (SIT) è stato istituito con Legge regionale n. 16/2004 ed è incardinato presso l'Area Generale di Coordinamento Governo del Territorio. Le funzioni che esso svolge sono in sintesi quelle di acquisire, organizzare, standardizzare e distribuire i dati territoriali, ai vari Settori regionali, agli Enti strumentali della regione Campania ed agli Enti Locali, per una più efficace programmazione e pianificazione negli ambiti di interesse dell'amministrazione regionale. Il SIT è parte integrante del Settore Monitoraggio e Controllo degli Accordi di Programma, quale struttura per la gestione e il coordinamento dei sistemi informativi territoriali in regione Campania (La struttura organizzativa del SIT è diretta dal dirigente Anna Martinoli e comprende al suo interno i funzionari: Orlando Battipaglia, Francesco Panebianco, Giuseppe Papa, Vincenzo Parità).

I compiti svolti dal SIT sono:

- rapporti con Enti ed Istituti operanti nell'ambito della cartografia e dei sistemi informativi territoriali. In particolare con: Istituto Geografico Militare (I.G.M.), Agenzia delle Entrate (ex catasto), Servizio Geologico Nazionale, Agenzia per l'Italia Digitale (ex DigitPA), Ministeri competenti in materia di informazione territoriale;
- partecipazione al CISIS quale organo tecnico della Conferenza delle Regioni e delle Province autonome in materia di sistemi informatici, geografici e statistici che opera per le singole materie di competenza esclusiva e congiunta attraverso tre Comitati Permanenti: Comitato Permanente Sistemi Informatici (CPSI), Comitato Permanente

Fig.1 - Catasto incendi boschivi

- **Omogeneizzazione dei dati territoriali alle specifiche comuni ai fini della condivisione dell'informazione geografica.**
Intervento di adeguamento alle specifiche tecniche Intesa-Gis dei dati geografici acquisiti presso i Settori regionali, gli Enti territoriali e gli Enti strumentali della Regione Campania;
- **Realizzazione del repertorio dei metadati regionali**
Acquisizione ed aggiornamento dei metadati delle singole banche dati geografiche e tematismi dei vari Settori regionali nel catalogo di metadati, nello standard ISO 19115;
- **Centro Tematico Repository**
Acquisizione hardware e software e infrastrutture di comunicazione per la gestione del Centro Tematico Repository del SIT regionale (vedi Fig.2);

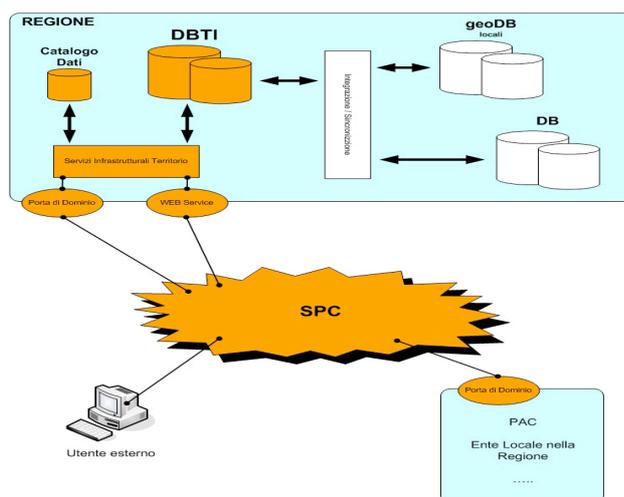


Fig.2 - Centro tematico repository

- **Rete di stazioni permanenti GPS**

Realizzazione e gestione della rete di 13 Stazioni GPS permanenti per i servizi di distribuzione di dati GPS di posizionamento in post-processing (PPDB), o in tempo reale (NRTK). Le stazioni di Carinola (CE) e San Marco di Castellabate (SA) rientrano nella rete dinamica nazionale, gestita dall'I.G.M. (vedi Fig.3);

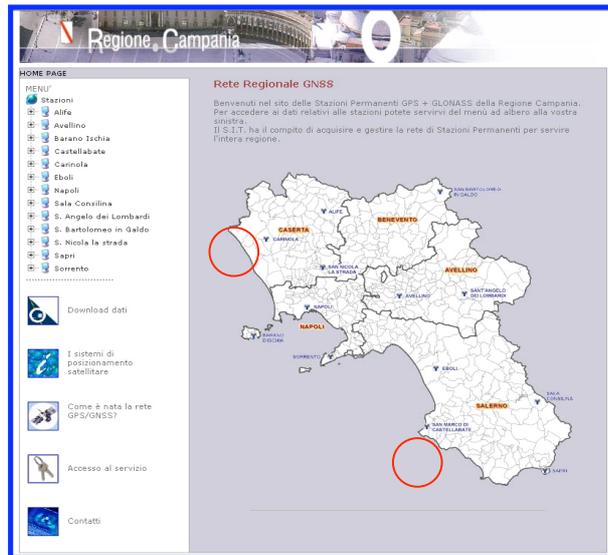


Fig.3 - Rete di stazioni permanenti GPS

- **DB Topografico alla scala 1:5.000**
Realizzazione del Data Base Topografico (DBT5k) mediante conversione della Carta tecnica regionale numerica anno 2004/2005 secondo le specifiche di cui al D.M. 10 novembre 2011 (vedi Fig.4);

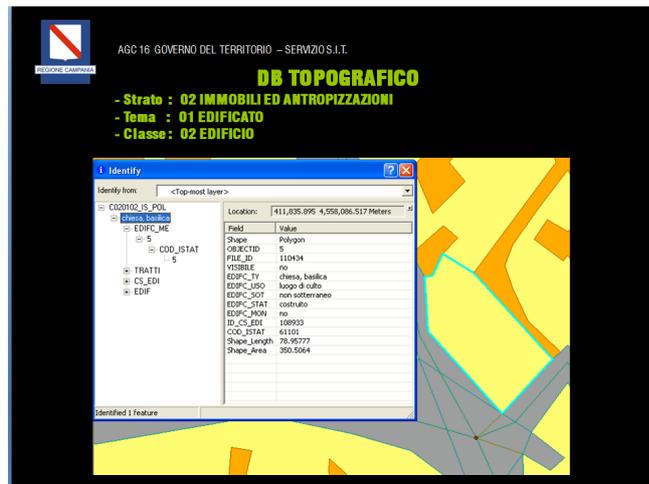


Fig.4 - Data Base Topografico

- **Servizio di consultazione delle riprese aerofotogrammetriche eseguite dalla regione Campania**

Realizzazione del web gis di consultazione on-line delle riprese aerofotogrammetriche della regione Campania (vedi Fig.5);

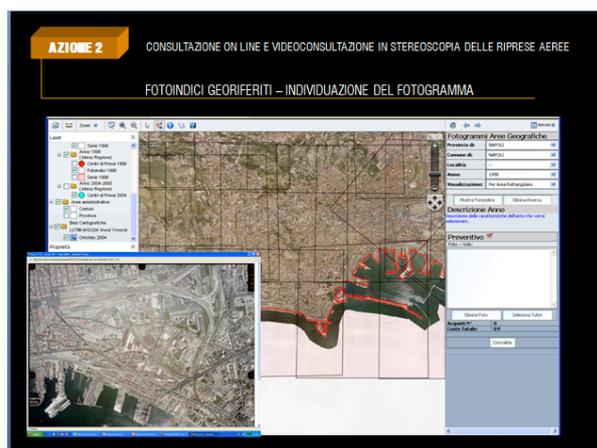


Fig.5 - Web gis riprese aerofotogrammetriche

- **Servizi di consultazione delle basi dati catastali (riuso SIGMATER)**

Digitalizzazione delle mappe catastali delle provincie di Caserta e Benevento, semina dei punti in doppia coordinata dei punti fiduciali catastali per la trasformazione di coordinate, realizzazione del web gis di consultazione dei dati catastali e servizi di interscambio, in cooperazione applicativa, con l'Agenzia delle Entrate (vedi Fig.6);

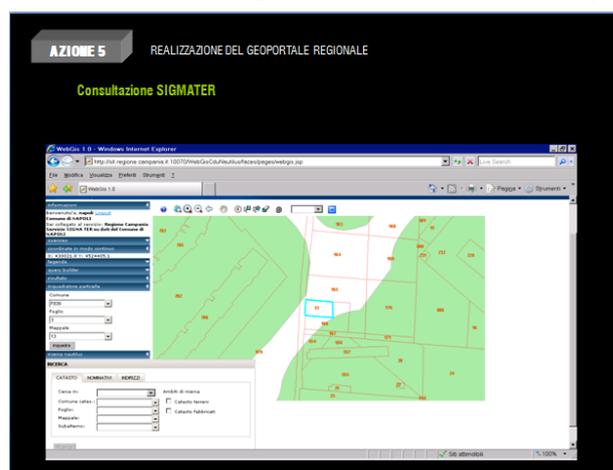


Fig.6 - web gis consultazione dati catastali

- **Realizzazione delle linee livellazione di alta precisione in accordo con l'I.G.M.**

Verifica e riattazione dei capisaldi delle linee di livellazione geometrica di alta precisione sul territorio della regione Campania (circa 1000 km);

- **Portale Cartografico Regionale**

Il GeoPortale, realizzato su piattaforma Jboss portal, comprende tutti i servizi realizzati nell'ambito del progetto oltre a specifiche aree di utilità, quali: forum, news, sondaggi, upload, download, conversione formati shape, content management system, etc.. Sono inoltre disponibili web-services conformi alle specifiche OGC (WMS per i profili 1.1.0, 1.1.1, 1.3.0.) (vedi Fig.7). Sono presenti i link ai principali portali tematici disponibili in regione Campania, quali: contrasto all'abusivismo edilizio (Mistrals), difesa del suolo, grotte e cavità della Campania e cavità di Pianura 3D.

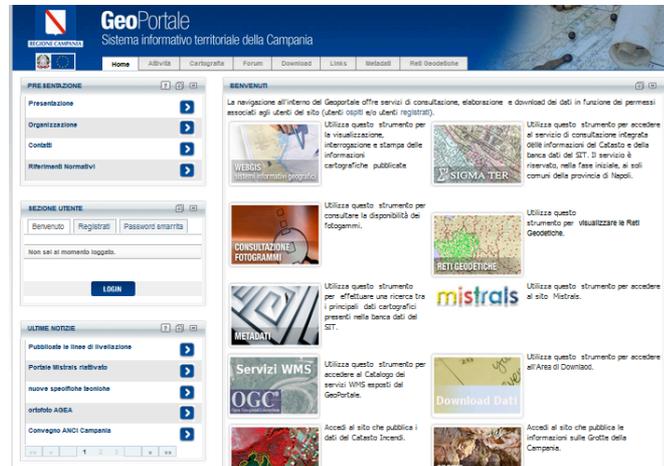


Fig.7 - Portale cartografico regionale

Il progetto PR5SIT è stato ultimato e collaudato nel giugno 2008.

3. Sviluppi futuri

Allo stato attuale è in corso di approvazione una proposta di estensione dei servizi, da finanziare mediante il riutilizzo delle economie di gara, per le seguenti azioni:

- **Adeguamento alla GeoUML methodology**
Adeguamento delle nuove specifiche tecniche delle strutture dati e delle procedure del data base topografico (DBT5k) della regione Campania per il passaggio alla GeoUML methodology. Migrazione dei dati del DBT5k nella nuova struttura attraverso GeoUML methodology;
- **Derivazione del Data Base Topografico 25K dal DBT5k**
Riuso delle procedure applicative implementate dalle altre regioni nell'ambito del progetto PR5SIT (previo raffittimento e correzioni del grafo della viabilità secondaria e della viabilità primaria del DBT5k e della correzione delle sconessioni del grafo delle ferrovie e dei corsi d'acqua);
- **Servizio di semina dei punti in doppia coordinata dei punti fiduciali catastali delle altre province della regione Campania**
Completamento dei servizi di consultazione della banca dati catastale per tutti i comuni della regione Campania;

- **Servizi per il miglioramento delle procedure di apposizione/verifica dei vincoli alle aree incendiate per il catasto degli incendi boschivi da sviluppare mediante cooperazione applicativa con il Sistema Informativo della Montagna (SIM) realizzato dal Ministero delle Politiche Agricole e Forestali**
- **Servizi di raffittimento e riconfigurazione della rete di stazioni permanenti GPS e degli apparati del centro di controllo**
- **Servizi di monitoraggio dei programmi di Housing sociale e di riqualificazione urbana per le analisi di compatibilità con la pianificazione territoriale e per la valutazione delle ricadute sociali ed economiche e resi disponibili sul GeoPortale per la consultazione**
- **Servizi di monitoraggio per il contrasto del fenomeno dell'abusivismo edilizio, basato sui dati forniti dai segretari comunali, georeferenziati sulla base cartografica del DBT5k della regione Campania e resi disponibili sul GeoPortale per la consultazione.**

Analisi regionale per il riconoscimento dell'iride

Donato Impedovo, Giuseppe Pirlo, Lorenzo Scianatico

Università degli studi di Bari "A. Moro"

Via Orabona, 4, 70125 Bari

fimpedovo@gmail.com; giuseppe.pirlo@uniba.it; scianaticolorenzo@gmail.com}

Abstract. *Recent studies have explored a various set of techniques, methods and algorithms for iris recognition. Segmentation algorithms, normalization processes and well-defined distances measures are employed ad decision methods to compose the outline of all currently deployed systems, including commercial ones. So far, all system make use of the entire iris pattern of any individual, assuming that a fixed threshold applied on the distance can partially avoid problems like noise of information and acquisition errors. In this paper, a new approach is presented that uses only selected sectors of the iris and a simple analysis strategy, for verification purpose. Therefore, it differs form all previous approaches that uses the entire iris pattern. In addition, three Bayes-based similarity measures have been considered for iris image matching.*

The experimental results, carried out using the CASIA database, are encouraging and demonstrate that the proposed approach supports several directions for further research.

Keywords: Biometrics, personal verification, iris recognition.

1. Introduzione

Negli ultimi anni l'importanza della biometria è cresciuta notevolmente nella società, e sono stati realizzati differenti sistemi biometrici. Il vantaggio di questa tecnologia è la sua potenzialità nel poter sostituire codici PIN, password, chiavi o altri sistemi di identificazione simili, utilizzando tratti del corpo umano facilmente distinguibili. Se ne possono tracciare di due tipi: comportamentali o fisiologici. Il primo riguarda il modo che hanno le persone di portare a termine determinati compiti, come ad esempio, firmare, scrivere al pc o camminare. Il secondo è più specificatamente una misura presa sul corpo di una persona, come l'impronta digitale, il viso, o il DNA [Roberts, 2005].

L'occhio umano offre diverse possibilità di tratti biometrici utili alla verifica di un'identità, come retina, iride e sclera [Berggren, 1985; Roberts, 2005; Ostaff, 2005; Zhi et al., 2012]. La struttura dei vasi sanguigni della retina in particolare

può essere oggetto di analisi per un sistema biometrico, ottenendo ottimi risultati. Tuttavia, vanno considerati gli svantaggi di tali sistemi. Uno scanner per la retina può essere molto costoso, molto più di uno scanner per l'iride, e inoltre viene visto come meno "user-friendly" rispetto quest'ultima. Esistono comunque altri problemi. Alcune condizioni cliniche possono indurre il sistema in errore, come cataratte o gravi astigmatismi. Infine, lo strumento fisico dell'iris scanner viene visto dagli utenti come molto invasivo [Roberts, 2005; Ostaff, 2005; Zhi et al., 2012]. Per contro, l'iride risulta essere la parte dell'occhio più interessante e attraente per l'identificazione.

L'iride si trova al di sotto della cornea, protetta da agenti esterni, come polvere o pulviscolo, dalle ciglia e dalle palpebre. Al centro è forata dalla pupilla, che regola la quantità di luce che deve passare attraverso l'occhio, ed è caratterizzata oltre che dal colore più o meno raro o diffuso anche dalla sua struttura, altamente intricata e ricca di dettagli. Non ci si deve quindi sorprendere se per molti anni l'iride è stata giustamente presa in considerazione per effettuare una misura biometrica [Bergren, 1985].

Grazie alla ricerca, sono state realizzate numerose tecnologie di riconoscimento. Un primo esempio è I.R.I.S. (Iris Recognition Immigration System) usato nel Regno Unito per monitorare l'immigrazione, usato dal 2004 al 2011 [Keegan, 2012]. Google utilizza questa tecnologia per gestire gli accessi ad alcuni dei suoi datacenter [Google Inc., 2011]. La più interessante è il ritrovamento di Sharbat Gula, la celebre "Ragazza Afgana" fotografata da Mc.Curry per il National Geographic nel 1984 e ritrovata 18 anni dopo. Il sistema che ha riconosciuto la donna è stato realizzato da John Dougman. In futuro si pensa di adottare questa tecnologia su piattaforme web come Facebook o eBay.

Attualmente esistono numerosi approcci. Wildes [Wildes, 1997] usa per la segmentazione dell'immagine una tecnica basata sulla trasformata di Hough, algoritmo divenuto standard in visione artificiale per rilevare oggetti circolari. Successivamente utilizza una tecnica di registrazione d'immagine per allineare l'iride con quelle precedentemente ottenute. Utilizzando un filtro numerico, come un Laplaciano di una Gaussiana, ottiene una codifica dell'iride e il confronto avviene con la correlazione normalizzata.



Figura 1: Sharbat Gula, la "Ragazza Afgana"

Una strategia alternativa è stata proposta [Boles e Boashash, 1998]. Invece di una tecnica di registratura d'immagine, l'iride viene scalata e portata ad un diametro costante, stabilito rispetto un immagine di riferimento. La codifica viene effettuata prelevando i valori di intensità lungo dei cerchi concentrici. Successivamente viene applicato lo zero-crossing id una wavelet 1D. Oltre questa, [Zhu et al. 2000] è stata proposta come metrica per la valutazione della similarità di due iridi la distanza euclidea pesata:

$$WED(k) = \sum_{i=1}^n \frac{(f_i - f_i^k)^2}{\delta_i^k}$$

dove un'iride incognita viene confrontata con tutti gli esempi del database e il match viene trovato al valore minimo rispetto k. Il contributo più importante al riconoscimento dell'iride è stato dato da Daugman [Daugman, 2002]. Il suo sistema è la base di tutti i sistemi commerciali finora sviluppati. Innanzitutto, l'iride viene segmentata facendo uso del seguente operatore integro-differenziale [Mahboubeh et al., 2009]:

$$\max_{(r, x_0, y_0)} \left| G_\sigma(r) * \oint_{(r, x_0, y_0)} \frac{I(x, y)}{2\pi r} ds \right|$$

dove I è l'immagine e G è una funzione gaussiana di sfocatura. Dopo la normalizzazione, una versione 2D del filtro di Gabor viene usata per la codifica, mentre la distanza di Hamming viene usata per il confronto [Daugman, 2005; Vatsa et al. 2004]:

$$HD = \frac{1}{N} \sum_{i=1}^N X_i \otimes Y_i$$

Recentemente, [Liu et al. 2012] è stato presentato un approccio per la segmentazione basato sulla trasformata di Hough efficace per la determinazione della regione della pupilla e per il limite dell'iride. Inoltre consente di rilevare efficacemente le palpebre utilizzando la trasformata di Hough lineare.

In questo articolo viene presentato un nuovo approccio per il riconoscimento di iride. Invece di usare l'intero pattern dell'iride, l'immagine viene processata per trovare le regioni più stabili dell'iride. Quindi, verrà usato un ranking di queste regioni per valutare la stabilità dei vari settori di iride, di cui solo un certo numero verrà usato per la fase di confronto, utilizzando misure di similarità. Il lavoro è organizzato come segue: la sezione 2 riporta la descrizione del sistema. La sezione 3 presenta i risultati sperimentali. Le conclusioni e alcune considerazioni sulla ricerca futura sono discusse nella sezione 4.

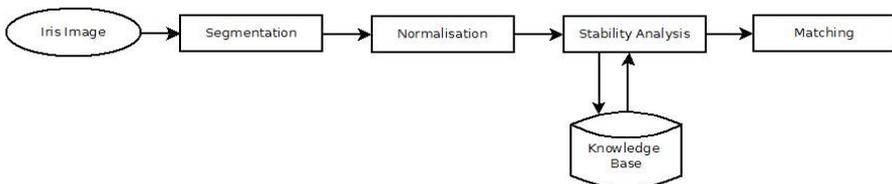


Figura 2: Schema del sistema implementato

2.Descrizione del sistema

Il sistema proposto è descritto schematicamente in figura 2. Questa sezione spiegherà ogni fase del sistema nel dettaglio. La sottosezione 2.1 mostra la procedura di segmentazione; la sottosezione 2.2 mostra il processo di normalizzazione di Dougman; la sottosezione 2.3 mostra come l'Ssim sia stato utilizzato per valutare la stabilità delle regioni; la sottosezione 2.4 descrive le funzioni di distanza utilizzate.

2.1 Segmentazione

Nella fase di segmentazione, la corona circolare dell'iride va isolata dal resto dell'immagine dell'occhio. A tal proposito è necessario localizzare la porzione dell'immagine situata tra il limbo (il bordo tra la sclera e l'iride) e la pupilla. Inoltre è importante includere la parte di iride non coperta dalle palpebre, superiori o inferiori. Infatti le palpebre possono coprire parte del limbo, mentre il bordo della pupilla in genere è sempre ben definito. Naturalmente quest'ultima è più scura dell'iride. In base alla pigmentazione della pelle e dell'iride, il contrasto con le palpebre può essere più o meno marcato. Le ciglia possono rendere il contorno più o meno irregolare.

Di conseguenza, un metodo per localizzare l'iride deve essere sensibile ad una grande variabilità di range del contrasto e deve essere robusto rispetto ai bordi irregolari.

Per questo lavoro, seguendo l'approccio di Wildes, per rilevare l'iride nell'immagine è stata usata la trasformata di Hough, seguendo l'equazione del cerchio:

$$x_c^2 + y_c^2 - r^2 = 0.$$

Questo comporta l'applicazione dell'algoritmo di Canny per generare i contorni.

La figura 3 mostra un esempio di segmentazione. Si possono applicare metodi per gestire palpebre e ciglia, per sopprimerli ed eliminarli dall'immagine. Le figure mostrano anche come sia necessaria una seconda trasformata per tracciare e rimuovere la pupilla dalle immagini [Masek, 2003]. Talvolta l'algoritmo può fallire nel tracciare correttamente i contorni sia dell'iride che della pupilla, come mostrato in figura 4.

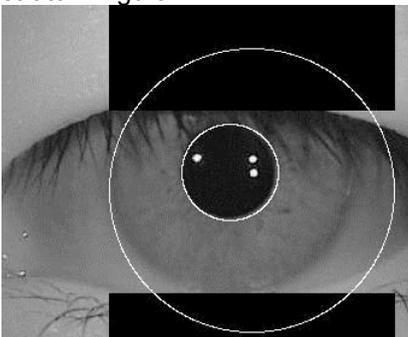


Figura 3: segmentazione corretta

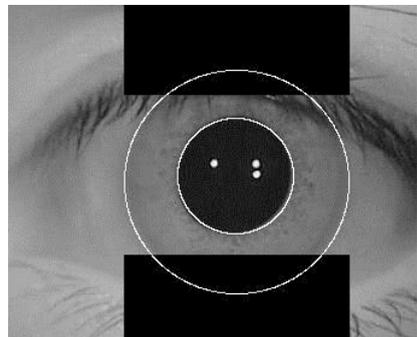


Figura 4: segmentazione errata

2.2 Normalizzazione

La normalizzazione mantiene l'immagine il più inalterata possibile e semplifica i passi successivi, trasformando l'immagine originale in un'altra immagine. Come descritto da Daugman [Daugman, 2002], il pattern deve essere invariante per dimensione e deve tenere conto della pupilla e della relativa posizione. Inoltre è necessario per la corretta esecuzione nel sistema di riconoscimento tener conto della trasformazione non-lineare a cui l'iride è sottoposta. Questo è un compito di importanza critica. Infatti, è necessario notare che la pupilla non è esattamente concentrica con l'iride, ma è leggermente scostata verso il naso. Ciò implica che un modello matematico che deve normalizzare l'iride deve considerare tutti questi aspetti. In letteratura sono noti diversi modelli [Boles e Boashash, 1998; Daugman, 2002]. Per questo lavoro verrà considerato il modello Rubbersheet di Daugman, rappresentato in figura 5, che di fatto è il modello più completo ed efficiente.

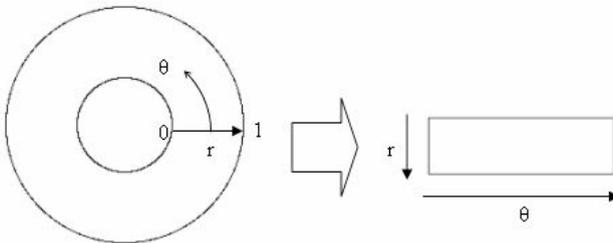


Figura 5: Il modello Rubbersheet di Daugman

Il modello usa un sistema di coordinate proiettate pseudo polari. In questo caso non è necessario che le griglie siano concentriche, considerando lo scostamento precedentemente descritto. L'angolo –la variabile polare- è adimensionale e la variabile radiale varia dal limite della pupilla al limbo. La dilatazione e il restringimento della pupilla vengono modellate da questo sistema di coordinate come il restringimento di un foglio di gomma con la topologia di una corona circolare ancorata sul perimetro esterno con tensione controllata da un anello più interno di raggio variabile. Con queste coordinate è possibile correggere la deformazione elastica del pattern dovuta alla variazione della dimensione della pupilla.

Il modello di Daugman rimappa ogni punto nella regione dell'iride in un sistema di coordinate polari. Questa operazione è modellata come:

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta)$$

avendo:

$$\begin{aligned} x(r, \theta) &= (1 - r)x_p(\theta) + rx_l(\theta) \\ y(r, \theta) &= (1 - r)y_p(\theta) + ry_l(\theta) \end{aligned}$$

dove:

- I è l'immagine dell'iride;
- (x, y) sono le coordinate cartesiane;
- (r, θ) sono le coordinate polari;

- (x_p, y_p) e (x_l, y_l) sono le coordinate del bordo della pupilla e dell'iride lungo la direzione dell'angolo θ .

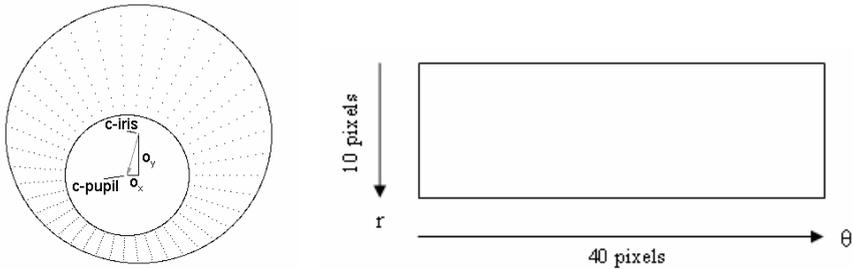


Figura 6: Lo scostamento della pupilla rispetto l'iride

Molto importante per il sistema è la modellazione della posizione della pupilla operata da Daugman. Questo dettaglio è mostrato in figura 6, dove lo scostamento è stato accentuato per motivi di chiarezza.

Infine, come mostrato in figura 7, questo modello consente il remapping di un'immagine su di un template di dimensioni prefissate, che è ottimo per il matching tramite la distanza di Hamming [Daugman, 2002] o con altre misure di distanza [Wildes, 1995; Boles e Boashash, 1998].

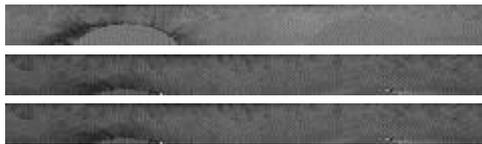


Figura 7: i pattern normalizzati

2.3 Indice SSIM per la selezione delle regioni

In letteratura il riconoscimento dell'iride è portato a termine facendo uso dell'intera immagine della stessa. L'idea proposta in questo lavoro considera ed usa effettivamente solo una parte dell'immagine più efficace per il riconoscimento. Per valutare la qualità dei settori dell'immagine viene preso in

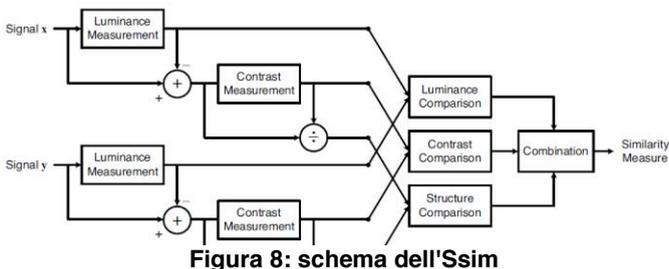


Figura 8: schema dell'Ssim

considerazione un indice statistico di basso livello [Wang et al. 2004]. Quest'indice è noto come Ssim, e sarà usato per valutare la similarità tra i settori dell'iride tra diverse immagini. La figura 8 illustra lo schema dell'indice Ssim, che valuta ogni pixel considerando la luminosità, il contrasto e la struttura.

La comparazione di questo indice è la seguente:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

Questa statistica è calcolata in un intorno di 8 pixel del punto considerato. L'immagine può essere traslata per ogni pixel dell'immagine ma gli autori hanno proposto di usare solo un sottogruppo di possibili finestre per ottimizzarlo. L'indice varia da -1 a 1, valore ottenuto solo da immagini perfettamente identiche. In questo lavoro l'indice viene usato per creare un rank di settori di iride. Più precisamente per ogni autore, l'iride viene divisa in 16 settori di uguale dimensione. Successivamente, per ogni settore viene calcolato il valore medio dell'Ssim, che verrà usato come punteggio. Si può notare che questo indice assegna un punteggio più basso ai settori che di solito sono coperti da ciglia o palpebre, mentre altri ricevono un punteggio alto. Di conseguenza, scegliendo un certo numero di settori, è possibile applicare una misura per la fase di matching.

2.4 Misure indotte dalla regola di decisione di Bayes

Di recente, sono state derivate dalla regola di decisione di Bayes nuove misure per il matching delle immagini, e la loro efficacia è stata mostrata in alcune specifiche biometrie [Liu, 2007]. Per poterle usare vanno precisate delle precondizioni:

- La probabilità condizionale delle funzioni di densità di tutte le classi è multivariata normale.
- La probabilità a priori di tutte le classi è uguale.
- La matrice di covarianza di tutte le classi è identica alla matrice di covarianza di tutti gli esempi senza considerare la classe di appartenenza.
- I vettori di pattern in rumore bianco sono normalizzati in norma unitaria.

La quarta assunzione è la più importante. Se questa non viene mantenuta, il classificatore diventa la distanza di Mahalanobis, che è leggermente meno potente di un altro classificatore, il Whitened Cosine, ottenibile rispettando le quattro condizioni:

$$\delta(U, V) = \frac{(W^t U)^t (W^t V)}{\|W^t U\| \|W^t V\|}$$

Con:

$$W = \Phi \Lambda^{-1/2}$$

Dove la matrice di trasformazione in rumore bianco è ottenuta fattorizzando con SVD la matrice di covarianza degli esempi $\Sigma = \Phi \Lambda \Phi'$.

L'autore del lavoro definisce due nuove misure tralasciando la terza assunzione e mantenendo la quarta.

Per la prima delle due nuove misure, la trasformata in rumore bianco è definita come:

$$W = \Phi \Delta^{-1/2}$$

$$\Delta = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_d^2\}$$

$$\sigma_i^2 = \frac{1}{L} \sum_{i=1}^L \left\{ \frac{1}{N_k - 1} \sum_{j=1}^{N_k} (y_{ij}^{(k)} - m_{ki})^2 \right\}$$

E prende il nome di PWC (PRM Whitened Cosine). La seconda misura è il WWC (Whitin-class Whitened Cosine), ottenuto fattorizzando sempre con SVD la within-class scatter matrix:

$$W_w = \Phi_w \Lambda_w^{-1/2}$$

Entrambe le due nuove trasformazioni in rumore bianco vengono applicate sulla similarità del coseno.

3. Risultati sperimentali

I risultati sperimentali sono stati ottenuti con immagini di iridi nella sezione ‘Lamp’ del database CASIA [CASIA, 2010]. Il processo di analisi di stabilità è stato usato per ordinare le varie regioni seguendo i ranking di stabilità assegnati ad ognuno. Nei test sono stati utilizzati i quattro settori migliori (rank più alto) per la fase di matching. Per questo motivo, le tre misure proposte (WC, PWC, WWC) e i risultati connessi sono stati valutati considerando le relative AUC (Area Under Curve) dei test ROC (Receiver Operating Characteristic). Le figure 9 e 10 riportano due esempi.

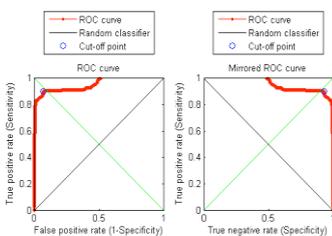


Figura 9: high performance

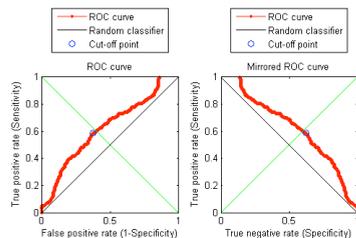


Figura 10: low performance

In tabella 1, 2 e 3 invece si riportano di valori di AUC, di FAR e FRR delle misure testate. Dai valori di curva AUC nella tabella 1 si nota immediatamente come il WWC nella maggior parte dei casi raggiunge un’area che statisticamente non è differente da 0.50, che indica valori simili a quelli di un classificatore casuale. Ciò significa che il WWC fallisce nel classificare correttamente i dati. Invece il WC e il PWC sono più efficaci e in qualche caso raggiungono buoni risultati di classificazione. Il FAR e l’FRR vanno comparati tra le due misure.

Autore	1	2	3	4	5	6	7	8	9	10
WC	0.68	0.89	0.95	0.71	0.62	0.61	0.81	0.57	0.77	0.50
PWC	0.62	0.89	0.95	0.71	0.63	0.61	0.82	0.58	0.76	0.50
WWC	0.50	0.53	0.99	0.50	0.50	0.63	0.53	0.52	0.51	0.50

Tabella 1: valori di curva AUC

La tabella 2 riporta i valori di FAR mentre la tabella 3 riporta i valori di FRR per WC e PWC. Come si può notare, le variazioni di valori suggeriscono che non ci sono differenze significative in termini di potenza discriminante tra le due misure. Questa considerazione è stata confermata da T-test che ha rivelato che non sono presenti differenze significative tra le due misure in termini di FAR e FRR.

Autore	1	2	3	4	5	6	7	8	9
WC	23.3	5.7	6.9	25.4	37.4	48.9	27.3	26.3	24.8
PWC	20.3	5.7	6.9	25.4	36.5	48.9	26.5	25.4	25.7

Tabella 2: valori FAR

Si può quindi concludere che il Whitened Cosine e il PRM Whitened Cosine hanno prestazioni simili nell'analisi regionale dei settori dell'iride. Nella sezione successiva verranno discussi i risultati insieme alle possibilità di sviluppi futuri.

Autore	1	2	3	4	5	6	7	8	9
WC	53.5	15.5	10.0	31.5	41.0	30.5	20.0	45.0	24.0
PWC	50.0	15.5	10.0	31.5	40.0	30.5	21.1	44.1	24.9

Tabella 3: valori FRR

4. Conclusioni

In questo lavoro è stato presentato un approccio regionale per il riconoscimento dell'iride. La tecnica proposta analizza le regioni più importanti, dipendenti dagli utenti, per effettuare il riconoscimento, tramite le tre misure Bayes-derivate per la fase di matching: il Whitened Cosine, il PRM Whitened Cosine e il Whitin-Class Whitened Cosine. I risultati sperimentali mostrano che tale approccio ha delle potenzialità, ma è comunque importante valutare possibili migliorie in tutte le fasi di riconoscimento per migliorare il più possibile i risultati, valutando l'uso di altri indici statistici per l'analisi di stabilità, tecniche di coregistrazione per allineare i template o altre distanze per il matching.

Bibliografia

Berggreen L., Iridology: A critical review, Acta Ophthalmologica, 63(1): pp. 1-8, 1985.

Boles W., Boashash B., A human identification technique using images of the iris and the wavelet transform, IEEE Transactions on Signal Processing, Vol. 46, No. 4, 1998, pp. 1185-1188.

CASIA Iris Image Database, available at: <http://biometrics.idealtest.org/> , 2010

Daugman J., How Iris Recognition Works, Proceedings of 2002 International Conference on Image Processing, Vol. 1, 2002, pp. 1-10.

Daugman J., Results from 200 billion iris cross-comparison, University of Cambridge, Computer Laboratory, UCAM-CL-TR-635, June 2005, pp. 1-8.

Google Inc., <http://www.youtube.com/watch?v=1SCZzqfdTBo> ,2011

Keegan M., Manchester Airport eye scanner scrapped over delays, <http://www.manchestereveningnews.co.uk/news/greater-manchester-news/manchester-airport-eye-scanners-scrapped-682100> , 2012

Liu C., The Bayes Decision Rule Induced Similarity Measures, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, June 2007, pp.1086-1090.

Liu C., Verma A., Jia J., Iris recognition based on robust iris segmentation and image enhancement, Int. J. Biometrics, Vol. 4, No. 1, 2012, pp. 56-76.

Mahboubeh S., Puteh B. S., Subariah B. I., Abdolreza R.K., Fast Algorithm for Iris Localization Using Daugman Circular Integro Differential Operator, Proc. of the 2009 International Conference of Soft Computing and Pattern Recognition (SOCPAR '09), 2009, pp. 393-398.

Masek L., Recognition of Human Iris Patterns for biometric Identification, BEng Thesis, The University of Western Australia, Western Australia, 2003.

Ostaff C., Retinal Scans Do More Than Let You In The Door, available at: <http://phys.org/news6134.html> , 2005

Roberts C., Biometrics, 2005

Vatsa M., Singh R., Gupta P., Comparison of Iris Recognition Algorithm, IEEE ICISIP, 2004, pp. 354-358.

Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P., Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing, vol. 13, no. 4, Apr. 2004, pp. 600-612.

Wildes R., Iris recognition: an emerging biometric technology, Proceedings of the IEEE, Vol. 85, No. 9, 1997, pp. 1348-1363.

Zhi Z., Du E. Y., Thomas N.L., Delp E. J., A New Human Identification Method: Sclera Recognition, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions, May 2012, Vol. 42, Issue 3, pp. 571-583.

Zhu Y., Tan T., Wang Y., Biometric personal identification based on iris patterns. Proceedings of the 15th International Conference on Pattern Recognition, Spain, Vol. 2, 2000, pp. 2801-2804.

Un sistema di autenticazione biometrica basato su tag RFID e riconoscimento del viso

Filippo Battaglia¹, Giancarlo Iannizzotto¹, Lucia Lo Bello²

¹Dipartimento delle scienze cognitive, della formazione e degli studi culturali
Università di Messina, Via Concezione 6-8, 98100 Messina
E-mail: filbattaglia@libero.it, ianni@unime.it

²Dipartimento di Ingegneria Elettrica, Elettronica ed Informatica (DIEEI)
Università di Catania, V.le A.Doria 6, 95125 Catania
E-mail: lucia.lobello@dieei.unict.it

Abstract. *This work proposes a Multi Factor Authentication system based on a dual-stage cascading classifier for biometrics identification (face recognition) and an RFID tag. A novel approach has been devised in order to automatically determine the thresholds for the two stages. The system records the biometric data of each user directly on his/her RFID tag and offers several benefits. First, as the biometrics data are stored in a database embedded in the RFID data, there is no need for a centralized database or server storing biometrics data and this solves a number of legal and technological issues relevant to data privacy. Second, the size of the embedded database used for authentication is significantly reduced, to the benefit of both speed and accuracy of the identification process. Finally, the proposed system is able to operate at very low resolution (down to 20x15 pixel) in near-real time and therefore it is suitable for application to restricted areas access control (being relatively cheap, it may be applied even to house access control), authentication for confidential services such as electronic payments, and homeland security. The system has been realized on a board for embedded systems. This paper describes the authentication system and discusses its performance in terms of accuracy and authentication time.*

Keywords: Biometrics, Security, Authentication, RFID tags.

1. Introduzione

I meccanismi di protezione d'accesso (o autenticazione) hanno lo scopo di assicurare che una determinata risorsa o servizio sia accessibile in ogni istante solo agli utenti autorizzati, e prevenire l'accesso da parte di impostori. Il più semplice meccanismo di autenticazione è, com'è noto, la digitazione di password [Jain et al, 2006], ma dover eseguire continuamente tale operazione per accedere a risorse o ad aree ad accesso limitato (edifici, uffici, strumenti), è noioso, inefficiente e soggetto a severe falle di sicurezza a causa della possibilità che la password venga letta da terzi mentre viene digitata o venga

Congresso Nazionale AICA 2013

involontariamente svelata. Anche l'impiego di *token* (badge, key), per quanto dotati di tecnologie molto sicure per la tutela dei dati in essi memorizzati (criptazione, chiavi pubbliche/private, etc.) è soggetto a falle di sicurezza, perché un token non è indissolubilmente legato al suo possessore e può essere usato da un impostore. Per tale ragione un considerevole sforzo è stato profuso negli anni per lo sviluppo di sistemi di autenticazione basati su dati biometrici, che hanno il vantaggio di essere sempre disponibili nel luogo in cui si trova l'utente da autorizzare e di essere difficili da falsificare. Sebbene esistano diverse caratteristiche che possono essere usate per il riconoscimento biometrico (impronta digitale, iride, geometria della mano, voce), il riconoscimento del viso offre il vantaggio di richiedere dei sensori poco costosi, di non richiedere il contatto fisico tra utente e sensore e di essere utilizzabile anche per scopi di videosorveglianza e nelle fasi non conclusive di indagine giudiziaria. Usate da sole, le tecniche di autenticazione biometrica hanno il problema di richiedere un database centralizzato o distribuito che contenga i dati biometrici di tutti gli utenti da autenticare, in modo da poter operare l'autenticazione tramite confronto fra le caratteristiche biometriche dell'utente e quelle memorizzate nel database. Inoltre non è ancora stato dimostrato che il riconoscimento facciale automatizzato possa raggiungere accuratezze del 100% per un database di dimensioni arbitrariamente grandi e questo pone dei dubbi sull'impiego di queste tecnologie. Un approccio recentemente molto investigato è quello dell'autenticazione multifattore (Multi Factor Authentication) che combina più approcci per ottenere risultati più affidabili. Ad esempio in passato sono state proposte alcune soluzioni che integrano tag RFID (quindi dei token a radiofrequenza) con sistemi di riconoscimento facciale [Min et al, 2011] [Jing et al, 2009] [Nguyen et al, 2012]. Tuttavia, esse memorizzano il set di pose dei soggetti autorizzati in un database centralizzato, e si limitano a memorizzare sul tag RFID solo poche informazioni come un codice identificativo del soggetto (identità dichiarata). Queste soluzioni pongono, oltre al problema delle dimensioni del database, anche problemi di natura legale: le autorità competenti nazionali e sovranazionali hanno più volte dichiarato che i dati biometrici sono dati personali e quindi il mantenimento di un database centralizzato presso il provider del servizio protetto non è in generale lecito, a meno che non esistano importanti e dimostrabili esigenze di sicurezza da preservare (principi di finalità, necessità e proporzionalità). Inoltre è stata dichiarata la preferenza per soluzioni *serverless*, in cui il dato biometrico viene memorizzato su un chip che l'utente autorizzato può portare sempre con sé [Art. 29 WP, 2003].

Il sistema presentato in questo articolo, chiamato **VisilabFaceRec**, integra la tecnologia di riconoscimento a radiofrequenza (RFID) con un sistema di riconoscimento facciale in modo da ottenere un sistema in cui:

- i dati biometrici dell'utente sono criptati e memorizzati nel suo badge RFID, quindi non ci sono database centralizzati nè database di dimensioni arbitrariamente grandi;
- gli algoritmi impiegati consentono un'accuratezza elevata sebbene i dati memorizzati rientrino nella capacità di un tag RFID industriale, riducendo enormemente costi e complessità architetturale rispetto

all'impiego di altre soluzioni dotate di database centralizzati.

Il principale contributo di questo lavoro riguarda l'accuratezza del sistema di autenticazione proposto, il calcolo automatico delle soglie di decisione per i due stadi in cascata ed il fatto di non ricorrere ad un database centralizzato, il che elimina i problemi di tutela della privacy. Il paper è organizzato come segue: la Sez. 2 descrive l'architettura di VisilabFaceRec, la Sez. 3 descrive i risultati sperimentali e la Sez. 4 riporta conclusioni e futuri sviluppi.

2. VisilabFaceRec

In VisilabFaceRec le pose del soggetto vengono memorizzate su un tag RFID. Quando l'utente si avvicina al varco di autenticazione il contenuto del chip viene letto; nello stesso momento, una webcam acquisisce l'immagine del viso e, se viene rilevato un livello di somiglianza sufficiente, il soggetto viene autenticato. Il sistema fa uso di un algoritmo consistente in due stadi di autenticazione in configurazione a cascata, ottimizzati per operare su immagini a risoluzione tanto bassa da poter essere salvate nelle memorie di piccola dimensione tipiche dei tag RFID (4-32 KB). Il primo stadio adotta l'algoritmo *2D-Principal Components Analysis* (2DPCA) [Yang et al, 2004] per una prima valutazione del livello di similarità del viso, mentre il secondo stadio impiega l'algoritmo *Scalar Image Feature Transform* (SIFT) [Lowe, 2004] per confermare il responso ottenuto dal primo.

2DPCA è un algoritmo di riconoscimento facciale di tipo *template-matching* che, dato un database di immagini (pose) appartenenti a soggetti (classi) differenti con k pose per soggetto, ed acquisita la posa di un soggetto di identità incognita, trova nel database l'identificatore della posa che è più simile (in base ad un'adeguata funzione di similitudine) al soggetto di identità incognita.

Data un'immagine I di $w \times h$ pixel del viso di un soggetto di identità incognita, è possibile associare ad essa un *template* costituito da un set di $Z_{2DPCA} \leq w$ componenti vettoriali (detti *vettori di decomposizione*):

$$(\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_{Z_{2DPCA}-1}) \quad Z_{2DPCA} \leq w \quad \mathbf{w}_s \in R^h \quad (1)$$

dove il numero di componenti Z_{2DPCA} viene scelto in fase di configurazione del sistema in base alla frazione della quantità di informazione dell'immagine originale che si vuole mantenere. In modo analogo è possibile eseguire una decomposizione 2DPCA su ciascuna delle m pose del database B ed è possibile determinare la posa più simile ad I in B (e quindi l'identità associabile alla posa in analisi I) per mezzo di un criterio di massima verosimiglianza applicato sui rispettivi template e basato sulla metrica definita in (2):

$$DIFS(\mathbf{I}, B) = d(\mathbf{I}, \mathbf{W}_i) = \sum_{s=0}^{Z_{2DPCA}-1} \|\mathbf{w}_s - (\mathbf{w}_i)_s\|_2 \quad i \in \{0 \dots (m-1)\} \quad (2)$$

in cui \mathbf{W}_i indica la i -esima posa del database B , \mathbf{w}_s è l' s -esimo vettore componente associato alla posa acquisita I , $(\mathbf{w}_i)_s$ è il s -esimo vettore componente associato alla posa \mathbf{W}_i di B . La (2) prende il nome di *distance in the feature space* (DIFS).

Per eseguire la decomposizione, in 2DPCA si moltiplica la matrice \mathbf{A} contenente le intensità dei pixel dell'immagine acquisita I per ciascuno degli

$Z_{2DPCA} \leq w$ autovettori principali della matrice di covarianza 2DPCA \mathbf{G}_t definita dalla (3):

$$\mathbf{G}_t = \frac{1}{m} \sum_{j=0}^{m-1} (\mathbf{A}_j - \bar{\mathbf{A}})^T (\mathbf{A}_j - \bar{\mathbf{A}}) \quad (3)$$

dove \mathbf{A}_j rappresenta la matrice contenente le intensità dei pixel della j -esima posa del database B. Se si opera con un database B costituito dalle pose del solo soggetto autorizzato ($h=1$; $m=k$), detto database delle pose principali del soggetto, è possibile verificare se l'immagine I acquisita appartenga o meno a quest'ultimo semplicemente stabilendo una soglia ρ di decisione sulla distanza DIFS (2). Un principio generale ben noto in letteratura [Flach, 2003] è che nei sistemi a soglia singola non è possibile minimizzare contemporaneamente sia il *false rejection rate* (FRR) che il *false acceptance rate* (FAR): la soglia di decisione ρ deve pertanto essere frutto di un compromesso. Un'ulteriore difficoltà è insita nel fatto che tale soglia dipende dalle pose del soggetto acquisito ovvero dal database B (per via della (3)). Nel nostro sistema si prevede che l'immagine acquisita I che supera il primo stadio venga passata ad un secondo stadio basato sull'algoritmo SIFT. Quest'ultimo appartiene alla categoria degli algoritmi *feature-matching*, che operano sulle caratteristiche locali dell'immagine. Scopo dell'algoritmo è determinare una serie di punti "di interesse" (*SIFT keypoint*) e definire un criterio di corrispondenza robusto ed affidabile tra di essi (criterio di Lowe) [Lowe, 2004]. Data l'immagine acquisita I e la posa più simile $I_{n_{pose}}$ determinata dallo stadio 2DPCA, lo stadio SIFT determina il numero di caratteristiche corrispondenti (*matching SIFT feature*) fra le due immagini e, se questo è superiore ad una soglia di decisione σ_{SIFT} , il soggetto viene autenticato, altrimenti viene respinto. Questo lavoro dimostra che l'architettura qui utilizzata consente di ottenere sia un basso FRR che un basso FAR.

2.1 La fase di enrollment

Quando un nuovo utente autorizzato viene iscritto al servizio di autenticazione, vengono acquisite $n_{enrolldb}$ pose del suo viso. Viene eseguita la decomposizione 2DPCA delle pose acquisite e successivamente applicato un algoritmo di clustering di tipo K-means che fa uso della metrica (2): il set delle pose viene così suddiviso in $n_{main,poses}$ cluster. Le pose più vicine ai centroidi dei cluster trovati vengono assunte come pose principali rappresentative dell'utente e registrate nel tag usando un algoritmo di compressione lossless (LS-JPEG).

Si noti che il nostro sistema, a differenza di altre soluzioni proposte, usa più pose per singolo soggetto, memorizzate a 256 toni di grigio. Nel tag sono memorizzate le pose e non i template 2DPCA poichè i vettori di decomposizione (1), scegliendo di operare con un numero elevato di componenti Z_{2DPCA} per non inficiare le prestazioni in riconoscimento, occuperebbero molto più spazio delle immagini originali. Pertanto, si preferisce memorizzare le pose in formato compresso e ricalcolare *on-the-fly* la matrice di covarianza (3) ed i rispettivi autovalori ed autovettori al momento della lettura del tag RFID. Tale ricalcolo non è particolarmente oneroso e non incide in maniera decisiva sulle prestazioni dell'algoritmo presentato.

Partendo dalle pose acquisite è necessario determinare una coppia ottimale di soglie (ρ_{2DPCA} , σ_{SIFT}) personalizzata per il soggetto. Il compito non è semplice, perchè il comportamento dello stadio 2DPCA influenza il comportamento dello stadio SIFT. Per risolvere tale problema VisilabFaceRec usa un algoritmo dedicato: non rientra negli scopi di questo lavoro dimostrare che il modello proposto è ottimale per il sistema aggregato costituito dai due stadi, ma l'adeguatezza delle prestazioni ottenute rispetto all'applicazione descritta è supportata dai risultati sperimentali mostrati in Sez. 3.

Per determinare la coppia (ρ_{2DPCA} , σ_{SIFT}) cercata si considera un *threshold optimization database* (TOdb) costituito da h_{TOdb} righe e k_{TOdb} colonne. Sulla prima riga si pongono le $k_{TOdb} = n_{enrolldb} - n_{main,poses}$ pose dell'utente autorizzato; sulle righe successive si pongono le k_{TOdb} pose di ($h_{TOdb}-1$) impostori, prelevate da un archivio di pubblico dominio come il database VidTIMIT [Sanderson, 2002]. Viene calcolata la matrice di covarianza (3) a partire dalle pose principali dell'utente per il quale si sta creando il tag, quindi viene simulato il comportamento dello stadio 2DPCA per ciascuna delle pose I_n (n in $[0..h_{TOdb}, k_{TOdb}-1]$) del TOdb. Per ciascuna posa viene registrata la distanza $DIFS(I_n, B)$ secondo 2DPCA (vedasi (2)) dal set delle pose principali B . Infine, il valore massimo della distanza d_{max} riscontrato durante l'elaborazione viene registrato. I dati acquisiti consentono di tracciare una curva detta **ROC (Receiver Operating Characteristic) primaria**. In un piano (FRR, FAR) la curva ROC primaria mostra come varierebbero il tasso di false rejection (FRR) ed il tasso di false acceptance (FAR) per un determinato valore di ρ (ρ in $[0..d_{max}]$) come soglia per il primo stadio (si noti che si ha sempre $FRR(\rho=d_{max})=0$ e $FAR(\rho=0)=0$). Come soglia ρ_{2DPCA} per il primo stadio viene quindi scelto il valore di ρ che minimizza la funzione di costo (4):

$$C(\rho) = C_{FRR} FRR(\rho) + C_{FAR} FAR(\rho) \quad (4)$$

dove C_{FAR} e C_{FRR} sono i costi assegnati euristicamente rispettivamente all'eventualità di false acceptance e di false rejection.

Una volta determinata la soglia ρ_{2DPCA} ottimale, viene rieseguita una scansione del threshold optimization database. Per ciascuna delle pose I_n per le quali $DIFS(I_n, B) < \rho_{2DPCA}$ vengono calcolate le SIFT feature e si verifica il numero s di feature che trovano corrispondenza nella posa $I_{n,pose}$ più simile tra le I_n nel database B delle pose principali. Infine, viene registrato il numero massimo s_{max} di SIFT feature corrispondenti rilevato su tutto il database. Possiamo a questo punto tracciare una seconda curva, che chiameremo **ROC secondaria**, che mostra come varierebbero i tassi di false rejection (FRR) e di false acceptance (FAR) del sistema aggregato se, per le pose I_n che passano il primo stadio in quanto $d < \rho_{2DPCA}$, imponessimo un certo valore di σ come soglia per il secondo stadio (con σ in $[0..s_{max}]$). La ROC secondaria è dunque una curva parametrica che descrive le prestazioni del sistema aggregato al variare di σ .

Per minimizzare contemporaneamente sia il FAR che l'FRR si è scelto il valore di σ che minimizza la funzione:

$$\sqrt{[FRR^*(\rho_{2DPCA}; \sigma)]^2 + [FAR^*(\rho_{2DPCA}; \sigma)]^2} \quad \sigma \in [0..s_{max}] \quad (5)$$

dove **FRR*** e **FAR*** sono i valori di FRR e FAR normalizzati tra 0 ed 1. La

coppia di soglie (ρ_{2DPCA} , σ_{SIFT}) così determinata viene memorizzata nell'immagine binaria del tag: quest'ultima viene infine crittografata con l'algoritmo AES-128 e registrata in un chip RFID ad alta capacità (4-32 KB).

2.2 La fase di autenticazione

In fase di autenticazione il soggetto di identità incognita si avvicina al varco: una videocamera ne acquisisce l'immagine frontale. L'area contenente il viso viene isolata tramite un algoritmo di face detection [Viola e Jones, 2001] e sottoposta a *scaling* in modo da normalizzarne le dimensioni. Quindi viene applicato un compensatore della luminosità [Chen et al, 2006] e poi le intensità dei pixel dell'immagine che sono associate al background vengono poste a 0, ottenendo così l'immagine I da passare ai due stadi 2DPCA/SIFT. Parallelamente, una routine legge e decrittografa il database B delle pose principali del soggetto contenuto del tag RFID e calcola la matrice di covarianza (3).

3. Misura delle prestazioni

Per valutare le prestazioni in termini di FAR e FRR dell'approccio proposto è stata impiegata la procedura "A priori performance type A" basata sul database VidTIMIT e proposta in [Sanderson, 2002]. Il database VidTIMIT fornisce un set di pose appartenenti a 43 soggetti, distinti in due gruppi: 35 soggetti autorizzati (*true claimants*) ed 8 impostori (*impostors*). I test sono stati eseguiti su 2 configurazioni diverse, caratterizzate da 2 e 4 pose per ciascun utente da autenticare, rispettivamente. Per ciascuna configurazione le pose di ogni soggetto sono divise in 3 sessioni, la prima delle quali viene usata per creare un modello del soggetto. Nel sistema in esame, questo gruppo di pose è stato usato per determinare un set di $n_{\text{main,poses}}=2,4$ pose (a seconda della configurazione di test) secondo la procedura di clustering descritta in Sez. 2.1.

La seconda e la terza sessione del db VidTIMIT vengono invece usate per una procedura in 2 fasi (A e B). Durante la fase A, la seconda sessione viene usata per la procedura di *threshold optimization* (che determina la coppia di soglie (ρ_{2DPCA} , σ_{SIFT}) sulla base del criterio descritto in Sez. 2.1), mentre la terza sessione viene usata per la procedura di testing, che determina le prestazioni del sistema (FRR e FAR) con le soglie scelte. Durante la fase B, i ruoli della seconda e terza sessione vengono scambiati. Le misure di FAR ed FRR ottenute durante le fasi A e B per ogni singolo soggetto vengono mediate.

Alla fine dell'elaborazione si determinano i valori avgFRR e avgFAR come media dei valori ottenuti per tutti i soggetti; questi valori costituiscono una stima delle prestazioni del sistema a quella risoluzione e per quella configurazione. È possibile mostrare che con questa procedura, per ciascuno dei 35 true claimants si simulano 204 tentativi di autenticazione autorizzati e 1632 tentativi da parte di impostori.

I test prestazionali sono stati eseguiti alle risoluzioni di 20x15, 40x30 e 80x60 e 160x120 pixel; il numero di componenti 2DPCA scelto per il primo stadio è $Z_{2DPCA}=w$ (il massimo possibile in base alla teoria). Prima di iniziare, è stato necessario determinare il valore del rapporto C_{FAR}/C_{FRR} (vedasi (4)) che assicura le migliori prestazioni. Il risultato è mostrato in Fig.1, dove si è assunto

che tale rapporto possa essere determinato per la risoluzione target di 80x60 pixel e successivamente applicato anche a risoluzioni inferiori.

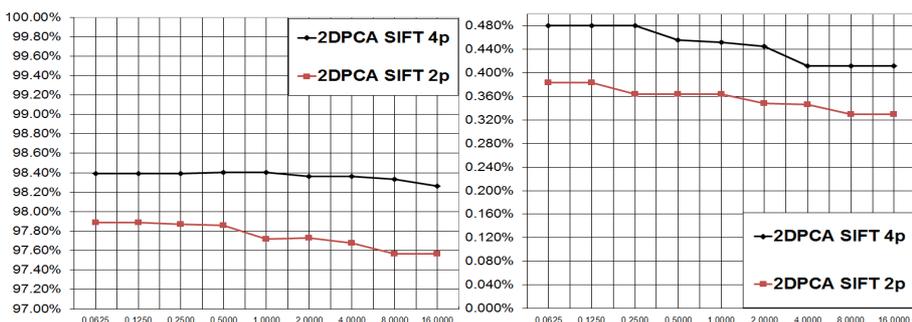


Fig. 1 – Recognition rate medio (a sinistra) e false acceptance rate medio (a destra) al variare del rapporto C_{FAR}/C_{FRR} . Risoluzione 80x60 pixel.

Allo scopo di scegliere il rapporto C_{FAR}/C_{FRR} che garantisce le prestazioni migliori è stato applicato un metodo che consiste nel riportare in un piano $[0;1] \times [0;1]$ i valori di FAR^* e di FRR^* normalizzati ad 1. In questo modo, per ciascun valore del rapporto suddetto, si ottiene un punto che rappresenta le prestazioni ottenibili. L'obiettivo è minimizzare sia il FAR che l'FRR: a tale scopo si sceglie il rapporto C_{FAR}/C_{FRR} che si trova più vicino all'origine degli assi nel piano normalizzato. Come è mostrato in Fig.2, per le configurazioni a 2 e 4 pose il rapporto ottimale è pari rispettivamente a $C_{FAR}/C_{FRR}=2$ e 4. Questo dato è comunque dipendente dal numero di pose: per valutare le prestazioni del sistema al variare della risoluzione, è stato scelto un valore medio di C_{FAR}/C_{FRR} pari a 3, assumendo che esso possa fornire buone prestazioni per tutte le configurazioni.

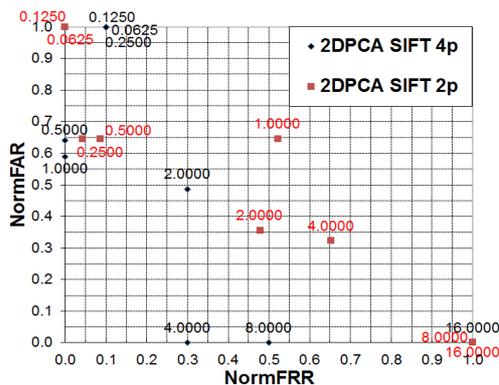


Fig. 2 – False rejection rate e false acceptance rate normalizzati in $[0;1]$ al variare del rapporto C_{FAR}/C_{FRR} . Risoluzione 80x60 pixel.

La Fig. 3 mostra le prestazioni del sistema al variare della risoluzione di lavoro per le configurazioni a 2 e 4 pose per tag.

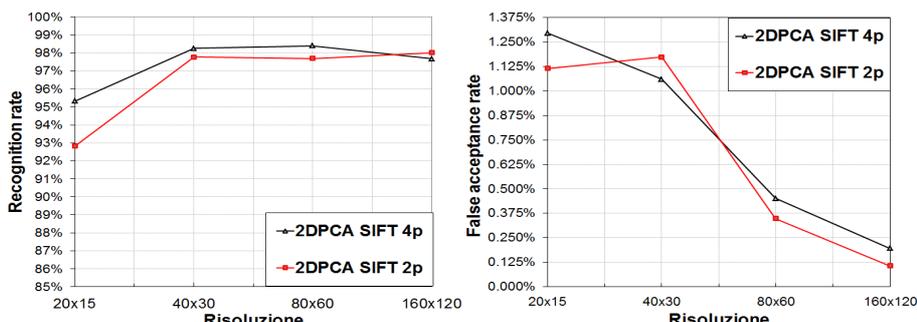


Fig. 3 – Recognition rate medio (a sinistra) e false acceptance rate medio (a destra) al variare della risoluzione di lavoro (configurazione 2 e 4 pose per tag).

Il sistema proposto assicura un RR>97.69% per tutte le risoluzioni maggiori di 20x15 pixel ed un FAR<0.45% per risoluzioni maggiori o uguali a 80x60 pixel.

Nell'architettura a due stadi lo stadio SIFT consente di mantenere basso il FAR pur mantenendo elevato il recognition rate RR (si ricordi che RR=1-FRR). Per misurare quale sia il vantaggio effettivo in termini di riduzione dei FAR a parità di RR, è stato elaborato il seguente esperimento. Per ciascuno dei soggetti del database VidTIMIT, viene considerata la coppia di valori ($exFRR^{dblstage}$, $exFAR^{dblstage}$) (expected FRR, expected FAR) determinata dal sistema a doppio stadio sulla base delle curve ROC primaria e secondaria descritte in Sez. 2.1. Si considera ora la curva ROC primaria relativa al solo stadio 2DPCA, e si sceglie una nuova soglia $\rho_{2DPCA}^{onestage}$ tale che:

$$exFRR^{onestage}(\rho_{2DPCA}^{onestage}) = exFRR^{dblstage}(\rho_{2DPCA}^{twostage}, \sigma_{SIFT}^{twostage}) \quad (6)$$

In altri termini, si configura un sistema a singolo stadio 2DPCA in modo da ottenere il medesimo *expected recognition rate* ($exRR = 1 - exFRR$) del sistema a doppio stadio. A questo punto si esegue una simulazione sul testing database in modo da determinare il RR ed il FAR reali del sistema a singolo stadio. Il risultato è illustrato in Fig. 4.

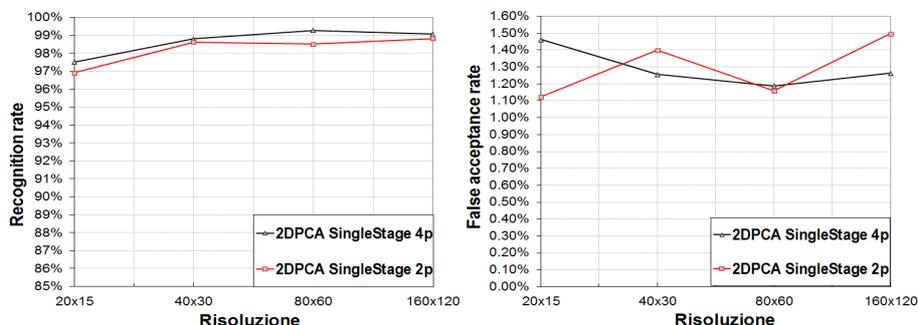


Fig. 4 – Recognition rate medio (a sinistra) e false acceptance rate medio (a destra) al variare della risoluzione di lavoro per i sistemi a singolo stadio 2DPCA.

Come prevedibile, la curva che descrive l'andamento del recognition rate del sistema a singolo stadio al variare della risoluzione di lavoro appare molto

simile a quello del rispettivo sistema a doppio stadio, a causa della (6). Il recognition rate del sistema a singolo stadio è anzi leggermente superiore, perchè nel rispettivo sistema a doppio stadio, il classificatore SIFT respinge non solo molti impostori (diminuendo il FAR), ma anche alcune pose dei soggetti autorizzati (aumentando il FRR ossia diminuendo il RR). Il vantaggio del sistema a doppio stadio rispetto a quello a singolo stadio si vede dalla curva dei FAR al variare della risoluzione: il singolo stadio mostra un tasso di false acceptance significativamente più elevato per tutte le risoluzioni maggiori di 20x15 pixel (al di sotto di tale risoluzione il secondo stadio non riesce ad operare) e ciò dimostra la correttezza del nostro approccio.

		4p 20x15	4p 40x30	4p 80x60	4p 160x120	2p 20x15	2p 40x30	2p 80x60	2p 160x120
Tempo di lettura dei tag RFID t_{rfid} (s)									
Necessario per la lettura del tag RFID, decrittazione e decompressione delle immagini principali	Avg	1.0072	3.6983	12.7322	50.1272	0.5517	1.9477	6.4282	25.1802
	StdDev	0.0787	0.2021	0.6568	2.5363	0.0395	0.1156	0.3333	1.2846
Tempo di rilevamento del viso t_{fdet} (s)									
Necessario per il rilevamento della presenza di un viso tramite detector Viola-Jones	Avg	0.0605	0.0650	0.0638	0.0640	0.0605	0.0650	0.0638	0.0640
	StdDev	0.0062	0.0067	0.0071	0.0067	0.0062	0.0067	0.0071	0.0067
Tempo di ricostruzione del face space t_{space} (s)									
Necessario per il calcolo della matrice di covarianza $G(3)$, degli autovalori, degli autovettori e dei vettori di decomposizione 2DPCA delle pose principali (1)	Avg	0.1208	0.1593	0.2549	0.5812	0.1129	0.1320	0.1887	0.3940
	StdDev	0.0066	0.0093	0.0172	0.0540	0.0048	0.0064	0.0178	0.0787
Tempo di riconoscimento del viso t_{frec} (s)									
Necessario perché il sistema esegua la decomposizione 2DPCA dell'immagine I acquisita seguita dal confronto con le pose principali (2)	Avg	0.0002	0.0007	0.0031	0.0133	0.0002	0.0007	0.0028	0.0122
	StdDev	0.0000	0.0000	0.0001	0.0004	0.0000	0.0000	0.0001	0.0004
Tempo di calcolo caratteristiche SIFT t_{sift} (s)									
Necessario per calcolo delle caratteristiche SIFT dell'immagine I acquisita dalla videocamera seguita dall'applicazione del criterio di Lowe	Avg	0.0060	0.0147	0.0393	0.1088	0.0057	0.0143	0.0397	0.1097
	StdDev	0.0014	0.0019	0.0059	0.0218	0.0013	0.0018	0.0059	0.0214
$t_{min} = t_{rfid} + t_{fdet} + t_{space} + t_{frec} + t_{sift}$		1.1948	3.9382	13.0933	50.8946	0.7309	2.1597	6.7232	25.7601
$t_{min}^* = t_{fdet} + t_{space} + t_{frec} + t_{sift}$		0.1876	0.2398	0.3611	0.7674	0.1793	0.2120	0.2950	0.5799

Fig. 5 – Tempi operativi del sistema di autenticazione VisilabFaceRec (in sec.)

La Fig. 5 riporta il tempo di riconoscimento minimo t_{min} del sistema di autenticazione proposto. I valori sono stati misurati su una board ASRock E350M1 dotata di processore Amd Fusion dual core a 1,6 Ghz. Alla risoluzione di 80x60 pixel nella configurazione a 4 o 2 pose per tag, per il riconoscimento è necessario un tempo minimo rispettivamente di 13.09 s e 6.72 s. Questi valori, tuttavia, sono comprensivi del tempo t_{rfid} ; una volta eseguita la lettura del tag, se il riconoscimento sul primo frame fallisce, quello sui successivi viene eseguito in tempi $t_{min}^* = t_{min} - t_{rfid}$ molto più bassi (0.36 s e 0.29 s rispettivamente). Questi valori, inferiori ad 1 s, e la loro deviazione standard molto ridotta suggeriscono che l'algoritmo da noi sviluppato possa essere usato anche per applicazioni in near- real time.

4. Conclusioni

Questo lavoro ha presentato VisilabFaceRec, un sistema multi fattore per la autenticazione degli utenti per il controllo di accesso a servizi ed aree riservate, che combina tag RFID e riconoscimento biometrico (face recognition). Il sistema è specificatamente pensato per lavorare a bassissima risoluzione e questo consente di archiviare i dati sensibili dell'utente (immagine del viso) direttamente nel tag RFID senza la necessità di un database centralizzato.

Per quanto a nostra conoscenza non esistono attualmente altri sistemi che ottengano risultati simili in termini di corretto bilanciamento FAR/FRR a parità di risoluzione e di vincoli operativi (impiego di RFID per l'archiviazione dei dati

sensibili). Il sistema proposto è stato realizzato e testato su una board per sistemi embedded ed ha evidenziato tempi di esecuzione abbastanza bassi da consentirne l'applicazione in casi quali il controllo di accesso ad aree riservate, per i quali non è pensabile avere tempi di autenticazione elevati o richiedere hardware costoso.

Ulteriori sviluppi comprenderanno un ulteriore miglioramento del sistema di calcolo delle due soglie di decisione dei due stadi in cascata e la sperimentazione di altri algoritmi per i due stadi.

Bibliografia

[Jain et al, 2006] Jain A.K., Ross A., Pankanti S., Biometrics: A Tool for Information Security, IEEE Transaction on Information Forensics and Security, 1, 2, 2006.

[Yang et al, 2004] Yang J., Zhang D., Frangi A.F., Yang J.Y., "Two-dimensional PCA: a new approach to appearance-based face representation and recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, 26, 1, 2004, 131-137.

[Lowe, 2004] Lowe D., Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60, 2004, 91-110.

[Min et al, 2011] Min D.G., Kim J.W., Jun J.S., The Entrance Authentication System in Real-Time using Face Extraction and the RFID Tag, Intl. Conf. on Ubiquitous Computing and Multimedia Applications, 2011, 20-24.

[Jing et al, 2009] Jing B.Z., Yeung D.S., Ng W.W.Y, Ding H.L., Wu D.L., Wang Q.C., Li J.C., RFID Access authorization by face recognition, Eighth Intl. Conf. on Machine Learning and Cybernetics, Baoding, 2009, 302-307.

[Nguyen et al, 2012] Nguyen T.D., Quang L.D., Van N.C., Thanh L.T., Hoang T. M., De Souza-Daw T., An efficient and reliable human resource management system based on a hybrid of face authentication and RFID technology, Fourth Intl. Conf. on Communications and Electronics (ICCE) 2012, 333-338.

[Art. 29 WP, 2003]. Documento di lavoro sulla biometria 12168/02/IT WP 80. Available online: <http://www.garanteprivacy.it/garante/document?ID=443868>

[Flach, 2003] Flach P.A., The Geometry of ROC Space: Understanding Machine Learning Metrics through ROC Isometrics, Twentieth Intl. Conf. on Machine Learning, 2003, 194-201.

[Sanderson, 2002] Sanderson C., The VidTIMIT Database, Available online at: <http://publications.idiap.ch/index.php/publications/show/710>

[Viola e Jones, 2001] Viola P., Jones M., Rapid Object Detection Using a Boosted Cascade of Simple Features, IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, 2001, 511-518.

[Chen et al, 2006] Chen W., Er M.J., Wu S., Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain, IEEE Transaction on Systems, Man, and Cybernetics, 36, 2, 2006, 458-466.

Local Stability of Static Signatures using Optical Flow Analysis

G. Pirlo, D. Impedovo

Dipartimento di Informatica, Università degli Studi di Bari "Aldo Moro"

Via Orabona, 4 – 70125 Bari

giuseppe.pirlo@uniba.it

In this paper optical flow is used to obtain information on local stability in static signature images. The approach is useful both for extracting information about personal behaviour in the signature apposition process and for performing signature verification. Processes by a C4.5 Decision Tree Classifier. The experimental results, carried out on the signatures of the GPDS database, demonstrate the effectiveness of the new approach.

Index Terms— *Biometrics, Static Signature Analysis, Optical Flow.*

1. Introduction

Handwritten signature is one of the most interesting biometric traits since people are familiar with the use of signatures in their daily life. In addition, signature is generally well recognized as a legal means of verifying an individual and it is well-accepted by administrative and financial institutions [Impedovo and Pirlo, 2007].

Depending on the data acquisition method, two categories of signature verification systems can be identified in the literature [Impedovo and Pirlo, 2007]: dynamic (on-line) and static (off-line) systems. Dynamic systems use on-line acquisition devices that generate electronic signals representative of the signature during the writing process. In this case dynamic systems involve the treatment of a spatio-temporal representation of the signature. Of course, the number of possible applications for on-line signature verification is continuously growing. Examples can be found in health care applications, distributed working communities, online banking, monetary transactions and retail Points-Of-Sale (POS) [Vielhauer, 2005].

Static systems perform data acquisition after the writing process has been completed. The results is that static systems involve the treatment of the spatio-luminance representation of a signature image, in which no dynamic information is available. Static signature verification is very important for many application fields, like automatic bank-check processing, insurance form processing, document validation and so on [Vielhauer, 2005].

Unfortunately, a handwritten signature is a complex pattern that originates from a ballistic-like movement depending on the psychophysical state of the signer and the conditions under which the signature apposition process occurs. Variability of handwritten signatures – even if they are generated by the same signer - is one of the most evident effects of the complexity of signature generation and implementation processes [Djioua and Plamondon, 2009].

When dynamic signatures are considered, local stability can be estimated using Dynamic Time Warping (DTW) to match a genuine signature against other authentic specimens [Pirlo, 1994; Impedovo and Pirlo, 2007]. In this case Direct Matching Points (DMPs), that are unambiguously matched points of the genuine signature, are used to indicate the presence of a small stable region of the signature, since no significant distortion has been locally detected. The local stability of a point of a signature is determined as the average number of time it is a DMP, when the signature is matched against other genuine signatures. This procedure allows to detect low- and high-stability regions [Huang and Yan, 2003; Dimauro et al., 2002; Dimauro et al., 2004] and used for selecting the reference signatures [Congedo G. et al., 1995; Di Lecce et al., 1999] and improving verification strategies [Di Lecce et al., 2000].

A client-entropy measure based on local density estimation by a HMM, was used to group and characterize dynamic signatures in categories that can be related to signature variability [Garcia-Salicetti et al., 2008; Houmani et al., 2009].

Other approaches estimate the stability of a set of common features and the physical characteristics of signatures which they are most related to, in order to obtain global information on signature repeatability which can be used to improve the verification systems [Impedovo et al., 2009; Impedovo et al., 2011]. In general, these approaches have shown that there is a set of features that remain stable over long time periods, while there are other features which change significantly in time [Schomaker and Plamondon, 1990; Impedovo et al., 2009]. This aspect is very useful to determine parameter-updating approaches to avoid deterioration of verification performances in time [Guest and Fairhurst, 2006; Impedovo and Pirlo, 2010].

Concerning static signatures, pattern-matching [Impedovo and Pirlo, 2010, Impedovo and Pirlo, 2011] and cosine-distance strategy [Impedovo et al., 2012] have been recently proposed to determine - at local level - the degree of stability of each region of a signature.

In this paper the optical flow was considered as a tool useful to estimate the local stability of the signature images. The basic idea of the approach is that each signer has its own ideal model of signature and all specimens he/she generates can be considered as distorted versions of the unique ideal model.

2. The Optical Flow

In the literature, optical flow is generally defined as the distribution of apparent velocities of movement of brightness patterns in an image I [Fleet and Jepson, 1990, O'Donovan, 2005]. In the past, several approaches have been considered for computing the optical flow, based on differential techniques,

region-based matching [Horn and Schunck, 1980; Lucas and Kanade, 1981; Nagel, 1989], energy-based [Singh, 1992; Bigun et al., 1991] and phased-based techniques [Fleet and Jepson, 1990].

In this paper optical flow is applied to the analysis of regional stability of static signatures. Let $\{I_n^g \mid n=1,2,\dots,N\}$ be N genuine signatures of a writer. After normalization to a fixed rectangular area we compute

$$[u_{ij}^{gg}(x,y), v_{ij}^{gg}(x,y)]^T, \quad i,j=1,2,\dots,N, i \neq j \quad (1)$$

that is the optical flow between I_n^g and I_j^g .

In particular the Horn and Shunck approach [Horn and Schunck, 1980] is considered for optical flow estimation, that determined through the minimization of the energy functional [Impedovo and Pirlo, 2011]:

$$E = \iint \left[(I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2) \right] dx dy \quad (2)$$

where

- I_x, I_y and I_t are the derivatives of the image intensity values along the x, y and time dimensions, respectively;
- $[u_{ij}(x,y), v_{ij}(x,y)]^T$ is the optical flow vector;
- α is the regularization parameter.

Now, for each pixel (x,y) , the sets of the horizontal and vertical components of the optical flow are defined as:

$$U(x,y) = \{u_{ij}(x,y) \mid i, j=1,2,\dots,N; i \neq j\} \quad (3a)$$

$$V(x,y) = \{v_{ij}(x,y) \mid i, j=1,2,\dots,N; i \neq j\} \quad (3b)$$

The local stability at point (x,y) of the genuine signatures of the writer can be estimated as [Impedovo and Pirlo, 2011]:

$$S(x, y) = \sqrt{\sigma_U(x, y)^2 + \sigma_V(x, y)^2} \quad (4)$$

being $\sigma_U(x,y)$ and $\sigma_V(x,y)$ the standard deviation of $U(x,y)$ and $V(x,y)$, respectively.

3. The Signature Verification System by Optical Flow

Local stability is here used to support a Signature Verification System. The system exploits the consideration that some variations are typical of the signer whereas other variations are not. Therefore, in the verification process, local variability is considered as a powerful feature typical of the user and anomalous local variability is considered as the evidence of a forged specimen. A C4.5 classifier was used for signature

classification [John et al., 2004; Quinlan, 1993]. Let be:

- I_n^g the set of N genuine signatures of a writer, $n=1,2,\dots,N$;
- I_m^f the set of M forgeries, $m=1,2,\dots,M$.

In the enrollment stage the classifier is trained by using the optical flow vectors concerning intra-class and inter-class variability:

- $[u_{ij}^{gg}(x,y), v_{ij}^{gg}(x,y)]^T$: optical flows between I_i^g and I_j^g , $i,j=1,2,\dots,N$, $i \neq j$ (intra-class variability);
- $[u_{ij}^{gf}(x,y), v_{ij}^{gf}(x,y)]^T$: optical flow between I_i^g and I_j^f , $i=1,2,\dots,N$, $j=1,2,\dots,M$ (inter-class variability).

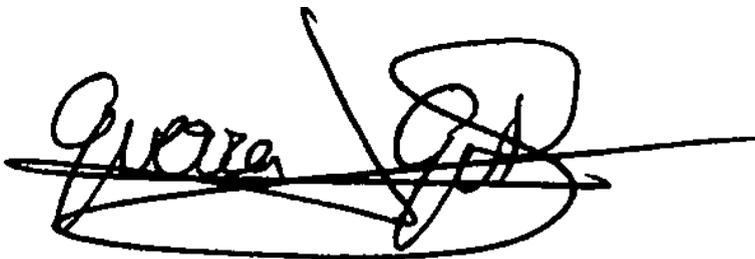
In the testing stage, the optical flows $[u_i^t(x,y), v_i^t(x,y)]^T$ between the test signature I_t and each genuine signature I_i^g , $i=1,2,\dots,N$, are computed. Each one of the N optical flows is provided to the classifier that returns a verification results $r_i(t)$, $i=1,2,\dots,N$. The N results are finally combined according to a majority vote strategy, in order to define the final verification result for the test signature.

4.Experimental Results

The GPDS gatabase was considered for the test [Vargas et al., 2007]. The database contains 16200 signatures from 300 individuals. Figures 1a and 1b present two genuine signature of the same signer, whereas Figure 2 presents the optical flow computed between the two signature images.



(a).



(b)

Fig. 1. Examples of two genuine signatures

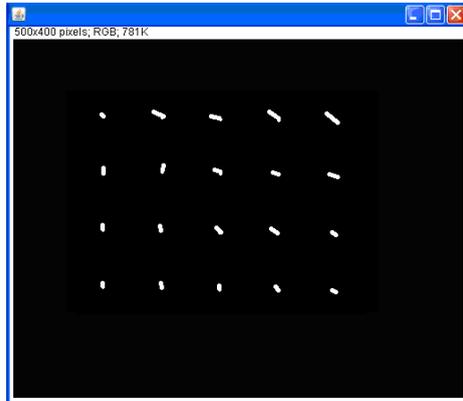


Fig. 2. Optical Flow: Genuine vs Genuine

In order to estimate local stability in handwritten signatures, $N=5$ genuine signatures ($I_i^g, i=1, \dots, 5$) have been considered for each signer and the local stability has been computed according to eq. (4).



Fig. 2. Stability Analysis: ■ High, ■ Medium, ■ Low

Figure 2 shows the result of stability analysis for the signature in Figure 1. In this case three different levels of stability have been considered: low, medium, high. It is easy to verify that pictorial patterns are generally more variable than character-based patterns. In general, the initial part (left) of the signature is more stable than the central and the final part (right). This evidence indicates that as the process of signature apposition proceeds, the signers are generally less and less stable.

In the second part of the experimental test, the use of the optical flow for automatic signature verification is analysed. For each signer 20 genuine signatures and 20 forgeries were considered for the test. In the enrollment stage, $N=6$ genuine signatures ($I_i^g, i=1, \dots, 6$) and $M=6$ forgeries ($I_i^f, i=1, \dots, 6$) were considered for training.

Table I reports the performance of the system in terms of Equal Error Rate (EER) compared to other approaches in the literature, carried out on the same

database. The result demonstrates the effectiveness of the proposed approach when compared to other techniques in literature. In fact, the equal error rate of the proposed approach on the GPDS database is equal to 5.6%, that is better than other results reported the literature, on the same dataset.

TABLE I. Performance

Approaches	EER
[Yilmaz et al., 2011]	15.4%
[Vargas et al., 2009]	6.2%
[Vergas et al., 2011]	18.9%
[Aguilar et al., 2004]	20.2%
<i>This approach</i>	5.6%

5. Experimental Results

In this paper optical flow is considered as a tool for static signature analysis and verification. In the first part of the paper local stability in static signatures is analyzed by optical flow analysis. In the second part, optical flow vectors among between the test signature and the genuine/forged specimens are considered to verify the authenticity of a test signature, using an Alternate Decision Tree. Some results carried out on static signatures of the GPDS database demonstrate the new approach is worth consideration for further research. In particular, further research is necessary to verify the capability of the Optical Flow in recognizing short-term and long-term variability as well as for evaluating the extent to which local stability depends on the signature type and signer characteristics.

References

- Aguilar, J. F. , Hermira, N. A. , Marquez G. M. , Garcia, J. O. , An offline signature verification system based of local and global information, LCNS 3087, 2004, 295-306.
- Bigun, J. , Granlund, G. , Wiklund, J., Multidimensional orientation estimation with applications to texture analysis and optical flow, IEEE Trans. PAMI, 13, 1991, 775-790.
- Congedo, G., Dimauro, G., Forte, A.M., Impedovo, S., Pirlo, G. , Selecting Reference Signatures for On-Line Signature Verification, Proc. of the 8th International Conference on Image Analysis and Processing (ICIAP-8), Series: Lecture Notes in Computer Science, 974, Springer-Verlag Berlin, Heidelberg, C. Braccini, L. De Floriani and G. Vernazza (Eds.), San Remo, Italy, Sept. 1995, 521-526.
- Di Lecce, V., Dimauro, G., Guerriero, Impedovo, S., Pirlo, G., Salzo, A., Sarcinella, L., Selection of Reference Signatures for Automatic Signature Verification, Proc. of the 5th International Conference on Document Analysis and Recognition (ICDAR-5), Bangalore, India, 1999, 597-600.
- Di Lecce, V., Dimauro, G., Guerriero, A., Impedovo, S., Pirlo, G., Salzo, A., A Multi-Expert System for Dynamic Signature Verification, Proc. of the 1st International Workshop, Multiple Classifier Systems (MCS 2000), Series: Lecture Notes in Computer

Science, Springer-Verlag Berlin Heidelberg, J. Kittler and F. Roli (Eds.), Vol. 1857, Cagliari, Italy, 2000.

Dimauro, G., Impedovo, S., Modugno, R., Pirlo, G., Sarcinella, L., Analysis of Stability in Hand-Written Dynamic Signatures, Proc. of the 8th International Workshop on Frontiers in Handwriting Recognition (IWFHR-8), Niagara-on-the-Lake, Canada, 2002, 259-263.

Dimauro, G., Impedovo, S., Lucchese, M. G., Modugno, R., Pirlo, G., Recent Advancements in Automatic Signature Verification, Proc. of the 9th International Workshop on Frontier in Handwriting Recognition, Tzukuba, Japan , 2004, 179 – 184.

Djioua, M., Plamondon, R., Studying the Variability of Handwriting Patterns using the Kinematic Theory, Human Movement Science, 28, 5, 2009, 588-601.

Fleet D.J. , Jepson, A. D., Computation of oponent image velocity from local phase information, Int. Journal Computer Vision, 5, 1990, pp. 77-104.

Garcia-Salicetti, S., Houmani, N., Dorizzi, B., A client-entropy measure for on-line signatures, Proceedings of the IEEE Biometrics Symposium (BSYM '08), Tampa, Fla, USA, 2008, 83-88.

Guest, R. M. , The Repeatability of Signatures, Proc. of the 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR-9), Kichijoji, Japan, 2004, 492-497.

Guest, R., Age dependency in handwritten dynamic signature verification systems, Pattern Recognition Letters, 27, 10, 2006, 1098-1104.

Guest, R., Fairhurst, M., Sample Selection for Optimising Signature Enrolment, Proc. of the 10th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR 10), La Baule, France, October 2006.

Horn B.K.P., Schunck, B.G., Determining Optical Flow, MIT Press, A.I. Memo n. 572, April 1980.

Houmani, N., Garcia-Salicetti, S., Dorizzi, B., On assessing the robustness of pen coordinates, pen pressure and pen inclination to time variability with personal entropy, Proc. of the IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, 2009 (BTAS '09), Washington, DC, 2009, 1–6.

Huang, K., Yan, H., Stability and style-variation modeling for on-line signature verification, Pattern Recognition, 2003, 36, 10, 2253-2270.

Impedovo, S., Pirlo, G., Verification of Handwritten Signatures: an Overview, Proc. of the 14th International Conference on Image Analysis and Processing - ICIAP 2007, IEEE Computer Society Press, Modena, Italy, 2007, 191-196.

Impedovo, D., Pirlo, G., Automatic Signature Verification – The State of the Art, IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Review, 38, 5, 2008, 609 – 635.

Impedovo, D., Pirlo, G., Stasolla, E., Trullo, C.A. , Learning Local Correspondences for Static Signature Verification, Proc. of the 11th Int. Conf. of the Italian Association for Artificial Intelligence (AI*IA 2009), Reggio Emilia, Italy , 2009.

Impedovo, D. , Pirlo, G. , On the Measurement of Local Stability of Handwriting - An application to Static Signature Verification, Proc. Biometric Measurements and Systems for Security and Medical Applications (BIOMS 2010), Taranto, Italy, 2010, 41-44.

Impedovo, D., Pirlo, G., Stability Analysis of Static Signatures for Automatic Signature Verification, Proc. of the 16th International Conference of Image Analysis and Processing (ICIAP 2011), Lecture Notes in Computer Science, 6979, Springer Publ., 2011, 241-247.

Impedovo, D., Pirlo, G., Static Signature Verification by Optical Flow Analysis, Proc. of the 1st International Workshop on Automated Forensic Handwriting Analysis, Beijing, China, 2011, 31-35.

Impedovo, D., Pirlo, G., Sarcinella, L., Stasolla, E., Trullo, C.A., Analysis of Stability in Static Signatures using Cosine Similarity, Proc. of the XIII International Conference on Frontiers in Handwriting Recognition (ICFHR 2012), Monopoli, Bari, Italy, 2012, 231-235.

John, G. H., Kohavi R., Peger, K., Irrelevant features and the subset selection problem, Proc. of 11th Int. Conf. Machine Learning, San Francisco, 2004, 121-129.

Lei, H., Govindaraju, V., A comparative study on the consistency of features in on-line signature verification, Pattern Recognition Letters, 26, 2005, 2483-2489.

Lucas B., Kanade, T., An iterative image registration technique with an application to stereo vision, Proc. DARPA, 1981, 121-139.

Nagel, H. H., On a constraint equation for the estimation of displacement rates in image sequences, IEEE Trans. PAMI, 11, 1989, 13-30.

O'Donovan, P., Optical Flow: Techniques and Applications, The University of Saskatchewan, T.R. 502425, April 2005.

Pirlo, G., Algorithms for Signature Verification, in Proc. of NATO-ASI Series Fundamentals in Handwriting Recognition, S. Impedovo (Ed.), Springer-Verlag, Berlin, 1994, 433-454.

Quinlan, J. R., C4.5: Programs for Machine Learning, San Mateo, US, Morgan Kaufmann, 1993.

Schomaker, L. R. B., Plamondon, R., The Relation between Axial Pen Force and Pen-Point Kinematics in Handwriting, Biological Cybernetics, 63, 1990, 277-289.

Singh, A., Optic Flow Computation: A Unified Perspective, IEEE Computer Society Press, 1992.

Vargas, J.F., Ferrer, M.A., Travieso, C.M., Alonso, J.B., Off-line Handwritten Signature GPDS-960 Corpus, Proc. Of the 9th ICDAR, 2007, 764-768.

Vargas, J. F., Ferrer, M. A., Travieso, C. M., Alonso, J. B., Offline Signature Verification Based on Pseudo-Cepstral Coefficients, Proc. of the 10th Int. Conf. on Document Analysis and Recognition, 2009, 126-130.

Vargas, J. F., Ferrer, M. A., Travieso, C. M., Alonso, J. B., Off-line signature verification based on grey level information using texture features, Pattern Recognition, 44, 2011, 375-385.

Vielhauer, C., A Behavioural Biometrics, Public Service Review: European Union, 9, 2005, 113-115.

Yilmaz, M. B., Yanikoglu, B., Tirkaz, C., Kholmatov, A., Offline signature verification using classifier combination of HOG and LBP features, Proc. International Joint Conference on Biometrics (IJCB), 2011, 1-7.

Analysis of Stability and Complexity in Dynamic Signatures

D. Impedovo, G. Pirlo, F. Pomes

Dipartimento di Informatica, Università degli Studi di Bari "Aldo Moro"

Via Orabona, 4 – 70125 Bari

giuseppe.pirlo@uniba.it

Handwritten signature is a very special biometric trait since it is the most widespread and well-accepted way for personal authentication in the modern society. This paper presents a new approach for estimating signature stability and complexity. For the purpose a technique based on Dynamic Time Warping is illustrated that allows the analysis of stability and complexity in diverse domains of representation of the signature (position, velocity, acceleration, pressure, etc.). The experimental results show relevant insights about collective and individual characteristics of the signing process that can be useful for the improvement of adaptive and personalized signature verification systems.

Index Terms— *Biometrics, On-line Handwritten Signatures, Dynamic Signature Analysis, Dynamic Time Warping.*

1. Introduction

In the information society, in which security issues are an area of growing interest, biometrics is becoming more and more important. Among the different biometric traits, signature occupies a very special place since it is the most widespread means for personal verification and people are familiar with its use. In addition, signature is accepted as a legal means for personal verification by institutions and private and public companies [Boyer et al., 2007].

Of course, handwritten signature is a very complex pattern that convey a lot of personal information deriving from the “motor model” stored in the signer’s brain, the writing system (neuromuscular system) and the writing device (pen or pencil, paper, etc.). In addition, psychological and physical conditions of the writer can modify significantly the process of signature apposition [Impedovo and Pirlo, 2008].

In order to treat adequately the problem of variability in handwritten signatures, research has been focused on the analysis of signature stability. Pattern-matching [Impedovo et al., 2009; Impedovo and Pirlo, 2010; Impedovo and Pirlo, 2011], cosine-similarity [Impedovo and Pirlo, 2012] and optical flow [Impedovo and Pirlo, 2011] have been proposed to determine - at local level -

the degree of stability of each region of static signature, when dynamic signatures are considered. A client-entropy measure has been proposed to group and characterize dynamic signatures in categories that can be related to signature variability and complexity [Garcia-Salicetti et al., 2008]. The measure, that is based on local density estimation by a HMM, can be used to access whether a signature contains or not enough information to be successfully processed by any verification system [Houmani et al., 2009]. Stability of dynamic signatures can be also estimated by considering a set of features and the physical characteristics of signatures which they are most related to, in order to obtain global information on signature repeatability which can be used to improve the verification systems [Guest, 2004; Guest and Fairhurst, 2006]. In general, these approaches have shown that there is a set of features that remain stable over long time periods, while there are other features which change significantly in time [Guest, 2006, Djioua and Plamondon, 2009, Impedovo and Pirlo, 2008].

A lot of research has been also carried out to estimate complexity in a signature. Some approaches, consider complexity of the signature as a measure of the difficulty for its imitation, obtained as the result of the estimated difficulty in perceiving, preparing and executing each stroke of the signature itself [Brault and Plamondon, 1993].

In this paper we consider Dynamic Time Warping (DTW) for the analysis of signature stability and complexity. The approach is used for a systematic analysis of dynamic signatures of the SUSIG database, in order to derive information on handwritten signatures useful to adapt and personalize next generation signature verification systems.

The paper is organized as follows: Section 2 presents the technique for stability and complexity analysis by DTW. Section 3 illustrates the SUSIG database of handwritten online signatures, that has been used for the experimental tests. The experimental results are discussed in Section 4. Section 5 presents the conclusion of the paper.

2. Stability and Complexity of Dynamic Signatures by DTW

In this paper DTW is used for estimating local stability and complexity in dynamic signatures. In particular let $S = \{ S^r \mid r=1,2,\dots,n \}$ be a set of n handwritten signatures of the same writer. Each S^r can be described as $S^r = (z_1^r, z_2^r, \dots, z_i^r, \dots, z_{M^r}^r)$, where, for each $i=1,2,\dots,M^r$, we have that z_i^r describes the i -th sample (i.e. the sample at time i of the writing process) of the r -th signature. It is worth noting that the sample can be a value in the different domains of representation of the signature, like for instance: position, velocity, acceleration, pressure, etc. Now, let be S^r, S^t two genuine signature of the set ($S^r, S^t \in S$). A warping function between S^r and S^t is any sequence of couples of indexes identifying points of S^r and S^t to be joined [Huand and Yan, 2003; Congedo et al., 1995]:

$$W(S^r, S^t) = c_1, c_2, \dots, c_K, \tag{1}$$

where $c_k=(i_k,j_k)$ (k,i_k,j_k integers, $1 \leq k \leq K$, $1 \leq i_k \leq M^r$, $1 \leq j_k \leq M^t$). Now, if we consider a distance measure $d(c_k)=d(z^r(i_k), z^t(j_k))$ between points of S^r and S^t , we can associate to $W(S^r,S^t)$ the dissimilarity measure

$$D_{W(S^r,S^t)} = \sum_{k=1}^K d(c_k) . \tag{2}$$

The elastic matching procedure detects the warping function $W^*(S^r,S^t)=c^*_1,c^*_2,\dots,c^*_{K^*}$ which satisfies the monotonicity, continuity and boundary conditions, and for which it results [Congedo et al.,1995]:

$$D_{W^*(S^r,S^t)} = \min_{W(S^r,S^t)} D_{W(S^r,S^t)} . \tag{3}$$

From $W^*(S^r,S^t)$ we identify the *Direct Matching Points* (DMP) of S^r with respect to S^t [Huang and Yan, 2003]. A DMP of a signature S^r with respect to S^t is a point which has a one-to-one coupling with a point of S^t . In other words, let $z^r(p)$ be a point of S^r coupled with $z^t(q)$ of S^t ; $z^r(p)$ is DMP of S^r with respect to S^t iff:

- (a) $\forall p=1,\dots,M^r$, $p \neq p$, it results that: $z^r(p)$ is not coupled with $z^t(q)$;
- (b) $\forall q=1,\dots,M^t$, $q \neq q$, it results that: $z^t(q)$ is not coupled with $z^r(p)$.

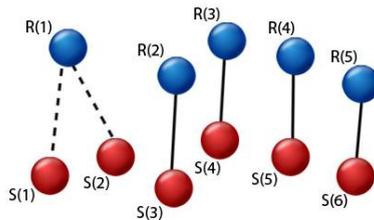


Figure 1. Direct Matching Points

Figure 1 shows an example of two small sequences of samples. The first sequence R(1), R(2), R(3), R(4) and R(5) are of the first signature. The second sequence S(1), S(2), S(3), S(4), S(5) and S(6) are of the second signature. From the DTW it results that the DMP of the first and second signatures are respectively R(2), R(3), R(4), R(5) and R(3), R(4), R(5), R(6).

Now, a DMP indicates the existence of a region of the r-th signature which is roughly similar to the corresponding region of the t-th signature (in the domain specified by the distance used for the elastic matching procedure). Therefore, for each point of S^r , a score is introduced according to its type of coupling with respect to the points of S^t [Congedo et al., 1993]:

$$\forall p=1,2,\dots,M^r: \text{Score}^t(z^r(p)) = \begin{cases} 1 & \text{if } z^r(p) \text{ is a DMP} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The local stability function of S^r is defined as:

$$\forall p = 1, 2, \dots, M^r: I(z^r(p)) = \frac{1}{n-1} \sum_{\substack{t=1 \\ t \neq r}}^n \text{Score}^t(z^r(p)). \quad (5)$$

In other words Direct Matching Points (DMPs), that are unambiguously matched samples of the genuine signature, are used to indicate the presence of a small stable region of the signature, since no significant distortion has been locally detected. The local stability of a point of a signature is determined as the average number of time it is a DMP, when the signature is matched against other genuine signatures. This procedure allows to detect low- and high-stability regions.

Similarly, in this paper DTW is also used to estimate signature complexity. This is performed by matching a genuine signature S^g against a set of forgeries $S^f = \{ S^{f,r} \mid r=1,2,\dots,m\}$. In fact, in this case the presence of DMP means that, for a small region of the signature, a forger has been extremely precise in imitating the genuine sample. Therefore, also in this case, complexity is considered as a measure of difficulty in imitating a specific signature.

3. The SUSIG Database

For the aim of this research we have considered the SUSIG database of online handwritten signatures [Kholmatov and Yanikoglu, 2009]. The database consists of two parts that are collected using different pressure-sensitive tablets (one with and the other without an LCD display). A total of 100 people contributed to each part, resulting in a database of more than 3,000 genuine signatures and 2,000 skilled forgeries. The genuine signatures in the database are real signatures of the contributors. In collecting skilled forgeries, forgers were allowed to practice in imitating genuine signatures. Furthermore, for a subset of the forgeries (highly skilled forgeries), this animation was mapped onto the LCD screen of the tablet so that the forgers could *trace over* the mapped signature. Forgers in this group were also informed of how close they were to the reference signature, so that they could improve their forgery quality. For each signature, the following data is available: position (x,y coordinates), pressure, time stamp, pen down/pen up signal.

It is worth noting that, from position information, we can derive velocity and acceleration information. Therefore, for the experimental tests, we have considered the following four domains of representation for online signatures: Position (P), Velocity (V), Acceleration (A), Pressure (Pr).

4. Experimental Results

Stability and complexity analysis has been carried out on the signatures of the SUSIG database (both genuine and forgery signatures).

For each signature, stability analysis has been performed on the four representation domains: position, velocity, acceleration and pressure. Figures 2a,b show the result of stability analysis in the position domain on a genuine signature. In particular Figure 2a shows the space-oriented representation of low (green), medium (yellow) and high (red) stability regions. Figure 2b shows the time-oriented representation of the stability regions.

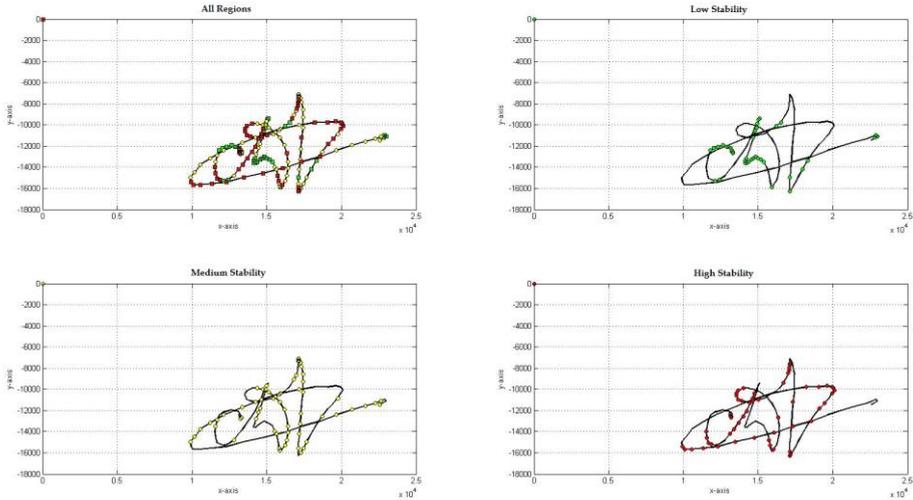


Figure 2a: Space-oriented representation of Stability Regions of a Genuine Signature

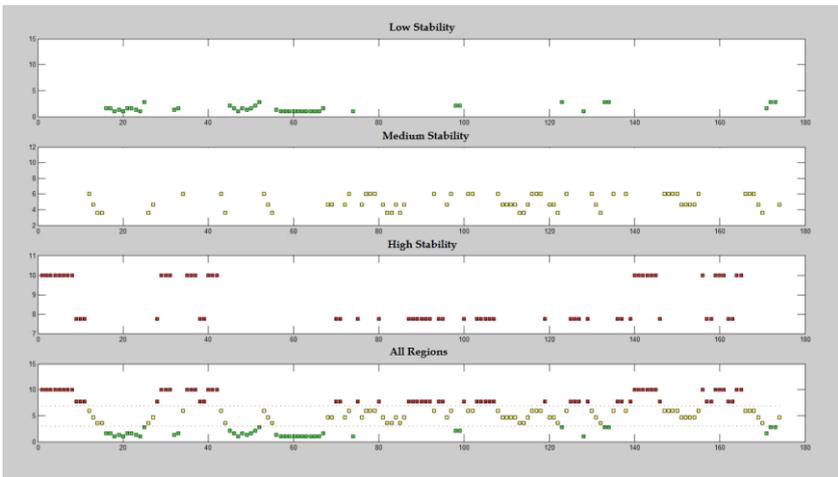


Figure 2b: Time-oriented representation of Stability Regions of a Genuine Signature

Furthermore, complexity analysis has also been performed for each signature considering the four representation domains. Figures 3a,b show the result on a genuine signature. In particular Figure 3a shows the space-oriented representation of low (green), medium (yellow) and high (red) complexity regions. Figure 3b shows the time-oriented representation of the complexity regions.

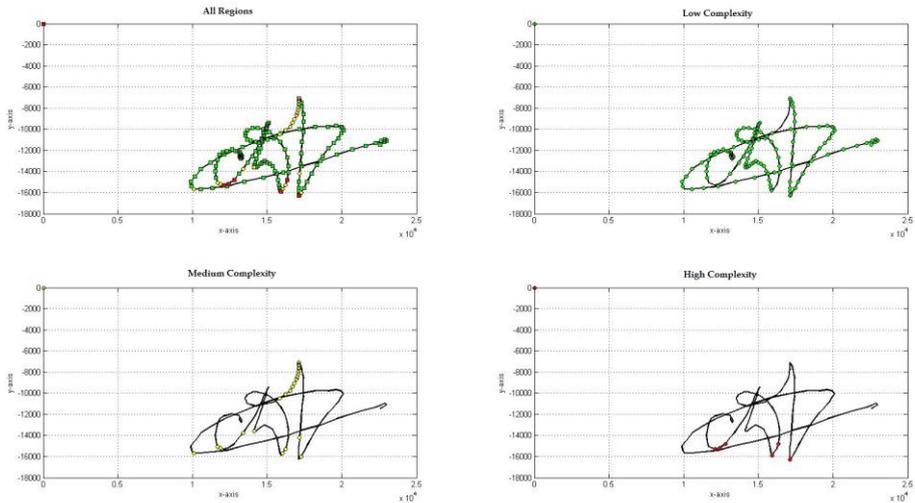


Figure 3a: Space-oriented representation of Complexity Regions of a Genuine Signature

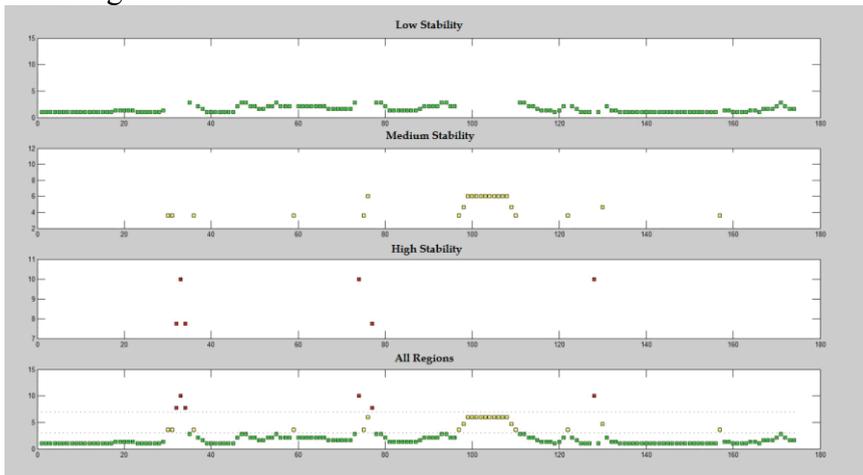


Figure 3b: Time-oriented representation of Stability Regions of a Genuine Signature

The distribution of stability and complexity in a signature is analysed in Figure 4, obtained from the analysis of all genuine signatures of the SUSIG database. For each domain of representation three parts of the signatures are considered (according to their length defined as the total number of samples): the Initial part (I), the Central part (C) and the Final part (F).

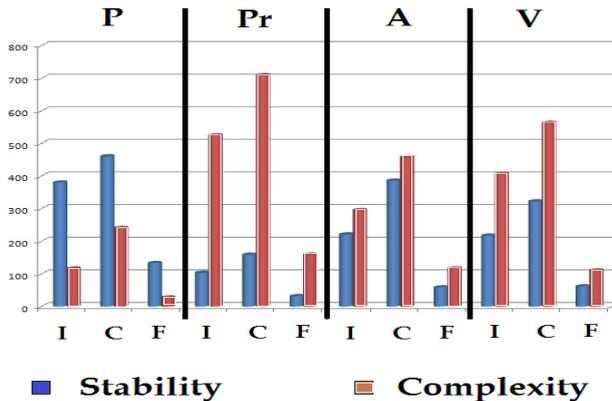


Figure 4. Stability and Complexity Distribution

Figure 4 shows that position is the most stable domain of representation, followed by velocity and acceleration. Concerning signature complexity, pressure is the most complex domain of representation. In other words this means that it is extremely difficult for a forger imitating the pressure profile of the genuine signer. Also velocity and acceleration are difficult domains of representation for a forger imitating a genuine signature.

Furthermore, this analysis points out that stability is mainly concentrated in the initial and central part of the signatures. Complexity is much more evident in the central part of the signature. It is quite interesting that the final part of the signature is generally not stable nor complex. In general, this appears to be not useful for verification aims.

5. Conclusion

In this paper a novel approach to estimate stability and complexity in handwritten dynamic signatures is presented. The approach is based on DTW and exploits potential of multiple matching strategy.

The experimental results have been carried out using the signatures of the SUSIG database. The results points out some interesting characteristics related to the signature apposition process that can be useful to improve, introducing advanced adaptive and personalized strategies, next generation systems for dynamic signature verification.

References

Boyer, K.W. , Govindaraju, V. , Ratha, N.K. , (Eds.), *Special Issue on Recent Advances in Biometric Systems*, IEEE Trans. on Syst., Man and Cybernetics – Part B, 37, 5, 2007.

Brault J.-J., Plamondon, R., A Complexity Measure of Handwritten Curves: Modeling of Dynamic Signature Forgery, *IEEE Transactions on Systems, Man, and Cybernetics (T-SMC)*, 23, 2, 1993, 400-413.

Congedo, G., Dimauro, G., Forte, A.M., Impedovo, S., Pirlo, G. , Selecting Reference Signatures for On-Line Signature Verification, Proc. of the 8th International Conference on Image Analysis and Processing (ICIAP-8), Series: Lecture Notes in Computer Science, 974, Springer-Verlag Berlin, Heidelberg, C. Braccini, L. De Floriani and G. Vernazza (Eds.), San Remo, Italy, Sept. 1995, 521-526.

Djioua, M., Plamondon, R., Studying the Variability of Handwriting Patterns using the Kinematic Theory, *Human Movement Science*, 28, 5, 2009, 588-601.

Garcia-Salicetti, S., Houmani, N., Dorizzi, B., A client-entropy measure for on-line signatures, Proceedings of the IEEE Biometrics Symposium (BSYM '08), Tampa, Fla, USA, 2008, 83-88.

Guest, R. M. , The Repeatability of Signatures, Proc. of the 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR-9), Kichijoji, Japan, 2004, 492-497.

Guest, R., Age dependency in handwritten dynamic signature verification systems, *Pattern Recognition Letters*, 27, 10, 2006, 1098-1104.

Guest, R., Fairhurst, M., Sample Selection for Optimising Signature Enrolment, Proc. of the 10th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR 10), La Baule, France, October 2006.

Houmani, N., Garcia-Salicetti, S., Dorizzi, B., On assessing the robustness of pen coordinates, pen pressure and pen inclination to time variability with personal entropy, Proc. of the IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, 2009 (BTAS '09), Washington, DC, 2009, 1–6.

Huang, K., Yan, H., Stability and style-variation modeling for on-line signature verification, *Pattern Recognition*, 2003, 36, 10, 2253-2270.

Impedovo, S., Pirlo, G., Verification of Handwritten Signatures: an Overview, Proc. of the 14th International Conference on Image Analysis and Processing - ICIAP 2007, IEEE Computer Society Press, Modena, Italy, 2007, 191-196.

Impedovo, D., Pirlo, G., Automatic Signature Verification – The State of the Art, *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Review*, 38, 5, 2008, 609 – 635.

Impedovo, D., Pirlo, G., Stasolla, E., Trullo, C.A. , Learning Local Correspondences for Static Signature Verification, Proc. of the 11th Int. Conf. of the Italian Association for Artificial Intelligence (AI*IA 2009), Reggio Emilia, Italy , 2009.

Impedovo, D. , Pirlo, G. , On the Measurement of Local Stability of Handwriting - An application to Static Signature Verification, Proc. Biometric Measurements and Systems for Security and Medical Applications (BIOMS 2010), Taranto, Italy, 2010, 41-44.

Impedovo, D., Pirlo, G., Stability Analysis of Static Signatures for Automatic Signature Verification, Proc. of the 16th International Conference of Image Analysis and Processing (ICIAP 2011), Lecture Notes in Computer Science, 6979, Springer Publ., 2011, 241-247.

Impedovo, D., Pirlo, G., Static Signature Verification by Optical Flow Analysis, Proc. of the 1st International Workshop on Automated Forensic Handwriting Analysis, Beijing, China, 2011, 31-35.

Impedovo, D. , Pirlo, G. , Sarcinella, L. , Stasolla, E. , Trullo, C.A., Analysis of Stability in Static Signatures using Cosine Similarity, Proc. of the XIII International Conference on Frontiers in Handwriting Recognition (ICFHR 2012), Monopoli, Bari, Italy, 2012, 231-235.

Kholmatov, A., Yanikoglu, B., SUSIG: An On-line Signature Database, Associated Protocols and Benchmark Results, Pattern Analysis and Applications, 12, 3, 2009, 227-236.

CARATTERIZZAZIONE DI MALATTIE RARE MEDIANTE ANALISI MORFOLOGICA DELLE ALTERAZIONI DEL VOLTO

Giuseppe Mastronardi, Cinzia Accettura, Domenico Daleno,
Roberto F. Depasquale, Massimiliano Dellisanti Fabiano Vilardi ¹

¹ DEI - Dipartimento di Ingegneria Elettrica e dell'Informazione
Politecnico di Bari – Via Orabona 4 – 70125 Bari
mastrona@poliba.it

Abstract. *Biometrics, in the medical field, helps to provide a response to the ever increasing need to automate the procedures for early diagnosis of diseases. The proposed method is undertaken by a collaboration between Politecnico di Bari and the Hospital "Casa Sollievo della Sofferenza" by S. Giovanni Rotondo, aimed at the characterization of RD (Rare Diseases). Among these, the need has arisen to compare the faces belonging to particular subjects suffering from syndromes such as, for example, Williams syndrome, characterized by the presence of morphological alterations of the face. This method can be used on stored images and is an innovative approach based on biometric comparison with respect to normalized face templates.*

Keywords: *Biometrics, Soft Computing, Rare Diseases.*

1. Introduzione

Le biometrie, in campo medico, contribuiscono a fornire una risposta alla sempre più crescente esigenza di automatizzare le procedure di diagnostica precoce di patologie. La metodica proposta, nasce da una collaborazione intrapresa tra il Politecnico di Bari e l'Ospedale "Casa Sollievo della Sofferenza" di S. Giovanni Rotondo, finalizzata alla caratterizzazione di MR (Malattie Rare). Tra queste è sorta la necessità di confrontare i volti appartenenti a soggetti affetti da particolari sindromi come, per esempio, la sindrome di Williams, caratterizzate dalla presenza di alterazioni morfologiche del volto. Utilizzabile su immagini già archiviate, il metodo è basato su un nuovo approccio di visione artificiale finalizzato alla classificazione biometrica rispetto a modelli di volto normalizzati. Infatti, dopo aver acquisito le tre proiezioni fondamentali del volto del paziente (frontale, laterale sinistra, laterale destra), su queste, mediante appositi software (SI FaceGen Modeller e SISCA), realizzati per attività investigative in ambito forense, si costruisce il modello normalizzato del volto partendo dai punti fisionomici più comuni che vanno a caratterizzare la maggior parte delle proporzioni più regolari del volto umano. A tal fine, si imposta la ricostruzione del modello del volto di riferimento in modo da ottenere comuni caratteristiche fisionomiche in analogia al paziente per sesso, razza ed età.

Una volta realizzato il modello, le patologiche differenze fisionomiche vengono quantificate confrontando tra loro alcuni indici. Le forme presenti in un volto (arcate sopraccigliari, profilo del naso, cavità acustiche e oculari, conformazione delle labbra, etc.), possono, infatti, essere caratterizzate da luoghi geometrici (es. triangoli) ottenuti collegando tra loro i diversi punti di repère, opportunamente selezionati per tipologia di alterazione. Due tra questi indici hanno fornito la maggiore capacità di discriminazione: l'indice di compattezza ($I=A/P^2$) e l'indice di somiglianza (S), ottenuto dalla matrice di accumulazione della GHT (Trasformata Generalizzata di Hough). Infine, per riconoscere esiti di strabismo, si è usato un metodo che combina diversi tipi di algoritmi noti in letteratura: i classificatori Haarcascade, utilizzati per l'individuazione degli occhi nel volto umano, il Principal Component Analysis (PCA) combinato con il Linear Discriminant Analysis (LDA), utilizzati per ridurre la "dimensionalità" dell'immagine, ed infine le Catene Nascoste di Markov (HMM), utilizzate per classificare e riconoscere tra gli occhi analizzati quelli affetti da strabismo. Il sistema è stato testato per il momento su un campione modesto di 40 individui di cui 23 con anomalie di strabismo, ottenendo un risultato alquanto soddisfacente

Questo metodo ha, quindi, lo scopo di evidenziare le alterazioni genetiche presenti sul volto, attraverso il confronto del volto del paziente affetto da una sindrome rara ed un volto normalizzato, costruito a partire dal volto del paziente stesso. Le diverse fasi del metodo vengono di seguito riportate.

2. Creazione dei Modelli Normalizzati

La prima fase è la creazione del volto normalizzato, a partire dalle acquisizioni del volto del paziente, raffigurato nelle tre proiezioni fondamentali (frontale, laterale sinistro, laterale destro).

Per costruire un volto normalizzato si parte, quindi, dall'acquisizione fotografica del volto del paziente. Grazie a SI FaceGen Modeller 3.5, un adeguato software free, è possibile ottenere, già dalle sole acquisizioni, un volto normalizzato, che può essere ulteriormente elaborato, attraverso la specifica dell'età, del sesso e della razza del paziente. Inoltre, per andare a correggere eventuali alterazioni rimaste, è possibile modificare alcuni parametri facciali, tra i quali vi sono quelli maggior rilevanza, rappresentati da labbra, tempie, mento, occhi, mascella e naso. Inoltre, al fine di rendere più evidenti singole alterazioni, partendo dal modello normalizzato, è anche possibile costruire tanti modelli, fedeli all'originale, quanti sono le singole alterazioni fisionomiche da considerare. Anche questi modelli modificati possono essere oggetto di confronto con il modello generale normalizzato per consentire, durante la sperimentazione, di verificare l'andamento degli indici e la loro affidabilità dal punto di vista della determinazione della discrepanza.

Per esempio, come nella Sindrome di Williams [1], può essere necessario concentrare l'attenzione su alcune delle alterazioni fisionomiche caratteristiche, quali:

- Ipoplasia mascellare
- Riduzione distanza bi-temporale

- Strabismo
- Naso infossato
- Labbra grosse e pendenti.

In altri termini, partendo dalle tre proiezioni del volto normalizzato, è possibile creare sul modello l'alterazione voluta sempre attraverso la modifica di alcuni parametri facciali.

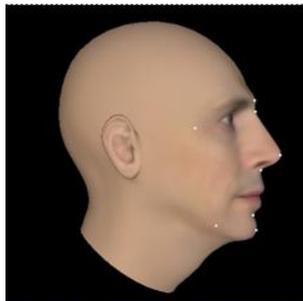
Una volta costruiti i modelli è possibile quantificare le differenze fisionomiche tra il modello normalizzato e i modelli rappresentanti le diverse alterazioni, attraverso gli indici ottenuti mediante il software SISCA (by eBIS) [2].

È possibile ottenere questi indici andando a fissare dei punti, detti punti di repère sui modelli, comunque ottenuto e sul volto del soggetto (Fig.1), facendo in modo che:

- la nuvola di punti di repère si riferiscano a siti omologhi sia sul modello rappresentante l'alterazione sia sul modello normalizzato con cui viene successivamente fatto il confronto, tenendo presente che i punti di repère cambiano a seconda della tipologia di alterazione considerata;
- l'unione dei punti di repère fissati sul volto devono andare a formare una rete di triangoli.



a) Modello normalizzato frontale



b) Modello normalizzato laterale



c) Labbra grosse e pendenti



d) Ipoplasia mascellare



e) Naso infossato



Fig.1 - Punti di repère a confronto

3. Calcolo e confronto degli indici

Il passo successivo consiste, quindi, nel fissare due set di punti di repère omologhi sui due volti da confrontare. Si ottengono, così, due set di indici omologhi che vanno a caratterizzare le differenze tra i due volti messi a confronto [3]. Infine, si procede con il confronto tra i due set, basato sul calcolo del coefficiente di correlazione. Tra i vari indici ottenuti, l'indice di compattezza ($I=A/P^2$) e l'indice di somiglianza (S) sono risultati maggiormente discriminanti e, perciò, più significativi ai fini del confronto.

L'indice di compattezza, anche indicato come fattore di forma, caratterizza la forma di un poligono, nel nostro caso si applica ai triangoli ottenuti dal congiungimento dei vari punti di repère. E' rappresentato dal rapporto tra l'area ed il quadrato del perimetro della superficie considerata. Ottenuti i due set omologhi di indici di compattezza, rispettivamente dal volto normalizzato e dal volto del paziente (o suo modello specifico), si procede al confronto basato sul calcolo del coefficiente di correlazione, che per essere significativo deve risultare minore di 0,96.

L'indice di somiglianza stabilisce, invece, il grado di somiglianza che esiste tra i due volti messi a confronto (volto normalizzato e volto del paziente o suo modello specifico). Questo confronto è possibile grazie alla GHT (Trasformata Generalizzata di Hough) che permette di individuare forme che non sono descritte da nessuna equazione analitica. La GHT si compone di una prima fase che consiste nella costruzione del modello della forma da individuare e di una seconda fase che consiste nell'individuazione proprio della forma [4]. L'idea di partenza della GHT è che per ogni punto di contorno è possibile ottenere un'informazione del gradiente, e dato che il gradiente è una misura della variazione in una funzione, si ricercano i punti di contorno tra quelli caratterizzati da valori elevati del gradiente. Nella prima fase, la forma da individuare viene descritta scegliendo un punto di riferimento $P'(x',y')$ e valutando i punti di contorno della figura $P(x,y)$. Per ogni punto di contorno valutato, viene aggiornata una tabella, chiamata tabella di riferimento, in cui vengono indicati i valori dell'orientamento del gradiente (φ) ed il valore (r) della distanza del punto rispetto all'origine, che viene salvato come funzione di φ . Inoltre, si alloca una matrice di punti di accumulazione (accumulatori) che

possiede le stesse dimensioni dell'immagine ed è costituita da celle che rappresentano le posizioni attese del punto di riferimento. Una volta aggiornata la tabella, il punto di contorno va ad incrementare il rispettivo accumulatore. Al termine di questa operazione, eseguita su entrambe i volti messi a confronto, è possibile passare alla seconda fase che consiste nell'analizzare la matrice di accumulazione per definire il valore del picco massimo. Questo valore, rapportato agli altri valori presenti nella matrice di accumulazione, consente di calcolare l'indice di somiglianza, che rappresenta l'istanza della forma cercata. Ovviamente, minore è l'indice di somiglianza e maggiore è lo scostamento dalla norma. Per essere significativa la discrepanza occorre ottenere valori minori del 75%.

MODELLO	INDICE DI COMPATTEZZA (COEFFICIENTE DI CORRELAZIONE)	GHT (INDICE % DI SOMIGLIANZA)
c) Labbra grosse e pendenti	0,78265 < 0,80	7,56% < 25%
d) Ipoplasia mascellare	0,94441 < 0,95	66,20% < 75%
e) Naso infossato	0,95990 < 0,96	62,50% < 75%
f) Strabismo	0,97965 < 0,98	87,60% < 90%
g) Riduzione distanza bitemporale	0,95390 < 0,96	51,00% < 75%

Tabella di confronto

Analizzando questi due indici, è possibile ottenere delle indicazioni circa l'affidabilità dei risultati ottenuti e di conseguenza del metodo utilizzato per ottenerli. Ad eccezione del caso "f", negli altri casi, è possibile ritenere affidabili i risultati ottenuti.

Il caso "c" è quello più significativo, in cui l'indice di compattezza è minore di 0,80 e l'indice di somiglianza è molto minore del 25%.

Risultati comunque affidabili, sebbene più alti rispetto al caso "c", si ottengono nei casi "d", "e", "g", che hanno indice di somiglianza che non supera il 75%, associato ad un indice di compattezza che non supera lo 0,96.

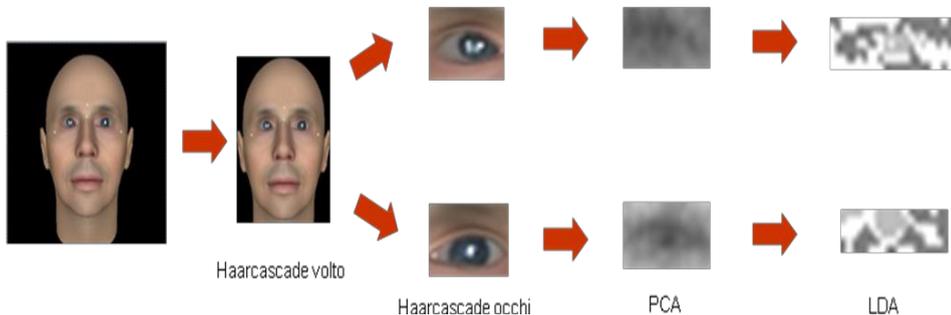
Diverso è il caso "f", in cui il valore dell'indice di compattezza si aggira attorno allo 0,98 e l'indice di somiglianza tende al 90%. In questo caso non è, quindi, possibile ritenere affidabili i risultati ottenuti, per cui il metodo utilizzato è stato integrato con un approccio basato su algoritmi di soft-computing.

4. Analisi dello Strabismo

Il sistema proposto per la valutazione dello strabismo è basato su un metodo ibrido caratterizzato da una fase di pre-processing in cui l'immagine viene elaborata da tre sottosistemi collegati in cascata: Haarcascade [5], Principal Component Analysis (PCA) [6], Linear Discriminant Analysis (LDA) [7].

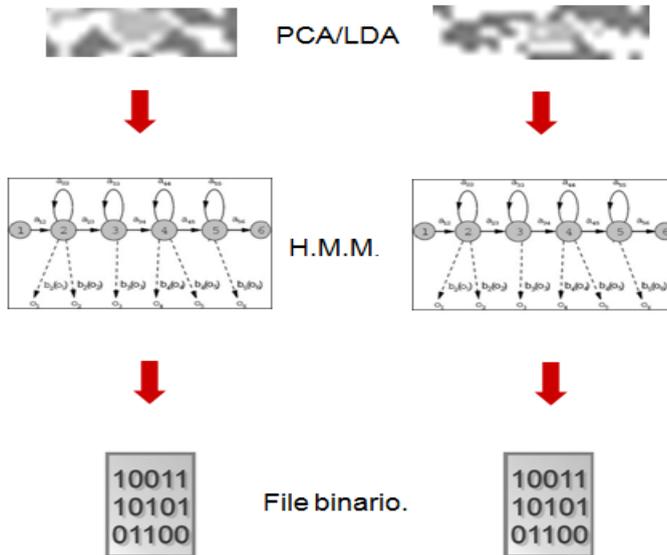
Acquisendo un'immagine da fotocamera o da un archivio database, la si converte in scala di grigi affinché la si possa elaborare attraverso l'algoritmo Haarcascade al fine di rilevare la presenza, prima del volto umano ed in seguito degli occhi, e restituire due immagini ridimensionate contenenti singolarmente i due occhi in primo piano. Queste immagini, successivamente, vengono processate in cascata da altri due metodi: il PCA e la LDA. Il PCA è un algoritmo utilizzato per la semplificazione dei dati applicata nell'ambito della statistica multivariata. Ha come fine primario la riduzione di un numero più o meno elevato di variabili, che rappresentano le numerose caratteristiche del fenomeno analizzato, in poche variabili latenti. Il parametro fondamentale di questo algoritmo è il numero delle componenti con il quale è possibile aumentare o diminuire la riduzione della "dimensionalità" dei dati.

Successivamente, le immagini processate dal PCA vengono elaborate dalla LDA. Questo algoritmo, effettua una riduzione lineare della "dimensionalità" supervisionata dell'immagine, in modo da rendere i dati risultanti più idonei alla classificazione.



Dopo il pre-processing, le immagini vengono convertite in file binario, caratterizzato da autovettori, necessario all'elaborazione successiva che è effettuata mediante gli HMM [8]. Essi sono modelli probabilistici in cui le sequenze vengono generate da due processi stocastici concomitanti. Il primo è costituito dal modello di Markov, che possiamo, in tutta generalità, considerare del primo ordine, rappresentabile graficamente come un insieme di stati connessi che simboleggiano le probabilità di transizione. Il secondo processo è l'emissione da parte di ogni stato di un carattere di un alfabeto secondo una distribuzione di probabilità che dipende solo dallo stato. La scelta di utilizzare gli HMM come nucleo fondamentale di classificazione biometrica, è stata motivata dagli studi condotti, che hanno evidenziato come in letteratura, l'impiego degli HMM [9] [10] risulti essere uno dei sistemi più affidabili per il riconoscimento biometrico. Alla base degli HMM è presente la teoria delle probabilità: essi permettono un approccio più robusto attraverso la descrizione dell'incertezza sui parametri, mediante insiemi di distribuzioni. Nello sviluppare un HMM, si deve decidere un certo numero di stati nascosti prima di formare un modello. Quindi, si procede con la fase di addestramento che consiste in una prima operazione di stima delle matrici delle transizioni e di output e, come seconda operazione, la fase di vero addestramento delle matrici in base a un training set

ottenuto col metodo standard di segmentazione di Viterbi e gli algoritmi di Baum-Welch. Dopo che l'HMM è stato addestrato, la fase di riconoscimento consiste nel recupero della sequenza di osservazione e della probabilità d'uscita di un'osservazione che determina la classe di appartenenza. Addestramento e riconoscimento vengono applicati ad un database contenente occhi strabici e non strabici.



La fase di testing del sistema ibrido proposto, lavorando sul numero delle componenti del PCA in modo da migliorare la classificazione biometrica e sul numero degli autovettori dell'HMM, seppure applicato al momento a un set limitato di campioni (23 casi di strabismo su 40 volti archiviati di diversa età e genere), ha prodotto risultati alquanto interessanti facendo registrare il riconoscimento di tutti i casi affetti da strabismo.

5. Conclusioni

La sperimentazione effettuata in campo medico diagnostico ha evidenziato, ancora una volta, come sia necessario ricorrere a diversi metodi di identificazione e classificazione, essendo particolarmente variegato lo scenario delle singole patologie anche quando queste, desumibili da immagini di diversa natura, sono caratterizzate dalla presenza di semplici alterazioni morfologiche (organi, volto, arti, etc.).

Tale convinzione è stata rafforzata dal caso specifico preso in considerazione in questo lavoro, poiché si è dovuto fare riferimento alla generazione di modelli ricavati dall'originale, evidenzianti ciascuno una singola alterazione morfologica, soprattutto al fine di valutare l'affidabilità delle soluzioni algoritmiche adottate.

In altri termini, è meno importante valutare la bontà di un algoritmo su un ventaglio di applicazioni rispetto alla necessità di doverne utilizzare più di uno, al fine di definire la soluzione anche ibrida più attendibile in funzione dei risultati ottenuti.

6. Riferimenti bibliografici

- [1] B. Dallapiccola, S. Vicari. "LA SINDROME DI WILLIAMS - Genetica, clinica e riabilitazione", Franco Angeli (2012)
- [2] SISCA (Shape Investigation System by Combined Analysis), Software sviluppato e commercializzato da eBIS srl (2011)
- [3] G. Mastronardi, et al. "Personal Identification of Bank Robbers by Morphometric Image Analysis - An Italian Experience". In: IASTED-ISLAT Conference on Law and Technology. ISBN: 0-88986-333-4, pp. 108-111, Boston, November 6-7 (2002)
- [4] D.H. Ballard, "Generalizing the Hough Transform to detect arbitrary shapes", Pattern recognition, Vol. 13, No. 2, pp. 111-122 (1981)
- [5] R. Padilla, C. F. F. Costa Filho, M. G. F. Costa. "Evaluation of Haar Cascade Classifiers Designed for Face Detection". World Academy of Science, Engineering and Technology 64 (2012)
- [6] A. P. M. Turk. "Eigenfaces for Recognition". Journal of Cognitive Neuroscience. Vol. 3.1. (1991)
- [7] S. R. O. Duda. "Pattern Classification". Wiley-Interscience Publication (2000)
- [8] A. V. Nefian, M. H. Hayes. "Face Detection and Recognition Using Hidden Markov Models". Proc. IEEE Int'l Conf. Image Processing, Vol.1 (1998)
- [9] V. Bevilacqua, G. Mastronardi, A. Pedone, G. Romanazzi, D. Daleno. "Hidden Markov Models for Recognition Using Artificial Neural Networks". Proc. ICIC 2006, pp. 126-134 (2006)
- [10] V. Bevilacqua, L. Cariello, G. Carro, D. Daleno, G. Mastronardi. "A face recognition system based on Pseudo 2D HMM applied to neural network coefficients". On Soft Computing - A Fusion of Foundations, Methodologies and Applications, Vol. 12, Issue 7, February (2008)

A facial image based age estimation system: TAGe

E. Vitulano, D. Impedovo, G.Pirlo
Dipartimento di Informatica - Università degli studi di Bari
Via Orabona 4, 70100 Bari
evitulano@gmail.com

Abstract. *In this work we introduce a facial image based age estimation system, named TAGe. We use discrete cosine transformation applied on locale feature pattern of facial image and SVR to obtain age estimation. We discriminate between male and female using a SVM to obtain gender classification.*

Keywords: Age estimation system, face analysis, FG-NET aging.

1. Introduzione

Un sistema di stima dell'età effettua il processo tramite cui l'età di un individuo viene stimata a partire da caratteristiche biometriche estratte dall'immagine del suo volto. L'obiettivo minimo è fornire in tempo reale una stima sull'età che sia precisa almeno quanto la stima fornita da un uomo; la direzione di ricerca è quella di ottenere, se possibile, una precisione maggiore rispetto alla stima effettuata da un umano.

Le sembianze del volto umano variano in maniera considerevole in seguito all'invecchiamento. Gli effetti dell'invecchiamento sono generalmente attribuiti alla crescita delle ossa e alla deformazione della pelle seguita dalla comparsa di rughe e alla riduzione della tensione muscolare. Di solito la crescita delle ossa del cranio è un fenomeno che avviene per lo più durante l'infanzia, mentre durante l'età adulta i cambiamenti maggiori indotti dall'invecchiamento sono quelli della struttura del viso.

In questo lavoro introduciamo un sistema di stima dell'età basato sull'analisi dell'immagine del volto chiamato TAGe. Abbiamo seguito un approccio basato sull'apprendimento automatico usando come elemento caratteristico la trama del volto. Abbiamo usato la trasformata discreta del coseno per estrarre le informazioni da pattern locali dell'immagine e macchine a vettori di supporto per effettuare la classificazione. Un sistema di stima trova la sua applicazione in svariati settori, tra cui:

- controllo ad accesso condizionato per contenuti fisici e/o virtuali;
- personalizzazione dell'interfaccia utente in base all'età dell'utilizzatore;
- creazione di sistemi di ritrovamento di immagini basato sulle età dei soggetti raffigurati;

- personalizzazione di campagne pubblicitarie su apparecchiature elettroniche munite di dispositivi di acquisizione di immagini (webcam o telecamere di sicurezza);
- statistiche sull'età della clientela di locali pubblici e attività commerciali;
- identificazione di materiale pedopornografico all'interno di piattaforme di condivisione video per adulti;
- aumento della velocità e della precisione di sistemi di riconoscimento facciale, eliminando la possibilità di usare immagini obsolete nei processi di verifica.

2. Approccio usato

Fra i fattori che caratterizzano l'invecchiamento, oltre ad i fattori primari come la crescita ossea e la comparsa di rughe, ci sono anche elementi secondari, come il sesso e l'appartenenza a diverse etnie [Farkas, 1994]. Si è scelto pertanto di effettuare la stima dell'età distinguendo il processo di stima in base al sesso. All'inizio del processo quindi, un classificatore SVM è stato usato per distinguere il sesso del soggetto raffigurato e solo in seguito è stata effettuata la stima delle diverse età utilizzando la SVR. All'analisi della forma del viso, si è preferito procedere all'analisi delle caratteristiche della trama della pelle, caratteristico dell'invecchiamento in età adulta. Si è scelto di seguire l'approccio fornito da [Steiner, 2010]: il vettore delle caratteristiche è stato popolato usando i coefficienti della trasformata discreta del coseno applicati ai parametri di aspetto locale; il vettore di caratteristica così ottenuto è stato usato nella classificazione:

- del genere tramite una macchina a vettori di supporto (SVM);
- nella stima dell'età delle donne e degli uomini tramite la regressione a vettori di supporto (SVR).

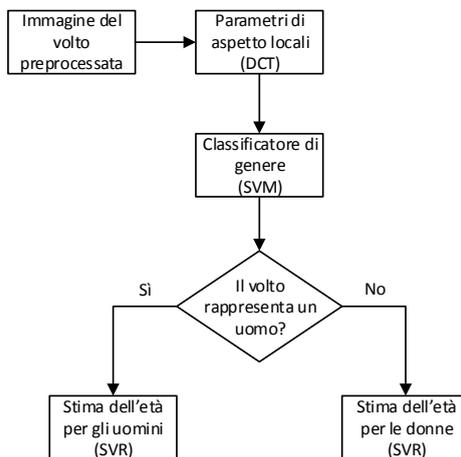


Fig.1 - Diagramma a blocchi del sistema proposto

3. TAGe

TAGe è un sistema di stima dell'età dall'analisi del volto che è stato sviluppato seguendo prevalentemente l'approccio di [Steiner, 2010]. Esso è stato implementato in Python, benché le funzioni dal costo computazionale più elevato siano svolte da librerie C++ alle quali è connesso tramite interfacce a basso livello. Fra queste, per l'elaborazione delle immagini, il calcolo della trasformata discreta del coseno, il rilevamento dei volti e degli occhi, è stata usata la nota libreria di computer vision open source e multiplatforma OpenCv [Geng et al, 2006]. Per l'uso dei classificatori SVM e SVR è stata usata la libreria Scikit-learn [Edwards et al, 1998] che si interfaccia a basso livello con la nota libreria LIBSVM [Chang e Lin, 2011].

3.1 Pre-elaborazione delle immagini

Le immagini fornite in input al sistema di analisi sono convertite in bianco e nero, normalizzate e scansionate per la ricerca del volto: per ogni volto identificato si procede alla localizzazione delle coordinate degli occhi. In queste due fasi, per la localizzazione del viso e degli occhi, viene usato il classificatore a cascata con le relative Haar-feature, fornito da openCv.

Il sistema procede quindi rotando l'immagine del volto rilevato, in modo tale che gli occhi siano sulla stessa linea: ciò si rende necessario per diminuire la variazione delle caratteristiche del volto in esame, causate dal cambiamento della posa, dell'angolo di inclinazione e della scala dell'immagine. L'allineamento avviene con una trasformazione nello spazio euclideo, e per effettuarlo vengono prese come riferimento le coordinate degli occhi: l'immagine ottenuta viene poi ritagliata e scalata alla dimensione di 64x64 pixel in modo tale da avere una parte fissa e omogenea di superficie al di sopra degli occhi e ai lati. Il fattore scelto sperimentalmente per selezionare l'area di interesse dell'immagine del volto al di sotto degli occhi (orizzontalmente e verticalmente) è del 21%.

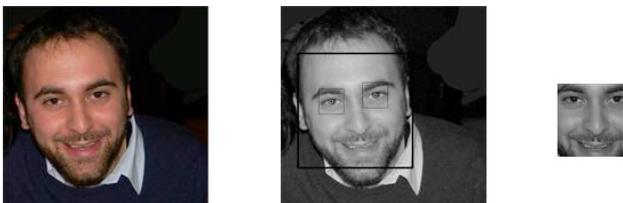


Fig. 2 - Esempio di immagini di volti pre elaborati

3.2 Estrazione del vettore di caratteristica

Per la rappresentazione del volto in un vettore di caratteristica, viene usato un approccio di estrazione di parametri di aspetto locali: l'immagine del volto viene mappata in un reticolo di blocchi non sovrapposti e l'estrazione delle caratteristiche viene effettuata su questi blocchi, piuttosto che sull'intera

immagine. L'immagine del volto può presentare delle occlusioni, delle zone a diversa illuminazione, o semplicemente rappresentare diverse espressioni, il che induce a cambiamenti nell'aspetto; quando vi è un cambiamento di aspetto, l'utilizzo di blocchi locali offre il vantaggio che solo la regione correlata del blocco o dei blocchi è interessata. Invece nel caso in cui l'estrazione delle caratteristiche venga effettuata sull'intero volto, l'intera rappresentazione verrebbe influenzata dai cambiamenti [Ekenel e Stiefelhagen 2010].

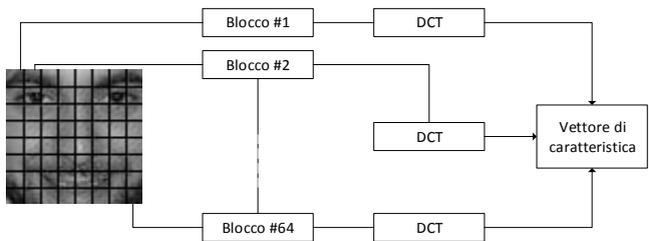


Fig. 3 - Schema di estrazione delle caratteristiche

L'immagine di un volto pre elaborato (cfr. 3.1) viene diviso in blocchi di dimensioni 8x8 pixel: tale dimensione viene scelta come grandezza in applicazioni di riconoscimento facciale perché è abbastanza piccola per fornire stabilità all'interno del blocco e, benché la complessità della trasformazione sia banale, sufficientemente grande per fornire un livello opportuno di compressione. Essendo l'immagine di dimensione 64x64 pixel, la divisione in blocchi di dimensione 8x8 porta alla creazione di 64 blocchi non sovrapposti. Quindi a ciascun blocco, viene applicata la trasformata discreta del coseno. Il primo coefficiente della matrice ottenuta viene eliminato dalla rappresentazione, in quanto corrisponde alla media del valore di intensità del blocco. Usando la scansione a zig-zag (cfr. Fig. 4) i primi 5 coefficienti vengono estratti e concatenati con il vettore di caratteristica dell'immagine.

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

Fig. 4 - Ordine coefficienti nella scansione a zig-zag in una matrice 8x8

Il vettore di caratteristica dell'immagine sarà dunque composto da 5 coefficienti per ciascun blocco. La sua dimensione totale è quindi pari a $8 \cdot 8 \cdot 5 = 320$ coefficienti. L'operazione è descritta in fig. 3.

3.2.1 Trasformata Discreta del Coseno

La trasformata discreta del coseno (Discrete Cosine Transform, DCT) [Jain 1989] è una funzione che provvede alla compressione spaziale, capace di rilevare la variazione del peso informativo tra un'area e quella contigua di un'immagine digitale, trascurandone le ripetizioni. La DCT è preferita in applicazioni in tempo reale grazie al suo costo computazionale basso. È dimostrato che la rappresentazione DCT è robusta ai cambiamenti di luce e alle variazioni di scala grazie alla sua capacità di scomposizione.

La trasformata discreta del coseno in due dimensioni, per una matrice quadrata, può essere così descritta:

$$t(i, j) = c(i, j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{\pi(2y+1)i}{2N} \cos \frac{\pi(2x+1)j}{2N}$$

dove f , N , t hanno notazione analoga del caso ad una dimensione, e $c(i, j)$ è dato da: $c(0, j) = \frac{1}{N}$, $c(i, 0) = \frac{1}{N}$, e $c(i, j) = \frac{2}{N}$ per $i \neq 0$ e $j \neq 0$.

Il primo coefficiente rappresenta il valore medio di intensità dell'immagine causato da variazioni di illuminazione.

3.3 Addestramento del sistema

Per effettuare l'addestramento del sistema è stato scelto il database FG-NET Aging, distribuito con annotati, per ogni immagine, 68 punti caratteristici fra cui le coordinate degli occhi: ciò permette di poter effettuare una fase di pre-elaborazione più accurata in quanto le coordinate degli occhi sono espresse e non devono essere ricercate all'interno dell'immagine del volto. Come abbiamo visto nel paragrafo 3.1, nella fase di pre-elaborazione le coordinate degli occhi sono fondamentali per poter procedere all'allineamento e alla trasformazione dell'immagine: avere l'immagine di un volto non allineata causerebbe del rumore nell'insieme di addestramento, il che è evidentemente deprecabile.

Per ogni immagine del database FG-NET viene estratto il vettore delle caratteristiche, così come mostrato nel paragrafo 3.2, e memorizzato in una struttura dati insieme ad un'etichetta indicante l'età e un'altra indicante il genere. Quindi i valori del vettore delle caratteristiche vengono normalizzati affinché ricadano nell'intervallo $[-1, 1]$ [Chang e Lin, 2011].

A questo punto vengono istanziati tre classificatori:

- il classificatore SVM C_1 per il riconoscimento di genere;
- il classificatore SVR C_2 per effettuare il processo di regressione sull'età degli uomini;
- il classificatore SVR C_3 per effettuare il processo di regressione sull'età delle donne.

Per ciascuno dei tre classificatori si è scelto di usare il kernel Radial Basis Function (RBF) e i parametri C e γ sono stati ottenuti mediante la tecnica di cross-validation e grid-search.

3.4 Classificazione del sesso e stima dell'età

Durante la fase di inizializzazione, il sistema provvede a caricare in memoria i classificatori precedentemente addestrati (cfr 3.3) e resta in attesa di ricevere nuove immagini da analizzare. Alla ricezione di una nuova immagine, viene effettuata all'interno della stessa la ricerca di volti. Se vengono trovati l'immagine viene pre elaborata (cfr. 3.1) e, in seguito, dall'immagine viene estratto il vettore di caratteristica (cfr. 3.2). A questo punto il vettore di caratteristica viene sottoposto al classificatore C_1 per individuarne il genere: se il classificatore C_1 classifica il vettore come appartenente all'immagine del volto di un uomo, la classificazione dell'età verrà effettuata tramite il classificatore C_2 ; in caso contrario, ovvero se il classificatore C_1 identifica il vettore di caratteristica come appartenente all'immagine di una donna, il classificatore C_3 viene usato per stimare l'età della donna.

4 Sperimentazione

4.1 Approccio

Dopo aver realizzato il sistema, per valutare il suo funzionamento si è usato un approccio sperimentale perché il problema di stima dell'età dall'analisi del volto non ha una specifica formale che consenta un altro tipo di valutazione. La sperimentazione oltre ad essere effettuata su dati di test indipendenti da quelli di addestramento, dovrebbe essere validata in un contesto reale. Per questo motivo si è scelto di creare un database che ha reso possibile ciò: TAGd. Esso contiene immagini di volti fornite spontaneamente dagli utenti dell'applicazione "The Age Guesser" con l'obiettivo di riceverne la stima dell'età. Il numero elevato di immagini presenti in TAGd, le condizioni di ripresa non controllate e la variazione di illuminazione e in termini di posa del capo fanno delle immagini in esso presenti una rappresentazione fedele di un contesto reale. Ogni volto è etichettato con età e sesso.



Fig. 5 - Alcuni volti in TAGd

4.2 Risultati ottenuti

Per valutarne i risultati ottenuti dal classificatore di genere C_1 usiamo la precision (π), il recall (ρ) e la F_1 -measure.

Per misurare le prestazioni globali del sistema di stima dell'età abbiamo fatto uso di altri due metodi statistici: la media degli errori assoluti (*Mean Absolute Error, MAE*) ed il punteggio cumulativo (*Cumulative Score, CS*).

MAE rappresenta la differenza mediata fra l'età stimata e l'età reale. Si calcola con la seguente formula:

$$MAE = \frac{\sum_{i=0}^n |EA_i - RA_i|}{n}$$

dove EA_i è il valore dell'età stimato, RA_i è il valore reale dell'età del i -esimo esempio fra gli n valutati.

Il Cumulative Score di una differenza di età d descrive la percentuale di valutazioni che hanno un errore di stima minore o uguale di d anni. Ovvero:

$$CS(d) = \frac{N(|EA_i - RA_i|) \leq d}{n} \times 100$$

dove $N(|EA_i - RA_i|) \leq d$ è il numero di valutazioni con un errore di stima minore o uguale ad una certa soglia d .

I risultati ottenuti sono sintetizzati nelle tabelle di seguito riportate:

	Precision	Recall	F-measure	# Elementi
Donne	0.78	0.43	0.55	1088
Uomini	0.40	0.76	0.52	551
Media	0.65	0.54	0.54	-

Tab. 1 - Risultati sperimentali del classificatore di genere eseguito su TAGd

	MAE
Uomini	10.01
Donne	10.31
Combinato	10.21

Tab. 2 - Mean Absolute Error sulla stima dell'età eseguita su TAGd

$CS(d)$	3	5	10	15
Uomini	21,77	35,57	60,8	76,77
Donne	17,37	30,7	57,9	74,9
Combinato	18,85	32,34	58,87	75,53

Tab. 3 - Cumulative Score(d) sulla stima dell'età eseguita su TAGd

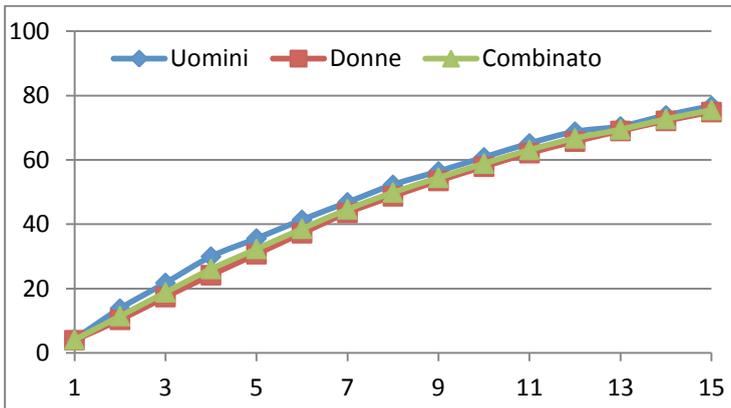


Fig. 6 - Grafo Cumulative Score(d) per la stima dell'età eseguita su TAGd

5. Conclusioni

Nella stima dell'età abbiamo privilegiato l'analisi della struttura del volto, piuttosto che l'analisi della forma. Un possibile sviluppo futuro potrebbe prevedere la combinazione del vettore di caratteristica ottenuto concatenando i coefficienti della trasformata discreta del coseno calcolata sui pattern locali di aspetto con parametri che tengano conto della forma del viso, quali per esempio i parametri del modello di apparenza attivo (Active Appearance Model, AAM [Cootes et al., 1998]).

Inoltre, nella fase di sperimentazione si è evidenziato come la buona riuscita del processo, dipenda in maniera fondamentale nella fase di pre-elaborazione delle immagini. Infatti la corretta identificazione dei volti all'interno delle immagini e la conseguente localizzazione degli occhi in essi contenuti fa sì che la zona su cui viene estratto il vettore delle caratteristiche sia quanto più omogenea possibile fra tutte le immagini fornite al sistema. Rafforzando questo processo, le prestazioni del sistema migliorano significativamente.

Bibliografia

[Chang e Lin 2011] Chang, Chih-Chung and Lin, Chih-Jen, LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology. 2011, Vol. 2, 3.

[Steiner 2010] M. Steiner, "Facial Image-based Age Estimation". Study Thesis at the Faculty of Computer Science Institute for Anthropomatics Facial Image Processing and Analysis. 2010

[Cootes et al. 1998] Cootes T.F., Edwards G.J., Taylor C.J., Active appearance models, Proc. European Conference on Computer Vision, 1998.

[Ekenel e Stiefelhagen 2010] Ekenel, H.K., Stiefelhagen, R. Automatic frequency band selection for illumination robust face recognition, In: Proc. of Intl. Conf. on Pattern Recognition. 2010.

[Edwards 1998] Edwards G. J., Lanitis A., Cootes C.J., Statistical face models: Improving specificity. Image and Vision Computing. 1998.

[Farkas 1994] Farkas, L. G., Anthropometry of the Head and Face, Raven Press, New York, 1994.

[Geng et al. 2006] Xin Geng, Thi-Hua Thou, Yu Thang, Gang Li, Honghui Dai., Learning from facial aging patterns for automatic age estimation., Proc. of 14th ACM International Conference on Multimedia, Santa Barbara, CA, USA 2006.

[Jain 1989] Jain, A.K., Fundamentals of Digital Image Processing, Prentice Hall, s.l. 1989.

The “CSCT Living Lab” for Computer Science and Computational Thinking

Claudio Demartini, Fabrizio Lamberti, Marina Marchisio¹, Claudio Pardini²,
Amelio Patrucco³

Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino

claudio.demartini@polito.it, fabrizio.lamberti@polito.it

¹*Università di Torino, Via Carlo Alberto 10, 10124 Torino,*

marina.marchisio@unito.it

²*IIS Carlo Anti, Via Magenta 7, 37069 Villafranca di Verona,*

dirigente@carloanti.it

³*Fondazione Torino Wireless*

Via Francesco Morosini 19, 10128 Torino

amelio.patrucco@torinowireless.it

Abstract *This work describes the “CSCT Living Lab”, where a new approach to learning, based on computational thinking and inquiry-based science, is practiced and shaped. Ten high schools, together with seven technical institutions and other three, chosen among vocational ones, have been involved. All these members, represented by teachers interested in joining the process, gave life to the lab, where enhanced scenarios for learning are planned, implemented and traced. An abstraction model is also proposed, to promote a collaborative learning context for teachers and students, derived from effective experiences carried out at University and Politecnico di Torino. A specific example, coming from process control domain, is described to show part of the scenario adopted to make inter-disciplinary learning a sustainable and effective experience.*

Keywords: Problem Posing&Solving, Computer Science, Computational Thinking, Process Control, IDEF0, Advanced Computing Environment

1. Introduction

The Project PP&S100 [Palumbo and Zich, 2012], Problem Posing & Solving, planned for 100 schools, grown to 150 in the meanwhile, is an initiative promoted by the General Directorate of the Italian Ministry of Education, Research and University, aimed at tracking innovations that have recently been introduced in the curricula adopted in school at the upper secondary level.

The purpose of this project, whose stakeholders include the Italian Association for Automatic Computing (AICA), the National Research Center (CNR), the National Industrial Unions, the University and the Politecnico di Torino, is to sustain the framework’s tuning proposed at regulatory level,

Congresso Nazionale AICA 2013

pursuing the transition from "programs specified by the Ministry of Education" to national guidelines for high schools, technical institutions and vocational organizations complying with the EQF [Cedefop, 2012]. This change affects also teachers, who are now required to express a clearer role so that a more responsible and independent design of educational paths be possible for achieving learning outcomes that guidelines have established for the whole education system.

The project is focused on problem posing and solving and aims at sustaining computer science as an enabler of innovation in society and education. Computer science, by its own nature, has an immediate impact in any other scientific domain (physics, chemistry, natural sciences) and also brings a relevant potential within all other knowledge areas, even those placed in the socio-economic field, thanks to the basis provided by trans-disciplinary languages and information representations [Booch, 2013].

This already takes place in ordinary paths when regulations formally decree that mathematics and computer science must be joined together and combined when delivered in a class. Main drawbacks stem from the organization context, which lies on teachers, whose competence profile is often unavailable at school or, when present, unaware to face that kind of task, which requires clear specific skills and experience in the field.

2. Living Lab and Computational Thinking

To deal with that drawback, a "Living Lab" has been founded, which involves ten high school organizations, together with seven chosen from technical institutions and other three got among vocational ones. These institutional members, represented by teachers involved in the process, gave life to the lab, where a new approach to learning is planned and implemented.

Fig. 1 depicts the model drawn to introduce the rationale of the computer science discipline to be proposed in the secondary school, adopting as a background the "Inquiry based Science Learning model" and its specialization, represented by the "Computational Thinking" scenario [Wing, 2008][Yinnan and Chaosheng, 2012]. The first layer in the model hosts the highest abstraction process, where methodologies

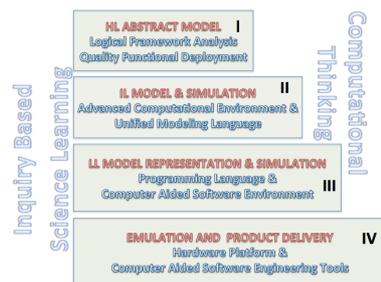


Fig. 1 – The Abstraction Hierarchy

like the Logical Framework Analysis (LFA) [EU, 2004] investigate the essence of the problem, moving from instances promoted by the reference stakeholders, using brainstorm as a tool. They dig into the cause-effect relationships among the various facets that the complex structure of reality often shows, transforming it into the means-aims set in the solution domain. Given that, a further step leads to the release of the objective-purpose-results schema, which, at the end, brings to the activity scheduling and resource planning.

Still at the first layer (I), taking advantage of the "problem-posing" task, partly carried out in the previous phase, a further activity - rooted on a methodology which addresses issues related to the solution specification to be developed in the following layer (II) - can be found. To this purpose the Quality Function Deployment method, used to trace relationships between "customer voice" and "technical features" of the product/ service/ application, is suggested.

At the second layer, which is devoted to the solving perspective, target processes are first identified and analyzed to outline loosely dependent activities. The latter are given Specific functions in order to sustain process interactions, as planned by the IDEF0-based specification [IDEF0, 28/8/2013]. Once the functional model has been fully described, the next envisaged phase, in the same layer, concerns the Advanced Computation Environment (such as MAPLE or MATLAB [Chonacky and Winch, 2005]), where the IDEF model can be drawn to refine the mathematical perspective. Layer III and IV concerning language descriptions and prototype implementation will be dealt with in future works.

To validate the model, the Living LAB focused its attention first on a problem easily perceived by everybody and immediately understood in terms of effects, so that a reference framework could be shaped using available tools widely spread in the CS community. To this purpose a "process control system" was chosen and, in the following, a general abstraction model is shown taking advantage of the IDEF0 graphical notation.

3. Conceptual Modeling of a Process Control System

The choice of "*control theory*" [Doyle and Francis, 1990] is justified by the fact that any kind of process in real life is a subject of control. Even in sociological contexts this scenario happens normally, since control theory itself underlines how weak bonds between individuals and society make people free to go sometimes against the law, or other people having weak ties can engage themselves in crimes to gain benefit. This can be overcome when strong bonds make deviance more costly. Anyway to keep things simpler, a more technical domain has been taken as a stage to trace the example.

Fig. 2 depicts a controlled system taken in climate control domain. The " $\{I_{cs}|$ *Reference Temperature*", being I_{cs} the set of potential inputs for the process control system (CS meaning *Controlled System*), represents the target temperature the controlled system should hold. In other terms it should be the ideal temperature to be achieved within a $\{R_{cs}|$ *Room*", being R_{cs} the set of potential supplies/resources for the whole process, even contrasting disturbances, $\{C_{cs}|$ *windows, doors*) - with C_{cs} representing the set of potential constraints to be applied to the process - which may occur, such as windows or doors left open or other accidents, able to induce a temperature change. The latter, in fact, can be somehow caught at the output, whenever the sampled temperature may not be coherent with the expected value given at the input. So to understand how a controlled system works, the hypothesis is that an appropriate solution should be able to perform repeated reading of the real

temperature achieved by the room under control. This means that the actual temperature should be taken by means of a suitable “{R_{cs} | Sensor}” placed somewhere in the “{R_{cs} | Room}” space. The picture gives more details about the sensor, able to catch the temperature, as “{O_{cs} | CV (Controlled Variable)}”, and to bring it back towards the input. It is interesting to see the representation of the { R_{cs}|Room} assumed in the model description, which, in our case, is simply the room space together with its own heating system, { R_{cs} | Heating System }, without any control or regulation applied. Its input is labeled “{I_{ss}|MV (Manipulated Variable)}” - being I_{ss} the set of inputs for the *Simple System* - and could be assigned, as a physical quantity, to the open angle of the valve regulating the fuel flow to the burner of the heating equipment. The abstraction concept represented here is that of “composition association” where a “Controlled System” is viewed as the composition of other subunits, the “Simple System” without control, which is the environment equipped with its own heating system, the “Sensor” to get samples of the temperature (output), and other units, all of them connected through arrows which represents specific information and/or material flows.

To give an example of the whole system, it is supposed to set the desired input, “{I_{cs} | Reference Temperature}”, to k C° degrees. To get the output temperature, possibly aligned with the desired input specified above, the Controlled System should measure the current value of the output temperature “{ O_{cs} |CV}” shown by the simple system, constituted of a “{R_{cs} | Room}”. A specific sensor, “{R_{cs}|Sensor}”, transforms sampled temperatures into something else, easier to manage, such as a voltage signal. The higher the temperature, the higher the voltage brought back by the sensor to be used for other simple operations. To give an idea of the sensor sensitiveness it could be requested to equip the feedback line with a sensor able to convert a change of 1 C° of the sampled temperature into a variation of 0,01 V for the measured output.

The abstraction depicted graphically in Fig. 2 draws a picture of the model whose development is in progress. For the controlled system to do its job, it is needed to get the voltage fed back by the sensor compared against the desired temperature given as {I_{cs} | Reference Temperature}, in its turn translated into a voltage signal: this comparison can be performed using a simple “Comparator” able to detect the difference between the two physical quantities. The comparator provides the so-called “{E_{cs}|Error Signal}”, able to give the “feeling” of the current difference between the desired input and the actual value experienced by the system to be controlled. The role of the error signal is fundamental, being it the real driver of the regulation/control action exercised on the simple system. The model drawing reported in the picture is completed with the last unit, which is the key to perform the regulation of the simple system (Room).

The picture makes it clear that the Controller takes the error signal as its input and transforms it into its output - actually the simple system input - labeled “{I_{ss}| Manipulated Variable}”, which is the signal able to exercise the control on the simple system. In other words it takes the running difference between the

desired temperature and the actual temperature of the simple system, the error signal, to process it in order to shape an output signal able to act on the valve controlling the fuel to be provided to the burner. The higher the differences computed between the temperatures, the higher would be the signal used to change the position of the valve. Of course much more is required in terms of

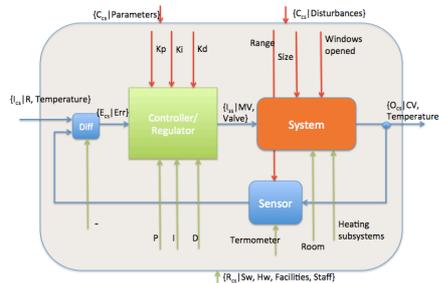


Fig. 2 - A Controlled System

real equipment to be chosen to perform what the model suggests. As an example, if we suppose that a motor is used to exercise control on the open valve, or more simply tune the speed of a fan, attention should be paid to translate the output signal made available by the Controller, coping with the power needed to pilot the chosen actuator. In many cases the Controller is implemented in such a way that both the comparator and

the Controller devices be joined in a single unit. According to this perspective, the controller is expected to carry out two main activities: computing the error signal and applying the computed control effort to the simple system. Looking at the model it is quite immediate to realize that if the gain computed along the forward links - starting from the input error up to the output - is large enough, a small error experienced now might imply fairly large outputs. Let's imagine that, because of the controller, the valve be opened too much with respect to the computed error: as a consequence it is expected that a steep growth of the output temperature might occur. Hence, a strategy should be applied to the choice of the gain values, in fact it is needed to avoid unstable behavior of the Controlled System: in some cases the Controller itself, instead of exercising control, could become the cause of the simple system destruction. Anyway, there is a certain level of coherence with the choice of having high-level gains, since that makes it possible to get a fast recovery of the output with respect to very small errors experienced at the controller input.

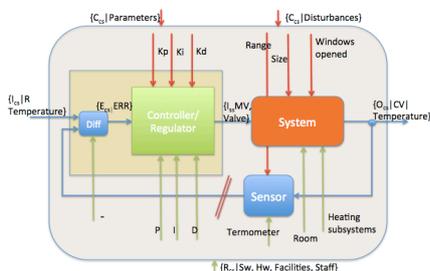


Fig. 3 - Closed and Open loop Control

3.1 Diving more into the Conceptual Model

The control action will be exercised either to vary the heat inflow or intervene on the leaks which determine the heat outflow from the simple system, this should be done to keep the temperature level at the set-point stated in the controlled system. Fig. 3 depicts the general schema of the model for achieving those objectives.

As it can be seen in the picture, the controlled system process can be

represented by a closed loop. A specific sensor monitors the simple system output, "CV Temperature", and the measured signal is fed-back to a comparator, represented by the block called *Diff*, sited close to the input of the controlled system. The second input provided to the comparator is the Reference (R) Temperature, being the comparator output the difference or error signal E computed on the two inputs. The regulator/controller, represented just as a black box, will provide the appropriate correction to maintain the process at its set point even in case of disturbances that may affect the heating system.

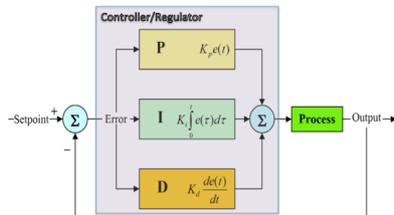


Fig. 4 - The PID Controller

controlled variable. This action is called *open loop operation*.

The best known form of control is given by the so-called "*PID controller*" [Åström and Hägglund, 1995], Proportional, Integral, Derivative control which, as depicted in Fig. 4, is based on the mathematical modeling of those three main actions, usually exercised intuitively in any ordinary individual experience when a behavior correction is required. As an example of this, it is immediate to recall what happens when a correction is to be performed by an exhausted driver as its car in a road is approaching the guardrail. At first the driver evaluates the relative position of the vehicle with respect to the guardrail and operates proportionally with the lateral distance using short turning of the steering wheel (proportional action). The proportional correction is usually applied repeatedly, in that performing what is called the integral action, which takes care of the previous history of the driver reaction. These corrections are also tuned by means of the speed applied to any single action, which should anticipate in short the next behavior of the car (derivative action).

Hence the effective calculation algorithm implemented in the Controller involves three separate parameters, as reported in Fig. 4: the proportional, the integral and the derivative values, denoted P, I, and D. Simply put, these values can be interpreted in terms of time: P depends on the present error, I on the accumulation of past errors, and D is a prediction of future errors, based on current rate of change. The weighted sum of these three actions is used to tune the process via a control element such as the position of a control valve, a damper, or the power supplied to a motor or a heating plant.

Given the absence of knowledge of the underlying system, a PID controller has long been considered as the best controller obtained by tuning the three parameters relevant in the PID controller algorithm, planned to devise control actions specifically designed for the identified process requirements. The features of the controller can be represented in terms of the responsiveness it

shows with respect to an error, given in terms of the degree with which the controller passes the indicated set-point, and the degree of oscillation which the system experiences at its output. On the base of this issue, it should be underlined that the PID algorithm does not guarantee optimal control of the system or even system stability.

3.2 Mathematical Perspective

The mathematical representation of the PID controller, in a form widely used in the industrial domain, is the so-called canonical form, where the K_p gain is applied to the integral and derivative terms, so that, being MV the manipulated variable, the following equivalence holds:

$$MV(t) = K_p \left(e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{d}{dt} e(t) \right)$$

where T_i is the integral time and T_d the derivative time. In this form, the parameters K_p , T_i and T_d have a precise physical meaning; in fact the inner summation computes a new error value, which is compensated taking care of past and future errors. In particular referring to the sum of the proportional and derivative components effectively predicts the error value looking at T_d seconds, or samples in the future. The integral component adjusts the error value to compensate for the sum of all past errors, by planning to cancel them in T_i seconds (or samples). Hence, the resulting compensated error value is scaled by the gain K_p .

There is also the so-called ideal parallel form in literature, given below:

$$MV(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{d}{dt} e(t)$$

where the relationship with the parameters shown in the previous one are the following: $K_i = K_p / T_i$ and $K_d = K_p * T_d$.

3.3 How can it be made understandable to young students?

The simplest way is given by the digital representation of phenomena where the analysis for designing a digital implementation of a PID controller in a microcontroller device requires the standard form be discretized. Approximations for first-order derivatives are made by backward finite differences. The integral term is discretized, with a sampling time Δt , as follows:

$$\int_0^{t_k} e(\tau) d\tau = \sum_{i=1}^k e(t_i) \Delta t$$

while the derivative term is approximated as:

$$\frac{de(t_k)}{dt} = \frac{e(t_k) - e(t_{k-1})}{\Delta t}$$

Thus, a velocity algorithm for implementing the discretized PID controller in a microcontroller is yielded by differentiating $u(t)$ and also using the numerical definitions of the first and second derivative, then solving it for $u(t_k)$ getting the final result:

$$u(t_k) = u(t_{k-1}) + K_p \left[\left(1 + \frac{\Delta t}{T_i} + \frac{T_d}{\Delta t} \right) e(t_k) + \left(-1 - \frac{2T_d}{\Delta t} \right) e(t_{k-1}) + \frac{T_d}{\Delta t} e(t_{k-2}) \right]$$

The physical meaning of this last expression is easily understandable since it computes the new value for the manipulated variable, $u(t_k)$ as the old one, $(u(t_{k-1}))$, compensated with the previous values of the error samples ($e(t_k)$, $e(t_{k-1})$, $e(t_{k-2})$), - got on a temporal window extended to two only previous samples, besides the current one -, suitably weighted, because of the derivative and integral contributions.

3.3 The System Model and its Simulation

Once a simple mathematical expression has been obtained to describe what seemed so complex at a first glance, the next step concerns the way a model can be perceived according to the typical sensing process human brain puts in place, whenever an abstract representation has to be validated. © MapleSim comes to the aid providing simulation scenarios on the basis of the many and various mathematical descriptions for several devices and phenomena made available by still growing and rich libraries.

MapleSim is a "multi-domain" system software simulator that allows analysts to create and simulate physical systems representations. Create a model means building the approximate description of a given physical system with a notation based on the combination of its main components, most of them often retrieved from already existing system libraries. That model is then able to generate its own mathematical representation in terms of a set of equations, solving that mathematical system, after proper simplification, leads to the definition of the physical quantities featuring the system in its temporal dynamics.

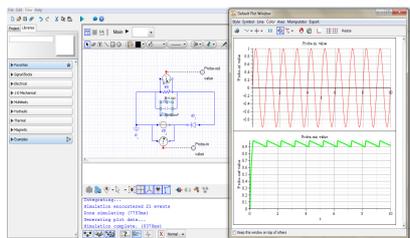


Fig. 5 - MapleSim user interface

MapleSim takes advantage of the © MAPLE [Chonacky and Winch, 2005] system software that provides a rich set of numerical and symbolic features becoming the real "engine" for any provided service. Maple and MapleSim together provide also support functionalities, such as graphical display of results, capability to extract specific analysis from mathematical descriptions such as, for example, other models given in a descriptive or programming languages such as C, further templates, and so forth. In particular MapleSim is able to describe and simulate systems having different physical nature: electrical, mechanical, thermal, hydraulic, to mention the most important. For this reason it is often referred to as a multi-domain environment.

Anyway, before deepening the description of the test-bed, a brief presentation of the user interface of the system is given, as depicted in Fig. 5, to underline the effectiveness of its organization, which makes interaction easy ensuring immediate understanding of operations. To this purpose it is organized into several sections, labeled panels, each devoted to a specific task and command. The central zone is the working area, where the analyst can build the model, taking the basic components from libraries organized by type. At the top are sets of controls grouped for handling model simulation providing a

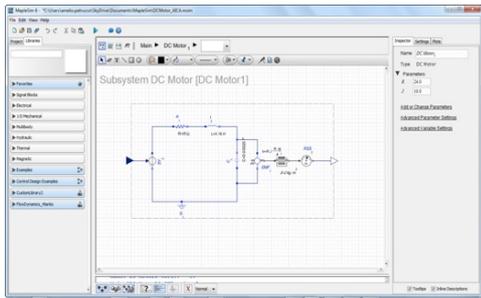


Fig. 6 – DC motor model in MapleSim

the results of a simulation.

Coming back to the construction phase of the control system model for a DC motor, the activity sequence starts with the description of the model for the engine to be controlled as shown in Fig. 6.

Fig. 7 reports a simulation result related

hierarchical navigation. Editing commands for the contents are also included: on the right side, in the figure covered with the window showing signals graphical images, the View menu together with a set of parameters associated with any component can be found. The window showing waveforms displays

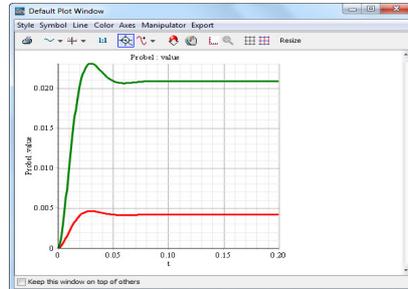


Fig. 7 - DC motor simulation results

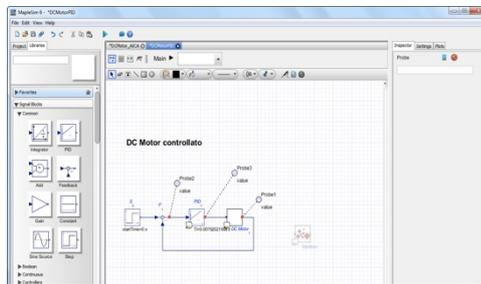


Fig. 8 - PID control for the DC motor

first linearized, then a control block, the PID controller, is built. Of course the linearization makes use once again of the link Maple-MapleSim through the creation of a specialized "template" for the generation of the model, Fig. 8 reports the result of this operation.

to a small motor tested for two different values of power (1 and 5 volts respectively), in this specific case the angular velocity of the shaft is plotted in rad / sec.

The introduction of a PID control is carried through two steps as described in this example, the mathematical model of the motor is

4. Conclusions

The “CSCT Living Lab” has been presented as a technological and social context where a new approach to learning is being experienced, working on the base of the conceptual stage offered by computational thinking and inquiry-based science. The Lab consists of teachers interested in joining the process gathered from ten high schools, seven technical institutes and three vocational ones, which have also been involved according to their role of institutions, by the Ministry itself. In the Lab all attendees, following a collaborative learning approach, promote enhanced scenarios for learning,

which should impact on 400 students for this year, laying the foundation to deploy the experience on traditional high school in the next years, increasing considerably the whole impact of the process.

An abstraction model was also proposed to help teachers and students share working contexts and experiences carried out at University and Politecnico di Torino.

The model was drawn to introduce the rationale of a new computer science discipline to be proposed in the secondary school, adopting as a background the "Inquiry based Science Learning model" and its specialization, the "Computational Thinking" scenario. A specific example, coming from process control domain, was also described to show part of the scenario adopted to make inter-disciplinary learning a sustainable and effective experience, so that the conceptual model could be validated.

The choice of "*control systems*" was justified by the fact that any kind of process in life is a subject of control. Anyway to keep things simpler, a more technical domain, related to climate control, has been selected as a stage to trace an example, which has been dealt with following the given abstract model and focusing mainly on its second layer, where more detailed specifications and simulations can be carried out using a set of suitable tools and methodologies.

References

[Palumbo and Zich, 2012] Palumbo, C., Zich, R., *Matematica e informatica: costruire le basi di una nuova didattica*, Bricks, Anno 2, n. 4, ISSN 2239-6187, 2012, pp. 10-19.

[Wing, 2008] Wing, J. M., *Computational thinking and thinking about computing*, *Phil. Trans., R. Soc.*, 2008, 366, 3717-3725.

[Chonacky and Winch, 2005] Chonacky, N., Winch, D., *3MS for Instruction: MAPLE, MATHEMATICA, and MATLAB*, *Computing in Science & Engineering*, Volume: 7, Issue: 4, July/August, 2005.

[Cedefop, 2012] Cedefop, *Analysis and overview of NQF developments in European countries*, Annual report 2012, © Cedefop, 2013.

[IDEF0, 28/8/2013] IDEF0, <http://www.idef.com/IDEF0.htm>, access: 38 August 2013.

[EU, 2004] EU Commission. 2004. *Project Cycle Management Guidelines*. pp 57-94.

[Doyle and Francis, 1990] Doyle, J., Francis, B., Allen Tannenbaum, *Feedback Control Theory*, Macmillan Publishing Co., 1990.

[Åström and Hägglund, 1995] Åström, K. J., Hägglund, T., *PID Controllers: Theory, Design, and Tuning*, ISA, Research Triangle Park, North Carolina, 1995.

[Booch, 2013] Booch G., *From Minecraft to Minds*, *On Computing*, IEEE Software, April, 2013.

[Yinnan and Chaosheng, 2012] Yinnan, Z., Chaosheng L., *Training for Computational Thinking Capability on Programming Language Teaching*, *Computer Science & Education (ICCSE)*, 2012 7th Int. Conf., 14-17 July 2012, pp. 1804-09. ISBN: 978-1-4673-0241-8.

La parola agli studenti: le ICT a scuola

Isabella Donato
Phd Student LCTIC (Lingue Culture e Tecnologie dell'Informazione e della Comunicazione)
Scuola di Nuove Tecnologie per le Scienze Umane e Sociali
Università degli Studi di Genova
isabella.donato@alice.it

This contribution reports the results from surveys designed to measure high school student review on the use of educational technology. The responses attested to the interest of the student towards the technologies applied to teaching, but also the limits of ICT.

Keyword: digital technology; digital natives; teachers

1. Introduzione

Negli ultimi anni il Ministero dell'Istruzione ha avviato un Piano nazionale per la scuola digitale, investendo 100 milioni di euro nell'acquisto di più di 34.800 LIM e nella formazione di 83.000 docenti; inoltre ha finanziato 416 classi su tutto il territorio nazionale con "Azione cl@sse 2.0", coinvolgendo 1.350 docenti e 13.500 studenti, con un impegno finanziario di 3.750.000 euro.

Ascoltando la voce dei docenti, però, emerge anche un'altra realtà: infatti i dati sulla diffusione dell'Information and Communication Technologies (d'ora in poi, nel testo, ICT) andrebbero correlati al loro uso effettivo. La formazione rivolta ai docenti appare, nella pratica, inferiore rispetto alle effettive esigenze (l'ISTAT conta nel 2011 più di 680.000 docenti): perciò accanto a esperienze di eccellenza nell'uso delle tecnologie, convivono situazioni in cui, per esempio, le LIM non sono utilizzate o lo sono in modo limitato. Uno dei punti cruciali, quindi, del gap tra la diffusione e l'utilizzo degli strumenti tecnologici appare essere proprio la formazione: la maggior parte degli insegnanti coinvolti nel progetto di digitalizzazione, pur esprimendo un giudizio positivo sull'introduzione delle tecnologie e sui cambiamenti conseguenti degli studenti, osservati in termini di impegno e motivazione, individua da un lato la necessità di personale di supporto, dall'altro la mancanza di incentivi economici [OECD, 2013], facendo emergere diversi errori in questa politica di innovazione e diffusione degli strumenti informatici [Parigi, 2009].

Questa è la situazione vista dalla parte degli insegnanti, ma dalla parte degli studenti? Come dei liceali giudicano l'introduzione di strumenti tecnologici in classe? Percepiscono dei cambiamenti nella didattica dei docenti? Il loro parere è importante non solo perché diretti interessati e protagonisti della scuola, ma anche in quanto espressione di una generazione. Infatti si tratta di studenti tra i 14 e i 19 anni, nati, cioè, negli anni Novanta. Già la generazione dei ragazzi degli anni Ottanta è stata etichettata con varie espressioni: "generazione Y", termine apparso nel '93 sulla rivista Advertising Age, "Millennials" [Strauss e Howe, 1992], "Net Generation" [Tapscott, 1998], "nativi digitali" [Prensky, 2001] o "I-Generation" [Rosen, 2010]. Le caratteristiche di apprendimento sono differenti, perché modellate e prodotte dall'uso degli strumenti tecnologici; e i cambiamenti sono incessanti, tanto che ora si parla di "digital wisdom" [Prensky, 2009]. Parallelamente a questa serie di trasformazioni sarà interessante osservare quando il gap tra nativi e immigrati digitali sarà azzerato, cioè quando entreranno nel mondo della Scuola docenti nati dopo gli anni '80. Ora, infatti, considerato che gli ultracinquantenni nella Scuola di Secondo Grado sono oltre il 50 % [EACEA et al, 2012; OECD, 2013a], la diffusione di pratiche innovative nella didattica è ostacolata non solo dalla mancanza di formazione e aggiornamento degli insegnanti, ma dalle loro abitudini nelle relazioni con la tecnologia. Infatti, se consideriamo i dati ISTAT (<http://www.istat.it/it/archivio/78166>), dopo i 54 anni la percentuale di individui che utilizza il personal computer e internet è inferiore al 50 %, mentre quella dei giovani tra gli 11 e i 34 anni è di, rispettivamente, 78 e 76%. Il divario sembra confermato anche dall'indagine conoscitiva sulla condizione dell'Infanzia e dell'Adolescenza in Italia nel 2012, realizzata da Telefono Azzurro ed Eurispes: solo il 17% del campione degli studenti afferma che in classe la LIM è utilizzata quotidianamente.

2. Scenario

Questo contributo nasce dall'intento di riflettere su come un gruppo di studenti, di età ed esperienze diverse, vive l'utilizzo delle tecnologie in un Liceo di Savona che ha puntato molto sulla dotazione tecnologica: nel giro di qualche anno, le LIM in aula sono passate da 3 (dato del settembre 2012 da Scuola in chiaro http://www.datiopen.it/it/opendata/Attrezzature_multimediali_nelle_scuole_italiane?metadati=showall#dettagli-sul-dato) a 41 per l'inizio dell'anno scolastico 2013/2014.

La parola agli studenti: le ICT a scuola

In particolare questa indagine si propone di riflettere sulle strategie didattiche, presupponendo che esista una relazione tra uso delle ICT e andamento scolastico degli studenti. Infatti, notando prestazioni inferiori degli allievi, evidenziate da valutazioni insufficienti, proprio in parallelo all'incremento dell'uso di LIM e computer, si è ipotizzato che tra i due elementi vi fosse un rapporto che doveva essere ripensato nelle forme fino a quel momento proposte, riorientando la metodologia, a partire da un questionario con domande finalizzate a conoscere l'opinione degli studenti sull'uso delle ICT nell'attività didattica.

Certamente è difficile parlare di dati oggettivi nel rendimento scolastico, perché il riferimento a eventuali miglioramenti in termini di votazione non è un valore assoluto, dipendendo dal giudizio del singolo docente. Per esempio le indagini in corso sulla valutazione dei progetti di innovazione che sono curati dalla Fondazione Agnelli e dalla Compagnia di San Paolo, faranno riferimento ai punteggi dei test, tra prove di ingresso e uscita degli esami nella scuola media. Sull'uso delle tecnologie nella didattica abbiamo il punto di vista degli insegnanti, che, certo, ne misurano l'efficacia sulla base dei feedback degli studenti, anche se sull'attendibilità dei giudizi forniti dai docenti si può leggere un quadro riassuntivo in Parigi [Parigi, 2009].

Chi scrive ha fatto riferimento alle classi in cui insegnava: ciò può costituire, anche se in modo empirico, un ulteriore elemento di valutazione, conoscendo il tipo di didattica applicata alle classi, gli strumenti a disposizione e le caratteristiche generali dell'utenza. D'altra parte, la stessa considerazione può essere valutata come un limite alla ricerca, perché la rende circoscritta: sarebbe necessario estenderla ad altre scuole, dello stesso territorio o di zone diverse.

E' stato scelto di sottoporre un questionario al campione composto da 101 studenti (su una popolazione di più di 900 iscritti), appartenenti a 5 classi (4 del primo biennio e 1 del secondo biennio del Liceo), di diversi indirizzi, ma ciò non comporta differenze rilevanti, in quanto nel piano di studi non sono presenti materie come Informatica, Linguaggi e comunicazione, che potrebbero condizionare alcune risposte.

Per la classe del secondo biennio (una IV) sono stati distinti i risultati, ritenendo che l'età e il percorso scolastico potessero fornire risposte interessanti. Questa classe, infatti, come è normale che sia in considerazione all'età, non solo ha maturato abilità e competenze maggiori rispetto a studenti di 14 e 15 anni, ma ha avuto esperienze dirette di uso delle ICT (alcuni ragazzi partecipano a un Gruppo su Google).

Chiameremo Gruppo 1 quello formato dagli studenti del primo biennio (84 elementi) e Gruppo 2 quello formato dagli studenti del secondo biennio (17): pur non essendoci equilibrio nel numero dei gruppi, può essere utile ragionare sui casi in cui le risposte si discostino in modo netto.

3. Questionario

Il questionario on line è stato somministrato tra la fine di febbraio e i primi giorni di marzo alle classi, che hanno risposto utilizzando un link in un post protetto da password sul blog del docente. La scelta di inserire 10 domande chiuse di due tipologie (dicotomica e con molteplici risposte), escludendo inevitabilmente alcuni temi e quesiti, è derivata dalla necessità di creare un questionario semplice e snello, compilabile in pochi minuti per non impegnare eccessivamente gli studenti.

Per il rispetto della privacy, non sono stati richiesti alcuni dati come il sesso, la classe specifica di appartenenza, il rendimento scolastico.

Le domande poste sono state organizzate su 4 argomenti:

- a. Tecnologia a scuola: sì o no? (3 domande)
- b. Studio e Tecnologie: la mia esperienza (3 domande)
- c. Suggerimenti e proposte – LIM (2 domande)
- d. Suggerimenti e proposte – Aula di informatica (2 domande).

Per l'argomento a) si chiedeva di esprimere un parere sull'utilità delle ICT a scuola, indicando le motivazioni sia delle risposte affermative che di quelle negative.

Per il b) gli studenti dovevano ragionare su eventuali miglioramenti nello studio domestico e nel rendimento scolastico, conseguenti all'uso delle ICT, e su quali attività hanno ritenuto più efficaci in classe.

Per gli argomenti c) e d) gli intervistati hanno avuto la possibilità di dare suggerimenti sul modo più efficace di usare i due principali strumenti a loro a disposizione, la LIM e i computer dell'aula di informatica.

4. Osservazioni

L'analisi descrittiva dei dati del questionario mostra, per l'argomento a), come la grande maggioranza degli studenti (88 su 101) si schierino a favore delle tecnologie a Scuola. Interessante è ragionare sui 13 che le ritengono inutili, perché 12 appartengono al Gruppo 1 e solo 1 al Gruppo 2 (Fig. 1): quindi sono gli studenti più giovani a essere maggiormente scettici. Essendo i questionari anonimi, anche per garantire una maggiore libertà nelle risposte, è impossibile stabilire se l'andamento scolastico e la motivazione personale degli studenti abbiano influito: si può ipotizzare che, in qualche caso, uno studente poco motivato e con risultati deludenti sotto il profilo scolastico (il che può essere anche dovuto a un orientamento sbagliato, visto che il dato più significativo appartiene alle classi del primo biennio) abbia un atteggiamento negativo verso l'intero mondo della scuola.



Fig. 1 – Grafico dei risultati delle risposte alla domanda 1 “Credi che le tecnologie possano essere utili a scuola?”

Un risultato interessante è che le tecnologie sono causa di distrazione durante la lezione, soprattutto per le classi del primo biennio (Fig. 2). Eppure questi strumenti non dovrebbero produrre nessun tipo di curiosità né suscitare l'effetto "wow" [Beauchamp G. e Parkinson J., 2005] per i nativi digitali. Si può pensare che gli studenti vivano la Scuola come un mondo separato, in cui è bene mantenere i ruoli e le conoscenze distinte? Uno studio su giovani lavoratori [Rapetti e Cantoni, 2010] ha prodotto riflessioni analoghe: un ambiente formalizzato (come può essere quello scolastico o lavorativo) non produce spontaneità e gli strumenti tecnologici sono utilizzati in modo diverso rispetto al tempo libero.

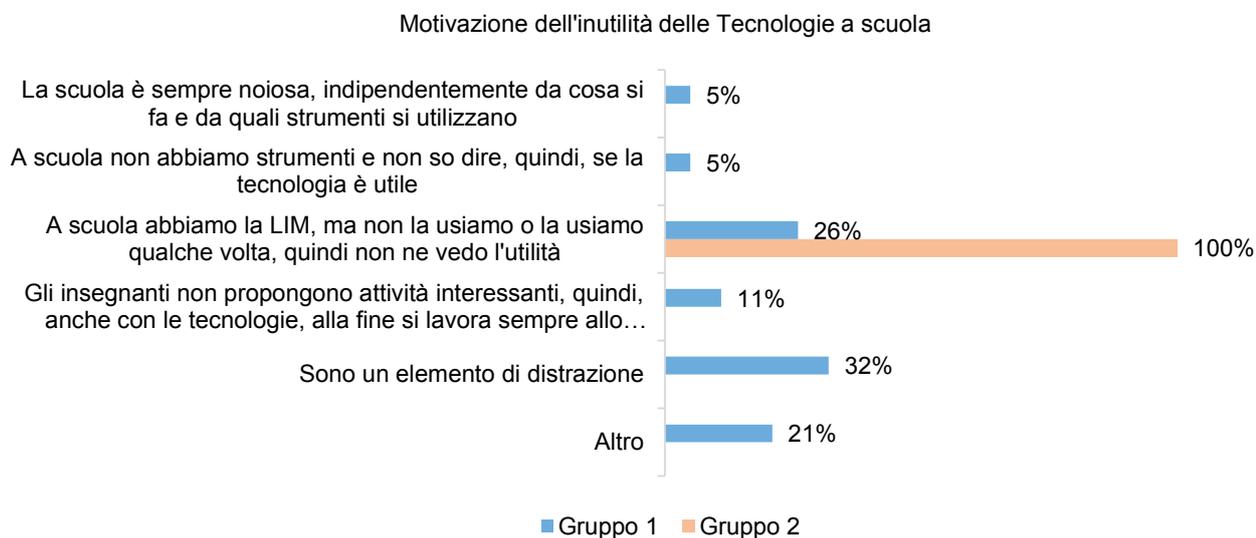


Fig. 2 – Grafico dei risultati delle risposte alla domanda 2 “Se hai risposto no alla domanda 1, scegli le motivazioni”

Forse la componente dell'età prevale comunque, anche quando alcuni comportamenti dovrebbero essere standardizzati e abituali, come l'uso degli strumenti tecnologici dai parte dei nativi: quindi questi studenti, prima ancora di essere nativi digitali, sono adolescenti poco concentrati e con la tendenza a distrarsi.

Non possono escludersi neppure ipotesi riguardanti il pregresso percorso scolastico, cioè l'aver avuto o meno esperienze di uso delle ICT (aspetto che non è valutabile sia per l'anonimato, come già detto, sia perché la provenienza degli studenti è eterogenea).

Un altro aspetto emerso riguarda l'impreparazione di alcuni docenti, che non sanno utilizzare gli strumenti. Questo, come già evidenziato, è uno dei problemi di fondo del piano di digitalizzazione della Scuola italiana (è quello che viene rimproverato anche dall'UE, cioè una politica di acquisto di strumenti a cui non si accompagna un'educazione all'uso rivolta agli insegnanti. Anche per questo nasce il progetto europeo Teaching and Training Quality Management, TQM <http://www.tqmproject.eu> volto a promuovere la cultura del miglioramento della qualità nel sistema educativo). Di conseguenza è difficile per gli studenti vedere le potenzialità degli strumenti tecnologici e averne una visione diversa, cioè impiegata anche in un contesto che non sia quello personale dello svago e del gioco.

Tra gli studenti che hanno testimoniato l'utilità delle ICT nella didattica (Fig. 3), la motivazione è essenzialmente di carattere pratico: durante la lezione si sentono più coinvolti e a casa studiano più facilmente, perché possono scaricare materiali e appunti. Dalle risposte del Gruppo 2 emerge una maggiore consapevolezza del metodo di studio, perché gli studenti si distraggono meno e, stando più attenti, capiscono meglio i concetti; l'altro gruppo non sottolinea tanto questo aspetto, quanto quello di sentire la scuola più vicina al mondo giovanile.

Viene confermata la tendenza degli studenti di utilizzare individualmente i materiali offerti. Anche quando sono a disposizione strumenti che possono prevedere l'interazione virtuale tra i membri del gruppo (scrittura di post sul blog o sul gruppo di Google), pochi la realizzano, confermando la propensione a un uso individuale delle ICT nell'ambito didattico, che continua a essere registrata anche all'Università [Albano et al, 2011].

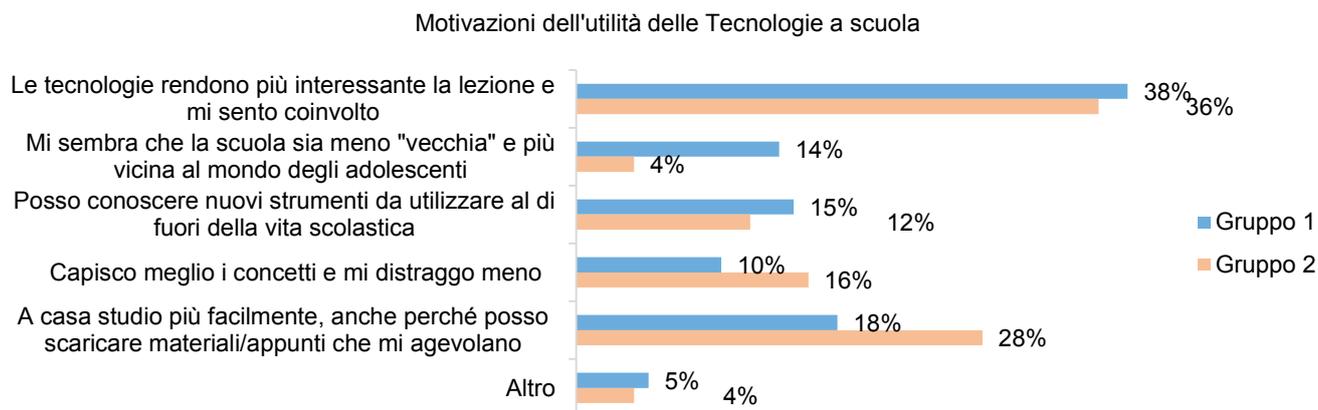


Fig. 3 – Grafico delle risposte alla domanda 3 “Se hai risposto sì alla domanda 1, scegli le motivazioni”

Emerge, poi, una delle peculiarità delle ICT, cioè poter rendere partecipativa, interattiva e stimolante la lezione. Alcuni studenti sottolineano come imparare l'uso di strumenti e applicazioni può essere utile anche al di fuori della scuola: si può osservare che, pur essendo “nativi digitali”, la conoscenza degli strumenti informatici è abbastanza limitata, come emerge dall'osservazione delle attività svolte in classe (per esempio è stato necessario realizzare un tutorial per l'uso dei Gruppi di Google). Quindi sembra di poter concordare con coloro che sono più critici rispetto all'ottimismo verso le capacità tecnologiche degli adolescenti, che, da prove empiriche, non posseggono quelle competenze descritte come tipiche dei nativi digitali [Martini et al, 2011]. Quindi bisogna riflettere su cosa intendiamo per nativi digitali: adolescenti che hanno maggiore facilità nell'uso di questi strumenti, un approccio più spontaneo perché fanno parte della loro quotidianità, ne capiscono l'uso con la pratica a differenza degli immigrati che cercano di capire prima come funziona uno strumento [Ferri, 2011], ma hanno da imparare qualcosa anche dai docenti che conoscono le ICT.

Il fatto che gli studenti italiani siano lontani da quelli americani descritti da Prensky emerge anche da ricerche su studenti universitari: nonostante la familiarità con internet, l'utilizzo che questi ne fanno è legato, in ordine, a ricerche, posta elettronica e infine didattica [Cinque e Pensieri, 2009].

Per quanto riguarda lo studio domestico (argomento b, Fig. 4), metà degli studenti ha la percezione di studiare più facilmente perché ha già i materiali organizzati, anche se il tipo di attività svolta in classe è determinante (cioè solo per alcuni argomenti l'utilizzo dei materiali presentati e/o prodotti in classe è di aiuto). Questo aspetto è sottolineato dal Gruppo 2, che dimostra una maggiore consapevolezza riguardo allo studio; una parte dell'altro gruppo non nota differenze (17%), ma alcuni hanno più difficoltà perché si distracono (7%) e hanno problemi a gestire i materiali (4%). Possiamo immaginare che gli studenti che hanno segnalato queste difficoltà siano da identificare con coloro che inizialmente hanno affermato che la tecnologia a scuola è inutile.

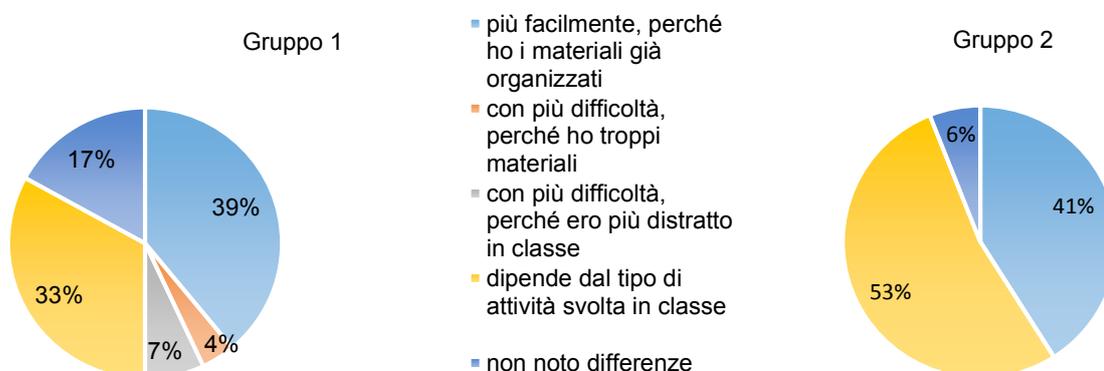


Fig. 4 – Grafico delle risposte alla domanda 4 “Quando durante la lezione sono usati gli strumenti informatici, hai la percezione di studiare a casa”

Gli strumenti a disposizione degli studenti (LIM e computer nel laboratorio di informatica) risultano ugualmente utili per diverse attività: la LIM consente di produrre appunti e cercare informazioni in internet durante la lezione; il laboratorio di informatica serve per fare ricerche, ma gli studenti indicano come poco efficace la lezione in cui non hanno il controllo del computer (per esempio mentre il docente sta mostrando una presentazione).

Gli studenti, dovendo scegliere le attività che ritengono più efficaci, rispondono diversamente a seconda dell'età: il Gruppo 1 vede come aspetto principale prendere appunti, utilizzare internet anche per fare ricerche; il Gruppo 2 ritiene più utile visionare materiale, svolgere attività guidate, prendere appunti e solo dopo utilizzare internet, che è associato, spesso, a una maggiore libertà, che sono gli studenti più giovani a ricercare. L'uso della LIM per fare esercizi non sembra interessare, contravvenendo all'idea che la tecnologia e i suoi strumenti generino sempre collaborazione e cooperazione.

A seconda delle attività svolte i due gruppi evidenziano miglioramenti nell'andamento scolastico (Fig. 5). Capire se questi siano prodotti o influenzati dalle tecnologie è difficile da stabilire, utilizzando come parametro le valutazioni ottenute dagli studenti, proprio, come già scritto, per gli elementi soggettivi presenti all'interno del processo valutativo. Gli studenti possono giungere a considerazioni empiriche e registrare progressi, soprattutto nel metodo di studio, che può significare non solo una crescita delle valutazioni, ma semplicemente un'organizzazione diversa e più funzionale del tempo dedicato a casa allo studio, una minore fatica a fronte dell'introduzione di elementi come appunti digitali, schemi e mappe. I numeri di chi non nota miglioramenti potrebbero essere accostati a quelli che hanno mostrato scetticismo sull'uso delle tecnologie a scuola.

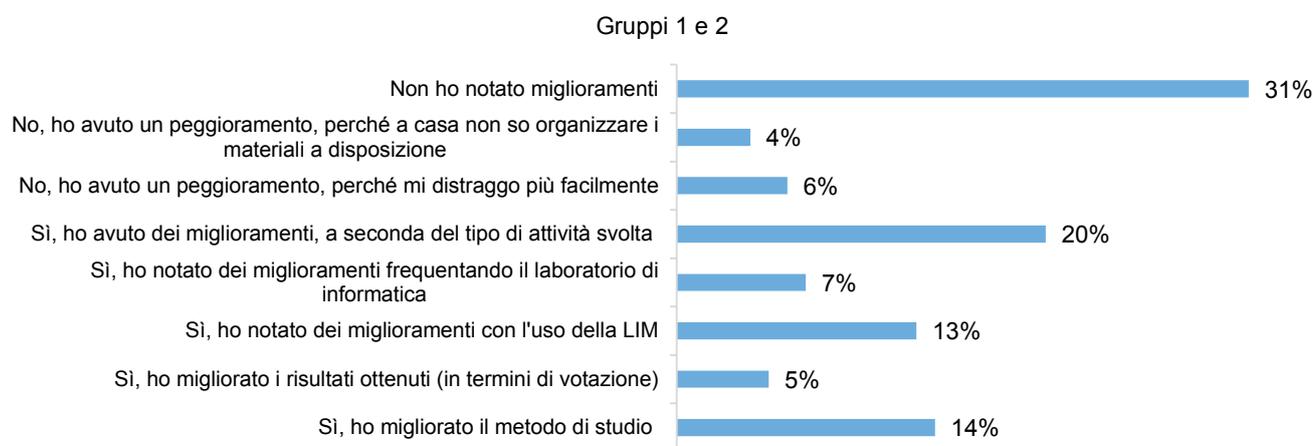


Fig. 5 - Grafico del totale delle risposte alla domanda 6 "Sulla base della tua esperienza, puoi associare un miglioramento del tuo rendimento scolastico all'uso delle tecnologie in classe?"

Per quanto riguarda l'uso della LIM (argomento c), i giudizi degli studenti concordano: l'utilità consiste nel fare schemi, ricerche, archiviare il materiale e creare appunti; la possibilità di vedere film è scelta dagli studenti del Gruppo 1, mentre si presenta come ultima opzione per l'altro gruppo. Sembra confermata la tendenza a prediligere l'aspetto visivo e soprattutto la preferenza verso schemi e appunti fatti dal docente e ciò denota negli allievi una scarsa capacità di autonomia in queste operazioni.

Sui fini didattici della LIM (Fig. 6), le risposte si differenziano: per gli studenti del Gruppo 1 rimangono confermate la produzione di appunti, la ricerca in internet, la possibilità di riprendere concetti spiegati in precedenza o anche registrare la lezione; gli studenti del Gruppo 2 vedono invece l'aspetto accattivante e coinvolgente della lezione, seguito dalla possibilità di confrontare e integrare gli appunti, attività che nella classe IV gli studenti fanno autonomamente.

Queste risposte testimoniano la differenza nella metodologia di studio: gli studenti più giovani non si sono ancora appropriati di pratiche come quella di prendere appunti e cercano un aiuto; quelli che posseggono già questa abilità, vogliono controllare ciò che hanno scritto e si sentono responsabili del loro apprendimento.

Per quanto riguarda l'uso del laboratorio di informatica (argomento d), le risposte, anche in questo caso, si dividono: per il Gruppo 1 l'aspetto più importante è il lavoro a gruppi, testimoniando come sia necessario quel cambiamento nella didattica, che proprio le tecnologie ci spingono a fare, per esempio, con la didattica laboratoriale. L'idea di cercare autonomamente in internet supplisce, forse, un uso domestico non del tutto libero e una voglia di indipendenza. Per il gruppo 2 la preferenza per la ricerca in internet è legata a un argomento dato, per realizzare prodotti multimediali: è chiara la tendenza a essere anche produttori di materiali e sperimentare una maggiore autonomia.

In generale le risposte testimoniano una compilazione ragionata del questionario. Anche in questo caso, per l'aula di informatica, il Gruppo 1 vede più utile la possibilità di fare ricerche, mentre il 2 l'utilità delle presentazioni per seguire le lezioni (Fig. 7). Per tutti la possibilità di disporre di un collegamento internet è un elemento importante.

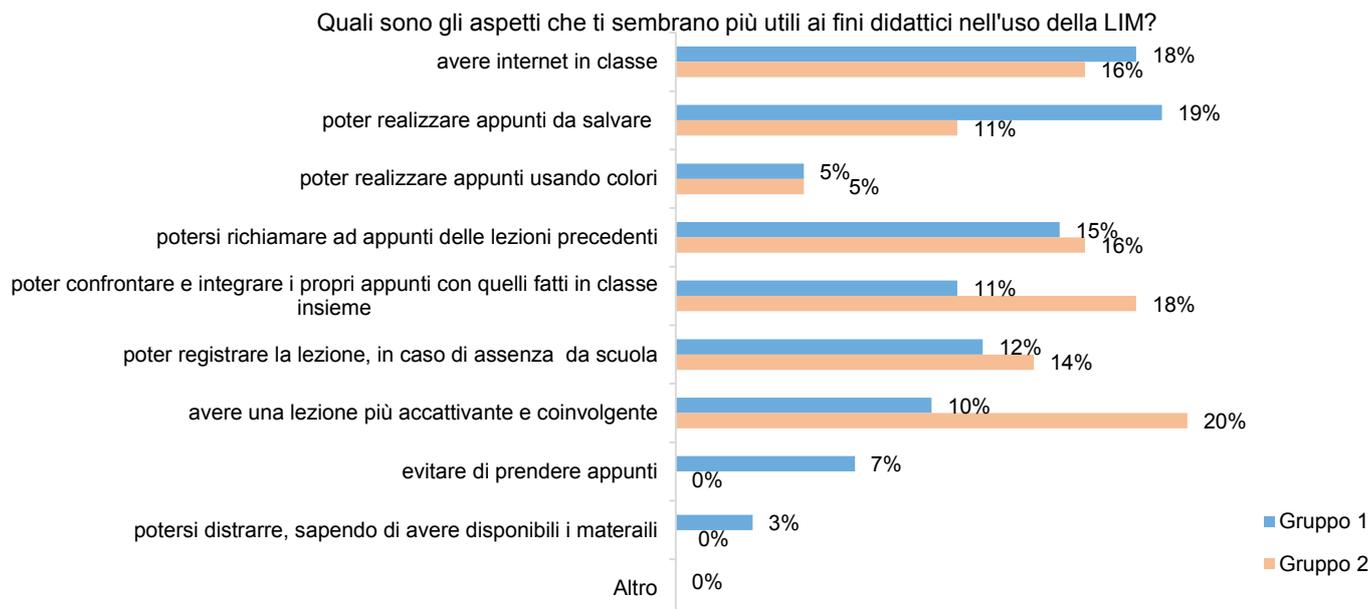


Fig. 6 - Grafico delle risposte alla domanda 8

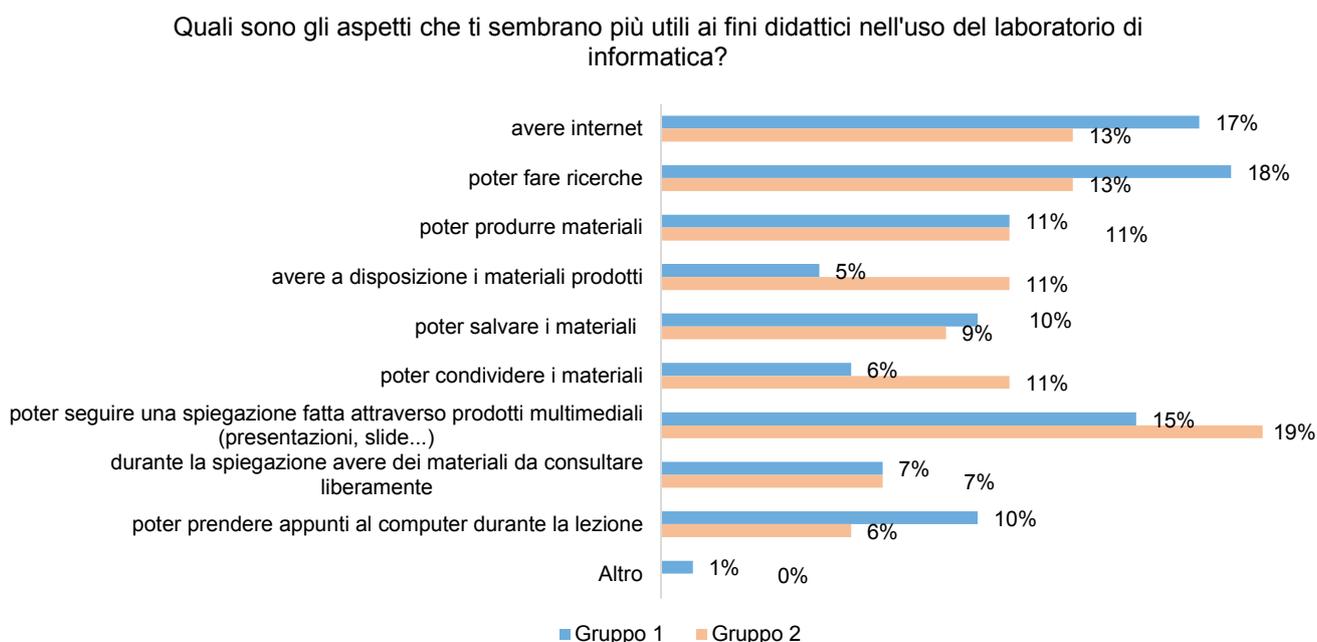


Fig. 7 - Grafico delle risposte alla domanda 10

5. Conclusioni

L'obiettivo di questa ricerca era conoscere l'opinione di un gruppo di studenti sull'introduzione delle ICT a scuola; il fatto che siano nativi digitali non deve portare a escludere a priori un giudizio critico sulle loro competenze digitali e sulla qualità delle relazioni che stabiliscono con gli strumenti.

L'utilizzo delle tecnologie sembra associarsi al conseguimento di risultati positivi, come diverse ricerche testimoniano [Pedrò 2006, 2008; CERI-OECD, 2010]. In questo caso i dati del questionario non hanno confermato questo legame, dal momento che solo una parte degli studenti ritiene che si sia verificato un miglioramento, mentre la maggioranza mantiene posizioni più caute. Quello che si può stabilire è che il semplice uso delle tecnologie non è sufficiente per fare una didattica innovativa con ricadute positive sugli studenti (anche solo in relazione a motivazione e interesse): è indispensabile integrare questi strumenti nella didattica che deve essere ripensata, proprio in funzione delle potenzialità che offrono le ICT, non solo in termini tecnologici, ma anche pedagogici. La dotazione tecnologica della Scuola deve essere accompagnata a un uso consapevole e pedagogico degli strumenti.

La parola agli studenti: le ICT a scuola

Il giudizio da parte degli studenti è critico, consapevole e ragionato: le tecnologie non vengono chieste e accettate semplicemente perché sono ormai largamente diffuse, ma è richiesto un uso funzionale alle esigenze degli studenti. Sembra che questi vivano la tecnologia a scuola in modo passivo, cioè come spettatori, per questo manifestano la necessità di avere più autonomia per lavorare, apprendendo proprio dall'esperienza, in linea con quanto emerge dall'Indagine OCSE-PISA-CERI New Millennium Learner [Pedrò, 2008].

La LIM e il computer (ma anche il tablet, di cui non si è parlato perché non presente nella scuola dove è stato individuato il campione di studenti) non garantiscono partecipazione, collaborazione e miglioramento del rendimento scolastico degli allievi: anzi, se usati in modo superficiale, possono determinare l'effetto opposto, come ha indicato il dato sulle tecnologie fonte di distrazione.

Soprattutto gli studenti del primo biennio necessitano di un'attenzione particolare nella scelta della metodologia per evitare il rischio di overflow di informazioni. Per esempio per evitare che la possibilità di realizzare facilmente appunti e schemi si trasformi in uno svantaggio, può essere utile stabilire i materiali obbligatori, quelli essenziali per gli obiettivi stabiliti dal docente, distinguendoli nettamente da quelli prodotti per esercizio o per approfondimento, in modo da non disorientare gli studenti che non posseggono ancora un valido metodo di studio e di organizzazione.

Si rende necessario un forte segnale di cambiamento nella didattica, proponendo laboratori che possano, attraverso l'esperienza, andare incontro agli stili di apprendimento degli studenti. Un primo passo può consistere nella drastica riduzione della lezione frontale a favore di un approccio che prediliga una partecipazione più consapevole e attiva degli allievi, proprio attraverso le ICT, potenziando per esempio l'uso della LIM, che riveste un ruolo fondamentale per le attività, con ricadute anche sullo studio domestico degli studenti.

Bibliografia

[Albano et al, 2011] Albano G., D'Aprile M., Maccarini R., Opinioni di studenti su pratiche di blended learning. TD Tecnologie Didattiche, 19 (3), 2011, 189-194.

[Beauchamp G. e Parkinson J., 2005] Beauchamp G. Parkinson J., Beyond the 'wow' factor: developing interactivity with the interactive whiteboard, School Science Review, 86, 2005, 97-103

[CERI-OECD, 2010] CERI-OECD, Are New Millennium Learners Making the Grade? Technology Use and Educational Performance in PISA, CERI-OECD, Paris 2010 (disponibile all'indirizzo www.oecd.org/edu/ceri/45000441.pdf ultima consultazione 20/07/2013)

[Cinque e Pensieri, 2009] Cinque M., Pensieri C., Il progetto Campus We-Com. L'attitudine degli studenti universitari verso l'innovazione didattica. Je-LKS, 5, 1, 2009, 57-65.

[EACEA et al, 2012] EACEA, Eurydice, Eurostat, Key Data on Teaching Languages at School in Europe, 2012

[Ferri, 2011] Ferri P., Nativi Digitali, Bruno Mondadori, Milano, 2011.

[Martini e Cinque, 2011] Martini A., Cinque M., Social networking come supporto alla didattica universitaria: quali i benefici di Ning? Journal of e-Learning and Knowledge Society, 7, 1, 77-87.

[OECD, 2013], OECD, Review of the Italian Strategy for Digital Schools, 2013.

[OECD, 2013a], OECD, Education at a Glance 2013: OECD Indicators, OECD Publishing, 2013.

[Parigi, 2009] Parigi L., Interfaccia a misura di scuola? La Lavagna Interattiva Multimediale nella percezione degli insegnanti. A. Andronico, L. Colazzo (Eds.) DIDAMATICA, 2009.

[Pedrò, 2006] Pedrò, F., The New Millennium Learners. What Do We Know about The Effectiveness of ICT in Education and What We Don't, Espoo, 2006 disponibile all'indirizzo www.oecd.org/edu/ceri/37172511.pdf (ultima consultazione 20/07/2013).

[Pedrò, 2008] Pedrò, F., The New Millennium Learners. A project in Progress, disponibile all'indirizzo <http://www.oecd.org/dataoecd/39/51/40554230.pdf>, pp. 19-20 (ultima consultazione 20/07/2013).

[Prensky, 2001] Prensky, M., Digital Natives, Digital Immigrants. On the Horizon, 9 (5) consultabile all'indirizzo <http://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf> (ultima consultazione 20/07/2013)

[Prensky, 2009] Prensky M., H. Sapiens Digital. From Digital Immigrants and Digital Natives to Digital Wisdom. Innovate. Journal of Online Education, 5 (3), 2009.

[Rapetti e Cantoni, 2010] Rapetti E., Cantoni L., "Nativi digitali" e apprendimento con le ICT. La ricerca "Gen Y @ work" in Ticino, Svizzera. Je-LKS, 6, 1, 2010, 43-53.

[Rosen, 2010] Rosen L.D., Rewired: Understanding The I-generation and the Way The Learn, New York, Palgrave Macmillan, 2010.

[Strauss e Howe, 1992] Strauss, W., Howe N., Generations : The History of America's Future, 1584 to 2069, William Morrow & Co, 1992.

[Tapscott, 1998] Tapscott D., Growing up digital: The rise of the Net Generation, New York, McGraw-Hill, 1998.

e-Learning Platforms: A comparative analysis

G. De Serio, D. Impedovo, G. Pirlo, L. Scianatico
Dipartimento di Informatica, Università degli Studi di Bari "Aldo Moro"
Via Orabona, 4 – 70125 Bari
{sirius_91@live.it; impedovo@gmail.com; giuseppe.pirlo@uniba.it;
scianaticolorenzo@gmail.com}

Abstract. *Platforms for e-learning plays a fundamental role in the modern society, in which education is the key aspect for social and economic growth. This paper presents a comparative analysis among four main open source platforms: ATutor, Claroline, Moodle and Plone. The analysis allows to better explain the characteristics of the diverse platforms and their capability in supporting high quality e-learning activities.*

Keywords: Distance-Learning, e-Learning Processes, e-Learning platforms.

1. Introduction

In modern society, e-learning is a key factor for social and economic development. Many educational institutions and companies offer e-learning solutions not only as a complement to classroom training but also as a learning path aimed at users with difficulty to attend to classroom activities.

Of course, efficiency and effectiveness of e-learning activities is greatly relevant and therefore e-learning evaluation is rightly considered as a key aspect of any e-learning activity, that allows the definition of the most effective tools and methodologies for its adoption with respect to the e-learning products (e-learning courses, etc.) and processes (course design, production and use, etc.) [Willis, 2000, Kirkpatrick, 1994, Rossi et al, 1999].

In order to "measure" quality of an e-learning activity, several types of finalities can be considered, like client satisfaction, real learning, behavior modification, and efficacy of investment [Greco et al, 2005]. Furthermore, several evaluation strategies can be used for the evaluation of educational programs [Greco et al, 2006]:

- *Objectives-oriented* (it determines the extent to which program and instructional objectives have been met);
- *Management-oriented* (it servers to decision-makers in order to make decisions about the reallocation of funds);
- *Consumer-oriented* (it focuses on the development of information on products, that is essential for the appeal of distance education programs);
- *Expertise-oriented* (it mainly depends upon professional expertise to judge an educational program);

- *Adversary-oriented* (that attempts to use both positive and negative views into the evaluation itself);
- *Participant-oriented* (it is used in qualitative research studies to evaluate and match all opinions).

Whatever approach is considered, a fundamental role in the e-learning processes is played by the Learning Management System (LMS), which is the technological platform that manages the distribution and use of training: it is a management system that lets you monitor the frequency of courses and activities user training. All information on courses and users are stored in the database of the platform: this feature allows the user to access their provision effectively from any computer connected to the Internet. The user is hence completely free and by virtue of what its easier access to their training on the server model, in any place and at any time.

While the platform turns out to be a key component for distance learning, it still allows interaction between users. In general, a platform should provide tools that facilitate instant communication via chat, shared whiteboards (interactive whiteboards) and video conferencing and so on. The environment of collaboration software can also manage asynchronous learning (which does not require the presence of users at the same time): discussion forums, document repositories, access to educational materials or support materials.

In the following, some of the most widespread open source platforms are discussed and comparatively analyzed. Some characteristics of each platform are highlights that allows to determine the most profitable application domain. In particular, four platforms are considered: ATutor, Claroline, Moodle and Plone.

2. ATutor

ATutor is an Open Source Web-based Learning Content Management System (LCMS). It is used in various contexts, including online course management, continuing professional development for teachers, career development, and academic research [Rovai, 2004]. The software is cited as unique for its accessibility features, (useful to visually impaired and disabled learners); and for its suitability for educational use according to software evaluation criteria established by The American Society for Training and Development (ASTD). ATutor is used internationally and has been translated into over fifteen languages with support for over forty additional language modules currently under development.

ATutor is the first LCMS to comply completely with the accessibility specifications of W3C WCAG 1.0 at the AA+ level, allowing the access to all the included content of the system at all levels of user-privilege, including administrator accounts. Its conformity with XHTML 1.0 is intended to ensure that ATutor is presented and displayed consistently in any compatible technology.

ATutor's developers assert that it is the only fully accessible LCMS software on the market, including for-fee and closed source software. This aspect of ATutor has been affirmed in at least one University-published review of the software. ATutor is also cited in numerous technical reviews and scholarly articles; and

many third-party extensions have been developed and distributed for use with the software [Chee-Kit, 2005; Williams, 2003; St.Amant, 2007; Qing, 2004]. (Fig.1)



Fig.1 ATutor

Two of many accessibility features in the system are text alternatives for all visual elements, and keyboard access to all elements of the program. With these features, a blind person can listen to the entire interface of the system with the help of a screen reader, and he or she can access the system without needing a mouse. These features also allow ATutor to adapt to a wide variety of technologies including cell phones, personal data assistants (PDAs), and text-based Web browsers, and much more.

ATutor includes a content authoring tool that prompts content developers to create accessible learning materials. Such prompts encourage authors to add a text alternatives if they forget to include one when adding an image, for instance. The authoring tool also includes a Web service that evaluates the accessibility of authored content against various international standards. In addition to creating accessible content, the tool is itself accessible, allowing a blind user to create content themselves.

3. Claroline

Claroline is a collaborative eLearning and eWorking platform (Learning Management System) released under the GPL Open Source license. It allows hundreds of organizations worldwide ranging from universities to schools and from companies to associations to create and administer courses and collaboration spaces over the web. The platform is used in more than 100 countries and is available in 35 languages.

The Claroline platform is organized around the concept of space associated to a course or a pedagogical activity. Each course space provides a list of tools enabling the teacher to: (Fig.2)

- Write a course description;
- Publish documents in any format (text, PDF, HTML, video...);
- Administer public and private forums;
- Develop (also by e-mail);
- Propose assignments to be handed in online;

- See the statistics of the users activity;
- Use the wiki to write collaborative document.



Fig.2 Claroline

Claroline is compatible with GNU/Linux, Mac OS and Microsoft Windows. It is based on open-source technologies like PHP and MySQL.

Claroline is based on pedagogical principles coming from the literature on the advantages of technologies for training. Since 2000, Claroline development teams have been focusing on the code stability and the development of features in line with the user's needs. The main concern of the developers is not to build up a large number of new features, but to concentrate upon a few elaborated tools concerning the pedagogical approach and the interface offered to the users. Besides, a large worldwide community of users and developers contributes to Claroline's development and diffusion extensively [Consortium Claroline, 2006].

4. Moodle

Moodle (acronym for *Modular Object-Oriented Dynamic Learning Environment*) is a free software e-learning software platform, also known as a Learning Management System, or Virtual Learning Environment (VLE). As of June 2013 it had a user base of 83,008 registered and verified sites, serving 70,696,570 users in 7.5+ million courses with 1.2+ million teachers.

Moodle was originally developed by Martin Dougiamas to help educators create online courses with a focus on interaction and collaborative construction of content, and is in continual evolution [Impedovo et al, 2006]. The first version of Moodle was released on 20 August 2002.

Moodle has several features considered typical of an e-learning platform, plus some original innovations (like its filtering system). Moodle is very similar to a learning management system. Moodle can be used in many types of environments such as in education, training and development, and business settings.

Some typical features of Moodle are:

- Assignment submission;
- Discussion forum;

- File downloads;
- Grading;
- Moodle instant messages;
- Online calendar;
- Online news and announcement (College and course level);
- Online quiz;
- Wiki.

Developers can extend Moodle's modular construction by creating plugins for specific new functionality. Moodle's infrastructure supports many types of plugins, like:

- activities (including word and math games);
- resource types;
- question types (multiple choice, true and false, fill in the blank, etc.);
- data field types (for the database activity);
- graphical themes;
- authentication methods (can require username and password accessibility);
- enrollment methods;
- content filters.

Moodle users can use PHP to write and contribute new modules. Moodle's development has been assisted by the work of open source programmers. This has contributed towards its rapid development and rapid bug fixes.

By default Moodle includes the TCPDF library that allows the generation of PDF documents from pages.

Moodle has continued to evolve since 1999 (since 2001 with the current architecture). Major improvements in accessibility and display flexibility were developed in 1.5. The current version can be seen at the top of the table of releases, below. It has been translated into 82 different languages.

The development of Moodle continues as a free software project supported by a team of programmers and an international user community, drawing upon contributions posted to the online Moodle Community website that encourages debate and invites criticism. As of July 11, 2012, Moodle is developing Moodle Mobile on HTML5 and Phonegap. It is planned to be released at the end of 2012 [Dougiamas, 2012].

5. Plone

Plone is a free and open source content management system built on top of the Zope application server. In principle, Plone can be used for any kind of website, including blogs, internet sites, web shops and internal websites. It is also well known as a document publishing system and groupware collaboration tool. The strengths of Plone are its flexible and adaptable workflow, very good security, extensibility, high usability and flexibility.

Plone runs on the Zope application server (Fig.3), which is written in Python. Plone by default stores all information in Zope's built-in transactional object database (ZODB). It comes with installers for Windows, Mac OS X, and Linux, along with other operating systems. New updates are released regularly on Plone's website. Plone is available in over 35 languages. Since version 3.0, it claims conformance to (the now outdated) WCAG 1.0 AA and U.S. section 508, which allows people with disabilities to properly access and use Plone. An update to make Plone conform to WCAG 2.0 has been planned. A major part of Plone is its use of skins and themes. When working with Plone, templates can be used to customize a website's look. These templates are written with Cascading Style Sheets. In addition, Plone comes with a user management system called Pluggable Authentication Service. Introduced in Plone 2.5, "PAS" is used to properly sort actions from different users to their respective folders or accounts. PAS is also used to search for users and groups in Plone. Most importantly, PAS covers the security involved for users, requiring authentication in order to login to Plone. This gives users an increase in both security and organization with their content. A large part of Plone's changes come from its community. Since Plone is open source, the members of the Plone community regularly make alterations or add-ons to Plone's interface, and make these changes available to the rest of the community via Plone's website. The name Plone comes from a band by that name and "Plone should look and feel like the band sounds" [Plone.org, 2012].

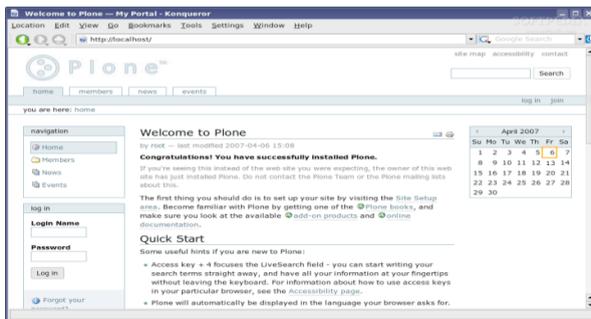


Fig.3 Plone

The community supports and distributes thousands of addons via company websites but mostly through PYPI and www.plone.org. There are currently 2149 packages available via PYPI for customizing Plone.

Since its release, many of Plone's updates and add-ons have come from its community. Events called "Plone sprints" consist of members of the community coming together for a week and helping improve Plone. The Plone conference is also attended and supported by members of the Plone community. In addition, Plone has an active IRC channel to give support to users who have questions or concerns. Up through 2007, there have been over one million downloads of Plone. Plone's development team has also been ranked in the top 2% of the largest open source communities.

6. e-learning Platforms: A comparative analysis

Figure 4 shows the main utilities and tools of the e-learning platforms described above:

	Platforms			
Tools & Utility	ATutor	Claroline	Moodle	Plone
Open Source	x	x	x	x
Course management	x	x	x	x
Academic research	x		x	x
Forums		x	x	
Groups		x	x	
Chat		x	x	
Blog		x	x	
Wiki		x	x	
Live Search system				x
Professional development	x	x	x	x
Functions for disabled	x			x

Fig 4. Platforms's tools and utility.

The four platforms indicated in the table are all open source, absolutely free of charge and available on their websites, which offer the opportunity to follow one direction or distance course courses for the professional development of various fields (sometimes even refresher courses).

The areas of academic research are well considered by systems ATutor, Moodle and Plone, which has meant a significant use of these platforms over the years by research centers and especially universities. Claroline prefers instruments related to forums, chat, groups, blogs and wikis, and most useful tools for the social sphere, a field in which Moodle is also very well developed.

Plone is the only one that presents a tool for system Live Search, making available for use for all users using the search box, and it is the only one, along with ATutor, which has tools and features for disabled users and blind users.

7. Conclusions

In this paper some of the most widespread platforms for e-learning are discussed and analyzed comparatively. The results demonstrate that some platforms (Moodle, Claroline) seems to be complete for supporting effectively and efficiently e-learning processes for a variety of application domain.

In particular, the presence of a complete set of functionalities makes these platforms mature for their application in schools, from minor schools up to universities, so that they can train people from an early age to relate to the world of technology.

References

Chee-Kit, Looi, *Artificial Intelligence in Education Supporting Learning Through Intelligent and Socially Informed Technology*. IOS Press. ISBN 1-58603-530-4, 2005.

Consortium page, Claroline.net, 2007.

Dougiamas, Martin , "Martin in black and white Mobile Moodle app moving to HTML5". Retrieved 13 July 2012.

Greco, N., Dimauro, G., Ferrante, A., Impedovo, S., Lucchese, M.G., Modugno, R., Pirlo, G., Sarcinella, L., "e-learning activities at the University of Bari: the PROTEO Project", Proc. of the WSEAS Int. Conf. on "Applied Informatics and Communications (AIC'05), Malta, pp. 143-147, 2005.

Greco, N., Dimauro, G., Ferrante, A., Impedovo, S., Lucchese, M.G., Modugno, R., Pirlo, G., Sarcinella, L., "The PROTEO Project: New Advancements in e-learning activities at the University of Bari", WSEAS Transactions on Communications, Issue 1, Vol. 5, pp. 23-30, 2006.

Impedovo, S., Modugno, R., Pirlo, G., "A Feedback-based Framework for Continuous Evaluation of e-Learning Activities", Proceedings of the 3rd WSEAS/IASME International Conference on Engineering Education (EE '06), Vouliagmeni, Athens, Greece, July 11-13, pp. 258-262, 2006

Kirkpatrick, D., "Evaluating Training Programs. The four levels", Berrett-Koehler, San Francisco, USA, 1994.

Plone.org, "Plone Sites — Plone CMS: Open Source Content Management". Retrieve, 2012.

Qing, Li, *New Horizon In Web-based Learning*. World Scientific. ISBN 981-256-029-7, 2004.

Rossi, P. H., Freeman, H. E., Limpsey, M.W., "Evaluation: a systematic approach", Newbury Park, CA: Sage, 1999.

Rovai, A.P., "A constructivist approach to online college learning", *Internet and Higher Education*, Vol. 7, Issue 2, pp. 79-93, 2004.

St.Amant, Kirk, *Handbook of Research on Open Source Software: Technological, Economic, and Social Perspectives*. Idea Group Inc (IGI). ISBN 1-59140-999-3, 2007.

Williams, Roy, 2nd European Conference on E-Learning Glasgow Caledonian University, Glasgow, 6–7 November 2003. Academic Conferences Limited. ISBN 0-9544577-4-9, 2003.

Willis, J., "The maturing of constructivist instructional design: Some basic principles that can guide practice", *Educational Technology*, Vol. 40, Iss. 1, pp.5-16, 2000.

Progettare e realizzare l'e-learning nella Pubblica Amministrazione: il caso della Provincia autonoma di Trento

Nicola Villa
tsm-Trentino School of Management
via Giusti, 40, 38122 Trento
nicola.villa@tsm.tn.it

The aim of this paper is to describe the process that led to the Autonomous Province of Trento to adopt e-learning initiatives for their employees. In particular, after the first experiments in 2009/2010 on small groups of users, since 2011 the Distance Learning has become a method formally used as the "traditional" classroom. In this paper we describe the entire process, both highlighting quantitative and qualitative variables and the e-learning platform used.

Keywords: e-learning, Public Administration, learning objects.

1. Introduzione

La Pubblica Amministrazione italiana è da anni impegnata in un processo di modernizzazione, atto da un lato a semplificare i processi organizzativi e decisionali e dall'altro a contenere i costi per una gestione più efficace e efficiente delle proprie risorse interne.

Anche la Provincia autonoma di Trento (da questo momento indicata come P.A.T.) ha adottato una serie di provvedimenti in tal senso. Da citare ad esempio il processo di riorganizzazione della struttura dei Dipartimenti e Servizi, il processo di *spending review* [Hinna & Marcantoni, 2012], l'utilizzo delle nuove tecnologie per la semplificazione e miglioramento dei processi organizzativi, anche in rapporto con il cittadino.

Obiettivo di questo contributo è però quello di focalizzare l'attenzione su un particolare ambito di innovazione: la formazione del personale interno all'amministrazione. Per la P.A.T. la formazione costituisce una leva strategica fondamentale per l'accrescimento della professionalità e delle competenze del proprio personale, che deve essere in grado di fare fronte alle continue modifiche e aggiornamenti del metodo lavorativo e organizzativo.

Da molti anni la P.A.T. ha investito nel miglioramento continuo delle proprie risorse, gestendo il processo formativo in una prima fase autonomamente, per poi successivamente affidare la formazione a tsm-Trentino School of Management.

Congresso Nazionale AICA 2013

Da qualche anno, oltre alle iniziative formative d'aula di tipo "tradizionale", l'indirizzo dell'Amministrazione è stato quello di avviare un cambiamento delle modalità di svolgimento dei corsi tramite l'inserimento dell'e-learning come metodo formativo. Dapprima mediante una sperimentazione, biennale (2007-2008) [Casagrande & Tomasini, 2008], che ha coinvolto uno stretto numero di utenti; in questa fase si è valutato l'impatto delle nuove metodologie formative in termini di efficienza ed efficacia [Casagrande *et al.*, 2009]. A fronte dei buoni risultati dei primi due anni, l'Amministrazione ha deciso di scegliere la Formazione a Distanza come nuovo "modo di lavorare" e di formarsi dei propri dipendenti.

In questo lavoro si descriverà questo processo, suddividendo le prossime sezioni in questo modo: nella sezione 2 si presenterà il ruolo di tsm-Trentino School of Management all'interno del sistema formativo provinciale. La sezione 3 descriverà l'e-learning nella P.A.T., concentrando l'attenzione in primo luogo sulle tipologie di processi formativi e le modalità di progettazione/erogazione, riepilogando alcune informazioni relative alla piattaforma informatica utilizzata per le attività in FAD. Nella sezione 4 si presenteranno i risultati ottenuti in questi anni proponendo una riflessione sugli scenari futuri.

2. Il ruolo di tsm-Trentino School of Management

La Provincia autonoma di Trento, per la realizzazione di programmi di formazione rivolti ai propri dipendenti e a quelli dei suoi enti strumentali, si avvale di tsm-Trentino School of Management.

tsm-Trentino School of Management è una Scuola costituita dalla Provincia autonoma di Trento, dalla Camera di Commercio Industria Artigianato e Agricoltura di Trento e dall'Università degli Studi di Trento¹. Il punto focale nelle strategie di tsm è l'attenzione alla qualità della didattica; l'innovazione e la sperimentazione nei contenuti e nei metodi rappresentano un impegno prioritario della Scuola. Ogni attività si avvale di una rete di interlocutori in grado di rappresentare le prospettive più avanzate della ricerca sulla formazione a livello nazionale e internazionale. A tal fine, le azioni formative previste propongono:

- un metodo formativo "attivo", con grande attenzione all'individuo quale soggetto della formazione;
- l'attenzione al miglioramento continuo dei risultati anche tramite metodologie di *customer satisfaction*;
- la valorizzazione del rapporto tra docenti, referenti di corso e utenti;
- la ricerca come base di conoscenza per ogni azione formativa;
- la formazione in rete.

Annualmente tsm pubblica con la collaborazione del Servizio del Personale della P.A.T. un Piano contenente tutte le attività formative proposte ai dipendenti; in particolare il Programma Annuale di Formazione 2013 si inserisce all'interno del Piano Triennale di Formazione 2013-2015, dividendosi in due sezioni:

¹ www.tsm.tn.it

- Formazione manageriale
- Formazione per il personale

Con **Formazione per il personale** si intendono tutte le azioni che hanno come destinatari i dipendenti della P.A.T. (inclusi anche i dirigenti e i direttori), concentrandosi su quattro aree specifiche:

- *Giuridico – economica*
- *Competenze trasversali*
- *Sicurezza*
- *Informatica*

Per ciascuna di queste aree sono proposte una serie di iniziative formative, così come sarà descritto nella sezione 3.2.

3. e-Learning nella Provincia autonoma di Trento

La P.A.T. nel 2009 ha iniziato un percorso di sperimentazione per l'adozione dell'e-Learning come metodologia formativa per i propri dipendenti. Il primo progetto, chiamato ESPERTO (E-learning per lo Sviluppo dell'E-procurement nella Provincia autonoma di Trento) prevedeva l'abbinamento di formazione d'aula e a distanza riguardante la centrale di e-procurement a servizio di tutta la Pubblica Amministrazione trentina².

Conclusa positivamente la prima fase, si è passati tra 2009 e il 2010 a sperimentare questa metodologia in nuovi percorsi formativi all'interno delle aree formative Sicurezza e Giuridico-Economica del Piano Annuale di Formazione.

Nel 2011 le attività sperimentali si sono concluse e, grazie ai buoni risultati ottenuti [*ibidem*], il progetto si è ampliato su tutte le aree di formazione proposte all'interno del piano formativo del sistema provinciale (Provincia ed Enti a essa collegati). Nel 2011 circa 500 persone hanno concluso corsi di formazione in modalità *full online* e *blended*, per un totale complessivo di 254 ore di formazione erogata.

L'offerta attuale si basa prevalentemente su proposte in autoformazione. Questi moduli hanno obiettivi comuni con quelli che si svolgono tradizionalmente in aula, così come lo stesso valore in termini di riconoscimento della frequenza (è possibile infatti valutare in modo opportuno la frequenza dei partecipanti alle singole lezioni, così come all'intero corso). All'interno di ciascun percorso sono messi a disposizione oggetti didattici contenenti audio o video lezioni, con momenti di autovalutazione in itinere o a conclusione delle attività svolte.

Per ciascun corso il partecipante ha la possibilità di utilizzare una piattaforma di formazione a distanza (chiamata *L3 – LifeLong Learning*, vd. Sezione 3.3) progettata specificatamente per rispondere alle esigenze dell'Amministrazione provinciale; inoltre è sempre disponibile il supporto di un e-tutor e, laddove necessario, del docente portatore delle conoscenze.

² www.esperto.provincia.tn.it

L'adozione dell'e-learning come metodologia formativa "a regime", ha consentito di raggiungere notevoli vantaggi, che possono essere così riassunti, come descritto in [Visintainer, 2012]:

- soluzioni di apprendimento flessibili, cioè personalizzabili e facilmente fruibili dal dipendente utilizzando la propria postazione di lavoro;
- autonomia nella gestione del tempo da dedicare alla formazione a seconda delle esigenze professionali e personali, sia a livello di fruizione del prodotto didattico, sia a livello di scelta dei singoli argomenti da approfondire;
- possibilità di accedere agli oggetti di apprendimento multimediali e al materiale didattico, anche dopo la chiusura del corso, per ulteriori approfondimenti o chiarimenti legati ad esigenze professionali;
- applicazione di economie di scala, grazie alla riusabilità del materiale didattico;
- economia di tempi e spostamenti, dato che il lavoratore non è costretto a recarsi nelle apposite aule per seguire i corsi;
- creazione di una cultura generalizzata della formazione continua, trasformando il luogo di lavoro in un luogo di apprendimento.

La partecipazione ai corsi da parte dei dipendenti provinciali è regolamentata direttamente dal Servizio per il Personale; in particolare è previsto che il dipendente abbia la possibilità di svolgere la formazione in orario di lavoro, direttamente dalla propria postazione informatica predisponendo un apposito giustificativo.

3.1 Modalità di progettazione/erogazione

Negli ultimi due anni tsm, per conto della P.A.T., ha adottato un metodo organizzativo e didattico per l'erogazione di corsi in e-learning. In particolare la costruzione del piano formativo segue il percorso "classico", prevedendo le fasi di *analisi dei requisiti*, *progettazione*, *realizzazione* dell'intervento formativo e *valutazione*. In ciascuna delle fasi le tecnologie informatiche svolgono un ruolo fondamentale.

L'**analisi dei requisiti** viene svolta in più momenti, dall'analisi del fabbisogno formativo (costruita da una combinazione di interviste e questionari da compilare online), all'analisi del fabbisogno di ogni singolo corso (per la quale vengono inviati ai partecipanti questionari per sondare le aspettative e le richieste da approfondire durante il percorso formativo).

La **progettazione** degli interventi in e-learning viene condotta tenendo conto delle specificità della metodologia adottata. I corsi in aula non vengono semplicemente "trasferiti" in FAD, bensì la progettazione parte proprio dal presupposto che un corso fruito a distanza debba essere pensato e costruito in modo specifico. In questa fase riveste un ruolo importante la figura dell'*instructional designer*, il quale con la collaborazione dell'esperto dei contenuti è in grado di strutturare il percorso didattico.

In particolare i corsi in e-learning possono prevedere diverse modalità di erogazione, a seconda del tipo di azione formativa:

- *Percorsi formativi di tipo "formale"*, appositamente pensati per soddisfare le esigenze del personale coinvolto.
- *Percorsi formativi di tipo "permanente"*, all'interno dei quali la formazione è a disposizione dei partecipanti per un periodo molto lungo e gli oggetti formativi rimangono a disposizione del dipendente anche dopo la conclusione del corso, come accompagnamento delle proprie attività lavorative.
- *Percorsi di accompagnamento a attività formative svolte in aula*. In questo caso durante i percorsi tradizionali vengono raccolte dai partecipanti le esigenze di approfondimento, rielaborate con il docente e messe a disposizione online al termine del corso come moduli di *follow up*.
- *Percorsi di aggiornamento*. In questo ultimo caso è possibile rispondere in modo immediato alle esigenze formative di un ampio *range* di partecipanti, offrendo online oggetti di breve durata e riguardanti magari piccoli ma essenziali aggiornamenti (es. della normativa, di funzionalità di un applicativo), evitando di riconvocare in aula grandi quantità di partecipanti.

La **realizzazione** del corso, ovvero la predisposizione degli oggetti didattici e dello spazio formativo da fruire online, viene svolta quasi interamente all'interno della struttura, con la collaborazione di fornitori esterni quando è richiesta una qualità del materiale molto alta (es. corsi sulla comunicazione).

Ciascun intervento formativo a distanza può prevedere l'adozione di diverse tipologie di oggetti didattici, a seconda degli obiettivi e del target. tsm ha sviluppato, ad esempio

- *audio/videolezioni*: il partecipante può seguire l'intervento del docente (in audio e/o video) similmente a come avviene in aula.
- *videotutorial*: sviluppati principalmente per spiegare l'utilizzo di applicativi informatici (es. SAP, pacchetto OpenOffice e Office).
- *casi di studio interattivi*: il partecipante può interagire con l'oggetto didattico per meglio comprendere gli argomenti trattati.
- momenti di autovalutazione.

Una scelta precisa è stata quella di non basare le lezioni online soltanto su dispense o, in generale, su oggetti didattici testuali (e quindi senza audio) perché considerati di difficile fruizione da parte dei partecipanti. In questi primi anni nei quali questa metodologia formativa si è sempre più incrementata, l'obiettivo è stato quello di migliorare sempre di più la qualità delle lezioni, sia dal punto di vista visivo che uditivo, scegliendo di volta in volta il miglior approccio per consentire un efficace passaggio delle conoscenze (non sempre, infatti, può essere necessario vedere il docente in video, così come, in altri casi, l'espressione e gestualità del docente è essenziale per comprendere alcuni aspetti comunicativi che altrimenti andrebbero persi).

Gli interventi formativi sono progettati a stretto contatto dei docenti, così da costituire un "**team della conoscenza**" in grado di realizzare il corso; entrano in gioco infatti l'*instructional designer*, il referente del corso, il docente e l'e-tutor.

La **valutazione** nei corsi in e-learning viene svolta online, sia per quanto riguarda il gradimento del percorso svolto dal partecipante che per l'autovalutazione delle conoscenze acquisite.

3.2 Tipologie di percorsi formativi

Le attività in Formazione a Distanza all'interno di tsm, come specificato precedentemente, sono iniziate nel 2010 (coinvolgendo inizialmente un piccolo gruppo di partecipanti); nel grafico che segue è riportato l'andamento delle partecipazioni a corsi in e-learning nel 2011 e 2012, con una proiezione nel 2013.

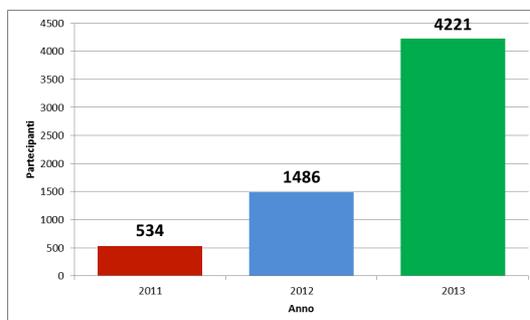


Fig 1: partecipazioni a corsi in e-learning di tsm nel 2011, 2012 e 2013

Come evidenziato, sono attualmente iscritti a percorsi in FAD più di 4000 dipendenti (a luglio 2013 sono già state formate più di 1500 persone), suddivisi in quattro aree formative.

L'area **Giuridico-Economica** presenta percorsi nell'ambito dei lavori pubblici e contratti e dell'atto amministrativo. Si proseguono le attività nell'ambito della privacy e del codice dell'amministrazione digitale iniziate tra il 2011 e 2012.

Per il primo anno sono avviate iniziative formative nell'ambito delle **Competenze trasversali**, con moduli in formazione a distanza che possono essere utilizzati sia come preparazione e accompagnamento agli interventi d'aula sia come percorsi autoportanti.

L'area **Sicurezza**, da sempre centrale all'interno della P.A.T., si concentra sulla formazione generale dei lavoratori e del personale tecnico di cantiere.

Infine tutti i corsi all'interno dell'area **Informatica** sono svolti in e-learning, prevedendo anche l'utilizzo di oggetti didattici provenienti dalla convenzione di riuso con la Regione Toscana³. In quest'area l'approccio di "*formazione permanente*" è applicata in modo totalizzante, in quanto i vari percorsi sono a disposizione dei partecipanti per sempre.

³ www.progettotrio.it

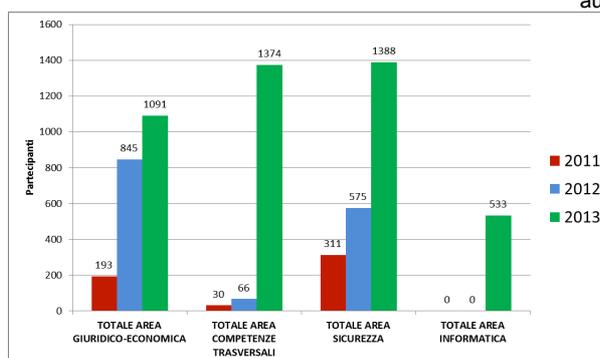


Fig 2: suddivisione in aree dei partecipanti a corsi in e-learning di tsm nel 2011, 2012 e 2013

I corsi attualmente pianificati all'interno di ciascuna area sono:

<p>AREA GIURIDICO ECONOMICA</p> <p>Il Codice dell'Amministrazione Digitale (CAD) La redazione dell'atto amministrativo – modulo base Il trattamento dei dati personali da parte dei soggetti pubblici – moduli base e avanzato informatico La tracciabilità Flussi Finanziari SAP - Bilancio impegni anagrafica SAP - Documento di spesa e liquidazioni SAP – Gestione richieste d'acquisto e gestione cespiti</p>	<p>AREA COMPETENZE TRASVERSALI</p> <p>I fondamenti della comunicazione La comunicazione assertiva Curare la comunicazione nelle relazioni lavorative Formazione d'ingresso Scrivere e-mail efficaci La gestione della comunicazione telefonica</p>
<p>AREA SICUREZZA</p> <p>Formazione dei lavoratori - modulo generale Formazione dei lavoratori – agg. rischio basso Sicurezza cantieri - modulo 40 ore e agg. 5 ore Accesso in sicurezza ai cantieri</p>	<p>AREA INFORMATICA</p> <p>Openoffice.org Writer, Calc, Impress Microsoft Office Word, Excel, PowerPoint Windows 7 Lotus Notes eWorks</p>

Le attività in e-learning all'interno del piano 2013 rappresentano circa un terzo delle attività formative svolte da tsm nell'Area Pubblica Amministrazione.

3.3 L3 – LifeLong Learning: la piattaforma FAD utilizzata dalla P.A.T.

L3-LifeLong Learning [Colazzo *et al.*, 2009] è il *Learning Management System* (LMS) basato sulla metafora delle comunità virtuali di apprendimento [Villa *et al.*, 2007] adottato dalla P.A.T. per l'erogazione dei corsi in FAD; la piattaforma è stata realizzata dall'Università degli Studi di Trento sulla base di precedenti esperienze di successo all'interno del contesto accademico [Colazzo & Molinari, 2007].

Attraverso L3 è possibile erogare e gestire i processi di apprendimento che vengono svolti a distanza.

Possono accedere alla piattaforma tutti i partecipanti iscritti ai corsi in formazione a distanza (provenienti dalla P.A.T. o da altri Enti collegati). Attualmente sono iscritti circa 9000 utenti con una media di accessi mensile in costante aumento (legata anche all'accrescere delle iniziative formative), così come presentato dal grafico che segue.

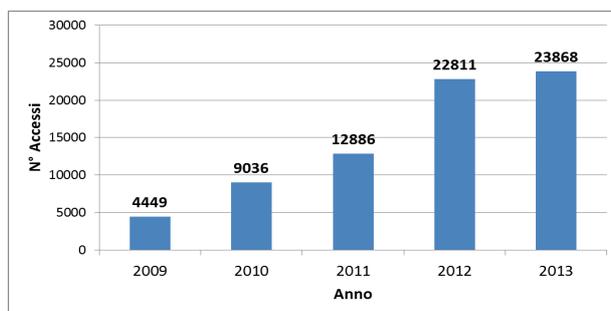


Fig 3: accessi (tramite identificazione) a L3 (dati aggiornati al 30 agosto 2013)

La P.A.T. ha scelto di utilizzare una piattaforma di e-learning non solo per consentire il caricamento/fruizione degli oggetti didattici, ma all'interno di un contesto più ampio nel quale tutti gli attori che partecipano al processo formativo (docente, discente, eTutor, progettista didattico, committente) possano dialogare costantemente. È possibile quindi non solo costruire degli spazi di formazione "formale", ma anche spazi di collaborazione all'esterno del contesto formativo (come, ad esempio, la gestione di gruppi di lavoro per l'esecuzione di progetti).

Come descritto in precedenza sono anche presenti spazi di formazione permanente, che quindi non si esauriscono allo scadere del corso ma, al contrario, continuano nel tempo; in questo caso riveste un ruolo ancora più fondamentale il docente che si trasforma da detentore di conoscenza ad aggregatore e cogeneratore di conoscenza insieme ai partecipanti; si cerca di coinvolgere il docente a supportare i partecipanti rispondendo alle loro domande, raccogliendo le loro esperienze e costruendo materiali di aggiornamento. Esempio ancora diverso rivestono le comunità relative ai gruppi di lavoro, focalizzate alla condivisione di informazioni ed esperienze slegate completamente dalle logiche formative.

Per soddisfare queste esigenze all'interno di L3 sono presenti *servizi di didattica "tradizionale"*, come ad esempio la possibilità di utilizzare materiale didattico (di ogni tipologia), glossari, *servizi di didattica "virtuale"* (fruizione di audio/video lezioni rispondenti allo standard SCORM), *servizi di "comunicazione"* tra il docente/tutor del corso e i partecipanti (sincroni come la *chat*, asincroni come *forum*, bacheca, FAQ) e *servizi "collaborativi"* tra i partecipanti, o tra il partecipante e il docente (*wiki*, *blog*, quaderno delle attività).

4. Risultati e scenari futuri

tsm prevede il monitoraggio costante delle iniziative formative, in particolare rilevando il gradimento dei partecipanti riguardo i contenuti trattati, l'organizzazione dei corsi e le metodologie adottate. Tale rilevamento è certificato attraverso il Sistema di Gestione Qualità interno, conforme ai requisiti della norma UNI EN ISO 9001:2008, in particolare per il campo applicativo "*Progettazione, erogazione e valutazione di servizi di alta*

*formazione e formazione continua per il comparto pubblico e privato. Attività di ricerca finalizzata all'innovazione nella formazione e nella valutazione*⁴.

Nel 2012 sono stati erogati dieci corsi in e-learning (full online o blended), per un totale di 38 edizioni che hanno formato circa 1500 partecipanti.

Si riportano di seguito le medie ponderate delle principali variabili riguardanti il livello di soddisfazione dei corsi erogati (dove 1 è il valore minimo e 5 il valore massimo).

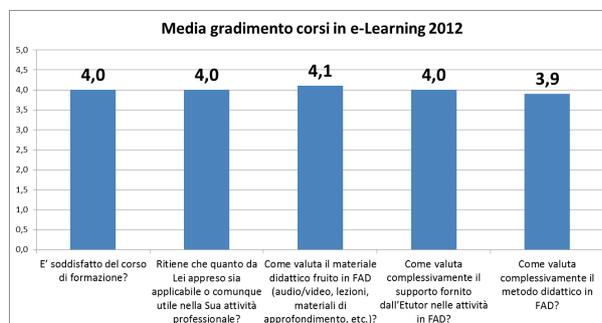


Fig 4: gradimento medio corsi in e-learning di tsm - 2012

Il gradimento medio dei corsi in e-learning è in linea con quello dei corsi svolti in modalità tradizionale in aula; le iniziative erogate nel 2013 fino a questo momento confermano sostanzialmente i risultati qui presentati.

Nel 2012 sono state implementate una serie di azioni migliorative per la progettazione e realizzazione dei materiali didattici rispetto ai punti di attenzione messi in luce all'interno della valutazione raccolta. In particolare per l'attività di progettazione didattica è stata posta attenzione a:

- Analisi e definizione dei bisogni formativi
- Profilazione dell'utenza
- Analisi e vincoli del progetto
- Strutturazione degli obiettivi formativi
- Definizione dei prerequisiti alla partecipazione
- Modalità di valutazione
- Valutazione delle strategie e metodologie didattiche

mentre per l'attività di realizzazione dei materiali didattici in e-learning si è lavorato per migliorare la produzione dei Learning Object, i costi di produzione e le competenze di produzione Learning Object. Per ciascuno dei punti sono state implementate azioni di miglioramento, che sostanziano quanto descritto in questo lavoro.

Attualmente è in fase di predisposizione il Piano Annuale di Formazione 2014, il quale prevede un sempre maggiore coinvolgimento dell'e-learning nelle attività formative di tsm per la Pubblica Amministrazione Trentina, riprendendo gran parte delle iniziative proposte nel 2013 e integrandole con altri corsi, in particolare nell'area Giuridico-Economica e delle Competenze Trasversali.

⁴ http://www.tsm.tn.it/interne/sistema_di_gestione_qualita/C3%A0_interna2.ashx?ID=18538

Conclusioni

In questo lavoro si è descritto il processo che ha portato la Pubblica Amministrazione trentina ad adottare un approccio metodologico per la progettazione, erogazione e fruizione di corsi in e-learning rivolti al proprio personale. Si sono evidenziati i risultati sia di tipo quantitativo, rappresentati da un incremento significativo di partecipazione nell'arco del triennio 2011-2013, sia di tipo qualitativo con un gradimento dei corsi a distanza in linea con quelli in presenza. Le attività si svilupperanno sempre più nel prossimo futuro concentrandosi non solo su corsi di tipo "formale" a distanza ma avviando una serie di progetti di tipo "informale" e rivolti ad un pubblico più ampio (non solo per i lavoratori); ne sono un esempio le attività avviate in ambito SAP e Gare telematiche e il percorso informativo/formativo rivolto ai Presidenti di seggio delle prossime elezioni provinciali.

Ringraziamenti

In conclusione un ringraziamento in primo luogo al gruppo che attualmente lavora al progetto FAD in tsm (Alessia Moser e Sara Tomasini), al Dipartimento Organizzazione e Gestione del Personale P.A.T., all'Università degli Studi di Trento e ad Informatica Trentina S.p.a. che hanno partecipato a vario titolo alla realizzazione del processo di adozione dell'e-Learning in P.A.T.

Bibliografia

Casagrande, M., Tomasini, S. (2008) *ESPERTO: sperimentazione di un modello blended nella Pubblica Amministrazione*. In Atti del V Congresso Annuale della Società Italiana di e-Learning, Trento, 8-11 ottobre 2008.

Casagrande M., Molinari A., Pegoretti V., Tomasini S. (2009), *Progettazione e intervento nell'e-learning per la Pubblica Amministrazione: dalla sperimentazione all'analisi dei costi*. In Sle-L 2009, Salerno: Società Italiana di e-Learning.

Colazzo L., Molinari A., Villa N. (2009), *Lifelong Learning and Virtual Communities in the Public Administration: a case study*. in International Conference on E-learning in the Workplace (ICELW) Columbia University.

Colazzo L., Molinari A. (2007), *Dai tradizionali LMS ai sistemi di Comunità virtuali*. Journal of e-Learning and Knowledge Society, v.3, n.2 (2007).

Hinna L., Marcantoni M. (2012), *Spending Review*. Donzelli Editore, Roma. ISBN: 9788860367556.

Villa N., Colazzo L., Conte F., Molinari A. (2007), *Real communities vs. Virtual communities: structural adaptation of a Learning Management System*. In Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007, Chesapeake (VA): AACE.

Visintainer S. (2012), *L'approccio metodologico di tsm nell'ambito dell'e-learning*. Economia trentina, Anno LXI, n. 1/4-2012. ISSN: 0012-9879.

Per un Museo Nazionale dell'Informatica

Giovanni A. Cignoni*, Fabio Gadducci*, Giuseppe Lettieri**, Carlo Montangero*

**Dipartimento di Informatica, Università di Pisa*

Largo B. Pontecorvo, 3 – 56127 Pisa

{cignoni, gadducci, montangero}@di.unipi.it

***Dipartimento di Ingegneria dell'Informazione, Università di Pisa*

Via G. Caruso, 16 – 56122 Pisa

g.lettieri@iet.unipi.it

Abstract. *An essential factor for the development of a smart society, in which information technologies are deployed in order to improve the quality of our lives, is a clear and common understanding of what computer science is. First, we discuss the importance of the history of computing in improving such understanding, with a focus on the role played by museums. Then we illustrate the activities and collections of the Museo degli Strumenti per il Calcolo of the University of Pisa, supporting its potential as a vehicle for a national institution devoted to computing.*

Keywords: computing history, museums, digital society.

1. Introduzione

L'aggettivo "smart" caratterizza, con una certa enfasi [Holland, 2008], una direzione nello sviluppo della società. Dalle *smart city* in avanti, le tecnologie dell'informazione (TI) sono viste come opportunità di prosperità.

Un prerequisito di tale programma è la conoscenza diffusa delle TI, o meglio, dell'*informatica*, termine che include i risultati tecnologici e le basi scientifiche. Tale conoscenza è il punto di partenza per superare il *digital divide*: l'informatica non deve essere competenza di pochi, ma parte della cultura di tutti.

La costruzione di una conoscenza diffusa dell'informatica passa anche attraverso la divulgazione della sua storia: come scienza è ancora anagraficamente giovane, ma ha già un passato molto ricco. AICA ha promosso i primi corsi di Storia dell'Informatica in Italia, con un'attenzione dimostrata anche dalle sessioni organizzate nelle recenti edizioni del Convegno Nazionale e di Didamatica.

La storia dell'informatica non è solo un esercizio culturale: può servire ad avvicinare gli studenti alle discipline scientifiche e tecnologiche. Nello sviluppo delle comunità di domani, la storia della disciplina serve tanto sul fronte culturale quanto su quello scientifico e tecnologico. L'hardware e il software del passato raccontano le sfide, gli investimenti, i risultati della ricerca e le sue applicazioni industriali, ma anche il modo di comunicare le notizie, raccontare le storie, passare il tempo libero, coltivare i rapporti personali.

Ma l'informatica d'epoca, con il suo fascino e le sue storie, può suscitare curiosità legate al funzionamento di macchine e programmi: esempi accattivanti, semplici e a "grandezza naturale", di principi e meccanismi ancora in uso nei calcolatori di oggi. Spiegano le basi delle scienze dell'informazione e stimolano a studiare informatica ed elettronica; rendono tutti utenti e consumatori consapevoli delle tecnologie, attori partecipi della società digitale.

Con questa prospettiva, l'articolo presenta il *Museo degli Strumenti per il Calcolo* dell'Università di Pisa [MSCWeb, 2013] candidandolo come Museo Nazionale dell'Informatica. La possibilità di tale ruolo deriva sia dal valore dei materiali conservati che dal successo delle esperienze divulgative, oltre che dal recente confronto con altre realtà simili. L'importanza e la responsabilità della proposta ne richiedono la condivisione con la comunità italiana degli informatici.

2. Perché un museo dell'informatica (e come)

Uno dei punti di svolta nella concezione della divulgazione della scienza come fattore di sviluppo delle società moderne è stato il cosiddetto *rapporto Bodmer* [PUS, 1985]. Il documento delinea le strategie per un sistema divulgativo capace di costruire e mantenere nel pubblico una solida cultura scientifica di base, ovvero quel *public understanding of science* che dà il titolo al rapporto. A tale sistema partecipano più attori:

- la scuola, ai suoi diversi livelli, che fornisce l'istruzione di base avvalendosi di una struttura e di un ruolo sociale consolidati;
- la comunità scientifica, che deve raccontare le sue sfide e le sue scoperte suscitando partecipazione e giustificando gli investimenti pubblici in ricerca;
- l'industria tecnologica, che deve chiedere competenze e non usare l'innovazione solo come slogan per la promozione dei propri prodotti;
- i mass-media, che, forti della loro posizione, devono costruire interesse e curiosità per la divulgazione scientifica.

Un ruolo speciale in questo sistema lo hanno i musei. Non più luoghi di conservazione di cimeli per un pubblico di esperti, ma attori di una divulgazione che sfrutta il fascino, la rarità, la capacità di attrazione delle collezioni per incuriosire e interessare, per presentare scienza e tecnologia al pubblico, per spiegarle e approfondirle. A quasi trent'anni dal rapporto Bodmer il ruolo dei musei come luoghi di divulgazione della scienza è ormai consolidato.

Ai musei dell'informatica è stata dedicata la conferenza ospitata dallo *Science Museum* di Londra e organizzata dallo *IFIP WG 9.7* [IFIPWG9.7, 2013], dalla *Computer Conservation Society* [CCS, 2013] e da Google; vi hanno partecipato i più importanti Musei del settore. Gli atti usciranno a breve e gli interventi sono già sulle pagine web della conferenza [MCHR, 2013]. Nel seguito proponiamo un breve resoconto centrato sui i due principali temi affrontati durante i lavori: i criteri da seguire per la divulgazione dell'informatica come scienza e come tecnologia e le strategie da usare per attirare e coinvolgere il pubblico.

2.1 Criteri per la divulgazione per l'informatica

L'informatica è un caso particolare fra le scienze: è una disciplina con un forte connotato tecnologico e soprattutto con applicazioni che spaziano dalle infrastrutture nazionali all'intrattenimento personale. Tuttavia, alcuni criteri di base della divulgazione moderna sono comuni con gli altri settori. In particolare:

- si deve evitare una presentazione deterministica della storia; è facile presentare il progresso scientifico come una naturale catena di successi letta con toni epici e retorici; si enfatizzano i “primati”, trascurando che la ricerca è un percorso al quale contribuisce anche chi sbaglia o arriva dopo;
- conviene approfondire gli aspetti umani della ricerca, perché coinvolgenti, ma non si deve insistere troppo sui personaggi, riconducendo gli scienziati moderni al cliché del genio rinascimentale; raramente i successi sono individuali, da un paio di secoli la ricerca e i risultati tecnologici sono quasi sempre il risultato di progetti collettivi o processi di raffinamento di idee;
- è necessario mantenere un atteggiamento obiettivo; la ricostruzione deve basarsi su più fonti e controlli incrociati; i ricordi personali hanno finora svolto un ruolo predominante, ma a volte sono di parte, influenzati da opinioni, fatti contingenti o riletture successive; inoltre, in un passato recente, molti successi tecnologici sono stati proposti e vissuti con enfasi nazionalistica.

Questi criteri, alla base di una rigorosa ricostruzione storica, assumono particolare valore per la comprensione della scienza da parte del grande pubblico. Veicolano l'idea moderna di una scienza che non segue vie predestinate, ma è fatta di lavoro e di ricerca, di collaborazioni internazionali; che ha bisogno di risorse e impegno costanti e non della fortuita apparizione di geni risolutori.

Oltre a questi criteri generali, valgono per l'informatica considerazioni particolari, legate al suo rapido evolversi, alla doppia natura hardware e software, alle ricadute commerciali. È infatti opportuno:

- evitare confronti anacronistici; il rischio è proporre come rilevante la differenza fra le prime macchine, grandi e difficili da usare, e i dispositivi di oggi, piccoli e *user-friendly*; occorre invece sottolineare l'importanza della apparizione di una tecnologia più della sua successiva evoluzione;
- presentare il software come metà (almeno) dell'informatica; il software, dalla macchina universale di Turing in poi, è ciò che rende l'informatica un *unicum* come tecnologia che trova applicazioni in una moltitudine di campi diversi;
- distinguere tra le applicazioni dell'informatica che sono strumenti di lavoro e quelle finalizzate all'intrattenimento; in entrambi i casi il software è il complemento dell'hardware, ma l'informatica di intrattenimento, così come ad esempio il cinema, cade anche all'interno di categorie artistiche;
- separare i risultati scientifico-tecnologici da quelli commerciali; all'informatica di consumo si deve gran parte dell'interesse per il settore, ma il successo di un prodotto dimostra più l'efficacia del *marketing* che un traguardo scientifico; occorre evitare le pubblicità gratuite e le concessioni ai *fanboy*;
- trovare letture originali della storia dell'informatica; dal cifrario di Cesare alla macchina di Turing, la storia dell'informatica e del calcolo si presta a essere raccontata in più modi; ogni museo deve trovare il suo, valorizzando le proprie collezioni, senza stravolgimenti e dichiarando le eventuali lacune.

I Musei non sono le sole istituzioni destinate alla preservazione dei cimeli informatici. Per esempio il *bit-saving*, il salvataggio dei dati digitali, è una priorità che dovrebbe coinvolgere altre strutture tradizionalmente preposte alla conservazione dei contenuti: biblioteche e archivi. I supporti originali vanno conservati come cimeli, i bit travasati al più presto in formati digitali perpetuabili.

2.2 Misurarsi con il pubblico

Un tipico indicatore dell'efficacia di un museo è l'affluenza di pubblico alle sue sale e alle sue attività (visite guidate, laboratori, eventi). Spesso, tale successo ha conseguenze economiche: dirette, come la vendita di biglietti, o indirette, essendo l'affluenza un indicatore utile quando si tratta di partecipare a bandi, accedere a finanziamenti pubblici, cercare sponsor.

Diventa quindi inevitabile trovare un compromesso fra missione divulgativa, che richiede di "imporre" percorsi scientificamente ineccepibili, e ricerca del successo, che invita ad assecondare le inclinazioni del pubblico, per esempio allineandosi alle mode che caratterizzano lo sviluppo dell'informatica di consumo.

Esistono però accorgimenti che aiutano a costruire esposizioni attraenti:

- “grandi e strani” piacciono; i vecchi calcolatori, ingombranti e riconoscibili nel loro aspetto retrò, sono *museum-friendly*, mentre prodotti più recenti, piccoli e chiusi, rischiano di essere mostrati solo come oggetti di design, sviando l'attenzione del visitatore sull'estetica dell'involucro che è apprezzabile, ma nasconde il valore scientifico e tecnologico di ciò che c'è all'interno;
- oggetti funzionanti catturano l'attenzione, l'interazione è appagante; i cimeli devono rimanere fuori della portata del visitatore, raramente sono operativi e comunque non conviene esporli a rischi; tuttavia, le macchine possono essere mostrate in funzione tramite filmati; repliche e simulatori sono ottime soluzioni per far sperimentare l'interazione con un calcolatore del passato;
- un “pezzo” deve richiedere pochi secondi per catturare l'attenzione contando su presenza scenica e allestimento; in questo tempo deve fornire tutto il contenuto divulgativo che deve rimanere a ogni visitatore stimolando i più interessati a fermarsi per fruire i contenuti di approfondimento;
- per coinvolgere, aiuta mostrare le tecnologie insieme ai loro effetti sul modo di vivere delle persone; ancora per diversi anni si può giocare sulla nostalgia per macchine riconosciute come vecchie compagne di lavoro o di gioco;
- i videogiochi, oltre a essere una parte dell'informatica di importanza sempre crescente, sono attrazioni irresistibili; devono però essere giocabili (altrimenti generano delusione), con tutti i problemi di gestione di macchine e giocatori;
- i figuranti in costume rendono più attraenti le visite dei più giovani; declinate in brevi rappresentazioni o interviste impossibili, possono incuriosire un pubblico adulto dando alla divulgazione ulteriori spunti di riflessione.

La difficoltà di progettare allestimenti accattivanti mantenendo alto il livello della divulgazione è evidente. È arduo catturare l'attenzione su un cimelio software: è facile solo con un gioco immediatamente riconoscibile, con tutti i *caveat* del caso. È determinante anche la scelta del visitatore da privilegiare: le scuole e i gruppi che partecipano a visite guidate e laboratori hanno esigenze diverse da quelle di famiglie e visitatori che girano liberamente per le sale.

3. Musei di informatica nel mondo e in Italia

La necessità di musei dedicati all'informatica è molto sentita all'estero. Confrontarsi con queste realtà è disarmante per la grandiosità delle installazioni e la straordinaria disponibilità di mezzi e di personale specializzato su hardware e software d'epoca. I musei esteri riescono sia a fornire una visione globale della storia dell'informatica sia a esaltare il particolare contributo della nazione che li ospita. Gli esempi che seguono sono significativi, ma l'elenco non è esaustivo.

3.1 Musei nel mondo

Lo *Science Museum* di Londra ha una vasta sezione dedicata all'informatica che copre il calcolo analogico, il calcolo digitale meccanico (con esempi didattici dei meccanismi), quello elettromeccanico e elettronico, i grandi calcolatori (il Ferranti Pegasus). Importanti sono l'area dedicata a Babbage con la ricostruzione della Macchina alle differenze, e la mostra *Code Breaker*, dedicata al contributo di Alan Turing a Bletchley Park e al progetto ACE.

Proprio a Bletchley Park il *National Computing Museum* ospita, ricostruite e mostrate in funzione, le macchine protagoniste della guerra dei codici durante il II conflitto mondiale (Tunny, Heat Robinson, Colossus). La ricca collezione del museo include anche *mainframe*, personal computer e calcolatori meccanici.

Restando in Inghilterra, citiamo anche il *Museum of Science & Industry* di Manchester, dove è possibile vedere in funzione la replica della Manchester Baby, in assoluto il primo calcolatore a programma memorizzato.

In Germania, a Paderborn, troviamo l'*Heinz Nixdorf MuseumsForum*, candidato a essere il più grande museo di computer del mondo. Tutta la storia dell'informatica è illustrata con ricostruzioni e originali, spesso mostrati in funzione. Sono anche in corso la digitalizzazione degli archivi di Konrad Zuse (in cooperazione con il *Deutsches Museum*) e le ricostruzioni di molte delle sue macchine al *Deutsches Technikmuseum Berlin* e al *Konrad-Zuse-Museum* di Hünfeld.

Al *Polytechnical Museum* di Mosca, attualmente in ristrutturazione, si possono ammirare, all'interno di una ricca collezione, i *mainframe* di produzione russa come lo Strela, lo Ural-1 (completo e funzionante), il BESM-6.

Completiamo questa veloce carrellata citando brevemente, anche in virtù della loro esposizione mediatica, alcuni musei negli Stati Uniti: il *Computer History Museum* a Mountain View, un altro candidato al titolo di più grande museo di computer al mondo, e il *Living Computer Museum* di Seattle, dove le macchine esposte sono tutte restaurate e funzionanti.

3.2 Musei in Italia

Possiamo distinguere due gruppi principali di musei per la storia dell'informatica: i musei di scienza e tecnologia, in cui l'informatica è un tema tra tanti a cui viene dedicato uno spazio proporzionalmente modesto, e i musei specializzati, a volte tenuti da appassionati, in genere piccoli e focalizzati.

Nel primo gruppo rientra il *Museo Nazionale della Scienza e della Tecnologia "Leonardo da Vinci"* a Milano, che recentemente, in occasione dei cento anni dalla nascita di Alan Turing, ha allestito una mostra temporanea (giugno 2012 – febbraio 2013) dal titolo "Le tecnologie che contano".

In questo gruppo, il *Museo delle Tecnologie Elettriche* a Pavia è dedicato al “patrimonio storico della tecnica elettrica e il suo impatto su tutti gli aspetti della vita quotidiana”. In questo contesto, ai calcolatori e all’informatica è riservato un percorso che include anche alcuni esemplari di calcolatrici meccaniche.

Annoveriamo ancora nel primo gruppo il *Museo Galileo* a Firenze, che espone una ricca collezione di strumenti analogici, ma, esclusi un paio di pregiati pezzi del 1600, non tratta lo sviluppo del calcolo digitale e dell’informatica.

Nel secondo gruppo il *Museo degli Strumenti per il Calcolo* di Torino, oltre a strumenti scientifici e per il calcolo analogico, ha una collezione di macchine digitali che spazia dalle calcolatrici meccaniche fino ai mini/micro computer.

Sempre in questo gruppo citiamo il *Laboratorio-Museo Tecnologico@mente* a Ivrea dedicato alle produzioni Olivetti, dalle calcolatrici di Capellaro alla Programma 101, per la quale fornisce anche un servizio di restauro. Il Museo ha anche una peculiare offerta didattica rivolta ai più piccoli.

Il *Museo del Computer* di Novara ha una ricchissima collezione di mini, micro, personal computer e accessori prodotti dagli anni ’70 a oggi. La visita è su appuntamento e il visitatore può non solo osservare, ma come sottolineano i curatori, “toccare, accendere, utilizzare pressoché tutte le macchine”.

Simile in spirito, ma su scala più ridotta, è il *Museo dell’Informatica Funzionante* a Palazzolo Acreide (SR), che fornisce anche l’accesso remoto (tramite telnet) alle macchine storiche rimesse in funzione o replicate tramite simulatori.

A Trieste, la mostra permanente *Percorsi Storici dell’Informatica e del Calcolo* espone un centinaio di pezzi dal XVIII secolo a oggi relativi soprattutto il calcolo personale. Molti sono funzionanti e dimostrabili durante le visite guidate.

Come esempio di museo “verticale”, a Savona *All About Apple* raccoglie una collezione molto ampia dei prodotti del marchio che, oggi, è probabilmente il più rappresentativo nella storia dell’informatica di consumo.

Infine, a metà tra i due gruppi collochiamo il *Museo del Calcolo Mateureka* a Pennabilli (RN), nel quale la storia del calcolo e dell’informatica sono funzionali a un discorso più ampio sulla matematica, rivolto ad un pubblico giovane.

Esistono anche diversi musei virtuali italiani. A titolo di esempio citiamo quello di AICA/CILEA, che prelude alla creazione di un museo tradizionale, e quello dell’Università di Udine, collegato all’allestimento di mostre temporanee.

4. Il Museo degli Strumenti per il Calcolo di Pisa

Il Museo degli Strumenti per il Calcolo dell’Università di Pisa nasce nel 1993, con la costituzione di una commissione ministeriale e un primo finanziamento. La proposta voleva proseguire l’esperienza del *Centro per la Conservazione e lo Studio degli Strumenti Scientifici* (CCSSS) del Dipartimento di Fisica sfruttando la disponibilità di spazi espositivi offerti dal Comune. L’obiettivo dichiarato fu di “*allestire un moderno museo di rilevanza nazionale finalizzato alla conservazione e allo studio di esemplari di calcolatori e, più in generale, di tutto ciò che è stato realizzato e scritto nell’area informatica*” [DM, 1993].

I dettagli della storia successiva sono complessi. In breve, istituito il Museo, la sua gestione fu affidata alla *Fondazione Galileo Galilei*, riattivata allo scopo nel 1998. Purtroppo, i finanziamenti del Ministero sono via via a diminuiti così

come gli spazi concessi dal Comune. Il Museo è stato per molti anni visitabile solo su richiesta o aperto in rare occasioni. Nonostante le difficoltà, grazie in particolare all'impegno di Roberto Vergara Caffarelli, nel corso dei primi anni, fra recuperi, donazioni e acquisizioni, il Museo è riuscito a mettere insieme una collezione di grande rilevanza anche a livello internazionale.

Nel 2011, con finanziamenti ottenuti dal Dipartimento di Informatica, sono stati finalizzati in nuovi allestimenti i risultati di una ricerca storica originale. Da allora il Museo è aperto regolarmente e propone un'offerta consistente di attività e laboratori didattici. La posizione si è ulteriormente consolidata quest'anno con l'istituzione del *Sistema Museale di Ateneo*, con il quale il Museo trova oggi una collocazione di primo piano all'interno dell'organizzazione dell'Università di Pisa.

4.1 Le collezioni

Intendendo "calcolo" nell'ampio significato contemporaneo, il Museo si propone specificatamente come museo dell'informatica, con collezioni centrate sul calcolo digitale, del quale copre la storia dalle prime macchine meccaniche dell'Ottocento ai PC odierni, passando dai grandi calcolatori degli anni 1950-70.

La collezione di macchine meccaniche digitali è uno dei principali vanti del Museo e rappresenta tutta l'evoluzione del calcolo meccanico, dai primi aritmetri De Colmar fino ai successi Olivetti protagonisti del mercato fino alla fine degli anni '60. La ricca collezione di calcolatrici Olivetti dà rilievo a un capitolo nazionale di indubbio interesse sia tecnologico che di storia industriale.

La CEP (1961) e l'Olivetti 9104 CINAC (1966, con parti originali del FINAC del 1955) sono testimoni unici dei primi passi dell'informatica scientifica nazionale. Il Bull Gamma 3 (1953) e l'Olivetti ELEA 6001 (1961), oltre al valore tecnologico (il Bull come passaggio fra tabulatrici e calcolatori moderni) rappresentano la storia Olivetti dalla commercializzazione di macchine estere al suo prodotto più significativo come maturità tecnologica e penetrazione di mercato. Fra gli altri molti cimeli ricordiamo icone assolute come i supercalcolatori Cray.

Per i PC, manca l'Apple I, spesso considerato dal pubblico come un *must*, e l'Altair 8800 (una vera pietra miliare), ma la collezione è capace di mostrare tutta l'evoluzione del PC, dalla Olivetti Programma 101 fino agli anni Novanta con molti pezzi, alcuni curiosi, come il *luggable* Osborne protagonista dell'omonimo effetto, altri già rari, come la coppia di antagonisti Amiga 1000 e Sinclair QL, che segnò l'apice e l'inizio del declino del fenomeno dello home computing.

Il Museo ha anche una collezione di strumenti analogici. Include diversi pezzi di pregio (il planimetro di Gonnella, il regolo Washington) ed è sufficiente a mostrare lo sviluppo parallelo di questi strumenti prima che il calcolo digitale si affermasse. Considerata la ricca collezione di strumenti analogici presente nella vicina Firenze, conviene pensare di instaurare collaborazioni fra musei vicini, per sfruttarne le singole specificità, senza creare inutili sovrapposizioni.

Infine, afferiscono al Museo anche le originarie collezioni del già citato CCSSS, con un'ampia dotazione legata a Antonio Pacinotti. È auspicabile che tali materiali siano in futuro valorizzati in una collocazione propria nel contesto del nuovo Sistema Museale, anche considerando che per le sole collezioni legate al calcolo gli spazi espositivi oggi disponibili sono di gran lunga inferiori alla quantità di pezzi che affollano i magazzini del Museo.

4.2 La ricerca

A partire dal 2006, presso il Dipartimento di Informatica è attivo il progetto di ricerca HMR [HMR, 2013] dedicato alla storia dell'informatica, italiana in particolare. HMR ha obiettivi storici e informatici: da un lato studia gli avvenimenti che hanno accompagnato lo sviluppo dei primi calcolatori; dall'altro vuole far rivivere le tecnologie d'epoca attraverso simulazioni e repliche.

Il significato dell'acronimo, *Hackerando la Macchina Ridotta*, ricorda il primo importante risultato del progetto: la riscoperta della Macchina Ridotta (MR), il primo calcolatore progettato e costruito in Italia, realizzato nel 1957 dall'Università di Pisa. La MR fu realizzata nell'ambito dello stesso progetto che, nel 1961, completò la più nota *Calcolatrice Elettronica Pisana* (CEP). Negli anni successivi la MR fu però dimenticata, sottovalutata e a volte ignorata dalla letteratura storica. HMR, ne ha identificato la grande rilevanza tecnologica, aggiungendo nuove pagine alle vicende del progetto CEP (si veda [Cignoni e Gadducci, 2012a] e la bibliografia ivi contenuta). La nuova lettura della storia del progetto CEP è stata protagonista della mostra "La CEP prima della CEP" inaugurata nel novembre 2011 in occasione del 50esimo della CEP [CEP50, 2013]. Questa mostra ha segnato anche la ripresa delle attività del Museo sia come apertura al pubblico sia come offerta didattica per le scuole.

La ricerca di HMR, oltre a seguire i metodi tradizionali basati sulla raccolta e l'analisi di evidenze documentali e testimonianze dei protagonisti, si distingue per applicare metodi di *archeologia sperimentale* [Cignoni e Gadducci, 2012b]: le macchine sono ricostruite per sottoporle a una completa valutazione, dove ci sono lacune nei progetti originali (molta documentazione si è persa) si procede per ipotesi verificate sperimentalmente con ricostruzioni hardware o simulatori software. Nel nome del progetto il riferimento alla cultura hacker, sottolinea la determinata curiosità richiesta dall'applicazione di questo metodo.

La MR offre ancora molti spunti di ricerca e il lavoro continua, ma HMR ha già allargato i suoi orizzonti con escursioni nella storia dei primi calcolatori Olivetti o sull'evoluzione del calcolo personale dai primi aritmometri ai portatili di oggi. Ai calcolatori Olivetti sono state dedicate le giornate sul design degli ELEA curato da Ettore Sottsass e sull'Agenda illustrata da Bruno Caruso nel 1961. Sul calcolo personale è stato realizzato il percorso "Dall'aritmometro al PC, personalmente" in fase di allestimento al Museo come sala permanente, ma già presentato al pubblico in occasione dell'Internet Festival 2012 [IF, 2012].

HMR persegue anche il coinvolgimento di altri attori nelle attività di ricerca mirate a fornire contenuti originali al Museo. Collaborazioni sono attive con altri Dipartimenti dell'Università (Ingegneria dell'Informazione e Fisica), con istituzioni di ricerca pisane (INFN e ISTI e IIT, istituti del CNR), con Musei (Museo del Computer di Novara, Museo delle Tecniche Elettriche di Pavia).

4.3 Didattica e divulgazione

L'offerta divulgativa del Museo è costruita utilizzando i risultati del progetto HMR, sia storici che strumentali, come le ricostruzioni virtuali della MR o la replica hardware del suo addizionale, protagoniste dei laboratori didattici. Le attività sono principalmente rivolte alle scuole medie e superiori, in alcuni casi

adattabili alle classi degli ultimi anni delle elementari. Le attività offrono approfondimenti sulle vicende italiane e sui calcolatori realizzati all'Università di Pisa, ma tutta la storia dell'informatica è affrontata nei suoi molteplici aspetti: scientifici, tecnologici, sociali e culturali. Tutta l'offerta ha come obiettivo comune la divulgazione scientifica: le tecnologie di oggi, in particolare la simulazione software, sono usate per far rivivere i calcolatori di ieri; nella loro semplicità, i calcolatori di ieri sono ottimi esempi per spiegare le tecnologie di oggi.

Le attività sono organizzate in moduli propedeutici che affiancano l'insegnamento dell'informatica a più livelli di approfondimento per instaurare con le scuole e con i docenti un percorso divulgativo pluriennale. La collaborazione con gli insegnanti e, in generale, con il pubblico del Museo è perseguita sia con strumenti tradizionali, quali il "libro delle visite" e i questionari di gradimento, sia animando sui social network una comunità partecipe alle iniziative del Museo.

I moduli di attività si dividono in *incontri*, *visite guidate* e *laboratori*. Gli incontri servono a introdurre, incuriosire e destare interesse stimolando domande e offrendo spunti di discussione. Le visite guidate usano l'unicità delle collezioni del Museo per legare i concetti alla concretezza e al fascino dei cimeli visti da vicino. I laboratori, basati sull'interazione con le ricostruzioni, propongono sfide per consolidare le nozioni con l'esperienza pratica. Per approfondimenti si rimanda a [Cignoni e Gadducci, 2013], dove sono descritti i dettagli dell'offerta didattica e l'uso di cimeli (nelle visite guidate) e repliche (nei laboratori) come testimoni affascinanti per spiegare i principi e i meccanismi dell'informatica.

Oltre alle attività proposte alle scuole, il Museo partecipa alla vita culturale cittadina e regionale. Ad esempio, gli incontri sui calcolatori Olivetti sono stati realizzati con il Museo della Grafica di Pisa. Eventi particolari sono stati organizzati per la Notte dei Musei, per gli Open Day della ricerca, per l'Internet Festival e per Pianeta Galileo. Iniziative che continueranno nelle prossime edizioni di tali manifestazioni. La visibilità del Museo è curata anche sui social network e sulle comunità nazionali di blogger; per esempio sono state organizzate visite speciali in occasione di iniziative quali *Invasioni Digitali* e *Blog Tour*.

5. Un progetto nazionale?

Alla diffusione del *public understanding* dell'informatica, fattore importante per lo sviluppo di una società moderna come crediamo debba essere quella italiana, può dare un contributo essenziale la divulgazione della sua storia, da realizzare anche tramite un'efficace rete di musei dedicati.

Da questo punto di vista, le differenze fra noi e l'estero – in termini di spazi, personale, risorse per gli allestimenti, finanziamenti, coinvolgimento delle istituzioni – sono grandi e sfavorevoli. A titolo di esempio: al National Museum of Computing di Bletchley Park contano su 50 persone fra storici e tecnici per la ricerca, il restauro e la manutenzione delle macchine; al Museo di Manchester dopo la ricostruzione integrale nel 1998 della Baby (un calcolatore con una storia alla quale quelle della MR e della CEP si avvicinano) hanno sostenuto 170 riparazioni per continuare a mostrarla in funzione. In generale, in Europa sono più attenti di noi, inclusi paesi quali Russia, Ucraina e Polonia. I grandi musei statunitensi di Mountain View e Seattle restano punti di confronto irraggiungibili.

Nonostante l'estremo divario di risorse a disposizione, o forse proprio per questo, la situazione italiana lascia spazio a una certa soddisfazione: considerati gli handicap, la ricerca storica e la qualità delle proposte didattiche sono paragonabili alle migliori in Europa. Da questo punto di vista, come abbiamo cercato di mostrare nelle sezioni precedenti, il Museo di Pisa ha raggiunto negli ultimi anni risultati ragguardevoli, i quali, assieme alla completezza delle collezioni e ai cimeli particolari, legati alla storia del settore – la CEP *in primis* – lo pongono in posizione privilegiata, anche come responsabilità, in un'iniziativa per l'istituzione in Italia di un Museo Nazionale dell'informatica.

Questo articolo vuol essere l'assunzione pubblica di tale responsabilità, con l'indicazione di far diventare il Museo dell'Università di Pisa un'istituzione di rilevanza nazionale che convogli e amplifichi attenzione, risorse intellettuali e, naturalmente, finanziamenti per raggiungere l'obiettivo di un'apprezzamento collettivo della natura dell'informatica e, quindi, della comprensione del suo ruolo nella società moderna, e in quella italiana in particolare.

D'altra parte, da addetti ai lavori, proprio la coscienza del ruolo dell'informatica nella società contemporanea ci impone un'iniziativa *smart*, che insieme ad AICA, che ci ospita e che si è più volte mostrata sensibile al tema, chiami alla collaborazione tutti gli attori pubblici e privati interessati.

Riferimenti

[CCS, 2013] Computer Conservation Society, www.computerconservationsociety.org

[CEP50, 2013] La CEP prima della CEP, cep50.di.unipi.it

[Cignoni e Gadducci, 2012a] Cignoni G.A., Gadducci F., Experimental Archaeology of Computer Science. Atti della Società Toscana di Scienze Naturali, Serie B 119, 2012.

[Cignoni e Gadducci, 2012b] Cignoni G.A., Gadducci F., Rediscovering the Very First Italian Digital Computer. Proceedings of 3rd IEEE HISTory of ELeCTro-technology CONference (HISTELCON). IEEE Computer Society, 2012.

[Cignoni e Gadducci, 2013] Cignoni G.A., Gadducci F., Calcolatori d'epoca per la didattica dell'Informatica al Museo degli Strumenti per il Calcolo. Atti del Convegno su Didattica e Informatica (DIDAMATICA). AICA, 2013.

[DM, 1993], Decreto del Ministro dell'Università e della Ricerca Scientifica e Tecnologica, 16 aprile 1993.

[HMR, 2013] Hackerando la Macchina Ridotta, hmr.di.unipi.it

[Hollands, 2008] Hollands, R.G., Will the real smart city please stand up? City: analysis of urban trends, culture, theory, policy, action 12(3), 2008, 303-320.

[MHCR, 2013] Making the History of Computing Relevant, IFIP Conference, London Science Museum, 17-18 Giugno 2013, historyofcomputingjune2013.wordpress.com

[MSC, 2013] Museo degli Strumenti per il Calcolo dell'Università di Pisa, www.fondazionegalileogalilei.it

[PUS, 1985] The Public Understanding of Science. The Royal Society, 1995.

[IFIPWG9.7, 2013] History of Computing, www.comphist.org

Formazione e Certificazione: possibili risparmi per il Sistema Italia, per le regioni, per i cittadini e la soddisfazione degli utenti.

Antonio Piva¹, Michele Baggi²

¹ Università di Udine, AICA nord-est, antonio@piva.mobi

² AICA nord-est, michele.baggi@tiscali.it

Abstract. *This survey, emerging from more than 160 Italian local governments, refers to the opinion polls submitted to IT managers concerning technical support costs, which can be saved through an adequate training of local governments employees. These collected opinions are compared with the experimental survey led together by AICA and SDA Bocconi. The present article introduces employees assessment and satisfaction about ECDL training and certification; estimates potential savings for local governments, the whole Italian system, and costs imposed on every single citizen.*

Keywords: costs savings, ECDL training, government efficiency.

1. Introduzione

In questo articolo vengono presentate le indagini sulla percezione dei responsabili dei sistemi informativi delle Pubbliche Amministrazioni Locali e dei dipendenti delle stesse P.A.L., utilizzatori finali degli strumenti informatici, circa il numero di interventi tecnici evitabili in seguito ad adeguata formazione del personale dipendente degli Enti Locali.

Il dato rilevato dai referenti dei sistemi informativi degli Enti Locali è stato confrontato con la ricerca sperimentale già condotta da AICA e SDA Bocconi (presentata in [1], al paragrafo 2.3): il numero di interventi tecnici evitabili in seguito a formazione del personale pubblico, infatti, risulta essere maggiore rispetto alla percezione fornita da parte dei responsabili dei sistemi informativi dei Comuni.

Per la ricerca in esame sono stati intervistati i responsabili dei sistemi informativi di oltre 160 Comuni del Triveneto; questi referenti sono stati invitati a rispondere a domande circa l'infrastruttura informatica lato utente (per conoscere il numero di postazioni utente disponibili) e il tipo di software installato per la produttività individuale.

La ricerca mette in evidenza importanti risultati economici ottenibili a fronte di un'adeguata formazione del personale dipendente mediante certificazione ECDL: è stato possibile, infatti, calcolare il possibile risparmio medio annuo per ogni singolo dipendente, riguardante la riduzione dell'assistenza tecnica informatica in seguito alle competenze acquisite. Con questo dato è stato inoltre possibile calcolare il risparmio medio annuo degli Enti Locali, a livello regionale

e nazionale. Infine si presentano i risultati riguardanti il risparmio medio annuo per ciascun cittadino italiano, per ogni singolo contribuente e per ogni famiglia italiana.

Dalle opinioni dei dipendenti della Pubblica Amministrazione Locale che hanno partecipato alla formazione ECDL emerge che vi siano stati benefici economici riguardanti in particolare l'aumento della produttività e la diminuzione degli interventi tecnici richiesti.

Non si tratta tuttavia di soli benefici economici: i dipendenti intervistati hanno infatti evidenziato una notevole soddisfazione in merito alle nuove competenze informatiche acquisite in seguito alla formazione ECDL. Inoltre la produzione di documenti di miglior qualità, fattore di cui i dipendenti pubblici sono ben consci, contribuisce ad una maggior soddisfazione da parte del cittadino, il quale avrà a disposizione modulistica e documentazione maggiormente fruibile e molto più comprensibile.

2. La valutazione dei responsabili dei sistemi informativi degli Enti Pubblici locali circa i costi dell'assistenza tecnica informatica e i possibili risparmi dovuti alla formazione dei dipendenti delle P.A.L.

In questa indagine si è voluto innanzitutto conoscere la valutazione dei referenti per l'informatica dei Comuni riguardo al numero di interventi di assistenza tecnica richiesti dai loro dipendenti a causa della loro scarsa conoscenza e capacità di utilizzo degli strumenti informatici; la ricerca è stata eseguita mediante intervista ai responsabili dei sistemi informativi degli Enti Pubblici Locali e/o referenti dell'assistenza tecnica dei Comuni. In secondo luogo si è voluta stimare la percezione degli intervistati circa l'efficacia della formazione e certificazione ECDL quale possibile soluzione per ridurre il numero di tali interventi, con un conseguente risparmio di costi per gli Enti.

Mediante questa indagine sono stati raccolti i dati di 164 Comuni del Friuli Venezia Giulia e del Veneto. Per focalizzare al meglio la significatività del campione analizzato, si pensi che soltanto per il Friuli Venezia Giulia sono stati sottoposti ad intervista i referenti del 75% dei Comuni regionali. I Comuni censiti coprono una popolazione di 1.118.005 abitanti, pari al 91% del totale regionale; visto anche la grande eterogeneità dei Comuni presi in considerazione (si va dai Comuni al di sotto dei 200 abitanti fino a Comuni che superano i 100.000 residenti), possiamo affermare che il campione analizzato risulta essere sufficientemente significativo.

I quesiti proposti sono serviti a valutare la percezione dei responsabili dei sistemi informativi riguardo al numero di interventi di assistenza tecnica dovuti alla poca dimestichezza dei dipendenti degli Enti Locali con gli strumenti informatici. È stato interessante inoltre conoscere la previsione del numero evitabile di tali richieste di assistenza a fronte di una formazione e certificazione ECDL dei dipendenti degli Enti Locali.

L'intervista consisteva nelle seguenti due domande:

1. *Su 100 richieste di assistenza tecnica informatica da parte dei dipendenti, quante secondo lei sono dovute a poca dimestichezza da parte dell'utente nei confronti degli strumenti informatici utilizzati.*
2. *Di queste richieste di assistenza tecnica, quante se ne potrebbero evitare se il personale dipendente fosse formato e certificato ECDL?*

2.1 Risultati dell'intervista ai responsabili dei Comuni

Rielaborando le risposte fornite al primo quesito, i responsabili dei sistemi informativi dei Comuni reputano che il 39% delle richieste di assistenza è dovuto alle scarse conoscenze e capacità di utilizzo degli strumenti informatici da parte degli utenti. Per quanto concerne il secondo quesito, riguardante il numero di interventi tecnici evitabili in seguito a formazione ECDL del personale, si ottiene che gli intervistati prevedono una riduzione del 45% degli interventi qualora il personale dipendente ricevesse una adeguata formazione e certificazione ECDL.

Analizziamo innanzitutto le risposte al primo quesito, riguardanti la percentuale di interventi tecnici dovuti a scarsa conoscenza degli strumenti informatici da parte dei dipendenti delle P.A.L.. Secondo la rilevazione sperimentale [2], la percentuale effettiva di questi interventi tecnici nei Comuni soggetti all'indagine, calcolata mediante rilevazione sul numero effettivo di interventi tecnici, sfiorava il 54%, con punte dell'82,9% nel comune di Caorle e nella rete di comuni circostanti Valdobbiadene e dell'81,6% nei Comuni del Medio Friuli, mentre si riscontrava il 38,3% nel Comune di Udine.

La Tabella 1 riepiloga i risultati della ricerca sperimentale.

Comuni	Interventi evitabili	Interventi totali	% interventi evitabili
Comune di Udine	1800	2904	38,3%
Comuni del Medio Friuli	1670	2047	81,6%
Comuni di Caorle e Valdobbiadene	491	592	82,9%
Totale	3961	7343	54%

Tabella 1: Numero di interventi tecnici eseguiti presso i Comuni campione nel periodo Aprile 2009 - Marzo 2010, suddivisi per categoria e totali.

Confrontando i dati ottenuti dall'elaborazione delle risposte al primo quesito con i dati rilevati nella ricerca sperimentale, notiamo **che la stima fornita dai responsabili circa il numero di interventi di assistenza tecnica dovuti alla poca conoscenza degli strumenti informatici da parte dei dipendenti (39%) è sensibilmente inferiore al valore medio riscontrato dalla ricerca sperimentale (54%).**

Analizzando inoltre il dato sulle risposte al secondo quesito, inerente alla percentuale di interventi tecnici evitabili in seguito a formazione e certificazione ECDL del personale delle Pubbliche Amministrazioni Comunali, emerge che

secondo i responsabili dei sistemi informativi il 45% degli interventi potrebbe essere evitato in seguito ad adeguata formazione del personale. Confrontando tale valore con i dati emersi dalla ricerca [2], si nota **che la percezione dell'efficacia della formazione ECDL sull'abbattimento delle richieste di assistenza dovute alla scarsa conoscenza degli strumenti informatici è anch'essa inferiore all'efficacia riscontrata sperimentalmente**. Nella ricerca sperimentale infatti si calcolava una riduzione di tali richieste di assistenza di oltre il 64%, ben al di sopra del 45% stimato dagli intervistati: ciò significa che i responsabili dei sistemi informativi comunali sottovalutano di quasi il 20% l'effettiva riduzione degli interventi tecnici dopo formazione ECDL. **Quindi anche in questo caso si nota come i responsabili sottostimino il numero di interventi (e conseguentemente dei costi) che potrebbero essere evitati in seguito alla formazione ECDL del personale delle Pubbliche Amministrazioni Comunali.**

3. La valutazione degli utilizzatori dei sistemi informatici negli Enti Pubblici locali riguardo al valore della formazione ECDL

Sono stati organizzati corsi di formazione sui 7 moduli ECDL *core* presso i Comuni di Caorle, Udine, la rete dei 10 Comuni gravitanti attorno a Codroipo e la rete dei 4 Comuni del circondario di Valdobbiadene. Ai corsi ha partecipato circa un centinaio di dipendenti degli Enti Locali, molti dei quali si sono successivamente certificati ECDL. Dopo il periodo di formazione, ai corsisti è stato sottoposto un questionario per la raccolta di opinioni e considerazioni inerenti al corso seguito, in particolare riguardo alle effettive competenze acquisite in virtù della formazione ricevuta; scopo di questa ulteriore raccolta di dati era quello di valutare la percezione degli utilizzatori veri e propri degli strumenti informatici all'interno dei Comuni circa i reali benefici da essi riscontrati dopo alla formazione e certificazione ECDL.

Analizzando i risultati dell'indagine scopriamo che alla domanda "Ritieni che la formazione abbia ridotto il numero di interventi tecnici?" è stato risposto in modo affermativo nel 59% dei casi, pertanto dopo la formazione e certificazione ECDL la maggior parte degli utenti dei sistemi informatici negli Enti Locali ritiene che avere buone competenze nell'uso di strumenti informatici possa portare dei risparmi anche per quanto concerne la diminuzione del numero di interventi tecnici; il supporto tecnico informatico non sarà più contattato per problematiche riguardanti l'uso degli strumenti di informatica d'ufficio, quali ad esempio problemi con gli allegati delle e-mail, rimozione destinatario da una e-mail, problemi con la formattazione del testo in Word, disabilitazione del controllo di ortografia in Word, difficoltà nella gestione delle tabelle di Excel, eccetera.

Altra considerazione di interesse è che l'85% dei dipendenti intervistati ritiene di aver migliorato la propria produttività in seguito alla formazione e certificazione ECDL: un risparmio notevole di tempo per i dipendenti delle P.A.L., valore che potrebbe a sua volta essere quantificato e tradotto in un notevole risparmio economico per gli Enti Locali. Interessante infine notare che quasi tre dipendenti su quattro ritengono che le competenze acquisite in seguito a formazione ECDL siano servite per la produzione di documenti

qualitativamente migliori; questo dato evidenzia una ricaduta sociale della formazione ECDL, in quanto l'aumento di qualità nei documenti prodotti implica un grande beneficio alla cittadinanza che in ultima istanza dovrà utilizzare questi documenti, resi ora più comprensibili e meglio redatti.

I risultati sono presentati in Tabella 2 e illustrati graficamente in Figura 1.

Quali sono i principali vantaggi della formazione e certificazione ECDL?	Migliorare la produttività	85,71%
	Creare documenti migliori	73,47%
	Aiutare i colleghi in difficoltà	57,14%
	Ridurre gli interventi tecnici	59,18%

Tabella 2: I principali vantaggi riscontrati dai dipendenti dei Comuni che hanno seguito il corso di formazione sui moduli ECDL core.

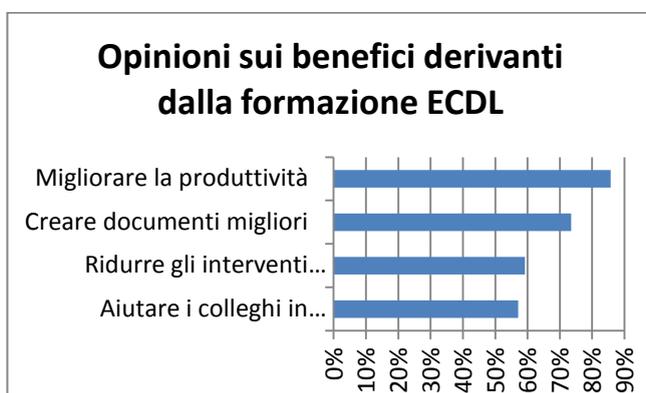


Figura 1: Le opinioni dei corsisti sui benefici derivanti dalla formazione e certificazione ECDL.

Se andiamo ad analizzare la seconda parte del questionario, riguardante le considerazioni dei dipendenti pubblici circa la formazione ECDL, scopriamo nuovi dati di grande interesse. In primis viene riconosciuta la positività dell'iniziativa dell'Amministrazione Comunale circa la proposta di un corso di formazione informatica per i dipendenti (quasi 94%). Si tratta di un dato assai importante, in quanto denota una **grande soddisfazione da parte dei dipendenti comunali nei confronti dell'iniziativa dell'Amministrazione.**

Un altro dato interessante, evidenziato da quasi il 90% dei dipendenti comunali, è che viene ritenuto estremamente utile il conseguimento della certificazione ECDL da parte del personale delle P.A.L.. Il dato assume notevole valenza in conseguenza alle considerazioni emerse nella prima parte dell'intervista (miglioramento della produttività e riduzione interventi tecnici, Tabella 3), considerando inoltre che i dipendenti comunali partecipanti alla formazione non erano "digiuni" nell'utilizzo degli strumenti informatici. In altre parole, **la quasi totalità dei dipendenti delle Pubbliche Amministrazioni**

Locali afferma l'utilità delle competenze ECDL, compresi coloro che già utilizzano con costanza strumenti informatici di vario genere.

Notiamo infine la dichiarazione del 75% dei dipendenti, i quali sarebbero disposti a seguire dei corsi di formazione più approfondita (ad esempio ECDL *Advanced*) per migliorare ulteriormente le proprie capacità informatiche.

I risultati di questa parte sono illustrati nella Tabella 3 e nella Figura 2.

Dopo aver partecipato alla formazione ECDL ritieni che...	L'iniziativa dell'Amministrazione sia stata molto buona.	93,88%
	Il corso abbia stimolato altri colleghi a richiedere questa formazione.	48,98%
	La certificazione ECDL sia generalmente molto utile.	89,80%
	Hai suggerito ad altri (amici, parenti) di conseguire l'ECDL.	51,02%
	Faresti dei corsi avanzati in materia (tipo ECDL <i>Advanced</i>).	75,51%

Tabella 3: Prima parte del questionario riservato ai soli dipendenti dei Comuni che hanno seguito il corso di formazione sui moduli ECDL core.

Volendo trarre alcune conclusioni dopo questa ulteriore indagine, possiamo affermare che gli "addetti ai lavori" all'interno degli uffici comunali accolgono con estremo favore iniziative di formazione sulle tematiche legate all'uso degli strumenti informatici, avvertendone essi per primi l'esigenza e l'utilità. In secondo luogo una formazione di questo tipo potrebbe portare **vantaggi economici alle Pubbliche Amministrazioni Locali principalmente in termini di risparmio di tempo/denaro, attraverso la riduzione del numero di interventi tecnici, l'aumento della produttività e, non da ultimo, miglioramenti nella qualità dei documenti prodotti, motivo di apprezzamento da parte dei cittadini.**

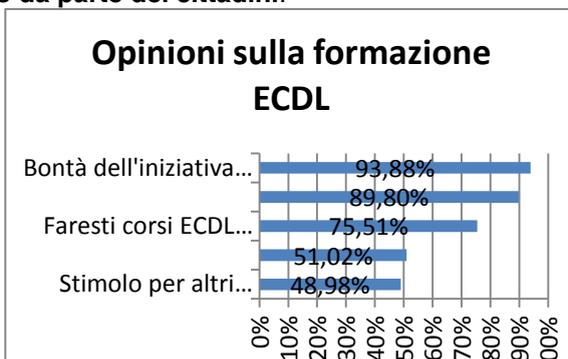


Figura 2: Opinioni dei corsisti. Si nota un apprezzamento dell'iniziativa formativa, oltre a considerazioni generali genericamente molto positive.

4. I possibili risparmi complessivi nella P.A.L., in particolare a livello regionale, e l'impatto sui contribuenti italiani

In questo paragrafo vengono calcolati i possibili risparmi per gli Enti Locali, ottenibili con formazione ECDL del personale dipendente. Viene così eseguita la proiezione a livello regionale, al fine di permettere il calcolo del risparmio medio annuo per ogni singolo contribuente italiano.

4.1 I dati della Pubblica Amministrazione Locale a livello regionale

Prendendo come "punto di partenza" la regione Friuli Venezia Giulia, il costo annuo per l'assistenza tecnica evitabile (descritto in [2]) è pari a 1.681.445 €, ed aggiungendo il costo annuo per le licenze d'uso proprietarie di software per l'informatica d'ufficio sostituibili dall'open source (993.800 € il costo annuo), si ottiene il costo annuo globale pari a 2.675.245 euro.

È interessante a questo punto comprendere quale sia il costo annuo per ogni dipendente degli Enti Locali (potenzialmente risparmiabile). Il valore, ottenuto dividendo il costo annuo per il numero totale di dipendenti negli Enti Locali del Friuli Venezia Giulia (15.078), è risultato pari a **177,43 Euro** per ciascun dipendente pubblico. A questo potenziale risparmio annuo per dipendente va aggiunto il costo ignoranza informatica ed il risparmio derivante dall'aumento produttività del dipendente [1], per un risparmio complessivo di **4.005,43** euro annui per dipendente (Tabella 4).

Risparmio medio annuo complessivo per ciascun dipendente della P.A.L. in F.V.G.	Risparmio annuo per licenze e Assistenza evitabile, per singolo dipendente	Risparmio annuo miglioramento della produttività individuale	Risparmio annuo complessivo risparmiabile per dipendente
		177,43 €	3.828 €

Tabella 4: Il risparmio medio annuo per gli Enti Pubblici Locali in Friuli Venezia Giulia è ottenuto dalla somma del risparmio medio annuo per l'assistenza tecnica evitabile e per le licenze d'uso del software per l'informatica d'ufficio e il risparmio per il miglioramento della produttività.

Avendo a disposizione i dati del Friuli Venezia Giulia, è interessante stimare il possibile risparmio medio annuo in ciascuna regione italiana. Per il calcolo possiamo effettuare una nuova proiezione dei dati del Friuli Venezia Giulia, al fine di quantificare il possibile risparmio medio annuo per ogni singola regione italiana, moltiplicando il risparmio medio annuo complessivo per ciascun dipendente per il numero di dipendenti di Comuni, Province e Regioni (Tab. 5).

	Dipendenti di Comuni, Province, Regione	Risparmio medio annuo complessivo
Totale	575.162	2.303.771.130 €

Tabella 5: Risparmi medi annui per ciascuna Regione italiana, basandosi sul numero di dipendenti di Comuni, Province e Regioni.

Se ai 575.162 dipendenti di Comuni, Province e Regioni aggiungiamo anche i dipendenti di Comunità Montane e Unioni di Comuni, a livello nazionale il numero sfiora le 595 mila unità: si tratta infatti di 594.855 dipendenti di enti locali [1], senza considerare l'Amministrazione Pubblica Centrale (i dipendenti dei palazzi ministeriali, di Camera e Senato, eccetera), la Pubblica Sicurezza, le Forze Armate, i dipendenti scolastici e i dipendenti del Sistema Sanitario Nazionale. Essendo il possibile risparmio medio annuo di 4.005,43 euro per ciascun dipendente pubblico, la proiezione a livello nazionale per i soli Enti Locali raggiunge l'esorbitante valore di 2.382.650.063 di euro annui.

Ci si può chiedere, a questo punto, quale sia l'incidenza di tali costi su ogni singolo cittadino. Dividendo il possibile risparmio annuo totale per il numero di cittadini italiani, otteniamo un "peso" annuo di 39,30 € a cittadino. (Tabella 6)

Risparmio totale annuo Italiano nella P.A.L.	Numero di abitanti residenti in Italia (Fonte: censimento ISTAT 2011)	Risparmio potenziale annuo per ogni cittadino italiano
2.382.650.063 €	60.626.442	39,30€

Tabella 6: Costo dovuto a scarsa formazione ECDL del personale delle P.A.L.

A livello regionale, oltre ai possibili risparmi derivanti dalla formazione del personale dipendente degli Enti Locali, è possibile considerare anche il risparmio riguardante i dipendenti della Sanità. Considerando i dati emersi e pubblicati in un precedente articolo di Mondo Digitale,[2] riguardanti il potenziale beneficio per ogni cittadino derivante dall'aumento di produttività del personale sanitario in seguito a formazione ECDL e dalla diminuzione di interventi tecnici informatici, è stato possibile integrare i dati sul risparmio complessivo a livello regionale. Il calcolo, effettuato in modo del tutto simile a quanto esposto nei paragrafi precedenti, quantificava il possibile risparmio, riguardante il personale sanitario, in 39,50 € per ogni cittadino italiano. Possiamo quindi proiettare il valore del risparmio, considerando il numero di cittadini di ogni regione italiana [3], calcolando così l'intero possibile risparmio derivante dalla adeguata formazione del personale a livello regionale. I dati sono esposti in Tabella 7.

Notiamo che la somma tra il risparmio medio annuo per la Pubblica Amministrazione Locale (2.303.771.130€) e il risparmio per il Servizio Sanitario (2.410.770.676 €) potrebbe raggiungere la cifra di 4.714.541.810 € all'anno (Il potenziale beneficio per ciascun cittadino italiano è di circa 77,80 €). Pertanto **oltre quattro miliardi e mezzo di euro si potrebbero risparmiare ogni anno se il personale delle Pubbliche Amministrazioni Locali e dei Servizi Sanitari Regionali avesse le adeguate competenze riguardanti l'uso degli strumenti informatici, rispetto agli standard ECDL.**

Regione	Risparmio medio annuo: sanità	Risparmio medio annuo: P.A.L.	Risparmio regionale medio annuo complessivo
Abruzzo	53.055.018 €	47.624.563 €	100.679.581 €
Basilicata	23.183.064 €	27.160.821 €	50.343.885 €
Calabria	79.408.825 €	79.728.084 €	159.136.909 €
Campania	239.957.839 €	216.773.872 €	456.731.711 €
Emilia Romagna	175.478.790 €	157.457.458 €	332.936.248 €
Friuli Venezia Giulia	48.807.109 €	59.997.336 €	108.804.445 €
Lazio	226.945.512 €	199.686.707 €	426.632.219 €
Liguria	63.878.097 €	75.430.257 €	139.308.354 €
Lombardia	395.265.045 €	311.225.916 €	706.490.961 €
Marche	61.872.247 €	60.385.863 €	122.258.110 €
Molise	14.599.674 €	14.619.820 €	29.219.494 €
Piemonte	176.266.064 €	167.378.909 €	343.644.973 €
Puglia	161.578.266 €	109.043.826 €	270.622.092 €
Sardegna	66.162.342 €	68.621.027 €	134.783.369 €
Sicilia	198.667.225 €	271.307.801 €	469.975.026 €
Toscana	148.291.019 €	147.992.628 €	296.283.647 €
Trentino - Alto Adig	41.061.593 €	96.134.325 €	137.195.918 €
Umbria	35.857.350 €	37.566.928 €	73.424.278 €
Valle d'Aosta	5.070.852 €	18.252.745 €	23.323.597 €
Veneto	195.364.749 €	137.382.244 €	332.746.993 €
Totale	2.410.770.676 €	2.303.771.130 €	4.714.541.810 €

Tabella 7: Calcolo del possibile risparmio complessivo dovuto a formazione ECDL suddiviso per regioni (personale del Servizio Sanitario e degli Enti Locali).

4.2 Ritorno economico della formazione ECDL e impatto sul contribuente e sulle famiglie

I possibili risparmi derivanti dalla diminuzione dell'ignoranza informatica da parte dei soli dipendenti delle P.A.L., dall'aumento di produttività degli utenti e dal costo delle licenze si ripercuotono in misura non indifferente sulle tasche del contribuente: considerando i soli contribuenti IRPEF, che in Italia ammontano a 41.802.902 persone [4], il risparmio annuo pro capite potrebbe essere di 58,58 euro. Considerando anche i dati del Servizio Sanitario, secondo i quali il risparmio per ogni contribuente potrebbe essere di 57,67 euro, otteniamo un possibile **risparmio complessivo di 116,25 € per ogni singolo contribuente italiano.**

Considerando infine che in Italia il numero di famiglie (dati censimento ISTAT 2011, consultabili dal sito dati.istat.it) è di 24.618.071, otteniamo un costo annuo medio di 99,48 euro per ciascuna famiglia italiana, considerando soltanto i possibili risparmi derivanti dalla Pubblica Amministrazione Locale. Se a valore aggiungiamo i 97,93 euro derivanti dal risparmio nel Sistema Sanitario, **otteniamo ben 197,41 euro in più ogni anno di potenziale risparmio per ogni singola famiglia italiana.**

5. Conclusione

In un momento storico caratterizzato da "tagli" causati dalla dilagante crisi economica, per le Pubbliche Amministrazioni Locali si affaccia una possibilità di risparmio e aumento di produttività del personale. La chiave di lettura per questa "boccata d'ossigeno" è rappresentata dalla formazione ECDL di base del personale delle P.A.L., espandibile eventualmente a tutte le altre categorie di dipendenti statali o degli Enti. L'indagine condotta presso i responsabili dei sistemi informativi degli Enti Locali ha messo in luce la consapevolezza degli stessi circa i possibili risparmi nell'assistenza tecnica dovuta a scarsa conoscenza degli strumenti informatici da parte dei dipendenti pubblici. Si tratta tuttavia di una percezione inferiore rispetto a quanto riscontrato dalla ricerca sperimentale, secondo la quale l'aumento di produttività dei dipendenti pubblici formati secondo lo standard ECDL è molto più elevato rispetto alla stima fornita dai responsabili. Questo fattore è da tenere in considerazione in sede di formazione del personale, in quanto la formazione ECDL del personale pubblico avrebbe notevoli ripercussioni sulla produttività e sulla soddisfazione degli stessi dipendenti comunali. Inoltre si potrebbe tenere in considerazione la formazione informatica di un potenziale nuovo dipendente, il quale dal momento della sua assunzione garantirebbe all'Ente Locale una maggiore produttività.

Per quanto concerne la formazione del personale, gli stessi dipendenti delle P.A.L. avvertono l'utilità di una maggiore formazione informatica, al fine di ridurre il numero di interventi tecnici e aumentare notevolmente la produttività degli uffici comunali. L'iniziativa delle diverse Amministrazioni Comunali, riguardante la realizzazione di corsi di formazione ECDL, è stata infatti accolta con grande favore dai dipendenti stessi, i quali hanno colto immediatamente l'opportunità per accrescere le proprie competenze informatiche. Oltre alla produttività, gli utenti dei sistemi informatici delle P.A.L. avvertono un sensibile margine di miglioramento nella qualità dei documenti prodotti, a maggior beneficio degli uffici comunali e, soprattutto, del cittadino.

In base ai dati rilevati è stato possibile calcolare un risparmio di oltre 4.000 Euro annui per ogni singolo dipendente della P.A.L., valore che se tradotto in termini nazionali significherebbe risparmiare mediamente 2,3 miliardi di euro all'anno. Se a questo valore aggiungiamo i 2,4 miliardi di euro potenzialmente risparmiabili dai Servizi Sanitari Regionali in seguito ad analoga formazione ECDL del personale sanitario, otteniamo un potenziale risparmio globale che si aggirerebbe attorno **ai 4,7 miliardi di euro all'anno**. Nel caso in cui tale somma di denaro dovesse essere destinata ad un minor carico fiscale, il contribuente italiano e **le famiglie risparmierebbero rispettivamente importi medi fino a 120 € e 200 € annui**. Infine i dati emersi nelle già citate ricerche sull'ignoranza informatica portano ad un potenziale risparmio annuo al Sistema Italia di oltre **18 miliardi di Euro** calcolando anche gli utenti del settore privato.

Pertanto, attraverso la formazione e certificazione ECDL del Personale della Pubblica Amministrazione Locale e del Sistema Sanitario, ma anche della Pubblica amministrazione centrale e degli Enti, si potrebbero cogliere appieno le innovazioni dell'Agenda Digitale, aumentando la produttività e la competitività dell'intero Sistema Paese.

Fuzzy clustering in user segmentation

Anna Maria Fanelli, Leo Iaquinta¹, Giovanni Semeraro, M. Alessandra Torsello¹
University of Bari - Dep. of Computer Science
Via Orabona 4, 70125 Bari (Italy)
{fanelli, semeraro}@di.uniba.it

¹ University of Milano-Bicocca - Dep. of Informatics, Systems and Communication
Viale Sarca 336, 20126 Milano (Italy)
{iaquinta, torsello}@disco.unimib.it

Abstract. *E-Government is becoming more attentive towards providing personalized services to citizens so that they can benefit from better services with less time and effort. To develop citizen-centered services, a fundamental activity consists in mining needs and preferences of users by identifying homogeneous groups of users, also known as user segments, sharing similar characteristics. Since the same user often has characteristics shared by several segments, in this work we propose an approach based on fuzzy clustering for inferring user segments that could be properly exploited to offer personalized services that better satisfy user needs and their expectations. User segments are inferred starting from data, gathered by questionnaires, which essentially describe demographic characteristics of users. For each derived segment a user profile is defined which summarizes characteristics shared by users belonging to that segment. Results obtained on a case study are reported in the last part of the paper.*

Keywords: user segmentation, fuzzy clustering, user-centered services.

1. Introduction

The rapid development of new telecommunication and computerized technologies in the last decades has offered new possibilities for providing citizens and businesses with better and more efficient services. E-Government (e-Gov) refers to the use of information and communication technologies, particularly the Internet, as a tool to achieve better government. Recent research is focusing on the benefits that can derive from endowing governments with some kind of personalization to promote their communication opportunities with citizens and to provide services in a user-centered way [Guo and Lu, 2007] [Lu et al, 2010]. In other terms, user needs have to be placed at the center of the development and the provision of public services. To do this, in

e-Gov domain a crucial activity consists in acquiring extensive knowledge about target users in order to mine their preferences and needs for services by identifying homogeneous groups of users, also known as user segments, sharing similar characteristics.

Cluster analysis has long been employed as a tool for deriving user segments [Punj and Stewart, 1983] [Namvar et al, 2010]. The analysis is often performed in a crisp partition form, i.e. user segments are disjoint, thus each user may belong to only one group. However, in real scenarios, based on their characteristics, users are not well distributed and separated. Often, a user may share characteristics with several user groups. Thus, one user does not always belong exclusively to a unique segment, but she/he may often belong to two or more segments with different membership degrees. This scenario can be better addressed by cluster analysis techniques that infer fuzzy partitions, i.e. (partially) overlapping user segments.

To address the fuzzy nature that characterizes segment boundaries, we propose an approach based on the adoption of a fuzzy clustering technique to infer user segments starting from data gathered by questionnaires. The derived segments are intended to offer personalized e-Gov services tailored to user needs and their expectations. Specifically, a number of variables related to user characteristics are identified as segmentation bases, i.e., variables used in the effective process of user segment derivation. Successively, user data are gathered by asking users to fill a questionnaire. Then, a fuzzy clustering process is performed on the collected data in order to infer a number of *segments*, i.e. groups of users sharing similar characteristics. Finally, each segment is endowed with a profile that encodes the main characteristics of users belonging to each segment.

The rest of the paper is organized as follows. Section 2 describes the adopted user segmentation approach. Preliminary results obtained by applying the approach to a case study within the Italian research project SMART (Services & Meta-services for smART e-Government) are presented in section 3. Section 4 closes the paper by drawing some conclusions and outlining future work.

2. The User Segmentation Approach

Starting from user data, the aim of the approach is to infer a number of user segments, i.e., groups including users that share similar characteristics. In the adopted user segmentation approach, we may distinguish four main steps: Segmentation Variable Definition, User Data Gathering, User Segment Derivation and Segment Profile Definition (see Fig. 1). Each step is briefly described in the following.

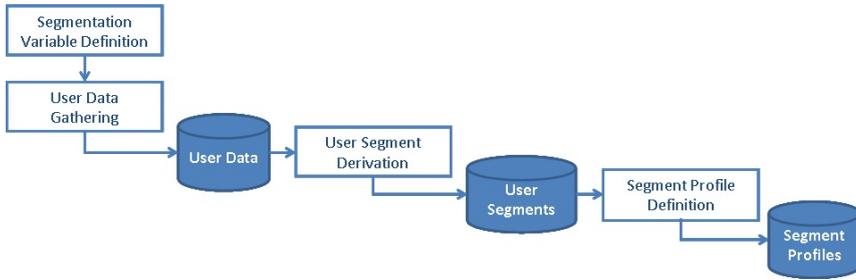


Fig. 1 - The user segmentation approach

2.1 Segmentation Variable Definition

The main goal of the first step consists in defining a set of variables, concerning user characteristics, that can be employed as a basis for segmentation, i.e., variables used in the clustering process for grouping together similar users.

A wide number of variables could be selected as segmentation bases. Among these, the most employed ones concern user features related to geographic, demographic, psychological, psychographic and behavioral characteristics [Lamb et al, 2011].

Broadly speaking, some approaches segment users by exploiting only one kind of variables, often simple to obtain, such as in geographic segmentation or demographic segmentation. On the other hand, more sophisticated approaches combine different types of variables in order to obtain better characterizations, thus allowing the inference of more significant segments. In life-style segmentation, for instance, demographic and psychological variables are combined together to characterize users.

In this work, demographic and geographic variables are selected as a basis for segmentation. More specifically, selected variables are ‘age’, ‘gender’, ‘marital status’, ‘family’, ‘children number’, ‘occupation’, and the following features of the city of residence: ‘size’, ‘chief town’, ‘altimetric zone’, ‘tourist vocation’. This choice is mainly due to the fact that, in the considered domain, demographic and geographic variables may be gathered in a simpler way than other information, and nevertheless they allow to discriminate users belonging to different segments.

2.2 User Data Gathering

The main aim of the second step is to acquire data about users according to the variables selected in the previous step. Several methods can be employed to gather user data. They may be coarsely distinguished into two main classes based on the degree of user involvement: methods that require a strong involvement of users and methods that acquire data without asking for the explicit intervention by users [Lamb et al, 2011]. Typically, methods of the first class collect user data by means of questionnaires, telephone interviews, and

focus groups. Conversely, the second class of methods implicitly gathers data from different sources that collect information exchanged by users when they interact to share opinions, perceptions, etc. For example, behavioral and psychographic data can be gathered by mining user generated content shared in blogs, social networks, etc. Other important data sources are logging systems that, by collecting information on user interactions during their running, allow to mine data about users and their behavior.

In this work, user data are gathered by means of a questionnaire devised to collect not only demographic data of users, but also perceptions about their usage experiences with a set of services on some aspects, such as time spent to request the service, time to obtain the service, cost of the service, service ability to satisfy user needs, service quality, etc.

In addition, starting from the city of residence of users, gathered by questionnaires, geographic data related to features concerning the city of residence are obtained by exploiting data of national statistics.

2.3 User Segment Derivation

The goal of this step consists in deriving user segments by grouping together similar users on the basis of the selected segmentation variables.

Cluster analysis is one of the most popular and widely applied techniques for user segmentation. Clustering techniques attempt to identify groups (clusters) of similar users by maximizing the homogeneity within each group and the differences between different groups through the application of suitable distance measures able to evaluate the (dis-)similarities existing between each pair of users. Broadly speaking, cluster analysis can be performed to obtain crisp and fuzzy partitions. In the first case, data are partitioned into disjointed clusters. Conversely, in fuzzy partitions, clusters are not well separated but they are partially overlapped. This means that a same data may belong to more than one cluster with different membership degrees.

To better capture the fuzzy nature underlying the user segment boundaries, in this work, we perform a fuzzy cluster analysis. Specifically, we apply Fuzzy C-Means (FCM), a well-known fuzzy clustering algorithm [Bezdek, 1981] to derive user segments. Such algorithm is an iterative technique that implicitly partitions the available user data into a set of k clusters by minimizing the distances between each pair of users. For each cluster, FCM determines the center that summarizes user data belonging to that cluster in the derived partition. The iterative process ends when distance values stabilize or a prefixed maximum number of iterations is reached.

Summarizing, as a result, FCM provides:

- k cluster centers that are represented as n -dimensional vectors where n represents the number of the employed segmentation variables;
- a fuzzy partition matrix where components represent the membership degrees of users to segments.

Based on clustering algorithm results, we may determine the membership of each user to several segments according to the values included in the fuzzy partition matrix.

2.4 Segment Profile Definition

The last step is aimed to describe user segments previously derived in order to provide extensive knowledge about users belonging to each segment. Such knowledge can be useful to select the target segments and to plan the most appropriate strategies that effectively address needs of the targeted users. Specifically, in this step, for each user segment we define a profile that summarizes the relevant aspects of users within the segment.

Typically, a segment profile includes information such as an identifier to unambiguously recognize the segment, a textual description to synthesize the most important characteristics of users within the segment and a collection of properties on the set of segmentation variables.

To derive profiles, the components of each cluster center are used to describe the characteristics of users belonging to the segment in terms of actual values taken by the corresponding variables.

3. A Case Study

To show the suitability of the adopted approach, we applied it to a case study that has represented an experimental experience within the SMART project. Particularly, the approach was applied in order to derive and characterize segments of entrepreneurs interested in services that are useful to open public businesses in Italy such as café and Bed & Breakfast.

To gather user data, we contacted a total number of 450 entrepreneurs from different Italian regions and, among these, 102 entrepreneurs have accomplished the questionnaire. Before segmenting entrepreneurs, we mapped values of nominal variables of gathered data into numerical values to obtain a representation of data in a more suitable form to be exploited in the clustering process. Successively, FCM was applied on the available user data to segment entrepreneurs.

We performed several runs with different k ($k = 2, \dots, 9$). To establish the most appropriate number of segments k for the gathered entrepreneur data, we employed some validity indexes (Xie-Beni index, Partition entropy and Separation index [Halkidi et al, 2001]) that are usually used in literature to this aim. In particular, the best partitions are obtained in correspondence of the smallest values for Xie-Beni index and Partition entropy and the highest values for Separation index. Fig. 2 reports the average index values obtained in the several performed runs with respect to the cluster number k . As can be observed, in correspondence of $k = 3$ we obtained the minimum value for Xie-Beni and Partition entropy indexes and the maximum value for the Separation index. Thus we chose 3 as the most appropriate number of segments to partition entrepreneurs.

Table 1 reports the component values of cluster centers obtained by FCM for the three derived segments. Each row contains the values in a cluster center in correspondence of the segmentation variables indicated in the column headers.

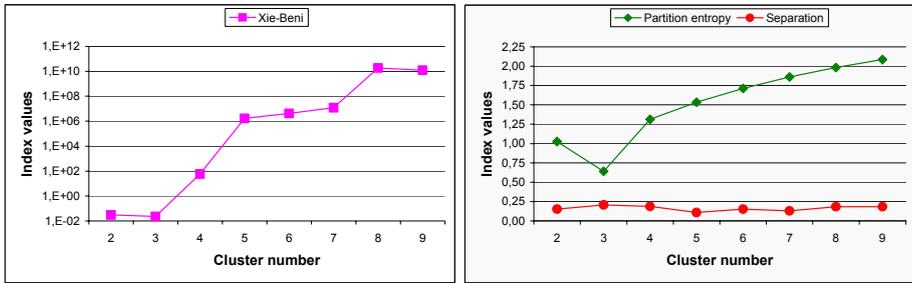


Fig. 2 - The obtained index values

Table 1 - The three cluster centers obtained by FCM

	Gender	Age	Marital status	Family	Children number	Occupation	Residence city: chief town	Residence city: alluvial zone	Residence city: littoral	Residence city: mountain	Residence city: tourist vocation	Residence city: size
S1	1.60	3.47	1.90	3.18	0.94	3.56	0.54	2.60	0.58	0.18	0.17	4.99
S2	1.54	2.99	1.76	3.48	1.02	3.12	0.31	2.37	0.38	0.24	0.16	4.15
S3	1.81	3.19	1.84	2.16	0.30	3.44	0.75	2.69	0.75	0.16	0.12	5.50

Finally, starting from cluster center components, we defined a profile for each derived segment. An intuitive interpretation of the defined profiles is as follows:

- S1:** Males aged 40-50 years, professionals. Married with an average family of 3 people. They live primarily in medium-large sized chief-towns located in plains of some tourist interest.
- S2:** People aged 36-45 years, involved in employee jobs. Living with an average family of about 3 people and one dependent. They live in small-medium sized cities that are not of tourist interest located in internal hilly area.
- S3:** Mostly male of about 36-45 years, self-employed. Conjugated with an average family with 2 members without dependents. They live in large municipalities that are not of tourist interest located in coastal hilly area.

The fuzzy partition matrix obtained by FCM allows to determine the membership of each entrepreneur to the derived segments. In particular, based on the membership degrees in partition matrix, a hard association of entrepreneurs to segments can be stated by establishing that each entrepreneur belongs to the segment with the highest membership degree. In this way we obtain a distribution of the considered entrepreneurs in the three segments that is respectively 29.9%, 27.3% and 42.9%.

Nevertheless, the main benefit of fuzzy approach to the cluster analysis consists in the possibility to determine a soft association of entrepreneurs to

segments by exploiting membership degrees of partition matrix. Precisely, based on these values, each entrepreneur does not exclusively belong to a unique segment, but we may state that she/he belongs with a strong possibility to one segment and, at the same time, with weaker possibilities to other segments. This can be particularly useful to provide more user-centered services that are able to satisfy needs characterizing several segments to which users belong with different membership degrees.

4. Conclusions and Future Works

This work has described an approach based on fuzzy cluster analysis for user segmentation in e-Gov domain. Users were segmented on the basis of demographic and geographic variables. The suitability of the approach was shown by considering data about entrepreneurs interested in opening public businesses. The obtained preliminary results encourage the application of the approach to wider scenarios involving e-Gov services.

As future work, other segmentation variables related to behaviors and attitudes of users could be fruitful to infer segments able to better capture needs and expectations of users useful to design more user-centered services.

References

- [Bezdek, 1981] Bezdek J. C., Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York, 1981.
- [Guo and Lu, 2007] Guo X., Lu J., Intelligent e-government services with personalized recommendation techniques. *Int. J. Intell. Syst.*, 22, 5, 2007, 401-417.
- [Halkidi et al, 2001] Halkidi M., Batistakis Y., Vazirgiannis M., On clustering validation techniques. *J. Intell. Inf. Syst.*, 17, 2-3, 2001, 107-145.
- [Lamb et al, 2011] Lamb C.W., Hair J.F., McDaniel C., *Essentials of Marketing*. South-Western College Publ., 2011.
- [Lu et al, 2010] Lu J., Shambour Q., Xu Y., Lin Q., Zhang, G., BizSeeker: A hybrid semantic recommendation system for personalized government-to-business e-services. *Internet Research*, 20, 3, 2010, 342-365.
- [Namvar et al, 2010] Namvar M., Gholamian M. R., KhakAbi S., A Two Phase Clustering Method for Intelligent Customer Segmentation. *International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, 2010, 215-219.
- [Punj and Stewart, 1983] Punj G., Stewart D. W., Cluster Analysis in Marketing Research: Review and Suggestions for Application, *Journal of Marketing Research*, 20, 2, 1983, 134-148.

Envisioning Smart Disclosure in the Public Administration

Giuseppe Ciaccio, Antonio Pastorino and Marina Ribaudò

DIBRIS, Università di Genova
Via Dodecaneso 35, 16146 Genova
giuseppe.ciaccio@unige.it
marina.ribaudò@unige.it
antonio.pastorino@gmail.com

Abstract. *The current generation of Open Data has to date provided only static datasets in which no data concerning specific individuals could be included, due to obvious privacy issues. Public administrations hold a great deal of personal data, as do many private entities. The lack of such personal data in the Open Data realm, and the static nature of the released datasets, are weaknesses of the current generation of Open Data. By segregating or “protecting” our personal data, those public and private entities become the “owners” of our data and hold a monopoly on services. By unleashing personal data “into the wild”, such a monopoly would collapse and a new ecosystem of services based on these data could flourish. Of course nobody wants personal data to enter the public domain without any control. We argue that an appropriate policy for online “smart disclosure” of personal data is one where the individuals are restored to their role of “data owners” and are allowed to exert online control over data accesses being performed by third parties. Based on the above matters, we propose a possible implementation of “smart disclosure” based on the OAuth 2.0 authorization framework together with a possible use case based on the tax return task.*

Keywords: Open Data, Smart Disclosure, OAuth.

1. Introduction and motivation

After the *Memorandum on Transparency and Open Government* signed by the US president Barack Obama [Obama, 2009] that marked the official onset of the so called **Open Data** movement, several public administrations started releasing massive amounts of Open Data in the form of aggregated datasets made available on their websites (see <http://data.gov> and <http://data.gov.uk> for instance). The current wave of Open Data released by

public administrations is largely made of formatted datasets of a static nature, i.e., they will not reflect changes occurring after the release date. Such datasets are roughly of two kinds, namely: aggregated and anonymized data (e.g. number of children in each school of the region); and identification data of public entities (e.g. names and addresses of restaurants in the region). No data concerning individuals have been released because of obvious privacy reasons. As already stated in a position paper of ours [Ciaccio and Ribaud, 2012], such a lack of **personal data** in the Open Data realm, along with the static nature of the released datasets, are weaknesses of the current wave of Open Data. Without personal data and without timeliness, it is indeed impossible to build useful services tailored to the actual needs of a given individual at a given time.

Many of the data managed by public administrations as well as private entities are of a personal kind. As these data are not in the Open Data domain, those public and private entities may act as the “owners” of our data. This means they hold a monopoly on services while we, the legitimate owners of the data, must abide by their terms and conditions concerning how our data are treated and used. By unleashing personal data “into the wild”, such a monopoly would collapse and a new ecosystem of personal services based on these data could flourish. In such a scenario, the various administrations holding our data are responsible for ensuring data authenticity and integrity, preventing any unauthorized access, yet allowing what is called a **smart disclosure** of personal data to the web.

A recent white paper from the UK Government [UK Government Cabinet Office, 2012] stresses the importance of smart disclosure as an enhancement of the current Open Data movement. The 'midata' initiative [UK Government, 2011] and the Smart Disclosure initiative [White House, 2012] are two programmes aimed at promoting smart disclosure of customer's personal data held by companies and providers, so as to allow people to make better choices.

It might be argued that adding personal data to the Open Data heap might jeopardize our privacy, if done in the wrong way. However, this risk is also present with the current process of releasing massive anonymized datasets. By definition, these datasets leak personal information, and information from many datasets may be jointly mined in search of individual profiles. The inferred profiles may sometimes be linked to real identities, leading to statistical de-anonymization or “identity disclosure through mosaic effect” [Hoffman et al., 2012]. The whole Open Data movement would immediately come to an end, should these confidentiality concerns prevail over the individual and social benefits of transparency and smart disclosure. A balance between privacy and transparency must clearly be sought, with the information technology playing a key role.

A proper technology for a “sufficiently smart” disclosure of data remains to be identified, along with a number of practical use cases working as an informal definition of what a smart disclosure is. In this paper we propose one use case, and we advocate the use of the OAuth 2.0 authorization framework [OAuth Working Group, 2010] to achieve smart disclosure. On the basis of this approach, individuals are restored to their role of **resource owners** while

administrations (public or private ones) are stripped of their de-facto ownership of personal data and keep a role of bare **resource managers**. A resource owner (namely, an individual) may grant online authorization to any third party application to use a given item of personal data located on a given resource manager, in exchange for a useful personalized online service that the application is expected to provide using that data item. The process of granting authorization is based on unforgeable cryptographic tokens released by a trusted **authorization server** with which individuals, applications, and resource managers, are all registered.

We cite here the related work by Falcão-Reis and Correia [Falcão-Reis and Correia, 2010] who suggest to couple Electronic Health Records (EHRs) with an extended version of OpenID [Recordon and Reed, 2006] in an effort to implement a user-controlled system of Health Digital Identity for Portuguese citizens. They also propose leveraging OAuth 2.0 as an authorization technology for user-controlled access to EHRs, thus anticipating smart disclosure in the medical care field.

2. OAuth 2.0

OAuth 2.0 [OAuth Working Group, 2010] [IETF 2012b] [Hammer-Lahav, 2010] originated from a complete redesign of the previous OAuth 1.0 specification. OAuth 2.0 is too high-level to be defined as a protocol specification. It should be considered as a blueprint of a protocol, within which many implementations are feasible, although possibly not interoperable with one another. The interest around this technology is huge: the IETF OAuth working group includes members like Google, Facebook, Microsoft, Twitter, Deutsche Telekom, and Mozilla [Hammer-Lahav, 2010], and OAuth 2.0 has already been adopted by Google [Google Inc., 2012], and Facebook [Facebook, 2012], just to cite a few.

OAuth stems from a typical use case with online social networks. Suppose a third party client application is offering a user a personalized service making use of that user's personal features provided by a social network (e.g. pictures, contact list, posting a comment) via a RESTful API. The naive approach of requiring users to release their credentials (username and password) to the third party client so that the latter could get those features from the social network is highly risky. A more appropriate solution is to release cryptographic **proof of authorization**, issued by the user (at least in principle) to the third party application, and to let the latter subsequently spend such authorization proof at the social network API in order to get the required features.

Abstractly, OAuth 2.0 identifies four actors exchanging information in an ordered way (Figure 1). These actors are the **client application** or just “client”, the **resource owner** or “owner” or “user” for short, an **authorization server** (AS, the trusted entity), and one or more **resource servers** (RS) hosting data or services to be smartly disclosed. The AS and the various RSs may each belong to a distinct administration domain.

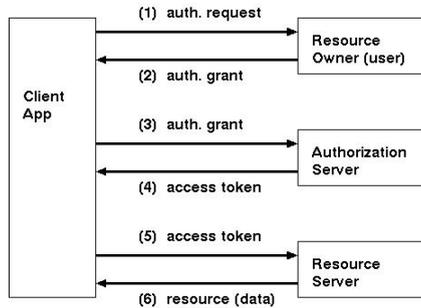


Fig. 1 - OAuth 2.0 abstract protocol, with the four actors (client application, resource owner, authorization server, and resource server) and their interplay. Actions are ordered by increasing number.

The process is started by the client requesting authorization to access a given **resource** (e.g. a set of personal data), subject to a set of constraints called **scope** (e.g. read vs. write access to specific fields of personal data during a specified time window). The resource owner is shown the request and, if they agree with the scope, they yield an **authorization grant** bound to the resource and scope. Such a grant is then exhibited by the client to the AS, which validates it and returns an **access token** valid for the required resource at a specific RS subject to the scope. The client finally passes the access token to the given RS along with the resource request; upon token validation and scope verification, the RS returns the resource to the client.

In practice, the authorization grant can be obtained in four possible ways. These four ways are called **flows** in OAuth jargon, and are nothing but instantiations of the abstract protocol scheme discussed above. Due to a lack of space we herein only describe the first and most important of the flows (the others are described in [IETF, 2012b] and [Ciaccio et al., 2013]). In such flow (Figure 2) the grant is obtained indirectly, with the AS acting as an intermediary under the control of the resource owner. With this flow the client is typically a web application on a remote server, and the resource owner is typically a human with a browser initially pointed to the web application. The grant is represented by an **authorization code** issued by the AS after obtaining consent from the resource owner, and is delivered to the client through the browser of the resource owner via HTTP redirection.

In the current IETF draft, TLS protection is not required when the AS redirects the browser to the client after the authorization step; in other words, it is legal for the client redirection URI (Fig. 2) to be an HTTP endpoint of the web application instead of an HTTPS one. This shortcoming is meant to cover client applications that are unable or unwilling to provide TLS endpoints due to lack of resources, for instance. The downside is that an authorization code sent to a non-TLS endpoint is transmitted in plain text and could therefore easily be eavesdropped. The stolen authorization code could then be used by an attacker

client to obtain an access token from the AS. But if the client application has a persistent identity registered at the AS (a “confidential” client, in the OAuth jargon), then the AS will prompt the client application to authenticate itself before converting the authorization code into an access token, thereby preventing the use of stolen authorization codes by rogue clients. As an additional security measure, the access token itself may be bound to the specific client identity (a so called “proof token” [IETF, 2012a], as opposed to an anonymous “bearer token”). In contrast, if the client does not hold a secure and persistent secret registered with the AS (a so called “public” client) then the flow is insecure, unless the client redirection URI is an HTTPS endpoint. A security analysis of OAuth 2.0 can be found in [IETF, 2012a].

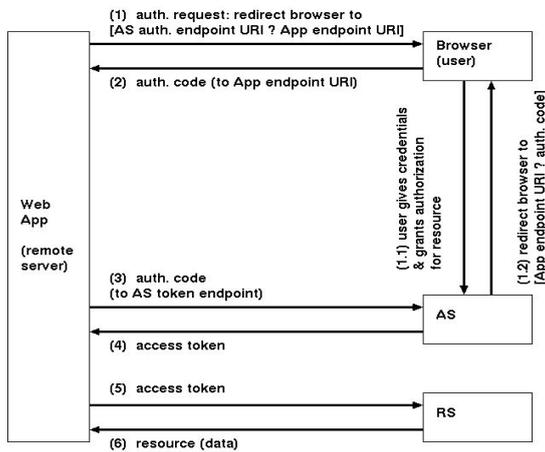


Fig. 2 - OAuth 2.0 Authorization Code flow, for web applications. The “?” denotes an HTTP GET parameter

From the above it is reasonable to deduce that resource owners, clients, and RSs should establish a relation with the AS before engaging in any OAuth 2.0 protocols. After registration, the client is given a unique identifier that is valid at the AS.

3. A use case: the annual tax return

Scenario. Bob needs to fill in his annual tax return. As usual, he visits his business accountant bringing lots of pieces of paper: his annual salary, medical expenses, documents concerning his properties, other expenses he can deduct from his income, and so on. The other possibility is to use a web application provided by the government and fill in an online form by manually copying all the data printed into the various paper documents.

How OAuth 2.0 could help. Figure 3 shows Bob in front of an innovative online service. He still has access to a web application to fill in the online form for his annual tax payment. This year, however, most of Bob's income information comes directly from various remote databases where it is scattered, upon a simple online authorization by Bob himself. Filling in the form is simpler, quicker and less prone to mistakes. Bob is the owner of the data. He applies for the authorization grant (arrows 1 and 2) so that the web application he is connected to can directly access his restricted access resources stored in different databases (the Employer Registry, the Estate Registry, the Medical Expenses Registry shown in the figure). He then fills in the form adding only data that could not be obtained via online third parties. The discussion is very summarized here for space constraints but, as we shall discuss in Section 5, there are indeed some open challenges even in such a simple use case.

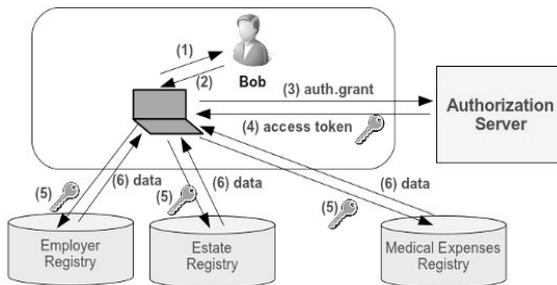


Fig. 3 - Tax payment with OAuth 2.0

4. OAuthwo: a free OAuth 2.0 prototype server

For our experiments we have implemented an OAuth 2.0 authorization server, that we have called OAuthwo. In this section we describe the main features of OAuthwo, especially those ones regarding how it handles multiple resource servers with possibly multiple credentials of the same user. OAuthwo is free software written in PHP as a Zend module and available for download at GitHub (github.com/andou/oauthwo_zend/).

In OAuthwo the authorization server AS keeps information about resource owners, clients and resource servers.

Information about clients. Information provided by a client C to the AS during the registration phase and retained by the AS are:

- *C_{id}*, a unique ID for C
- *C_{secret}*, a shared secret between AS and C, present if C is a confidential client and used for client authentication
- *C_{profile}*, the profile of C, with possible values *web*, *user-agent* or *native* ([IETF, 2012b], Section 2.1)

- *CredirectionURI*, the client redirection URI as specified in Section 2
- *Cname*, a human-readable name for C.

In OAuth2, public clients as well as confidential clients are required to register exactly one complete redirection URI to the AS, to prevent the "Authorization Code Redirection URI Manipulation" and mitigate the "Client Impersonation" and the "Open Redirectors" problems described in [IETF, 2012a].

Information about resource servers. Information for a resource server RS retained by the AS are:

- *RSID*, a unique ID for RS
- *RSsecret*, a secret key shared between AS and RS, to be used for encrypting access tokens
- *RSid_method*, the way by which RS identifies users within its realm, e.g. by email, social security number, phone number or others; see paragraph below
- *RSEndpointURI*, the URI from which RS's data could be accessed.

Furthermore, AS retains information about allowed scopes on each RS. All these information are gathered during trust establishment between each resource server and the authorization server.

Information about resource owners. Each RS employs one out of several available user identification methods (e.g. email, social security number, phone number, etc.). Thus the same user can have different IDs at different RSs and even another different ID at the AS. AS must therefore build and maintain an **identity equivalence class** for each resource owner. This is necessary because, for instance, when AS issues an access token T to be used at an RS to retrieve information on the resource owner U, T must contain an identification key that is valid at the RS, so that RS could properly identify the user U.

In order to build the identity equivalence class, at the registration phase the user is prompted with a list of RSs already known by the AS and is asked for an identification key for each of these. The AS should verify that the given identifiers do actually belong to the user. Some of the identifiers are easily verifiable online (e.g. email, social network nickname), others may require a third party digital certification (e.g. social security number).

Access tokens. An access token represents the authorization originally given by a resource owner for accessing a certain set of resources, subject to given constraints (scope and lifetime).

The OAuth 2.0 specification does not specify the format and structure of access tokens. Yet, two main categories can be distinguished. The access token may be an opaque identifier, used by the RS to retrieve the authorization information at the AS; or it may self-contain such information in a verifiable form. The latter kind of token requires RS to be able to interpret cryptographic

messages and implies the use of short-living access tokens due to impossibility to revoke a token once issued. As the access information is self-contained, this solution requires less communication and is therefore more scalable.

OAuthwo uses self-contained bearer tokens [IETF, 2012c]. When a client seeks authorization for protected resources stored in multiple resource servers RS_1, \dots, RS_n , the OAuthwo AS issues a composite access token T made of several chunks T_1, \dots, T_n . The generic chunk T_i is to be spent at resource server RS_i .

Access tokens are served to clients through the access token parameter in the successful response (see [IETF, 2012b], section 5.1). In OAuthwo this parameter is a base 64 encoding of a JSON structure [IETF, 2006] which serializes an associative array where the keys are the URIs of resource servers RS_1, \dots, RS_n and the values are the chunks T_1, \dots, T_n . Each T_i consists of a triple $\langle U_i, S_i, P_i \rangle$. S_i , the scope, instructs the client on what actions are allowed on what data at resource server RS_i , using U_i as endpoint and P_i as access token for that particular resource server. P_i is thus a monolithic self-contained access token for a single resource server. Being self-contained, P_i carries all the information needed by the resource server RS_i in order to provide the requested resources, namely:

- $P_{(i,user_reference)}$, a unique identifier of the user recognized at RS_i
- $P_{(i,scopes)}$, a space separated list of the scopes allowed by the token
- $P_{(i,validity)}$, the token lifetime.

These information are serialized as claims in a JSON Web Token [OAuth Working Group, 2012] digitally signed by the AS with a JSON Web Signature [JOSE Working Group, 2012b].

The user identifier $P_{(i,user_reference)}$ valid at RS_i is obtained by the AS using the identity equivalence class (see above). Such identifier must be considered as a sensitive information; to prevent disclosing it to the client [IETF, 2012c], each serialized and signed P_i is finally encrypted with the secret key shared between the AS and RS_i , using JSON Web Encryption [JOSE Working Group, 2012a].

5. Conclusions

We have presented OAuthwo, a server side implementation of the OAuth 2.0 framework, discussing how this framework could be used to enhance the Open Data realm with personal data, but without making these data public. Currently, a proof of concept in the annual tax return domain (Section 3) has been developed as a master thesis. The prototype uses OAuthwo to build a web application (<https://dione.disi.unige.it>) supporting a simplified yet representative example of tax return, namely, the italian *modello 730* (devoid of its graphical skin for simplicity) in the case of a single income plus medical expenses and the property of a single residential apartment, along with anagraphic personal data.

Even in such a simple use case there are some requirements that put OAuth 2.0 at challenge. We could in fact identify entities that need to access data without being the owners, for instance the Judiciary or the Inland Revenue Office in case of lawsuits related to tax evasion. In a medical scenario like the one of [Falcão-Reis and Correia, 2010] the owner might not be able to give any authorization, perhaps because unconsciousness during a rescue. Also in this case there must be another entity which, without being the data owner, has enough privileges to grant access rights. This requirement seems not explicitly covered by OAuth 2.0. We are investigating about the use of a recent evolution of OAuth 2.0 called User-Managed Access (UMA) [IETF, 2013].

Another use case that requires investigation is the release of aggregated data which is desirable for statistical purposes but is not covered by the current OAuth 2.0 flows which indeed grant access to specific records of an individual user.

Last but not least, in the annual tax return scenario there is also another subtle point. The same resource, for instance a receipt of payment, involves two people. For example the invoice of a payment involves at least two entities, namely, the issuer and the receiver. To the former the invoice represents income, while for the latter it is an expense; there is only one resource but it has two distinct resource owners. How to deal with such a kind of multi-owned resources is not explicitly specified by OAuth 2.0, especially concerning how to involve the various owners when only one of them issues an authorization for creating or modifying the resource.

References

[Ciaccio et al., 2013] Ciaccio, G., Pastorino, A. and Ribaldo, M., Open Data and Personal Information: A Smart Disclosure Approach based on OAuth 2.0 (extended version). Tech. rep. DIBRIS-TR-13-01, DIBRIS, Università di Genova, Italy, 2013.

[Ciaccio and Ribaldo, 2012] Ciaccio, G. and Ribaldo, M., Open Data for the Masses: Unleashing Personal Data into the Wild, in Proc. of 8th Int.l Conf. on Web Information Systems and Technologies (WEBIST 2012), SciTePress.

[Facebook, 2012] Facebook login -- Facebook Developers, [online], <http://developers.facebook.com/docs/concepts/login/>, 2012

[Falcão-Reis and Correia, 2010] Falcão-Reis, F. and Correia, M. E., Patient Empowerment by the Means of Citizen-managed Electronic Health Records. *Medical and Care Compunetics* 6, 2010, pp.214-228, IOSPress.

[Google Inc., 2012] Google Inc., *Using OAuth 2.0 to Access Google APIs*, [online], <http://developers.google.com/accounts/docs/OAuth2>, 2012

[Hammer-Lahav, 2010] Hammer-Lahav, E., Introducing OAuth 2.0, [online], <https://hueniverse.com/2010/05/introducing-oauth-2-0/>, 2010

[Hoffman et al., 2012] Hoffman, W. et al., Rethinking Personal Data: Strengthening Trust, [online], World Economic Forum, <http://www.weforum.org/reports/rethinking-personal-data-strengthening-trust>, 2012

[Howard, 2012] Howard, A., Data for the Public Good, in Strata: Making Data Work, O'Reilly, 2012

[IETF, 2006] IETF, The application/json Media Type for JavaScript Object Notation (JSON), [online], <http://www.ietf.org/rfc/rfc4627>, 2006

[IETF, 2012a] IETF, OAuth 2.0 Threat Model and Security Considerations, [online], <http://tools.ietf.org/html/draft-ietf-oauth-v2-threatmodel-07>, 2012

[IETF, 2012b] IETF, The OAuth 2.0 Authorization Framework, [online], <http://tools.ietf.org/html/draft-ietf-oauth-v2-31>, 2012

[IETF, 2012c] IETF, The OAuth 2.0 Authorization Framework: Bearer Token Usage, [online], <http://tools.ietf.org/html/draft-ietf-oauth-v2-bearer-23>, 2012

[IETF, 2013] T. Hardjono (ed.), User-Managed Access (UMA) Profile of OAuth 2.0, [online], <http://tools.ietf.org/html/draft-hardjono-oauth-umacore-07>, 2013

[JOSE Working Group, 2012a] JOSE Working Group, JSON Web Encryption (JWE), [online], <http://tools.ietf.org/html/draft-ietf-jose-json-web-encryption-05>, 2012

[JOSE Working Group, 2012b] JOSE Working Group, JSON Web Signature (JWS), [online], <http://tools.ietf.org/html/draft-ietf-jose-json-web-signature-05>, 2012

[OAuth Working Group, 2010] OAuth Working Group, OAuth 2.0, [online], <http://oauth.net/2/>, 2010

[OAuth Working Group, 2012] OAuth Working Group, JSON Web Token (JWT), [online], <http://tools.ietf.org/html/draft-ietf-oauth-json-web-token-06>, 2012

[Obama, 2009] Obama, B., Transparency and Open Government, [online], http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment, 2009

[Recordon and Reed, 2006] Recordon, D. and Reed, D., OpenID 2.0: a platform for user-centric identity management, in Proc. of the 2nd ACM workshop on Digital identity management (DIM '06), ACM, 2006.

[UK Government, 2011] The midata vision of consumer empowerment, [online], <https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment>, 2011.

[UK Government Cabinet Office, 2012] UK Government Cabinet Office, Open Data White Paper: Unleashing the Potential, [online], <http://www.cabinetoffice.gov.uk/resource-library/open-data-white-paper-unleashing-potential>, 2012.

[White House, 2012] Informing Consumers through Smart Disclosure, [online], <http://www.whitehouse.gov/blog/2012/03/30/informing-consumers-through-smart-disclosure>

Organization and Governance of IT Outsourcing processes

Aurelio Tommasetti, Giovanni Vaia¹,
Università degli Studi di Salerno
Via Ponte don Melillo, 84084 Fisciano (SA) - Italy
tommasetti@unisa.it

¹Ca' Foscari Università di Venezia
San Giobbe - Cannaregio, 873, 30121 VENEZIA (IT)
giovanni.vaia@unive.it

Abstract. *The aim of this research is to provide an empirical support to define how to organize an IT outsourcing relationship in terms of the allocation of decision-making power, structure of roles, mechanisms of integration and coordination between customer and provider. The paper attempts to define a framework to analyze from different perspectives the complexity of outsourcing relations. Results show the significance of two elements setting up the IT outsourcing engagement: the first is related to the characteristics of the technology in terms of products, infrastructure and interfaces; the second is related to the object of the transaction. This approach has allowed to classify four possible scenarios based on the characteristics of the relationship, the organizational characteristics and the characteristics of governance tools.*

Keywords: IT outsourcing, Governance, customer-provider relationships.

1. Introduzione

Nel settore IT sono presenti due forti tendenze: l'abbandono di soluzioni tecnologiche legacy e la scelta di prodotti standard (come Oracle e SAP); l'implementazione di processi standard per la gestione dei servizi IT (ITIL, CobiT) e metriche standard per il monitoraggio dei servizi. Questo permette di trasferire con un linguaggio condiviso e comprensibile i dettagli delle procedure operative definite nei processi di gestione, le policy, ruoli e responsabilità, le periodicità del reporting, i path di escalation funzionale relativi ai service provider [Moretti, Vaia e Zirpoli, 2013].

L'adozione di standard organizzativi di settore però pone una serie di interrogativi relativamente alla possibilità di "modularizzare" i servizi IT e di gestirli totalmente o parzialmente in outsourcing. Affidare un servizio all'esterno significa tipicamente coinvolgere specialisti, proprietari delle conoscenze, per

Congresso Nazionale AICA 2013

erogare il servizio senza considerare le interdipendenze esistenti, in termini di conoscenze, con gli altri attori della catena. Spesso al task partitioning non corrisponde un efficace knowledge partitioning. D'altra parte, l'intangibilità dei servizi intensifica il conflitto tra clienti e fornitori rappresentando un ostacolo all'innovazione. Questo conflitto dipende principalmente dalle caratteristiche della tecnologia che si trasforma da artefatto in servizi. E se risulta semplice conoscere e predeterminare i livelli di servizio di un singolo componente IT, la definizione dei livelli di servizio di "sistema" è una questione non banale.

La ricerca affronta due problemi principali legati alla crescente complessità della tecnologia e al declino dello standard tecnologico come modalità di coordinamento: il primo legato alla governance delle relazioni cliente-fornitore nel settore IT; il secondo è legato all'utilizzo di architetture standard di settore.

Il punto cruciale diventa governare la complessità e dinamicità dei processi di outsourcing attraverso attività di codificazione e standardizzazione di processi e unità di misura, definizione delle interfacce di comunicazione, allocazione del potere decisionale, definizione dei meccanismi di sanzione e ricompensa.

In questo senso pochi studi analizzano in maniera esplicita le condizioni che favoriscono il crearsi di un rapporto cliente – fornitore che porti ad un outsourcing di successo [Vaia, Zirpoli, 2011].

Il lavoro cerca di definire uno schema di lettura per analizzare da più punti di vista la complessità di un processo di IT outsourcing da una prospettiva organizzativa e relazionale.

2. Problema di ricerca

L'approccio tradizionale all'organizzazione delle attività IT si è principalmente focalizzato sulle "strutture di governance" per la gestione dell'infrastruttura e la gestione e l'uso delle applicazioni.

In pratica si fa riferimento al concetto di IT governance come la distribuzione dei diritti di decisione e responsabilità tra gli stakeholders, e le procedure e i meccanismi per prendere e monitorare le decisioni strategiche riguardo all'IT [Sambamurthy, Zmud, 2000].

Se l'obiettivo principale dell'IT management è quello di gestire in modo efficiente ed efficace le IT operations, servizi e prodotti, l'IT governance contribuisce contemporaneamente alle attuali business operations e a posizionare la funzione IT per soddisfare i futuri bisogni di business. Questo serve ad indicare che l'IT governance è sia orientata all'interno che all'esterno, tra presente e futuro.

Il termine governance è stato definito, in modo ampio nella letteratura di matrice organizzativa, come la "modalità di organizzare" le transazioni [Williamson, 1979]. Governance è un concetto che va al di là del concetto di controllo in senso stretto ed include elementi di inizio e strutturazione delle relazioni di scambio così come aspetti di monitoraggio e rinforzo [Heide, 1994].

La governance, in questo lavoro, viene definita come l'insieme dei meccanismi finalizzati all'orientamento dell'azione di coordinamento inter-organizzativo date le caratteristiche della relazione.

In riferimento agli studi sui canali distributivi due paradigmi hanno inizialmente caratterizzato le ricerche: microeconomico e comportamentale.

Il paradigma microeconomico attribuisce ai canali un ruolo funzionale agli obiettivi di minimizzazione dei costi e le decisioni di esternalizzazione sono legate alla possibilità di ottenere benefici di scala grazie alla specializzazione di canale. In pratica le decisioni di governance sono strettamente legate a decisioni di internalizzazione o di esternalizzazione. In modo parallelo l'Economia dei Costi di Transazione [Williamson, 1979] usa lo stesso approccio ma con un focus sui costi diverso.

Il paradigma microeconomico è stato fortemente criticato in quanto analizza il canale come un "bundle of function" e non include l'analisi delle relazioni che esistono alla base delle interdipendenze funzionali di canale. In questo senso il paradigma comportamentista focalizza la propria analisi sul disegno dei meccanismi di controllo delle performance degli attori del canale. In questo senso la governance è un problema di creazione e gestione del potere, finalizzato a coordinare gli sforzi del canale distributivo.

I due paradigmi quindi sottopongono all'analisi due aspetti fondamentali: il controllo delle risorse, attraverso l'internalizzazione, e l'efficienza economica, attraverso la specializzazione delle attività nel canale.

Come accennato, la governance viene definita come l'insieme dei meccanismi finalizzati all'orientamento dell'azione di coordinamento inter-organizzativo date le caratteristiche della relazione. In questa definizione si evidenziano i concetti di orientamento, coordinamento e caratteristiche della relazione. L'approccio usato nello studio prevede, infatti, un inquadramento non meccanicistico del nesso tra tool utilizzati per la gestione della relazione interorganizzativa ed output della relazione, consci della natura socio-tecnica delle relazioni inter-impresa e del ruolo degli attori organizzativi nell'esecuzione dei processi organizzativi. Da qui il concetto di "orientamento".

Quanto ai tool stessi di governance, al fine di una maggiore schematicità analitica si è praticata una distinzione tra "caratteristiche della relazione" e "meccanismi di coordinamento".

Per quanto riguarda le caratteristiche della relazione risulta difficile individuare un insieme di condizioni per definirle esattamente. Sulla base delle teorie analizzate è possibile porre l'attenzione su alcune di esse per individuare analogie e diversità.

L'idea di fondo è che le relazioni, nel breve o medio periodo, siano caratterizzabili sulla base di alcune dimensioni quali, ad esempio, l'asset specificity (human, site, asset, physical), l'incertezza, la frequenza, le caratteristiche del prodotto (standard, complesso, modulare, etc.), la routinizzazione vs. la flessibilità (frequenza, variazione, nesso con i meccanismi contrattuali) [Becker, 2004], l'esistenza di mutua dipendenza, conflittualità o fiducia.

Tali dimensioni nello schema analitico utilizzato rappresentano caratteristiche non modificabili dagli attori nel breve periodo e, quindi, di

conseguenza condizionano la relazione a prescindere dalle intenzioni strategiche degli attori stessi.

Diversamente da quanto accade per le caratteristiche della relazione, gli attori possono, almeno in misura maggiore, scegliere e modulare l'utilizzo di diversi meccanismi di coordinamento. Tra questi i meccanismi strutturali come il livello dello scambio informativo (direzionalità, frequenza, qualità, tipologia di informazione, utilità, etc.) [Dyer, Singh, 1998], il tipo di contratto (durata, format, formalizzazione, ricorso a terze parti, uso di meccanismi di profit sharing) [Bidault et al., 1998], monitoraggio delle performance [Poppo, Zenger, 1998], e gli strumenti di knowledge integration (shared meanings, shared norms).

Il modello utilizzato, quindi, inquadra la governance in stretta relazione con le scelte di make or buy (divisione del lavoro inter-impresa), ed il livello di diffusione della tecnologia e come caratterizzata sia dai meccanismi di coordinamento selezionati sia dalle caratteristiche della relazione.

Lo scopo della ricerca è quello di fornire un supporto empirico per definire, in base ad un efficace knowledge-task partitioning, le modalità di organizzazione di una relazione di IT outsourcing in termini di allocazione del potere decisionale, scomposizione dei ruoli, meccanismi di integrazione e coordinamento tra le parti.

Il lavoro cerca di definire uno schema di lettura per analizzare da più punti di vista la complessità di un processo di IT outsourcing da una prospettiva organizzativa e relazionale (vedi Fig. 1). L'obiettivo è quello di offrire ai practitioner uno strumento di analisi e lettura del processo di outsourcing. D'altra parte il tentativo è quello di coprire un gap nella letteratura IT riguardo l'organizzazione e gestione delle relazioni di fornitura.

Seguendo lo schema di definizione del problema di ricerca le domande principali alle quali questa ricerca tenta di dare una spiegazione sono:

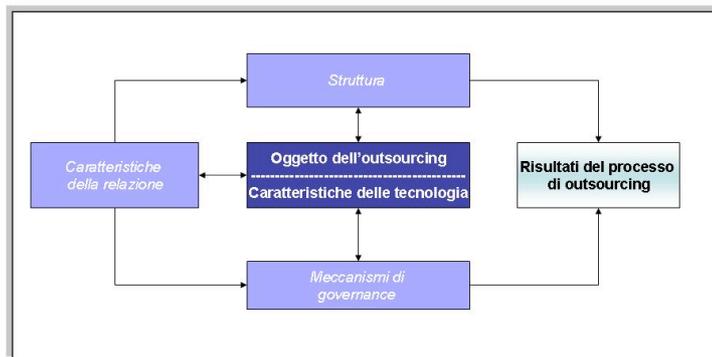
RQ1: In che modo le aziende dovrebbero organizzare le loro attività IT in modo da perseguire gli obiettivi di business?

RQ1.a Qual è il rapporto tra divisione della conoscenza e divisione dei compiti nella catena di fornitura?

RQ1.b: Qual è l'impatto delle architetture standard di settore nella gestione dei servizi IT?

RQ2: Quali sono le caratteristiche delle relazioni di fornitura nel caso di un rapporto cliente-fornitore che abbia ad oggetto l'acquisto di un servizio IT?

RQ2.a Quali sono i meccanismi di coordinamento utilizzati per gestire una relazione di IT outsourcing?



3. Metodo

La survey, composta da un questionario di 13 pagine con domande chiuse in scala di Likert, è stata distribuita durante una conferenza annuale italiana del settore dei servizi IT: L'economia dei servizi. Motore per vincere la sfida dei nuovi mercati "Investiamo sul nostro futuro". La conferenza si è tenuta presso Hotel Melià a Milano il 30 novembre 2006 e ha visto la partecipazione di 500 practitioner del settore. Il questionario è stato incluso nel materiale della conferenza e restituito al desk informazioni. Il tasso di risposta è stato del 10 %.

Nel lavoro di ricerca è stata privilegiata l'analisi multidimensionale. L'interesse per questo tipo di analisi è stato determinato dalla possibilità di studiare simultaneamente sia caratteri quantitativi sia caratteri qualitativi attraverso un questionario (poi raggruppate in classi come variabili qualitative ordinali). Il questionario somministrato è stato sottoposto con una finalità esplorativa, nella consapevolezza che nessuna analisi esplorativa è svincolata dall'obiettivo finale, per studiare il "possibile" contributo di circa 236 variabili alla gestione dei processi di IT outsourcing.

Questo approccio, nel produrre delle dimensioni (fattori) attraverso le quali semplificare, sintetizzare e rappresentare il fenomeno, permette l'esplorazione di un fenomeno e la produzione di elementi di conoscenza importanti per la formulazione di nuovi modelli o per la verifica di teorie. Le tecniche multidimensionali diventano utili quando è necessario analizzare ex novo vasti insiemi di informazioni di base, al fine di ricavare – senza forzose ipotesi a priori – delle sintesi preliminari, utili come base per ulteriori trattamenti e/o per porre le basi per disegnare modelli causali o processi d'analisi confermativa (superando il livello esplorativo dell'analisi) [Franco, 2001].

Il passaggio da una matrice di dati raccolti da un questionario ad una matrice adatta all'applicazione dell'ACM ha richiesto un processo di trasformazione dei caratteri quantitativi in variabili qualitative, suddividendolo in classi non vuote e la ricodifica delle modalità a bassissima frequenza in classi più ampie, compatibilmente con alcune proprietà statistiche (equivalenza distributiva, Lebart) e un processo di ricodifica delle variabili qualitative (codifica disgiuntiva completa). Infatti, il campo d'applicazione per eccellenza dell'ACM, indicato in letteratura è l'analisi di dati provenienti da questionari (anche se viene applicata sempre più spesso anche a dati provenienti da archivi amministrativi).

Tale tecnica ha permesso di produrre, come output, una serie di assi fattoriali su cui è possibile proiettare le singole modalità delle variabili per studiarne l'associazione. Dallo studio delle prossimità tra le modalità sul grafico è stato possibile individuarne immediatamente l'associazione e, tramite l'osservazione dei cosiddetti "contributi" delle modalità è stato possibile "caratterizzare" gli assi.

Questo ha permesso di studiare prima singolarmente le quattro sezioni del questionario (Informazioni Generali, Caratteristiche della Fornitura e della

Tecnologia, Caratteristiche Organizzative e Meccanismi di Coordinamento, le Relazioni Cliente-Fornitore) dove sono state individuate 43 variabili più rappresentative del fenomeno oggetto di studio, che incrociate tra di loro hanno permesso di spiegare il contributo delle diverse variabili organizzative e relazionali al successo dei processi di IT outsourcing, confermando le ipotesi iniziali.

4. Analisi dei risultati

Sui dati rilevati, come accennato, è stata applicata l'analisi delle corrispondenze multiple (Cor.Mu.) attraverso il software SPAD. Sui dati è stata fatta un primo lavoro di analisi parziale su ogni sezione del questionario per studiare le relazioni tra le variabili al fine di escludere quelle più ridondanti. Questa fase ha permesso di selezionare 43 variabili che sono state incrociate tra di loro. In SPAD il metodo utilizzato è stato il Corem, che non è nient'altro che l'analisi delle corrispondenze multiple dove le mancate risposte (M.R) sono state considerate illustrative, non contribuendo alla spiegazione del fenomeno.

L'output complessivo (vedi Fig. 2) ha mostrato una forte associazione tra le modalità considerate. Questo ha spinto a studiare le quattro sezioni in modo analitico.

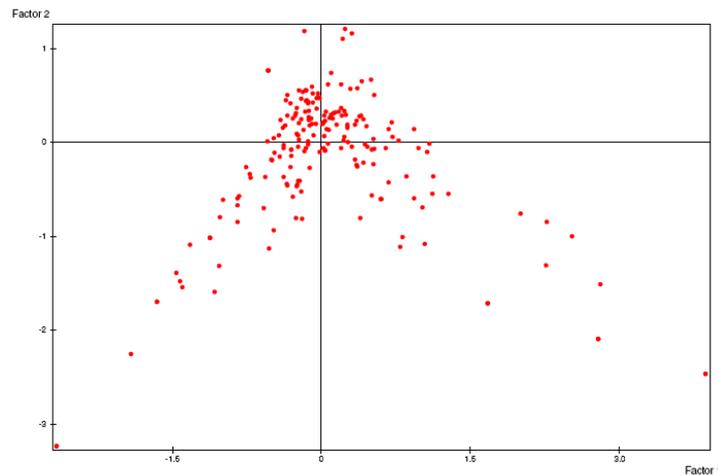


Fig.2 – Output SPAD complessivo

Dall'analisi del primo quadrante si evidenzia che nel caso di un basso utilizzo di regole e procedure standard e prodotti tecnologici con bassa standardizzazione vi è un mancato raggiungimento degli obiettivi in termini di innovazioni introdotte in azienda. L'oggetto della fornitura sono i servizi strategici. In questo caso i clienti tendono a non mantenere in house le attività di sviluppo software. Il cliente non è in grado di gestire le attività di diagnosi in house e quelle di change. La relazione è basata su contratti lunghi, dove non è possibile individuare le responsabilità del partner attraverso misure di

performance. In questo contesto sono presenti comportamenti opportunistici nella relazione.

D'altra parte si nota una forte associazione tra una tecnologia che presenta interfacce standard con una struttura organizzativa coordinata attraverso regole e procedure con la presenza di un'unità specifica che coordina il processo di scomposizione e assegnazione delle attività tra i partner. In questa situazione le persone hanno tutte le conoscenze per svolgere la propria attività. Il meccanismo di governance più utilizzato è il contratto. La relazione si delinea come collaborativa anche se vi è difficoltà a trasferire le informazioni. Cioè la collaborazione non si spinge mai verso una reale integrazione. Per quanto riguarda le attività, quelle strategiche sono mantenute in house. A questo tipo di configurazione è associato un livello alto di raggiungimento degli obiettivi definiti nel rapporto. Inoltre in questa situazione le aziende si dichiarano pronte a implementare lo standard di processo ITIL.

Nel secondo quadrante vi è una forte associazione tra le caratteristiche standard dei prodotti (i prodotti e servizi esternalizzati sono "non core") e standard di processo come ITIL, condivisi in tutta la catena di fornitura. Lo standard di processo permette di definire misure di performance standard utili per monitorare puntualmente le performance dei fornitori (dove sono tipicamente pagate penali a fronte del monitoraggio). I clienti tipicamente disegnano le interfacce dei sistemi e mantengono in house le attività di analisi dei problemi e disegno delle soluzioni, mentre sono esternalizzate altre attività di supporto strategico e decisionale. Questo implica un continuo rapporto di scambio informale e di condivisione, caratterizzato da fiducia reciproca e assenza di comportamenti opportunistici (la fiducia reciproca risulta essere un forte fattore di lock in).

Nel III quadrante c'è una forte associazione tra prodotti e infrastruttura personalizzata. In questa situazione le attività di problem management (diagnosi) e change sono mantenute in house, dove il change ha piena autorità decisionale sui cambiamenti all'infrastruttura che permette di introdurre innovazioni nei processi di business. In questo caso le persone hanno tutte le conoscenze per svolgere l'attività e adottano lo stesso standard di processo della catena. Disegnano l'architettura ma non le interfacce tra i sistemi, che sono standard. La governance è praticamente interna, non esiste una unità specifica per il task partitioning che viene basato sull'architettura organizzativa di ITIL. Anche in questo caso le misure di performance hanno un impatto positivo sulla fiducia.

Il quadrante IV è caratterizzato da un'associazione tra caratteristiche specifiche della tecnologia e oggetto della fornitura: servizi strategici esternalizzati. Questo caso si associa a ad un alto potenziale di opportunismo e ad una relazione poco collaborativa. Il cliente mantiene il governo della gestione economico finanziarie e in generale dei cambiamenti tecnologici (change management). Non vi è una unità specifica che permette l'allocazione dei task o uno standard di processo come ITIL, le attività sono regolate da processi informali. Le informazioni circolano con difficoltà e non sono usati data base condivisi. La conoscenza e condivisione reciproca dei processi organizzativi, investimenti specifici non ICT sono elementi di difficoltà per lo switching.

5. Discussione e conclusioni

I risultati dell'analisi hanno evidenziato la centralità di due elementi determinanti per l'impostazione dei processi di IT outsourcing: il primo legato alle caratteristiche della tecnologia nei termini di prodotti, infrastruttura e interfacce; il secondo all'oggetto dell'outsourcing ossia alle finalità legate all'IT dell'azienda. Questo approccio ha permesso di arrivare ad una possibile classificazione di quattro scenari rispetto ai quali è possibile individuare le caratteristiche della relazione, le caratteristiche organizzative in termini di allocazione e governo delle attività, le caratteristiche dei tool di governance (vedi Fig. 3).



Fig.3 – Organizzazione e Governance dell'IT outsourcing

I risultati hanno confermato che il concetto di standardizzazione della tecnologia e modularità non trova sempre applicazione. Uno standard tecnologico perde le sue caratteristiche di condivisione e apertura nel momento in cui viene immerso nelle stratificazioni dell'infrastruttura. Cioè le infrastrutture sono un risultato eterogeneo in cui la sua varietà è determinata dai modi diversi con cui le funzioni logiche possono essere implementate (spesso ciò è dovuto ad incorporazioni successive di sub-infrastrutture nate e sviluppatesi in modi diversi) [Virili, 2003]. E' importante, a questo punto, definire e gestire, come risultato, le performance di sistema, dato dal prodotto dei suoi singoli componenti.

Dal punto di vista organizzativo, così come ha evidenziato la letteratura sui prodotti complessi e sui sistemi [Brusoni e Principe, 2001], l'adozione di prodotti standard e modulari non portano automaticamente ad una struttura modulare "loosely coupled". La modularità della tecnologia può apportare dei benefici grazie ad un buon livello di divisione del lavoro riducendo il grado di interdipendenza, e i costi di comunicazione relativi, tra le parti del sistema.

Questi benefici, così come evidenziato dalla ricerca, si valorizzano insieme all'adozione di architetture organizzative standard che diventano uno strumento di progettazione organizzativa per il controllo e come strumenti di comunicazione a differenza dell'accezione Thompsoniana. Infatti, secondo Thompson la standardizzazione è il meccanismo di coordinamento più appropriato in presenza di interdipendenze deboli, situazioni relativamente semplici, attività routinarie e tecnologiche cosiddette di mediazione. Viene utilizzata come meccanismo di coordinamento in presenza di basso fabbisogno di integrazione in ambiente relativamente semplice e stabile.

Nel caso dei servizi ad alta intensità di conoscenza è difficile parlare di ambiente stabile e semplice. Quindi lo standard di processo, come ITIL per esempio, diventa un elemento che "abilita" la governance della relazione. E' interessante ribadire i che i buoni risultati ottenuti dal processo di outsourcing sono associati ad una particolare combinazione di caratteristiche iniziali, assetto organizzativo e meccanismi di governance.

E' da evidenziare, in ultima analisi, che gli attori che si coordinano seguendo una procedura standard conoscono i micro-obiettivi dettati dalle regole ma potrebbero ignorare il fine di efficacia complessiva delle azioni coordinate che la procedura persegue.

La scomposizione della conoscenza non sempre coincide con i confini definiti dalle scelte di make-or-buy, cioè vi è uno scollamento tra il task partitioning e il knowledge partitioning.

Sarebbe necessario un sforzo di coordinamento e di integrazione svolto da una unità ad hoc, un system integrator, figura che manca in questa tipologia di architetture o forme organizzative. Il change management, prototipo di una figura di system integrator, è ancora troppo orientato ad una attività di valutazione formale delle soluzioni di cambiamento proposte. Non dispone di un set di conoscenze e strumenti adatti a valutare il reale impatto di un cambiamento sulle "performance complessive di sistema". Tuttavia, l'attività di controllo sul design e coordinamento delle attività che viene svolto tipicamente dall'interno permette di avere un continuo controllo sull'infrastruttura e sulla crescita stratificata.

Il problema dell'integrazione della conoscenza è da approfondire soprattutto in riferimento al fatto che l'architettura organizzativa quasi in automatico determina il knowledge partitioning. Non è una mappatura uno-ad-uno, a parer di chi scrive, ma i principi che governano questa associazioni non sono chiari ancora, dove i risultati mostrano che i link non sono a senso unico ma hanno una mutua influenza.

In conclusione, questa ricerca ha voluto contribuire alla letteratura sull'IT outsourcing offrendo una chiave di lettura organizzativa al fenomeno dell'IT outsourcing. D'altra parte ha cercato di offrire ai practitioner uno strumento per indirizzare le scelte di IT outsourcing in chiave sistemica.

I limiti della ricerca, consci che si tratta di uno studi pilota, sono legati principalmente all'analisi dei pochi dati a disposizione. Inoltre la tecnica utilizzata ha permesso di fare delle associazioni, sottoposte ad interpretazione, tra le variabili senza dare un contributo alla spiegazione delle relazioni di causa-effetto.

Bibliografia

Becker, M.C., Organizational routines: a review of the literature. *Industrial and Corporate Change*, vol.13, n. 4, 2004.

Bidault, F., Despres, C., Butler, C., The drivers of cooperation between buyers and suppliers for product innovation. *Research Policy*, 26, 1998, 719-732.

Brusoni, S., Prencipe, A., Pavitt, K., Knowledge specialisation, organisational coupling, and the boundaries of the firm: why firms know more than they make? *Administrative Science Quarterly* 46, 2001, 597-621.

Dyer, J. and Singh, H., The relational view: cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review*, 23(4), 1998, pp. 660-679.

Feldman, M.S., Pentland, B.T., "Reconceptualizing Organizational Routines as a Source of Flexibility and Change". *Administrative Science Quarterly*, vol. 48, 2003.

Franco, M., *Competenze e sviluppo professionale nelle imprese del terziario avanzato in Molise*. Cuen, Napoli, 2001.

Heide, J., Interorganizational Governance in Marketing Channels. *Journal of Marketing*, vol. 58, 1994, pp. 71-85.

Moretti, A., Vaia, G., Zirpoli, F. (2013), The role of organizational standards in IT Outsourcing relations, in Ilan Oshri, Julia Kotlarsky and Leslie Willcocks (eds), *Advances in global sourcing: Models, Governance and Relationships*, SPRINGER, 2013.

Poppo, L. and Zenger, T., Testing alternative theories of the firm: transaction cost, knowledgebased, and measurement explanations for make-or-buy decisions in information services. *Strategic Management Journal*, 19, 1998, pp. 853-877.

Sambamurthy, V. and Zmud, R., Research commentary: the organizing logic for an enterprise's IT activities in the digital era—a prognosis of practice and a call for research. *Information Systems Research*, 11(2), 2000, pp. 105-114.

Vaia G., Zirpoli F. Le relazioni informali nella governance dei processi di IT outsourcing, in *Economia & Management*, vol. 2, 2001, pp. 67-84.

Williamson, O. E., Transaction-cost economics: the governance of contractual relations. *Journal of Law and Economics*, 22(2), 1979, pp. 233-261.

Il Cloud Computing nella strategia europea: l'armonizzazione delle regole tecniche e giuridiche

Caterina Flick

Via del Babuino, 141 (00167 – Roma) - c.flick@gmflick.it

Vincenzo Ambriola

L.go B. Pontecorvo 3 (56127 - Pisa) – ambriola@di.unipi.it

Dipartimento di informatica - Università di Pisa

Abstract. *Cloud computing is an emerging technology that aims at reducing the cost of software services and resources. The key idea is to move computational power and data storage from the client to the server, thus improving quality of service, enhancing security levels, and allowing better allocation of resources. The widespread diffusion of internet access and the strong acceleration of the use of mobile smart devices have shown that cloud solution are not only possible but somehow inevitable. In this new scenario the European Union has produced a strategy for developing cloud computing and exploiting its potentialities. In this paper we present the EU strategy addressing two relevant issues: standards and rules. The second issue is of paramount importance for the deep impact that cloud computing can have on our society. We will discuss the following aspects related to data stored in the cloud: access and integrity, property, storage and transfer. Special attention will be devoted to the adoption of cloud computing in the public administration.*

Keywords: *cloud computing EU strategy, standards and rules, public administration.*

1. Introduzione

Per *Cloud computing* o nuvola informatica (da ora *cloud*) si intende la centralizzazione di infrastrutture, piattaforme e programmi informatici e la loro redistribuzione agli utenti finali attraverso internet. La centralizzazione degli archivi di dati e dei processi per la loro fornitura consente economie di scala a cui perfino le più grandi organizzazioni da sole non possono arrivare. È per questo che il *cloud* può comportare risparmi considerevoli nei budget destinati all'IT, e rappresentare la fine dei problemi connessi con i vecchi sistemi.

Allo stato attuale in Europa il *cloud* è una tecnologia semi sconosciuta alla gran parte dei cittadini: meno di un quarto degli europei, infatti, utilizza servizi di *cloud*, nonostante questo sia uno dei settori IT a più forte crescita nel mondo. Inoltre, l'uso del *cloud* è molto più diffuso per scopi personali che per fini lavorativi.

Per la aziende, d'altra parte, il passaggio al *cloud* è visto come un'occasione importante per abbassare i costi legati alle infrastrutture IT. Il risparmio deriva in larga parte dalla possibilità di acquistare da terzi la capacità necessaria, senza dover acquistare e mantenere gli impianti. Secondo uno studio commissionato da Microsoft (*IDC – Cloud computing's role in job creation*) i ricavi legati al *cloud* potrebbero essere elevati (832 mld di euro nei prossimi tre anni) e numerosi i posti di lavoro che potrebbero essere generati (152mila in Italia, per un incremento pari a 125%).

La sostenibilità del *cloud* passa per la semplificazione e l'armonizzazione, sia sul piano tecnico che sul piano giuridico: strumenti facilmente accessibili, strumenti contrattuali flessibili e accessibili, controllo dei dati (per ragioni di sicurezza e di privacy), esercizio effettivo dei diritti.

2. La strategia europea per sviluppare il potenziale del cloud computing in Europa

La strategia adottata dalla Commissione Europea il 27 settembre 2012 è diretta ad accelerare e incrementare l'uso del *cloud* nell'economia. Tale strategia è il risultato di un'analisi ad ampio raggio dei profili politici, regolamentari e tecnologici nei Paesi membri e di un'ampia consultazione volta a individuare cosa occorre fare per sviluppare al meglio il potenziale offerto dal *cloud*. La strategia individuata si sviluppa attraverso tre direttrici.

La prima direttrice punta a semplificare ed eliminare la "giungla" di standard tecnici, con l'obiettivo di permettere agli utenti di godere di interoperabilità, portabilità dei dati e reversibilità. La Commissione lavorerà con il supporto di ENISA e altri enti per fornire assistenza allo sviluppo di schemi di certificazione volontaria, predisponendo un elenco entro il 2014.

La seconda direttrice punta a sviluppare modelli di sicurezza e condizioni contrattuali accessibili (*fair*), uniformi e di facile applicazione e interpretazione.

La terza direttrice punta a individuare una partnership europea per condurre l'innovazione e la crescita del settore pubblico, che metta insieme esperti dell'industria e utenti del settore pubblico per lavorare su requisiti comuni di acquisto per il *cloud*, in maniera aperta e pienamente trasparente. Il settore pubblico, infatti, ha un ruolo chiave nello sviluppo del mercato del *cloud*; ma con un mercato del settore pubblico frammentato, tali requisiti

hanno un impatto minimo, l'integrazione dei servizi è poca e i cittadini non ottengono il migliore risultato a basso costo (miglior rapporto qualità/prezzo).

Il Commissario per l'Agenda digitale Neelie Kroes ha sottolineato più volte l'importanza del *cloud* per la crescita economica europea e ha annunciato una European Cloud Partnership (con un investimento iniziale di 10 mln di euro) per cominciare a creare una solida base comune per gli appalti *cloud* da parte delle autorità pubbliche. Secondo i dati forniti dalla Commissione il *cloud* può generare 800mila posti di lavoro e benefici economici per oltre 200mld di euro. I vantaggi del *cloud* sono innumerevoli, soprattutto per i servizi pubblici, che possono essere resi più efficaci ed economici, con grande vantaggio per le amministrazioni.

L'esistenza di un mercato unico digitale coeso e globalmente integrato è importante per garantire agli utenti libertà di scelta tra i diversi servizi e ai provider di sfruttare le opportunità di crescita al di fuori dei confini europei. Entro fine luglio la Commissione dovrebbe indicare un quadro regolamentare armonizzato a livello europeo, volto a promuovere l'utilizzo del *cloud* nelle imprese e nelle amministrazioni pubbliche. Le difficoltà nella individuazione di una strategia europea sul *cloud* si annidano nella molteplicità di visioni e leggi diverse tra i singoli Paesi membri, in particolare su alcuni elementi importanti, quali sicurezza, privacy e trasparenza.

Tra gli aspetti che è necessario inquadrare, dunque, vi sono quelli relativi alla privacy, alla conservazione dei dati, al diritto applicabile, alle responsabilità e alla protezione dei consumatori. Altre questioni di rilievo sono quelle legate alla interoperabilità, alla standardizzazione, alla portabilità dei dati, al furto di identità. Per quanto riguarda il settore pubblico, inoltre, si pone il tema della scelta del fornitore nell'affidamento e nella gestione dei servizi di *cloud*.

3. La strategia per un cloud sostenibile: semplificazione degli standard tecnici

La semplificazione degli standard tecnici ha l'obiettivo di fornire agli utenti interoperabilità, portabilità dei dati e reversibilità. Un passaggio di questo genere, tuttavia, deve andare di pari passo con la corretta regolamentazione delle norme sulla proprietà intellettuale e le licenze d'uso.

L'esistenza di una molteplicità di standard tecnici, unita all'utilizzo di regole rigide sulla proprietà intellettuale, può infatti comportare effetti restrittivi sulla concorrenza, sia frenando l'abbassamento dei prezzi, che limitando o controllando la produzione, il mercato, l'innovazione e lo sviluppo tecnologico.

In primo luogo, se i produttori sono coinvolti in un confronto anti-competitivo nel contesto degli standard vi è il rischio che venga meno la concorrenza sui prezzi, facilitando accordi fra imprese al di fuori del mercato.

In secondo luogo, l'individuazione di standard che prevedano specifiche tecniche (spesso molto dettagliate) per un prodotto o un servizio, può limitare lo sviluppo e l'innovazione tecnologica. Nel momento stesso in cui uno standard è in corso di sviluppo, tecnologie alternative possono competere per essere incluse negli standard. Una volta che sia stata scelta una tecnologia e sia stato individuato lo standard, tecnologie diverse e produttori potrebbero incontrare delle limitazioni all'ingresso e potrebbero potenzialmente essere esclusi dal mercato.

In terzo luogo, la standardizzazione può portare a risultati contrari alla concorrenza, ove si pongano limitazioni ad alcune imprese rispetto alla possibilità di ottenere un accesso effettivo ai risultati del processo che ha condotto all'individuazione dello standard.

Negli ultimi anni la Commissione Europea ha rivisto le regole per il funzionamento degli accordi di cooperazione orizzontale alla luce della normativa europea sulla concorrenza, includendovi le linee guida sull'applicazione dell'articolo 101 del trattato sul funzionamento dell'Unione Europea negli accordi di cooperazione orizzontale. La Commissione ha inoltre affrontato il tema delle licenze per l'ICT e infine, nel novembre 2012, il tema dell'implementazione degli standard Open Source. Un gruppo di lavoro per l'individuazione di schemi di certificazione ha iniziato i lavori nel febbraio 2013.

Per quanto riguarda la semplificazione, si deve partire dall'approfondimento delle esigenze. Per portabilità dei servizi si intende la possibilità di effettuare una migrazione "semplice" delle applicazioni, macchine virtuali e dei dati da un ambiente *cloud* ad un altro. Perché ciò sia fattibile è necessario che l'ambiente di partenza e quello di arrivo siano caratterizzati dall'essere interoperabili, ovvero che l'ambiente applicativo sia effettivamente replicabile presso diversi fornitori.

Per interoperabilità, invece, si intende la possibilità di condividere gli stessi strumenti di gestione, macchine virtuali e altre risorse di *cloud*, tra una pluralità di fornitori di servizi e piattaforme di *cloud*. La portabilità dovrebbe rendere i dati e il codice "comprensibili" anche da un sistema ricevente, messo a disposizione da un diverso fornitore di servizi *cloud*, indipendentemente dalle specifiche caratteristiche delle piattaforme hardware e software utilizzate.

4. La strategia per un cloud sostenibile: regole sicure, uniformi e semplici

Nella conferenza europea sul *Cloud Computing* che si è tenuta il 7 marzo 2013 a Bruxelles ci si è chiesti quale quadro giuridico serve per creare un mercato dei servizi di *cloud*. Gli ostacoli alla regolamentazione uniforme di termini e condizioni di contratto per l'offerta di questi servizi derivano, infatti, dalle differenze esistenti in questo ambito tra le legislazioni nazionali.

È necessario dunque sviluppare un modello di condizioni contrattuali che copra gli aspetti non regolati dalla normativa europea sui contratti, quali: la conservazione dei dati dopo la scadenza del contratto; l'accesso e

l'integrità dei dati; la collocazione e il trasferimento dei dati; la proprietà dei dati; la responsabilità diretta e indiretta nella gestione del servizio in capo ai fornitori di *cloud* e ai loro sub-appaltatori.

L'identificazione e la diffusione di *best practices* relativamente alle condizioni contrattuali potrà accelerare lo sviluppo del *cloud* grazie al maggiore affidamento che si potrà creare in capo ai consumatori.

5. Accesso e integrità dei dati (sicurezza)

La disponibilità dei dati, e la loro accessibilità in qualunque momento, richiede in primo luogo la certezza sui livelli di qualità della connettività forniti da internet. In caso contrario, il servizio potrebbe essere degradato da picchi di traffico o, addirittura, reso indisponibile da eventi anomali (guasti, ad esempio). Questo aspetto dipende solo in parte dai fornitori di servizi *cloud*, mentre coinvolge in pieno gli organi politici e di governo, responsabili delle scelte di organizzazione, diffusione e gestione di internet.

L'accessibilità dei dati (personali, documentali) richiede per altro verso di preservarne l'integrità, assicurandoli rispetto alla cancellazione o al danneggiamento; è necessaria inoltre l'adozione di misure di sicurezza idonee a tutelare la riservatezza dei dati, rispetto alla visibilità o all'utilizzo da parte di soggetti non autorizzati. In estrema sintesi, dal punto di vista giuridico, occorre fare riferimento alle norme in materia di *privacy*, ovvero al d.lgs. 196/03 e alle direttive europee, in particolare la direttiva 95/46/CE - in corso di modifica, proprio con riferimento alla gestione dei dati su internet - e il recentissimo regolamento 611/2013, in vigore dal 25 agosto 2013, sull'obbligo di comunicazione delle violazioni alla *privacy*.

Il rispetto della riservatezza dei dati dipende in gran parte dai meccanismi di sicurezza adottati dal fornitore, tuttavia ciò non esime i soggetti pubblici e privati che se ne avvalgono per la gestione del proprio patrimonio informativo o - a maggior ragione - di dati e informazioni per conto di terzi, dalle responsabilità loro attribuite, nella misura in cui debbano attivarsi per l'adozione di opportune misure volte alla sicurezza dei dati e dei sistemi informatici all'interno della propria struttura. Nel caso del trattamento di dati personali l'utente di un'infrastruttura di *cloud* resta responsabile dell'adozione di misure di sicurezza, nella misura in cui riveste la qualifica di "titolare" del trattamento. Si pensi, inoltre, agli obblighi in materia di sicurezza informatica connessi con la commissione di reati informatici nella gestione dell'attività - da cui deriva una responsabilità diretta a carico delle imprese - nel quadro delle misure previste dalla Convenzione Europea del 2001 sul *cybercrime*.

Da parte del fornitore è essenziale garantire in modo trasparente la sicurezza dei dati, curandone correttamente la trasmissione e la conservazione, ivi compresa l'adozione di meccanismi sicuri di salvataggio periodico. Qualunque aspetto tecnico e organizzativo dell'infrastruttura *cloud* - dalla sua progettazione globale, al processo di sviluppo dei servizi erogati, alla configurazione dei sistemi di trasmissione, all'impostazione dei contratti con utenti e subappaltatori, ai sistemi di controllo delle richieste di accesso - ha un ruolo essenziale nell'obiettivo di garantire la sicurezza dei dati.

L'esigenza di tutelare i dati da intrusioni e utilizzi illeciti - che richiede di limitare la quantità di informazioni relative agli utenti, al traffico, alla struttura, che restano utilizzabili - può porsi in contrasto con la necessità sia di rendere i dati rapidamente e continuamente disponibili agli utenti, sia di reagire prontamente in caso di emergenza. Questa esigenza si contrappone anche con le esigenze di riservatezza (o meglio con l'interesse a non essere controllati) di coloro che sono autorizzati ad accedere. In caso di difficoltà, infatti, l'accuratezza e la tempestività della diagnosi e del rimedio sono strettamente connesse con la completezza delle informazioni disponibili. Inoltre, poiché molti errori e anomalie coinvolgono fornitori diversi, la risoluzione di un problema può richiedere la collaborazione tra più soggetti, potenzialmente residenti in Stati diversi.

In sintesi, un sistema che tutela la sicurezza e la riservatezza dei dati ha due obiettivi: da un lato deve essere in grado di individuare un attacco con accuratezza e il più rapidamente possibile e deve reagire altrettanto rapidamente, per salvaguardare i livelli di qualità dei servizi erogati; dall'altro deve garantire la riservatezza degli utenti coinvolti e le cui azioni sono monitorate. A fronte di una riduzione dei dati relativi agli utenti si può aumentare il livello di riservatezza, pur garantendo le operazioni di ripristino dell'infrastruttura; analogamente, all'aumentare dei dati disponibili, può essere elevato il livello di protezione. Una soluzione in grado di sciogliere questo dilemma si basa sull'uso di tecniche avanzate di crittografia.

6. Proprietà dei dati

L'affidamento del trattamento dei dati al fornitore di servizi di *cloud* richiede di definire preliminarmente il rapporto di titolarità - o proprietà - con i dati affidati. A questo proposito si pone, in primo luogo, il tema dei diritti di proprietà intellettuale e industriale, con le possibili difficoltà di farli valere a fronte dell'evidenziata difficoltà di individuare il luogo in cui tali dati vengono immagazzinati.

Con particolare riferimento alle opere creative - ancora oggi disciplinate in Italia dalla Legge sul diritto d'autore 633/1941, che solo in parte tiene conto delle peculiarità di internet - vale segnalare che la dematerializzazione dei supporti e la conservazione delle opere stesse presso il fornitore (su supporti di proprietà del fornitore) può cambiare le modalità di fruizione delle opere stesse, che potrebbero essere messe a disposizione di diversi utenti,

attraverso un accesso *on demand* da più postazioni, su autorizzazione dell'autore dell'opera, ma anche a sua insaputa.

Per altro verso, la riunificazione di dati nelle mani di singoli operatori di mercato che non intervengono nel mettere a disposizione su internet opere creative ma che si limitano a consentirne la diffusione richiede, sotto il profilo della tutela dell'autore dell'opera, la necessità di abbattere l'ampio ambito di irresponsabilità del fornitore del servizio, che consente a quest'ultimo di non rispondere per i contenuti caricati dagli utenti. È evidente, infatti, la difficoltà per l'autore dell'opera di individuare i singoli che, in ipotesi, abbiano violato i suoi diritti (alla quale si aggiungono problemi di coordinamento con la disciplina in materia di privacy) e costi considerevoli che un'indagine su singoli soggetti comporterebbe. Il conflitto fra i diversi interessi è stato portato nelle aule giudiziarie italiane, con la cd. vicenda Peppermint, che ha visto contrapporsi la tutela dei diritti d'autore, di cui era portatrice la casa discografica Peppermint, con la tutela della privacy degli utenti di internet.

In mancanza di modifiche normative, tra gli obblighi che gravano su un prestatore di servizi di telecomunicazione vi è solo quello di informare l'autorità giudiziaria o amministrativa di vigilanza, relativamente a notizie attendibili ricevute in ordine alla violazione di diritti d'autore compiute attraverso la propria rete di telecomunicazione e non anche l'obbligo di interrompere il servizio prestato a favore di soggetti di cui abbia avuto notizia in ordine ad atti illeciti di messa a disposizione di contenuti lesivi di diritti d'autore.

Un altro profilo di grande importanza da considerare in merito alla proprietà dei dati attiene alle esigenze connesse al trattamento dei dati personali; ciò sia durante il rapporto contrattuale, sia al termine di esso, in relazione agli obblighi di distruzione dei dati detenuti dal fornitore. In linea generale è sufficiente dire che i principi previsti in materia di tutela dei dati personali devono essere adattati all'utilizzo dei servizi di *cloud*, tenendo cioè conto delle problematiche connesse con le peculiarità di tali sistemi. Più in particolare è necessario affrontare il tema degli obblighi posti in capo al fornitore, in quanto soggetto diverso dal titolare del trattamento, e nominato da quest'ultimo responsabile, in presenza di vasti poteri nel trattamento in concreto dei dati stessi.

La questione deve essere affrontata con particolare attenzione nel caso in cui i dati affidati alla gestione del fornitore dell'infrastruttura da parte dell'utente – titolare del trattamento - riguardano terzi interessati: è questo il caso, ad esempio, della pubblica amministrazione, che tratta i dati dei cittadini, oltre a quelli dei propri dipendenti.

7. Collocazione e trasferimento dei dati

Il tema della collocazione e del trasferimento dei dati è di estrema importanza, in particolare nel caso in cui le risorse di calcolo siano fisicamente allocate in paesi situati al di fuori dei confini dell'Unione Europea.

La normativa comunitaria in tema di protezione dei dati personali – già citata in precedenza - consente la trasmissione di dati in paesi terzi solo se è garantito un adeguato livello di protezione (eventualmente sarà necessario avere il consenso dell'autorità nazionale che dovrà valutare l'adeguatezza della tutela che le parti si apprestano a garantire sulla base degli accordi negoziali pattuiti, ai sensi della vigente normativa europea sul trattamento dei dati personali). Le verifiche e gli adempimenti in tal senso - che dovranno precedere il trasferimento dei dati - saranno agevoli ove vi sia un rapporto diretto tra l'utente e il fornitore, o sub-fornitore, residente al di fuori del territorio dell'Unione Europea, saranno più complessi ove intervengano più soggetti nella fornitura di servizi *cloud*.

Un altro aspetto che merita attenzione è connesso con la possibilità che un'ingente quantità di dati - forniti da diversi titolari, portatori di interessi diversi - siano trattati da un limitato numero di società multinazionali, con possibili rischi di commistioni e conflitti (i cd *big data*).

8. Responsabilità diretta e indiretta nella gestione del servizio in capo a fornitori e subappaltatori

L'importanza di disciplinare nel dettaglio le modalità e le condizioni del servizio è di estrema rilevanza in particolare per il fatto che nel *cloud* entrano in gioco diversi soggetti: il gestore del servizio, i fornitori di accesso a internet. Da ciascuno di questi sarà necessario ricevere opportune informazioni e garanzie contrattuali, circa i parametri qualitativi del servizio prestato.

Nella migrazione ai servizi di *cloud*, infatti, l'utente diventa completamente dipendente dall'adeguatezza della professionalità dei fornitori. Ogni (anche temporanea) inutilizzabilità e/o inefficienza dei servizi può avere un impatto fortemente negativo e tradursi non solo in perdite economiche, ma anche in notevoli danni di immagine. In quest'ottica è fondamentale introdurre nei contratti clausole che prevedano un risarcimento dei danni, che descrivano, con la massima precisione, le prestazioni che l'utente si attende dal fornitore e che chiariscano come determinate prestazioni siano di cruciale interesse.

In linea generale i prestatori di servizi di *cloud* sono assimilabili agli altri fornitori di servizi internet, i cui obblighi si fondano, essenzialmente, sul d.lgs. 70/2003 che disciplina il commercio elettronico, in recepimento della direttiva europea 2000/31/CE e, per quanto riguarda la tutela dei consumatori, sul d.lgs. 206/2005, anch'esso di derivazione comunitaria. Sempre in linea generale occorre tenere conto del fatto che la fornitura di servizi telematici è una operazione contrattuale complessa, articolata in una prima fase costituita dalla conclusione del contratto e in una successiva fase di esecuzione, destinata ad evolversi in tutta la durata del rapporto e anche dopo la cessazione dello stesso. Gli obblighi del fornitore dipendono dalla tipologia di servizi offerti e dalle attività svolte nell'esecuzione del contratto.

In concreto le responsabilità derivanti dalla fornitura di servizi di conservazione digitale devono essere previste da un'analisi a monte che rifletta nel relativo contratto i termini delle responsabilità poste in capo a vari soggetti: al fornitore del servizio; a eventuali intermediari (che concorrono all'erogazione del servizio finale); al responsabile della conservazione (inteso quale persona giuridica a cui delegare parte dei processi, compresi quelli relativi all'archiviazione delle informazioni nell'infrastruttura di *cloud*), poiché la conservazione dei dati in luoghi geografici differenti potrebbe avere riflessi sia sulla normativa applicabile in caso di contenzioso tra il titolare del dato e il fornitore, sia in relazione alla specifica legge nazionale che disciplina il trattamento, l'archiviazione e la sicurezza dei dati. Di conseguenza, nel gestire al meglio la contrattualizzazione del servizio di *cloud* applicato ai processi di conservazione digitale risulta fondamentale l'applicazione di quel concetto di "interoperabilità intellettuale" tra legali, informatici, archivisti e coloro che, in qualità di responsabili (interni o esterni all'organizzazione), gestiscono il processo adottato per la conservazione digitale dei documenti.

Un buon contratto di erogazione di servizi *cloud* (e di conservazione digitale dei documenti) dovrà dunque essere il frutto dell'applicazione di norme che regolino la responsabilità civile del fornitore, nonché i processi di sicurezza e di tutela della riservatezza delle informazioni, stabilendo con quali modalità e da chi viene garantita la sicurezza dei dati. Nello stabilire le previsioni contrattuali si dovranno anche tenere presenti tutti gli obblighi richiamati da discipline specifiche e di settore come, ad esempio, quella che consente la deresponsabilizzazione dei vertici apicali in caso di commissione di reati informatici, attraverso la predisposizione di un idoneo modello organizzativo e l'adozione di specifiche procedure per l'ottimizzazione del servizio, da condividere con il fornitore. Nel contratto si dovrà inoltre tenere in conto e prevedere l'applicazione di norme tecniche internazionali o standard ISO (Come ad esempio la ISO 27001, in termini di garanzie dall'estero nei contratti stipulati con i terzi).

È possibile inserire all'interno dei contratti clausole finalizzate a garantire gli obblighi di riservatezza, la cui violazione comporta la previsione di penali. Per prevenire accessi abusivi, o comunque non consentiti, è inoltre opportuna la previsione dell'impiego della crittografia per i dati di transito oggetto di trasferimento e il ricorso a sistemi adeguati di autenticazione, in modo tale da assicurare la certezza circa l'identità dei soggetti legittimati all'accesso ai dati e alla ricezione degli stessi. Sempre in tema di flussi di dati trasfrontalieri è possibile fissare le garanzie minime necessarie per attenersi alle prescrizioni della normativa comunitaria.

Tra gli obblighi da porre a carico del fornitore dovrebbe esserci anche quello di consentire la portabilità dei dati e l'interoperabilità dell'infrastruttura con le risorse di calcolo dell'utente.

Per individuare eventuali profili di responsabilità del soggetto che gestisce i dati, è necessario chiarire l'esatta qualificazione giuridica del fornitore di un servizio di *cloud*. Si è ritenuto che questi, anche laddove mantenga propri margini di autonomia, non debba essere considerato un *controller*, ma un *processor*, cioè, in particolare, ove le modalità di gestione dei dati siano concordate dall'utente e dal fornitore mediante specifiche clausole.

In effetti, il fornitore di servizi svolge un'attività di gestione limitata solo a determinati dati e non possiede generalmente quelle competenze specifiche e adeguate a svolgere un ruolo predominante nel loro trattamento; in capo al fornitore rimane tuttavia un margine, assai ampio, di autonomia e conseguente responsabilità (quale *controller*) rispetto ai dati di traffico inerenti la circolazione delle informazioni, nell'infrastruttura o verso le risorse di calcolo dell'utente.

La ripartizione dei ruoli delle responsabilità tra fornitore e utente non può tuttavia essere stabilita rigidamente; i modelli di *cloud* esistenti possono infatti essere integrati fra di loro, sia nel caso in cui il fornitore offra un servizio integrato, sia nel caso in cui più fornitori concorrano a offrire un servizio completo. Basti pensare, a titolo di esempio, al caso in cui il servizio offerto comporti l'elaborazione e la gestione dei dati (archiviazione, effettuazione di copie, trasmissione di dati a terzi ecc.), nel quale occorre distinguere tra le operazioni di elaborazione dati effettuata direttamente dall'applicazione informatica e quella posta in essere dal gestore, in quanto solo in quest'ultimo caso si potrebbero ravvisare profili di responsabilità di quest'ultimo, restando le altre operazioni riconducibili al fornitore del servizio.

Particolare attenzione - con l'individuazione di specifiche garanzie - va posta nel caso in cui il fornitore sia legittimato ad appaltare a terzi parte dei servizi, anche solo la gestione/allocazione delle risorse fisiche su cui risiedono i dati. In questo caso occorre prevedere nel contratto specifiche garanzie relative al sub-appalto, imponendo al fornitore di rendere noto all'utente l'affidamento dell'appalto a terzi.

La legge applicabile alla fornitura di servizi di *cloud* nella pratica è spesso quella del luogo in cui è stabilito il fornitore; ove questi sia una società con sede all'interno dell'Unione Europea non si pongono ostacoli, dato che si applica la normativa comunitaria. Tuttavia, la problematica attinente alla normativa applicabile in caso di controversie sorte tra il fornitore del servizio e l'utente può presentarsi in termini più complessi, al verificarsi di situazioni nelle quali il fornitore interagisce con altri soggetti.

In un'infrastruttura di *cloud* i dati spesso vengono memorizzati in diversi *data center*, che ben possono trovarsi materialmente in più nazioni; il fornitore del servizio può inoltre ricorrere a terzi per lo scambio di risorse di calcolo (si pensi, ad esempio, al fornitore che non ha a disposizione sufficiente capacità in termini di supporti di memorizzazione e si affida perciò ad altri). Questi scambi tra più soggetti determinano un flusso continuo dei dati, così che non è facile individuare con esattezza il soggetto che li gestisce in un determinato momento, né la loro esatta localizzazione; l'utente si limita ad accedere al servizio e sarà il fornitore a reperire i dati e a riaggiargarli.

Una pluralità di relazioni negoziali ha inoltre luogo quando l'azienda fornitrice mette a disposizione un servizio a sua volta ottenuto da altri fornitori, sempre in modalità *cloud*.

Dunque, in considerazione del fatto che diversi soggetti potrebbero intervenire nella gestione dell'infrastruttura, occorre disciplinare l'ipotesi in cui il fornitore si avvalga per il compimento di parte di essi di terzi fornitori non residenti nel territorio dell'Unione Europea. Le possibili soluzioni proposte vanno dall'impiego di clausole comunitarie tipo, tra l'utente che usufruisce dei servizi e il sub-fornitore; al mandato rilasciato dall'utente al fornitore, affinché quest'ultimo stipuli da sé accordi con il sub-fornitore; alla previsione di specifici accordi contrattuali tra le parti.

La predisposizione del contratto può tuttavia non essere sufficiente a porre l'utente al riparo da rischi, dato che i fornitori possono affidare ad altri operatori alcune attività (ivi comprese la gestione delle risorse fisiche che mantengono i dati) e possono essi stessi essere soggetti a vicende societarie (fusioni o acquisizioni) che portano a cambiamenti rilevanti, quali, ad esempio, la sede legale e, conseguentemente, la disciplina normativa applicabile. In sintesi, anche dopo la stipula di un valido contratto, potrebbe essere difficile per l'utente esigere il rispetto degli obblighi che sono stati disciplinati nel contratto stesso, specie nel caso in cui il fornitore non risieda nel medesimo paese ove risiede il contraente destinatario dei servizi.

9. Conservazione dei dati (dopo la scadenza del contratto)

Quanto sopra evidenziato in tema di rapporti con il fornitore e il sub-fornitore è da tenere in considerazione anche con riferimento alla conservazione dei dati dopo la scadenza del contratto. Modalità e tempi di conservazione dei dati, in particolare, devono formare oggetto di specifiche previsioni sin dal momento della stipula del contratto.

10. Conclusioni

La frontiera tecnologica si muove rapidamente, cambiando stili di vita e assunzioni ritenute assolute. Nel caso del *cloud computing* gli effetti sono ancora poco visibili e poco percepiti. Come cittadini diamo per scontati molti servizi offerti su internet (*home banking*, sistemi di prenotazione e acquisto *on-line* di titoli di viaggio, posta elettronica, memorizzazione remota di dati) senza sapere che si basano su infrastrutture *cloud*. Sempre più la nostra società dipende dal corretto e continuo funzionamento di queste tecnologie.

Diversa è la velocità di adeguamento dell'impianto giuridico al progresso tecnologico, con il rischio (ma spesso la certezza) che l'introduzione di nuovi e necessari servizi non sia accompagnata da una solida normativa che tuteli gli interessi degli utenti ma anche di chi ha investito in tecnologia.

Forse è questa la novità del terzo millennio: un mondo in cui le leggi devono inseguire la realtà e non viceversa.

Bibliografia

- Bollier D., *The Promise and Peril of Big Data*, The Aspen Institute 2010
- Coleman N., Borrett M., *Cloud Security, who do you trust?*, IBM 2010
- Data Protection Working Party – Art. 29, Working Document 1/2009 on pre-trial discovery for cross border civil litigation, 2009
- DigitPA, *Raccomandazioni e proposte sull'utilizzo del cloud computing nella Pubblica Amministrazione*, ver. 2.0, 2012
- ENISA, *Priorities for Research on Current and Emerging Network Technologies*, 2012
- ENISA, *Security & Resilience in Governmental Clouds - Making an informed decision*, 2011
- Enter the Cloud, *Cloud survey 2013: lo stato del cloud computing in Italia*, www.enterthecloud.it, 2013
- European CIO Association, *Users recommendations from the European CIO Association for the success of the cloud computing in Europe*, ARES 2012
- European Commission, *Commission plans guide through global internet policy labyrinth*, ec.europa.eu/digital-agenda 2013
- European Commission, *Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (COM(2012) 11 final)*, 2012
- European Commission, *Un'agenda digitale per l'Europa*, COM(2012) 245 def/2, 2010
- Flick C. e Ambriola V., *Dati nelle nuvole: aspetti giuridici del cloud computing e applicazione alle amministrazioni pubbliche*, *Federalismi.it*, 6, 2013
- Forum PA – Quaderni, *La PA sulla nuvola – G-Cloud: innovare per guadagnare efficienza e ridurre i costi*, 2012
- Gantz. J.F., Minton S., Toncheva A., *Cloud computing's role in job creation*, IDC 2012
- Mantelero A., *Processi di outsourcing informatico e cloud computing: la gestione dei dati personali e aziendali*, *Dir Inf.* 2010, 673

Il *Cloud Computing* nella strategia europea: l'armonizzazione delle regole tecniche e giuridiche

Schubert L., *The Future of Cloud Computing. Opportunities for European cloud computing beyond 2010*, Expert Group Report, 2009

Visualizing Internet-Measurements Data for Research Purposes: the NeuViz Data Visualization Tool

Giuseppe Futia¹, Enrico Zimuel², Simone Basso¹, Juan Carlos De Martin¹

¹Nexa Center for Internet & Society, Politecnico di Torino (DAUIN),
Corso Duca degli Abruzzi, 24, 10129 Torino (TO), Italy
{giuseppe.futia, simone.basso, demartin}@polito.it

²R&D Department, Zend Technologies Inc.,
19200 Stevens Creek Blvd, Cupertino, CA 95014, USA
enrico@zend.com

Abstract. *In this paper we present NeuViz, a data processing and visualization architecture for network measurement experiments. NeuViz has been tailored to work on the data produced by Neubot (Net Neutrality Bot), an Internet bot that performs periodic, active network performance tests. We show that NeuViz is an effective tool to navigate Neubot data to identify cases (to be investigated with more specific network tests) in which a protocol seems discriminated. Also, we suggest how the information provided by the NeuViz Web API can help to automatically detect cases in which a protocol seems discriminated, to raise warnings or trigger more specific tests.*

Keywords: data visualization, network performance, big data.

1. Introduction

The Internet is a cornerstone of our societies and has been enabling unprecedented levels of social interaction, content sharing, business creation, as well as innovation in many fields. As Frischmann argues convincingly, one of the main reasons why the Internet is so relevant for us is that the Internet is an *infrastructural resource*, i.e., a shared piece of infrastructure that is typically managed as a *commons* in a non-discriminatory way [Frischmann, 2012].

However, the Internet is not an infrastructural resource as a fact of nature, or because of an immutable, technological law; the current status of the Internet is, instead, the consequence of specific choices, both private and public, that could very well change over time. For example the policy decision of who (the State or the Internet Service Providers) should finance (and under which conditions) the so-called 'Next Generation Networks' (NGNs) has the potential of radically changing the landscape.

In fact, many parties (including the authors of this contribution) believe that, if States allow the Internet Service Providers (ISPs) to implement premium services to collect more money and finance NGNs, the infrastructural-resource characteristics of the Internet may become less relevant, and the Internet may

lose part of its *generativity* (i.e., the property of enabling more and more people to write and distribute software and/or media content [Zittrain, 2009]).

To be fair, there is little empirical evidence supporting most policy positions on both sides of the debate. On the one hand, for instance, it is hard to prove empirically *ex ante* that allowing ISPs to implement premium services will reduce the generativity of the Internet. On the other hand, there is surprisingly little evidence backing the ‘bandwidth hogs’ argument (i.e., the argument that there is a little number of people that consume most bandwidth). The Internet policy debate, in general, is so ill informed by poor data, by missing data, and by data provided by one single stakeholder that – we agree with Palfrey and Zittrain – there is a need for more, better data to anchor the debate to solid foundations and move forward [Palfrey and Zittrain, 2011].

This is indeed starting to happen: more and more network measurement tools and visualizations, in fact, are being developed by researchers and companies worldwide. Many of such tools and visualizations are hosted by Measurement Lab (<http://measurementlab.net/>), an umbrella project run by the Open Technology Institute and the PlanetLab Consortium, and supported by academic partners and companies such as Google.

In this paper, in particular, we propose NeuViz (Neubot Visualizer), an architecture that allows us to process and visualize the data collected by Neubot, the network neutrality bot [Basso et al, 2011a], one of the tools hosted by Measurement Lab. Neubot – a project of the Nexa Center for Internet & Society – is a centrally-coordinated bot that runs the in background on the user computer and periodically runs network-performance tests that currently emulate HTTP and BitTorrent, and, in future, will emulate other protocols.

The purpose of NeuViz is to visualize and navigate Neubot data through its Web user interface, to search for cases (to be investigated with more specific network tests) in which a protocol seems discriminated. Also, NeuViz is designed to help, in the future and with a more advanced Neubot architecture, to automatically detect cases in which a protocol seems discriminated, to raise warnings or trigger more specific tests.

Many existing visualization architectures are based on cloud services and allow to query the data on demand using SQL-like query languages; compared to such visualization tools, NeuViz is much more optimized for the specific purpose of visualizing network measurement data. We designed, in fact, a robust, scalable backend architecture to support special-purpose, complex data analysis, in which the query (or the filtering algorithm) is executed in advance on the network-experiments dataset, and in which the result is stored in one (or more) NoSQL database(s), for fast data access.

We evaluate our work by loading into NeuViz the results of two Neubot network tests (Speedtest and BitTorrent) collected in the January 2012 - May 2013 period. We show that NeuViz helps us to effectively navigate Neubot data to identify cases in which a protocol seems discriminated. Also, we suggest that the information provided by the NeuViz Web API can help to automatically detect cases in which a protocol seems discriminated.

The rest of this paper is organized as follows. In Section 2 we describe related network measurement tools and visualizations. In Section 3 we describe Neubot and the Neubot data that we used in this paper. In Section 4 we

describe the NeuViz architecture. In Section 5 we describe our implementation choices. In Section 6 we describe what we learnt from browsing Neubot data with NeuViz. In Section 7 we draw the conclusions, and we describe future developments.

2. Related Work

In this section we mention the related tools and visualizations. Some of the tools that we mention (including Neubot) are hosted by Measurement Lab (M-Lab) [Dovrolis et al, 2010], a distributed server platform that also provides advanced services (e.g., the possibility of querying the hosted tools data using BigQuery, a RESTful service to query big datasets using an SQL-like query language – <https://developers.google.com/bigquery/>).

We mention four tools similar to Neubot: (i) Glasnost, which is an M-Lab based tool that compares a certain protocol flow (e.g., BitTorrent, Emule) with a *reference flow* to detect traffic shaping and its cause (e.g., the port number, the payload) [Dischinger et al, 2010]; (ii) NDT, the Network Diagnostic Tool, which measures the download and upload speed between the user computer and a Measurement Lab server [Carlson, 2003] and records the TCP state using Web100 [Mathis et al, 2003]; (iii) SpeedTest.net (<http://www.speedtest.net/>), which estimates the broadband speed of the user's connection using parallel HTTP flows; (iv) Grenouille, which measures the round trip delay, the download speed, and the upload speed (<http://grenouille.com/>).

Differently from Glasnost Speedtest.net, and NDT – which run on-demand tests – Neubot and Grenouille run tests in the background; however, Neubot uses diverse protocols, while Grenouille focuses on the performance only.

We mention four visualizations similar to NeuViz: (i) the two visualizations of Glasnost data developed by, respectively, the Syracuse University School of Information Studies (<http://dpi.ischool.syr.edu/MLab-Data.html>), and the Open Knowledge Foundation (<http://netneutralitymap.org/>); (ii) the many BigQuery-based NDT visualizations by the M-Lab team (e.g., <http://goo.gl/m9WbS> and <http://dmadev.com/2012/11/19/>); (iii) the SpeedTest.net visualizations available at NetIndex.com (<http://www.netindex.com/>); (vi) the Grenouille visualizations available at <http://grenouille.com/>.

Similarly to the NDT visualizations NeuViz is based on the world map; however, NeuViz is optimized for complex data analysis and uses precomputed data, while the NDT visualizations are more interactive and fetch the data from BigQuery on demand. Also, the aim of NeuViz is similar to the aim of the Glasnost visualizations; both, in fact, aim at making access networks more transparent by, respectively, showing anomalies and alleged shaping.

3. Neubot and Neubot data

Neubot is a free-software Internet bot that performs active, lightweight network-performance tests [Basso et al., 2011a]. Once installed on the user's computer, Neubot runs in the background and every 30 minutes performs active transmission tests with servers hosted by Measurement Lab. To coordinate the botnet composed of all the Neubot instances worldwide, there is the so-called

Master Server, which suggests each Neubot the next test to run as well as the default test parameters. Currently, the Master Server does not optimize the suggestions returned to each Neubot; however, as we will show, the information returned by NeuViz could help the Master Server to implement more dynamic policies.

Neubot implements three network performance tests: Speedtest, BitTorrent, and RawTest. Speedtest measures the network performance using the HTTP protocol, BitTorrent measures the network performance using the BitTorrent protocol, and the RawTest test measures raw, TCP-level performance (hence the name of the test). In this paper we only describe the Speedtest and the BitTorrent tests, because we are mainly interested to use NeuViz to find cases in which a protocol seems discriminated.

Speedtest is an HTTP-based test – originally inspired to the test of SpeedTest.net, hence the test name – that downloads and uploads data using a single HTTP connection [Basso et al., 2011b]. The test measures the download and the upload speed at the application level. Also, the test estimates the base Round Trip Time (RTT) using as a proxy the time that the connect system call takes to complete (later indicated as *connect time*). The test transfers a number of bytes that guarantees that each phase of the test (download, upload) lasts for about five seconds.

The BitTorrent test is similar to the Speedtest test, except that it uses the BitTorrent protocol (http://www.bittorrent.org/beps/bep_0003.html) instead of the HTTP protocol. However, while Speedtest makes a single GET request for a large-enough amount of data, BitTorrent – to better emulate the BitTorrent protocol – downloads many small chunks in a request-response fashion and, to approximate a continuous transfer, makes many back-to-back requests at the beginning of the test.

Measurement Lab (which hosts Neubot on its servers) periodically collects the Neubot experiments results saved on its servers and publishes such results on the Web (<http://measurementlab.net/data>) under the terms and conditions of the Creative Commons Zero 1.0 Universal license. We mirrored the data provided by Measurement Lab, and we converted such data in CSV format, generating CSV files that contain one month of data each. To prepare this paper, we imported into NeuViz the CSV files from January 2012 to May 2013 (reading 5,383,376 test, from 4,037 Neubot clients worldwide, for a total of 1.5 GB).

Each CSV file contains the following fields (the type is indicated in parentheses): client address (str); connect time, in second (float); download speed, in byte/s (float); Neubot version (str); operating system platform (str); server address (str); test name (str: “speedtest” or “bittorrent”); timestamp of the test, i.e., the number of seconds since 1970-01-01 00:00 UTC (int); upload speed, in byte/s (float); unique identifier of the Neubot instance (str).

4. Description of the NeuViz Architecture

Fig. 1 shows the NeuViz architecture, which is a pipeline that processes data provided by *Producers*, and which organizes the data such that *Consumers* can visualize (or further process) such data. The pipeline is composed of a *Backend*

Visualizing Internet-Measurements Data for Research Purposes: the NeuViz Data Visualization Tool and a *Frontend*: the Backend receives data from many Producers and processes such data to allow for efficient visualization; the Frontend is a Web interface that visualizes the data. In the middle there is a Web API.

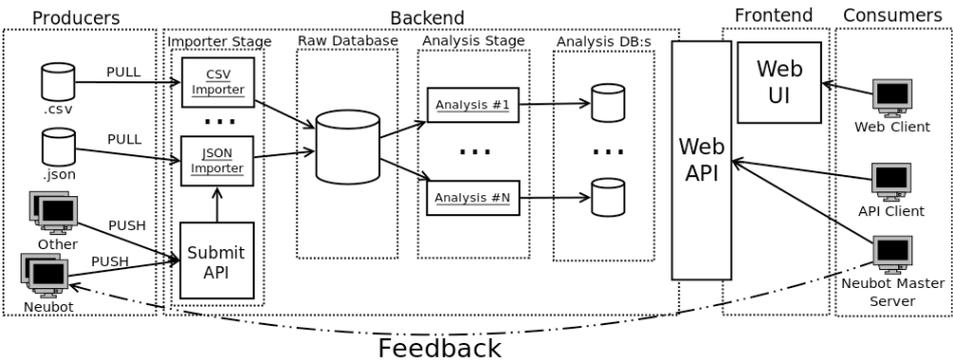


Figure 1: NeuViz architecture

4.1 Producers

As a first approximation a Producer is a static dataset. For example, in this paper we used Neubot data expressed in CSV format and in the future we may want to import datasets from other projects (e.g., SpeedTest.net) and encoded in many formats (e.g., JSON).

NeuViz also includes a Submit API, which allows network-experiment tools (e.g., Neubot and possibly other network-measurement tools) to push the result of their experiments just after the experiments are run. We added the API because we want to create a feedback loop in which data posted by Neubot is processed by NeuViz and consumed by the Master Server to provide better suggestions to Neubot instances.

4.2 The Backend

The Backend is composed of two processing stages (the *Importer Stage* and the *Analysis Stage*), each followed by a database stage (respectively, the *Raw Database* and the *Analysis Databases*).

The Importer Stage is composed of many modules: each module imports data from a specific network-measurement tool (e.g., Neubot, Grenouille) and format (e.g., CSV, JSON). Also, network-measurement tools can push their data via the Submit API. The problem of whether (and how) to authenticate the tool submitting its data is not discussed in this paper.

The Raw Database receives heterogeneous data organized in a uniform format (e.g., JSON) by the Importer Stage. It is not practical to reduce all the input data to the same schema, because each network experiment saves different metrics; therefore, NoSQL seems the best technology to implement the Raw Database. Also, the Raw Database shall be scalable-enough to handle continuous streams of data posted on the Submit API by Neubot (and possibly by other network measurement tools).

The Analysis Stage is a collection of modules that periodically fetch data from the Raw Database and process it to produce the aggregate data needed for the visualizations. We already implemented a world-map-based visualization and we are working on other visualizations. Of course, the Analysis Stage needs to be scalable, because we need to process multiple times the raw data stored into the Raw Database. Also, the Analysis Stage should minimize the computational cost of adding the results of new experiments to NeuViz.

The Analysis Databases are a number of conceptually-separated databases that store data which is ready to be visualized on the NeuViz Frontend with minimal computational cost. We want, in fact, to allow the user to visualize and browse the data as seamlessly as possible.

4.3 The Web API, the Frontend, and Consumers

The Web API connects the Backend and the Frontend (i.e., the Web interface of NeuViz). A Web browser that uses the NeuViz Web interface is the default Consumer; however, other Consumers are possible. For example, as mentioned above, as part of our future work we plan to extend the Master Server to fetch data from NeuViz to provide better suggestions to the Neubot instances.

In this paper we don't discuss whether and how the access to the Web API and to the Frontend should be restricted. This will possibly be the subject of a future work.

5. Implementation Choices

In this section we describe the implementation of the first NeuViz prototype (<https://github.com/neubot/neuviz>), and we explain our implementation choices.

5.1 Backend

We implemented the Importer Stage and the Analysis Stage in Python. The Importer Stage converts the input into a JSON document and adds geolocation information, if needed. The Analysis Stage prepares the data for a world-map-based visualization that allows the user to navigate the spatial level (countries, and cities) and the temporal level (hour of the day). The output of the Analysis Stage is a set of JSON documents.

For both the Raw Database and the Analysis Databases we use MongoDB, a NoSQL database that indexes JSON documents and that is very often deployed in big data scenarios [Moniruzzaman and Akhter, 2013]. To improve performance, we exploit the *indexes* feature of MongoDB to speed up the query execution. We imported 1.5 GiB of Neubot CSV files, from January 2012 to May 2013, and we stored such data into a MongoDB database, processing about 5.3M samples in less than 60 minutes. We run the Importer on a laptop with an Intel Core i7 CPU at 2.0 Ghz, with 8 GB of RAM, and a 256-GB SSD, running GNU/Linux 3.5.0.

We geolocated the incoming data using the GeoLite Free Geolocation Database (<http://dev.maxmind.com/geoip/legacy/geolite/>). As explained in the GeoLite website, when the database is not up-to-date, the geolocation loses

1.5% of accuracy each month because IP addresses are re-assigned. To minimize the damages caused by out-of-date GeoLite databases, we never used databases older than two months when we geolocated Neubot results.

In the Analysis Stage we computed the median number of tests, the median number of Neubot instances, the median connect time, the median download speed, and the median upload speed, for both Speedtest and BitTorrent, for each hour of the day, for each month, and along the geographical dimension (i.e., country, city), and the business dimension (i.e., ISP). We decided to use the median, which is a common index used to analyze network traffic, to avoid the risk that few outliers could dominate our indexes.

For future scalability we designed the code in a way that potentially allows us to use MapReduce techniques on cloud services (e.g., Amazon Elastic MapReduce, <http://aws.amazon.com/elasticmapreduce/>). To this end we divided the Importer and the Analysis code into a *map* step and a *reduce* step.

5.2 The Web API and the Frontend

To access the NeuViz API, the user sends the following HTTP/1.1 request: GET /neuviz/1.0/<viz>/<params>, where <viz> is the name of the visualization, and <params> is a placeholder for (possibly-empty) parameters. The returned JSON contains a recursive set of dictionaries that represent the geographical dimension (country, city) and the business dimension (ISP). The leaves are dictionaries that contain the following hour-wide median statistics for the Speedtest and the BitTorrent tests: download speed, upload speed, connection time, number of Neubot instances, number of tests.

The Web Interface, written using D3.js (<http://d3js.org/>), allows the user to explore different network measurement performances at different geographic dimensions (country, cities, and ISPs). For simplicity, and since it does not seem to cause any performance issue, we currently use the Web interface to compute some statistics, e.g., the difference between the median Speedtest download speed and the median BitTorrent download speed that we use in Section 6.2 to compare the performance of BitTorrent and Speedtest.

6. Results

In this section we report what we learnt from using NeuViz to browse Neubot data, both in terms of number of tests and in terms of performance.

6.1 Number of Neubot Tests

Fig. 2 shows the visualization of the number of tests per country and per hour. The alpha channel of the country color indicates the median number of tests per country. The visualization, in particular, shows the median number of tests performed between 9:00 PM and 10:00 PM (local time) in April 2013. The selected country is Canada, in which the median number of tests performed is indicated by the number in the bottom right corner (1084).

By selecting other countries in the visualization, we have seen that the countries with more median tests per hour between 9:00 PM and 10:00 PM in

April 2013 are: the US (4223); Italy (2866); Germany (2285); and Canada (1084). Other countries have less tests per hour.

The availability of the number of tests per country is interesting because, by knowing the number of tests per country, the Master Server could maximize the test coverage; e.g., it can increment the frequency of testing on countries where there are few Neubot users.

6.2 Comparison of Speedtest and BitTorrent performance

Before studying the visualization that shows the difference between the Speedtest and the BitTorrent test download and upload speeds, we checked that the Speedtest and the BitTorrent connect times were 'comparable'. To this end we arbitrarily define 'comparable' two median connect times whose difference is smaller than five millisecond, in our experience a reasonable threshold for this kind of analyses.

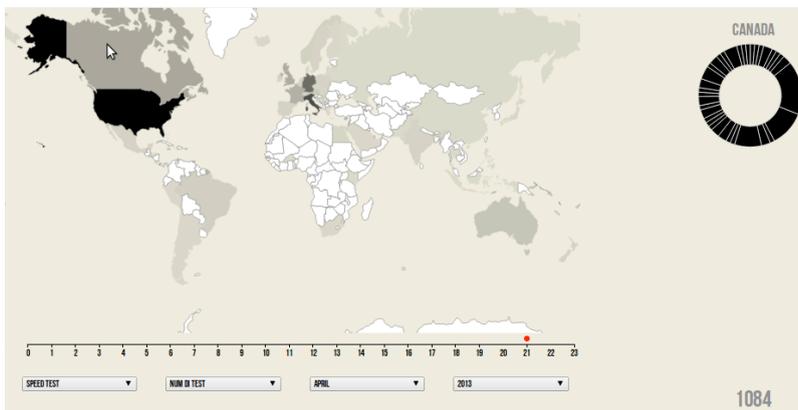


Figure 2: NeuViz interface of the worldwide map for Neubot data of April 2013

The visualization of the difference between the median BitTorrent connect time and the median Speedtest connect time shows, surprisingly, that in Italy such difference is always positive and often greater than five millisecond (i.e., the Speedtest connect time is typically lower). Italy is the only country in which, for 2013 data, we noticed this behavior.

Also we noticed interesting things from the comparison of the median upload speed in countries in which the median connect times are comparable. We noticed, in fact, that in 2013 the median upload difference between Speedtest and BitTorrent in Canada was very often positive, while the same difference was very often negative in the US (see Fig. 3).

Moreover, when comparing the download speeds in countries in which the connect times are comparable, we also noticed that the US Speedtest download speed is always lower (in median) than the BitTorrent one for every hour of the day and for every month of 2013. Interestingly, instead, the download speeds are comparable in Italy, in which – as we have seen – there is a connect time bias in favor of Speedtest.

The above observations lead us to speculate that: (a) BitTorrent is slightly faster than Speedtest; (b) in Italy the two tests are comparable because of the connect-time bias that we observed; (c) the BitTorrent upload speed seems to be discriminated in Canada. Of course, these are only hypotheses that need to be verified (or contradicted) by more detailed experiments.

6.3 Concluding Remarks

Despite being still in beta stage, NeuViz allowed us to discover the three diverse network anomalies we described in 6.2. In the future, a more advanced Master Server could learn, from the NeuViz API, about similar anomalies and ask Neubot instances that are near the anomalies to gather more information needed to investigate the anomalies (e.g., by capturing packets to gather RTT samples useful to understand whether there is a connect-time bias).

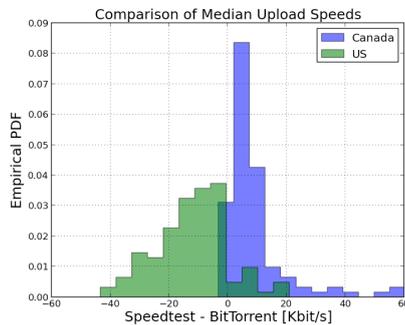


Figure 3: the Empirical Probability Density Function (PDF) of the difference of the median upload speed of US and Canada

7. Conclusion and Future Work

In this paper we described NeuViz, an architecture that allows us to process and visualize the data collected by Neubot, the active, network-measurement tool developed by the Nexa Center for Internet & Society. The purpose of NeuViz is to visualize and navigate Neubot data through its Web user interface, to search for cases (to be investigated with more specific network tests) in which a protocol seems discriminated.

Differently from other visualization architectures NeuViz is much less flexible and much more optimized, on purpose. NeuViz, in fact, executes the queries in advance and the result is stored into one or more NoSQL databases (using MongoDB), for fast data access. The Backend of NeuViz, written in Python, is structured to ease the task of porting it to a cloud-based MapReduce solution, for future scalability. The Web interface Frontend of NeuViz shows a world-map-based visualization of Neubot results implemented using the D3.js library.

To evaluate NeuViz we loaded one-year-and-a-half records collected by two network tests periodically run by Neubot, called Speedtest (based on HTTP) and BitTorrent. We showed that NeuViz effectively helped us to identify cases (to be investigated with more specific network tests) in which a protocol seems

discriminated. In our discussion we also suggested how the Web API of NeuViz can help to automatically detect cases in which a protocol seems discriminated, to raise warnings or trigger more specific tests (by cooperating with the Master Server of Neubot). As part of our future work we plan to extend NeuViz to automatically raise warnings and to cooperate with the Master Server of Neubot to trigger more-specific network experiments.

Acknowledgments

The first prototype of the NeuViz project has been developed as final project of the BigDive course 2013 (<http://www.bigdive.eu>). We would like to thank Christian Racca of the TOP-IX Consortium and all the staff and teachers of the BigDive course for their support during the development of this project.

References

[Basso et al, 2011a] Basso S., Servetti A., De Martin J. C., The network neutrality bot architecture: A preliminary approach for self-monitoring of Internet access QoS, in Proc. of the Sixteenth IEEE Symposium on Computers and Communications, Corfu, Greece, 2011.

[Basso et al, 2011b] Basso S., Servetti A., De Martin J. C., The hitchhiker's guide to the Network Neutrality Bot test methodology, in Proc. of Congresso Nazionale AICA 2011, Torino, 2011.

[Carlson, 2003] Carlson R., Developing the Web100 Based Network Diagnostic Tool (NDT), In Proc of the Passive and Active Measurement Conference, 2003.

[Dischinger et al, 2010] Dischinger M., Marcon M., Guha S., Gummadi K. P., Mahajan R., Saroiu S., Glasnost: Enabling End Users to Detect Traffic Differentiation, in Proc. of USENIX Symposium on Networked Systems Design and Implementation, 2010.

[Dovrolis et al, 2010] Dovrolis C., Gummadi K. P., Kuzmanovic A., Meinrath S., Measurement Lab: Overview and an Invitation to the Research Community, ACM SIGCOMM Computer Communication Review, 40, 3, 2010, 53–56.

[Frischmann, 2012] Frischmann B. M., Infrastructure: The Social Value of Shared Resources, Oxford University Press, 2012.

[Mathis et al, 2003] Mathis M., Heffner J., Reddy R., Web100: Extended TCP Instrumentation for Research, Education and Diagnosis, ACM SIGCOMM Computer Communication Review, 33, 3, 2003, 69–79.

[Moniruzzaman and Akhter, 2013] Moniruzzaman A. B. M., Akhter H. S., NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison, International Journal of Database Theory and Application, Vol. 6, No.4, 2013.

[Palfrey and Zittrain, 2011] Palfrey J., Zittrain J., Better Data for a Better Internet, Science, 334, 6060, 2011, 1210-1211.

[Zittrain, 2009] Zittrain J., The future of the Internet--and how to stop it., Yale University Press, 2009.

Una Panoramica delle Piattaforme Cloud e Cloud Pattern

Beniamino Di Martino¹, Giuseppina Cretella², Antonio Esposito³

*Department of Industrial and Information Engineering,
Second University of Naples
Via Roma 29, Aversa*

¹*beniamino.dimartino@unina.it*

²*giuseppina.cretella@unina2.it*

³*antonio.esposito7@studenti.unina2.it*

Abstract. *Cloud computing is a very highly interesting topic in both academia and industry and many big players of the software industry are interested to become protagonists of such panorama. This paper presents a comparative overview of cloud services offered by different cloud provider, in particular IBM, Oracle, Amazon and Microsoft, focusing on their PaaS offers and the available tools developed by the vendor themselves to enable their resources and services management. The analysis goes through the comparison of cloud pattern defined from specific vendor and agnostic pattern defined by independent efforts.*

Keywords: *Paas Cloud Service, IBM SmartCloud, Windows Azure, Oracle Cloud, Cloud Pattern*

1. Introduzione

Il cloud computing è un argomento molto discusso sia nel mondo accademico che commerciale, tanto che i maggiori protagonisti delle industrie software si sono addentrati nello sviluppo di servizi cloud.

I fornitori di piattaforme per il Cloud offrono i loro servizi secondo differenti modelli: infrastructure as a service (IaaS), platform as a service (PaaS), e software as a service (SaaS). IaaS è il modello di base offerto dalla maggior parte dei provider, PaaS e SaaS estendono e astraggono rispetto ad esso. IaaS fornisce le capacità computazionali, di archivio e di rete basilari, per cui è più flessibile, ma richiede agli utenti di fornire tutto il resto, comprese le applicazioni, il middleware e il database. PaaS fornisce piattaforme computazionali complete, che comprendono ambienti di sviluppo o di esecuzione, con supporto a diversi linguaggi di programmazione, rendendo lo sviluppo più semplice e veloce con un grado di flessibilità, tuttavia, molto più basso. Il livello SaaS fornisce invece applicazioni e software completi e funzionanti, eventualmente componibili. Il contesto PaaS, a differenza dell'IaaS

ha subito un avvio più lento; inizialmente l'unico grande fornitore nell'ambito PaaS era Microsoft, con la sua piattaforma Windows Azure lanciata sul mercato il primo febbraio 2010. In seguito anche altre grandi organizzazioni come IBM, Oracle e Amazon si sono fatte spazio con la loro piattaforma per lo sviluppo di applicazioni sul cloud. In questo articolo analizzeremo le offerte PaaS dei cloud provider IBM, Amazon, Oracle e Azure, sottolineandone le differenze e descrivendo il loro approccio all'utilizzo dei pattern, un concetto ereditato dalla letteratura che nel cloud computing assume un significato ben specifico, ossia quello di fornire una guida per aiutare gli sviluppatori di applicazioni cloud a comporre tra di loro funzionalità e servizi per realizzare applicazioni che abbiano determinate caratteristiche.

In ambito industriale, il concetto di cloud pattern viene a volte utilizzato in maniera fuorviante, impedendo un posizionamento competitivo dei servizi sul mercato del cloud computing in quanto i cloud provider forniscono come cloud pattern una soluzione architetturale preconfigurata basata sui loro prodotti. In questo articolo verrà esposta prima una panoramica delle piattaforme cloud e dei servizi offerti a livello PaaS dai fornitori IBM, Oracle, Amazon e Windows e poi si illustrerà il concetto di cloud pattern, mostrando come tale concetto viene adottato con sfumature diverse in diversi contesti.

2. Una panoramica delle offerte PaaS

Questa sezione propone una panoramica delle principali offerte dei cloud provider a livello PaaS, evidenziando le principali caratteristiche e funzionalità che contraddistinguono le singole piattaforme.

2.1 L'offerta PaaS di IBM: SmartCloud Application Services

IBM SmartCloud è un portfolio integrato di servizi che rappresenta l'intera pila dei servizi cloud offerto da IBM. Esso è composto da tre famiglie di servizi: **SmartCloud Foundation** comprende hardware e software che possono essere usate per costruire cloud privati autonomi o che possono essere connessi a cloud privati preesistenti, in una configurazione ibrida; **SmartCloud Services** include l'offerta pubblica di SmartCloud Enterprise e il servizio cloud privato di SmartCloud Enterprise+; **SmartCloud Solutions**, che fornisce servizi SaaS basati sulle applicazioni fornite da IBM.

IBM ha introdotto **SmartCloud Application Services** [SCAS] come parte del portfolio di SmartCloud, creando un'estensione logica dei servizi infrastrutturali di SmartCloud Enterprise (SCE) ed Enterprise+, definendo un PaaS integrato che permette di massimizzare i benefici dell'uso del cloud per applicazioni e ambienti diversi. Esso può essere visto come un servizio aggiuntivo fornito da SmartCloud Enterprise, per cui tutti gli utenti SCE possono utilizzare SCAS immediatamente, accedendovi attraverso lo stesso portale.

I servizi PaaS sono rappresentati da pattern applicativi, che operano sul cloud IaaS. IBM descrive i pattern come componenti software predefiniti e disegnati per velocizzare lo sviluppo e il rilascio di applicazioni in cloud, basati su architetture predefinite.

I primi servizi resi attivi sono il **Collaborative LifeCycle Management** e il **Workload Service**. Il Collaborative Lifecycle Management service consente ai manager dello sviluppo di aggiungere utenti e ruoli attraverso il portale SCE, permettendo così il monitoraggio e la disponibilità dei dati. Esso utilizza il set di tool di sviluppo di IBM **Rational**: in particolare, il ciclo di vita del processo di sviluppo è gestito attraverso Rational Team Concert, Rational Requirements Composer e Rational Quality Manager. Workload Service fornisce la possibilità di regolare la scalabilità e la gestione delle applicazioni in base a meccanismi automatici definiti da policy, attraverso il **Workload Deployer**, cuore del servizio. Il servizio offerto può essere suddiviso in cinque aree funzionali, di seguito elencate, che possono essere adoperate indipendentemente o in maniera cooperativa.

Application LifeCycle: basata sulla tecnologia IBM Rational, esso comprende tool di sviluppo e di collaborazione. I primi migliorano la velocità di creazione e rilascio delle applicazioni, permettendo agli sviluppatori di ridurre i task di programmazione ripetitivi, rimpiazzandoli con template pre-configurati. I secondi favoriscono le comunicazioni tra gli sviluppatori, permettendo anche agli stakeholder privi di conoscenze tecniche di verificare il codice e tenere traccia dei risultati. Gli obiettivi principali sono di migliorare l'agilità e l'innovazione a livello business, migliorando al tempo stesso la qualità del software e riducendo gli errori.

Application Resources: l'era del cloud ha portato con sé l'idea che l'IT non deve essere necessariamente sviluppato in proprio, ma le aziende possono trarre beneficio da servizi e pacchetti applicativi condivisi senza perdere competitività. Application Resources permette di ridurre i costi e semplificare lo sviluppo di risorse applicative comuni, come servizi di database e di backup, eliminando le spese di manutenzione e funzionamento connesse, fornendo al tempo stesso immediata disponibilità e scalabilità. Esso fornisce un servizio centralizzato e condiviso per usufruire di risorse applicative e middleware comuni, come **database as service** (basato su DB2 Enterprise Server).

Application environments: il servizio facilita e automatizza il rilascio di applicazioni, attraverso un'interfaccia semplificata. L'utente seleziona un pattern preconfigurato per il tipo di applicazione che deve essere rilasciato e seleziona la configurazione adeguata alle specifiche esigenze di performance. Il PaaS crea e pubblica la macchina virtuale allocando le risorse appropriate. Agenti di controllo integrati tengono traccia delle prestazioni, gestendo automaticamente le risorse per mantenere i livelli di prestazione richiesti.

Application Management: per molte aziende la linfa vitale del business è un sistema di pianificazione delle risorse come SAP. Spesso si esita a muovere sistemi così critici nel cloud temendo interruzioni o malfunzionamenti delle operazioni. Per questi business, IBM ha lanciato i servizi di Application Management, focalizzati sul supporto ad applicazioni di terze parti. Il servizio fornisce pattern specifici per l'applicazione SAP e configurazioni che permettono agli utenti di rilasciarne i workload critici direttamente nel cloud.

Integration: per massimizzare la produttività e l'efficienza, gli utilizzatori necessitano di accedere a componenti comuni da altre applicazioni, il che richiede la pubblicazione di numerose interfacce. Sfortunatamente

un'integrazione di questo tipo richiede tempo e risorse impiegati nella stesura del codice opportuno. Le funzionalità di Integration permettono di integrare applicazioni nuove e legacy, senza scrivere codice aggiuntivo, attraverso connettori per software esistenti e template di costruttori.

2.2 Oracle PaaS

La piattaforma PaaS di Oracle [Oracle] è un portfolio di prodotti utilizzabili per costruire applicazioni che possono essere pubblicate come servizi sia su cloud privati che pubblici. La piattaforma è basata sulle tecnologie Oracle grid, oltre che su applicazioni che includono WebLogic Server, Coherence-in-memory datagrid e Jrockit JVM. Dal punto di vista infrastrutturale, il PaaS si appoggia all'offerta IaaS di Oracle, comprendete Oracle Solaris, Oracle Enterprise Linux e Oracle Vm per la virtualizzazione, Sun SPARC e Storage. Sia la piattaforma PaaS che quella IaaS sono gestite da **Oracle Enterprise Manager**, che fornisce sistemi integrati di gestione dell'intero ciclo di sviluppo per applicazioni destinate al cloud o ad uso tradizionale.

Di seguito sono elencati i principali servizi e funzionalità messi a disposizione dalla piattaforma Oracle.

Storage: I prodotti Open Storage di Sun combinano software open-source con hardware industriali standardizzati per fornire una piattaforma di storage aperta e scalabile. La Sun fornisce reti virtuali per computazioni su larga scala attraverso **InfiniBand**, permettendo di creare griglie di computazione di dimensioni notevoli. Gli **Oracle Exadata Storage Server** forniscono caratteristiche di software intelligence, con una particolare affinità con l'Oracle Database.

Server e Sistemi Operativi: Oracle offre diversi server Sun (Netra, Blade Servers, SPARC Enterprise, X64) e supporto a sistemi operativi standard (Solaris, Linux, Windows), che forniscono un ampio spettro di infrastrutture fisiche ottimizzate per la natura virtualizzata e distribuita delle implementazioni cloud.

Server Virtualization: Oracle VM offre supporto sia alle architetture x86 che SPARC, rendendo possibile la pubblicazione delle applicazioni in ambienti eterogenei. Gli utenti possono sfruttare Oracle VM per consolidare server, rilasciare rapidamente software, recuperare rapidamente da fallimenti di sistema e associare capacità a carichi di lavoro.

VM Templates e Assemblies: **Oracle VM Templates** sono immagini di macchine virtuali che contengono software aziendali pre-installati e configurati, che possono essere usati per sviluppare, impacchettare e distribuire applicazioni in favore di una loro rapida pubblicazione. Il concetto di **VM Assemblies** è ad un livello superiore. Le applicazioni aziendali non sono sempre auto-contenute, ma comprendono diverse entità virtualizzate complesse, che richiedono software multi-tier, diverse macchine virtuali, l'uso di middleware differenti. L'assembly, come l'immagine virtuale di un'applicazione, è essenzialmente un file che contiene le immagini di tutte le applicazioni che la costituiscono, insieme a configurazioni, connessioni e sequenze di inizializzazione.

Database e Storage Grid: Oracle Database ha offerto capacità di grid computing dal rilascio di Oracle Database 10g nel 2003. Da allora, Oracle ha continuato a migliorare le capacità di grid computing nell'area clustering con **Oracle Real Application Clusters (RAC)**, nella virtualizzazione e manutenzione dello storage con Automatic Storage Management (ASM) e nell'ambito delle performance dei database con InMemory Database Cache. Tuttavia, nel caso in cui siano necessari servizi più leggeri, sono disponibili database come Oracle Berkeley DB e MySQL, attivamente sviluppati e supportati.

Application Grid: in maniera simile all'architettura grid in Oracle Database e storage, **Oracle Fusion Middleware** supporta robuste funzionalità di grid nel middle tier, con un gruppo di applicazioni chiamate Oracle Application grid. Le tecnologie chiave alla base dell'application grid sono: **Oracle WebLogic Server** come application server principale e cuore della tecnologia Java EE nel middleware Fusion; **Oracle Coherence** che fornisce servizi grid di memoria interna, per oggetti Java, .NET e C++; **Jrockit JVM** che fornisce ambienti di runtime Java veloci e leggeri, sfruttando le abilità di clustering dinamico offerte da WebLogic; **Oracle Tuxedo** per il monitoraggio e la gestione delle transazioni.

SOA e Business Process Management (BPM): Oracle SOA Suite fornisce una base di partenza completa e facile da usare per creare componenti riutilizzabili all'interno del cloud. La rapida creazione di componenti e la loro successiva composizione è supportata dalle caratteristiche drag-and-drop di Jdeveloper e SCA. **Oracle Service Bus** fornisce un modo semplice per rendere i componenti disponibili agli sviluppatori attraverso il PaaS. La possibilità di tracciare i componenti end-to-end le attività di monitoraggio, offerte da **Oracle Business Activity Monitoring**, forniscono metriche di valutazione che supportano sia le attività IT che riguardano il mantenere attivo e funzionale il PaaS, sia quelle di Business interessate agli indicatori di performance aziendale. In aggiunta alle componenti SOA, le aziende potranno includere componenti di business process, gestiti dal framework BPM, offerto come parte del PaaS. **Oracle BPEL Process Manager** fornisce la capacità di creare componenti di processo BPEL, così come di definire asset legacy continuando a supportare la flessibilità, permettendo alle entità interessate di inserire i propri componenti BPEL nei loro workflow.

User Interaction: una libreria centralizzata di componenti UI fornisce ai vari dipartimenti aziendali un punto di partenza importante per la creazione di soluzioni proprietarie, fornendo al tempo stesso alle funzioni centrali IT un elevato livello di controllo sulla consistenza delle applicazioni nell'azienda. Oracle WebCenter Suite fornisce diversi portali e capacità di interazione utente, ideali per creare componenti UI riutilizzabili come parte del PaaS.

Identity Management: Oracle Identity and Access Management Suite fornisce meccanismi per gestire accessi e sicurezza in ambiente PaaS. **Oracle Access Manager** supporta directory condivise e accessi singoli. **Oracle Entitlements Server** fornisce politiche per i controlli degli accessi, rivolte ad ambienti PaaS decentralizzati. **Oracle Identity Manager** permette di automatizzare i processi di aggiunta, rimozione e aggiornamento degli account

utente per applicazioni e directory. **Oracle Identity Federation** fornisce un server multi-protocollo e auto-contenuto con i sistemi di gestione e accesso predefiniti dall'utente già pronti.

2.3 Servizi Amazon PaaS

I servizi offerti da Amazon come PaaS possono essere classificati in base al loro contributo rispetto all'architettura di un'applicazione [AWS].

Application Logic-as-a-Service: la logica delle applicazioni sviluppate viene tipicamente scritta manualmente attraverso linguaggi di programmazione come Ruby, Java, PHP o C#. Ogni linguaggio presenta framework o librerie di supporto. Il servizio offerto da Amazon per ospitare ed eseguire logica applicativa è **Elastic Beanstalk**. Il servizio si concentrava originariamente sull'esecuzione di applicazioni scritte per la Java Virtual Machine all'interno del motore per servlet di Apache Tomcat. Recentemente esso è stato esteso per supportare PHP. Beanstalk gestisce automaticamente le configurazioni necessarie a lanciare i software, oltre che a garantire scalabilità, persistenza, integrazione e sicurezza.

Database-as-a-Service: Amazon offre tre database tra cui scegliere. Il primo ad essere introdotto è stato **SimpleDB**, che rappresenta un modo semplice per immagazzinare informazioni nella forma chiave/valore. La semplicità di questo database è data dal fatto che non fornisce alcuna funzionalità aggiuntiva tipica di altri database, pur fornendo servizi di backup e scalabilità automatica delle risorse. Se invece diventa necessario un database di tipo relazione, è possibile sfruttare il **Relational Database Service (RDS)**. A differenza di SimpleDB, RDS non supporta la scalabilità automatica, ma permette di istanziare, configurare e personalizzare un database di tipo relazionale, che può essere basato su diverse versioni di MySQL e Oracle, specificando le dimensioni (Cpu e memoria) delle macchine su cui esso lavorerà. Il terzo database offerto è dato da **DynamoDB**, considerato di tipo "NoSQL", in quanto non sfrutta SQL per la definizione dei dati o la loro manipolazione. Esso viene visto spesso come un'estensione di SimpleDB.

Caching-as-a-Service: Amazon offre un'implementazione di clustered cache basata su una soluzione open source piuttosto popolare, nota come **Memcached**. Gli utenti possono creare una cache in maniera autonoma utilizzando le API offerte o attraverso il portale di Amazon. Il servizio **Elasticache** di Amazon offre la possibilità di associare il software di caching a differenti tipologie di servizi EC2, consentendo di gestire la scalabilità delle risorse cache in maniera automatica o sulla base di allarmi generati dal servizio stesso.

Integration-as-a-Service: l'offerta PaaS di Amazon prevede due tipologie di servizi di integrazione che gestiscono il disaccoppiamento di sistemi e lo scambio di messaggi. Il **Simple Notification Service (SNS)** fornisce un'implementazione del paradigma pubblicazione/sottoscrizione. Esso gestisce automaticamente la scalabilità, mentre nasconde agli utenti tutti i dettagli relativi alle librerie di messaggi utilizzate per fornire le funzionalità fornite. Inoltre è possibile invocare il servizio tramite SOAP o REST, specificando i protocolli da

utilizzare. SNS non offre garanzie sulla notifica dei messaggi né sui tempi di consegna: si tratta di un servizio “best-effort”. Il **Simple Queue Service (SQS)**, disaccoppia moduli software che desiderano comunicare tra loro utilizzando delle code, permettendo così di gestire componenti che lavorano a velocità differenti. Come per SNS, i messaggi possono essere gestiti tramite SOAP e REST e inoltre possono essere raggruppati. Il servizio non offre code FIFO o con priorità, ma lascia che siano gli utenti a specificare il comportamento della coda che vogliono usare.

2.4 I modelli di esecuzione di Windows Azure

Windows Azure [Redkar et al, 2011] è la piattaforma Microsoft per il cloud pubblico. Questa piattaforma può essere usata in modi differenti a seconda delle esigenze degli sviluppatori, coprendo un ampio range di utilizzo. Uno dei più immediati è la costruzione di applicazioni web che vengono eseguite e memorizzano i propri dati nei data center di Windows Azure, ma la piattaforma può essere usata anche solo per memorizzare i dati che usa un'applicazione eseguita on-premise (ossia al di fuori del cloud pubblico) o per ottenere macchine virtuali utili allo sviluppo e al testing delle applicazioni.

Windows Azure fornisce diversi modelli di esecuzione per le applicazioni. L'esecuzione attraverso macchine virtuali rappresenta essenzialmente l'approccio IaaS di Azure, che permette di selezionare sistemi operativi da un catalogo o di caricare il proprio Virtual Hard Drive. Quella attraverso macchine virtuali è la modalità di esecuzione di applicazione meno trasparente in quanto la gestione dell'ambiente di sviluppo e la responsabilità di amministrazione delle macchine virtuali è lasciata all'utente. Se si vuole sviluppare applicazioni o siti web delegando al provider la cura della parte amministrativa la soluzione è il modello di esecuzione Windows Azure Web Site, che offre sia un portale di gestione sia API. L'esecuzione delle istanze dell'applicazione può essere scalata dinamicamente a richiesta mentre il load balancing delle richieste viene gestito automaticamente dalla piattaforma. Questa soluzione presenta comunque molti limiti e non è adatta per applicazioni che richiedono ad esempio l'installazione di software arbitrari.

Una soluzione che offre una certa flessibilità, pur rimanendo trasparente rispetto ad aspetti affidabilità e amministrazione è **Windows Azure Cloud Service**. Questo modello rappresenta la vera offerta PaaS di Windows Azure permettendo l'esecuzione di applicazioni con supporto alla scalabilità, affidabilità e senza dover curarsi di dettagli amministrativi. In generale un'applicazione sviluppata per la piattaforma Windows Azure è legata a due principali tipi di entità in qualche modo combinate tra loro: istanze di tipo *web role* (destinato principalmente per l'esecuzione di applicazioni Web-based) e *worker role* (progettato per eseguire una mole molto grande di calcoli, ad esempio simulazioni o testing o per eseguire un servizio in background). Un'applicazione di tipo cloud service viene resa disponibile generalmente in un processo a due fasi: prima lo sviluppatore sviluppa e carica l'applicazione con i file necessari sulla piattaforma e successivamente usando il portale mette in esecuzione l'applicazione. Questa divisione tra produzione e messa in scena è

utile durante l'upgrade delle versioni, in quanto la nuova versione viene aggiornata senza interruzione del servizio. Il servizio di monitoraggio della piattaforma stessa è capace di individuare fallimenti non solo dell'hardware ma anche dell'applicazione e rieseguire le istanze dei worker role e web roles che hanno subito dei guasti. La natura PaaS del servizio Azure Cloud Service ha una forte implicazione, ossia che l'intera applicazione deve essere progettata e implementata in modo che possa continuare a funzionare correttamente anche se una qualsiasi istanza di un worker role o web role dovesse fallire. Per far sì che ciò avvenga c'è bisogno che l'applicazione memorizzi esplicitamente il proprio stato nei servizi di data store di Azure, poiché lo stato non può essere salvato persistentemente sulle macchine virtuali su cui i servizi sono eseguiti data la natura stessa del servizio platform. Il rispetto di questo vincolo rende le applicazioni sviluppate facilmente scalabili e resistenti ai fallimenti. La scalabilità delle applicazioni può essere facilmente gestita sia manualmente che settando dei parametri che guidano le operazioni di scaling dei web role e worker role per assecondare il workload, ad esempio l'uso della CPU o lo stato delle code.

Oltre all'esecuzione delle applicazioni, un altro aspetto importante è la fruizione dei dati utili all'applicazione. L'uso di un modello di esecuzione basato su IaaS permette l'installazione e l'uso di un qualsiasi tipo di sistema, dai database relazionali fino alle soluzioni NoSQL, ma anche in questo caso, per facilitare il management del database Windows Azure offre delle soluzioni di tipo platform; in particolare SQL Azure per offrire storage di tipo relazionali, Tables per database NoSQL scalabili e Blob per la memorizzazione binaria non strutturata. Queste funzionalità sono offerte dalla piattaforma Azure come servizi e possono essere utilizzate sia da applicazioni in esecuzione sulla piattaforma stessa, sia da applicazioni che girano in diversi datacenter.

Un altro aspetto fondamentale per l'esecuzione delle applicazioni è l'interazione tra diversi componenti che avviene attraverso servizi di messaging. Windows Azure offre due meccanismi di messaging, uno più semplice basato su code (**Azure Queue**) con un paradigma di tipo publisher/subscriber e uno più complesso, chiamato **Service Bus** che permette di implementare oltre ai meccanismi one-to-one e publish-subscribe anche un sistema di sottoscrizione a topic, che permette di implementare la comunicazione one-to-many tra diverse applicazioni oltre ad offrire altre funzionalità relative alla sicurezza della comunicazione.

3.Cloud Patterns

Nel panorama attuale del cloud computing ci sono così tante alternative offerte da diversi vendor che risulta difficile decidere quale soluzione o servizio usare nel contesto di una specifica applicazione. L'iniziativa di promuovere i cloud pattern nasce proprio dall'esigenza di fornire la descrizione di soluzioni a problemi ricorrenti nella definizione di architetture per applicazioni cloud. I diversi cataloghi che si stanno sviluppando, data la loro diversa natura, offrono contenuti a diversi livelli di dettaglio ed astrazione, con scopi diversi. Alcuni di essi, come quelli elencati in [AWSP] e [WAP] sono legati a una piattaforma cloud specifica e per questo sono più dettagliati in termini di componenti da

usare per implementare il singolo pattern, ma la soluzione offerta è dipendente dalla piattaforma stessa.

Invece, i cataloghi che provengono da studi accademici come quelli definiti in [Fehling et al, 2011], non legati all'industria, descrivono funzionalità generali proponendo un modello architetturale meno vincolato alle specifiche soluzioni.

Secondo una definizione fornita dagli stessi esperti di IBM [IBM2], un pattern è una collezione di elementi che descrivono una soluzione software completa, che può coinvolgere diversi sistemi interconnessi o una singola entità. Tutta la conoscenza necessaria per creare, configurare e supportare ogni aspetto della soluzione è già inclusa nel pattern. In IBM un virtual application pattern rappresenta una collezione di componenti applicativi, politiche di comportamento e relativi collegamenti. Un servizio di workload che sfrutti questo genere di applicazione può costruire automaticamente l'infrastruttura e le risorse di middleware necessarie a far funzionare e gestire l'applicazione virtuale, che diventa dunque un'istanza del pattern. L'uso dei virtual application pattern avviene attraverso i servizi offerti dalla piattaforma Workload Deployer e essere sfruttati sia attraverso l'IBM PureApplication System (cloud privato), sia tramite lo Smart Cloud Application Workload Service (cloud pubblico), basato sull'offerta IaaS di Smart Cloud Enterprise.

Oracle Fusion Applications User Experience (UX) Design Patterns [Oraclep] è un set contenente più di 150 modelli d'interazione tra utenti e flussi di dati, applicati alle Oracle Fusion Applications. I vari pattern di Oracle sono esposti secondo un template comune, che riassume: il problema o l'obiettivo affrontato, il contesto in cui il pattern viene applicato, una o più soluzioni che risolvono il problema. In particolare vengono mostrati esempi di progettazione che mostrano possibili implementazioni del pattern, generalmente facendo riferimento al middleware Fusion.

Gli AWS Cloud Design Pattern [AWSP] sono una collezione di soluzioni e idee di design che utilizzano le tecnologie cloud AWS per risolvere problemi di progettazione comuni. Per ogni pattern è presente una descrizione che indica le problematiche che hanno spinto alla creazione del pattern e quali difficoltà possono essere risolte attraverso la sua implementazione. Inoltre è descritta la parte implementativa, in cui viene descritta passo passo la procedura di applicazione del pattern in ambito AWS, specificando in maniera precisa quali componenti devono essere utilizzate e come ciò deve avvenire.

Quello fornito per Windows Azure [WAP] non può essere considerato come un catalogo vero e proprio, quanto piuttosto come una raccolta di link ad articoli o pubblicazioni digitali in cui sono descritti alcuni cloud pattern o sono illustrate delle tecniche, basate sui servizi offerti dalla piattaforma Azure stessa, per la gestione di particolari funzionalità in ambito cloud. I pattern presentati in questo catalogo di Microsoft sono tutti strettamente legati alle implementazioni possibili sulla piattaforma Azure, al punto che la loro descrizione avviene presentando direttamente il servizio, la funzionalità o il tool che risolvono il problema affrontato dal pattern stesso.

La visione dei pattern dei vari provider è quindi differente: per alcuni come IBM è limitata essenzialmente alla personalizzazione tramite politiche e vincoli di soluzioni preconfigurate ed esistenti; i pattern Oracle in generale non sono

propriamente pensati per il cloud, ma provengono da raccolte di soluzioni a pattern ricorrenti nelle Oracle Fusion Applications; Amazon offre un catalogo di pattern che anche essendo espressi in termini di servizi proprietari risultano facilmente generalizzabili; infine Microsoft descrive i pattern come problemi che però sono già risolvibili dalla piattaforma in modo trasparente.

4. Conclusioni e sviluppi futuri

Dalla breve panoramica dei servizi PaaS fornita si evince quanto siano vaste e variegate le offerte dei diversi cloud vendor. Tale diversificazione si riflette nelle profonde differenze esistenti tra i diversi cloud pattern che, se in alcuni casi sono neutrali rispetto alle piattaforme esistenti, ponendosi ad un livello di astrazione piuttosto elevato, in altri sono profondamente legati ad esse. Lavori futuri potrebbero concentrarsi sulla formalizzazione dei cloud pattern sia proprietari che non proprietari, oltre che sulla definizione di pattern di carattere generale in grado di accogliere entrambe le categorie, in modo da rendere facilmente riconducibile a soluzioni particolari problemi di carattere generale.

5. Ringraziamenti

Questo lavoro è stato supportato dal Progetto di Rilevante Interesse Scientifico e Tecnologico anno 2009 dal titolo "Fruizione assistita e context aware di siti archeologici complessi mediante dispositivi mobili".

Bibliografia

- [AWS] An Overview of the Amazon PaaS. Transcend Computing. 2012
- [AWSP] "Aws cloud design patterns", <http://en.cloud.designpattern.org>.
- [Fehling et al, 2011] Fehling, C.; Leymann, F.; Mietzner, R.; Schupeck, W. A Collection of Patterns for Cloud Types, Cloud Service Models, and Cloud-based Application Architectures. Technical Report No. 2011/05; University of Stuttgart: Stuttgart, Germany, 2011.
- [IBMP] IBM Design Patterns, www-01.ibm.com/software/ucd/designpatterns.html
- [IBMP2] <http://expertintegratedsystemsblog.com/index.php/2012/07/getting-back-to-the-basics-what-is-a-pattern/>
- [Oracle] Oracle, "Platform-as-a-Service Private Cloud with Oracle Fusion Middleware", Oracle White Paper, <http://www.oracle.com/us/036500.pdf>
- [Oraclep] Oracle Fusion Applications User Experience Patterns and Guidelines, <http://www.oracle.com/webfolder/ux/applications/fusiongps/patterns/index.htm>
- [Redkar et al, 2011] Redkar, Tejaswi, Guidici. Windows Azure Platform. Apress, 2011.
- [SCAS] <http://www.ibm.com/cloud-computing/us/en/paas.html>
- [WAP] Windows Azure Architecture and Patterns, <http://www.windowsazure.com/en-us/develop/net/architecture/>

Digital Citizenship and Social Responsibility of Computer Professionals

¹Norberto Patrignani, ²Marco De Marco

¹*Catholic University of Milano
Largo Gemelli 1, 20123 Milano
norberto.patrignani@unicatt.it*

²*Università Guglielmo Marconi
Via Plinio 44, 00193 Roma
m.demarco@unimarconi.it*

Abstract. *Digital revolution is taking us into the Information Society so quickly that the entire Information and Communication Technology (ICT) world has no time to reflect on new social and ethical issues. In particular this dramatic change put computer professionals in a position full of opportunities but also with new responsibilities. Computer professionals and Information Systems managers have in front of them a "rainbow" of risks and opportunities, they face a collection of ethical dilemmas and social issues that requires a deep reflection and debate among all stakeholders: universities, ICT professional organizations, ICT industry, policy makers, users and society at large. A code of ethics for computer professionals could be a good starting point for defining a collection of guidelines.*

Keywords: Social responsibility, Computer professionals, Computer ethics, Code of ethics.

1. Introduction

In this paper we would like to address one of the themes of the AICA 2013 National Congress, "*digital citizenship*", from the particular point of view of computer professionals and Information Systems managers. They are the real people in charge for the design, development and governance of ICT infrastructures at the basis of our Information Society, they are the real providers of this "*digital citizenship*". How are changed their roles and responsibilities among the years? What are the most important changes from *mainframe era* to Distributed Computing era and finally to Cloud Computing (and Big Data) era? What kind of new social issues and ethical dilemmas are they facing everyday? What kind of instruments have they for facing these

issues? What kind of relationship should be in place between "*digital citizens*" and the people that are in charge for ensuring that the "*digital platforms*" run smoothly, respect basic human rights, and minimize the impact on the environment? What are the relationships between the computer professionals (at all levels) inside an organization and the Corporate Social Responsibility strategy of the organization itself? What kind of support can computer professionals societies (like AICA in Italy, British Computing Society, BCS in UK, etc.) provide to their members? What contribution could provide a *Code of Ethics* for computer professionals? In this paper we would like to set up the stage for finding some possible answers to these questions.

2. From Mainframe to Cloud Computing and Big Data

Computing evolution can be shortly described with three main eras: the (centralized) *Mainframe Computing era*, the *Distributed (Personal) Computing era*, and the *Cloud Computing era*. In each era, computer professionals and Information Systems managers played different roles and faced many responsibilities.

In the (centralized) mainframe computing era (1950s-1970s) the basic technologies were "dumb" terminals (with just input, output and network) connected to big computer (mainframes) inside machine rooms. Any application was designed to run on mainframe platforms. The main issues in front of computer professionals were to properly design the hardware and software platforms to run inside the computer room. The main problems for Information Systems organizations was to control the access to computer rooms (physical and logical) and to guarantee the reliability of the systems. The end users had little or no role in this scenario, they were just "consuming" ICT applications entering their inputs on keyboards or punched cards and visualizing output on printers or screens [Williams, 1997].

In the distributed (personal) computing era (1970s-2010s) the novelty was based on the processing power available to any department and user, and the networking capabilities extended across countries. The computer professionals were involved into the design of "distributed computing" infrastructures and "client-server" applications. The Information Systems role grew in complexity due to the management issues related to networked domains and applications running also at department and user levels [Couloris et al., 2011].

In the Cloud Computing era (2010s-now) the novelty is the re-centralization of computing power and storage ("*in the cloud*"), the decreasing role of end-users' devices (with the local processing and storage capabilities mostly unused), and the network extended at global scale. Computer professionals develop Web-based applications (or simply *apps*) to be "consumed" on touchscreen devices with little or no computing autonomy.

In this scenario the *cloud ICT infrastructure* becomes global with the following characteristics: it is (broadband) *network-based*, the servers are shared-platforms (*multitenancy*), the processing power and storage capabilities are easily scalable (*scalability* and *elasticity*), all resources consumption are

measured (e.g. for billing purposes) and the users can allocate resources *on-demand* in a *self-service* way [NIST, 2013].

Then there are different Service Models in this scenario:

- *Infrastructure as a Service* (IaaS) where Cloud Providers just rent IT physical components to Cloud Users (in this case Cloud Users organizations have the control of the Applications and shares the Virtual Machines with the Provider, whilst the Provider controls Server, Storage and Network levels);
- *Platform as a Service* (PaaS) where the Cloud environment is used mainly for developing and testing very large applications (in this case Cloud Users shares the Applications and the Services with the Provider, whilst the Provider has the control of Server, Storage and Network levels);
- *Software as a Service* (SaaS) where the entire application is on the Cloud Provider side (in this case Cloud User organizations have no control on any level, whilst the Cloud Provider controls all the Applications, Services, Server, Storage and Network levels) [Mather et al., 2009].

Of course many organizations are implementing Cloud Computing solutions "inside" their domain for using ICT resources more efficiently. In these cases the term used is "Private Cloud", an evolution of the "intranet" concept [Nuttgens et al., 2011], there are no data crossing public networks. In these cases, *"data produced and subsequently distributed within an organization is not only a strategic resource to achieve/maintain a competitive advantage but it is also used as a tool to develop and expand the ability of an enterprise to abruptly respond to unexpected generated by the external environment"* [Zardini et al., 2011; p. 390].

In the standard (Public) Cloud solutions, the data of organizations crosses public networks and are stored "outside" the administration border controlled by Information Systems managers. In these cases, often referred as "Public Cloud" there are many interesting and difficult issues that arise. Information Systems and Chief Information Officers (CIOs) roles are now very critical since the pressure from Chief Financial Officers (CFOs) about moving ICT services towards the *pay-per-use* model of Cloud Computing is growing: with no more investments needed, the move from CAPEX (Capital Expenditures) model to OPEX (Operational Expenditures) model for ICT services is becoming very attractive for CFOs [Patrignani and Kavathazopoulos, 2013].

The risk for CIOs and IT services is to become just *"service brokers"*: to chase the best offers from "Cloud Providers" market and offer them to the organization's employees [Ricciardi and De Marco, 2012]. What will happen to all workers inside the IT organizations? Of course all the activities related to design, develop, deploy, maintain all IT services (and the computer room itself) will disappear as all those tasks and responsibilities will be shifted outside to Cloud Providers. Even the task of managing the internal workplaces, personal computers, and laptops is disappearing, since many organizations are encouraging employees to use their daily personal device also at work, it is the so called phenomenon of BYOD, Bring Your Own device [Casey, 2013].

But there are also many other important issues related to Cloud Computing that CIOs have to address like:

- *governance* (in particular in the Service Model like SaaS, the Cloud Provider is delegated to control the entire "stack" of ICT layers, Application, Services, Server, Storage, and Network);
- *de-perimeterisation* (the traditional network boundaries between organization domains will disappear, information storage and processing are outside the control of the organization that still keeps only input, output and network, a kind of "return" to mainframe era);
- *contractual obligations* (there will be the case where the Cloud Provider itself is not the owner of the resources it is "selling"; there will be "ICT brokers", in these cases who will be responsible of what?);
- *problems of many-hands* (when there are several administrators - "*many hands*" - that control mission critical infrastructures, what will happen if, for example, a cloud administrator stops a service for maintenance? There will be cases where both sides – Providers and Users of the cloud – must agree with a kind of "*four-hands-authorization*");
- *risk management plans* (in case of an incident – ICT people know that "*something can always go wrong*" – will the cloud be "*traceable*"? It will emerge the need for storing events in encrypted secure logs, time-stamped and digitally signed, and the need to agree on *joint risk-management-plans*);
- *compliance* (some Cloud Users will need, for compliance purposes, to know the physical location of their data);
- *open market* (what will be the data formats? What kind of standards will be adopted for guaranteeing portability? Can Cloud Users easily change Cloud Provider?).

Computer professionals, Information Systems managers, and CIOs has to deeply reflect on all these issues before signing contracts with Cloud Providers. But there are also many others social and ethical issues related to the management of ICT infrastructures in the current Information Society scenario. Here there are some examples:

- *e-democracy* (what is the correct definition of digital citizenship? What are the rights and duties of citizens online, what kind of decisions are we going to make online?) [Ricciardi and Lombardi, 2010];
- *e-Inclusion* (are we providing the proper interfaces also to elderly people or people with disabilities?);
- *digital divide* (is the access to ICT infrastructures available everywhere? For example, in a "*smart city*" environment, who and how will have access to these kind of "*public digital services*"? Is all the data collected by the smart city applications provided to other organizations in an "*open access*" way?);
- what will be the impact of ICT on workplaces, on information systems users?
- how will change the definition and management of intellectual property in the digital world?
- what will be the approach to *privacy* issues when most of the data and applications will be "outside" the organization? [Patrignani and DeMarco, 2012];
- what kind of reliability of ICT infrastructures and applications will be guaranteed in the Cloud computing scenario?

- what kind of initiatives and procedures will CIOs put in place in order to minimize the environmental impact of ICT? [Patrignani, 2009];
- how will the organization be prepared to face the "*digital tsunami*" of BigData, where Billions of Gigabytes (Exabytes) will have to be processed and stored?

3. ICT landscape: a network with many stakeholders

In the previous part of the paper we have discussed the complex scenarios in front to computer professionals at all levels. These scenarios include several social and ethical issues. Now we introduce some tools or instruments that could help these people in facing these problems that cross the traditional technology borders and reach in many aspects the "*digital citizens*" and the society. Indeed the traditional strategies of the majority of organizations are concentrated only on business opportunities and on short-term profit goals. ICT was always considered just as a "*business enabler*", until 1990s. Since then, with the growing role of *e-business* and online engagement, transaction, fulfillment and service management, ICT is becoming *the (core) business* in itself: without ICT business activities are simply impossible [Rossignoli et al. 2009]. Also, the growing need for organizations to have a clear Corporate Social Responsibility strategy raised the need of a kind of alignment between these strategies and the need to address the social and ethical issues of computing, a clear *computer ethics* strategy. It is now clear how the two domains should be strictly aligned in order of transparently distributing roles and responsibilities inside the organization, and for a complete alignment between *Business Ethics* and *Computer Ethics* strategies [Chartier and Plante, 2013].

Nevertheless computer professionals need some specific tools for facing the immense social and ethical issues and the "*rainbow*" of risks and opportunities in front of them.

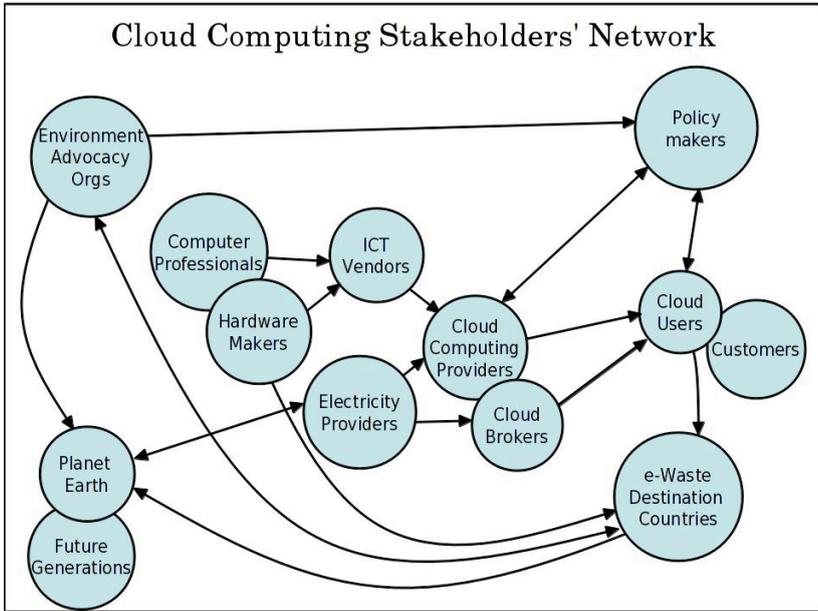


Fig.1 – Example of ICT Stakeholders Network

The *ICT stakeholders network* has been proposed as a powerful tool for reasoning and dialoguing about these difficult choices (see fig.1 – Example of ICT Stakeholders Network – applied to the recent Cloud Computing scenario) [Patrignani and Kavathatzopoulos, 2013]. Within this network, it is possible to identify all the stakeholders and relationships related to a specific ICT scenario. The simple construction of this network is already a good help for identifying *conflicts* between stakeholders and missing relationships usually not considered into the customary ICT landscape. In some way it could help also in *ethical decision making* [Laaksoharju, 2010].

Computer professionals and people usually involved just in "technical" decisions are rarely exposed to the concept of artifacts (including information systems), as "*socio-technical systems*", or to the concept that artifacts embed values [Johnson, 1985]. It is difficult for them to see ICT systems as artifacts that at development-time incorporated the values of the designers, or to see systems as a result of a *Value Sensitive Design* [Friedman, 1996].

This is the main reason of our next proposal: in order to really face their social and ethical responsibilities, as "*digital citizenship providers*", computer professional, or in general, ICT people involved along the entire ICT value-chain need to introduce a reflection inside their professional organizations, this reflection will enrich their "technical" skills with some deontological competences that may prepare the road towards a *code of ethics for computer professionals*.

4. Towards a Code of Ethics?

Around the world many computer professionals organizations are providing their members with resources, events, and working groups related to the ethical issues of ICT. One of the most established ones is the Ethics Group inside British Computing Society (BCS) in the UK, one of the oldest ICT professional organizations in the world. The BCS Ethics Group has a challenging responsibility: "... *is responsible for promoting awareness and engagement with the ethical issues associated with the advancement of Information Technology science and practice and ensuring that Ethics is fully embedded in everything the Institute says and does*" [BCS-a, 2013]. Along the years, they have also defined a "*BCS Code of Conduct*" for providing support to the people in the field [BCS-b, 2013]. One of the oldest Code of Ethics in ICT is the one defined in the 1990s by the Association for Computing Machinery (ACM) [ACM, 2013].

Maybe that these "codes" will not be able to provide answers to any possible question or ethical issue one can face, but for sure they will provide useful guidelines and suggestions in order to *be prepared*, in order to improve one status from "*technician*" to a real *professional*.

Also in Italy, the "*Italian Computing Society*" (AICA, Associazione Italiana per l'Informatica ed il Calcolo Automatico) has recently setup a Working Group on ICT and Ethics (Progetto ETIC, Etica e Tecnologie dell'Informazione e della Comunicazione), coordinated by professor Ivo De Lotto of the University of Pavia. There are several ongoing activities within this working group and many of them are preparing the background for a discussion about a *code of ethics for computer professionals* also in Italy [AICA, 2013].

5. Conclusions

The evolution of ICT happened so quickly that gave little time to reflections on different points of view with respect to the common vision of the computer as a technology that can solve most of the problems of society. The risk that we face is to be too fascinated by the wave of technological innovation losing sight of the sea of changes and controversial issues below the sea surface. Yet the widespread diffusion of computers in society, and the indispensable role played by networks of computers in almost all economic activities, induces a series of ethical reflections, not just professionals or experts in ICT, but also to end users and the entire society.

In conclusion we would like to propose some simple recommendations:

- for Universities: to introduce interdisciplinary courses in Computer Science and Computer Engineering courses that could prepare the future generations of computer professionals to face also the social and ethical issues related to ICT;
- for Policy makers: encourage and promote public discussions for decisions regarding the use of ICT in critical systems for the society (e.g. "*smart cities*"), encourage the use of technology for improving quality of life, *well-being* and *well-living* of human beings;

- for ICT industry: define clear Corporate Social Responsibility strategies that involve the analysis of the entire stakeholders network of ICT market;
- for computer professionals organizations: inform the public about the intrinsic limits of ICT systems reliability, question the presupposition that technology can (alone) solve all social and political problems, be involved in national and international debates about the (social, ethical, legal) professional issues related to ICT, and design systems in order to minimize the environmental impact of ICT.

With this paper we would like to support a reflection, in particular among computer professionals and people involved in the ICT value-chain, about their social responsibility as main providers of the "digital citizenship" to the society. We would also like to provide a contribution to the preparation of people with complex skills and knowledge, people that are not just ICT experts, but people that are also able to understand and make right evaluations about social and ethical implications of ICT.

Bibliography

ACM, ACM Code of Ethics and Professional Conduct, adopted d by ACM Council 16 October 92, available at: <http://www.acm.org/about/code-of-ethics> (accessed 12 June, 2013).

AICA, Progetto ETIC, Etica e Tecnologie dell'Informazione e della Comunicazione, available at: <http://www.aicanet.it/attivita/gruppi-di-progetto/progetto-etic> (accessed 10 July 2013).

BCS-a, The Chartered Institute for IT - Enabling the Information Society - Ethics Group, available at: <http://www.bcs.org/category/8620> (accessed 12 July 2013).

BCS-b, BCS Code of Conduct, available at: <http://www.bcs.org/category/6030> (accessed 12 June 2013).

Casey K., Six Risks Your BYOD Policy Must Address, Information Week, 19 November 2012.

Chartier A., Plante B., IS/IT ethical issues as a Corporate Social Responsibility: revisiting strategic business planning under the lens of IS/IT ethical preoccupations, in Ward Bynum T., Fleishmann W., Gerdes A., Moldrup Nielsen G., Rogerson S. (eds.) The Possibilities of ICT, Proceedings of 13th International Conference ETHICOMP2013, University of South Denmark, 2013.

Coulouris, G., Dollimore J., Kindberg T., Blair G. (2011). Distributed Systems: Concepts and Design (5th Edition), Addison-Wesley, Boston, 2011.

Johnson, D.G., *Computer Ethics*, 1st Edition 1985, 4th Edition, Pearson International Edition, Prentice Hall, 2009.

Friedman, B., Value Sensitive Design, *Interactions*, November/December 1996.

Laaksoharju M., Let us be philosophers! Computerized support for ethical decision making, Department of Information Technology, Uppsala University, Uppsala, 2010.

Mather T., Kumaraswamy S., Latif S., *Cloud Security and Privacy - An Enterprise Perspective on Risks and Compliance*, O'Really Media, 2009.

NIST (National Institute of Standards and Technology) Cloud Computing Program 2013, Available at: <http://www.nist.gov/itl/cloud/> (accessed 1 July 2013).

Nuttgens M., Gadatsch A., Kautz K., Schirmer I., Blinn N., (eds), *Governance and Sustainability of Information Systems - Managing the transfer and diffusion of IT*, Proceedings of the IFIP WG 8.6 International Working Conference, Hamburg, Germany, September 2011.

Patrignani N., *Computer Ethics. Un quadro concettuale*, Mondo Digitale, n.3, Settembre 2009.

Patrignani N., DeMarco M., The road to a responsible and sustainable e-Business, Proceedings of the International Conference on e-Business, ICE-B2012, Rome, July 2012.

Patrignani N., Kavathazopoulos I., The Brave New World of Socio-Technical Systems: Cloud Computing, in Ward Bynum T., Fleishmann W., Gerdes A., Moldrup Nielsen G., Rogerson S. (eds.) *The Possibilities of ICT*, Proceedings of 13th International Conference ETHICOMP2013, University of South Denmark, 2013.

Ricciardi F., Lombardi P., Widening the Disciplinary Scope of eParticipation. Reflections after a Research on Tourism and Cultural Heritage. In: Tambouris E., Macintosh A., Glassey O. (eds.), *Electronic Participation. Second IFIP International Conference, ePart 2010*. Lausanne, Switzerland. *Lecture Notes in Computer Science*, Springer, 2010.

Ricciardi, F., De Marco, M., The challenge of Service Oriented performances for Chief Information Officers, in Snene M. (ed.), *Exploring Service Science*, Third International Conference, IESS 2012, Geneva, Switzerland. *Lecture Notes in Business Information Processing*, Springer, 2012.

Rossignoli, C., Carugati, A., Mola, L., The strategic mediator: a paradoxical role for a collaborative e-marketplace, *Electronic Markets*, 19(1), 55-66, 2009.

Williams M.R., *A History of Computing Technology*, 2nd Edition, IEEE Computer Society Press, Los Alamitos, CA, 1997.

Zardini, A., Mola, L., Vom Brocke, J., Rossignoli, C., The Role of ECM and its Contribution in Decision-Making Processes, *Journal of Decision Systems*, 19, 4, 389-406, 2010.

L'“arcobaleno dei diritti della cittadinanza digitale” alla prova

Andrea Trentini, Fiorella DeCindio
Dip.to Informatica - Università degli Studi di Milano
Via Comelico, 39 - 20135 MILANO - ITALY
(andrea.trentini|fiorella.decindio)@unimi.it

We live in a society shaped by information and communication technologies, a continuous interplay between what happens in the physical world and what happens online. This inevitably extends the citizenship concept which becomes digital, where rights and obligations are properly declined to meet technology opportunities. These opportunities, however, challenge the very idea of citizenship and the exercise of underlying rights. The authors, along with Leonardo Sonnante, proposed a layerization called “The rainbow of digital citizenship rights” [De Cindio et al., 2012] to slice all aspects of digital citizenship in conceptual levels. The framework consists of abstraction layers spanning from basic network access up to the highest “right to active involvement in policy-making.” This article describes the framework first application. The authors, with their students (“Digital Citizenship and Technocivism” course) used it to analyze the “Public consultation on the fundamental principles of the Internet” promoted by the Italian Ministry of Education.

Keywords: digital citizenship, e-democracy, e-participation, open government, open data

1. L'arcobaleno dei diritti digitali

L'“arcobaleno dei diritti digitali” [De Cindio et al., 2012] è un framework concettuale che prende ispirazione dai lavori di Clement e Shade [Clement and Shade, 2000] e Caddy e Vergez [Caddy and Vergez, 2001], declinando i colori dell'arcobaleno secondo livelli di diritto “digitale” via via crescenti:

LIVELLO 0: diritto all'accesso alla rete (the net)

LIVELLO 1: diritto all'accesso al servizio universale (access)

LIVELLO 2: diritto ad una educazione consapevole (education)

LIVELLO 3: diritto a usufruire di servizi online, pubblici e privati (e-services)

LIVELLO 4: diritto alla trasparenza (be informed, transparency & opendata)

LIVELLO 5: diritto a informare (inform: content provided by users)

LIVELLO 6: diritto ad essere ascoltati e consultati (consultation)

LIVELLO 7: diritto al coinvolgimento attivo (active participation)

L'idea, e anche il numero, dei livelli non può non evocare agli informatici il modello OSI (Open Systems Interconnection, [Zimmermann, 1980]) per le architetture di rete. La similitudine corrisponde al fatto che se non viene garantito il diritto di "cittadinanza digitale" a livello N, il livello N+1 si fonda su basi precarie. Per la descrizione accurata dei livelli rimandiamo i lettori all'articolo [De Cindio et al., 2012] su Mondo Digitale.

2. La consultazione del MIUR sui Principi Generali di Internet

In concomitanza con l'Internet Governance Forum (IGF) Italia, in programma a Torino nell'ottobre 2012 in preparazione di quello previsto per il mese di novembre a Baku, il Ministero dell'Istruzione, Università e Ricerca (MIUR) ha implementato una consultazione pubblica, aperta a tutti i cittadini, sul tema della governance di Internet. La consultazione, ora chiusa, ma ancora consultabile al sito: <http://discussione pubblica.ideascale.com>, è stata gestita con Ideascale, un software proprietario di idea gathering, e articolata in un preambolo e cinque aree tematiche, che riprendevano quelle presenti nel documento preparatorio "La posizione italiana sui principi fondamentali di Internet" (<http://www.governo.it/Notizie/Ministeri/dettaglio.asp?d=69257>).

Le cinque aree tematiche riguardavano: a) Principi Generali, che definiscono le caratteristiche principali dell'infrastruttura; b) Cittadinanza in rete; c) Consumatori e utenti della rete, cioè le questioni di competenze e identità digitali, riservatezza e gestione dei dati personali; d) Produzione e circolazione dei contenuti; e) Sicurezza in rete. La consultazione è rimasta aperta dal 18 settembre 2012 al 1° novembre, per una durata complessiva di 45 giorni; ha raccolto 159 proposte (e commenti alle stesse) da un bacino di utenti registrati al sito di 746 persone su tutta la popolazione italiana, che non costituisce un campione statisticamente significativo. Oltre alla bassa partecipazione di cittadini (che va comunque pesata rispetto alla scarsa comunicazione che la consultazione ha avuto e alla tematica settoriale) l'elaborato di Cuculo e Rasente [Cuculo and Rasente, 2013] (vedasi Sezione 3) ha evidenziato che se da un lato i gestori dichiaravano una certa "trasparenza di metodo": *"Questa consultazione intende raccogliere i contributi dei cittadini sul tema della governance di Internet al fine di arricchire e migliorare il documento che riassume la posizione italiana sui principi fondamentali di Internet in vista del prossimo Internet Governance Forum (IGF)." [...] "I dati e le proposte inviate verranno raccolti ed analizzati al fine di individuare i temi emersi con maggiore frequenza e le indicazioni più rilevanti. Questi verranno utilizzati per integrare i documenti ufficiali, successivamente pubblicati."* nulla era poi detto sui modi, sui tempi e sugli attori responsabili dell'utilizzo effettivo dei suggerimenti forniti dal "processo partecipativo" messo in atto. In altre parole il "contratto sociale" o "patto partecipativo" implicito nella consultazione era molto debole. Ed infatti, nonostante l'impegno dichiarato nella precedente citazione, nessuna analisi delle proposte raccolte, né il documento programmatico modificato sono stati poi pubblicati. Pur con questi limiti, le proposte raccolte forniscono comunque un'immagine significativa dei temi maggiormente all'attenzione degli utenti di Internet, come emerge dall'analisi che presentiamo nei paragrafi che seguono.

3. L'applicazione del framework e categorizzazione proposte

Per validare il framework serviva un caso di applicazione che potesse metterne alla prova tanto l'organizzazione in livelli a stratificazione quanto la completezza. L'occasione è venuta proprio con la “Consultazione pubblica sui principi fondamentali di Internet” proposta dal MIUR (Ministero dell'Istruzione, Università e Ricerca) durante il governo Monti. Una delle sezioni in cui la consultazione era articolata era infatti dedicata alla “Cittadinanza in rete”. Abbiamo quindi deciso di “cogliere la palla al balzo” e proposto un esercizio agli studenti del corso di Cittadinanza Digitale e Tecnocivismo (<http://tecnocivismo.di.unimi.it>), tenuto dagli autori del presente articolo. Durante il corso discutiamo e analizziamo le reciproche influenze, positive e negative, tra la tecnologia e l'essere cittadini al fine di sensibilizzare gli studenti di informatica verso quegli aspetti etici, sociali e politici delle tecnologie dell'informazione che restano a margine dei corsi di studi tradizionali. A fine corso, come lavoro d'esame, viene concordato un tema che gli studenti devono approfondire con una relazione scritta e una presentazione. Per questo articolo ci basiamo sull'elaborato preparato da Vittorio Cuculo e Sonia Rasente [Cuculo and Rasente, 2013] (disponibile sul sito del corso): hanno esaminato le proposte inserite sul sito della consultazione e le hanno classificate secondo i livelli dell'arcobaleno per fare un quadro della sensibilità tecnologico-politica dei partecipanti alla consultazione.

In questa sede riassumeremo le principali considerazioni scaturite dal lavoro di Cuculo e Rasente, ma in particolare ci concentreremo sul posizionamento delle proposte nei livelli dell'arcobaleno. Per valutare l'adeguatezza del framework proposto elencheremo tutte le proposte raccolte sul sito e ne commenteremo l'assegnazione ai vari livelli, analizzando soprattutto la casistica delle proposte che non riusciremo a catalogare per capire se e come eventualmente fare opera di fine tuning (aggiunta o rimozione di livelli, etc.) del framework.

3.1 Metodo di indagine

Come ricordato più sopra, la consultazione è stata svolta utilizzando Ideascale, un software proprietario di idea gathering, che tuttavia permette, autenticandosi, l'accesso al database delle proposte attraverso delle API REST (Application Programming Interface, REpresentational State Transfer) documentate (<http://support.ideascale.com/customer/portal/articles/1001563-ideascale-rest-api>), anche se questa funzione non è molto pubblicizzata per cui gli studenti l'hanno scoperta solo a lavoro inoltrato. L'analisi da loro fatta in [Cuculo and Rasente, 2013] e quella che segue sono basate su questo database, che mette a disposizione per ogni proposta inserita un IDentificatore numerico progressivo e una stringa di testo con il TITOLO della proposta, qui di seguito riportato rispettando la forma originale, inclusi errori, maiuscoli, etc. Gli studenti hanno passato in rassegna le proposte associando ciascuna ad un livello dell'arcobaleno (nel seguito indicato con $L <n>$), identificando quelle per le quali una assegnazione non era possibile, identificate con NA (“livello non

assegnato”). Per il presente articolo abbiamo fatto una revisione delle assegnazioni, come segue:

- in caso di riesame della proposta (rispetto all’assegnazione degli studenti) abbiamo corretto con la notazione $[NA|L<n>] \rightarrow L<m>$
- in caso di proposte troppo generiche, oppure palesemente fuori tema o comunque non rilevanti rispetto all’oggetto di questo lavoro, le abbiamo trascurate identificandole con la notazione *NA skip*
- in caso di proposte trasversali, che toccano più livelli, la non assegnazione potrebbe segnalare un caso degno di attenzione per la validazione del framework, ergo le abbiamo marcate con *NA** e discusse nel seguito
- le proposte che si pongono ad un “meta” livello (relative al processo della consultazione) sono invece raggruppate in fondo alla sezione

In sintesi, il formato con cui ogni proposta è elencata è il seguente:

<IDproposta>: “*<titolo originario>*” [*,commenti*]

<L0|...|L7|NA [tag]> → [*<nuovo>*]

[*eventuali citazioni (indentate e in corsivo) dalla proposta*]

3.2 Analisi delle proposte

Le assegnazioni senza ulteriori commenti sono state raccolte in “gruppi” perchè non rilevanti rispetto alla validazione del framework. In tal caso ciascuna delle proposte raggruppate è denotata dalla tripla (ID: titolo, livello). Va detto infine che gli ID da 1 a 4, 48, 50, 56, 59, 60, 70, 147, 149, 155-161 mancano dal database, mentre ID 5 è duplicato.

5: “Limiti della neutralità in rete mobile” _____ L0

6: “Obiezioni di Internet Society (Europa) alla proposta ETNO”, ampia e trasversale, copre molti livelli, dichiarazione di principio _____ NA skip

(gruppo) 7: “Dati della pubblica amministrazione”, L4 - 8: “Protezione soggetti deboli”, L2

9: “Tutela dei minori nell’accesso ad internet” _____ NA → L1

10: “governance trasparente e flessibile” _____ L4

11: “Internet e Scuola” _____ NA → L2

12: “La Costituzione non tutela esplicitamente diritti digitali” _____ NA → L0

(gruppo) 13: “CONTROLLO OPERATO P.A.”, L4 - 14: “Gratuità dati della P.A.”, L4 - 15: “Diritto ad Internet”, L0 - 16: “Libertà e neutralità di Internet”, L0

17: “Software Libero”, trasversale _____ NA *

(gruppo) 18: “Identità digitale” come libera scelta”, L1 - 19: “Certificazioni gratuite per le competenze informatiche di base”, L2 - 20: “prima decisione efficace”, L0 - 21: “L’ecosistema digitale e la gestione della conoscenza”, L2 - 22: “visibilità dei dati pubblici”, L4

23: “Un URI per ogni cittadino” _____ NA → L1

24: “Autori di opere digitali” o “Creatori di contenuti”? ”, semi trasversale _____ NA → L5

25: “E-LEARNING FOR PRISONERS” _____ NA → L2

(gruppo) 26: “Accesso a internet”, L0 - 27: “Mappa dei disservizi pubblici e delle soluzioni online”, L6 - 28: “Piattaforma di informazione iperlocale in ogni Comune italiano”, L5 - 29: “ACCESSO AD INTERNET DIRITTO

FONDAMENTALE DEL CITTADINO", L0 - 30: "Internet è di tutti", L0 - 32: "Referendum comunale online su decisioni importanti", L7

33: "Si può dire database", forse meta_____L4

(gruppo) 34: "Garantire connessioni simmetriche", L1 - 35: "Istituzione di un sito per le comunicazioni di emergenza", L4 - 36: "Riduzione Analfabetismo Digitale", L2

37: "CITIZEN EXPERT", trasversale o poco chiara_____NA skip

"IN ITALIA L'ISTITUZIONE DI PIATTAFORME PUBBLICHE GOVERNATIVE PER coinvolgere i cittadini alle politiche del territorio . Il cittadino giustifica il proprio benessere, attraverso la PIATTAFORMA descrive le proprie politiche, coerenti al ambiente in cui vive. Sulla piattaforma il cittadino esperto denuncia le politiche locali INCOERENTI che anno prodotto malessere. [...] QUESTO E IL PROGRESSO DI CUI ABBIAMO BISOGNO GRAZIE HA INTERNET e alla partecipazione dei cittadini ATTENTI e coinvolti nella politica Locale e Statale. Buon lavoro ispettori I RESPONSABILI POLITICI IGNORANTI DEVONO ESSERE PRIVATI DEL POTERE DECISIONALE."

38: "E la scuola"_____L2

39: "fiducia nella politica"_____NA

(gruppo) 40: "Bilanci delle pubbliche amministrazioni da consultare online", L4 - 41: "Agenda Digitale Italiana: S.O.S. competenze digitali", L2 - 42: "Accessibilità software didattici e opere multimediali", L1 - 43: "Accessibilità dei libri di testo scolastici e loro espansioni", L3 - 44: "Accessibilità di tutte le attività scolastiche on line", L3 - 45: "accesso e riutilizzo di dati pubblici", L4 - 46: "Minori e social network", L2

47: "Cittadino digitale", trasversale_____L1 → NA *

49: "Una nazione al passo con l'avanguardia"_____L0

"Internet è un mezzo che permette di collegare fra loro tutti i dispositivi elettronici che ci accompagnano ogni giorno [...]"

51: "La proprietà è inapplicabile"_____L0

"Internet, [...] ha il dovere di essere neutrale. [...]"

53: "accesso e cancellazione dei dati personali", consapevolezza_____L2

54: "Piccole proposte per un codice deontologico sull'uso di Fb", relativo a Facebook in particolare, poteva forse essere assegnato al Liv. 2_____NA skip

(gruppo) 55: "Ambiente e-learning per docenti di ogni ordine e grado", L3 - 57: "Servizio di istruzione gratuito di alto livello", L2 - 58: "Accesso minimo garantito", L0 - 61: "Box selfservice per rapporti con PA", L3 - 62: "Internet come diritto costituzionale", L0 L1 - 63: "fibra ottica statale", L0 - 64: "Obbligo di trasparenza decisionale almeno tramite Internet", L4 - 65: "Neutralità della rete", L0 - 66: "Istruzione adolescenziale all'utilizzo di Internet", L2 - 67: "Stato che investe e guadagna", L0

68: "uffici con orari uguali sul territorio", non inerente_____NA skip

69: "Gli operatori NON DEVONO fungere da controllori"_____NA → L0

71: "Sicurezza preservando il Diritto alla Segretezza", misto L.0 e 2_____L2

(gruppo) 72: "Formare sull'accessibilità e le disposizioni di legge", L2 - 73: "Camera di commercio on line", L3 - 74: "rendiconto e bilanci pubblici", L4

75: "Seggi votazione digitale via internet", così assegnato perché inteso come puro cambiamento di mezzo e non cambiamento di processo_____L3

76: "Ricette mediche"_____NA → L3

77: "Cartelle cliniche open per il paziente"_____L4

78: "Aboliamo i contanti", richiesta di digitalizzazione servizio_____NA → L3

79: "RIFORMA -> Politica -> INTERNET", troppo generica, trasversale, mal formulata, e probabilmente per questo ha voto totale negativo____NA skip
(gruppo) **80:** "COMPETENZE DIGITALI A SCUOLA: NON SOLO WYSIWYG", L2 - **81:** "Architettura aperta", L1

82: "password unica", trasversale, applicabile a ogni livello____NA *

(gruppo) **83:** "E' necessaria una legge che tuteli la neutralità della rete", L0 - **84:** "WiFi veramente aperto", L0

85: "Il Governo chieda una Direttiva EU a tutela della neutralità"__NA L0

(gruppo) **86:** "Identità digitale e limiti", L2 - **87:** "Istituire le "150 ore digitali"", L2 - **88:** "Open Access", L4 - **89:** "Linee guida per il "versioning" dei dati pubblici", L4

90: "la vision di Internet non solo strumento di comunicazione", generica
____NA skip

91: "Scuola digitale e competenze esitenti"____NA → L2

92: "Favorire uso di licenze libere su piattaforme online"____L1

93: "Intermediari per il riuso dei dati pubblici", trasversale____NA → L4

94, 95: "open data riutilizzo per finalità di lucro-licenze", L4 - "Tutela privacy: da costo a vantaggio", L1

96: "Riforma della proprietà intellettuale", trasversale____NA *

97: "Notice and take down: censura arbitraria? ", aspetti legali trasversali
____NA *

98: "Disponibilità di collegamento gratuito"____L0

99: "LA DIGITALIZZAZIONE", trasversale, dich. di principio____NA *

100: "Tutte le scuole sono al passo con i tempi? ", generale____NA → L2

101: "DIGITALIZZAZIONE SCOLASTICA/REGISTRI ON LINE" _NA→L3

102: "Scuola, solo libri digitali"____NA → L2/L3

103: "Extended School", dichiarazioni di principio, poco comprensibile
____NA skip

"OBIETTIVO GENERALE Estendere il tempo e lo spazio-scuola per trasformarla in un centro di aggregazione sociale, di produzione culturale e di offerta di servizi sul territorio. Inoltre, rendere disponibili le infrastrutture tecnologiche delle scuole al di fuori degli orari di lezione per fornire dei servizi al territorio giustifica e valorizza gli investimenti sugli istituti scolastici. [...]"

104: "Internet per E-Learning", relativamente generica____NA → L2

105: "CERTIFICAZIONE INTERNET", generica, forse meta____NA skip

"Seguendo lo stile del presente Forum di consultazione "solo se si è certi dell'identità di chi scrive si accettano i commenti". La liberà di parola è di espressione sancita dalla nostra costituzione e da difendere ma anche le politiche di privacy e di diritto all'informazione vanno tutelate. Un informazione corretta non può non far fede sulla certezza della fonte. Sono inaccettabili alcuni commenti sulla rete verso persone, aziende e opinioni che per esser impedito necessitano di tempi non proporzionati alla velocità con cui la diffamazione permette alle stesse di provocare i danni. [...]"

110: "Internet come strumento del popolo per il governo"____L7

118: "IL CONTRIBUTO DELL' INDUSTRIA CULTURALE ITALIANA", fuori tema
____NA skip

122: "7 Livelli per la Cittadinanza Digitale", meta, inserito dagli autori **proprio per proporre l'arcobaleno come metodo di categorizzazione.** Ci teniamo a sottolineare come ben 79 utenti l'abbiano votata (la seconda proposta più popolare dopo "Software libero") e 20 utenti abbiano commentato__NA skip

123: "Proteggere gli utenti digitali dagli algoritmi", trasversale, di principio
NA skip

"La nostra libertà digitale (ma non solo) è nelle mani degli algoritmi. Formule di calcolo, stabilite a priori da qualche ottimo programmatore, che organizzano, indicizzano, ordinano le informazioni. Gli algoritmi sono sia umani (perché c'è chi li stabilisce) sia informatici. Questa 'automatizzazione' della gestione dei dati rischia di avere impatti pesanti sia sulla dignità e sulla libertà degli individui, in relazione al trattamento di dati che li riguardano come soggetti passivi sia, al contrario, sulla loro capacità di informarsi come soggetti attivi. Libertà e dignità, prima di tutto. Pensiamo al caso della cosiddetta 'polizia predittiva' [...]"

124: "Il contributo della FAPAV", fuori tema _____ NA skip**125:** "education" _____ L2**126:** "Eliminazione delle "barriere"" _____ NA → L1**128:** "Contro il divario centri/periferie", si intende divario digitale _____ L0**130:** "open data per la crescita", trasversale, di principio _____ NA → L4**131:** "Cittadine/i attive/i in rete", trasversale, di principio _____ NA skip**136:** "Il DNS richiede una sicurezza "globale"", verticale _____ NA skip**137:** "Programmi scolastici e formativi su uso sicuro di Internet" _____ L2**138:** "Tutelare e rendere sicure le Identità Digitali "Soft"" _____ NA → L3**139:** "Protezione delle Infrastrutture di interesse nazionale", trasversale
NA skip**140:** "Tutela dei soggetti deboli", trasversale, forse Livello 2 ("campagna informazione e sensibilizzazione") _____ NA → L2**142:** "Internet libero per dipendenti pubblici" _____ NA → L0**143:** "UN P.C. con accesso ad internet presso ogni URP è del cittadino"
NA → L3**144:** "tre aspetti fondamentali preliminari per l'industry tlc", trasversale, di principio, intervento istituzional-politico per "fare presenza" _____ NA skip**145:** "importanza di internet", trasversale, di principio, generica _____ NA skip**146, 148:** "evitiamo la tragedy of the commons", L2 - "quale neutralità della rete", L0**(gruppo) 152:** "punti di accesso ad internet", L0 - **153:** "accesso e riutilizzo dei dati del settore pubblico", L4 - **154:** "gestione dei big data", L4 - **162:** "infrastrutture di interesse nazionale", L0 - **163:** "fiducia nella rete", L2**164:** "Cieli bui", fuori tema _____ NA skip**(gruppo) 165:** "PIANO CONTRO IL DIVARIO DIGITALE", L0 - **166:** "Identità digitale su server dedicato", L3 - **167:** "Diffondere la cultura della privacy online", L2 - **168:** "ALFABETIZZAZIONE DIGITALE E PROGRAMMI LIBERI", L2**169:** "Profili professionali standard per i professionisti del Web" _____ NA → L2**170, 171:** "Internet nella scuola dell'obbligo", L2 - "Cybersecurity", L2**(gruppo) 173:** "Garanzia di accesso da parte della PA", L5 - **174:** "qualità del servizio internet", L0 - **175:** "Formazione Tecnica", L2 - **176:** "Cultura digitale", L2 - **177:** "Open government", L4

Infine raggruppiamo qui le proposte di livello trasversale o le meta-proposte (proposte relative al processo di consultazione). Queste proposte non sono state ovviamente assegnate a nessun livello dell'arcobaleno: 5) "Lo sviluppo dei diritti umani ed il loro esercizio in internet", molto generale e trasversale, meta, dichiarazione di principio; 31) "Coccarda", trasversale/meta? 52) "Il modello che garantisce la ripresa", trasversale o meta? 106) "Internet/web

come stato a se stante”, generica, di principio, trasversale; 107) “Serietà nella descrizione della governance di Internet”; 108) “Serietà nella descrizione della governance di Internet II”; 109) “sul forum”; 111) “Infrastruttura di rete che permette la creazione delle reti”; 112) “Diritti fondamentali”; 113) “Fusione di due articoli su diritti umani? ”; 114) “Accorpamento di questo principio ad altri? ”; 115) “Cooperazione internazionale o diritto internazionale? ”; 116) “Divario digitale o competenze digitali? ”, sulla terminologia; 117) “Quali competenze digitali? ”, terminologia; 119) “Difendere Internet”, dichiarazioni di principio; 120) “Estensione dei principi di Internet”, dichiarazioni di principio; 121) “Accesso alle informazioni”; 127) “Diritto di accesso”; 129) “Garantire anche aree “libere da Internet””, meta? Richiesta negativa, strana; 132) “revisione titolo Tematica III - Consumatori e utenti della rete”; 133) “eterogenesi dell’online”, trasversale, di principio; 134) “segnalazione di un passaggio poco chiaro”; 135) “proposta di spostamento del tema in Competenze digitali”; 141) “Tassonomia nella descrizione dei Threat ed Incidenti”, terminologia; 150) “processo multistakeholder: come renderlo ottimale”, trasversale; 151) “quali misure per favorire l’accesso per tutti”, trasversale, generico; 172) “Standard Internazionali”, trasversale;

3.3 Validazione del framework

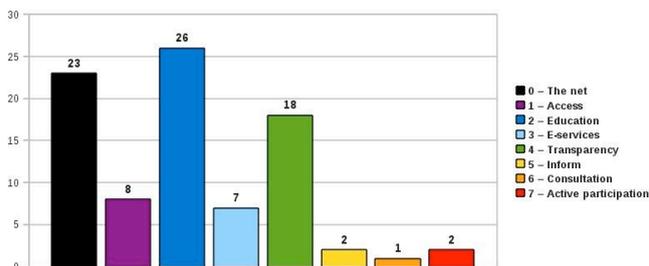


Fig.1 - distribuzione originale sui livelli

Le assegnazioni finali (comprendono le riassegnazioni rispetto al lavoro originale degli studenti visibile in Figura 1) ai vari livelli sono: L0=27, L1=11, L2=33, L3=13, L4=20, L5=3, L6=1, L7=2 da cui si evince la maggior propensione degli utenti della consultazione per i livelli bassi dell’arcobaleno mentre una notevole disaffezione o nemmeno presa di considerazione dei livelli più alti, quelli relativi alla partecipazione. Che questo fatto sia dovuto a puro disinteresse - quindi una **causa**, una mancanza di domanda - o a mancanza di fiducia nelle istituzioni - quindi un **effetto**, una mancanza nell’offerta di apertura della P.A. - al momento e con le poche informazioni a disposizione non è dato sapere.

Ovviamente ai fini della nostra analisi è interessante lo studio delle proposte che non hanno trovato posto nella categorizzazione definita dall’arcobaleno, troviamo 48 proposte non assegnate inizialmente dagli studenti (NA) di cui però 24 sono state riassegnate, ne rimangono 24, di cui 19 sono state marcate con

skip e non vanno quindi ulteriormente analizzate. Le rimanenti 5 sono: 17, 47, 82, 96, 97 e 99. Da questa operazione di (ri)assegnazione si evince che la maggior parte delle non assegnabili lo è non per mancanze del framework, ma per difficoltà di interpretare la proposta, a volte espressa in maniera generica/trasversale (quindi assegnabile contemporaneamente a più livelli) o relativa a principi generali o per errore (le proposte “sfuggite”) del valutatore.

Sofferamoci invece sulle proposte “asteriscate” in sezione 3.2, quelle che potevano essere (ri)assegnate o che hanno dato da pensare, cioè potevano eventualmente suggerire la modifica del framework stesso:

17: “Software Libero”, trasversale, licenze

47: “Cittadino digitale”, trasversale, identità

82: “password unica”, trasversale, applicabile a ogni livello, identità

96: “Riforma della proprietà intellettuale”, trasversale, proprietà intellettuale

97: “Notice and take down: censura arbitraria? ”, aspetti legali trasversali

99: “LA DIGITALIZZAZIONE”, trasversale, dichiarazioni di principio

Queste proposte ricadono in due aree: aspetti legali e gestione dell'identità digitale. Gli “**aspetti legali**” sono sicuramente trasversali al framework: il diritto d'autore, le licenze su software, hardware, dati, contenuti si possono applicare a molteplici livelli. Ad esempio si può parlare di diritto d'autore e licenze nell'ambito dei contenuti per l'education (L2), di licenze nell'ambito del software per la deliberazione online (L7) che deve essere verificabile e quindi disponibile con “licenze libere”, ancora di licenze nell'ambito dell'OpenData (L4) in cui i dati devono essere utilizzabili senza troppi vincoli, etc.. La “**identità digitale**” potrebbe essere rappresentata da un livello aggiuntivo intermedio fra L1 e L2, o anche dal Livello 1 stesso (intendendo l'identità digitale come uno dei “servizi minimi di cittadinanza digitale”). Ci pare di ravvisare questa ipotesi di livello aggiuntivo come unico “difetto” del framework: un livello mancante per promuovere l'attenzione sull'identità digitale e sulla sua difesa.

4. Conclusioni

La nostra “conclusione” (basata su questa prima applicazione, ovviamente serviranno ulteriori verifiche sul campo) è quindi che all'arcobaleno possa anche mancare la trattazione degli **aspetti legali**, ma che però tali aspetti sono trasversali, perpendicolari, applicabili a tutti i livelli dell'arcobaleno stesso e che quindi non sia possibile semplicemente aggiungere un livello. Mentre per l'**identità digitale** potrebbe servire un nuovo livello come meccanismo per elevare il rango di un aspetto (quello dell'identità appunto) importante della vita dei cittadini digitali.

Quindi alla domanda “manca davvero qualcosa in termini di livello? ” ci sentiamo di affermare che la risposta sia “ni” (più “no” che “si”): i livelli rappresentano un partizionamento sensato dell'insieme dei diritti della cittadinanza digitale, salvo decidere di estrapolare l'identità digitale per farne un livello a parte.

Metaforicamente parlando si potrebbe affermare che l'arcobaleno classifica una serie di “canali” (in senso lato) di comunicazione tra cittadini e stato, mentre gli aspetti legali rappresentano delle “modalità”, delle “politiche” (sia in senso di

policy che nel senso del termine italiano), dei pattern comportamentali applicabili a vari livelli. Si pensi ad esempio al concetto di “formato” delle informazioni: è un concetto generico e applicabile in moltissimi contesti e quindi a molti (se non tutti) livelli dell’arcobaleno. O anche al concetto di modalità (parola non a caso) “push/pull” che viene nominata in contesti molto diversi applicabili ai vari livelli ISO-OSI.

E’ quindi sensato ipotizzare eventualmente un’estensione (sulla quale stiamo lavorando), un’integrazione dell’arcobaleno nella direzione delle “modalità”: **iniziare una catalogazione e successiva classificazione degli aspetti legali, in termini astratti**. In concreto c’è già il codice, inteso come legge dello stato, sebbene la forma lasci molto a desiderare, naturalmente, si vedano ad esempio questi due interessanti post relativi alla legislazione USA: 1) “... The bill and the amendment were written in what amounts to a legal foreign language, making it impossible for citizens to know who was lying or telling the truth ...”[Potts, 2013]; 2) “Legislation is difficult to read and understand. So difficult that it largely goes unread.”[Vergottini, 2013]. Ma ovviamente per l’Italia vale lo stesso.

Bibliografia

[Caddy and Vergez, 2001] Caddy, J. and Vergez, C. (2001). Citizens as partners: Information. Consultation and Public Participation in Policy Making, report prepared for Organisation for Economic Cooperation and Development (OECD).

[Clement and Shade, 2000] Clement, A. and Shade, L. R. (2000). The access rainbow: Conceptualizing universal access to the information/communication infrastructure. Community informatics: Enabling communities with information and communications technologies, pages 32–51.

[Cuculo and Rasente, 2013] Cuculo, V. and Rasente, S. (2013). Analisi della consultazione del MIUR sui principi fondamentali di internet. Technical report, Dip.to di Informatica. Report.

[De Cindio et al., 2012] De Cindio, F., Sonnante, L., and Trentini, A. (2012). Cittadinanza digitale: un arcobaleno di diritti e opportunità. Mondo Digitale, nr. 42.

[Potts, 2013] Potts, R. (2013). Lying about the law? how can you tell? http://www.americanthinker.com/2013/04/lying_about_the_law_how_can_you_tell.html

[Vergottini, 2013] Vergottini, G. (2013). Transparent legislation should be easy to read. <http://legixinfo.wordpress.com/2013/07/08/transparent-legislation-should-be-easy-to-read>

[Zimmermann, 1980] Zimmermann, H. (1980). Osi reference model—the iso model of architecture for open systems interconnection. Communications, IEEE Transactions on, 28(4):425–432.

Ringraziamo gli studenti Sonia Rasente e Vittorio Cuculo per il lavoro di analisi effettuato come elaborato d’esame del corso di Cittadinanza Digitale e Tecnocivismo (<http://tecnocivismo.di.unimi.it>) su cui questo articolo è basato.

Integrated Mobility for Individuals in Smarter Cities: a Crowd-sourcing approach

Gianmario Motta¹, Daniele Sacco², Linlin You³, Ruan Chonghe⁴, Lu Zhang⁵

Università di Pavia

Via Ferrata 1, 27100, Pavia (PV)

¹ motta05@unipv.it

² daniele.sacco01@ateneopv.it

³ linlin.you01@ateneopv.it

⁴ ruan.chonghe01@ateneopv.it

⁵ lu.zhang01@ateneopv.it

We present an ongoing research on an Integrated Real-time Mobility Assistant (IRMA). IRMA is a software system that targets the personal mobility in a near future scenario, based on green, shared and public transports. IRMA architecture includes smartphone applications and a set of web services to gather and interpret any relevant source of information, that includes open data, crowd data and big data. The technology is SOA/EDA (Service Oriented Architecture / Event Driven Architecture) and the service will gather and interpret any relevant source of information on transport resources and their availability. Information includes user generated content through a crowdsourcing service and data from/to social networks. Our paper provides a review of current crowdsourcing approaches to a smart and collaborative mobility, and proposes a future application scenario. IRMA, after being proved on test cases, will be tested by the students of University of Pavia.

Keywords: smart city; mobility integrator; service system; crowdsourcing; open data

1. Introduction

Smart cities are the future, of course. According to Horizon 2020 transport future cities shall be smart, green, and integrated (<http://ec.europa.eu/research/horizon2020/>). To be smart, cities have to be wired and with an ubiquitous presence of terminals, as smartphones, smart TVs, in-vehicle devices, that enable users to be served by an integrated mobility information service. Also, the future city implies a new vision of mobility. In such perspective, a future smart city is the outcome of the integration of various systems, namely Traffic Management Systems, Transport Systems, User Systems, Vehicle Systems [Motta et al, 2013]. Finally, a smart city not only

needs a strong ICT infrastructure, but also human capital/education, social and relational capital and environmental interest. Smart cities include smart infrastructure, smart operation, smart service and smart industry, and have four basic characteristics [Hao et al, 2012]:

- (a) Interconnection between urban parts,
- (b) integration of urban information systems,
- (c) urban management and service cooperation
- (d) application of modern city management theory.

However, most researches and real world projects target the current city where car-based mobility overwhelms any other public transport option. By contrast, Horizon 2020 and our same survival needs a green mobility, where public and shared transport dominates and fulfills social inclusion.

To develop a vision of an explicitly sustainable transport system which has the power to motivate people to change, it is necessary to combine together substantial technology and behavioral changes, and to ensure ecology, economic viability and good quality of life. Technology change ensures that the remaining transport will be sustainable. The optimization and intelligent design of multimodal mobility concepts via information and communication technologies also play an important role [Schade et al, 2011].

The role of public and collaborative transport is essential to an urban sustainable mobility. An integrated transport can lead to an improved quality of life, socio-economic development and urban renewal. This integration also enables higher density development with positive benefits on the environment: more efficient land use, higher energy efficiency, reduced pollution, climate change mitigation and protection of open space through smarter growth pattern [Ribeiro et Mendes, 2010][Hensher et Wong, 2011].

Our paper intends to provide a review of current crowdsourcing approaches to a smart and collaborative mobility, and to show future application scenarios, by providing a proof of concept for individual mobility by public and shared transport.

2. Smart Mobility and Crowdsourcing

Our assumption is that citizens must be involved in the value creation process of their own city. Thus, further researches suggest that a crowdsourcing approach can be used for both smart mobility and smart communities.

We report the definition given by [Howe 2006]: “*crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers*”.

[Brabham, 2008] demonstrates that crowdsourcing is not only a buzz-word by reporting many success stories. Specifically, he defines the presence of non-experts and amateurs as the fundamental for crowdsourcing initiatives.

The coordination of user-generated content, web applications, and mobile devices can extend crowdsourcing beyond the digital domain and link it to tasks in the real world [Alt et al, 2010]. So, cities and municipalities can use crowdsourcing to identify and solve issues with more effectiveness thanks to the help of volunteers (i.e. citizens). [Erickson, 2010] identifies following reasons to use crowdsourcing in urban contexts: (a) cities are dense concentrations of people who can be easily involved in a crowdsourcing initiative, (b) residents have a deep knowledge of the place where they live, (c) residents have a practical interest in participating in systems that help their daily life, and (d) residents of a city or municipality identify themselves with networks of family, friends and communities around them.

For urban governance several applications exist, the most wide-spread ones follow:

- PDX Reporter is a crowdsourcing system that allows citizens to interact with the city of Portland government concerning problems or issues with publicly maintained infrastructure (<http://www.portlandoregon.gov/bts/53613>).
- CitizensConnect is used to alert the City of Boston about neighborhood issues such as potholes, damaged signs, and graffiti (<http://www.cityofboston.gov/doit/apps/citizensconnect.asp>).
- In Italy, Decoro Urbano/WeDu web site and related mobile application have an architecture very similar to PDX Reporter and CitizensConnect (<http://www.decorourbano.org/>).
- SeeClickFix goes further and allows users (individual or collective) to draw “public watch areas” so that alerts may be sent whenever a new project is started or a new issue is reported. It covers several cities in USA (<http://seeclickfix.com/>).

However, research and application of crowdsourcing to mobility governance is only at initial phases. [Balena et al, 2013] developed an internet-based mobile application to improve urban utilities management, with special regard to public open space maintenance and local transport. [Steinfeld et al, 2011] focus their system on crowd-sourcing acquisition of information about bus location and fullness, and reporting problems and positive experiences within the transport service. [Stenneth, 2011] implemented an incident reporting system where users can report an incident such as an accident or construction on a road network. Finally, [Liu et al, 2013] apply crowdsourcing for carpooling purposes. They developed MobiCrowd, which leverages private smartphone to collect individual trips for carpooling. It automatically crowd-sources daily trips and mobility models for each user, instead of tracking vehicles or asking users to input their trips.

Within a collaborative mobility, moving across a network of public and/or shared transport systems (shared bikes, pooled cars) requires that the users are guided by a computerized assistant, that supports them in defining the trip, and, en route, supports them to identify and report service disruptions. This is precisely the reference scenario for which IRMA (Integrated Real-time Mobility Assistant) has been conceived.

3. Proof of Concept

IRMA (Integrated Real-time Mobility Assistant) handles itineraries, performed by one person or a group, that are conceived as an oriented graph, whose connections are represented by nodes, while energy, pollution and cost requirements are modeled as resource constraints as it happens with project modeling techniques.

Work started in Department of Information and Industrial Engineering of Pavia University to develop a mobile application that could assist travelers to meet their schedule even with transport disruptions [Motta et al, 2011]. A SOA - EDA (Service Oriented Architecture – Event Driven Architecture) infrastructure enables services that access field information on mobility. Such collection involves various data sources.

IRMA uses itinerary as master information, that is integrated both before and during the trip by additional information about:

- Transport map, based on structured information about transport resources which are / will be available e.g. train or underground;
- Transport status, based on semi-structured and real-time information about future and current availability and delay; e.g. train timetable, train delay, underground load, road traffic and deviation from standard travel time;
- Crowd model, based on mining of un-structured information gathered from social network and sensor data (i.e. Big Data)
- User feedback, based on information gathered from the “crowd” (i.e. Crowdsourcing), handled as a text message (e.g. traffic jams, road bumps, etc.)

Through a communication layer, the system retrieves information on traffic and mobility, with various purposes, e.g. access to timetables, on demand access, real-time access.

IRMA intends to cover the whole mobility lifecycle for individuals. So, the system enables planning, execution, monitoring and control of individual mobility, as follows:

1. Inter-modal travel planning by considering data from transport providers, e.g. timetables, bus stops, etc.
2. Execution is focused on:
 - payment, that is a single payment for a multi-modal journey involving different transport companies.
 - user feedback provision about network and infrastructure condition.
3. Real-time travel monitoring and statistics (e.g. mean measured time of travel from A to B), by considering:
 - structured data, i.e. real-time data from transport providers.
 - unstructured data, e.g. mining of social network data, sensor data, etc.
4. Alternatives identification and journey re-planning, e.g. by compensating delays, identifying disruptions, strikes, etc.

Figure 1 shows the functional modules of IRMA.

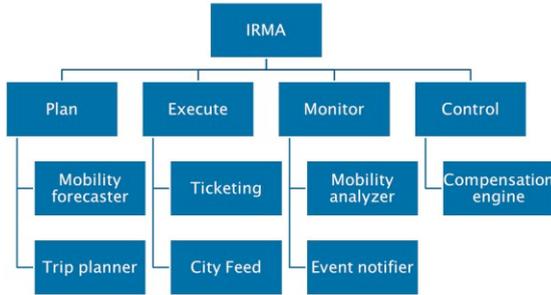


Fig. 1 - Functional modules of IRMA

The overall concept has been defined in a previous paper and a proof of concept for the Trip Planner module, based on transport data from Torino (Gruppo Torinese Trasporti), was performed [Motta et al, 2013].

Here we present an extension of Trip Planner and a proof of concept for City Feed module.

3.1 Trip Planner

Trip Planner stores a mobility map and related mobility data. The mobility map describes mobility resources within the urban area, by route, time, and mode.

It shall assist the end user to plan and configure his/her mobility across multiple mobility options. Request Handler processes the mobility request that may concern an individual trip or a calendar. The Handler helps the user to define the optimal mobility plan by accessing mobility timetables and mobility through the Information Retrieval sub service. The user will choose and confirm the ideal option as an individual itinerary.

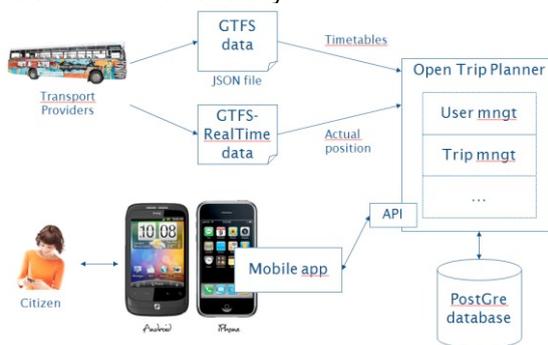


Fig. 2 – Architecture of Trip Planner module

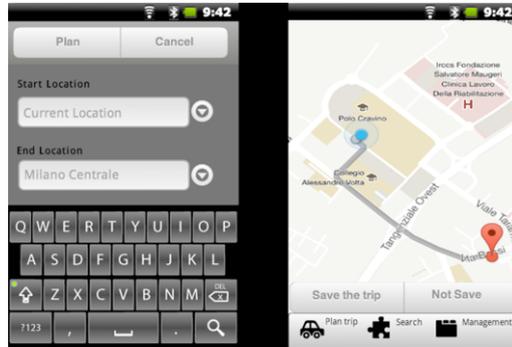


Fig. 3 – Trip Planner implementation in Android

To access mobility timetables, the Information Retrieval uses GTFS (General Transit Feed Specification). GTFS is an open and common data format for temporal schedule and spatial data that is used for transit trip itinerary planning. GTFS was initially developed collaboratively for Google Transit project (<https://developers.google.com/transit/gtfs/reference>). Over 400 transportation agencies are publishing their schedules and other information using GTFS format. Thanks to standardization, a tool designed for one city/agency can be applied elsewhere.

As a specific proof of concept, we have tested the Trip Planner module on Torino open data. We have extended Open Trip Planner (OTP), an open source platform that can be implemented by transit agencies or municipalities to provide services to travelers at a low cost. OTP relies on GTFS data.

Our extension provides new features to Open Trip Planner and also access through mobile devices, as shown in Figure 2. User management and trip management features allow users to store information about their journey information and preferences, so they can re-use them on different devices (e.g. laptop, smartphone, tablet, smart TV, etc.). Figure 3 shows the current implementation on Android.

3.2 City Feed

City Feed module for city management gathers crowd data with picture-and-text feeds and enables citizens to identify (a) infrastructure issues, e.g. roadway disruptions, not working traffic lights, inadequate bicycle and pedestrian facilities, work zones, etc., and (b) mobility issues, e.g. delays, parades, strikes, increased truck traffic, etc.

City feed intends to support the whole lifecycle of issues, from identification to repair. The process can be described as shown Figure 4. Residents identify issues during their mobility by providing a brief description and a picture taken with their mobile device (e.g. smartphone). Feeds are tagged according to a predetermined list. When the feed is uploaded to the system also its geo-location is stored thanks to GPS coordinates. Second step requires issue certification by contractors who go on the spot identified by the user feed. They

can easily identify a new feed because they receive an automatic notification (e.g. by email or in the system itself) as soon as a feed is uploaded. Finally, contractors open a remedy ticket to track the status of the issue. A notification of succeeded reparation is generated in the system as soon as the remedy ticket is closed.



Fig. 4 – City Feed process

City Feed shows all the open issues on the city map. The map can be enhanced by other variables (e.g. traffic flow). So, main information entities managed by City Feed are:

- Events: they describe potential issues by brief description and pictures. They are generated by residents.
- Issues: information about damages in terms of location, type and description. They are generated by contractors.
- Tickets: information about repair status of issues. They are generated by contractors.
- Analysis results: city performance indicators for events, issues and tickets. They are monitored by the municipality.

It may happen that a resident uploads pictures which contain people without their permission. In consideration of privacy protection, we implement an algorithm that automatically delete pictures containing human faces, thanks to face recognition technology.

The feeds reporting transport network or infrastructure status (recognized because of the classification tag) are automatically redirected to Trip Planner thanks to Event Driven Architecture, so users who intend to plan a trip have a comprehensive knowledge about transport condition in their city.

Finally, we studied integration with main social networks (e.g. Facebook, Twitter, Foursquare), in order to exploit the power of hash-tags and easily disseminate information about city status. The tag in the predetermined list can be transformed in an hash-tag and published in a social network. Fig. 5 shows the final architecture for City Feed module, in Fig. 6 its implementation.

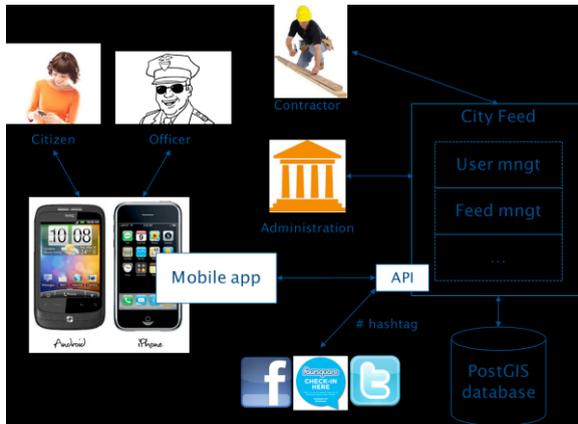


Fig. 5 - Architecture of City Feed module

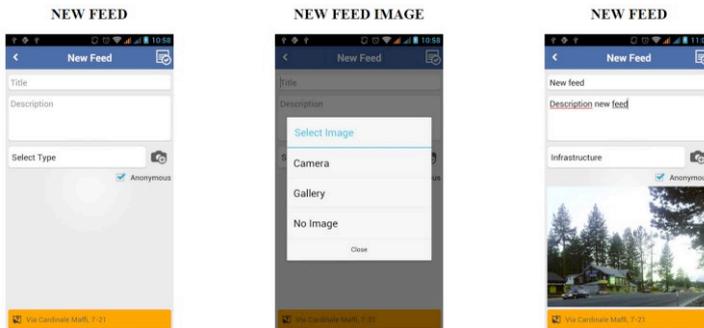


Fig. 6 – City Feed implementation in Android

4. Conclusion

We have illustrated the key concepts and modules of IRMA, i.e. an integrated mobility assistant, which we regard as a stage 2 of smart cities. IRMA is in an early implementation stage and only prototypes have been made to date.

IRMA implementation in Smart Cities is expected to impact environment, business and users. The service will lower energy and emissions, since it will ease the use of public/shared transport systems even with complex connections. Also, a variety of stakeholders can benefit from IRMA. In business perspective, municipalities and transport providers can analyze and forecast mobility in terms of time, route, connection, and mode, and, therefore, can optimize transportation resources. Mobile users can define their travel and can actually travel in an optimal way, without any previous knowledge of any transportation system. City Feed module allows the municipality to achieve

following purposes: (a) to make a cost-effective city management process, (b) to motivate residents to participate in city management, and (c) to improve the city environment. Finally, City Feed and Trip Planner integration allow users and municipalities to retrieve transport network and infrastructure status in real-time thanks to the collaborative approach given by crowdsourcing.

As future goal, with a cloud based service as IRMA, users will, ideally, plan their mobility in any urban area across the world. Also, a global coverage implies several advances in clouds, as interoperability among cloud offerings with federating applications running on different clouds, and combine pull and push approaches by integrating the new modules to be implemented by Service Oriented and Event Driven architectures. Finally, IRMA can be platform for a global virtual ticket. Simply, the users buy a mobility ticket, that is nothing else but a credit, alike a SIM recharge, that can be spent on the itineraries managed by IRMA.

References

- [Alt et al, 2010] Alt, F., Shirazi, A. S., Schmidt, A., Kramer, U., & Nawaz, Z., Location-based crowdsourcing: extending crowdsourcing to the real world, In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (pp. 13-22). ACM
- [Balena et al, 2013] Balena, P., Bonifazi, A., Mangialardi, G., Smart Communities Meet Urban Management: Harnessing the Potential of Open Data and Public/Private Partnerships through Innovative E-Governance Applications, Computational Science and Its Applications–ICCSA 2013. Springer Berlin Heidelberg, 2013, 528-540.
- [Brabham, 2008] Brabham, D. C., Crowdsourcing as a model for problem solving an introduction and cases, *Convergence: the international journal of research into new media technologies*, 14.1 (2008): 75-90.
- [Erickson, 2010] Erickson, T., Geocentric Crowdsourcing and Smarter Cities: Enabling Urban Intelligence in Cities and Regions, 1st Ubiquitous Crowdsourcing Workshop at UbiComp, 2010
- [Hao et al, 2012] Hao, L., Lei, X., Yan, Z., Li, Y. C., The application and Implementation research of Smart City in China, International Conference on System Science and Engineering, Dalian, China, 2012.
- [Hensher et Wong, 2011] Hensher, D. A., Wong, G., Different Approaches to Public Transport Provision, LTA Academy-Journeys , Land Transport Authority, 2011
- [Howe, 2006] Howe, J., Crowdsourcing: A Definition, Crowdsourcing: Tracking the Rise of the Amateur (weblog, 2 June), URL (accessed 24 November 2006): http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html
- [Liu et al, 2013] Liu, N., Feng, Y., Wang, F., Liu, B., Tang, J., Mobility Crowdsourcing: Toward Zero-Effort Carpooling on Individual Smartphone, International Journal of Distributed Sensor Networks, 2013.
- [Motta et al, 2011] Motta G., Barrorero T., Telese F., Design of performance aware service systems, The International Joint Conference on Service Sciences (IJCSS) 2011, Taiwan

[Motta et al, 2013] Motta, G., Sacco, D., Belloni, A., You, L., A System for Green Personal Integrated Mobility: A research in progress, IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), 2013, China

[Ribeiro et Mendes, 2010] Ribeiro, P., Mendes, J.F.G., The role of public transport in the achievement of urban sustainable mobility in mid-sized cities, Congresso Luso-Brasileiro para Planejamento Urbano, Regional, Integrado, Sustentável (Portugal), 2010

[Schade et al, 2011] Schade, W., Peters, A., Doll, C., Klug, S., Köhler, J., Kra, M., VIVER. A sustainable transport vision for Germany, Business Unit Transportation Systems Fraunhofer Institute for Systems and Innovation Research ISI, 2011

[Steinfeld et al, 2011] Steinfeld, A., Zimmerman, J., Tomasic, A., Yoo, D., Aziz, R. D., Mobile transit rider information via universal design and crowdsourcing, Proceedings from Transportation Research Board 2011 Annual Meeting

[Stenneth et al, 2011] Stenneth, L., Stone, W., Alowibdi, J., Reducing Travel Time by Incident Reporting via CrowdSourcing, ICOMP'11

Smart city: innovazione e occupabilità

Dunia Pepe, Marta Palombi¹
Isfol, Università Roma Tre, Corso d'Italia, n. 33 – 00198 – Roma
d.pepe@isfol.it

¹Marta Palombi
Palombi Editore, Via Gregorio VII, n. 224 – 00165 – Roma
marta.palombi@email.it

Abstract. *This paper is the result of a research on the relationship among smart cities, dynamics of innovation and processes of transition from training to work. It has assumed that the possibility of facilitating the transition from training to work is currently linked to innovative practices based simultaneously on several variables: training and guidance, work and firms, individuals and territories, attention to the person and digitization of processes. The existence of a close relation among these variables is due to the fact that most of the current socio-economic systems are taking on a smart dimension characterized by strong and significant connections among subsystems. This paper reports two innovative examples for work: The Porta Futuro Centre in Rome for guidance, training and introduction to work of unemployed people; the EnLabs Luiss Center, which offers, within the Termini Station in Rome, space, advice and resources for the creation of start-ups for young people.*

Keywords: smart city, innovation, employability

Introduzione

Questa ricerca su “Smart city, innovazione e occupabilità” si sviluppa intorno ad un’ipotesi fondamentale secondo cui i processi legati alle dinamiche di innovazione e di digitalizzazione sono destinati ad esercitare un’influenza positiva sui processi di transizione tra i sistemi dell’istruzione, della formazione e del lavoro. In particolare, le misure della digitalizzazione che dovrebbero maggiormente influenzare i meccanismi dell’inclusione sociale e della transizione al lavoro riguardano l’innovazione dei sistemi gestionali e organizzativi; la rapida diffusione delle informazioni e la facile circolazione dei saperi; una partecipazione attiva da parte della cittadinanza anche attraverso contesti virtuali di scambio delle informazioni; lo sviluppo di interazioni tra reti informative e reti territoriali che favoriscono il passaggio dell’informazione sui percorsi di costruzione del lavoro; lo sviluppo di competenze e di conoscenze originate dalla diffusione delle reti e dei sistemi di interconnessione e di interoperabilità; l’impulso dato alla promozione delle start up vale a dire di imprese ad alto contenuto tecnologico e costituite soprattutto da giovani.

La promozione di politiche attive del lavoro appare legata, nell’attuale contesto storico-sociale, all’esistenza di forti sinergie tra i contesti conoscenza, del lavoro, delle imprese, dei territori, delle realtà pubbliche e private. All’interno di questa prospettiva, la nostra ricerca si colloca inevitabilmente nel contesto delle interpretazioni delle città moderne in termini di “smart city”.

Le problematiche relative alle smart city sono affrontate facendo riferimento, da un lato, ad autori impegnati nel panorama internazionale per lo studio dei processi di costruzione e di evoluzione delle smart city quali Charles Leadbeater, Pelle Ehn e Tim Campbell. D’altro canto, si fa riferimento agli studi ed ai lavori condotti in Italia da Associazioni quali gli Stati Generali dell’Innovazione e da autori quali Luca De Biase, Andrea Granelli, Flavia Marzano, Miriam Ruggiero, Michele Vianello.

Ispirandosi alla classificazione proposta dall’Università di Vienna, Flavia Marzano descrive la smart city come il risultato della combinazione intelligente di sei fattori: economia, mobilità, governance, vita, cittadini, ambiente [Marzano, 2012]. La smart city è dunque una smart community: un luogo in cui le relazioni non si esauriscono nei confini fisici; la qualità della vita dei city-user è l’indicatore predominante e l’obiettivo principale; i servizi sono centrati sulle esigenze delle persone; le politiche sono intelligenti e caratterizzate da volontà di ‘apertura’ e ‘integrazione’.

La città intelligente viene concepita dalla maggior parte delle prospettive teoriche e degli autori, presentati in questo saggio, in termini sistemici. In particolare, una chiave importante di interpretazione delle dinamiche organizzative ed evolutive della smart city può essere data dalla teoria dei sistemi complessi introdotta nel contesto europeo ed internazionale intorno al 1980, da autori quali: I. Prigogine, I. Stengers, H. von Förster, E. Morin, U. Maturana, F. Varela, G. Bocchi, M. Ceruti [Bocchi G. e Ceruti M., 1985]. Questa teoria afferma che i sistemi della biologia, della società e della storia sono sostanzialmente sistemi viventi dotati di una forte vitalità interna e di autonomia. I sistemi complessi o autoorganizzatori si adattano al loro ambiente esterno ed alle loro necessità organizzative interne costruendo e selezionando dunque i comportamenti e le risposte che essi ritengono più adeguati alla loro storia ed alle esigenze della loro evoluzione.

1. Smart city e smart system

Nella "Carta d'intenti" degli Stati Generali per l'Innovazione, pubblicata nel Gennaio 2013, si sottolinea come il concetto ed il tema dell'innovazione non riguardino una sola area di intervento, uno specifico settore di sviluppo e di crescita, ma un modello di sviluppo economico e sociale che interessa al tempo stesso l'istruzione, la formazione ed il lavoro; la pubblica amministrazione e le imprese private; l'esistenza dei giovani e quella degli adulti; la qualità della vita nei territori urbani ed extra urbani [Stati Generali dell'Innovazione, 2013]. Le politiche per l'innovazione toccano dunque simultaneamente diverse problematiche e questo si deve, evidentemente, al fatto che la società della conoscenza è caratterizzata da forti interazioni tra le dimensioni ed i contesti dell'istruzione, della conoscenza, della formazione, del lavoro, delle imprese, della ricerca e dell'ambiente.

I principi per l'innovazione sopra delineati rispondono in ampia misura agli obiettivi posti dalla Strategia Europa 2020 e riguardano: una crescita intelligente, una crescita sostenibile e una crescita inclusiva. La scelta di porre al centro delle politiche dell'innovazione la qualità della vita ne comporta altre basate, a loro volta, su alcuni concetti fondamentali: la sostenibilità volta ad evitare lo sfruttamento eccessivo di risorse non rinnovabili; l'apertura intesa come possibilità di connessione tra dati, idee, proposte, progetti ed esperienze; la centralità territoriale che fa del territorio il fulcro delle politiche per l'innovazione. Per ciò che riguarda in particolare la centralità dei territori, nella Carta d'intenti per l'innovazione si evidenziano tre strategie: a) favorire lo sviluppo delle *smart city* inteso come sviluppo delle interazioni tra i diversi sistemi e sottosistemi come sanità, scuola, lavoro, ambiente, energia, mobilità; b) adottare il modello della *open innovation* teso a favorire gli scambi per l'innovazione tra le aziende e tra queste ed il contesto di riferimento; c) stimolare l'utilizzo del *cloud computing* come soluzione tecnologica che consente di ottimizzare i costi di gestione dei servizi, pur mantenendo il controllo delle loro politiche a livello dei singoli territori [Stati Generali dell'Innovazione, 2013].

Nel corso del Convegno Human Smart City, tenutosi a Roma nell'ambito dell'edizione Forum PA 2013, Pelle Ehn illustra il concetto di "Design Thinking in the City". "Il design, spiega Ehn, non è più inteso solo in riferimento all'oggetto ed alle sue forme" [Ehn, 2013]. Questo termine non è più interessato solo alla produzione industriale, all'architettura e alla comunicazione visiva, ma si occupa ormai da diversi anni anche di problemi quali lo sviluppo internazionale, la sanità, il design dei servizi pubblici rafforzando la convinzione che per la risoluzione di questi problemi debbono attivarsi soprattutto le organizzazioni pubbliche ed i governi. "Il *Design Thinking* e la progettazione partecipata, così come io li intendo, sono diventati sempre più degli asset in ambito pubblico e nella vita di tutti i giorni e ciò comporta un ri-orientamento che porta dalla 'democrazia al lavoro' all' 'innovazione democratica'. L'innovazione è stata, infatti, democratizzata attraverso un facile accesso a strumenti di produzione che rendono gli utenti dei nuovi esperti in grado di guidare l'innovazione" [Ehn, 2013].

Il Design Thinking, in relazione allo sviluppo delle smart city, deve andare dunque di pari passo con l'impegno pubblico. Esso rappresenta un modo intelligente per pensare a come le risorse possono funzionare meglio e guidare il cambiamento. Un concetto, fortemente legato al design thinking, riguarda le nuove tecnologie che rappresentano una dimensione chiave della smart city. Tuttavia la dimensione tecnologica è rilevante solo nella misura in cui viene messa in relazione con i problemi di natura sociale. Secondo la prospettiva della Human Smart City, la città è intelligente nella misura in cui mette insieme prima di tutto diversi attori e il loro potenziale umano.

Tim Campbell spiega per quali ragioni le città debbono essere riposizionate al centro dello sviluppo socio-economico. Da un lato, nella prospettiva di questo studioso, il livello territoriale della città è quello colpito più direttamente dai problemi riguardanti il cambiamento climatico, le catastrofi e le azioni di recupero, d'altro canto, le città stanno avanzando verso ruoli sempre più importanti in affari nazionali e internazionali, così come verso direzioni specifiche legate alla promozione del commercio ed allo sviluppo economico locale, ai rapporti di cooperazione con le università, gli ospedali e le organizzazioni di ricerca. In tutti questi campi le città sono in grado di agire e reagire più velocemente rispetto ai governi nazionali, e in ogni caso le città sono meno gravate da vincoli burocratici e diplomatici internazionali [Campbell, 2013].

Il concetto di intelligenza, in riferimento alla definizione delle città, fa riferimento il più delle volte alle connessioni ad alta tecnologia, ai sensori, alla possibilità di trattare simultaneamente grandi quantità di dati. Questa definizione di smart city, osserva Tim Campbell [Campbell, 2013], non è però esaustiva perché lascia fuori altre dimensioni estremamente importanti legate all'infrastruttura immateriale di organizzazioni civiche, ai gruppi di quartiere, al mondo accademico e alle comunità imprenditoriali come attori e partner nel processo di gestione della città. Inoltre, le definizioni convenzionali di "smartness" ignorano spesso l'importante processo di apprendimento collettivo; vale a dire l'imponente volume di conoscenze e di saperi che si costruiscono e si scambiano sia all'interno della città che tra la città ed il mondo esterno. All'interno di questa prospettiva, le Smart city sono in grado di fare il miglior uso di strumenti tecnologici attivando un processo di apprendimento attraverso il quale nuove idee vengono catturate e adattate per uso locale.

Proprio grazie ai processi di apprendimento, crescita ed innovazione si costruisce una città intelligente intesa in una prospettiva sistemica; vale a dire "sia come struttura connettiva (aperta, consapevole e finalizzata), sia come struttura adattiva, capace di generare dati e conoscenza e di far evolvere i propri comportamenti. In relazione al territorio, le comunità intelligenti identificano l'ambito in cui si esplicano la capacità di governo e i livelli di aggregazione amministrativa opportuni (area municipale o comunale, sistema metropolitano, area vasta territoriale/provinciale) rispetto alle tematiche gestibili in autonomia" [Stati Generali dell'Innovazione, 2013]. Il concetto di smart city sembra dunque ridefinirsi, da un punto di vista epistemologico, nei termini di un sistema auto-organizzatore o complesso dotato cioè di una forte vitalità interna, che ne definisce l'identità, l'autonomia e la creatività. Il sistema auto-organizzatore si configura come un insieme di elementi integrati all'interno di specifici livelli; esso possiede di conseguenza un notevole grado di autonomia sia per ciò che riguarda i suoi rapporti con

l'ambiente esterno che per ciò che riguarda i rapporti che mettono in connessione, al suo interno, i diversi livelli organizzativi. Gli input ambientali del sistema non si pongono come programmi deterministici, capaci di avere un'influenza decisiva sulle dinamiche sistemiche, bensì come perturbazioni casuali i cui effetti generano sempre nuove dinamiche di comunicazione tra i sottosistemi e nuove configurazioni dei livelli organizzativi che caratterizzano la vita del sistema stesso. Il concetto di autoorganizzazione permette infine di affrontare, secondo un'interessante prospettiva, anche rispetto al concetto di smart city, il problema dell'identità sistemica e quello, ad esso fortemente connesso, della creazione del significato.

Il sistema auto-organizzatore si configura come un sistema vitale che si adatta al proprio ambiente selezionando gli stimoli e costruendo le risposte, che esso ritiene più adeguate per vivere e per evolvere nel suo ambiente. Il sistema possiede infatti una capacità cognitiva che gli consente di dare delle risposte attive alle influenze del mondo esterno di reagire in maniera adeguata agli stimoli interni ed esterni mantenendo il suo equilibrio e cercando di realizzare la sua evoluzione. "Il dominio cognitivo di un sistema autonomo, osserva Mauro Ceruti, finisce per configurarsi come l'insieme delle relazioni in cui il sistema può entrare senza perdere la sua identità, cioè senza morire, se è un sistema vivente, o senza smettere di funzionare se è un sistema cognitivo individuale o sociale" [Ceruti, 1989]. L'evoluzione del sistema non è determinata in maniera necessitante dagli inputs del mondo esterno ma è piuttosto il risultato di una costruzione: vale a dire dell'elaborazione di eventi singolari che il sistema incontra nel corso della sua storia. Ne emerge una concezione del sistema inteso come un processo biologico e dunque vitale, continuamente mutevole e continuamente rinnovantesi grazie alla sua capacità di confrontarsi con gli inputs del mondo esterno, di svilupparsi e di evolvere nella scelta delle risposte e delle soluzioni più adeguate a questi inputs e nell'esigenza di adattarsi alle mutevoli condizioni del mondo circostante.

Nel corso del suo intervento al Forum PA 2013, Charles Leadbeater spiega come la smart city rappresenti il risultato di una combinazione ideale tra il sistema città e l'empatia. Queste due variabili sono rappresentate da rette che si intersecano perpendicolarmente: specificamente, la prima variabile si snoda lungo un continuum ai cui estremi si pongono le categorie low system ed high system mentre la seconda variabile esprime un continuum ai cui estremi si pongono le categorie low empathy e high empathy. Leadbeater illustra la possibilità di porre i diversi modelli di città nei diversi punti rilevabili all'interno dell'area disegnata dall'incrocio di queste due rette e, nella prospettiva dello studioso, la smart city si colloca proprio in spazi caratterizzati simultaneamente da alti livelli di organizzazione sistemica e di empatia [Leadbeater, 2013]. Le migliori città, spiega Leadbeater, sono progettate in modo che abbiano infrastrutture, trasporti, scambi energetici, sistemi di connessione reale e virtuale efficienti. Ma queste città debbono anche consentire lo sviluppo dei rapporti umani, di ampi spazi di convivialità e di possibilità di crescita della società civile. Tuttavia ed in ultima analisi, solo una logica di bottom up può rendere realmente possibile lo sviluppo di una città a partire dalla vita delle persone, consentendo alla smart city di prendere forma e fiorire nelle sue molteplici dimensioni [Leadbeater, 2013].

Nel trattare il problema relativo alla riformulazione della città e dell'urbanistica, Flavia Marzano utilizza, ispirandosi a William J. Mitchell, il termine e-topia: termine che fa riferimento alla creazione di ambienti virtuali, di interazione e di connessioni elettroniche tra edifici e spazi urbani. Il termine e-topia, chiarisce la studiosa, ne include altri quali la dematerializzazione, cioè lo sviluppo digitale delle città che conduce alla virtualizzazione di molti spazi; la demobilizzazione grazie alla quale la rete consente di ripensare l'utilizzo degli spazi e quindi anche le possibilità di spostamento; il funzionamento "intelligente" degli spazi urbani e conseguente personalizzazione di massa, con edifici interconnessi tra loro, in modo da costituire una sorta di sistema nervoso urbano con sensori e componenti elettronici di vario tipo e funzionalità, tali anche da far sì che le esigenze specifiche degli abitanti possano essere soddisfatte grazie all'interazione tra persone e oggetti [Marzano, 2012].

"Siamo noi che diamo forma ai nostri edifici, ma sono gli edifici che poi modellano la nostra vita e la nostra cultura. L'uomo dà forma alle città, che a loro volta forgianno l'uomo. Attraverso questa citazione di Winston Churchill, Vincenzo Barbieri sottolinea come una smart city sia una città che consente di valorizzare le caratteristiche del territorio, le sue unicità sulla base degli obiettivi definiti a livello politico. Le città, egli osserva, devono essere in grado di valorizzare le proprie caratteristiche, le proprie unicità per poter attrarre cittadini, imprenditori, investitori [Barbieri, 2012]. "Ogni città sedimenta al suo interno un'intelligenza collettiva che si stratifica nel tempo grazie alle tecnologie, le quali consentono ai cittadini di coordinarsi tra loro. Il capitale sociale, ambientale e fisico come strade, edifici, segnaletica, reti tecnologiche, dotazione ICT contribuiscono alla realizzazione di questo patrimonio che rappresenta l'infrastruttura su cui si poggia ogni città. Smart city è dunque un concetto complesso quasi filosofico che può essere espresso solo nel lungo periodo attraverso obiettivi chiari, strategie coerenti attraverso tecnologie adeguate in un continuo processo evolutivo..." [Barbieri, 2012].

"Il percorso per diventare una smart city, osserva Miriam Ruggiero [Ruggiero, 2012], è lungo e articolato perché ogni città ha una specifica condizione di partenza. Il punto di partenza per la costruzione della smart city è una profonda conoscenza della realtà locale, dei bisogni della collettività, delle criticità e della situazione che deve essere gestita. Per compiere il percorso verso la città intelligente, è necessario fare ricerca nei fattori e nelle tecnologie abilitanti, una ricerca interdisciplinare che si basi su forti competenze specifiche tecnologiche, economiche e sociali per arrivare alla definizione di una metodologia che possa sfruttare in modo coordinato tutte le competenze specifiche.

Gli obiettivi devono essere raggiungibili, quantificabili, condivisi tra tutti gli stakeholder e definiti nel tempo. Si deve poi passare all'elaborazione di un piano strategico e di una road map con una quantificazione degli investimenti e dei possibili ritorni e, infine, si deve costruire un sistema di indicatori per monitorare il progetto, "misurarne" le componenti, le lacune, i progressi, le tendenze positive, quelle negative e i passi che ancora restano

da compiere. La misurabilità deve monitorare performance, efficacia e sostenibilità. Per fare questo la città deve investire su professionalità qualificate e competenti, che siano in grado di gestire processi innovativi, che abbiano capacità relazionali e che sappiano guardare lontano con interventi basati su un approccio complessivo e non occasionale.

“La realizzazione di una smart city è il risultato dell’impegno e della collaborazione di diversi soggetti pubblici e privati, che detengono la conoscenza, condividono processi, producono innovazione. È così che si costituisce una “task force sinergica in cui tutti (enti pubblici, aziende, cittadini, banche, istituti di ricerca, università, ecc.) concorrono ad individuare soluzioni per la città, frutto di partecipazione e intelligenza collettiva” [Ruggiero, 2013].

Non esiste dunque un modello universale di smart city: si devono elaborare modelli innovativi, trasversali, realizzabili, misurabili, replicabili, flessibili e finanziabili, basati su caratteristiche intrinseche della città, su efficienza, crescita e vivibilità. È fondamentale poter bilanciare le due dinamiche - top down e bottom up - in modo da riuscire a ottenere i servizi migliori per le persone che vivono in città. Quindi è richiesta sia un’elevata capacità di valutare le singole situazioni mettendo in atto risposte specifiche, sia la capacità di elaborare protocolli, che riescano, successivamente, a prescindere dal particolare. Inoltre si deve uscire dai sistemi altamente qualificati ma verticali per entrare in un’ottica orizzontale, trasversale, che sappia coinvolgere tutti gli ambiti in modo integrato (IT, pianificazione territoriale, sociale, istruzione, ecc.) e che sappia ottimizzare costi e risorse.

Una smart city è una città senziente che conosce quello che ha e che prende decisioni sulla base di informazioni aggiornate, certe e condivise. La città intelligente è in grado di gestire una grande complessità di dati eterogeni. La città intelligente è un luogo dove dati cartografici digitalizzati vengono ‘mesciati’ e integrati con informazioni provenienti da diversi soggetti pubblici (Comuni, Agenzia del Territorio, Camere di Commercio, Aziende dei servizi, ecc.), con i dati rilevati dai sensori (centraline di raccolta dati meteo, di qualità dell’aria, contatori elettronici, sul traffico, videosorveglianza, ecc.), con i commenti su Facebook o taggati su Flickr o Twittati. L’integrazione di tali dati permette di ampliare le conoscenze e di ridurre i tempi di reazione rispetto all’accadere dei fatti sul territorio. Una smart city fa circolare la conoscenza e possiede un’intelligenza collettiva grazie alla quale è in grado di rispondere in qualche misura alle esigenze del territorio relativamente a diversi contesti - cittadini, lavoro, imprese e commercio, edilizia, patrimonio, strumenti urbanistici, imposte, verde, istruzione, strade e viabilità - e dei vari soggetti pubblici e privati che rappresentano il tessuto vitale della città [Ruggiero, 2013].

2. Modelli virtuosi del rapporto tra città, giovani e lavoro

Come si è detto, secondo la maggior parte delle teorie, ogni progetto per la costruzione di una città intelligente deve partire dalle persone ed, in questo senso, ancora una volta viene sottolineata l’importanza della componente ‘human’. La smart city non è infatti né fine a sé stessa né una mera applicazione dell’innovazione tecnologica, bensì un percorso mirato a rendere la vita delle persone migliore, a semplificarla, a consentire una focalizzazione verso il benessere complessivo, a creare opportunità di lavoro e di inserimento per i cittadini nel tessuto della città stessa. E per offrire opportunità, soprattutto ai giovani, le imprese devono crescere e assumere, presentando occasioni interessanti ed innescando un circolo virtuoso di sviluppo economico [Piaggio, 2012]. La smart city significa allora sostenere le imprese a creare prodotti attraenti e vendibili, lavorando in stretto collegamento con la ricerca, usando l’innovazione per migliorare la vita delle persone. Il cammino di una città verso un modello smart, come sottolinea Gloria Piaggio, coordinatrice del progetto “Genova smart city”, è rappresentato dalla capacità di creare un circolo virtuoso tra alcune caratteristiche fondamentali: qualità della vita, sviluppo economico, sostenibilità, ricerca e hi-tech, leadership locale, pianificazione integrata [Piaggio, 2013].

Proprio il Comune di Genova, al fine di consolidare e facilitare il processo di trasformazione in città intelligente, ha intrapreso, insieme ad alcuni partner fondamentali, Enel Distribuzione, l’Università degli Studi di Genova e l’Associazione Genova Smart City, un percorso virtuoso che stabilisce un’alleanza tra enti di ricerca, imprese, istituzioni, finanza e cittadini. Tutti debbono lavorare insieme per individuare obiettivi, metodi e strumenti per raggiungere una dimensione smart della città. L’obiettivo è quello di trasformare gli spazi, gli edifici, l’energia, la trasparenza, la mobilità in chiave ecocompatibile, integrando e captando finanziamenti, capacità industriali, mestieri, servizi e solidarietà. Il percorso verso la smart city cerca di tenere uniti gli elementi della qualità della vita della città con lo sviluppo economico della stessa, che sta a significare occasioni di crescita e di business per le aziende, che si trovano sul territorio genovese, con riflessi positivi sul fronte dell’occupazione di qualità, e dunque una città che funziona meglio per tutti [Comune di Genova, 2013].

È evidente come il tema relativo all’esistenza di un circolo virtuoso tra la città, l’ambiente, il territorio, i cittadini e la qualità della vita debba essere trattato oggi in relazione ad uno dei problemi più rilevanti della società: il problema dell’occupazione. All’interno di questa prospettiva, la maggior parte dei percorsi e dei progetti in diverso modo legati al tema delle città intelligenti si pongono in maniera significativa i problemi riguardanti il lavoro, l’employability, lo sviluppo dell’occupazione soprattutto giovanile.

Secondo la Camera di Commercio di Milano, il progetto teso a fare di questa città una smart city potrebbe portare alle imprese risorse fino a 1,5 miliardi di euro, grazie all’attivazione di nuovi servizi nell’ambito della tecnologie e della comunicazione, della mobilità e dell’efficienza energetica. Si parla, al riguardo, di banda larga e ultra larga fissa e mobile, dell’abilitazione ai pagamenti elettronici e dell’ottimizzazione dei consumi energetici. “Ma cosa chiedono le imprese milanesi al progetto Smart City? Innanzitutto il potenziamento della rete wi-fi, seguito da una serie di iniziative nell’ambito della Green Economy volte a favorire il risparmio di risorse, unitamente allo sviluppo di iniziative di responsabilità sociale per i lavoratori” [Barone, 2013]. Il Comune di Milano e la Camera di Commercio sono impegnati nella costruzione di partenariati stabili tra realtà diverse, nel campo della ricerca e

dell'innovazione sociale, del business e della finanza con l'obiettivo di costruire realmente una città intelligente verde ed inclusiva per i cittadini.

L'Europa che vuole parlare ai cittadini e risolvere i loro problemi a partire da quello dell'occupazione, scrive Il Sole 24 Ore [Il Sole 24 Ore, 29 giugno 2013], esce rafforzata dall'ultimo Consiglio europeo del 27 e 28 giugno. I ministri del Lavoro della UE, riuniti in un vertice a Berlino il 3 luglio del 2013, individuano in particolare cinque assi di intervento volti a rendere operativo il Piano Youth Guarantee: lo scambio di esperienze sulle misure che hanno il maggiore impatto sulla riduzione della disoccupazione giovanile, promuovere la mobilità dei giovani alla ricerca di lavoro; migliorare il funzionamento dei centri per l'impiego e dei sistemi nazionali che si occupano di educazione e formazione, prendere misure atte ad aumentare l'utilizzo di tutti gli strumenti finanziari disponibili [Pogliotti, 2013]. Grazie al Piano Garanzia i giovani sotto i 25 anni devono ricevere quindi dagli stati membri una buona offerta di lavoro, un corso di perfezionamento, l'apprendistato o il tirocinio entro 4 mesi dalla disoccupazione o dall'uscita dal sistema educativo. "I singoli stati membri devono occuparsi anche della prevenzione di lunghi periodi di disoccupazione o di inattività, e del miglioramento dell'occupabilità. Con la consapevolezza che non può esserci un unico approccio che possa andare bene per tutti, vista la varietà delle sfide con cui devono misurarsi i diversi Paesi [Pogliotti, 2013].

Il Programma Youth Guarantee viene dunque potenziato dal Consiglio europeo del giugno 2013 che offre all'Italia la possibilità di avere a disposizione una cifra complessiva di 1,5 miliardi da spendere nel biennio 2014 - 2015. Dopo aver ottenuto dall'Europa l'incremento dei fondi per la Garanzia Giovani, il Governo italiano intende muoversi potenziando in primo luogo i servizi per l'impiego e migliorando l'incontro tra domanda e offerta di lavoro, anche attraverso la partnership tra pubblico e privato [Il Sole 24 ore, 29 giugno 2013]. Gli altri tasselli della strategia del Governo per la lotta alla disoccupazione sono il rafforzamento dell'apprendistato; la staffetta generazionale tra giovani e anziani; un pacchetto di incentivi per favorire le assunzioni e le forme imprenditoriali più innovative; una nuova spinta alle politiche attive del lavoro vale a dire ai tirocini formativi ed agli stages; l'adozione di misure straordinarie e temporanee che potranno essere approvate per aumentare la flessibilità in entrata nel mercato del lavoro.

"I giovani, scrive Paolo Serreri [Pepe e al., 2012], devono affrontare oggi un numero crescente di sfide e di scelte difficili. Per iniziare a costruire la loro vita professionale hanno bisogno di informazioni relative ai percorsi formativi ed alle opportunità lavorative. È necessario dunque sviluppare l'offerta di servizi di qualità in materia di orientamento professionale e di assistenza alle prospettive di formazione e di impiego. Occorre fornire nuovi e complessi strumenti che consentano ai giovani di orientare le loro scelte, optando per percorsi coerenti con le aspettative personali ma anche con le esigenze del mercato del lavoro al fine di evitare circoli viziosi senza uscita". La crescita economica non può fare a meno dei giovani né i giovani della crescita. È necessario un impulso significativo alla facilitazione delle transizioni tra i sistemi educativi e il mercato del lavoro. Impulso che può essere attuato mediante strumenti quali un sistema adeguato di servizi per l'impiego, l'offerta di servizi di orientamento di qualità, la possibilità di intraprendere la strada dell'autoimprenditorialità e di acquisire esperienze lavorative già durante il ciclo scolastico.

2.1 Porta Futuro

All'interno di una prospettiva volta a favorire l'inserimento dei giovani, nel lavoro e nella società, si colloca la sfida di Porta Futuro, il nuovo centro di orientamento e di servizi per il lavoro che la Provincia di Roma ha voluto lanciare come piano di innovazione nel cuore della Capitale d'Italia. Porta Futuro rappresenta un investimento, soprattutto per i giovani, un servizio che vuole guardare oltre la crisi proponendo politiche attive per l'inclusione sociale e ponendo al centro di queste politiche la formazione, l'orientamento ed il lavoro.

L'inaugurazione del Centro di orientamento, formazione e accompagnamento al lavoro di Roma, *Porta Futuro* [www.portafuturo.it], è avvenuta il 5 luglio del 2011: giorno in cui la Provincia di Roma ha aperto le porte di questo incubatore tecnologico volutamente nel cuore della Capitale, nel quartiere Testaccio, e proprio in un momento di grave crisi economica ed occupazionale, per dimostrare come esso potesse rappresentare una sfida rivolta al domani ancor prima che un innovativo progetto. Nonostante la sua breve vita, la strategia di questo Centro è stata già stata notata e premiata dall'Unione Europea. 'Porta Futuro' è diventato infatti uno dei quattro hub europei per l'inserimento dei giovani nel mondo del lavoro grazie al progetto *Your first Eures job*. "Porta Futuro, scrive Paolo Serreri, è un nome che contiene una metafora augurale: una porta aperta sul futuro, sul futuro lavorativo dei giovani, giovani a cui oggi in Italia, tra precarietà e disoccupazione, il futuro è negato; sul futuro degli adulti maturi (tali sono per lo più quelli che frequentano, ad esempio, il servizio del Bilancio delle competenze in PF) il cui ricco patrimonio di competenze, piuttosto che essere rivalutato, rimesso in circolo e reinvestito in funzione della crescita, sempre più spesso, viene gettato in acqua, con sconsiderate manovre di alleggerimento: sia nel mare piatto della stagnazione, sia nella tempesta della recessione" [Pepe e al. 2012].

"Porta Futuro, osserva Carmelo Ursino [Ursino, 2012, 261], ... è un servizio pubblico fortemente innovativo che mira all'empowerment di cittadini e imprese... un luogo fisico, ma anche virtuale, dove l'utente è posto al centro dei servizi organizzativi e dove l'innovazione e le moderne tecnologie legate al mondo dell'ICT rappresentano leve strategiche rilevanti per il raggiungimento degli obiettivi. Porta Futuro investe su creatività, innovazione e sviluppo per fornire strumenti utili ad affrontare i cambiamenti economici, sociali, demografici e le variazioni strutturali del mercato del lavoro".

Porta Futuro offre servizi completamente gratuiti e mette a disposizione degli utenti milleottocento metri quadrati di area open space facilmente accessibile perché aperta al pubblico dalle 9 mattina fino alle 19 di sera, durante la

settimana, e sino alle 20 nel week end. Esso mette a disposizione servizi multimediali avanzatissimi visto che l'intera area è coperta dalla rete WiFi della Provincia di Roma, tali servizi permettono di erogare attività molto differenti l'una dall'altra ma fortemente integrate tra loro.

Portafuturo ha le sue radici nel modello spagnolo 'Porta 22', attivo a Barcellona, con ottimi risultati, dal 2003. Porta 22 è un'agenzia municipale di formazione, orientamento e occupazione che registra circa 66 mila utenti all'anno. "Circa il 64% dei disoccupati, per la maggior parte ultraquarantenni, trova una nuova occupazione grazie a Porta 22. Il cuore pulsante dell'attività è rappresentato da un sofisticato software al quale è affidato il compito di incrociare la domanda e l'offerta del mondo del lavoro. In pratica l'aspirante lavoratore, o anche chi voglia perfezionare le proprie abilità professionali, inserisce i propri dati indicando il percorso scolastico, quello professionale, le eventuali esperienze lavorative" [Ursino, 2012, p 264]. A seconda del profilo tracciato, al candidato vengono indicati, in ordine di affinità, alcuni mestieri, scelti fra gli 865 del database. "L'offerta formativa, precisa Ursino [2012, 264], è articolata in 4 punti: sviluppo delle competenze, strategia, gestione di un cambio professionale e dialogo con il mercato. A questi punti si aggiunge un'applicazione del software che consente di esercitarsi nei colloqui di lavoro e l'affiancamento di un tutor che attraverso i social media gestisce i progressi del frequentante".

In Italia Porta Futuro è il primo centro che raccoglie in un unico luogo i servizi dedicati al lavoro, all'orientamento e alla formazione. Questo Centro è dotato di un software in grado di unificare i database del mercato del lavoro di tutto il territorio, favorendo il matching tra domanda e offerta e di tracciare, per il cittadino, il profilo professionale più adatto alle proprie caratteristiche. Il progetto, unico sia per le finalità che per l'impianto tecnologico che ne è alla base, è stato realizzato dalla Provincia di Roma in collaborazione con il partner tecnologico ETT, già partner dello sviluppo del Sistema Informativo Lavoro (SIL) 'Bussola' della Provincia di Roma. Porta Futuro dispone dunque di un Sistema Informativo avanzatissimo collegato alla banca dati della Provincia di Roma e degli altri Centri per l'Impiego del territorio. Grazie a questa banca dati, il Centro può interoperare con i vari sistemi della rete dei servizi per il lavoro, a livello nazionale, e raccogliere le esigenze contingenti del tessuto economico locale, promuovendo azioni in grado di favorire la crescita in tema di formazione e lavoro. Il fatto di far riferimento ad un'unica banca dati fa sì che non vi siano rischi di duplicazione delle informazioni; che siano possibili la massima tracciabilità, un matching coerente e la possibilità di analizzare i dati secondo una procedura estremamente semplificata.

I dati raccolti dal Centro sono accessibili alle diverse tipologie di utenti grazie all'uso di più canali: il sistema informativo, lo staff di Porta Futuro, intranet, il portale del Centro ed i social media come Twitter, Facebook e LinkedIn. Il primo obiettivo di Porta Futuro è quello di insegnare agli utenti il "mestiere di trovare un lavoro". Gli strumenti attraverso i quali si insegna questo 'mestiere' sono: il Curriculum assistito; il video-curriculum in formato europeo; l'orientamento individuale e di gruppo; il bilancio delle competenze; i seminari su come affrontare un colloquio di selezione; l'inglese di base; la contrattualistica del lavoro; i percorsi di *personal branding*.

Nella misura in cui intende rispondere alle finalità espresse nella strategia *Europa 2020*, Porta Futuro risponde anche a specifici principi ed obiettivi delle politiche europee per la crescita del lavoro e dell'occupazione. Questa innovativa struttura è chiamata infatti a far fronte alle necessità di: costruire un mercato del lavoro più efficiente, ovvero in grado di mettere in contatto domanda e offerta in tempo reale; contribuire alla costruzione di una rete tra imprese, cittadini e mercato del lavoro; diventare un punto di riferimento per la formazione e l'apprendimento continuo. Si tratta di un luogo nel quale i lavoratori possono costantemente aggiornarsi, migliorare i propri curricula, incrementare i propri punti di forza e lavorare sui propri punti deboli, con il fine ultimo di sentirsi parte integrante di una comunità in continua evoluzione.

Un aspetto fondamentale di Porta Futuro è il suo intento di essere vicino ai giovani: ai giovani che non lavorano così come a quelli impegnati nella scuola e nell'università. Porta Futuro nasce con l'obiettivo prioritario di coinvolgere direttamente i ragazzi. "Chiunque si presenti per la prima volta a Porta Futuro viene accolto presso una reception, dove i consulenti fanno un primo screening che si può riassumere nella semplice domanda: 'Perché questa persona è qui? Di cosa ha bisogno?'. Tecnicamente si chiama 'analisi del fabbisogno', ed è il primo passo per capire verso quali servizi indirizzare il nuovo arrivato. Sempre alla reception viene compilato un modulo per ottenere immediatamente lo username e la password personali per accedere al software da quel momento in poi" [Repubblica degli stagisti, 10 maggio 2012]. La caratteristica di Porta Futuro infatti è la specificità del sistema informatico che permette di registrarsi, caricare il proprio cv e visualizzare un gran numero di opportunità sul territorio: offerte di lavoro e di stage, corsi di formazione, bandi di concorso. Dopo questa operazione, sulla base delle informazioni che ciascun utente ha inserito il sistema indica le aree professionali coerenti – fase di definizione del 'portafoglio competenze' - e poi mette direttamente in evidenza le offerte di lavoro compatibili.

Porta Futuro comprende dunque al suo interno un centro per l'impiego, che rende possibile l'iscrizione alle liste di inoccupazione e di disoccupazione e l'attivazione di stage e contratti di lavoro. La rosa degli altri servizi offerti da Porta Futuro prevede bilanci di competenze, sessioni di orientamento individuale e di gruppo chiamate 'jobclub', con la presenza di un orientatore ogni dieci persone. Due risorse si occupano specificamente di autoimprenditorialità, offrendo a chi aspira a mettersi in proprio una prima valutazione dei progetti di start-up e consulenza per la stesura del business plan. Una stanza è attrezzata per la registrazione dei videocurriculum: chi ne vuole fare uno si iscrive ed aspetta di essere chiamato per la videoregistrazione attraverso la strumentazione di Porta Futuro. Altri eventi frequenti nello spazio sono seminari, presentazioni aziendali, dibattiti su libri.

"L'80% delle possibilità di trovare lavoro, spiega Silvio Petrassi [2012], sono legate al meccanismo del passaparola ed oggi questo meccanismo agisce in maniera estremamente significativa nei canali della digitalizzazione. Oggi il passaparola avviene anche e soprattutto all'interno di Facebook, Twitter e LinkedIn; noi insegniamo agli utenti ad entrare ed a scambiare informazioni all'interno di questi canali, insegniamo loro a creare

ed a cercare relazioni tra le diverse piattaforme cui Porta Futuro è legata... Saper far uso delle piattaforme digitali significa essere facilitati nella ricerca del lavoro e noi insegniamo agli utenti sia a navigare all'interno di queste piattaforme che a presentarsi, eventualmente in un momento successivo, alle aziende con il giusto curriculum".

"A tre mesi dal taglio del nastro di Porta Futuro, scriveva il Sole 24 Ore del 12 ottobre 2011, sono oltre 2377 le persone iscritte, soprattutto giovani (946, il 40% nella fascia di età tra i 26 ed i 35 anni), ma anche un gran numero di adulti (545 di età compresa tra i 36 e i 45 anni e ben 357 oltre i 45 anni). La maggior parte degli iscritti ha un titolo di studio medio-alto ed è disoccupata o inoccupata [Il Sole 24 Ore, 12 ottobre 2011]. In occasione del Forum PA 2012, tenutosi a Roma dal 16 al 19 maggio, i rappresentanti di Porta Futuro espongono i dati aggiornati sulle attività e sulla frequenza del Centro. In questo momento, gli utenti registrati nella struttura sono 7763. Le fasce di età più rappresentate sono quella compresa tra i 26 ed i 35 anni di età (39%) e quella compresa tra i 36 ed i 45 anni di età (23%). Il 22% degli iscritti ha un'età inferiore ai 26 anni ed il 16% degli iscritti ha un'età superiore ai 45 anni. Per ciò che riguarda il grado di istruzione degli iscritti, sempre nel 2012, si rileva che il 39% di questi è in possesso del diploma di istruzione secondaria superiore mentre il 30% ha conseguito la laurea [Portafuturo, 2012].

2.2 Il Centro Luiss EnLabs

Si è tenuta il 4 aprile a Roma, alla stazione Termini, l'inaugurazione del Centro Luiss EnLabs [<http://www.luissenlabs.com/>], nato dalla joint venture tra l'università Luiss ed EnLabs, uno dei più importanti incubatori d'impresa in Italia, con la collaborazione di Wind. L'evento è stata l'occasione in cui protagonisti dell'innovazione si sono incontrati per uno scambio di vedute sul mondo dell'imprenditoria. Vi hanno partecipato attivamente esponenti delle istituzioni, dell'industria, della cultura e dell'informazione. A presentare il progetto ci sono stati il direttore dell'Università, Pierluigi Celli, il fondatore di Luiss EnLabs, Luigi Capello e il direttore delle Public Relations di Wind, Massimo Angelini.

Quella che può essere definita una vera e propria 'fabbrica delle startup' a livello nazionale e internazionale, ha l'obiettivo di fornire agli studenti un'importante opportunità nel mercato del lavoro e di trasformare le migliori idee di business in iniziative economiche di successo. Luiss EnLabs è aperto non solo agli studenti Luiss, ma a quelli provenienti dalle Università di tutta Italia, con particolare attenzione per chi arriva da facoltà come medicina, ingegneria, fisica, matematica, informatica, biologia.

Luiss EnLabs è uno spazio di oltre 1500 metri quadri - messo a disposizione dal Gruppo Ferrovie dello Stato - dotato di 120 postazioni di lavoro. Mini uffici, scrivanie e connessione a banda larga all'interno di box trasparenti, che a pieno regime possono accogliere fino a 50 startup, le quali hanno la possibilità di avviare le loro attività beneficiando della contaminazione positiva facilitata proprio dalla vicinanza offerta dall'incubatore, in una location strategica, architettonicamente pregiata e perfettamente attrezzata dal punto di vista tecnologico. Un incubatore, infatti, è uno spazio fisico dove persone possono affittare una scrivania per il coworking o partecipare ad un programma di incubazione per sviluppare la propria idea, uno spazio così progettato dà l'opportunità di confrontarsi con diverse realtà imprenditoriali e di condividere con loro i propri progetti e idee, guadagnando forza e sviluppando sinergie. L'obiettivo dell' incubatore è quello di fornire ai team gli strumenti necessari per avviare un'impresa. Luiss EnLabs accelera il lancio di start-up fornendo il finanziamento iniziale, il tutoraggio e il sostegno nella ricerca di ulteriori risorse finanziarie. Durante il periodo di incubazione le startup possono lavorare a tempo pieno sul progetto presso il Centro. I mentori si occupano di sostenere le squadre in tutto il percorso e di eseguire le revisioni settimanali.

La fabbrica delle startup mette a disposizione la sua rete di relazioni di co-lavoratori, dando loro la possibilità di partecipare agli eventi ospitati dal programma di incubazione. Co-working fornisce anche un risparmio economico, riguardo alle spese classiche che si affrontano in un ufficio privato, quindi elettricità, pulizie, spese condominiali, ecc.

È un modello basato sull'incontro di tre diversi soggetti: imprese, università e finanziatori con altrettanti fattori: spazi fisici, competenze e risorse economiche. Il programma prevede una preventiva selezione delle migliori idee imprenditoriali, che vengono successivamente inserite in un programma di accelerazione e incubazione della durata di sei mesi. Durante questo periodo le start-up hanno la possibilità di usufruire degli spazi di lavoro e di una serie di servizi tra cui un corso di imprenditoria e l'assistenza da parte di mentor qualificati, oltre a un investimento finanziario a fronte dell'acquisto di partecipazioni. Il supporto finanziario alle start up meritevoli, «fino alla dimensione che possa interessare il venture capital», è assicurato da LVenture Group, società quotata alla borsa di Milano. Inoltre Luiss EnLabs stringerà accordi con partner terzi, come gruppi industriali o banche, per accrescere le possibilità di finanziamento delle aziende incubate e per ottenere accordi di mercato favorevoli.

In questo modo, l'Università Luiss prosegue sulla strada della ricerca e dell'innovazione, già tracciata con successo nell'ambito della sua didattica, e sempre secondo quella cultura d'impresa che la caratterizza, nella convinzione che i giovani debbano oggi 'inventarsi' il proprio lavoro, oltre che affidarsi ai tradizionali canali di occupazione offerti dal mercato.

L'idea di questo ambizioso progetto nasce nel 2010, quando Luigi Capello, cofondatore di Italian Angels for Growth, il maggiore gruppo di Business Angels in Italia, attraverso la sua esperienza diretta a San Francisco ha avuto modo di apprezzare la struttura e il funzionamento degli incubatori della Silicon Valley. In un'intervista, rilasciata il 3 Aprile alla testata quotidiana on-line www.formiche.net, Capello dichiara: "Mi sono meravigliato di come questi incubatori creassero degli elementi positivi per la nascita di nuove aziende mentre in Italia questo tipo di attività erano in uno stato embrionale. Anziché portare le start up in America decisi allora di importare in Italia un

sistema già consolidato. Il 30 novembre del 2010, giorno dell'apertura di EnLabs si presentarono 300 persone, segno che su Roma c'era un tessuto pronto di persone dedite alla tecnologia e con l'ambizione di aprire una start up" [Capello, 2013].

È stato proprio Capello a presentare, durante l'evento di inaugurazione dell'incubatore, i dieci team di ragazzi che già occupano una postazione all'interno della stazione. Aziende come CoContest, piattaforma di crowd sourcing rivolta ad architetti e designer che permette a chi deve arredare casa di lanciare una gara fra i designer della community, cinquemila architetti di novanta paesi di tutto il mondo. Come Soundreef, che propone un sistema di gestione dei diritti d'autore alternativo alla Siae, la cosa più innovativa è il pay-per-play con cui mostra all'autore dove viene suonata la propria canzone, ed è così che un utente del Nevada con 15 brani è riuscito a incassare 1.300 dollari negli ultimi sei mesi. Come Bulsara, specializzata nella pubblicità in luoghi non convenzionali, tra cui i bagni dei cinema o dei centri commerciali. O come Atooma, applicazione che permette di creare speciali comandi per rendere più intelligente il proprio smartphone, con cose del tipo: leggimi i messaggi se sono in auto, spegni il wifi quando esco di casa. Ideata da un'ex studentessa della Luiss, la 27enne Francesca Romano. È nominata qualche settimana fa migliore app del mondo al Mobile World Congress di Barcellona.

Tutte società incubate o finanziate da LVenture. Ma in Luiss EnLabs c'è posto anche per aziende esterne alla scuderia. Non è necessario aderire al progetto di incubazione per essere ammessi, infatti questo è anche uno spazio di coworking, dove si può accedere pagando una quota di iscrizione. Le scrivanie sono a disposizione anche di professionisti attivi nel mondo della tecnologia, come avvocati o pubblicitari. Nel tentativo di creare un piccolo ecosistema digitale pronto ad accogliere diverse startup. I valori della fabbrica delle startup sono:

- Community, Luiss EnLabs infatti è parte di un network internazionale fortemente connesso alla Silicon Valley, in tal senso il Centro vuole essere una piattaforma per lo sviluppo di una comunità dinamica e innovativa a livello globale;
- cooperazione, per favorire l'apprendimento collettivo attraverso la cooperazione e il contatto tra le persone e lo spazio dell'incubatore, in particolare, l'ambiente aperto promuove l'interazione tra colleghi di lavoro e prevede sinergie, scambio di idee, servizi e molto altro,
- Innovazione e Collaborazione, tese a favorire la collaborazione tra persone che si impegnano per cambiare qualcosa nel mondo.

Le imprese partecipate fino ad oggi da Luiss EnLabs sono 18 di cui 11 provenienti dal programma di accelerazione e l'impatto dal punto di vista occupazionale effettivo coinvolge circa 150 persone. In soli due anni, dall'attivazione originaria del progetto, già 100 persone hanno trovato lavoro. È possibile sintetizzare il contributo che offre l'incubatore in tre elementi: semplicità e facilitazioni nel trasformare idee di business in idee di successo, mentori che mettono a disposizione le proprie conoscenze tecniche e imprenditoriali e possibilità per le aziende incubate di poter usufruire di finanziamenti.

Ed è così che la stazione Termini, luogo di partenze e luogo di arrivi, di saluti, di addii e di ritorni si trasforma per molti giovani nel luogo di nascita di nuove speranze, dove è possibile cominciare un nuovo viaggio con una valigia piena di idee e di speranze.

Conclusioni

Il nostro saggio ha cercato di descrivere un percorso possibile e realistico di evoluzione delle città verso un modello smart: modello che vede la costruzione di significative e complesse relazioni sistemiche tra istituzioni pubbliche e private; cittadini e territori; infrastrutture e sistemi di connessioni reali e virtuali; interventi relativi all'utilizzo ottimale delle risorse, al risparmio energetico ed allo sviluppo sostenibile; problemi legati al lavoro ed all'occupazione di giovani ed adulti; riflessioni ed azioni possibili sulla valorizzazione delle attività dell'uomo e del contesto in cui quelle attività si esplicano.

All'interno di questa prospettiva il percorso delineato dal nostro saggio sembra riprendere, seppur in termini metaforici, il percorso compiuto dall'Associazione Prima Persona attraverso il viaggio 'ItaliaViva'. "ItaliaViva è un percorso, intrapreso su tutto il territorio nazionale, alla scoperta di realtà umane, sociali, imprenditoriali e associative di rilievo. Esperienze positive, determinate da creatività e spirito innovativo. Esempi di coraggio e determinazione, guidati dalla volontà di costruire un mondo migliore" [http://www.primapersona.eu/?page_id=696]. Il viaggio di Prima Persona, Associazione presieduta da Luigi Pittella, è iniziato in Umbria il 15 luglio del 2012 e si è concluso a Roma il 28 luglio dello stesso anno. Il viaggio ha toccato gran parte delle regioni italiane: dalla Sicilia al Piemonte, dalla Puglia alla Lombardia.

Le storie incontrate e raccontate da ItaliaViva sono tutte differenti l'una dall'altra ma tutte "accumunate dal desiderio di mettere a frutto creatività e capacità innovative per trovare un proprio spazio, cosa non facile in un periodo complesso e in costante evoluzione come quello attuale. Ogni iniziativa è scaturita da un bisogno forte: reagire alla criminalità organizzata, contribuire alla salvaguardia dell'ambiente, dare ai cittadini la possibilità di scegliere come investire parte del denaro pubblico" [<http://www.primapersona.eu/?p=1139>]. Ed ogni storia ha reso partecipi di volta in volta rappresentanti di organizzazioni, aziende, enti locali, progetti culturali e imprenditoriali, associazioni di categoria, organizzazioni del terzo settore, esponenti istituzionali e altri soggetti legati ai territori ed alle tematiche affrontate, per individuare insieme nuove modalità di azione.

L'iniziativa ha coinvolto tante vicende e tante figure come, ad esempio, "Wainer Molteni, ex invisibile, oggi consulente al Comune di Milano per il reintegro dei senza tetto e guida dei Clochard alla riscossa. Come Giovanni Luca Ciaffoni, sviluppatore di un'applicazione rivoluzionaria per i non vedenti. Ciaffoni ha solo 36 anni, ed è stato

ringraziato per la sua opera da Tim Cook in persona, AD Apple” [http://www.primapersona.eu/?page_id=696]. L’iniziativa ha egualmente coinvolto i comuni virtuosi, come “Minervino di Lecce o Capannori, con amministrazioni attente alla salute e alla sicurezza dei cittadini. E poi ancora, attività imprenditoriali che costituiscono un vanto per il Paese, come Lodaracing Project di Terni, specializzata nella produzione delle moto da corsa. I suoi prodotti sono sulle piste dei principali MotoGP mondiali” [http://www.primapersona.eu/?page_id=696].

L’ultima tappa del viaggio ha toccato il Convitto Nazionale Vittorio Emanuele II di Roma, istituzione storica ma all’avanguardia nella costruzione di nuovi percorsi formativi [<http://www.primapersona.eu/?p=913>]. Il Convitto rappresenta un esempio di eccellenza e di avanguardia, in primo luogo, per ciò che riguarda il suo progetto educativo volto a formare i giovani come cittadini europei e come portatori di una forte coscienza civica; in secondo luogo, l’Istituto costituisce un modello paradigmatico nella misura in cui propone ed adotta sistemi attivi di apprendimento e di formazione tesi a fare degli studenti i protagonisti del loro percorso di costruzione delle conoscenze. Infine, l’adozione da parte del Convitto di sofisticati sistemi di digitalizzazione nei processi di insegnamento e di apprendimento ha l’obiettivo di preparare i giovani ad un modello di società innovativa in cui costruire nuove opportunità e nuovi spazi di lavoro e di vita.

A partire da questa prospettiva, il percorso di progettazione e di costruzione della città intelligente chiama in causa non solo fattori strutturali, infrastrutturali, legati all’ambiente, ai servizi, alla vita ed al lavoro dei cittadini, alle relazioni tra gli individui e tra gli individui e la città stessa. I problemi legati ai processi di evoluzione verso la smart city sembrano chiamare in causa in primo luogo ed in termini sostanziali l’evoluzione di modelli educativi, formativi e culturali. E’ impossibile concepire una qualsiasi forma di evoluzione della società della conoscenza prescindendo dall’importanza che, all’interno di questa società, rivestono modelli di conoscenza e di apprendimento che accompagnano l’individuo durante tutto il corso della sua vita e che permeano i diversi momenti della sua vita.

L’importanza riconosciuta ai modelli di apprendimento lifelong e lifewide rappresenta, in primo luogo, uno strumento per fronteggiare la complessità, favorire la democrazia e l’inclusione sociale. La capacità, da parte di ogni individuo, di saper costruire sempre nuove conoscenze e di rimettere in gioco e ridefinire continuamente le proprie competenze sono essenziali affinché questo individuo sia in grado di adattarsi e riadattarsi alle condizioni di estrema mutevolezza della società contemporanea. Questa stessa capacità di adattamento continuo è essenziale affinché, in un momento di grave crisi economica quale è quello attuale, l’individuo sia in grado di restare o di rientrare nel mercato del lavoro. Nel momento in cui il lifelong learning viene posto come un punto di vista attraverso il quale comprendere la complessità sociale e favorire lo sviluppo democratico, l’educazione permanente diventa la “condizione per i diritti di cittadinanza, strumento di convivenza civile e risorsa per lo sviluppo economico-sociale dei Paesi” [Alberici, 2008]. “Se è vero infatti che oggi, di fronte alle grandi sfide della vita, fame, pace, lavoro, inclusione/esclusione, solidarietà, libertà, equità è necessario puntare sullo sviluppo umano e sulla possibilità, per un numero sempre maggiore di donne e di uomini, di saper produrre pensiero riflessivo, divergente, innovativo, allora la formazione cambia radicalmente natura, genere, e diviene un processo finalizzato sempre più alla crescita di soggetti responsabili e autonomi, proattivi. Ne deriva la necessità di puntare sulla formazione ‘for all’, come valorizzazione delle risorse umane, facendo leva sulla centralità del soggetto, sui suoi saperi e sulle sue competenze di vita e di lavoro, sulla sua riflessività, sulle capacità sociali e di relazione, sulla responsabilità, su quelle che possiamo definire come le competenze strategiche per la vita in quanto è con il loro possesso, sviluppo e crescita lifelong che donne e uomini divengono capaci di affermare i loro diritti di cittadinanza sostanziale” [Alberici, 2005].

Bibliografia

Alberici A., Prefazione a S. Cerrai e Beccastrini S., Continuando a cambiare. Pratiche riflessive per generare e valorizzare le competenze nelle organizzazioni, Arpat, Firenze, 2005, VII.

Alberici A., La possibilità di cambiare. Apprendere ad apprendere come risorsa strategica per la vita, FrancoAngeli, Milano, 2008, 19.

Barbieri, V., Smart city o cittadini smart? Geoexperience, 8, 2012, 8.

Barone, T. Milano Smart City: tutti i vantaggi per le PMI, PMI.it, 13 maggio 2013, testo consultabile al sito: <http://www.pmi.it/economia/lavoro/news/65836/milano-smart-city-tutti-i-vantaggi-per-le-pmi.html>

Bocchi G. e Ceruti M. (eds), La sfida della complessità, Feltrinelli, Milano, 1985.

Campbell T., Oltre la Smart City, Intervista a Tim Campbell, Forum PA 2013, Roma 28 - 30 maggio, testo disponibile al sito <http://smartinnovation.forumpa.it/story/70506/oltre-la-smart-city-intervista-tim-campbell>

Capello L., Luiss Enlabs, la Silicon Valley italiana a Roma. Parla il fondatore Capello, Intervista rilasciata a Valeria Covato, 03 Aprile 2013, testo disponibile al sito: <http://www.formiche.net/2013/04/03/luiss-enlabs/>

Ceruti M., Una possibile reinterpretazione del concetto di sistema: i progetti della natura e della storia, in De Masi D. e Pepe D. (eds), I modelli organizzativi tra conoscenza e realtà, FrancoAngeli, Milano, 1989, 201.

Comune di Genova, Associazione Smart City: la città che pensa al suo futuro, 2013, articolo consultabile al sito <http://www.comune.genova.it/articoli/associazione-smart-city-la-cittagrade-che-pensa-al-suo-futuro>.

Ehn P., Design Thinking in the City, Un'intervista a Pelle Ehn a margine del convegno Human smart city, Forum PA 2013, Roma 28 - 30 maggio, 2013, testo disponibile al sito <http://saperi.forumpa.it/story/71796/design-thinking-city-unintervista-pelle-ehn-margine-del-convegno-human-smart-city>

Il Sole 24 Ore, 29 giugno 2013, 2.

Leadbeater, C., Intervento al Convegno Human smart city, Forum PA 2013, Roma 28 – 30 maggio, 2013.

Marzano F., Smart Community e inclusione sociale per l'eco-sostenibilità, Forum PA 2012, Roma 16 - 19 maggio, testo consultabile al sito <https://www.flavia-marzano-smart-city-e-inclusione-sociale-per-l'ecosostenibilita.pdf>

Pepe D., Serreri P. e Alampi G. Un progetto per la formazione e l'orientamento dei giovani: Portafuturo. Formazione e cambiamento, webmagazine sulla formazione, n. 74, novembre dicembre, 2012.

Petrassi S., Intervista rilasciata a D. Pepe il 27 luglio 2012.

Piaggio G., Human Smart Cities: come realizzarle, Forum PA 2013, Roma 28-30 Maggio 2013.

Piaggio G. Genova Smart, Rivista QualEnergia, giugno-Luglio 2012, 35.

Pogliotti G., Ecco le linee dell'azione congiunta della Ue contro la disoccupazione giovanile, Il sole 24 Ore, 3 luglio 2013, testo disponibile al sito: <http://www.ilssole24ore.com/art/notizie/2013-07-03/ecco-linee-azione-congiunta-172116.shtml?uuid=Ab5esvAI>

Pontillo A., A Roma il lavoro si cerca a Porta Futuro, ufficio pubblico a misura di giovani, Intervista rilasciata a La Repubblica degli stagisti, 10 maggio 2012, testo disponibile al sito <http://repubblicadeglistagisti.it/article/porta-futuro-provincia-roma-servizi-seminario>

Portafuturo, Dati e statistiche presentati al Forum PA 2012, Roma 16 – 19 maggio.

Repubblica degli stagisti, A Roma il lavoro si cerca a Porta Futuro, ufficio pubblico a misura di giovani", 10 maggio 2012, testo disponibile al sito <http://repubblicadeglistagisti.it/article/porta-futuro-provincia-roma-servizi-seminario>

Ruggiero M., Cosa vuol dire smart? Sostenibile, partecipata, senziente... Ecco la città che vorremmo, Forum PA 2012, Roma 16 – 19 maggio, 2012, testo disponibile al sito <http://smartinnovation.forumpa.it/story/69436/cosa-vuol-dire-smart-sostenibile-partecipata-senziente-ecco-la-città-che-vorrei>

Stati Generali dell'Innovazione, Carta d'Intenti per l'Innovazione, testo disponibile al sito http://www.statigeneralinnovazione.it/wiki/index.php?title=Carta_d'Intenti_per_l'Innovazione

Ursino C., Il lavoro e la formazione. Un 'brand' per creare 'valore pubblico', in Alessandrini G. e Pignalberi C. (eds), Le sfide dell'educazione oggi. Nuovi habitat tecnologici, reti e comunità, Pensa Multimedia, Lecce e Brescia, 2012.

www.luissenlabs.com/

www.portafuturo.it

www.primapersona.eu/italiaviva/

The Role of the Innovation Poles in the Innovation System of the Region of Calabria

C. Luciano Mallamaci, Domenico Sacca¹
Centro di Competenza ICT-SUD, Polo di Innovazione ICT
Piazza Vermicelli - Università della Calabria, 87036 Rende (CS)
luciano.mallamaci@gmail.com
¹Università della Calabria
Ponte Pietro Bucci, 87036 Rende (CS)
sacca@unical.it

Abstract. *Building an innovation system requires a strong cooperation among government, universities and enterprises (the three helices of a territorial innovation system). The adoption of suitable tools and policies for the integration of the strategies of the three helices produces relevant progresses in terms of innovation capacity of a territory. However, the difficulty of interaction of the above subjects may hinder the integration process. As a consequence, a territory may lose one or more helices of its “innovation engine” taking the risk to fail in creating its own innovation system. The presence in the territory of a “catalytic entity” may facilitate the interaction among government, universities and enterprises. In Calabria the catalyst has been identified in a network of eight Innovation Poles, each of which is a cluster of companies and research organizations operating in the regional strategic sectors: ICT, health, new materials, logistics, transportation, agriculture, cultural heritage, water resources. This paper illustrates organization and governance model of the Innovation System in Calabria.*

Keywords: innovation system, technology transfer, innovation poles, catalyzed triple helix, innovation intermediary, one-stop shop

1. Introduction

The innovation process is not anymore considered as a linear process (scientific discovery, technological application, industrial application) but rather a pervasive process that tends to introduce a permanent innovation in the territory. Therefore, a technology transfer process should be based on the economic, social and cultural environment in which the new technologies are going to be introduced. Such a process should not be based solely on scientific and economic studies and/or desk analysis, but it should directly involve the main public and private strategic operators that are present in the territory, such as local authorities, academic bodies, research centres, large organizations as

well as Small and Medium Enterprises (SMEs), entrepreneurial associations, development agencies, etceteras [Corti 1986, Corti e Mallamaci, 1986a, Corti e Mallamaci, 1986b].

The presence in a territory of a number of subjects performing high quality research activities is not a sufficient condition to promote the innovation at the local level. Indeed, with the progressive rise of the "Open Innovation" model [Chesbrough, 2003, Chesbrough, 2006], each company, especially SMEs, should carefully consider whether or not to invest in research because more and more often it is more convenient to develop *absorptive capacity* of external knowledge [Cohen and Levinthal, 1990] rather than producing internal knowledge. In any case, if we look at Calabria, a less favoured region in Southern Italy, we notice that the increased investment in R&D (thanks to the considerable availability in the last decades of national and EU research funds) does not seem to have produced any appreciable improvement of the regional innovation system.

Several studies have shown that the most advanced countries in terms of innovation capacity are those which have adopted suitable tools and policies allowing the integration of three fundamental components: governments, research institutions and enterprises, the so called "*helices*" [Etzkowitz and Leydesdorff, 1997, Etzkowitz, 2002]. According to the Triple-Helix Model, the relationships among the three helices have substantially changed over the time. While in the past they were acting separately according to their own strategies, nowadays they tend towards a continuous interaction. Any internal change that occurs in one of these three entities is rapidly transmitted to the others thus generating an "*endless transition process*" [d'Andrea et al., 2005]. Usually, under particular conditions, the above process is able to produce positive effects in terms of competitiveness of clusters of enterprises. However, triggering an effective innovation strategy in a wide territory (for example, at the regional or country level) requires a suitable **governance model** of the Triple-Helix system. Yet, this may require more and more complex and sophisticated governance models, which are strictly related to the economic, social and cultural environment of each territory.

The remaining of this paper is organized as follows. Section 2, firstly, analyzes possible causes and effects of the loss of some of the helices of the Triple-Helix model and, secondly, describes the role of the Innovation Poles in a Convergence Region, namely Calabria, in facilitating the process of integration of the strategies of the three helices. Section 3 illustrates the governance model of the regional Innovation System of the region of Calabria. Section 4 presents the Innovation Pole operating in the Information and Communication Technologies sector.

2. The Innovation Poles as Innovation Intermediary

The experience indicates that realizing an effective interaction among government, universities, research centres and enterprises is not an easy task in spite of the fact that only a territory that is equipped with a powerful "Triple-Helix engine" can rapidly and effectively move towards an innovation

ecosystem, meaning that the main actors of the R&D work as a *system*. Unfortunately, there are various reasons that may hinder the integration process. As a consequence a territory may lose the contribution of one or more main players (government, universities, enterprises). The remaining of this section illustrates some guidelines to mitigate the risk of failure in the process of integration of the strategies of the R&D actors.

2.1 Double-Helix and Single-Helix engines

When only one actor is missed, we say that the territory is driven by a double-helix engine. For example, if the public authorities are not part of the process, a strong collaboration could still be established between academic bodies and enterprise and this might have a positive impact on the development of a territory. However, only partial results could be obtained because if the process is not sustained by the public administration, it is most likely that research activities may be limited by shortage of public funds and/or by the missing of an appropriate regulatory framework fitting the enterprises needs.

A territory can also be “powered” by a single-helix engine, as in the case of a very innovative company that operates in a territory without cooperation with the other main actors (with regard to the R&D sector). The innovation may propagate towards other companies and trigger development processes of that territory.

Finally, the worst case is when the main actors act separately without any form of collaboration in the R&D sectors. In this situation various scenarios may occur and they need to be analyzed on a case-by-case basis. However, the most common scenario is that in which the whole mechanism seizes up and the territory doesn't take any step towards the creation of its own innovation system. Some of the main reasons that could hinder the sharing of common goals among the main actors are shortly discussed below (we wish to go deeper on this topic in a future work).

Language and competences

The main actors speak very different “*languages*” often based on their heterogeneous competences. Even in those cases where they wish to collaborate, they fail to have a deep understanding of the point of view of the others. And most likely this is made more complicated by the fact that they have different objectives (see below).

Change management

The various actors move at a very different *speed*. Business is changing at a very high speed for a very simple reason: technology is changing, market is changing, new competitors are coming up each day and so on. A company is driven to make rapid changes or risks to die [Toffler and Toffler, 2007]. The institutions, however, change very slowly: introducing a revolutionary technology may be less difficult than changing an institution. One good example is represented by the public education system: the economic and business system is quickly changing, while the education system remains behind and this

is a big problem for any advanced economy. The public administration is the slowest of all. It often takes decades to change laws to suit changes in the economic system.

Specific Objectives

Businesses are primarily interested in profits. They would like that universities concentrate on industrial research whose results can be commercially exploited. They also expect that governments fund development programs and simplify the bureaucracy. On the other side, however, researchers are primarily interested in scientific honours. They expect funding for public research by both the corporate and the public sector. Finally, governments expect to increase their political consensus, but this is difficult to achieve with technology policies because the related results often require a long time to be produced and also because they are less *visible* to the community with respect to other areas, such as welfare and transportation system.

2.2 A Catalyzed Triple-Helix engine to facilitate the interaction among the three helices

An effective and responsive Innovation System requires the *three helices* to run in sync. When this happens, the whole territory obtains great advantages in terms of competitiveness and economic development. Unfortunately, as discussed previously, the main actors of R&D have very different objectives, speak different languages and move at different speeds. Therefore, it is not an easy task to have them working on shared goals.

The prevention of the seizing up of the “triple-helix innovation engine”, can be simplified through the presence in the territory of a “**catalytic entity**” that may facilitate the interaction among government, universities and enterprises. The importance of the role of the so called “*innovation intermediary*”, i.e. organizations that act as an interface between two or more *helices*, has emerged from a number of different sources [Howells, 2006]. Intermediaries play a fundamental role in supporting technological transfer processes, in building collaboration networks and innovation communities, as well as in facilitating and coordinating the flow of information and knowledge among the various R&D actors.

The approach based on the *catalyzed engine* has been adopted by the Calabria Region. In fact, the “Regional Strategy for Research and Innovation 2007-2013” of the Region of Calabria, confers the role of *innovation intermediary* to the **Network of Innovation Poles**. The Poles, which are eight in all, were born in 2011 under the aegis of the Regional Department for Education and Research of the Region of Calabria with the following main objectives:

- stimulate, qualify and meet the local demand for innovative services;
- support the technology transfer process and the circulation of information and knowledge;
- **facilitate and support the cooperation between universities, businesses and local governments.**

The Innovation Poles are **clusters of independent companies and research organizations** operating in those sectors that the Regional Government has identified as strategic for the development of Calabria, namely: ICT, logistics, transportation, agriculture, new materials, cultural heritage, health, water resources. The eight Poles aggregate approximately 320 companies, the three universities of the region, some institutes of the National Research Council and ENEA.

Recently, the eight Innovation Poles have operatively started the building of an Innovation Network based on the concept of "*coopetition*" according to which they cooperate and compete at the same time with the aim of increasing the quality of the innovation services.

As part of the Regional Innovation System, they participate in the local innovation process according to the Governance Model illustrated in detail in the next section.

3. The Governance Model of the Regional Innovation System of the Region of Calabria

The Regional Innovation System of Calabria consists of: (i) a number of Operating Nodes, whose main objective is to provide services for innovation and to facilitate the cooperation of the local R&D actors; (ii) a Coordination Node, which is responsible for Planning, Monitoring and Supervising the innovation process; (iii) an Advisory Body, supporting the decision-making process. In the specific, entities that make up the Innovation System are:

- Coordination Node:
 - ✓ It is the (to be built) Regional Innovation Agency which should be responsible for Planning, Coordination, Supervision, Monitoring and Evaluation of the overall innovation process.

- Operating Nodes:
 - ✓ The eight Innovation Poles operating in the region (ICT, logistics, transportation, agriculture, new materials, cultural heritage, health, water resources);
 - ✓ The industrial Liaison Office of the local Universities;
 - ✓ The Technology Transfer Offices (TTOs) of the regional Chambers of Commerce;
 - ✓ The TTOs of the regional Entrepreneurial Associations.

- Conference for Scientific and Technological Innovation:
 - ✓ It is an advisory body made up of the main regional public and private stakeholders: Regional Government, Universities, Research Centres, Regional Public Bodies, Chambers of Commerce, Entrepreneurial Representations, Social partners and so on. The contributions of the stakeholder must be collected, analyzed and processed by the Coordination Node with the aim of supporting the

Calabria Region in decision-making in the research and innovation area.

The Network of Operating Nodes is organized on a market-basis as a **One-Stop Shop** for innovation services. The One-Stop Shop coordinates and integrates all the skills necessary to meet the demand for innovation, regardless of where those skills are spread among the members of the Network. Thanks to this, users of the One-Stop Shop do not have to worry about identifying to which node of the network they should submit a request for an innovation service.

The Coordination Node plays a key role for the efficiency and effectiveness of the overall Innovation System. In particular, it is responsible for:

- ensuring the meeting of supply and demand through a careful and periodic analysis of the regional demand of innovation and the subsequent orientation of the offer of the Operating Nodes towards the real needs of the regional public and private organizations;
- ensuring the constant raising of the quality of the services offered by the Operating Nodes;
- overseeing the process of decomposition of each user request to the One-Stop Shop in a series of more specific operation to be forwarded to the appropriate Operating Nodes;
- encouraging, in a logic of *coopetition*, integration and specialization of the services offered by the various nodes, avoiding duplication which would result in waste of public resources as well as in disorientation of the users;
- making more attractive and competitive the regional innovation system through appropriate communication of the available services, even beyond the regional borders;
- disseminating the best practices;
- monitoring the operation of the overall system on the basis of suitable performance indicators agreed with the Conference for Scientific and Technological Innovation.

The organization of the Regional Innovation System has been based on the experience gained by the Region of Calabria in the context of the European Pilot Actions RIS and RISI (respectively, Regional Innovation Strategies and Regional Information Society Initiative [Mallamaci, 2000]) with the primary purpose of ensuring the **governance** of the whole system. In this context, a key role is played by the Conference for Scientific and Technological Innovation which institutionalize the former Regional Council for Scientific and Technological Research that have already been successfully established and activated in the context of RIS and RISI Initiatives [European Commission, 2007].

4. The Innovation Pole in the Sector of Information and Communication Technologies

As stated in the previous section, the Regional Government has identified eight Innovation Poles among which the Innovation Pole in the sector of the Information and Communication Technologies (ICT) which plays a relevant role in the innovation process because of the enabling nature of ICT. The legal entity operating the ICT Innovation Pole is ICT-SUD, a Competence Centre founded in December 2006 under the aegis of the Italian Ministry for University and Research, as a non-profit consortium company [Mallamaci and Saccà, 2011].

ICT-SUD consists of 59 members (private companies, various universities, CNR, etc.) and operates as a network organization with local offices in Apulia, Campania, Sardinia, Sicily and the head office in Calabria. ICT-SUD has implemented the ICT Innovation Pole with the aim to broaden its offer of services for innovation and technology transfer in the ICT sector to non-member organizations [Mallamaci and Saccà, 2012].

The ICT Pole is a **cluster of 21 companies and research organizations**, which offer and/or demand services in the ICT sector. The main objective of the Pole is to stimulate innovative activities by promoting intensive interactions, sharing of facilities and exchange of knowledge and expertise and by contributing effectively to technology transfer, networking and information dissemination among the companies that make up the Pole [European Commission, 2006].

The ICT Pole has developed some basic innovation services, such as: Sharing of Research Infrastructure; Networking and Strategic Marketing; Exploitation of research results; etcetera.

The Pole has also developed a number of specialized services in the ICT sector, among which: **Vulnerability Assessment** (Analysis, Mitigation, Correction, Prevention), **Cloud Computing** (Design, Migration, Experimentation of the OpenStack Platform, etc.); **Knowledge Sharing** (Extraction and acquisition of knowledge, classification of contents and documents on the basis of a taxonomy of concepts related to innovation; distribution of contents and documents). The Pole has also realized an **Observatory** on ICT market and trends.

Finally, the ICT Pole has supported many public and private organizations in the preparation of project proposals in the ICT sector, as well as in the construction of project partnerships for the participation to National and Regional calls, especially within the Regional Operational Programme and the National Operational Programme.

Currently, the ICT Innovation Pole represents one of the leading networks of innovation in the ICT sector in Calabria.

5. Conclusion

The adoption of suitable tools and policies for the integration of the strategies of the main protagonists of the Triple-Helix Model (government, universities, enterprises) produces relevant progresses in terms of innovation

capacity of a territory. In this paper we have presented some causes and effects related to the difficulty of interaction of the above R&D actors. Moreover, we have illustrated the role of the *innovation intermediary* (in particular, the Network of Innovation Poles) in the implementation of the Regional Innovation System in Calabria. We have also presented the organization and the governance model of the Innovation System of the region of Calabria. Finally, we have illustrated in more details the organization of the ICT Pole whose growing importance is due to the enabling nature of the Information and Communication Technologies.

Acknowledgements

This work has been developed by ICT-SUD within the framework of the *Progetto Integrato Strategico Regionale* “Rete Regionale dei Poli di Innovazione”, ROP Calabria ERDF 2007-2013, which is co-funded by the Italian State, the Region of Calabria and the European Regional Development Fund (ERDF).

The authors wish also to thank the legal entities running the eight Regional Innovation Poles (Logistica Ricerca e Sviluppo, Biotecnomed, Cultura e Innovazione, Centro di Competenza ICT-SUD, Natura Energia e Territorio, Agrifoodnet, Calpark, Nautilus) as well as Fincalabria, in charge of the complementary Project “Calabrialnova”, for their valuable contribution to the debate on the implementation of the Regional Innovation System of the region of Calabria.

References

[Chesbrough, 2003] Chesbrough H., *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard Business School Press, Boston, 2003

[Chesbrough, 2006] Chesbrough H., *Open Business Models: How to Thrive in the New Innovation Landscape*, Harvard Business School Press, Boston, 2006

[Cohen and Levinthal, 1990] Cohen W. M., Levinthal D. A., *Absorptive capacity: A new perspective on learning and innovation*, *Administrative Science Quarterly*, Volume 35, Issue 1 pg. 128-152, 1990

[Corti, 1986] E. Corti, *La Diffusione e il Trasferimento di Tecnologie: il ruolo dei Parchi Scientifici e Tecnologici per lo sviluppo delle economie locali*, TECHNE Ed., Il Cardo, Venezia, 1996.

[Corti e Mallamaci, 1986a] Corti E., Mallamaci C. L., *The Networking of SMEs by the Information and Communication Technologies: the case of the Science and Technology Park of Calabria*, Proc. Of the Sixth International Conference on Economics of Innovation, Networks of Firms and Information Networks, Cremona, Piacenza, June 5-7, 1996.

[Corti e Mallamaci, 1986b] Corti E., Mallamaci C. L., *Tacit Knowledge Flowing Through Advanced Telecommunication Technologies: the case of the Science and Technology Park of Calabria*, Proc. Of the 6th International Forum on Technology Management, Amsterdam, October 15-18, 1996.

[d'Andrea et al., 2005] d'Andrea L., Quaranta G., Quinti G., *Manuale sui processi di socializzazione della ricerca scientifica e tecnologica*, CERFE, Roma, 2005.

[Etzkowitz and Leydesdorff, 1997] Etzkowitz H., Leydesdorff L., *Universities in the Global Economy: A Triple Helix of University-Industry-Government Relations*, Cassel Academic, London, 1997.

[Etzkowitz, 2002] Etzkowitz H., *The Triple Helix of University - Industry - Government. Implications for Policy and Evaluation*, Working Paper 2002/11, ISSN 1650-3821, SISTER, Swedish Institute for Studies in Education and Research, Stockholm, 2002.

[European Commission, 2007] European Commission, *Innovative strategies and actions - Results from 15 Years of Regional Experimentation*, European Union, Regional Policy, November 2007

[European Commission, 2006] European Commission, *Community Framework for State Aid for Research and Development and Innovation (2006/C 323/01)*, Official Journal of the European Union, C 323/10 (page 10), December 30, 2006

[Howells, 2006] Howells J., *Intermediation and the role of intermediaries in innovation*, Elsevier, *Research Policy* 35 (2006) 715–728, Amsterdam, 2006

[Mallamaci, 2000] Mallamaci C. L., *Una metodologia per la Pianificazione Strategica per lo Sviluppo della Società dell'Informazione in una Regione Obiettivo 1*, Atti del Convegno "Fondi Strutturali e Sviluppo Regionale", Pricewaterhouse Coopers, Aula Convegni CNR, Roma, Luglio 2000.

[Mallamaci and Saccà, 2011] Mallamaci C.L., Saccà D, *Strategy and Experience in Technology Transfer of the ICT-SUD Competence Center*, in Proc. of the VII Conference of the Italian Chapter of AIS, itAIS2010, Villa Doria d'Angri, Naples (Italy), October 8-9, 2010; in (book): D'Atri, A. et al. (Eds), "Information Technology and Innovation Trends in Organizations", page 291, ISBN: 978-3-7908-2631-9, Springer Physica Verlag, Heidelberg, 2011.

[Mallamaci and Saccà, 2012] Mallamaci C.L., Saccà D, *Towards a Digital Ecosystem to Increase Effectiveness of Technology Transfer Services*, in Proc. of the VIII Conference of the Italian Chapter of AIS, itAIS2011, Rome (Italy), October 7-8, 2011; in (book) "Information Systems: Crossroads for Organization, Management, Accounting and Engineering", M. De Marco, D. Te'eni, V. Albano, S. Za (Editors), page 275-283, ISBN: 978-3-7908-2788-0; Springer Physica Verlag, Heidelberg, 2012.

[Toffler and Toffler, 2007] Toffler A., Toffler H., *Revolutionary Wealth: How It Will Be Created and How It Will Change Our Lives*, Broadway Business, June 2007.

Cloud Agency: una Guida attraverso il Cloud

Rocco Aversa¹, Luca Tasquier¹, Salvatore Venticinque¹

¹Dipartimento di Ingegneria Industriale e dell'Informazione, Seconda Università degli Studi di Napoli

via Roma 29, 81031 Aversa - Italia

rocco.aversa@unina2.it, luca.tasquier@unina2.it, salvatore.venticinque@unina2.it

Abstract. *In the Cloud context, the interoperability among different providers becomes critical due to the vendor lock-in problem. The different technologies and the lack of standards make the selection of the best Cloud offers suitable for the user's application an hard work. Here we present a multi agent system that accesses, on behalf of the user, the utility market of Cloud computing to maintain the best resources configuration that satisfies the application requirements. It also offers management and monitoring facilities for the Cloud infrastructure in order to guide the user in all the phases of the application's lifecycle. Together with the platform, we present client-side tools that can be used to orchestrate agents' based services.*

Keywords: IaaS Cloud interoperability, Mobile Agents, Cloud monitoring.

1. Introduzione

Il Cloud Computing rappresenta un'opportunità per gli utenti riducendo i costi di gestione dell'infrastruttura e incrementando l'efficienza dei prodotti offerti fornendo un modo alternativo per utilizzare i servizi IT. L'elasticità del Cloud consente la riconfigurazione delle risorse computazionali quando i requisiti dell'applicazione cambiano dinamicamente e, allo stesso tempo, il modello di business basato sul *pay-per-use* permette di cambiare provider quando viene trovata un'offerta più conveniente. Ovviamente, data la scarsa interoperabilità tra soluzioni Cloud differenti e la scarsa portabilità delle applicazioni, sorgono varie problematiche nell'attuale mercato del Cloud Computing. Inoltre, anche se lo sviluppatore dei servizi è in grado di superare tali difficoltà, non è facile individuare le differenti proposte Cloud disponibili, verificare se esse possono soddisfare i requisiti del servizio ed effettuare un confronto tra le varie offerte. Attualmente non c'è nessuna ontologia comune che descriva i termini e i livelli di servizio, né in modo formale né in linguaggio naturale. Un altro problema riguardante è la gestione delle risorse acquistate. Anche in questo caso la mancanza di uno standard largamente condiviso per il Cloud Infrastructure Level (IaaS) mina la possibilità di poter optare per differenti

Congresso Nazionale AICA 2013

soluzioni commerciali e tecnologiche. Oltre alle appena citate problematiche, risulta essenziale monitorare l'utilizzo delle risorse Cloud acquisite. Dal punto di vista dei provider, il monitoring è necessario per controllare l'effettivo utilizzo delle risorse computazionali; per controllare i livelli di servizio garantiti da contratto agli utenti; per sviluppare strategie di gestione e negoziazione atte a massimizzare gli introiti. In questa accezione il monitoring è rivolto sia alle risorse hardware che a quelle virtuali. Dall'altro punto di vista, gli utenti hanno bisogno di controllare le loro risorse per verificare che i livelli di servizio garantiti siano continuamente rispettati e per identificare possibili sottoutilizzi o sovraccarichi per eventualmente rinegoziare nuovi contratti. Inoltre, nel contesto Cloud, il monitoring dei livelli di servizio diviene critico a causa del contenzioso che può verificarsi tra provider e utenti in caso di interruzione dei servizi.

In questo articolo verrà presentato un insieme di tool che consentono all'utente l'orchestrazione di servizi ad agenti che supportano l'individuazione, la negoziazione, la gestione ed il controllo delle risorse Cloud. Nella Sezione 2 verrà presentato lo stato dell'arte relativo alle problematiche affrontate. Nella Sezione 3 sarà descritto il ciclo di vita delle applicazioni Cloud a livello infrastrutturale. La piattaforma ad agenti progettata e sviluppata per supportare tale ciclo di vita verrà presentata nella Sezione 4, definendo un client per l'accesso ai servizi illustrati nella Sezione 5. Nella Sezione 6 verranno riassunte le conclusioni e i possibili sviluppi futuri.

2. Stato dell'arte

La progettazione e lo sviluppo di soluzioni per la governance di Cloud multipli ed eterogenei è una problematica affrontata sia in ambito di ricerca che in domini commerciali [Aversa et al, 2009]. Le attuali tecnologie Cloud offrono un supporto limitato alla negoziazione dinamica dello SLA tra i vari attori. Il lavoro presentato in [Kertesz et al, 2009] presenta un'architettura per la virtualizzazione delle risorse basata su SLA che fornisce una soluzione estesa per l'esecuzione di applicazioni utente su Cloud. Questo lavoro rappresenta il primo tentativo di combinare negoziazione basata su SLA con risorse virtualizzate in termini di fornitura di servizi on-demand. La descrizione dell'architettura si focalizza su tre argomenti: negoziazione dei contratti, brokering dei servizi e pubblicazione degli stessi utilizzando la virtualizzazione. Essa coinvolge broker multipli. Un'architettura multi-agente per la gestione di Cloud è proposta in [Cao et al, 2009]. Una più semplice architettura basata su agenti è proposta in [You et al, 2011]. SLA@SOI è il principale progetto che mira ad offrire una piattaforma open-source per la gestione degli SLA.

Relativamente alla gestione di Cloud eterogenei, la principale problematica riguarda l'interoperabilità. Dalla comunità scientifica sono arrivati molti di importanti contributi alla standardizzazione. Alcuni esempi sono OCCl (Open Cloud Computing Interface) della Open Grid Forum, e SOCCl (Service-Oriented Cloud-Computing Infrastructure) dalla ISO Study Group on Cloud Computing (SGCC). In particolare OCCl [Metsch et al, 2011] è una proposta di standard per IaaS Cloud. Essa definisce entità, relazioni e protocolli per tutti i tipi di

funzioni relative alla gestione delle risorse. Questa soluzione mira al soddisfacimento di tre requisiti: integrazione, portabilità e interoperabilità di funzioni comuni, tra le quali la distribuzione delle risorse, la scalabilità delle stesse ed il monitoring dell'infrastruttura. OpenNebula implementa già una interfaccia RESTFull che rispetta lo standard OCCI, e altre tecnologie, come Eucalyptus ed OpenStack si stanno adeguando al suddetto standard.

Il diffondersi dell'uso del Cloud Computing per gestire le applicazioni, i servizi e le risorse rende ancora più urgente la necessità di monitorare i parametri di QoS in modo da avere le performance desiderate [Clayman et al, 2010]. Inoltre il controllo delle risorse è molto importante quando un cliente vuole verificare la conformità dei servizi offerti dal provider con quelli che sono stati firmati nello SLA. Il monitoring delle risorse a livello di infrastruttura [Clayman et al, 2010(2)] ha come obiettivo principale la misurazione dei parametri di sistema relativi a servizi offerti all'utente (ad esempio CPU, RAM, o parametri di storage). Le tradizionali tecnologie di monitoring per macchine singole o cluster sono, in qualche modo, condizionate dalla località e dall'omogeneità degli oggetti monitorati e quindi non possono essere applicate al Cloud in maniera appropriata [Emeakaroha et al, 2010]. Una soluzione per il monitoring atta al rilevamento delle intrusioni in infrastrutture Cloud è proposta in [Ficco et al, 2012].

Molti tool disponibili hanno bisogno di essere eseguiti sulle macchine fisiche e usano i servizi dell'hypervisor. In un Cloud privato possono essere utilizzati per il management dell'infrastruttura, ma in un Cloud pubblico le risorse sono controllate dal Cloud vendor. Esistono molte soluzioni che offrono strumenti per il monitoring di Cloud, come Cloudkick, Nimsoft Monitor, Monitis, Opnet, RevealCloud. Tutti sono soluzioni proprietarie e non mirano a definire uno standard per l'interoperabilità del controllo delle performance. Alcune tecnologie per il controllo delle reti e degli host come sFlow sono state estese per trasporto favorire la trasmissione di informazioni di monitoring di risorse virtualizzate. Per esempio Host-SFlow [Host sflow] esporta metriche di performance relativa a server fisici e virtuali attraverso il protocollo sFlow [sflow]. Ganglia [Ganglia] e altri collector sono già compatibili con tale standard.

Siccome l'utilizzo delle risorse e il carico di lavoro del servizio può cambiare molto frequentemente, un report continuo e aggiornato risulta essenziale all'utente per monitorare in maniera efficace l'ambiente di elaborazione. Inoltre, dal momento che la flessibilità del Cloud consente di cambiare dinamicamente la configurazione del sistema, l'infrastruttura di monitoring distribuita deve adattarsi velocemente ai nuovi requisiti. Al fine di affrontare tutte queste problematiche viene proposto un servizio di monitoring basato su agenti. In [Illari et al, 2008] gli autori sostengono che un approccio basato su agenti software è il modo naturale per affrontare il controllo di sistemi distribuiti con le caratteristiche citate sopra. Gli agenti si muovono e si distribuiscono per compiere le azioni di monitoring ad essi assegnate. In [Liu e Chen, 2011] viene studiato un controllo ottimale di agenti mobili per il monitoring di reti artificial-immune-system-based (AIS-based).

3. Ciclo di vita delle applicazioni in un IaaS Cloud

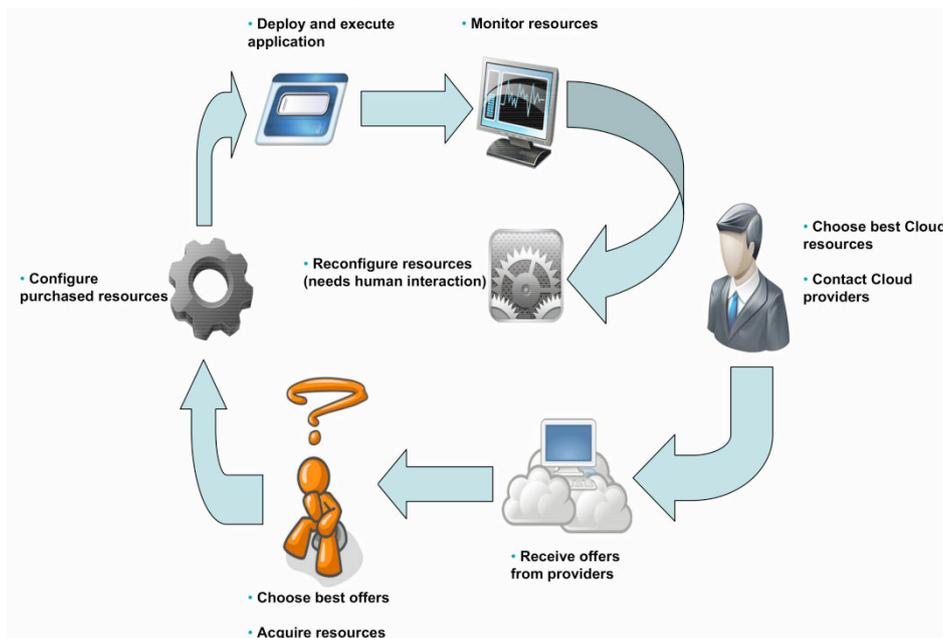


Figura 1 - Ciclo di vita di un'applicazione su IaaS Cloud

Come illustrato in Figura 1 il ciclo di vita di un'applicazione in esecuzione su un'infrastruttura Cloud si divide in tre fasi:

1. **Cloud Provisioning.** L'utente deve scegliere le migliori risorse Cloud per la propria applicazione (Virtual Machine, Storage, ecc.). In seguito ha la necessità di selezionare il miglior provider IaaS basando la propria scelta su molti parametri (costo-per-uso, quantità di memoria per le macchine virtuali, capacità dei dischi, banda, ecc.). Molto spesso questa scelta è difficile a causa della disomogeneità con la quale i Cloud Provider presentano le caratteristiche e i parametri delle risorse offerte. I vendor non hanno un'interfaccia comune e standardizzata per descrivere i parametri delle risorse, rendendo così arduo il confronto tra risorse anche dello stesso tipo.
2. **Cloud Configuration.** Dopo aver selezionato le migliori risorse per l'applicazione, il cliente deve firmare lo SLA con il Cloud provider. Fatto questo, devono essere espletate alcune attività di gestione dell'infrastruttura prima di poter mettere in esecuzione l'applicazione. Per esempio, le immagini del sistema operativo da utilizzare devono essere linkate e partire vanno attivate le istanze delle macchine virtuali acquistate. Per questo motivo, l'utente deve conoscere le operazioni consentite per la particolare risorsa e la specifica interfaccia ai servizi fornita dal corrispondente Cloud Provider. A questo punto l'applicazione può essere distribuita sull'infrastruttura e posta in esecuzione.

3. *Cloud Monitoring*. In questa fase l'utente configura una rete di sonde al fine di collezionare misure riguardanti parametri di performance delle risorse utilizzate. Per tenere sotto controllo continuo le performance dell'infrastruttura e la storia del comportamento delle risorse, risulta essenziale elaborare periodicamente un insieme di indici di performance e configurare trigger il cui compito è notificare condizioni critiche verificatesi sul Cloud. Ad es. è molto utile conoscere se il carico di lavoro dell'infrastruttura è differente da quello previsto, in modo da prevenire saturazione o sottoutilizzo delle risorse. Queste informazioni sono, inoltre, essenziali per pianificare una efficace riconfigurazione dell'infrastruttura al fine di adattare al meglio i requisiti correnti dell'applicazione e di ottimizzare performance e costi.

4. Cloud Agency

Cloud Agency (CA) [Ventiquattro et al, 2012] è un sistema multi-agente (MAS) che consente il provisioning, management e monitoring delle risorse Cloud che meglio soddisfano i requisiti dell'applicazione. Questo sistema è stato progettato e sviluppato nell'ambito delle attività di ricerca del FP7 mOSAIC project [Di Martino et al, 2010]. Esso integra le funzionalità comuni di gestione offerte da infrastrutture IaaS private e pubbliche implementando uno strato trasparente verso quattro servizi principali che consentono la gestione delle risorse Cloud: provisioning, management, monitoring e reconfiguration. In particolare, il *Resources Management Service* è utile per configurare e far partire le risorse prima della messa in esecuzione dell'applicazione e può essere anche usato per attivare, arrestare, migrare e riconfigurare dinamicamente le risorse durante il loro utilizzo.

Per consentire, in modo semplice, la specializzazione del comportamento degli agenti sono state definite delle interfacce astratte per la gestione delle risorse, il monitoring e gli altri servizi che la Cloud Agency offre. Queste interfacce possono essere anche utilizzate dagli sviluppatori che vogliono estendere la Cloud Agency con nuove tecnologie di monitoring, supportando nuovi Cloud Provider. Queste implementazioni diventano plugins della CA, che vengono automaticamente integrate ed usate dagli agenti attraverso la loro interfaccia generica.

Il *Mediator* e i *Broker* sono agenti che si occupano della negoziazione. Essi interagiscono con i client della Cloud Agency e con i *Vendor Agent*. Quest'ultimo rappresenta uno specifico provider: è usato per ottenere un'offerta nella fase di approvvigionamento delle risorse, per accettare o rifiutare l'offerta proposta, per avere informazioni riguardanti una risorsa o per eseguire un'azione sulla stessa; per quanto concerne il servizio di monitoring, esso si occupa di distribuire un'istanza della Cloud Agency su di una risorsa al fine di abilitare la migrazione degli agenti di monitoring su di essa e la comunicazione con le altre istanze della CA. Il servizio di monitoring è implementato dai *Meter* e da un *Archiver*, i quali sono responsabili delle misurazioni di performance di basso livello e della raccolta e gestione delle informazioni di monitoring. L'architettura generale della Cloud Agency è mostrata in Figura 2.

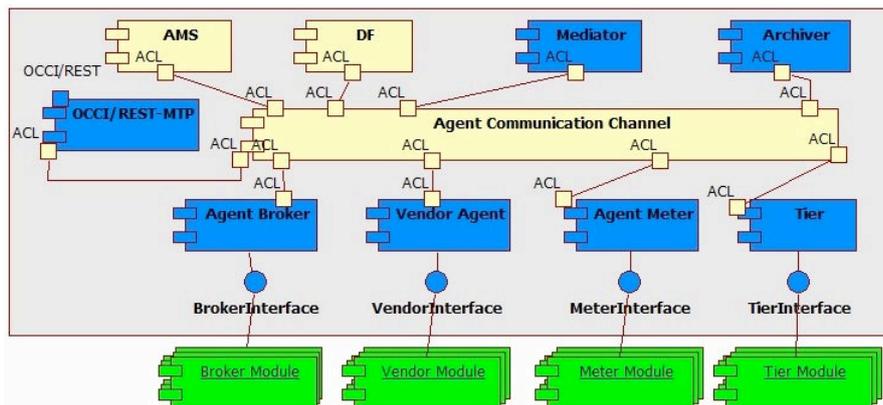


Figura 2 - Architettura della Cloud Agency

L'esecuzione degli agenti può essere orchestrata invocando i servizi offerti dalla Cloud Agency attraverso una interfaccia RESTFull conforme allo standard OCCl. L'interfaccia OCCl implementa due tipi di richieste: le *Asynchronous Service Requests* (ASR) sono usate per richiedere alla CA l'esecuzione di un servizio, ad esempio per far partire una negoziazione, per accettare o rifiutare uno SLA, ecc. Queste richieste sono non-bloccanti, una volta effettuate, il client può continuare la propria esecuzione e richiedere altri servizi. Le ASR genereranno eventi futuri, che verranno gestite dal richiedente. Le *Synchronous Service Requests* (SSR) sono utilizzate per richiedere informazioni. Per esempio un utente può richiedere la lettura di uno SLA, lo stato di una negoziazione, la lista dei vendor disponibili e così via. Queste interrogazioni sono sincrone e ritornano immediatamente il risultato se disponibile oppure un'eccezione in caso contrario. L'OCCI-MTP consente la comunicazione tra il client e la Cloud Agency; esso gestisce sia le richieste asincrone che quelle sincrone. Le *Cloud Agency Notifications* (CAN) vengono inviate asincronamente al client come risposte ad eventi generati da un ASR.

Il *Resources Monitoring Service* è usato per avere una conoscenza aggiornata delle performance relative all'infrastruttura Cloud. L'utente ha la possibilità di scegliere le metriche da monitorare su ogni risorsa. L'insieme delle metriche scelte per ogni risorsa rappresenta un agente Meter che è delegato ad effettuare le misure richieste. L'ASR contenente la richiesta di inizio misurazioni viene inoltrata all'Archiver che chiede di creare un'istanza di Cloud Agency sulla risorsa selezionata al relativo Vendor Agent, al fine di poter migrare l'agente Meter sull'infrastruttura ed iniziare le operazioni di raccolta delle misure.

5. Un tool per la gestione dell'infrastruttura Cloud: CA-Client

Per rendere semplice la fruizione dei servizi offerti della Cloud Agency, sono stati progettati ed implementati un insieme di tool che supporti l'orchestrazione dei servizi della CA da parte di diversi tipi di utente. Il cuore di tali strumenti sono le *CA-Client API*: esse mettono a disposizione un insieme di

API che supportano lo sviluppo di applicazioni client event-based per Cloud Agency. In accordo con il modello di interazione basato su eventi della CA, le API permettono di inviare richieste asincrone e sincrone e di ricevere notifiche asincrone. Il diagramma delle classi di una tipica applicazione client è mostrato in Figura 3.

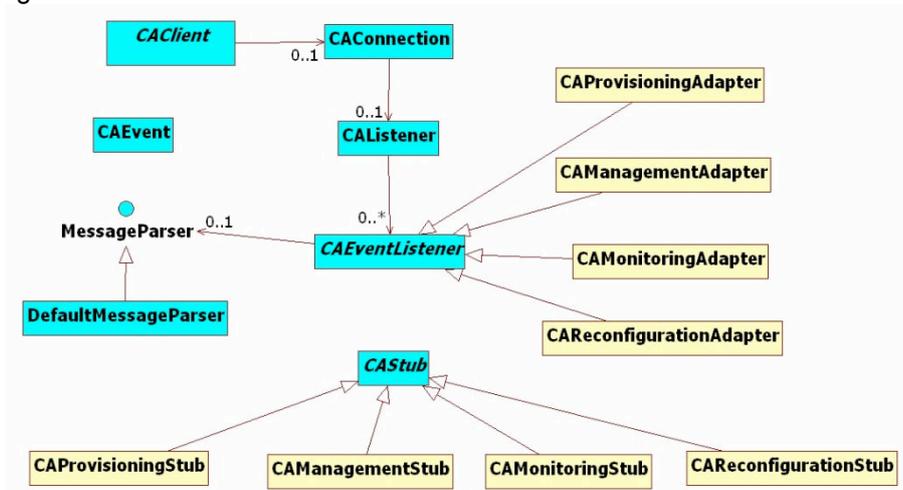


Figura 3 – Diagramma delle classi delle CA-Client API

Le classi colorate di blu rappresentano le *core* API. Le classi gialle rappresentano alcune implementazioni delle classi astratte del core. In particolare, ogni *stub* offre un insieme di metodi per invocare i servizi della Cloud Agency. Questi metodi inglobano le richieste REST con le quali è possibile richiedere alla CA di effettuare le operazioni messe a disposizione. In aggiunta agli stub vengono forniti degli adapter; un *adapter* consente di ricevere i messaggi asincroni provenienti dalla Cloud Agency. Ogni utente può personalizzare il comportamento del client all'arrivo di particolari eventi semplicemente estendendo la classe *CAEventListener* o riscrivendo un adapter esistente. Così è possibile implementare reazioni automatiche a particolari eventi semplicemente aggiungendo un nuovo adapter al core. Sulla base di tali API è stato progettato e sviluppato un tool grafico che mette a disposizione dell'utente tutti i servizi della Cloud Agency: il *CA-GUI*. Tale interfaccia grafica (vedi Figura 4) consente di inizializzare le transazioni di provisioning, di tenere sotto controllo le transazioni aperte, di ricevere, accettare, rifiutare e visualizzare le proposte ricevute.

Anche dal punto di vista della gestione delle risorse acquisite viene offerta un'interfaccia uniforme verso tutte le risorse fornite da qualsiasi Cloud Provider: è dunque possibile inizializzare e arrestare l'esecuzione di macchine virtuali, collegare dischi e richiedere informazioni sullo stato della risorsa.

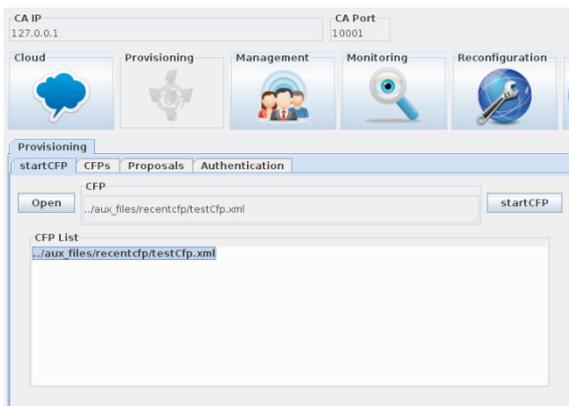


Figura 4 – CA-GUI

La console di monitoring consente la configurazione dell'infrastruttura di monitoring sulle risorse acquistate. Una volta configurati gli agenti Meter, è possibile tenere sotto controllo le performance del Cloud utilizzando il *Monitoring Tool*, (vedi Figura 5) anch'esso fruibile da interfaccia grafica. Esso permette la visualizzazione della lista delle metriche disponibili e l'elaborazione delle stesse. L'utente può scegliere di visualizzare il live-stream degli indici di performance selezionati per una determinata risorsa o di effettuare analisi sui parametri collezionati utilizzando diversi operatori messi a disposizione come media, massimo/minimo valore etc.

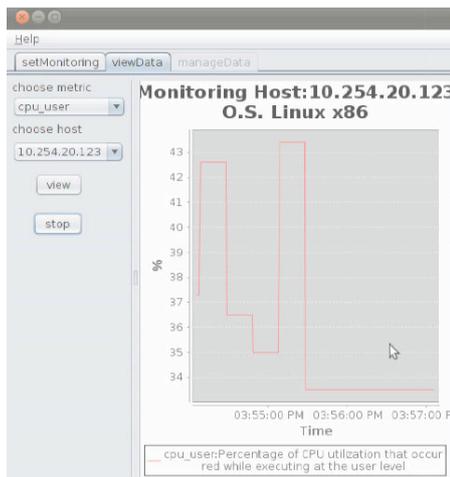


Figura 5 - Monitoring Tool

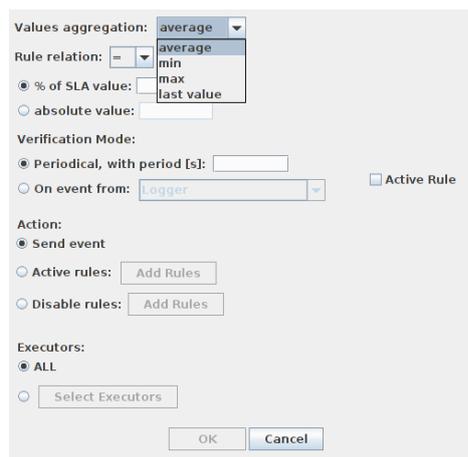


Figura 6 - Observer Tool

L'utente può inoltre creare dei trigger in modo da ricevere una notifica per un determinato evento (o condizione critica) in maniera asincrona. Attraverso l'*Observer Tool* (Figura 6) è possibile scegliere il parametro della risorsa relativamente al quale si vuole creare il trigger. Scelto il parametro, l'interfaccia consente di definire la regola: prima di tutto si può definire il tipo di

aggregazione relativa ai risultati di monitoring scegliendo tra diverse operazioni,(valor medio, ultimo valore ricevuto etc). E' inoltre possibile determinare il criterio attraverso il quale la regola risulti verificata agendo su due parametri: il valore di confronto e la relazione d'ordine. Per quanto riguarda il primo parametro, è possibile scegliere una percentuale del valore contenuto nello SLA relativo al parametro considerato o un valore arbitrario inserito dall'utente. La relazione d'ordine può essere selezionata tra vari valori come uguaglianza, diversità, minore ecc. Un altro parametro che è possibile settare è il periodo di verifica della regola; ogni regola può avere due possibili stati alla creazione del trigger: nello stato *active* la regola diventa attiva appena viene creata, mentre nello stato *idle* la regola viene sospesa e può essere attivata in seguito. Quando un trigger viene attivato dal verificarsi di una regola su un parametro, l'utente può decidere se essere allertato dal verificarsi dell'evento o di attivare/disattivare altre regole precedentemente definite e riferite a parametri di altre risorse dell'infrastruttura. E' possibile così definire trigger complessi attraverso la composizione di regole più semplici.

6. Conclusioni

All'interno delle attività del progetto di ricerca europeo FP7 mOSAIC è stata progettata e implementata una piattaforma multi-agente per provisioning, management, monitoring e reconfiguration di risorse Cloud. In questo lavoro è stata presentata l'infrastruttura ad agenti, dettagliando i ruoli delle singole entità e le metodologie di accesso ai servizi offerti. E' stata inoltre presentata un'API e un tool grafico che permette l'orchestrazione dei servizi ad agenti secondo il ciclo di vita adottato per la governance delle risorse Cloud. E' stato descritto come accedere alle funzionalità di provisioning e management e come configurare ed utilizzare l'infrastruttura di monitoring. In particolare, la funzionalità di triggering di alert al verificarsi di specifiche condizioni potrebbe trovare utilizzo nell'implementazione di policy di riconfigurazione basate sul verificarsi di condizioni critiche su uno o più parametri dell'infrastruttura Cloud. in futuro sarà possibile utilizzare tale funzionalità per progettare ed implementare agenti autonomici atti alla riconfigurazione delle risorse al fine di bilanciare l'infrastruttura e risolvere le condizioni critiche rilevate.

Ringraziamenti

Questo lavoro è stato supportato dal progetto mOSAIC (EU FP7-ICT programme, project under grant #256910) e da PRIST 2009, Fruizione assistita e context aware di siti archeologici complessi mediante terminali mobili.

Bibliografia

[Aversa et al, 2009] Aversa, R., Di Martino, B., Venticinque, S.: Integration of mobile agents technology and globus for assisted design and automated development of grid services. Pp. 118-125 (2009).

[Kertesz et al, 2009] Kertesz, A., Kecskemeti, G., Brandic, I.: An sla-based resource virtualization approach for on-demand service provision. In: Proceedings of the 3rd International workshop on Virtualization technologies in distributed computing, VTDC '09, pp. 27-34. ACM, New York, NY, USA (2009).

[Cao et al, 2009] Cao, B.Q., Li, B., Xia, Q.M.: A service-oriented qos-assured and multi-agent cloud computing architecture. In: Proceedings of the 1st International Conference on Cloud Computing, CloudCom '09, pp. 644-649. Springer-Verlag, Berlin, Heidelberg (2009).

[You et al, 2011] You, X., Wan, J., Xu, X., Jiang, C., Zhang, W., Zhang, J.: Aras-m: Automatic resource allocation strategy based on market mechanism in cloud computing. *Journal of Computers* 6(7) (2011).

[Metsch et al, 2011] Metsch, T., Edmonds, A., *et al.*: Open cloud computing interface-core and models. In: The Open Grid Forum Document Series, Standards Track, no. GFD-R (2011).

[Clayman et al, 2010] Clayman, S., *et al.*: Monitoring service clouds in the future internet. In: Towards the Future Internet, pp. 115-126 (2010).

[Clayman et al, 2010(2)] Clayman, S., Galis, A., Chapman, C., Toffetti, G., Roderomero, L., Vaquero, L.M., Nagin, K., Rochwerger, B.: Monitoring service clouds in the future internet. In: Future Internet Assembly, pp. 115-126 (2010).

[Emeakaroha et al, 2010] Emeakaroha, V.C., *et al.*: Low level metrics to high level sla- based framework: Bridging the gap between monitored metrics and sla parameters in cloud environments. In: HPCS, pp. 48-54 (2010).

[Ficco et al, 2012] Ficco, M., Venticinque, S., Di Martino, B.: mOSAIC-based intrusion detection framework for cloud computing. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 7566 LNCS n. 2, pp. 628-644 (2012).

[Host sflow] Host sflow. URL <http://host-sflow.sourceforge.net/>.

[sflow] sflow. URL <http://www.sflow.org/>.

[Ganglia] Ganglia. URL <http://ganglia.sourceforge.net/>.

[Illari et al, 2008] Illari, S., Mena, E., Illaramendi, A.: Using cooperative mobile agents to monitor distributed and dynamic environments. *Inf. Sci.* 178(9), pp. 2105-2127 (2008).

[Liu e Chen, 2011] Liu, W., Chen, B.: Optimal control of mobile monitoring agents in immune-inspired wireless monitoring networks. In: *Journal of Network and Computer Applications* 34(6), pp. 1818-1826 (2011).

[Venticinque et al, 2012] Venticinque, S., Tasquier, L., Di Martino, B.: Agents based cloud computing interface for resource provisioning and management. In: Proceedings of the 6th International conference on Complex, Intelligent and Software Intensive Systems, pp. 249-256 (2012).

[Di Martino et al, 2010] Di Martino, B., *et al.*: Building a mosaic of clouds. In: Proceedings of the 2010 conference on Parallel processing, Springer-Verlag (ed.), pp. 571-578 (2010).

Building a Cloud Enabled Data Center

Raffaele Stifani
Executive Architect, IBM Italia
Circonvallazione Idroscalo, 20090 Segrate (MI)
raffaele_stifani at it.ibm.com

Today organizations are looking to get more out of their IT infrastructure as they face new technological and economic pressures. Cloud Computing offers attractive opportunities to reduce costs, accelerate development, and increase the flexibility of the IT infrastructure, applications and services. Infrastructure as a Service is the typical starting point for organizations when moving to a cloud computing environment. Infrastructure as a Service can be used for the delivery of resources such as compute, storage and network services through a self-service portal. The “Cloud Enabled Data Center adoption pattern” of the IBM Cloud Computing Reference Architecture contains prescriptive guidance on how to architect, design and implement an Infrastructure as a Service solution. The Cloud Enabled Data Center adoption pattern contains four architectural macro-patterns and each architectural macro-pattern addresses a specific set of business needs for an Infrastructure as a Service solution.

1. Introduction

Today organizations are looking for ways to get more out of their IT infrastructure as they face new technological and economic pressures. They are also trying to satisfy a broad set of users (internal and external to the enterprise) who demand improvements in the quality of service (QoS), regardless of increases in the number of users and applications. Cloud Computing offers attractive opportunities to reduce costs, accelerate development and increase the flexibility of the IT infrastructure, applications and services.

Infrastructure as a Service (IaaS) is the typical starting point for organizations when moving to a cloud computing environment. IaaS can be used for the delivery of resources such as compute, storage and network services through a self-service portal. With IaaS, IT services are delivered as a subscription service, eliminating up-front costs and driving down ongoing support costs.

IBM has defined a Cloud Computing Reference Architecture (CCRA) based on years of experience of working with customers who have implemented Cloud Computing solutions. The IBM CCRA is a blueprint for architecting Cloud Computing implementations. It is driven by functional and non-functional requirements that are collected from many Cloud Computing implementations.

Congresso Nazionale AICA 2013

IBM CCRA provides guidelines and technical artifacts, such as services and deployment models, and has defined the overarching implementations as *adoption patterns*. An adoption pattern embodies the architectural patterns that represent the ways organizations are implementing Cloud Computing solutions. An adoption pattern can help in the definition of a Cloud Computing solution.

The “*Cloud Enabled Data Center adoption pattern*” of the IBM Cloud Computing Reference Architecture contains prescriptive guidance on how to architect, design and implement an IaaS solution, defines the core requirements and provides guidance on adding new capabilities as they are needed.

The Cloud Enabled Data Center adoption pattern contains four architectural macro-patterns. Each architectural macro-pattern addresses a specific set of business needs for an IaaS solution. This modular architecture allows the extension of an IaaS solution by adding new capabilities and components as needed.

2. Business drivers for a Cloud Enabled Data Center

The following key business drivers influence an organization to implement a Cloud Enabled Data Center:

- *Manage costs* - Cost is a key consideration in business decisions and organizations are challenged to find new ways to do more with less. This situation is important for IT organizations that need to ensure that their IT investments are real business enablers and not cost drivers.
- *Respond to changing business needs* - Businesses are faced with increasing pressure to launch new products and services ahead of the competition and to quickly respond to change. IT infrastructure underpins the business and must be agile to respond to changing business events. For example, as new products or services are launched, the supporting IT infrastructure must be able to scale-out quickly to meet increased transaction volumes.
- *Increased dependence on technology* - Businesses today rely more on technology to deliver services of value. The adoption of technology is increasing at a faster pace than ever before, and in some cases, it is giving rise to new businesses. For example, with an expanding number of sensors and devices, there is a growing need for reliable ways to collect and access and efficiently process and analyze information. The capabilities to satisfy these needs might be delivered by new technologies and services.
- *Faster time to deployment* - Compressing the amount of time required in realizing a business idea into a service or product is often a key strategic business objective. Organizations need to deliver services faster than ever before to stay ahead of the competition. Accelerating time-to-market is crucial to realize revenues faster, control market share, enhance brand image and retain customers. IT plays an important role by improving availability of IT systems that support business needs.

3. Benefits of a Cloud Enabled Data Center

A Cloud Enabled Data Center has the following main benefits:

- *Allow organizations to shift focus quickly* - With the quick availability of IT resources (such as servers and storage), the Cloud Enabled Data Center model allows organizations to focus on the core business and to provide value to their clients.
- *Lower cost of deployment* - By using high levels of automation, IT resources are made available faster, repeatedly and accurately. For example, IT does not have to incur the cost of configuring a server every time it needs one.
- *Provide faster deployment and retirement of systems* - With standardized services offered through the service catalog, IT can provision systems faster. In some cases, such as development and test workloads, users might be allowed to provision the servers and manage them through a web-based portal. The ability to automate the de-provisioning of the resources (after a period of use) is important. This approach ensures that resources that are no longer used are returned in the pool of available resources and can be reused for other purposes.
- *Establish a utility service* - A Cloud Enabled Data Center provides a *pay-as-you-use* model, which can be used to measure and charge individual lines of business for their IT usage.
- *Create predictable deployment* - A Cloud Enabled Data Center promotes standardization of components, which are made available through the service catalog and the description of the service itself. This approach ensures a predictable outcome every time a service is requested.
- *Support dynamic scaling* - A scale-up and scale-down infrastructure is required to support a business service. When a business experiences periods of high demand, a new IT infrastructure can be provisioned to meet temporary increases in workloads. The provisioned IT infrastructure can be de-provisioned when it is no longer required. This approach adds enormous flexibility to the IT environment and ensures optimal utilization of resources.
- *Greater level of control and visibility* - Every task performed by the Cloud Enabled Data Center solution is monitored and reported so that organizations have greater visibility to the use of IT resources.

4. Designing a Cloud Enabled Data Center

The Cloud Enabled Data Center adoption pattern provides a proven approach to define and implement IaaS solutions. IaaS is usually the entry point for organizations that are approaching Cloud Computing. Because IaaS is the starting point, IaaS must be supported by a modular and flexible architecture that easily allows the integration of additional and more robust capabilities. The Cloud Enabled Data Center adoption pattern provides the modular architecture to accomplish this goal by establishing the architectural framework for designing IaaS solutions. The Cloud Enabled Data Center adoption pattern includes several key concepts and typical user roles, use cases and requirements for a private IaaS.

4.1 Key concepts

The following key concepts are related to the Cloud Enabled Data Center adoption pattern:

- *Use-case packages* - A use-case package is a collection of related use cases that define interactions of an actor who performs a specific activity and uses specific Cloud Enabled Data Center functions or capabilities. For example, the set of use cases for interaction with virtualized storage or use cases to support identity and access management.
- *Micro-pattern* - A micro-pattern is a collection of related use-case packages. For example, the micro-pattern for virtualization is a collection of use-case packages for server virtualization, storage virtualization, network virtualization, and hypervisor management.
- *Macro-pattern* - A macro-pattern is a collection of micro-patterns that, when implemented together, permits the organization to improve the maturity level.

4.2 User roles

The Cloud Enabled Data Center establishes new user roles that interact with cloud services. A cloud service can enable the user role to create, offer, use, maintain, and run the Cloud Computing services. A user role is delineated by a distinctive set of typical tasks that can be performed by a single person.

Cloud user roles support the design and development of services, and they support the servicing process for the cloud service. Each role can be accomplished by one or more persons. For example, an organization can have a logical construct or a single human being perform the user role. Conversely, a person can fill a role, part of a role, or multiple roles.

5. Macro-patterns for a Cloud Enabled Data Center

To simplify the design of a Cloud Enabled Data Center solution, use case sets were aggregated together into micro-patterns. To drive this abstraction to a higher level, multiple micro-patterns are grouped into a macro-pattern. Macro-patterns provide incremental increases in capabilities and functionality and describe, from an architectural perspective, the solution to implement. Each macro-pattern represents a subset of the entire Cloud Enabled Data Center adoption pattern and addresses a business problem through a specific set of capabilities. The following macro-patterns, built on top of each other, are defined for the Cloud Enabled Data Center adoption pattern:

- *Simple Infrastructure as a Service* - This macro-pattern is the entry point in the IaaS cloud space; organizations can use it to start building a multitenant cloud infrastructure to deliver simple virtual machines (configured with the appropriate network and storage) that cover the most common business needs for cloud computing;
- *Cloud Management* - This macro-pattern complements the Simple IaaS macro-pattern by adding capabilities to manage requirements as Service Level Agreements (SLA), security, resiliency, and capacity planning; this

Building a Cloud Enabled Data Center macro-pattern helps to optimize the IT processes, manage complexity of virtualization and automation, and increase efficiency for both the infrastructure that provides the cloud and the cloud service itself. This macro-pattern addresses non-functional requirements in the area of reliability, availability, and basic security.

- *Advanced Infrastructure as a Service* - This macro-pattern is used to create a more sophisticated cloud infrastructure for delivery and management of complex and critical IaaS in highly demanding environments. It also can deploy services in more than one data center or scale out to off-premise public clouds. Dealing with the complex, production-like environments, this macro-pattern also addresses non-functional requirements in the area of performance, scalability, extensibility, and advanced security.
- *ITIL managed Infrastructure as a Service* - With this macro-pattern, the Advanced IaaS macro-pattern is integrated with the processes and best practices of Information Technology Infrastructure Library (ITIL). This macro-pattern defines a Cloud Computing environment for the integration with the existing enterprise applications, systems, and processes. This integration is accomplished by including the cloud infrastructure and services in the enterprise ITIL processes.

Figure 1 shows a schematic view of the four macro-patterns and their key capabilities; it also shows how these macro-patterns stack one on top of the other to build more sophisticated Cloud Enabled Data Center solutions.

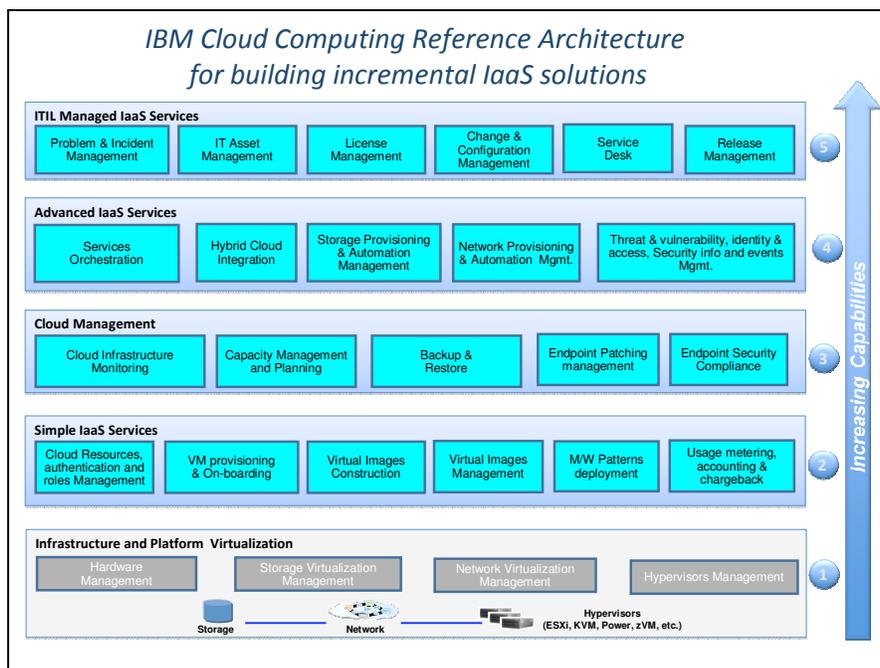


Figure 1 – Cloud Enabled Data Center macro-patters

6. Simple Infrastructure as a Service

The Simple Infrastructure as a Service macro-pattern, also called the virtual machines provisioning macro-pattern, focuses on automating the delivery of virtual machines using a Cloud Computing model. It is the starting point to a Cloud Enabled Data Center solution. This macro-pattern has the following main aspects:

- virtual machines provisioning which focuses on the following capabilities:
 - virtual machines self-service provisioning in a multitenant environment;
 - the possibility to attach virtual machines to preconfigured networks and storage;
 - enablement of on-boarding existing virtual machines under control of the cloud management infrastructure;
- metering and accounting services:
 - even if the customer is not a cloud service provider, metering and accounting can be used to measure the usage of cloud services and quantify potential savings.

The Simple IaaS macro-pattern contains the following functional blocks of key capabilities:

- Virtual machines provisioning and on-boarding - The existing virtualization capabilities are used to provide a simple interface to consumers of IT to provision virtual machines to satisfy the customers' needs. The underlying resources that are required by a virtual machine (such as processor, memory, storage, and network resources) are preconfigured and made available as resources that are attached to a virtual machine as requested. There might be a simple workflow to approve creation of a virtual machine. Virtual machines that exist in the data center might be discovered by this macro-pattern and on-boarded under the control of the cloud management infrastructure. Alternatively, they can be migrated from one cloud environment to another.
- Cloud administration - Cloud administration provides capabilities to administer a cloud environment, such as adding new storage or computational resources in the cloud pool or defining new segregated networks.
- Role and authentication management - The Cloud Enabled Data Center ecosystem has many users who perform different roles. This capability can assign users to roles and can ensure that authorized access to cloud resources exists to perform a task, such as approving the creation of virtual machines.
- Virtual machines image construction - This capability provides tools to build virtual images for deployment into the Cloud Enabled Data Center.
- Image management – Virtual machines tend to drift from the original copy (provided by the service catalog) as new security and operating system patches are installed. This capability provides tools to create new virtual machines, establish version control, search for and compare images, and delete images from your virtual images templates repositories.

- Usage metering, accounting and chargeback - Organizations need to meter the resources that are used by each workload running on the shared infrastructure. This capability provides tools to track the usage for each characteristic of a workload (such as processor, memory, storage, and network), rating each characteristic and generating consumption-based chargeback reports.

7. Cloud Management

As the Cloud Enabled Data Center solution matures and organizations begin to deal with business and mission-critical workloads, it is important to ensure that the underlying infrastructure performs well. Also, companies need to reduce the unexpected failures or overloads that could cause to miss the Service Level Agreements (SLA). Such workloads require better orchestration and support to recover quickly if a disaster occurs. Capacity planning becomes more important, because the business expects IT to be more agile and accommodate new workloads quickly.

These additional requirements are fulfilled by the set of management tools that are provided by the Cloud Management macro-pattern. The Cloud Management macro-pattern has various optional services and components that provide the following services:

- Health management of the cloud management environment and cloud-managed services;
- Capacity planning of cloud infrastructure;
- Backup and restore of cloud services;
- Event management;
- Patch management and security compliance.

The functional blocks in the Cloud Management macro-pattern deliver the following key capabilities:

- Virtual infrastructure monitoring - A highly virtualized infrastructure requires a virtualization monitoring tool that provides end-to-end visibility. Although the virtualization helps to reduce costs, improve availability and increase flexibility, it makes it difficult to determine the root cause of service performance degradation. This capability provides the tools to monitor virtual infrastructure utilization, ensure availability of business critical applications and quickly identify, isolate and remedy performance issues.
- Capacity planning - One promise of cloud computing is that virtualization reduces the number of servers that are needed. Therefore, companies must identify the balanced amount of cloud infrastructure that is required to meet the anticipated needs of users. This capability provides the tool for governance and helps capacity planners to define the right business context and requirements. The analysis is necessary so the Cloud Enabled Data Center can continue to meet the performance goals of a business application while contributing to the organization's financial goals.
- Event management - As the data center delivers services to support business-critical applications, operations must identify exception situations quickly and determine the root cause of the problem. This capability

provides the tools and best practices to collect events that come from the different elements of the infrastructure, correlate them and eventually trigger notifications or remediation actions.

- Backup and restore - This capability provides tools and governance for backup processes to enable service to be restored within agreed service levels to ensure that the business is protected.
- Patch management - This capability provides tools and governance to ensure that the provisioned virtual machines are patched to the right security and version levels.
- Endpoint compliance and management - This capability provides tools to ensure that all the deployed virtual machines are compliant with the organization's security and enterprise policies.

8. Advanced Infrastructure as a Service

The complexity of the virtualized infrastructure increases as the Cloud Enabled Data Center solution is used to deliver business-critical applications. These applications have typically strict security requirements. Therefore, they need to be configured with security elements such as firewalls. They usually support high volumes of data or users. As a result, they are required to attach high-quality storage to the virtual machines or to configure load balancers to optimize their utilization. There is increased pressure to eliminate application downtime from planned server maintenance. High availability might be required across entire virtualized infrastructure to withstand physical server failures. The compute clusters might require spare nodes to withstand single node failures. The main functional blocks that implement this macro-pattern deliver the following key capabilities:

- Storage provisioning and automation management - This capability provides functions for the discovery of storage resources, dynamic provisioning, storage resources allocation, monitoring, provisioning, and configuration capabilities across different heterogeneous devices. With this capability, organizations automatically provision different types of storage resources, making them available to the computing resources. These provisioned storage resources can be used to generate a new revenue stream by offering a new service such as storage-as-a-service.
- Network provisioning and automation management - This capability provides functions for management and discovery of network resources, dynamic provisioning and allocation, monitoring across heterogeneous devices. With this capability, companies could automatically configure a load balancer or allocate a new network that can be deployed as part of a cloud service and used by the virtual machines. These network resources can be used to generate a new revenue stream by offering them as a network-as-a-service.
- Cloud services orchestration - One of the most powerful capabilities provided by this macro-pattern is the automated arrangement, coordination, and management of complex computer systems, middleware, and services. Organizations can use this capability to deliver a complex business

application that consists of two or more virtual machines with attached storage and interconnectivity. Another example is to provision across data centers by using a single self-service interface. The self-service interface includes the following capabilities:

- Self-service provisioning of infrastructure elements such as virtual machines, storage or network elements with multi-tenancy;
- Self-service provisioning of complex integrated infrastructure services.
- Hybrid cloud integration - Hybrid cloud integration provides the capability to provision to external Cloud Computing environments (such as public clouds) to meet peak workloads. This optional capability makes it possible to provision virtual machines into a public cloud and to securely connect and manage from a private infrastructure. It can implement cloud bursting scenarios in case of workload peaks or for cost reasons.
- Advanced security - The advanced security capabilities support more complex environments, which require more sophisticated security to protect against threats and vulnerability. This support is accomplished by providing identity and access controls, security information, event management and log information management.

9. ITIL managed Infrastructure as a Service

Because the Cloud Enabled Data Center solution delivers complex cloud services, its management processes must be integrated with the organization's IT management processes such as change management and incident management.

This capability is typically required when delivering complex and composite services in environments similar to production environments. This macro-pattern allows for registration of cloud services in a Configuration Management Database (CMDB) and, for the services, to be placed under control of the major enterprise ITIL processes. The Infrastructure as a Service managed according to ITIL processes and best practices includes the following capabilities:

- Problem and incident management
- IT assets management
- Licenses management
- Change and configuration management
- Service desk
- Release management

This capability is also a prerequisite to move the Cloud Enabled Data Center solution to a Cloud Service Provider business model. In the Cloud Service Provider business model, governance and control capabilities are important to ensure the Quality of Service (QoS) and Service Level Agreements (SLA) that are typical of the Cloud Service Provider business.

10. Conclusions

The Cloud Enabled Data Center adoption pattern of the IBM Cloud Computing Reference Architecture provides a comprehensive business and architecture approach to companies that want to establish or extend an Infrastructure as a Service solution. By applying the Cloud Computing Reference Architecture and its adoption patterns, organizations can take advantages of the IBM Cloud Computing expertise garnered by working with companies around the world and by deploying Cloud Computing solutions internally.

The Cloud Enabled Data Center adoption pattern covers the business and technical needs and long-term plans for enterprises. It shows how to apply the essential requirements to support an initial deployment and explains how to grow the solution as business needs change and expand to support new business opportunities and business models.

The Cloud Enabled Data Center adoption pattern within the IBM Cloud Computing Reference Architecture has been implemented and deployed with a wide variety of customers. It provides a sound foundation for building an Infrastructure as a Service solution. Companies have created a wide range of Infrastructure-as-a-Service solutions from simple single purpose solutions to sophisticated deployments, aligned to best practices of Information Technology Infrastructure Library (ITIL). As Cloud Computing becomes a more integral part of the computing fabric, the flexible architectural model of the Cloud Enabled Data Center adoption pattern will provide a robust platform for business growth.

References

- [1] IBM Market Insights. Cloud Computing Research, July 2009.
- [2] 2010 IBM Tech Trends Survey. October, 2010.
www.ibm.com/press/us/en/pressrelease/32674.wss
- [3] The Open Group SOA Reference Architecture.
http://www.opengroup.org/soa/drafts/refarch.htm#_SOA_Reference_Architecture
- [4] Peter Mell and Tim Grance, The NIST Definition of Cloud Computing, Version 15, October 7, 2009. <http://www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf>
- [5] Raffaele Stifani, Cloud Computing Reference Architecture, Atti del Congresso Nazionale AICA 2011, AICA
- [6] Raffaele Stifani, Il Cloud Computing: sorge una nuova era, Atti del Congresso Nazionale AICA 2009, AICA
- [7] Raffaele Stifani, Le potenzialità del Cloud Computing, Atti del Congresso Nazionale AICA 2010, AICA

Applicazioni Smart per Città Smart: un Contributo all’Innovazione

Domenico Saccà, Simona Citrigno¹, Sabrina Graziano²
Università della Calabria
Viale Pietro Bucci 41C, 87036 Rende (CS)
sacca@unical.it

¹Centro di Competenza ICT-SUD
Polo Tecnologico Unical, Piazza Vermicelli, 87036 Rende (CS)
simona.citrigno@cc-ict-sud.it

²OKT Srl
Polo Tecnologico Unical, Piazza Vermicelli, 87036 Rende (CS)
sabrina.graziano@okt-srl.com

Abstract *This paper provides an illustration of main research activities and results of the PON project “TETRis – TETRA Innovative Open Source Services”. TETRis focuses on innovative services for Smart City/Smart Territory via the definition of technological tools and intelligent platforms which enable local organizations to acquire, represent and manage data and information gathered from sensors and devices by means of several communication systems deployed in the vest of add-value services. The communication system TETRA and its integration with other commercial mobile networks will enable the realization of complex Smart Environments that can provide advanced services to citizen and operators for environmental monitoring, urban mobility and emergencies management.*

Keywords: Urban Monitoring, Urban Mobility, Intelligent Platforms

1. Introduzione

Le attività descritte nel presente articolo riguardano l’identificazione di scenari per la realizzazione di servizi innovativi finalizzati ad una gestione intelligente di un territorio urbano ed extraurbano.

Tali attività hanno permesso l’identificazione di soluzioni e piattaforme tecnologiche innovative che consentono una nuova operatività agli enti presenti sul territorio (comuni, province, regioni, università, ecc), ai cittadini e agli operatori coinvolti. Viene infatti esteso e valorizzato il patrimonio informativo acquisito sul territorio attraverso l’utilizzo di dispositivi e sensori interconnessi mediante sistemi di comunicazione locali e remoti che supportano servizi ad alto valore aggiunto volti a migliorare la qualità del territorio in termini di vivibilità e sostenibilità anche attraverso il coinvolgimento dei cittadini che diventano i

principali tutori del territorio, i cosiddetti “sensori sociali”, per la rilevazione di situazioni critiche relative al contesto in cui risiedono.

Gli scenari d'innovazione descritti nel presente articolo sono stati realizzati nell'ambito del progetto PON 2007 2013 - Ricerca e Competitività “TETRis – Servizi Innovativi Open Source su TETRA” e si collocano all'interno dello schema di riferimento generale di “Internet delle Cose” a supporto di Smart City/Smart Territory, in cui l'acquisizione dei dati dagli oggetti può essere applicata ad ampie aree territoriali tramite la disponibilità diffusa delle reti di comunicazione [Commissione Europea, 2012], [<http://www.internet-of-things-research.eu/documents.htm>]. I dati raccolti, opportunamente valorizzati, alimentano servizi innovativi orientati alla produzione ed allo scambio di conoscenza fra i diversi attori interconnessi dalle reti urbane e territoriali. Lo sviluppo di tali servizi è stato realizzato mediante la cooperazione di dispositivi intelligenti, operatori e fruitori dei servizi stessi, tutti identificabili come Smart Object.

Le attività descritte nel presente articolo riguardano due scenari principali individuati all'interno di una Smart City: (A) *mobilità urbana*, (B) *monitoraggio, controllo e manutenzione del territorio*.

Il presente articolo è così composto: una introduzione riguardante il progetto TETRis e i suoi obiettivi; una descrizione delle applicazioni innovative individuate; una descrizione delle sperimentazioni attuate negli ambiti della Mobilità urbana e del Monitoraggio urbano.

2. Obiettivi del Progetto TETRis

Obiettivo del progetto TETRis è quello di realizzare servizi ad alto valore aggiunto nell'ottica delle Smart City e Smart Territory [Kanter e Litow, 2009], [Komninos et al, 2011] anche estendendo, laddove possibile, le funzioni del sistema di comunicazione TETRA e agendo lungo tre assi principali: L'evoluzione e l'apertura degli ambiti applicativi propri del sistema di comunicazione TETRA al fine di definire nuovi servizi informativi per gli operatori, che sfruttino nuovi modelli e strumenti Open Source per l'interconnessione di TETRA con altre reti, e di identificare nuove tipologie di dispositivi ottenuti attraverso l'interoperabilità del sistema TETRA con sensori e reti di sensori esistenti; La modellazione e prototipazione di un framework Open Source che permette di definire una specifica ed un modello di cooperazione di Smart Object e la loro gestione all'interno dei relativi Smart Environment; L'identificazione di scenari e modelli applicativi in ottica di servizi per Smart City /Smart Territory attraverso la definizione di Smart Environment e servizi Open Source ad alto valore aggiunto applicati ai domini del monitoraggio del territorio, della gestione delle emergenze, della mobilità urbana ed extraurbana e dei servizi ai cittadini.

Gli obiettivi specifici del progetto TETRis possono essere così formulati: Apportare benefici economici e sociali alla comunità attraverso interventi più mirati ed efficaci da parte degli operatori della Pubblica Amministrazione e della Pubblica Sicurezza in diversi scenari applicativi quali la gestione delle emergenze, la salvaguardia ambientale, la mobilità e i servizi ai cittadini, anche con il contributo dei cittadini stessi attraverso la condivisione di informazioni e l'uso di strumenti innovativi di social networking; Estendere la pervasività e l'efficacia dei servizi delle amministrazioni pubbliche, degli enti strumentali, delle polizie locali, degli operatori sanitari, delle aziende di trasporto nei territori di riferimento; Migliorare la qualità della vita e del senso di sicurezza dei cittadini per mezzo della diffusione di canali sicuri e affidabili di infrastrutture tecnologiche "always on".

Le attività di progetto sono organizzate in obiettivi realizzativi (OR) che prevedono un insieme di attività di Ricerca industriale e di Sviluppo Sperimentale .

Il presente articolo illustra gli OR di tipo sperimentale ed in particolare:

- "*TETRis Smart Environments per la mobilità*" - relativo alla realizzazione di un modello per il rilevamento delle problematiche della mobilità in area urbana tramite l'utilizzo di appositi Smart Object distribuiti sul territorio e l'impiego di una rete di sensori ad essi collegati. I dati rilevati vengono raccolti e aggregati in un data warehouse che alimenta una piattaforma di *Mobility Intelligence* definita progettando tecniche innovative di analisi di dati spazio-temporali e di mining di dati complessi, incluso traiettorie.
- "*TETRis Smart Environments per il monitoraggio del territorio e l'erogazione di servizi ai cittadini*" - relativo alla definizione di uno Smart Environment in cui vengono predisposti alcuni sensori fisici connessi a Smart Object e arricchiti da sensori "sociali", che rilevano in tempo reale lo stato del territorio. I dati rilevati vengono raccolti e aggregati in un data warehouse che alimenta una piattaforma di *Territory Intelligence* che consente l'estrazione e l'elaborazione delle informazioni relative al monitoraggio del territorio in base all'ambito di riferimento del dato stesso.

All'interno del progetto sono stati siglati degli accordi di collaborazione con il Comune di Cosenza, per quanto riguarda l'ambito della Mobilità Urbana, e con il Comune di Rende, per l'ambito del Monitoraggio Urbano. I comuni si sono dimostrati fortemente interessati alle sperimentazioni di tecniche e soluzioni informatiche innovative in quanto il loro intento è quello di perseguire la realizzazione di una Smart City, ossia un nuovo modello di città, vista come sistema intelligente che ruota intorno alle esigenze degli individui, che vigila sul

rispetto e sul controllo dell'ambiente e che gestisce in maniera ottimale le risorse attraverso l'utilizzo di infrastrutture tecnologiche e di strumenti ICT innovativi al fine di erogare servizi a valore aggiunto per i cittadini [<http://www.smart-cities.eu/>].

3. Progetto di Applicazioni Innovative per Smart City

Lo scenario di base che è stato individuato per una tipica applicazione a supporto di una Smart City comprende i seguenti elementi:

- **Attori** divisi in due categorie: **Enti**: un Ente Gestore dello Smart Environment e della rete degli Smart Object (EGSE) dislocati nel territorio; un certo numero di Enti Territoriali Esterni (ETE) interessati ai servizi dello Smart Environment e degli Smart Object con cui interagiscono; **Individui**: possono essere cittadini o operatori che svolgono i loro compiti nel territorio e sono alle dipendenze dell'EGSE o degli ETE;
- I **sistemi centrali** degli Enti: Il sistema dell'EGSE, che è responsabile del funzionamento complessivo dell'applicazione e controlla la rete degli Smart Object distribuiti sul territorio; I sistemi degli ETE, che interagiscono con gli Smart Object sulla base di protocolli di funzionamento concordati con l'EGSE e che possono anche interagire con il sistema centrale dell'EGSE;
- Le **applicazioni** su smart phone per individui;
- Gli **Smart-Object**: una rete di Smart Object distribuiti sul territorio sotto la gestione del sistema centrale dell'EGSE, che presentano due tipologie di comunicazione: Remota: (i) con il sistema centrale dell'EGSE attraverso TETRA (ii) con i sistemi centrali degli ETE attraverso protocolli commerciali di telecomunicazione (GSM, UMTS, ecc.) ed eventualmente anche con TETRA; Locale: (i) con gli operatori attraverso WiFi e tecnologia NFC per l'attivazione immediata e personalizzazione della interazione, (ii) con i cittadini attraverso WiFi, (iii) con oggetti fisici distribuiti su un'area territoriale contigua a uno Smart Object e collegati ad esso tipicamente attraverso protocollo ZigBee o RFID.
- I **Sensori**, che possono essere: classici sensori di rilevamento traffico collegati tra di loro con rete ad hoc con protocollo ZigBee, RFID e attestata ad uno Smart Object; classici sensori di rilevamento condizioni ambientali (temperatura, umidità, acustica, accelerometri, estensimetri, ecc) collegati tra di loro con una rete ad hoc; sensori sociali, ossia cittadini o operatori dotati di un dispositivo mobile e di applicazioni realizzate ad hoc in grado di segnalare e/o monitorare eventi che accadono sul territorio di interesse; dispositivi sensori/attuatori di urbotica (cioè la domotica a livello urbano) che interagiscono con lo Smart Object o direttamente tipicamente attraverso protocolli di basso consumo ma ad alte prestazioni quali ad esempio Cliffside e Wibree, o attraverso un apposito kit di controllo connesso allo Smart Object in WiFi.

In Figura 1 è rappresentato il quadro complessivo di uno scenario tipico di utilizzo di Smart Object all'interno di una Smart City con l'indicazione delle varie componenti e delle loro interazioni.

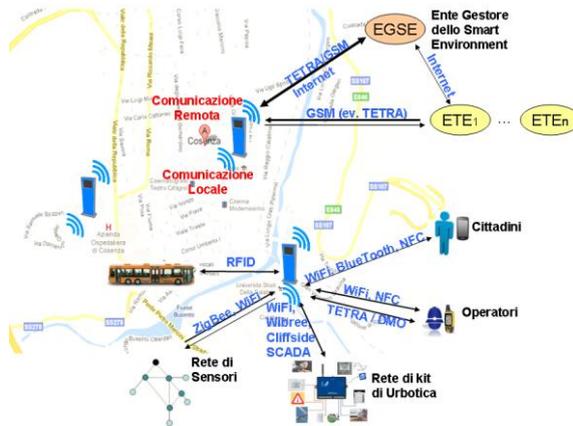


Fig 1. Scenario di utilizzo Smart Object

In Figura 2 è rappresentata un'architettura di massima dello Smart Object.

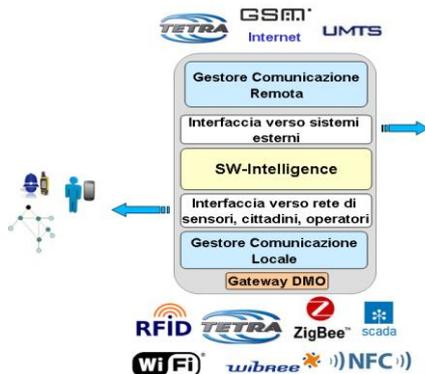


Fig 2. Architettura Smart Object

4. Mobilità Urbana

Le attività all'interno di tale ambito sono state mirate alla realizzazione di un modello basato sulla definizione di Smart Object e di una rete di sensori ad essi collegati attraverso il quale è possibile rilevare in tempo reale i dati relativi alla mobilità di un'area urbana. Tale modello prevede anche la possibilità di rilasciare servizi ad operatori e cittadini attraverso l'utilizzo di dispositivi mobili.

I dati rilevati dai sensori e dagli Smart Object distribuiti sul territorio sono stati raccolti e aggregati in un data warehouse collegato ad una piattaforma di

Mobility Intelligence capace di effettuare analisi on-line in modo da fornire in tempo reale informazioni agli operatori degli enti territoriali interessati e ai cittadini per la gestione e l'utilizzazione intelligente dei sistemi di mobilità.

La sperimentazione è stata avviata sul territorio del Comune di Cosenza e sono state individuate due principali macro aree di sviluppo:

- rilevazione e gestione del traffico urbano tramite Smart Object fissi, ossia immagini e video da webcam e sensori di rilevazione del traffico posti in punti critici della città di Cosenza;
- rilevazione e gestione delle infrazioni commesse dalle auto sulle corsie preferenziali degli autobus tramite Smart Object mobili installati sui mezzi pubblici dell'azienda municipalizzata AMACO di Cosenza.

La piattaforma di Mobility Intelligence elabora i dati in tempo reale, seguendo le linee più recenti di ricerca nell'ambito dei sistemi di analisi OLAP e di data mining e fornisce un sistema di supporto alle decisioni il cui principale scopo è quello di permettere di estrarre, in tempi brevi e in modo flessibile, le informazioni che servono a soddisfare i fabbisogni di mobilità dei cittadini [Commissione Europea, 2011]. La presenza di cruscotti decisionali offre la possibilità alle amministrazioni interessate di prendere decisioni strategico/operative e di pianificare al meglio gli interventi sulle infrastrutture stradali, sulle tratte di trasporto pubblico, sui parcheggi disponibili e sulle infrastrutture di comunicazione multimodale con i cittadini. L'analisi dei dati attraverso la piattaforma di intelligence permette inoltre di offrire un valido supporto alla definizione del piano del traffico urbano avviando un'attività di calibrazione del modello in modo da ottenere quello definitivo e più vicino alla realtà da osservare.

La piattaforma, grazie all'integrazione di strumenti di data mining e di analisi dei dati georeferenziati, permette di visualizzare dati spaziali, modelli e risultati su una mappa geografica reale.

Il modello ha previsto anche la realizzazione delle seguenti classi di applicazioni rese disponibili ai cittadini e agli operatori attraverso gli Smart-Object e la piattaforma di Mobility Intelligence: 1. App per cittadini, che forniscono informazioni sugli autobus, i tempi di arrivo, gli orari, le posizioni delle fermate più vicine; informazioni per la ricerca dei parcheggi disponibili, visualizzandone la posizione sulla mappa, il numero dei posti liberi ed il percorso per raggiungerli; informazioni sul traffico e segnalazioni di ingorghi; informazioni di carattere informativo sulla città: luoghi di interesse, ristoranti, alberghi e locali. Tutte le informazioni sono contestualizzate in base alla geolocalizzazione del cittadino. 2. App per operatori, che permettono di visualizzare lo stato del traffico, la circolazione dei mezzi pubblici, le segnalazioni da parte dei cittadini e riportare nuove segnalazioni quali ad esempio ingorghi, parcheggi in doppia fila, manifestazioni ecc; App Multe, tramite la quale è possibile registrare violazioni del codice della strada

geolocalizzando il luogo dell'infrazione e allegando delle foto a comprova della violazione.

Le stesse informazioni sono state rese disponibili anche tramite un portale web.

Inoltre sono state realizzate le seguenti analisi di supporto della generazione di modelli di mining per la mobilità urbana [Fosca Giannotti et al, 2011], [Roberto Trasarti et al, 2010], [Fabio Pinelli et al, 2009]: 1. Studio di come la mobilità privata nell'area urbana di Cosenza può essere servita dal trasporto pubblico evidenziando eventuali mancanze, sprechi di risorse e suggerendo miglioramenti/potenziamenti; 2. Studio dei comportamenti frequenti degli autobus dell'azienda municipalizzata AMACO durante le tratte urbane analizzando le tratte e i diari di bordo degli autobus; 3. Analisi di raggiungibilità della città e delle aree limitrofe, utilizzando le tratte del trasporto pubblico e/o i dati di mobilità pubblica, per capire quale sia la reale distanza in termini di tempo tra le varie zone della città di Cosenza e come questa vari durante il giorno; 4. Profilazione della popolazione presente, utilizzando i dati di telefonia mobile e analizzando il tipo di utenti che frequentano una certa zona e il comportamento di chiamata.

5. Monitoraggio Urbano e Analisi del Rischio

Le attività all'interno di tale ambito sono state mirate alla definizione di uno Smart Environment e alla sua integrazione con alcuni Smart Object distribuiti sul territorio per la gestione e la manutenzione del contesto ambientale urbano e per l'erogazione di servizi ai cittadini. Sono stati pertanto predisposti nel territorio della sperimentazione una serie di sensori fisici e "sociali" tramite i quali è stato possibile rilevare in tempo reale lo stato del territorio stesso e i rischi ad esso connessi. La piattaforma di Territory Intelligence realizzata, raccoglie da una parte i dati acquisiti dagli Smart Object e, dall'altra, a seguito di opportune fasi di integrazione ed elaborazione dati utilizzando tecniche avanzate di data warehouse, rende disponibili le informazioni ai diversi soggetti che con essa interagiscono al fine di un corretto monitoraggio delle aree urbane ed extra-urbane interessate.

Il modello realizzato ha previsto anche la realizzazione di applicazioni contestualizzate rese disponibili a cittadini e operatori, impegnati nella manutenzione del territorio, attraverso i propri dispositivi mobili. Tali applicazioni consentono di comunicare tempestivamente con l'amministrazione pubblica segnalando situazioni di criticità, malfunzionamento nelle reti urbane, dissesti stradali e necessità di interventi di ripristino di condizioni di buona vivibilità del territorio. Gli eventi da segnalare possono essere suddivisi nelle seguenti macrocategorie: Insidie su strada, Randagismo, Servizi e Reti, Arredo urbano/Verde, Deposito rifiuti. Le informazioni sono rese disponibili anche

tramite un modulo web che funge da console di gestione delle segnalazioni trasmesse dalle app in modo da poter assegnare il relativo intervento ad un operatore o ad un gruppo di operatori, monitorare lo stato dell'intervento stesso, visualizzare gli Smart Object distribuiti nel territorio e realizzare un sistema di safety operator.

La sperimentazione è stata avviata sul territorio del Comune di Rende (CS) e gli ambiti di intervento individuati sono stati i seguenti: monitoraggio delle aree urbane ed extra-urbane per ciò che concerne la piccola manutenzione stradale, manutenzione della pubblica illuminazione, dell'arredo urbano, del verde cittadino, monitoraggio degli ambienti di lavoro indoor e monitoraggio strutturale.

Disponendo in tempo reale di informazioni provenienti da reti di sensori connessi a Smart Object opportunamente distribuiti sul territorio, si possono alimentare i modelli attualmente utilizzati in modo da stilare nuovi piani d'intervento, o modificare gli stessi in modo che un'amministrazione locale possa meglio rapportare la programmazione delle risorse ai bisogni in maniera dinamica e puntuale. I sensori distribuiti sul territorio sono stati progettati anche per rilevare, rispetto alle tematiche prescelte, le condizioni di superamento delle "soglie" che richiedono modifiche contingenti ai piani di manutenzione/gestione del territorio da parte degli enti preposti.

All'interno dell'ambito del monitoraggio urbano è stato implementato anche un modello di Risk Analysis per la circolazione stradale in area urbana al fine di identificare e valutare i rischi che incombono sul sistema oggetto di osservazione e di attivare le necessarie contromisure. L'individuazione dei possibili scenari di rischio e l'attività di monitoraggio consentono in maniera più agevole e razionale ai vari enti locali di attuare i piani di intervento per scongiurare il verificarsi del rischio o per ridurre l'impatto, per rilevare prontamente i rischi in atto e allertare gli operatori per un loro pronto intervento coinvolgendo anche i cittadini sia nella rilevazione delle situazioni critiche sia nell'assistenza immediata .

Infine, l'attività di sperimentazione realizzata sul territorio di riferimento, ha anche previsto l'impiego di reti di sensori wireless in grado di monitorare l'ambiente circostante [Rosi et al, 2011] inoltrando le informazioni acquisite ad un centro raccolta (gateway o Smart Object) attraverso la costruzione di una rete ad-hoc multihop. I dati rilevati tramite la rete di sensori riguardano il monitoraggio dei livelli di CO₂, CO, polveri, temperatura e umidità di una determinata area che possono essere collezionati per diverse settimane e resi accessibili successivamente in real-time e sintetizzati in un report. Alla fine dell'attività di monitoraggio di un determinato sito la rete può essere smantellata e impiegata in un nuovo luogo (edificio, area urbana) in modo da ridurre i costi e ripetere le stesse analisi. Inoltre, un'ulteriore rete di sensori è stata utilizzata per il monitoraggio della statica degli edifici in modo da fornire indicazioni su quali

siano i valori all'interno dei quali la situazione può ritenersi "sotto controllo" e quali siano invece valori critici per i quali è necessario un intervento o l'invio di un segnale di allarme.

6. Conclusioni

Il lavoro svolto nell'ambito del progetto TETRis ha permesso di studiare e sperimentare soluzioni innovative per la rigenerazione di contesti urbani ed extra urbani secondo la visione strategica integrata del modello Smart City. L'attuazione della sperimentazione è stata resa possibile grazie agli accordi di collaborazione e al coinvolgimento degli enti locali che si sono resi soggetti promotori per la riqualificazione e il rinnovamento del proprio tessuto sociale al fine di offrire servizi sempre più efficaci ed efficienti ai cittadini. Il progetto ha previsto anche il coinvolgimento dei cosiddetti 'sensori sociali', ossia i cittadini stessi e gli operatori amministrativi che possono collaborare per comunicare in tempo reale con le istituzioni tutte quelle segnalazioni relative a situazioni critiche quali vandalismo e incuria, presenza di rifiuti, buche, disservizi vari. Grazie agli strumenti ICT è possibile quindi potenziare il filo diretto cittadini - Pubblica Amministrazione e sensibilizzare i cittadini stessi a diventare parte integrante delle buone pratiche amministrative.

Ringraziamenti

Questo lavoro è stato sviluppato all'interno del progetto MIUR PON 2007 2013 - Ricerca e Competitività TETRis – *Servizi Innovativi Open Source su TETRA - D.D. n.1/Ric del 18 gennaio 2010.*

I partner del progetto sono: Orangee srl (coordinatore), SelexElsag Spa, Centro di Competenza ICT SUD (insieme alle aziende socie Methodi Srl, Kaleidos Srl, Sirfin Spa, SCAI LAB Srl), ICAR-CNR, H2i Srl, Exeura Srl, Sinapsys Srl, TSC Consulting Srl, CNIT, CINI, Università della Calabria, Università Mediterranea di Reggio Calabria.

Gli autori dell'articolo ringraziano i seguenti partner e referenti di progetto per l'importante contributo apportato:

Andrea Vitaletti, Ugo Colesanti (Dipartimento di Ingegneria Informatica Automatica e Gestionale Antonio Ruberti della Sapienza Università di Roma), *Fosca Giannotti, Dino Pedreschi, Barbara Furletti, Lorenzo Gabrielli, Roberto Trasarti* (KDD Lab ISTI CNR - Pisa), *Rosario Curia, Loredana Sisca* (H2i Srl – Cosenza), *Michele De Buono* (SCAI LAB Srl), *Raffaele Bianco, Salvatore Pirruccio* (Sinapsys Srl – Soverato , CZ), *Roberto De Donato* (Sirfin SpA – Cosenza), *Sergio Scrivano* (Methodi Srl – Cosenza), *Giuseppe Musso, Francesco Scarpelli, Luigi Leonetti* (Kaleidos Srl – Cosenza), *Geppino De Rose, Leonardo Acri, Maria Rosaria Mossuto, Roberto Caruso* (Comune di Cosenza), *Luigi Mamone, Corrado Zoccali, Vincenzo Settino* (Comune di Rende).

Bibliografia

[1] Fosca Giannotti, Mirco Nanni, Dino Pedreschi, Fabio Pinelli, Chiara Renso, Salvatore Rinzivillo, Roberto Trasarti: Unveiling the complexity of human mobility by querying and mining massive trajectory data. VLDB Journal Special issue on Data Management for Mobile Services (2011).

[2] N. Komninos, H. Schaffers, M.Pallot, "Developing a Policy Roadmap for Smart Cities and the Future Internet", eChallenges e-2011 Conference

[3] Moss Kanter R., Litow S.S., "Informed and Interconnected: A Manifesto for Smarter Cities", Working Progress, 09-141, 2009

[4] PHC09: Fabio Pinelli, Anyang Hou, Francesco Calabrese, Mirco Nanni, Chris Zegras, Carlo Ratti, Space and time-dependant bus accessibility: a case study in Rome. Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, 2009

[5] A. Rosi, M. Berti, N. Biccocchi, G. Castelli, M. Mamei, A. Corsini, F. Zambonelli, "Landslide Monitoring with Sensor Networks: Experiences and Lessons Learnt from a Real-world Deployment", International Journal of Wireless Sensor Networks, Seattle, 2011, to appear.

[6] Roberto Trasarti, Fosca Giannotti, Mirco Nanni, Dino Pedreschi, Chiara Renso: A Query Language for Mobility Data Mining. International Journal of Data Warehousing and Mining (IJDWM) 2010

[7] <http://www.smart-cities.eu/>

[8] Commissione Europea, "Intelligent Transport Systems in Action, action plan and legal framework for the deployment of intelligent transport systems (ITS) in Europe" 2011

[9] Commissione Europea, "Smart Cities and Communities – European Innovation Partnership", 2012

[10] <http://www.internet-of-things-research.eu/documents.htm> - The Internet of Things 2012 - New Horizons - Cluster Book 2012

Un metodo di valutazione delle emissioni di CO₂ di un sito eCommerce B2C

Valerio Morfino¹, Salvatore Rampone^{1,2}

¹*Futuridea Innovazione Utile e Sostenibile*
Via Piano Cappelle, Benevento, Italy
valerio.morfino@ctcgroup.it

²*Dipartimento di Scienze e Tecnologie - Università del Sannio*
Via dei Mulini 59/A, Benevento, Italy
rampone@unisannio.it

The use of the Internet has grown continuously in recent years and it is expected it still continues to grow significantly. The eCommerce applications are also growing significantly. In this work the main elements that determine the emission of CO₂ by this specific type of applications are analyzed. We highlight the role played by hardware and software factors, but also by elements which are not strictly technical as Web Design and Web Marketing. Finally we describe an explicit method for calculating the CO₂ emissions.

Keywords: eCommerce; B2C; CO₂ emissions; Computing method.

1. Introduzione

Probabilmente è capitato a tutti di leggere in fondo ad un messaggio di posta elettronica un invito del tipo: "Rispetta il tuo ambiente: pensa prima di stampare questa e-mail". Facciamo volentieri questo piccolo gesto per difendere la natura. Il rapporto tra il nostro gesto e l'ambiente è immediato: non stampiamo il messaggio e-mail e così facendo inquiniamo di meno, utilizzando meno carta e meno inchiostro della stampante.

Non stampare un'e-mail non è però sufficiente a rendere completamente ecologico l'utilizzo del computer e di Internet. Tutte le operazioni che facciamo al computer producono delle emissioni di anidride carbonica (CO₂). Essa è indispensabile per la vita e per la fotosintesi delle piante, ma è anche responsabile dell'aumento dell'effetto serra [Macrini, 2010].

L'invio di una e-mail produce 4 grammi di CO₂ che diventano 50 con un allegato di grandi dimensioni e che possono moltiplicarsi fino a 4 volte se in copia conoscenza ci sono 10 destinatari [Ademe, 2011]. Una ricerca su Google

ne produce 0,2 g [Hölzle, 2009]. In un anno, Skype ne produce 24 milioni di tonnellate, Facebook 13.6 e tutta Internet 300 milioni [Atkinson, 2012].

Nel suo complesso l'Information e Communication Technology (ICT) contribuisce per circa il 2% delle emissioni globali di CO₂ [Petty, 2007] [Kiplier et al, 2011] [Chan et al, 2012]. Si tratta delle stesse emissioni di CO₂ dell'industria aeronautica [Hanle e Richter-Shalaby, 2010] [Petty, 2007] [Gunaratne et al, 2005]. Si stima che entro il 2020 le emissioni di CO₂ dell'industria ICT dovrebbero almeno raddoppiare [Chan et al, 2012].

Nel seguito dello scritto si cercherà di identificare i principali elementi che determinano emissione di CO₂ nel caso particolare dell'eCommerce Business to Consumer (B2C) e cioè la vendita su Internet rivolta a consumatori finali.

Questo particolare mercato sta infatti avendo una crescita costante in tutto il mondo negli ultimi anni. Ad esempio negli Stati Uniti ha raggiunto nel 2012 i 343 miliardi di dollari, crescendo del 14% rispetto al 2011; l'Europa occidentale nel 2012 ha raggiunto quota 173 miliardi di euro, con un incremento dell'11% rispetto all'anno precedente. In Italia nel 2012 il valore del transato è aumentato del 18%, sfiorando il totale di 11 Miliardi di euro. Gli utenti dei siti eCommerce sono cresciuti del 33% rispetto al 2011 toccando i 12 milioni. Quindi circa il 40% degli utenti Internet italiani [Rangone et al, 2012].

Dopo aver evidenziato, nella Sezione 2, come un sito eCommerce produca CO₂, nella Sezione 3 verrà illustrato come la maggiore o minore emissione di CO₂ dipenda da aspetti tecnici ma anche da aspetti non tecnici relativi al web-design ed al web-marketing. Infine, nella Sezione 4, sulla base delle valutazioni della precedente Sezione, verrà sintetizzato un esplicito metodo di calcolo dell'emissione di CO₂ di un sito eCommerce B2C.

2. Siti eCommerce e produzione di CO₂

Un sito eCommerce necessita di diverse componenti e servizi per il suo funzionamento, come hardware, software, connessioni di rete ed anche servizi logistici (ad. esempio le spedizioni). Alcuni di questi elementi - come le spedizioni - hanno una emissione diretta di CO₂, altri utilizzano direttamente - come l'hardware - od indirettamente - come il software - energia elettrica, la quale è prodotta, almeno per una quota, attraverso la combustione di materiale che emette gas serra.

Il consumo di energia associato ad un servizio informativo può essere convertito in emissioni di CO₂ utilizzando fattori di conversione normalmente forniti da agenzie governative su base nazionale o regionale [IEA, 2012] [DEFRA 2012]. Per l'Italia i valori possono essere recuperati dai rapporti mensili di Terna Rete Italia. Ad esempio in Italia a maggio 2013 la produzione di energia termoelettrica è stata pari al 56% della produzione totale [Terna, 2013].

Ovviamente anche per la produzione degli apparati viene emessa CO₂, ma il nostro scopo è comprendere quali siano gli elementi che hanno un ruolo diretto sul consumo di energia elettrica, e dunque sulla produzione di CO₂ nell'utilizzo di un sito eCommerce. In particolare si presterà attenzione a tutti gli elementi hardware, software e servizi legati all'applicativo eCommerce sul server. Non

vengono considerate le emissioni prodotte dalla periferica con la quale l'utente finale utilizza il sito (pc, tablet, ecc.).

3. Componenti di un sistema eCommerce

Per la valutazione delle emissioni totali di CO₂ di un sistema eCommerce ci concentreremo sull'analisi dei seguenti elementi:

- Hardware, Software di Base e Networking
- Software
- Web Design ed usabilità
- Web Marketing
- Altri elementi

Ad eccezione degli elementi esterni, in particolare il trasporto della merce, tutti gli altri elementi non producono direttamente CO₂. Il loro contributo in termini di emissione è riferito al consumo di energia elettrica. Si cercherà quindi di stimare quale sia il contributo degli elementi sul consumo di energia elettrica.

Nel caso delle componenti hardware il consumo di energia è diretto, in quanto assorbito direttamente dalle apparecchiature. Nel caso del software e delle altre componenti è indiretto, in quanto l'utilizzo di determinati algoritmi o procedimenti genera un maggiore utilizzo di determinate risorse di sistema.

Il modello di riferimento che utilizzeremo prevede che vi sia un consumo base del sistema a fermo (Idle) ed un consumo dovuto alle richieste a cui il sistema deve rispondere [Chan et al, 2012]. Il consumo della prima tipologia è stimato per i server in circa il 50%-60% della potenza di picco [Kansal, 2008], che può arrivare anche al 80%-90% nel caso degli apparati di rete [Mahadevan, 2009], sebbene sia auspicabile che, migliorando la tecnologia, questo valore tenda a diminuire. Il consumo della seconda tipologia, oggetto della nostra riflessione, dipende dall'utilizzo delle risorse del sistema. Esso è influenzato dalla qualità del software, dalle caratteristiche grafiche e di usabilità del sito e dalle strategie marketing adottate per pubblicizzare l'iniziativa eCommerce.

3.1 Hardware, Software di Base e Networking

Gli elementi hardware utilizzati da un sistema di eCommerce, eccettuato il client con cui l'utente finale visualizza il sito, sono i server e gli apparati di rete.

Il consumo di energia elettrica di un server è variabile in base alla sua architettura ed al Data Center che lo ospita. Inoltre sempre più spesso il server può essere virtuale od un sistema con architettura Cloud Computing. Per la valutazione del consumo del server esistono diversi modelli [Kansal, 2008] [Kooimey, 2007] e dei benchmark che permettono di valutare le prestazioni [SPEC, 2012] e di misurare l'energia computazionale del sistema o di singoli servizi [Goraczko et al, 2012]. Nel nostro modello indicheremo il consumo dovuto all'hardware ed al software di base in stato Idle, riferito alla generazione di una singola pagina, con un valore medio costante:

$$H_p (1)$$

Tale contributo può essere misurato utilizzando uno dei sistemi di benchmark indicati.

Per quanto riguarda le emissioni degli apparati di rete che connettono il server del sito eCommerce con la periferica dell'utente finale si rimanda ad un già citato lavoro [Chan et al, 2012] che fornisce specifici modelli di stima.

3.2 Software

Un sito eCommerce B2C è un sito web con la specifica funzione di permettere la vendita di prodotti o servizi ad un pubblico consumer.

Esso è composto da un insieme di pagine web che contengono informazioni sui prodotti, tipicamente generate dinamicamente da un Content Management System (CMS), e da pagine informative e di supporto.

Comunemente vengono utilizzati linguaggi di programmazione quali Php, Java, Perl, Asp, Aspx ed altri. Tali linguaggi vengono utilizzati per la generazione delle pagine HTML che contengono i contenuti del sito e che vengono visualizzate dal browser presente sul dispositivo del fruitore finale. Normalmente i dati utilizzati dal sito risiedono in parte su File System ed in parte su Data Base. Le pagine HTML e le risorse del sito vengono normalmente gestite da un Web Server che, attraverso il protocollo HTTP o HTTPS gestisce la loro trasmissione al Browser presente sulla periferica dell'utilizzatore.

Il sito di eCommerce può essere sviluppato a partire da prodotti già esistenti, commerciali od open source, framework applicativi, oppure utilizzando semplicemente il linguaggio di programmazione. Questo può comportare qualità e prestazioni del software finale molto diverse. In ogni caso, le pagine web vengono generate dinamicamente utilizzando:

- informazioni digitate dall'utente.
- informazioni presenti nel data base dei prodotti
- informazioni relative all'utilizzo ed alla navigazione recente del sito (es. lingua scelta, dispositivo utilizzato dall'utente, pagine già visitate, prodotti nel carrello).
- informazioni inviate dal browser del cliente e relative a navigazioni precedenti (es. prodotti già visti, userid e password, ecc.). Tali informazioni sono contenute generalmente nei Cookies del Browser dell'utente.

Nella generazione della pagina web, vengono utilizzate direttamente tre tipologie di risorse del server:

- CPU, per la produzione dinamica delle pagine
- memoria RAM, per la memorizzazione temporanea dei dati
- Disco Rigido, per la persistenza delle informazioni

Le risorse che hanno un impatto diretto e significativo sul consumo di energia elettrica in base alla specifica elaborazione in corso possono essere considerate principalmente la CPU ed il Disco Rigido [Goraczko et al, 2012]. Viene trascurata la RAM per l'impatto basso rispetto alle altre due risorse ed il Monitor in quanto non utilizzato su server Web.

La risorsa CPU viene impiegata in modo significativo durante la generazione delle pagine Web dinamiche. Maggiore è la complessità algoritmica, maggiore è l'utilizzo di questa risorsa [Kansal, 2008]. Anche l'esecuzione di query sul database comporta un largo utilizzo di CPU.

Ne segue che sul tempo di utilizzo della CPU nella generazione di una pagina Web hanno un impatto importante elementi quali, a titolo esemplificativo:

- la qualità e la complessità degli algoritmi utilizzati nelle pagine
 - il corretto dimensionamento dei tipi di dati utilizzati
 - la qualità del disegno del database, dei suoi indici e delle query
 - il corretto utilizzo di meccanismi di caching delle query e delle pagine web
- Indichiamo con

$$P_{cpu} \quad (2)$$

la potenza media erogata al secondo dalla CPU.

La risorsa disco rigido viene utilizzata principalmente dal Web Server per trasmettere elementi multimediali e dal Data Base. Quindi sull'utilizzo del disco hanno un impatto rilevante fattori quali:

- Utilizzo di risorse multimediali e binarie dimensionate e compresse opportunamente. In particolare immagini che siano della dimensioni appropriate rispetto all'uso del sito. Questo valore verrà considerato specificamente anche nel paragrafo dedicato al design
- Utilizzo opportuno del caching delle risorse sul browser [Google 2012]. Questo meccanismo consente di salvare localmente sul browser risorse comuni, quali immagini, ed evitarne il ricaricamento
- Evitare lo Swap su Disco, dimensionando opportunamente la memoria RAM ed evitando errori sul Data Base o nell'algoritmica della generazione delle pagine.

Indichiamo con

$$P_{hd} \quad (3)$$

la potenza media erogata al secondo dall'Hard Disk.

Da tutti i parametri indicati si evince che tecnologia e qualità con cui è progettata la parte informatica del sistema eCommerce può avere un ruolo importante nell'utilizzo delle risorse CPU e Disco.

Diremo T_{gc} il tempo di CPU necessario per la generazione del codice dinamico di una pagina e del suo caricamento (dal server al primo apparato di rete). Lo consideriamo composto da un tempo di generazione e caricamento medio ideale T_{gci} che va rapportato con un **tasso di efficienza energetica informatica del sito** η_i . Tanto meglio il sito è progettato, tanto più il tasso di efficienza energetica si avvicina ad 1, tanto peggio è progettato tanto più il valore tende a zero. Quindi

$$T_{gc} = T_{gci} / \eta_i \quad (4)$$

Un sito Web ideale avrà come tasso di efficienza 1 e $T_{gc} = T_{gi}$. Un sito che abbia una pagina con un loop avrà un tasso di efficienza $\eta_i = 0$ e quindi $T_{gc} = INF$.

Analogamente si può ricavare il tempo della risorsa Hard Disk T_{gd} per la generazione di una pagina come dato da un tempo di generazione ideale T_{gdi} rapportato al fattore di qualità del sito. Quindi

$$T_{gd} = T_{gdi} / \eta_i \quad (5)$$

Può essere ragionevole immaginare che un sito eCommerce sviluppato a partire da un prodotto, sia esso commerciale od Open Source, possa avere a priori un tasso di efficienza energetica η_i maggiore di uno sviluppo custom, in quanto soggetto ad un test e ad un utilizzo maggiore.

3.3 Web Design ed usabilità

Che l'emissione di CO₂ possa dipendere da elementi tecnici informatici, risulta abbastanza intuitivo. Tuttavia vi sono altri elementi che hanno impatto. La scelta del design di un sito può avere un ruolo non meno rilevante.

In primo luogo la scelta di immagini ed altre risorse multimediali, della loro dimensione e qualità ha un impatto diretto sull'utilizzo della risorsa Disco Rigido. Quindi quanto più le immagini sono numerose, di elevate dimensioni e di alta qualità tanto più verrà utilizzata la risorsa disco. L'utilizzo di immagini in numero e dimensione adeguata, tra l'altro, è anche oggetto di studi ormai consolidati sull'usabilità dei siti web [Nielsen e Loranger, 2006].

Per indicare qualità e quantità delle risorse multimediali utilizzeremo il parametro T_{md} , inteso come tempo medio del disco necessario a caricare le risorse multimediali di una pagina del sito. Analogamente al caso dell'efficienza energetica informatica del disco T_{gd} , si definisce

$$T_{md} = T_{mdi} / \eta_m \quad (6)$$

dove T_{mdi} è il tempo di caricamento ideale delle risorse e η_m il tasso di efficienza energetica relativo all'utilizzo delle risorse multimediali.

In secondo luogo il progetto della modalità di navigazione, del motore di ricerca, della struttura del sito, insomma di quei parametri che sono parte specifica dell'usabilità di un sito web [Nielsen e Loranger, 2006], hanno un impatto sul numero di pagine che vengono caricate dall'utente. In altre parole quante meno pagine il cliente deve attraversare per trovare il prodotto che cerca e per concludere l'acquisto, tanto minore saranno le pagine caricate e quindi le risorse utilizzate. Sono stati proposti dei modelli per valutare l'usabilità di siti web [Bruno et al, 2009]. Alcuni permettono di dare stime automatiche o semiautomatiche [Mastroianni e Vellutino 2002].

Utilizzeremo il parametro η_{nd} (efficienza di navigazione dovuta ad design) per indicare quantitativamente questo valore. Tale valore ha un impatto sul numero di pagine visitate.

Dalle considerazioni, risulta esistere una relazione tra l'usabilità e quindi la qualità del design del sito eCommerce e la sua prestazione energetica.

3.4 Web Marketing

Con Web Marketing si intendono una serie di attività che hanno lo scopo di portare visite su un sito e di convertirle in obiettivi. Nel caso di un sito eCommerce, quindi, il fine è portare visite e convertirle in vendite.

Le tecniche utilizzate sono varie [Maltraversi, 2013], tra cui il Search Engine Optimization (SEO), campagne di mailing diretto (DEM), PayPerClick, campagne virali sui social network, scambi di banner ed altre ancora.

Il SEO contempla una serie di attività che hanno lo scopo di mostrare il sito web più in alto nella Search Engine Result Page (SERP), ossia la pagina dei risultati del motore di ricerca. Quindi chi effettua una ricerca dovrebbero trovare più facilmente il link al sito e dunque raggiungerlo e navigarlo. Esso si basa su tecniche di ottimizzazione del codice e dei contenuti del sito (tecniche on site); tecniche di miglioramento del PageRank, che è l'algoritmo di Google che assegna un peso a ciascuna pagina in base alla sua notorietà; acquisto di parole chiave sui motori di ricerca.

Il Direct Email Marketing (DEM) consiste nell'invio di email a clienti o potenziali clienti del sito con indicazioni di offerte, promozioni, informazioni, buoni sconto, ecc. al fine di invogliare il cliente a registrarsi al sito od a concludere un acquisto.

Il PayPerClick (PPC) comprende tutti i sistemi, come AdWords di Google, ma anche tutti gli strumenti di tipo "trova prezzo", che inviano traffico sul sito web e che fatturano il singolo click che porta il visitatore al sito.

Le campagne virali sono tutte le attività che intendono riprodurre i benefici del passaparola utilizzando internet ed in particolare i social network.

Nel web marketing vanno considerati anche altri elementi, quali la politica commerciale dell'azienda (sconti, spedizioni e resi gratuiti, ecc.), le ricerche di mercato per il posizionamento dei propri prodotti, le politiche di gestione del cliente, i contenuti del sito, spesso indicazioni per il Web Design, ed altri fattori.

Per valutare l'impatto complessivo delle attività del Web Marketing si prende in considerazione il Tasso di Conversione del sito, che è dato dal rapporto tra le conversioni ottenute, C , ed il numero di pagine visitate, V . Quindi $T_c = C / V$.

Quindi più numerose saranno le visite al sito che producono una vendita, più efficiente viene considerato il lavoro complessivo del Web Marketing.

A questo punto se indichiamo con T_{ci} il tasso di conversione ideale potremmo definire η_{wm} il tasso di efficienza del Web Marketing. Quindi

$$T_c = T_{ci} * \eta_{wm} \quad (7)$$

Il tasso medio di conversione di un sito web è un dato difficile da stimare, perché molto variabile in base a diversi fattori, tra cui il settore merceologico ed il paese. Una rilevazione viene fornita annualmente dalla Forrester Research Inc. Per il 2013 il valore medio rilevato sul campione è del 3% [Mulpuru, 2013].

Questo valore è particolarmente rilevante dato che ha impatto su ogni visita al sito. Con un tasso di conversione del 3%, si ha che il 97% delle visite su un sito eCommerce producono emissioni di CO₂ senza produrre conversione.

3.5 Altri elementi

Altri elementi di interesse per la valutazione della CO₂ emessa da un sito di eCommerce sono quelli logistici.

Considereremo in particolare il trasporto della merce dal deposito del negozio al cliente, che si verifica solo nel caso di una vendita (Conversione).

Il valore dell'emissione dipende dalla logistica del negozio e dalla modalità di trasporto scelta. Indicheremo il valore di CO₂ come una costante *L*, dipendente dal particolare negozio.

4. Emissioni di CO₂ di un sistema eCommerce

Quindi l'emissione di un sistema eCommerce per pagina è data da:

$$Eecp = (Hp + Tgc * Pcpu + Tgd * Phd + Tmd * Phd) * G \quad (8)$$

dove *G* è il fattore di conversione tra Kwh e gas CO₂

Se consideriamo le pagine visitate durante una visita al sito *Ppv* abbiamo che l'emissione totale di una navigazione senza conversione è:

$$Eecnc = Eecp * Ppv \quad (9)$$

dove *Ppv* è il numero di Pagine per Visita (valore reperibile su sistemi di analisi della navigazione quali Google Analytics [Maltraversi, 2013])

Le emissioni totali di una navigazione con conversione *Eecc* sono date da:

$$Eecc = Eecp + L \quad (10)$$

L'emissione totale di un sito eCommerce nell'unità di tempo è data da:

$$Eec(t) = V(t) * ((100-Tct) Eecnc + Tct * Eecc) \quad (11)$$

Dove *V(t)* è il numero di visite nell'unità di tempo indicata. L'unità di tempo dovrà essere sufficientemente ampia da includere il tempo della navigazione del numero medio di pagine per visita.

5. Conclusioni e lavori futuri

Nel contributo sono stati messi in evidenza gli elementi che determinano emissione di CO₂ da parte di un sito eCommerce. Dall'analisi dei fattori si evince che vi sono elementi tecnici hardware e software oltre che elementi logistici che giocano un ruolo rilevante. Il fattore di qualità del software è risultato particolarmente importante per le emissioni dovute alla generazione di una singola pagina. Ma è risultato che vi sono anche altri elementi non strettamente tecnici, il web marketing ed il web design, che giocano un ruolo importante nell'emissione della CO₂. In generale appare che la qualità del software e dei servizi di Web Design e di Web Marketing sia legata alla prestazione energetica complessiva di un sito eCommerce B2C.

Il presente lavoro potrà essere sviluppato in diverse direzioni. Da una parte approfondendo e cercando di schematizzare i parametri di qualità relativi al software ed al web design. In particolare potranno essere effettuate delle misurazioni su diversi sistemi, con particolare riferimento a quelli Open Source vista la disponibilità del codice sorgente, per stabilire dei valori di riferimento, anche utilizzando metodologie di soft computing [Rampone, 2013].

Dall'altra parte potrà essere interessante rendere più granulare il modello inserendo delle variabili trascurate in prima battuta, quali l'effetto dei crawler dei motori di ricerca sul numero di pagine visitate.

Infine potrà essere utile sviluppare la relazione tra il tasso di conversione del sito ed il concetto di emissione di CO₂ senza corrispondente conversione.

Bibliografia

[Ademe, 2011] Lardillon C., Ménard C., Minier N., Communiqué de Presse 7-7-2011 - Courriers électroniques, requête Web, clé USB: quels impacts environnementaux?, Ademe (Agence de l'environnement et de la maîtrise de l'énergie), Angers Cedex, France, 2011

[Atkinson, 2012] Atkinson A., How Much CO₂ is created by..., General Electric Company, <http://visualization.geblogs.com/visualization/co2/>, 2012

[Bruno et al., 2009] Bruno G., Esposito E., Mastroianni M., Evaluation of Public E-Procurement Services Accessibility: A Multicriteria Approach in International Handbook of Public Procurement, Auerbach Publication Taylor & Francis Group, USA, 2009

[Chan et al. 2012] Chan C. A. , Gyax A. F. , Wong E., Leckie C. A., Nirmalathas A., Kilper D. C., Methodologies for Assessing the Use-Phase Power Consumption and Greenhouse Gas Emissions of Telecommunications Network Services. Environmental Science & Technology, 2013, 47 (1), pp 485–492

[DEFRA 2012] Department for Environment Food and Rural Affairs, 2012 Greenhouse Gas Conversion Factors for Company Reporting, <http://www.defra.gov.uk/publications/2012/05/30/pb13773-2012-ghg-conversion/>, 2012

[Google 2012] Google Developer Team, Google Developer Optimize caching. Upd. 28/03/2012, <https://developers.google.com/speed/docs/best-practices/caching>, 2012

[Goraczko et al, 2012] Goraczko M., Kansal A., Liu J., Zhao F., Joulemeter: Computational Energy Measurement and Optimization, Microsoft Research <http://research.microsoft.com/en-us/projects/joulemeter/>, 2012

[Gunaratne et al, 2005] Gunaratne C., Christensen K., Nordman B., Managing energy consumption costs in desktop PCs and LAN switches with proxying, split TCP connections, and scaling of link speed Int. J. Network Manage. 2005, 15 (5) 297– 310

[Hanle e Richter-Shalaby, 2010] Hanle H., Richter-Shalaby B., White Paper Green ICT, T-systems International GmbH, Frankfurt, Germany, 2010

[Hölzle, 2009] Hölzle U., Powering a Google search, Google Official Blog <http://googleblog.blogspot.it/2009/01/powering-google-search.html>, 2009

[IEA, 2012] CO₂ Emissions from Fuel Combustion 2012, International Energy Agency. CO₂ Emissions from Fuel Combustion, <http://www.iea.org/publications/freepublications/publication/name,4010,en.html>, 2012

[Kansal 2008] Aman Kansal, Feng Zhao, Fine-grained energy profiling for power-aware application design in ACM SIGMETRICS Performance Evaluation Review, Volume 36 Issue 2, September 2008 Pages 26-31 ACM New York, NY, USA

[Kiplier et al. 2011] Kilper, D. C.; Atkinson, G.; Korotky, S. K.; Goyal, S.; Vetter, P.; Suvakovic, D.; Blume, O. Power Trends in Communication Networks IEEE J. Sel. Topics Quantum Electron. 2011, 17 (2) 275– 284

[Kooimey, 2007] Kooimey J.G., Estimating total power consumption by servers in the U.S. and the world. Technical report, Lawrence Berkeley National Laboratory, USA, 2007

[Macrini 2010] Macrini, D., L'effetto coperta. <http://www.climatemonitor.it>, 2010.

[Mahadevan 2009] Mahadevan, P.; Sharma, P.; Banerjee, S.; Ranganathan, P. A Power Benchmarking Framework for Network Devices. In Proceedings of the 2009 International IFIP-TC 6 Networking Conference, Berlin, Germany, 2009

[Maltraversi 2013] Maltraversi M., SEO e SEM. Guida avanzata al Web marketing 2a edizione Edizioni FAG Milano 2013

[Mastroianni, Vellutino 2002] Mastroianni M., Vellutino D., DA.MA: A Model for Automatic Evaluation of Public Sector Web Sites in Proceedings of the 2nd European Conference on eGovernment, 2002

[Mulpuru, 2013] Mulpuru S., The State Of Retailing Online 2013: Key Metrics And Initiatives Forrester research inc., 60 Acorn Park Drive, Cambridge, Ma 02140 Usa, 2013

[Nielsen, Loranger, 2006] J. Nielsen, Hoa Loranger Web Usability 2.0. L'usabilità che conta, Apogeo, Milano, 2006

[Petty, 2007] Petty C., Gartner Estimates ICT Industry Accounts for 2% of Global CO₂ Emissions, Gartner Press Rel., <http://www.gartner.com/it/page.jsp?id=503867>, 2007

[Rampone, 2013] S. Rampone, Three-and-six-month-before forecast of water resources in a karst aquifer in the Terminio massif (Southern Italy). Applied Soft Computing, 2013, 13(10), pp 4077–4086.

[Rangone et al, 2012] Rangone A., Liscia R., Perego A. Mangiaracina R., Osservatorio eCommerce B2C 2012: I Consumi 2012 in Italia: -2% offline, +18% online... ma la partita è multicanale, School of Management Politecnico di Milano, Dipartimento di Ingegneria Gestionale, 2012

[SPEC, 2012] SPECpower_ssj2008 benchmark, Standard performance evaluation corporation, http://www.spec.org/power_ssj2008/index.html, 2012

[Terna, 2013] Rapporto mensile sul sistema elettrico consuntivo maggio 2013, Terna Rete Italia, <http://www.terna.it/LinkClick.aspx?fileticket=109432>, 2013

Un sistema di supporto al posizionamento e al dimensionamento dei punti di accumulo dei rifiuti urbani domestici

Paolino Di Felice, Leonardo D'Errico¹, Fabio Franchi¹
Dipartimento di Ingegneria Industriale e dell'Informazione e Economia
Università di L'Aquila
Via Giovanni Gronchi n. 18 – Nucleo Ind.le di Pile, 67100 L'AQUILA
paolino.difelice@univaq.it

¹BiTE s.r.l.
via F. Paolo Tosti, 15, 67100 L'Aquila
leonardo@bitesrl.it, fabio@bitesrl.it

Abstract. *The paper describes the potentiality of a method for solving the problem of locating and sizing the garbage bins for the separate accumulation of the household litter in urban areas. The solution, centered around a spatial database that integrates both geographic and descriptive data, has been thought as a supporting tool of the figures responsible for the management of the municipal solid waste to which it offers a dual-mode display of the results: one tabular (the standard format featured by relational databases) and another based on the metaphor of geographic maps, the latter particularly useful to highlight the spatial component of the problem.*

Keywords: household waste, garbage accumulation point, garbage bin, spatial database.

1. Introduzione

"Managing solid waste well and affordably is one of the key challenges of the 21st century, and one of the key responsibilities of a city government" [UN-HABITAT, 2010]. Una classificazione usuale dei Rifiuti Solidi Urbani consiste nel suddividerli in residenziali e commerciali. Il presente contributo fa esplicito riferimento ai RSU residenziali, perché essi sono la categoria di maggiore rilevanza quantitativa, [EEA, 2013].

Il primo passo nella catena di gestione della raccolta differenziata di RSU inizia allorché il cittadino deve liberarsi dei rifiuti domestici da lui stesso prodotti. Ad oggi quel che si registra passeggiando per le vie delle città italiane è la coesistenza di due forme alternative, oltre che complementari, di accumulo dei RSU: l'accumulo in bidoni di modesto volume forniti in dotazione alle singole famiglie e l'accumulo in cassonetti di grosso volume (oltre 1m³) dislocati lungo

le strade pubbliche urbane. L'accumulo differenziato dei RSU nei cassonetti pone l'ovvio problema della loro collocazione, mentre, il problema del dimensionamento del numero dei cassonetti/bidoni riguarda entrambe le forme di accumulo. Trattasi di due questioni non di lana caprina, che se sottovalutate fanno insorgere problemi di ordine pratico [Di Felice, 2013a].

Il seguito del presente articolo è organizzato come segue. La Sez. 2 fa un cenno allo stato dell'arte nel settore, mentre la Sez. 3 richiama un metodo di soluzione del problema della collocazione lungo strade pubbliche dei cassonetti per l'accumulo differenziato dei RSU e loro dimensionamento basato su due algoritmi da eseguire in sequenza apparsi molto recentemente [Di Felice, 2013a]. La Sez. 4 evidenzia come, a partire da tale metodo, sia possibile erogare un servizio informatico di supporto ai responsabili della gestione dell'accumulo e dello smaltimento dei RSU che sia *completo e flessibile*. Un'ampia sezione (5) conclude il lavoro.

Le ultime due sezioni di questo lavoro costituiscono il suo valore aggiunto rispetto a quanto già pubblicato dallo stesso autore. Esse chiariscono che algoritmi come quelli descritti in [Di Felice, 2013a] sono l'ossatura portante di un sistema di supporto alle decisioni da usare insieme con un webGIS in modo da restituire agli utenti del sistema delle mappe circa l'area da servire dalle quali sia agevole esaminare la soluzione individuata dal sistema software e, se ritenuto necessario, effettuare delle correzioni mirate.

2. Stato dell'arte

Per lungo tempo la comunità scientifica si è disinteressata del problema della dislocazione delle aree di accumulo dei RSU e del loro dimensionamento. I pochissimi lavori che hanno studiato tale problema, a partire dal 2006, sono citati in [Ghiani et al, 2012]. L'obiettivo di tali ricerche è stato minimizzare il numero delle aree di accumulo dei RSU con l'intento di ridurre i costi di acquisto dei cassonetti nonché i costi di raccolta dei RSU da parte degli operai comunali, oltre a limitare l'impatto negativo causato dalla presenza, a ridosso delle abitazioni, di tali aree di nessuna bellezza estetica.

Lo studio condotto da [Ghiani et al, 2012] si prefigge l'ulteriore obiettivo di accrescere la qualità del servizio di accumulo e raccolta dei RSU misurata come distanza massima tra ciascuna abitazione e l'area di accumulo ad essa più vicina. Tale servizio è percepito importante dai cittadini che pagano tasse comunali sempre più alte per sostenerlo e, conseguentemente, è inevitabile che pretendano una adeguata contropartita, ed inoltre per la crescente sensibilità civica verso la salvaguardia dell'ambiente.

Il contributo in [Di Felice, 2013a] si salda idealmente a questo lavoro, con due sostanziali differenze: la raccolta dei RSU è differenziata, inoltre, mentre costoro ipotizzano di conoscere, quale dato di partenza, un certo numero di aree urbane dove accumulare i RSU, nel nostro approccio tali aree possono essere localizzate in un qualunque punto lungo le strade cittadine e la loro posizione viene decisa autonomamente dal metodo risolutivo in base alla "geografia urbana" delle abitazioni da servire.

Sistema di supporto al posizionamento e al dimensionamento dei punti di accumulo dei rifiuti urbani

Evidentemente, il problema del posizionamento e dimensionamento dei cassonetti nelle aree urbane è di primario interesse per le amministrazioni comunali che però, ad oggi, sono sprovviste di soluzioni informatiche “robuste” per il ritardo con il quale la comunità scientifica si è interessata al problema. Quello cui oggi si osserva frugando nella rete è la proposta di soluzioni “miracolose” di questo problema.

Una testimonianza pertinente al riguardo è reperibile accedendo il sito www.rifiutilab.it della Labelab srl, dove si legge che trattasi del primo portale nazionale dedicato alla gestione dei rifiuti. Tra i servizi erogati dalla Labelab srl attraverso questo portale c'è quello per il dimensionamento del numero complessivo dei cassonetti in una certa area urbana per le varie tipologie di RSU (quali, ad esempio: vetro, carta, plastica, ...). Ad esempio la Fig. 1 fa riferimento all'indifferenziato. Dall'esame del portale non si trova traccia se esso fornisca ai gestori del servizio di raccolta e smaltimento dei RSU indicazioni sul posizionamento dei cassonetti lungo le strade cittadine, né tantomeno come tale problema venga affrontato dal punto di vista metodologico. Evidentemente trattasi di carenze gravi.

Parametro	Valore
Categoria	Raccolta stradale
Servizio raccolta	Indifferenziato
Popolazione (num. abitanti)	20000
Produzione specifica (kg/abit/giorno)	1
Peso specifico rifiuto (kg/m³)	80
Tipologia contenitore	bidone o cassonetto
Volume contenitore (l)	2400
Frequenza del servizio (gg/gg)	3 / 7
Massimo grado di riempimento (%)	80
Produttività squadra (contenitori/turno)	90
Autisti per squadra (num)	1
Operatori per squadra (num)	0
Turni giornalieri di utilizzo mezzo (num)	1
Tipologia mezzo	autocompattatore monopoperatore a carico laterale
Capacità cassone mezzo (m³)	24
Rapporto di compressione (x:1)	5
Esegui dimensionamento servizi	

Fig. 1. La pagina che consente all'utente del servizio di dimensionare il numero dei cassonetti per la raccolta differenziata dei RSU.

3. Una recente proposta di soluzione del problema dell'accumulo dei RSU domestici

Il problema dell'accumulo dei RSU nelle moderne città può essere formulato come segue: dato l'insieme delle case e delle strade pubbliche (*i dati spaziali*) che fanno parte di un qualche comune o una sua porzione, l'obiettivo è calcolare la posizione dei punti di accumulo dei RSU (Garbage Accumulation Points, nel seguito) nell'area, oltre a dimensionarne il numero dei cassonetti

(Garbage Bins, nel seguito). Ulteriori dati rilevanti del problema (i *dati descrittivi*) hanno a che fare con la categoria di RSU da depositare nei GAPs, la capacità dei GBs, la loro frequenza di svuotamento, il numero dei residenti in ciascuna abitazione del comune da servire, e la produzione pro capite giornaliera di immondizia.

Di seguito si elencano i dati del problema mantenendo le notazioni in [Di Felice, 2013a] per agevolare coloro che decidessero di leggere tale lavoro:

Comun_q denota il comune, o una sua frazione, da servire con i GAPs,

\mathbb{R} è l'insieme delle strade pubbliche che attraversano il Comun_q ,

\mathbb{H} è l'insieme delle abitazioni nel Comun_q , ove per abitazione s'intende un intero edificio. Nello studio ci si disinteressa del numero di piani di cui si compongono le abitazioni, ma non del numero degli occupanti,

\mathbb{GB} è l'insieme dei diversi tipi di GBs che compongono ogni GAP. Nello studio sono stati presi in considerazione i seguenti cinque tipi: *glass*, *plastic*, *paper*, *organic*, and *unsorted*, ma il metodo risolutivo è generale, pertanto esso può essere adattato ad operare con ipotesi diverse,

\mathbb{GAP} è l'insieme dei GAPs che devono essere dislocati nel Comun_q ,

dailyGlass, *dailyPlastic*, *dailyPaper*, *dailyOrganic* e *dailyUnsorted* denotano, in sequenza, la produzione pro capite giornaliera (in m^3) dei cinque differenti tipi di RSU cui si fa riferimento,

distance: denota il valore della distanza massima tra ogni abitazione ed il GAP ad essa più vicino.

In [Di Felice, 2013a], il problema dell'accumulo dei RSU è risolto nel rispetto dei seguenti vincoli:

- i GAPs devono essere posizionati su strade pubbliche,
- ogni abitazione deve avere ad una distanza (misurata lungo le strade pubbliche) non superiore a *distance* almeno un GAP,
- il numero dei cassonetti dei diversi tipi di rifiuti in ciascun GAP deve essere dimensionato tenendo conto della produzione pro capite potenziale giornaliera di RSU da parte dei residenti nel Comun_q in modo da prevenire il riempimento dei GBs prima che essi siano svuotati dagli operai comunali.

La soluzione del problema la si ottiene tramite gli algoritmi **LocationOfGAPs** e **SizingOfGAPs** da eseguire in sequenza, dei quali di seguito si specifica l'input e l'output.

Algoritmo **LocationOfGAPs**

Input: \mathbb{H} , \mathbb{R} , *distance*

Output: \mathbb{GAP}

Algoritmo **SizingOfGAPs**

Input: \mathbb{H} , \mathbb{GB} , \mathbb{GAP} , *dailyGlass*, *dailyPlastic*, *dailyPaper*, *dailyOrganic* e *dailyUnsorted*

Output: \mathbb{GAP} aggiornato

Sistema di supporto al posizionamento e al dimensionamento dei punti di accumulo dei rifiuti urbani

Il primo algoritmo individua il numero di GAPs da dislocare nel Comune e la loro posizione geografica, mentre il secondo, per ogni GAP, calcola il numero dei GBs delle cinque tipologie prese in esame che ne devono far parte.

L'implementazione dei due algoritmi è stata compiuta in due passi, [Di Felice, 2013b]. Il primo è consistito nella progettazione di una BD Spaziale (BDS nel seguito) e sua successiva realizzazione in PostgreSQL/PostGIS, seguita dal caricamento in essa dei dati spaziali e descrittivi pertinenti. Quindi (passo 2), si sono codificati (nel linguaggio PL/pgSQL) i due algoritmi in termini di funzioni utente.

I meriti principali dell'architettura software dell'implementazione realizzata sono due: a) essa si avvale della tecnologia dei DBMS spaziali la quale consente di sfruttare l'espressività del linguaggio SQL nell'interrogare la BDS facendo ricorso, quando necessario, a qualcuna delle funzioni utente progettate o di quelle di sistema per accrescerne ulteriormente l'espressività, b) fa uso esclusivo di tecnologia software open source, scelta quasi obbligata per le amministrazioni municipali cronicamente con i bilanci in rosso.

4. Verso un sistema di supporto completo e flessibile

Il supporto informatico che può agevolare il lavoro dei responsabili della gestione dei RSU investe, innanzitutto, le due fasi del posizionamento dei GPAs ed il loro dimensionamento (*completezza*); quindi, occorre consentire loro di compiere delle indagini a 360 gradi per maturare piena consapevolezza sulle scelte da effettuare, con visualizzazione dei risultati tanto in forma tabellare che di mappe tematiche (*flessibilità*). La soluzione descritta in questo lavoro soddisfa entrambi i requisiti. Circa la completezza non c'è bisogno di spendere altre parole, pertanto nel seguito ci si concentra sulla flessibilità.

Aver realizzato la BDS e averla corredata di viste e funzioni utente ad hoc rende agevole condurre elaborazioni di svariati tipi. La più naturale delle indagini consiste nell'eseguire i due algoritmi di sopra per diversi valori del vincolo di distanza massima abitazione-GAP al variare del metodo di visita delle strade pubbliche presenti nel comune di interesse. Ad esempio, la Tab. 1 riunisce i risultati di una campagna di esperimenti riferiti alla frazione di Cansatessa della città di L'Aquila, nell'ipotesi che le strade pubbliche siano visitate in ordine decrescente (seconda colonna) e crescente di estensione.

Tab. 1. I risultati della campagna di esperimenti.

<i>distance</i> (m)	Strade visitate in ordine decrescente	Strade visitate in ordine crescente
50	43	43
100	28	29
150	22	23
200	15	19
250	12	16
300	8	15
500	6	6

La Fig. 2, invece, mostra la posizione geografica e il numero dei GAPs per *distance* = 250m e *distance* = 300m.

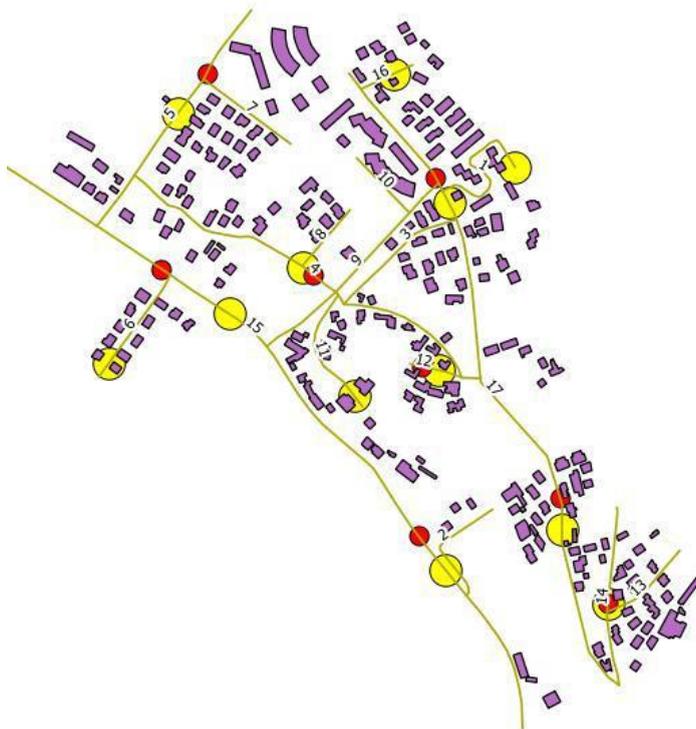


Fig. 2. I (12) cerchi gialli e gli (8) rossi denotano i GAPs per *distance* = 250m e *distance* = 300m, nell'ipotesi che le strade pubbliche siano visitate in ordine decrescente di estensione. I numeri sono gli identificatori univoci delle strade pubbliche di Cansatessa. La mappa è stata disegnata avvalendosi del software open source denominato Quantum GIS. Per ulteriori informazioni sui dati dell'esempio si rimanda a [Di Felice, 2013b].

Dall'esame congiunto della Tab. 1 e della Fig. 2 è possibile desumere che la migliore strategia di attraversamento delle strade pubbliche nel posizionare i GAPs debba essere la prima delle due perché ad essa corrisponde *sempre* il numero di GAPs più basso, ed inoltre perché essa nel posizionare i GAPs predilige le strade di maggiore estensione che, verosimilmente, sono anche quelle che offrono la migliore viabilità, aspetto, quest'ultimo, importante per le operazioni di svuotamento dei cassonetti da parte degli automezzi municipali.

La Tab. 2 mostra che la situazione *distance* = 300m è la più favorevole da questo punto di vista, essa, infatti, posiziona 5 degli 8 GAPs sulle tre strade pubbliche più lunghe di Cansatessa. Il posizionamento dei GAPs che si ottiene per *distance* = 250m (Fig. 2) è meno soddisfacente.

Altri dati che i gestori del servizio di accumulo dei RSU devono conoscere sono il numero dei cassonetti di cui si compone ciascun GAP. Questi dati, al termine dell'esecuzione dell'algoritmo **SizingOfGAPs**, sono depositati in una delle tabelle della BDS. Per portare tali dati in visualizzazione (Tab. 3) sarà sufficiente interrogare la BDS.

Tab. 2. id ed estensione delle strade dove andrebbero dislocati i GAPs per *distance* = 300m.

Id del GAP	Id della strada	Lunghezza della strada (m)
1	15	1493
2	15	1493
3	17	1046
4	17	1046
5	4	632
6	5	400
7	14	270
8	12	80

Tab. 3. La composizione dei GAPs al variare di *distance*.

<i>distance</i> (m)	GAP	Organic	Plastic	Paper	Glass	Unsorted	Totale
50	43	51	43	43	43	43	223
100	28	37	29	28	28	29	151
150	22	36	24	23	23	24	130
200	15	34	21	17	18	21	111
250	12	32	21	15	16	20	104
300	8	29	16	13	13	16	87

Segue un elenco non esaustivo di ulteriori elaborazioni che la soluzione proposta basata sull'uso di una BDS rende agevoli: a) calcolo della distanza minima media abitazione-GAP, b) calcolo del numero di abitazioni/cittadini servite/i da un GAP, c) visualizzazione della dislocazione geografica delle abitazioni servite da un GAP, d) visualizzazione di id e posizione dei GAPs che giacciono lungo una certa strada, e) calcolo del numero medio di abitazioni/cittadini servite/i da ciascun GAP entro un dato valore di *distance* (es. 250m), f) calcolo del numero medio di GBs per GAP, g) calcolo dell'area media di ingombro a terra dei GAPs, h) ...

5. Conclusioni

La soluzione richiamata in questo articolo può essere fruita dagli utilizzatori finali con un training on the job di qualche ora. Infatti, per effettuare le elaborazioni circa la localizzazione e dimensionamento dei GAPs al variare del valore del parametro *distance* è sufficiente invocare, dalla finestra di querying di PostgreSQL, il comando:

```
SELECT solve(distance) ;
```

dove `solve()` è la funzione utente che riunisce il codice PL/pgSQL che implementa i due algoritmi che insieme risolvono il problema. Analogamente, l'elaborazione di una qualsiasi delle queries fornite a corredo della BDS realizzata si effettua copiandola nella medesima finestra.

Tra i punti di forza del metodo richiamato in questo articolo c'è l'opportunità che esso offre di poter decidere il miglior trade-off circa la massima distanza abitazione-GAP, aspetto, quest'ultimo, che impatta sul numero complessivo dei GBs da impiegare (nel caso dei valori in Tab. 3, si registra uno scarto di ben 136 cassonetti tra i casi estremi - *distance* = 50m e *distance* = 500m) e, di conseguenza, sui tempi necessari per il loro svuotamento e, quindi, in ultima

analisi, sui costi che il gestore deve sostenere nell'avviare il servizio (a puro titolo informativo, si segnala che il costo di un cassonetto da 1m³ è superiore ai 100 euro - <http://www.greco-ecology.it/>), ma anche durante il suo esercizio. Evidentemente, la scelta tra le alternative che il metodo di calcolo restituisce rimane prerogativa esclusiva di coloro che hanno la responsabilità della gestione del servizio di raccolta dei RSU, ai quali il metodo proposto e la sua realizzazione su tecnologia open source intende offrire supporto.

Per invogliare i responsabili della gestione dei RSU ad usare il metodo richiamato in questo articolo, occorre farsi carico di un ulteriore sforzo di ingegnerizzazione consistente nello sviluppare un webGIS tramite il quale essi potranno invocare i servizi offerti loro, ed implementati interrogando la BDS, in condizioni di totale indipendenza dagli elementi architetturelli/tecnologici sottostanti. La Fig. 3 mostra il mockup della pagina web base destinata alla raccolta dei dati di input agli algoritmi di posizionamento e dimensionamento dei GBs nelle strade pubbliche di un qualsiasi comune italiano.

Fig. 3. La pagina base per l'introduzione dei dati di ingresso.

Pigiando il bottone “Solve WAP” si scatena l'esecuzione dei due algoritmi. L'utente potrà interagire con i risultati restituiti tramite la schermata di Fig. 4. In particolare, sarà per loro possibile conoscere la dislocazione dei GAPs sulla mappa dell'area di riferimento ed il numero dei GBs delle diverse tipologie di rifiuto che ne fanno parte.

Il sistema renderà fruibile un ampio repertorio di dati, ricavati a partire dai risultati immagazzinati nelle tabelle della BDS, quali, ad esempio: il numero totale dei GAPs da posizionare sul territorio urbano, il numero di GBs per ciascun GAP, la distanza media tra abitazione e GAP ad essa più vicino, ecc...

Sistema di supporto al posizionamento e al dimensionamento dei punti di accumulo dei rifiuti urbani
Lo sviluppo del sito è in corso presso la BiTE s.r.l. (<http://www.bitesrl.it>). Il proposito è rendere disponibile gratuitamente tale servizio a tutte le municipalità che vorranno servirsene, previa registrazione.

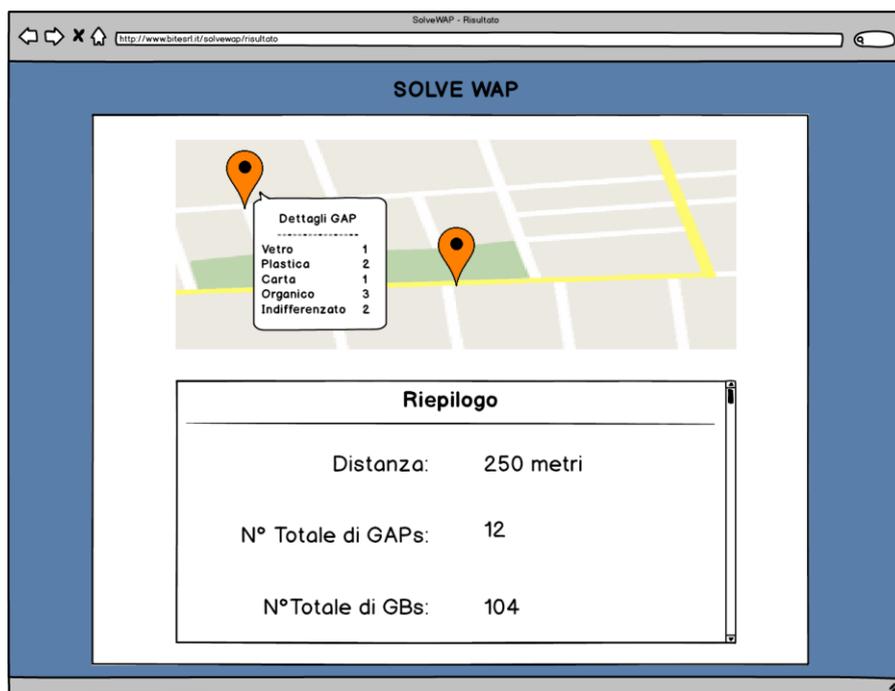


Fig. 4. Una possibile vista di dettaglio sui risultati del posizionamento e del dimensionamento dei GAPs.

5.1 Una nota di cautela

La rispondenza nel reale delle stime restituite da metodi di calcolo come il nostro dipende drammaticamente dai dati in input che esso riceve. Nello specifico, l'algoritmo di posizionamento dei GAPs dipende dalla disponibilità e dalla qualità dei dati territoriali, vale a dire le strade pubbliche del comune oggetto dello studio e le abitazioni che in esso insistono. La nostra piccola esperienza insegna che questi dati esistono presso le regioni e sono di buona qualità.

Circa i dati descrittivi necessari ad alimentare l'algoritmo di dimensionamento del numero dei cassonetti, viceversa, la situazione è molto più delicata. I dati in questione riguardano il numero degli occupanti le abitazioni censite nel comune e la loro produzione pro capite giornaliera di RSU delle diverse categorie. Disporre di dati affidabili a questo livello di dettaglio è arduo perché, come affermato in recentissimi studi di settore [Lebersorger e Beigl, 2011], tali valori dipendono da molte variabili, tra le quali il periodo dell'anno, il clima, il reddito delle famiglie, la dimensione delle loro abitazioni, il tipo di sistema di riscaldamento degli appartamenti, l'ammontare della tassa comunale pro capite da pagare per avere il servizio di raccolta dei RSU, etc.

Un'ulteriore criticità nell'acquisizione di dati attendibili da dare in pasto al secondo algoritmo risiede nella loro volatilità nel tempo. Infatti, quello cui si assiste è che, a causa anche di pressioni legislative crescenti che esistono in tutti i paesi della UE, le quote della raccolta differenziata dei RSU si accrescono di anno in anno. Per contro i dati a disposizione risalgono a diversi anni addietro. Ad esempio l'ultimo studio in ambito regione Abruzzo [BURA, 2011] riporta dati risalenti al 2009.

Ultima difficoltà pratica, ma non certo per importanza, discende dal fatto che non esiste un rapporto di trasformazione univoco tra peso dei rifiuti solidi (espresso in Kg) e loro volume (espresso in m³), mentre è proprio di questo dato che si ha bisogno per risolvere correttamente il problema discusso in questo lavoro (una conferma la fornisce la presenza del campo "Peso specifico rifiuto" nella Fig. 1) stante che la capienza dei cassonetti è espressa in m³. Né esistono degli studi al riguardo (per quanto ne sappiamo) da cui poter attingere. La rilevanza di questo aspetto è sostenuta anche in [Hoorweg e Bhada-Tata, 2012] dove si legge: "Although waste composition is usually provided by weight, waste volumes tend to be more important, especially with regard to accumulation".

6. Riferimenti bibliografici

[Di Felice, 2013a] Di Felice, P., Integration of spatial and descriptive information to solve the urban waste accumulation problem. Proc. of the 3rd International Conference on Integrated Information, Prague, Czech Republic, 2013a.

[Di Felice, 2013b] Di Felice, P. Integration of spatial and descriptive information to solve the urban waste accumulation problem: a pilot study. Proc. of the 3rd International Conference on Integrated Information, Prague, Czech Republic, 2013b.

[EEA, 2013] Managing municipal solid waste. A review of achievements in 32 European countries, Report by the European Environment Agency, 2013. <http://www.eea.europa.eu/publications/managing-municipal-solid-waste> (Recuperato in Giugno, 2013).

[Ghiani et al, 2012] Ghiani, G., Laganà, D., Manni, E. & Triki, C., Capacitated location of collection sites in an urban waste management system. Waste Management, 32, 2012, 1291–1296.

[Hoorweg e Bhada-Tata, 2012] Hoorweg, D. & Bhada-Tata, P., What a waste: A Global Review of Solid Waste Management. Urban Development & Local Government Unit, World Bank, Washington, DC 20433 USA, 2012. www.worldbank.org/urban.

[Lebersorger e Beigl, 2011] Lebersorger, S. & Beigl, P., Municipal solid waste generation in municipalities: Quantifying impacts of household structure, commercial waste and domestic fuel. Waste Management, 31, 2011, 1907–1915.

[BURA, 2011] Bollettino Ufficiale della Regione Abruzzo - Speciale Ambiente n.25, 15 Aprile 2011. http://bura.regione.abruzzo.it/2011/Speciale_25_15_04.pdf (Recuperato in Giugno, 2013).

[UN-Habitat, 2010] Solid waste management in the world's cities. United Nations Human Settlements Programme, 2010.

IL CONTRIBUTO DEL SETTORE FOTOVOLTAICO ALLA REALIZZAZIONE DI MODELLI ENERGETICI SMART

Galdi Vincenzo¹, Testa Mario²

¹DIIN, Facoltà di Ingegneria, Università degli Studi di Salerno
Via Giovanni Paolo II, 132, 84084, Fisciano (SA)
vgaldi@unisa.it

²DIIN, Facoltà di Ingegneria, Università degli Studi di Salerno
Via Giovanni Paolo II, 132, 84084, Fisciano (SA)
mtesta@unisa.it

Abstract. *In-depth analyses, core to current energy paradigms both at global and local scale, of the links between energy options, economic growth, protection of the environment and social consequences are an integral part of evaluating eventual impacts. The response to such a challenge can be glimpsed in the prospect of innovative development declined in what is commonly called the green economy. The present study aims to analyse the main implications of a technical, economic and environmental nature associated both to the production of photo-voltaic energy, and Smart Grid development.*

Keywords Smart Grid; Photovoltaic; Green Economy.

1. Background

L'orientamento della politica energetica europea verso la sostenibilità, intesa in termini di competitività economica, di tutela ambientale, di sicurezza degli approvvigionamenti e di equità sociale, unitamente al progressivo ridimensionamento, se non abbandono, del nucleare in molti paesi, ha impresso profondi cambiamenti negli scenari energetici, peraltro dominati tuttora dalle fonti fossili. In tale contesto, pertanto, le fonti di energia rinnovabili (FER) stanno acquisendo un ruolo sempre più significativo, quale strumento fondamentale della governance energetica europea, in linea con gli obiettivi del "Pacchetto Clima-Energia", ed in particolare della Dir 2009/28/CE, che fissa specifici target relativi alle fonti rinnovabili, il cui contributo ai consumi finali lordi di energia (media Ue) dovrà essere di almeno il 20%, entro il 2020, oltre ad una strategia di riduzione delle emissioni climalteranti del 20% (rispetto ai dati del 2005) e un miglioramento dell'efficienza energetica del 20%.

Nell'ambito del Piano di Azione Nazionale (PAN) per le Energie Rinnovabili, il Governo Italiano ha previsto, tra l'altro, che le FER soddisfino, entro il 2020, il 26,39% dei consumi elettrici e a tal fine è stato emanato il D.Lgs. 28 del 3/3/2011 che, oltre a ridefinire in maniera organica il quadro istituzionale e

finanziario, prevede l'adozione di una serie di strumenti e di meccanismi nonché di incentivi di natura finanziaria e fiscale a sostegno dell'energia rinnovabile.

Il lavoro evidenzia come la generazione elettrica da FER richiede lo sviluppo di Smart Grid e la diffusione di dispositivi elettronici per ottimizzare l'efficienza, la sicurezza e la flessibilità delle reti elettriche di trasmissione, affinché siano in grado di consentire ingenti scambi di dati tra i molteplici dispositivi intelligenti distribuiti nel sistema elettrico. A tal fine il lavoro parte dall'analisi del ruolo del settore fotovoltaico (FV), la cui produzione negli ultimi anni ha registrato notevoli tassi di crescita a livello nazionale, europeo e mondiale, per poi suggerire alcune metodologie, ormai in parte standardizzate dall'Iso, per elaborare corrette valutazioni in ordine agli impatti economici, ambientali e sociali associati alla filiera elettrica.

2. Smart Energy System: il ruolo del settore fotovoltaico

La diffusione di tecnologie energetiche basate su risorse rinnovabili richiede, in linea di massima, investimenti in capacità di generazione distribuita sul territorio, attraverso l'installazione di impianti di piccola e media taglia, impattando notevolmente sul sistema di trasmissione, in quanto tale generazione non può essere aumentata, nel caso di maggiore richiesta di elettricità, e quindi non è "dispacciabile". Pertanto, da un lato, occorre una profonda riorganizzazione di tipo strutturale del sistema elettrico, attraverso il mantenimento in esercizio di impianti di generazione programmabili (ad esempio, impianti termoelettrici, idroelettrici, etc.), ovvero di unità generative flessibili, in grado di compensare il fenomeno della intermittenza nella generazione elettrica da FER; dall'altro, appare necessaria una riorganizzazione di natura gestionale delle reti di trasmissione per rendere più resiliente l'intero sistema alla loro "variabilità". Tali cambiamenti, oltre al miglioramento della capacità di interconnessione, richiedono anche innovazioni nei sistemi di accumulo (*storage*) di energia elettrica, il ripensamento del *demand-side management* nonché la predisposizione di impianti di generazione programmabili [OECD/IAE, 2011]. Inoltre, l'affermazione delle FER segna la transizione da un modello di produzione centralizzata ad un altro, ben più articolato e complesso sia dal punto di vista gestionale che tecnologico, di generazione distribuita, che implica innovazioni nella rete di trasmissione e distribuzione che si evolve da rete passiva, connotata da flussi unidirezionali provenienti dai grandi impianti di produzione fino ai singoli consumatori, a *smart grid*, ovvero una rete attiva e intelligente. Essa, infatti, è in grado di gestire un sistema complesso di flussi elettrici discontinui e bidirezionali, provenienti da un fitto *network* di produttori/consumatori (*prosumer*).

3. Le sfide attuali e future del settore FV italiano

Recenti ricerche evidenziano [Bloomberg New Energy Finance, 2012] che la produzione di energia fotovoltaica (FV) potrebbe divenire competitivo rispetto alle fonti energetiche tradizionali nell'arco di pochi anni in vari paesi, anche se

con dinamiche differenti. Tra i molteplici fattori che contribuiscono maggiormente al raggiungimento della *grid parity* (ovvero la coincidenza tra costo del kWh FV rispetto al kWh ottenuto da fonti fossili) alcuni sono riconducibili ai costi crescenti dell'elettricità prodotta da impianti convenzionali e altri alla forte diminuzione del prezzo dei moduli FV, ridottosi negli ultimi tre anni di circa il 75%, per effetto della riduzione dei costi di produzione. Alcune stime indicano, per l'Italia, il 2014 quale anno in cui si raggiungerà la competitività dell'energia elettrica fotovoltaica rispetto a quella termoelettrica (*grid parity*), in seguito al calo dei prezzi che nel 2011 è stato di circa il 40% per i moduli FV tradizionali. Ciò in riferimento ad impianti di 200-400 kW [Cavallin F. et al., 2011]. Peraltro, al fine di realizzare nuovi paradigmi energetici, capaci di integrarsi in contesti caratterizzati da processi tesi ad armonizzare le problematiche energetiche con quelle socio-ambientali, occorre fin da ora varare corrette strategie, non solo per varare opzioni tecnologiche a favore di infrastrutture per la realizzazione di Smart Grid, ma anche per la gestione degli impatti ambientali associati al fine vita dei moduli fotovoltaici. L'impegno di smaltire in modo idoneo i moduli FV è prescritto nei dettami del Quarto Conto Energia (DM 5 maggio 2011), in riferimento agli impianti entrati in funzione successivamente al 30 giugno 2012, per i quali è richiesta una certificazione, da trasmettere al Gestore dei Servizi Energetici (GSE), rilasciata dal produttore o importatore dei moduli, che ne attesti l'adesione ad un sistema, o consorzio europeo, e ne garantisca il riciclo al termine della loro vita utile, pena l'esclusione dai benefici previsti dalle tariffe incentivanti del Quinto Conto Energia (DM 5 luglio 2012). Per una corretta gestione dei rifiuti FV è necessario analizzare, in primo luogo, il trend di crescita dell'energia FV non soltanto in termini quantitativi, ma soprattutto in relazione alle dinamiche evolutive delle innovazioni tecnologiche, che ovviamente influiscono sulla tipologia di rifiuto da trattare nonché sull'arco di vita utile dei moduli, che tende ad aumentare fino a circa 40 anni nel caso di alcune tecnologie di terza generazione.

Le significative quantità di rifiuti che saranno generati in Europa nei prossimi anni, peraltro facilmente stimabili e in gran parte riciclabili, richiede un improcrastinabile impegno verso la realizzazione della concreta sostenibilità di tale filiera. Ciò impone la necessità di programmare fin d'ora, accanto alla realizzazione di modelli energetici smart collegati alla produzione di energia fotovoltaica, anche la dismissione degli impianti FV alla fine del loro ciclo di vita, pianificandone le operazioni di demolizione, smantellamento, rimozione, smaltimento e/o riciclo dei differenti componenti.

4. Considerazioni conclusive

Il corretto smaltimento dei moduli FV, richiede preliminarmente la predisposizione di una banca dati per la tracciabilità degli stessi, al fine di ottenere le informazioni necessarie per un loro efficiente censimento, onde poter eseguire trasparenti analisi tecnico-economiche, ambientali e logistiche idonee alla individuazione delle migliori opzioni disponibili. Nel 2007 alcune organizzazioni europee operanti nel settore FV (produttori, importatori,

rivenditori, ma anche enti di ricerca) hanno dato vita ad un'associazione senza scopo di lucro - PV CYCLE - con sede a Bruxelles, con l'obiettivo di avviare un programma "volontario" per il ritiro e riciclo dei moduli a fine-vita. Occorre evidenziare che per elaborare corrette valutazioni in ordine agli impatti ambientali economici e sociali di una filiera produttiva, occorre avvalersi di una prospettiva *Life Cycle Management* (LCM), i cui strumenti operativi sono rappresentati dalla *Life Cycle Assessment* (LCA), dalla *Life Cycle Costing* (LCC) e dalla *Social Life Cycle Assessment* (S-LCA). La prima è una metodologia standardizzata (Iso 14040), basata sulla valutazione sistemica delle criticità ambientali di un prodotto (ma anche di un processo o di un'attività), attraverso l'identificazione e la quantificazione dei consumi di materie prime e di energia nonché degli impatti inquinanti associati a tutte le fasi del ciclo di vita: estrazione delle materie prime, trasformazione, trasporto, uso, riciclo e/o smaltimento, seguendo un approccio *from cradle to grave*, oppure *from cradle to cradle* nel caso di riciclo o riuso del prodotto, o nel caso di riciclo di alcuni suoi componenti soltanto (ad esempio, in riferimento ai moduli fotovoltaici: vetro, alluminio, plastiche, etc.), per realizzare una virtuosa "chiusura del cerchio" di tale ciclo. La seconda (LCC), invece, si propone di effettuare una valutazione economica complessiva delle scelte in un orizzonte temporale esteso fino la dismissione del bene. La terza (S-LCA), infine, rappresenta uno strumento che consente di avere una visione strategica e manageriale della sostenibilità, in relazione agli impatti sociali associati a prodotti e/o processi, onde rendere più trasparenti le decisioni e, soprattutto, incentivare l'adozione di prassi socialmente responsabili da parte degli operatori economici, rispetto ad una molteplicità di *stakeholders* (cittadini, utenti, lavoratori, etc. In estrema sintesi, l'obiettivo delle reti del futuro è quello di rendere più efficiente e sostenibile la filiera elettrica, attraverso l'implementazione di tecnologie orientate, da un lato, al risparmio energetico e, dall'altro, alla tutela ambientale in un'ottica LCA, ovvero durante tutte le fasi, dalla produzione alla trasmissione e distribuzione all'utente finale.

Riferimenti bibliografici e sitografici

Bloomberg New Energy Finance, 2012, Bazilian M., Onyeji I., Liebreich M., MacGill I., Chase J., Shah J., Gielen D., Arent D., Landfear D., and Zhengrong S., " Re-considering the Economics of Photovoltaic Power".

Cavallin F., Lorenzoni A., Sofia G., "Prospettive di parità per il fotovoltaico italiano", Qualenergia.it, 14 dicembre 2011, <http://qualenergia.it/articoli/20111214-prospettive-di-parita-grid-parity-del-fotovoltaico-in-Italia>

GSE, Rapporto Statistico 2011. Impianti a Fonti Rinnovabili, 9 ottobre 2012. <http://www.gse.it/it/Dati%20e%20Bilanci/Osservatorio%20statistico/Pages/default.aspx>

OECD/IAE, "ExecutiveSummary", 2011 http://www.iea.org/Textbase/npsum/Harness_Renewables2011SUM.pdf

Haptic Rendering of Deformable Surfaces in Medical Training

A. F. Abate¹, A. Casanova², M. Nappi¹, S. Ricciardi¹

¹DISTRA – Università degli Studi di Salerno

(abate, mnappi, sricciardi)@unisa.it

²Dipartimento di Scienze Mediche Internistiche "M. Aresu"

Università degli Studi di Cagliari

casanova@medicina.unica.it

Abstract. *Haptic systems applied to medical simulation and training can provide users with crucial kinaesthetic and tactile info otherwise impossible to convey. As the typical objects involved in this kind of simulation are very often deformable tissues, one of the main challenges in this area is the perceptually believable reproduction of these structures through haptic rendering techniques. In this paper we propose the use of a colour or grey scale bitmap to associate local deformability info to the geometry. This approach allows the visual-haptic engine to modulate the resistance to compression exerted by the simulated tissues based to a local parameter. By using multiple layers of textures or animated textures, locally non-linear or even dynamic behaviours can be simulated with a low computational load. Preliminary experiments based on a Immersion Cyberforce hand-based haptic device are encouraging.*

Keywords: haptic systems, haptic rendering, simulation and training.

1. Introduction

Haptic systems have a remarkable potential for many applications, specially those involving a specific tactile know-how. Indeed, medical applications such as tele-surgery or surgical simulators may particularly benefit from haptic interfaces, but their efficacy depend on the realism of the visual-haptic perceptions provided to the system's user. To this regard the simulation of the contact with different structures or tissues characterized by different deformability (i.e. reaction to the contact forces) can represent one of the key advantage of haptic based interaction compared to conventional visual interaction, as it provides user with a level of info not derivable otherwise. Unfortunately, contact modeling in medical simulation is a challenging problem. The way contacts are handled plays a very important role in the overall behavior

Congresso Nazionale AICA 2013

of the interacting objects. The kind of contact model adopted (including friction or not), highly influences the post-impact motion of the interacting objects. In most simulators with haptic feedback, the collision response of soft tissues with a virtual surgical instrument is assumed to be very local: the interaction only consider a single point [Mahvash and Hayward, 2004] [Mendoza et al., 2002].

The most popular approach is the penalty method which consists in defining a contact force $f = k$ at each contact point where is a measure of the interpenetration between a pair of colliding objects, and k is a stiffness parameter. This stiffness parameter must be large enough to avoid any visible inter-penetration, however, its value cannot be determined exactly. In addition, if an explicit time integration scheme is used, and k is large, very small time steps are required to guarantee the stability. The quick growth of energy in the haptic control loop induced by the method often leads to excessive damping in the provided solutions. A possible improvement over the penalty method can be achieved through the use of an implicit integration scheme [Meseure, 2003]. Yet, solving the resulting stiff and non-smooth system can be computationally prohibitive when the objective is to reduce as much as possible the inter-penetration distance. Some methods, developed for force feedback applications are based on the avoidance of visible inter-penetration through constraint-based techniques. The collision can be prevented by geometrical deformation constraints like in [Picinbono, 2002] or [Forest et al., 2004], or by god-object [Zilles and Salisbury, 1995] and proxy [Barbagli et al., 2003] methods. Another way is the use of Lagrange multipliers, which are appropriate for handling bilateral constraints [Galoppo et al., 2006]. However, contacts between objects intrinsically define unilateral constraints, which means that physics is not always verified when using techniques based on Lagrange multipliers. As a consequence, colliding objects could stay artificially stuck at the end of the time step. Improvements over constraint-based techniques are possible by using a Linear Complementary Problem (LCP) formulation. The solution of the LCP gives an accurate description of the contact forces needed to zero out the interpenetration, and prevents objects to stick together [Pauly et al., 2004]. By expanding the LCP, or by using a non-linear solver, the formulation can be extended to model both static and dynamic friction for rigid [Anitescu et al., 1999] and deformable [Duriez et al., 2006] objects. Computationally efficient methods for solving linear complementary problems are proposed [Murty, 1997], thus making such approaches appealing even for interactive simulations.

Anyway, though contact modeling is a major concern in haptic systems, most of the methods described so far compute reaction force according to a stiffness parameter often defined at an object level (e.g. the whole object surface has the same stiffness) and even in case local stiffness parameters are available they approximate the real surface behavior with a very coarse granularity. The same consideration may be applied to the law which approximates the typically non-linear behavior of the object's surface when compressed: whatever the law adopted, it is the same for the whole surface.

Unfortunately, the consistency of many organic tissues is complex to describe at a global level and is often related to their healthy or pathological condition, providing crucial info to the specialist during palpation. As one of the

main aims of visual-haptic simulators is to replicate the most faithfully possible the perceptual aspects of the interaction, we believe that a greater attention should be put to the simulation of those surface characteristics which may enhance the “haptic knowledge” of the system’s users.

2. Map-based approximation of deformable surfaces

We propose a simple yet effective and efficient way to represent detailed information on local stiffness of a simulated surface by means of a dedicated bitmap providing a much greater accuracy than usual object-level attributes or even coarse array of parameters.

The basic idea behind this proposal is to exploit texture mapping (typically used to simulate visual properties such as ambient and diffuse color, transparency, roughness, shininess, etc.) to associate local deformability data to 3D geometry. The local surface stiffness encoded by color at a texel level is then exploited by the haptic renderer to modulate the contact force feedback due to local tissue deformation resulting from interaction. More precisely, the deformability map is associated to mesh vertices through mapping coordinates in the form (u, v) , previously projected onto the surface.

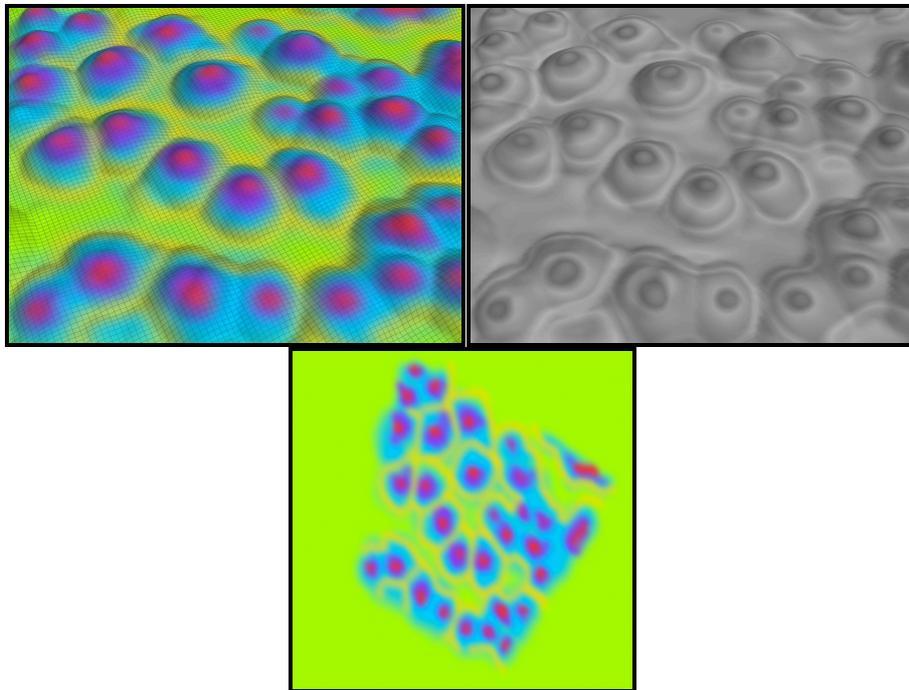


Fig. 2 – Color and grayscale deformability maps

The additional info can be represented through each pixel's RGB channels in a color texture or, in the simplest case, in a grayscale image, according to different arrangements offering a great flexibility of use. A deformability can be produced by a data driven methodology (for instance from image processing of diagnostic data or procedurally from anatomical models) or even by hand, by means of a 3D paint application.

In its simplest form an 8 or 16 bit grayscale image may encode the local stiffness parameters required to compute the reaction force at a texel level, thus providing a range of 256 or 65536 stiffness levels with a spatial granularity only depending by image's resolution. A 24 bit color image is able to arrange a more articulated set of data. Indeed it can store three 8 bit wide layers of stiffness data, enabling to simulate a non-linear surface behavior at a texel level. In other terms, it is possible to exploit the three (or even four if a 32 bit image is used) stiffness values associated to a given (u, v) position on the surface as a discrete approximation of the stiffness measured at a progressively greater depth. Where required, by using multiple textures the approximation can be easily improved.

As the deformability of a surface involves both visual and haptic feedback during simulation, but not necessarily these two channels (which in most applications are decoupled due to different frame-rate requirements) are supposed to share the same stiffness coefficient, mixed visual-haptic stiffness data may be embedded in the same 16, 24 or 32 bit deformability map (e.g. 8 bit visual + 8 bit haptic, or 16 bit visual + 8 bit haptic, or even 16 bit visual + 16 bit haptic). Moreover dynamic modification of the local stiffness can also be rendered by means of animated bitmaps or implementing a real time procedural processing of bitmap pixels in the RGB color space (e.g. a color shift may simulate a modification of the surface's deformation capability).

An additional advantage of this technique is related to its ultra-fast computing by means of modern Graphics Processing Units (GPUs). Indeed these vector processors can effectively process multiple very large (up to 8K pixel wide) textures on a single pass due to their highly parallel cores, while their dedicated VRAM (up to 4GB wide) could arrange many texture layers for each object in the virtual environment.

3. First results and future research

The proposed technique has been experimented on a visual-haptic platform which is part of a wider research project aimed to the training of obstetricians to delivery. A CyberForce® hand-based force feedback system by Immersion Corporation has been utilized in order to provide the user with haptic sensations during the simulated intervention (see Figure 1). The CyberForce is composed by an articulated exoskeleton anchored to the back of the user's hand which is devoted to the restitution of grounded forces to the user's arm-hand system within its operative volume. It ends up with a CyberGrasp system including a 22 sensors CyberGlove for each finger's flexion/abduction capturing. The CyberGrasp provides the transmission of contact sensations and object

resistance to grasp to the operator's fingertips. The test bed hardware included a dual quad-core Intel Xeon processor based Mac Pro workstation from Apple Inc., equipped with 8 Gigabytes of RAM and an Nvidia Quadro 5600 graphics board with 1,5 Gigabytes of VRAM. Overall the system's testers were positively impressed by quality of deformation of the simulated tissues (child's head and body, mother's perineal area), even if the quantity of deformation was sometimes inadequate to imitate actual plastic phenomenon due to the compression exerted on the child's head. Nevertheless the system capability to provide (via the deformability map and the related haptic rendering) different perceptions of the deformability of the structures handled through the haptic device, was appreciated and considered one of the strong point of the on-going research for a realistic delivery simulator. Besides performing an extensive testing of the proposed technique in different applicative environments, we are exploring the possibility to use a 24 bit depth image to store a quantized stiffness vector (each RGB component is a vector component, instead of a single scalar value) which could be useful for improve surface deformation calculations during inter-object contact.



Fig. 3 - A view of the visual-haptic delivery simulator test-bed, featuring an Immersion CyberForce hand-based haptic device.

References

- [1] M. Anitescu, F. Potra, and D. Stewart. Time-stepping for three-dimensional rigid body dynamics. *Computer Methods in Applied Mechanics and Engineering*, (177):183–197, 1999.
- [2] F. Barbagli, K. Salisbury, and D. Prattichizzo. Dynamic local models for stable multi-contact haptic interaction with deformable objects. *Haptic Interfaces for Virtual Environment and Teleoperator Systems 2003*, pages 109–116, March 2003.
- [3] C. Duriez, F. Dubois, A. Kheddar, and C. Andriot. Realistic haptic rendering of interacting deformable objects in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 12(1):36–47, 2006.
- [4] C. Forest, H. Delingette, and N. Ayache. Surface contact and reaction force models for laparoscopic simulation. In *International Symposium on Medical Simulation*, June 2004.
- [5] N. Galoppo, M. A. Otaduy, P. Mecklenburg, M. Gross, and M. C. Lin. Fast simulation of deformable models in contact using dynamic deformation textures. In *SCA '06*, pages 73–82, Switzerland, 2006. Eurographics Association.
- [6] M. Mahvash and V. Hayward. High-fidelity haptic synthesis of contact with deformable bodies. *IEEE Computer Graphics and Applications*, 24(2):48–55, 2004.
- [7] C. Mendoza, K. Sundaraj, and C. Laugier. Faithfull force feedback in medical simulators. In *International Symposium in Experimental Robotics*, volume 8. Springer, 2002.
- [8] P. Meseure. A physically based virtual environment dedicated to surgical simulation. In *Surgery Simulation and Soft Tissue Modeling (IS4TM)*, pages 38–47, June 2003.
- [9] K. Murty. *Linear Complementarity, Linear and Nonlinear Programming*. Internet Edition, 1997.
- [10] M. Pauly, D. K. Pai, and L. J. Guibas. Quasi-rigid objects in contact. In *SCA '04*, pages 109–119, Switzerland, 2004. Eurographics Association.
- [11] G. Picinbono, J.-C. Lombardo, H. Delingette, and N. Ayache. Improving realism of a surgery simulator: linear anisotropic elasticity, complex interactions and force extrapolation. *Journal of Visualisation and Computer Animation*, 13(3):147–167, July 2002.
- [12] C. B. Zilles and J. K. Salisbury. A constraint-based god-object method for haptic display. In *IEEE IROS '95: Proceedings of the International Conference on Intelligent Robots and Systems*, pages 31–46, 1995.

In store augmented reality: retailing strategies for smart communities

Maria Teresa Cuomo, Debora Tortora¹, Gerardino Metallo²

Università degli Studi di Salerno

Dipartimento di Studi e Ricerche Aziendali - Management & Information Technology

Via Giovanni Paolo II, 132, 84084 Fisciano (SA)

mcuomo@unisa.it

¹ Università degli Studi di Salerno

Dipartimento di Studi e Ricerche Aziendali - Management & Information Technology

Via Giovanni Paolo II, 132, 84084 Fisciano (SA)

dtortora@unisa.it

² Università degli Studi di Salerno

Dipartimento di Studi e Ricerche Aziendali - Management & Information Technology

Via Giovanni Paolo II, 132, 84084 Fisciano (SA)

gemetall@unisa.it

Abstract. *The multi-channel strategy becomes crucial for the competitiveness of retail system; it is based on open trustworthy relationships, on the support of new communication tools (e.g. social media), and on innovative devices (24/7 connections), which extend possibilities, process and moment of purchasing. Thanks to the support of innovative information technologies, the store constitutes a privileged area of integration between real and digital, producing ulterior narrative levels, where brand management strategies confront with new social spaces. Augmented reality, ranging from mobile to proximity marketing and augmented advertising, add different degrees of information to the consumers' sense. So as the augmented reality reshapes the commercial area, providing it for contextual information and activable by potential customers when needed; at the same time, the integration with the mobility reconfigures the mode of use, organizing new opportunities of connections with the user. In addition, the generation of contents both bottom-up and top-down makes the individual from tryer into buyer into advertiser, through social networking, generating greater value experiences and, therefore, additional sales.*

Keywords: augmented marketing, retailing, content brand perceptions.

1. Introduzione

La capacità di “contestualizzare” – proposte d’acquisto, beni di consumo, contenuti valoriali di *brand* ed interazioni tra individui/consumatori con i prodotti – si costituisce come tratto caratterizzante la formula del *retailing* pervasivo (non più invasivo, perché “attivato” dal consumatore), che attinge alla realtà

aumentata per identificare e definire inediti momenti di interazione. Di fatto, trovare promozioni o attività commerciali (o farsi rintracciare da altri utenti) attraverso informazioni geolocalizzate, visualizzare contenuti multimediali dedicati o indicazioni tratte dal *web* come dai *social network* (Riva, 2010; Cuomo et al., 2011), disporre di istruzioni “sovrapposte” al prodotto fisicamente presente nel punto vendita, seguire i *tweet* in maniera fisica oltre che telematica, taggare oggetti reali, segnalare notizie in e su luoghi specifici rappresentano attività in grado di modificare la percezione del consumatore, all’interno di un nuovo sistema di *social sensemaking*. Di più, la piena comprensione della portata strategica di tali cambiamenti e dei suoi possibili sviluppi – dove rapporti di fiducia basati su relazioni aperte, nuovi strumenti di comunicazione (p.e. *social media*), dispositivi innovativi e connessioni 24/7 (l’utente mobile è, per definizione, *always on*), estendendo possibilità, processo e momento d’acquisto, fanno della multicanalità un fattore imprescindibile per la competitività in ambito *retail* – può essere compresa pienamente solo alla luce della re-interpretazione del punto di vendita quale contesto socio-costruito, in grado di attivare un diverso rapporto con i pubblici. L’arena commerciale, infatti, si trasforma da *space*, ovvero area in cui prevalgono gli aspetti legati alla fisicità, a *place*, vale a dire un luogo disegnato dalle pratiche e dalle esperienze che lo definiscono (Accoto e Mandelli, 2012). In altre parole, nella sua integrazione con la realtà aumentata, il punto di vendita diviene un “contesto specifico” di una particolare comunità di pratica (Cova, 2003; 2008).

2. Augmented reality e nuovi scenari di marketing

Un miliardo di *smartphone* abilitati alla realtà aumentata entro il 2016, contro 150 milioni di unità al momento operativi (Editoriale, 2013a), già allo stato attuale nell’ambito del *retail commerce* costituiscono dispositivi atti a governare fattori abilitanti, deputati ad animare storie sui prodotti, suscitare interesse verso le offerte e generare emozioni/esperienze tra gli scaffali, proponendo modalità di interazione tra gli individui e la realtà circostante facilmente utilizzabili dagli utenti. Al contempo, l’evoluzione delle tecnologie (Frà et al., 2011, p. 22) – sia per l’introduzione di elementi analogici (*sound recognition, computer vision*) all’interno del mondo digitale, sia per ricreare nel mondo digitale realtà di interconnessione, analisi semantica e sintesi delle informazioni disponibili nel mondo reale – ed il proliferare di applicazioni con interfacce informatiche facilmente integrabili, consentono di comporre e fornire servizi finali sempre più specializzati e dedicati (Frà et al., 2011). Invero, proprio grazie al supporto di tali tecnologie informatiche innovative (già dagli anni ‘90 i *Quick Response Code* erano in grado di veicolare una quantità maggiore di informazioni sui beni presenti nell’area commerciale, offrendo la possibilità di connessione diretta al sito del produttore – Editoriale, 2013b) il punto vendita si costituisce come ambito privilegiato di integrazione tra reale e digitale, generando piani narrativi in cui le strategie di *brand management* si aprono a nuovi spazi relazionali e di socialità (Mandelli, 2011).

Comparsa in letteratura già negli anni ‘40, ma a diffondersi solo alla fine degli anni ‘80, l’espressione *augmented reality* (AR – coniata nel 1992 dal ricercatore

Thomas Preston Caudell, ingegnere della Boeing, per descrivere un sistema di nuova generazione che avrebbe aiutato nell'assemblaggio e installazione dei cavi elettrici negli aerei – Caudell e Mizell, 1992) indica «*a term for a camera enhanced view of a physical real-world environment, where virtual elements are merged with the real-life scene creating a “mixed reality” of virtual elements and the real world. The “virtual elements”, given their nature, can consist of anything. This could be 3D models, video, web information...anything. The point here is that your mind is the only boundary.*» (Trubow, 2011a, p. 4). Dunque, uno schermo ed una proiezione risultano sufficienti per integrare la realtà (fisicamente intesa) in un contesto virtuale.

Tuttavia, pur ponendosi come polo di un *continuum*, a differenza della *virtual reality* (VR), in cui le informazioni aggiunte o sottratte elettronicamente risultano preponderanti, inducendo nel consumatore una sensazione di “annullamento” dell'esplorazione sensoriale circa il contesto o la situazione, nella AR il sistema percettivo di tipo sensoriale dell'acquirente, attivo e protagonista, viene sollecitato dall'aggiunta di livelli informativi ed emozionali di varia natura (attraverso contenuti multimediali, tipo video, audio ed animazioni), fruibili “volontariamente”, appunto, ed allo stesso tempo trasferibili ad altri utenti – dopo essere stati modificati dal fruitore, che aggiunge a sua volta contenuti virtuali, i quali, transitando sulle piattaforme sociali, generano collegamenti con il *brand* – aggiungendo valore “auto-prodotto” ad una *shopping experience* di tipo immersivo (Pine e Gilmore, 2000; Tortora, 2007). *Output* evidente è la definizione di un modello di marketing di tipo partecipativo – basato sul contributo della comunità di riferimento, decretando la fine della centralità dei *mass media* nelle comunicazioni di prodotto/*brand* – che ridisegna il proprio interesse strategico intorno alla creazione di esperienze interattive e sociali in grado di conferire maggior valore all'offerta (Mardegan et al., 2012; Arvidsson e Giordano, 2013). Quest'ultimo riverbera dal suo significato strettamente economico, poiché «nasce all'interno di un'esperienza personale che viene condivisa e abbraccia le dimensioni di identità, di senso e di appartenenza ad una comunità» (Riva, 2012, p. 214), ridefinendo altresì il *brand equity* (Cuomo et. al, 2009). Dunque, per sopravvivere in questa “*new age of marketing*” le imprese devono individuare strategie di (*digital*) *engagement*, per proporre una progressione nell'interazione con i consumatori che progredisca dalla persuasione alla relazione, all'esperienza, alla condivisione (Scatena e Mardegan, 2012). Oggi, infatti, conoscenza, affidabilità e diffusione del marchio e distribuzione capillare dell'offerta non risultano più sufficienti ad assicurarne il successo, ovvero coinvolgere i consumatori e metterli al centro dell'esperienza di consumo, dal momento che la *governance* sulle interazioni di marca si ridefinisce di volta in volta all'interno di spazi specifici socio-costruiti dall'incontro tra ambiente, tecnologia, servizio e socialità (figura 1).

D'altra parte ricerche recenti (Scatena e Mardegan, 2012) segnalano una correlazione positiva e significativa tra i livelli di *engagement* promossi da un *brand* ed il fatturato ed il margine realizzati dall'azienda, specie in periodi di crisi, come quello attuale. Pertanto, la AR offre ai *marketer* inesplorate possibilità di avvicinamento e coinvolgimento del cliente, soprattutto in termini di *branding*.



Fig. 1 – Strumenti per promuovere il coinvolgimento del consumatore

Anzi, personalizzazione ed intimità del messaggio, tempestività, localizzazione e contestualità della *user experience* la stanno rendendo già oggi una piattaforma mediale pressoché essenziale per riposizionare le offerte nel processo di *sense and value making*.

3. Augmented in store experience

La possibilità di fruizione delle applicazioni di realtà aumentata attraverso gli internet standard *browser* suggerisce una serie di nuove opportunità nella relazione impresa-cliente (Trubow, 2011b, p. 5 – figura 2).

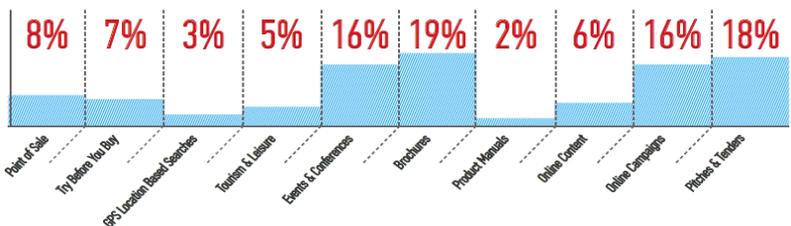


Fig. 2 – Uses based on current adoption by major brands

Emblematico in tal senso lo sviluppo registrato dall'*augmented advertising*, ovvero pubblicità e promozione in AR dei principali *corporate brand*. Si faccia riferimento, per comprendere la differenza in termini di partecipazione ed impatto cognitivo ed emotivo del messaggio, alla recente applicazione di AR Aurasma (relativamente alla *app economy*, Kim et al., 2011), che permette di riconoscere immagini nel mondo circostante, alle quali sovrappone in tempo reale un *layer* virtuale in grado di riprodurre contenuti multimediali relativi all'immagine catturata: così, quando l'utente registrato inquadra con la fotocamera del proprio *smartphone* la locandina di un film, Aurasma riconosce l'immagine (che deve essere contenuta nel *database* delle *tag* Super Anywhere dell'applicazione) e ne riproduce il *trailer*. Gli utenti possono anche creare degli *anywhere* personali, associando alla foto di un logo, piuttosto che di un oggetto, un contenuto multimediale a tema e condividerli con gli altri fruitori (Filardo et

al., 2012). Tale applicazione funziona, naturalmente, anche per i prodotti, “posizionando” letteralmente il *brand* nelle “mani” dei potenziali consumatori.

Anche i *retailer* iniziano a fare ampio ricorso alla AR “to bring to life point of sale”. In effetti, la tecnologia della AR trova diversi punti di applicazione nel processo di vendita (Trubow, 2011a – figura 3).



Fig. 3 – Augmented reality and sales strategy

Fornendo nuovo impulso all'interpretazione del modello AIDA, nel favorire la cattura dell'attenzione ed alimentando l'interesse del consumatore per l'offerta, attraverso una *selling proposition* in grado di mostrare “realmente” la capacità di *problem solving* dell'offerta, essa sostiene attivamente il consumatore nel suo processo di acquisto e consumo. Invero, uno dei maggiori vantaggi offerti dalla AR risiede proprio nella capacità di promuovere una “interazione diretta” con l'offerta, aspetto tipicamente relegato alla compresenza di acquirente e prodotto in un contesto fisico, il negozio, appunto, e soddisfatto attraverso il senso del tatto. Molti studi dimostrano, infatti, come l'esercizio della tattilità incrementa la *customer intention to buy* ed altresì la disponibilità a sostenere un sacrificio economico più elevato (Trubow, 2011a). Attraverso la AR è possibile, proprio all'interno della *trading area*, rendere l'interazione con l'offerta ben più profonda, rispetto al mero contatto tattile, potendo effettivamente realizzare per il frequentatore una esperienza di immersione (Addis, 2005) grazie al supporto di informazioni addizionali, animazioni, dettagli sulla componentistica, indicazioni sulla composizione dell'offerta, funzionamento, contestualizzazione o anticipazione di fruizione, in cui il potenziale acquirente è interprete e co-autore, e non solo beneficiario, del risultato raggiungibile (tabella 1).

	<p>Lego ha implementato alcuni terminali di realtà aumentata – Lego Digital Box – che consentono ai consumatori, posizionando la scatola davanti ad una telecamera, di osservare il giocattolo nei minimi particolari, senza dover aprire la confezione, guardandone la rappresentazione assemblata in 3D (www.korus.fr).</p>
	<p>Scaricando l'applicazione gratuita Aurasma Lite ed inquadrando la copertina del catalogo 2012 e altre pagine contenenti il logo “A”, il catalogo prende vita, con video ed immagini che raccontano l'universo Mercatone Uno, con notizie, dettagli e contenuti extra, grazie ad una consultazione interattiva, direttamente con lo smartphone (www.realta-aumentata.it/home.asp).</p>

 <p>i'm lovin' it®</p>	<p>Per aggirare il problema delle critiche su qualità, genuinità e provenienza dei cibi utilizzati è stata progettata un'applicazione di Realtà Aumentata, "Track My Macca", che mostra agli utenti australiani con animazioni molto originali gli ingredienti di provenienza ed alcune informazioni sull'agricoltore, sfruttando il riconoscimento di immagini e il GPS (www.realta-aumentata.it/home.asp).</p>
	<p>Per arricchire l'esperienza dei consumatori anche sul punto di vendita e valorizzare l'offerta di gamma delle sorprese, Kinder ha offerto per la Pasqua 2013 un ulteriore momento di gioco dedicato ai bambini. In esclusiva su alcuni ipermercati, infatti, nei pressi delle isole Kinder GranSorpresa allestite in occasione della Pasqua, è stato installato un pannello con un monitor. Ponendo uno tra le uova Kinder GranSorpresa di fronte alla webcam, questo si trasforma in una delle sorprese in esso contenute, proprio per enfatizzarne il valore e l'eccellenza qualitativa. Ogni uovo Kinder GS dà vita al mondo delle sorprese Kinder: un'occasione reale per i bambini di "immergersi" nel gioco (www.saporinews.com).</p>
	<p>In occasione dell'operazione "la moda 24h/24, 9-12.10.2012, le Galeries Lafayette Lille hanno installato nell'area vendita delle cabine interattive che consentono di provare un certo numero di abiti senza spogliarsi e molto velocemente. Lo specchio virtuale presente in camerino ricompono l'immagine del cliente che vi si posiziona di fronte e rileva automaticamente i suoi movimenti. Con un semplice gesto il consumatore cambia l'abito, il colore o la taglia, mentre i vestiti si adattano alle sue forme e alla sua gestualità (www.korus.fr).</p>

Tab. 1 – Augmented in-store experience: some examples

Utilizzata come parte di una strategia promozionale, o per stimolare il traffico sul punto di vendita, ancora, per incrementare la *brand awareness* e la *customer loyalty*, dilatando l'esperienza d'acquisto in ambienti di *mixed reality*, la AR in sintesi ridisegna lo spazio commerciale, offrendo un diverso modo di percepire il senso del luogo; influenzati strettamente anche dai processi culturali e sociali degli utilizzatori, chiamandone in causa anche aspetti antropologici e semiotici, le installazioni *stand alone* per i punti vendita diventano ora postazioni di interazione con i consumatori e di questi tra di loro e con altri soggetti non fisicamente presenti (valore di scambio). In tal senso, dunque, si può parlare di realtà socialmente aumentata: la *shopping experience* diviene, così, anche esperienza sociale.

L'arena commerciale, nella sua interpretazione quale contesto in cui si realizza la *shopping experience*, risponde ad una duplice categoria: di tipo situazionale, in cui, segnali anche in forma analogica (p.e. il riconoscimento di un'immagine o di un suono) costituiscono *input* per successive elaborazioni che permettono di sapere cosa accade, quali elementi siano presenti nella realtà fisica e a quali informazioni corrispondano nella realtà virtuale che circonda il consumatore – movimento *inbound* – e conversazionale, che permette di raccogliere informazioni dai *social media*, concentrandosi sulle comunicazioni attorno ad esso – movimento *outbound* (Frà et al., 2011).

Tuttavia, mentre nel movimento *inbound* (verso il soggetto) l'impresa continua ad interpretare un ruolo preminente, predisponendo ed orientando il processo di interazione soggetto-contesto, nel trasferimento *outbound* il controllo dei contenuti delle comunicazioni massmediali avviene da parte di un numero crescente di attori spazialmente diffusi e/o raccolti in comunità virtuali – *human broadcaster* (Riva, 2012), ridefinendo ed allargando i confini della

customer experience. Connessione (anche in mobilità), creazione di contenuti multimediali (soprattutto foto e video) e loro condivisione (insieme ad opinioni sulle offerte, ma anche recensioni ed approfondimenti su contenuti specifici), uso della tecnologia per rafforzare il senso di appartenenza alla comunità di riferimento caratterizzano la relazione consumatore-prodotto per la “Generazione C” (De Felice, 2011). Ciò nondimeno, soprattutto con riferimento al *mobile marketing*, ma estensibile più in generale alla tipologia di *shopping experience* promossa dalle tecnologie immersive, molti studi mettono in evidenza come la convenienza – unitamente al valore edonistico che se ne può ricavare – costituisca ancora un fattore preminente nell’“utilizzo dei servizi “aumentati” ed in mobilità (figura 4).

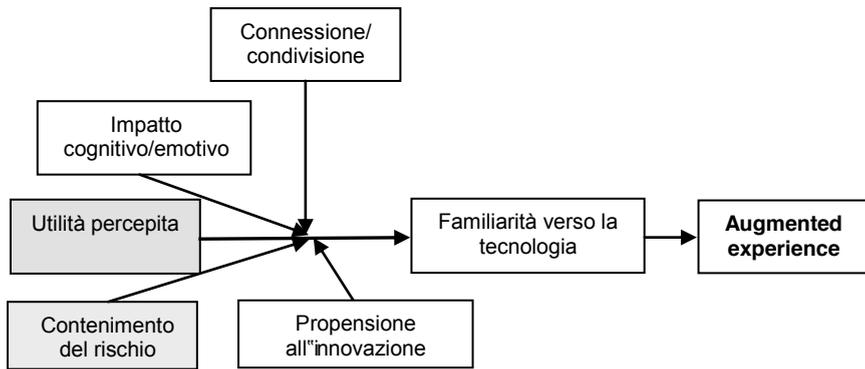


Fig. 4 – Un possibile modello concettuale di fruizione della AR

La coesistenza di fattori funzionali ed emozionali è, d’altra parte, condizione condivisa negli studi su *mobile internet*, *mobile data service*, *mobile multimedia service* e *location-based mobile service* (Varnali e Toker, 2010).

A completamento, è possibile individuare precise componenti che definiscono la *customer experience* negli *immersive store* (tabella 2 – Pantano e Servidio, 2012, p. 283).

Facilities	Fast response, secure transaction, system flexibility, entertainment
Product formation	Product variety, detailed product information, personalized information
Service	Product selection assistance, virtual sales person, online support
Convenience	Details about the firm, navigational efficiency, more realistic navigation and interaction
Appearance	Pleasant, attractive, more realistic appearance
Institutional factor	Requirements related to consumer’s profile, firm’s information, consumer’s position tracer while in the immersive store

Tab. 2 – Immersive store

In sintesi, la forza generativa di contenuti sia *bottom-up* (gestita dall’impresa/brand) che *top-down* (generazione e condivisione di contenuti multimediali), grazie ad una grammatica digitale in grado di garantire l’interazione tra le parti, trasforma il soggetto-consumatore:

- da *tryer*, fruitore della AR che sperimenta prima del potenziale acquisto l'offerta *augmented* in contesti *augmented*, in termini di qualità, rilevanza dell'apporto di tipo *problem solving* del prodotto, mantenimento delle promesse,

- a *buyer*, facendo leva sulla *value proposition* di incontro, esperienza contestualizzata, dinamiche di aggregazione, ma anche *privacy* ed accesso a transazioni sicure,

- ad *advertiser*, grazie alla *human broadcasting communication* che, attraverso il *social networking*, partecipa alla generazione e diffusione della *brand story*, motivando e supportando (se con orientamento soddisfacente) in maniera spontanea altri potenziali *tryer-buyer-advertiser*, potendo generare nel complesso esperienze di maggior valore e, dunque, maggiori vendite per l'impresa.

4. Brevi riflessioni critiche e sviluppi futuri di ricerca

Come detto, quindi, le nuove tecnologie immersive e gli strumenti digitali, in specie quelli sempre connessi, costituiscono un'occasione "vibrante" per promuovere il punto vendita, farsi rintracciare ed incrementarne la frequentazione: la AR, infatti, soprattutto grazie alla diffusione dei dispositivi mobili, presenta un grado sempre maggiore di accessibilità sia per i *brand* che per i consumatori (Rohm et al., 2012). Va però sottolineato che il suo utilizzo può nascondere alcune insidie di non poco conto che vale la pena di rimarcare in questa sede. Anzitutto, le informazioni addizionali, le animazioni, i dettagli, gli approfondimenti tecnici ecc. che caratterizzano un contesto di realtà aumentata potrebbero generare un caos informativo per il consumatore ovvero un disorientamento decisionale, generando fenomeni di distorsione dal legame *brand/consumer*. A riguardo, va ricordato il ruolo fondamentale della marca quale vettore di garanzia/affidabilità in grado di semplificare il processo di acquisto, che sembrerebbe non sempre attuabile in uno scenario aumentato dove sia l'eccessiva mole di informazioni in circolo, sia l'incapacità/difficoltà di selezionare quelle realmente utili potrebbero rivelarsi addirittura devianti e controproducenti. Probabilmente sarebbe opportuno strutturare gli "store di realtà aumentata" individuando quei fattori realmente capaci di catturare il cliente, nonché di fornirgli quei contenuti informativi necessari, efficaci e coadiuvanti della *shopping experience*. Non si tratta dunque di uno strumento indispensabile e utilizzabile a tutti i costi!

Inoltre, vanno messi in luce i limiti che l'uso della AR potrebbe generare in termini di auto-produzione di informazioni da parte dei consumatori con il rischio di "manipolazioni/alterazioni" delle informazioni stesse in ambienti *social* per effetto della forza generativa dei contenuti generati e condivisi (*logica top-down*). Infatti, affinché la AR sia in grado di dare impulso ad una *value experience* deve riguardare temi effettivamente rilevanti per gli acquirenti, tuttavia con incentivi-benefici immediati e riconoscibili, facilità di utilizzo nonostante i limiti tecnologici dei dispositivi mobili (Prunesti e Lalli, 2011), infine, deve promuovere cooperazione e sinergie da parte di tutti gli attori intercettati lungo la catena del valore, con un *focus* sempre centrato sulla interazione impresa/consumatore.

In maniera preordinata, però, è fondamentale per le organizzazioni poter disporre di un solido costrutto in grado di definire un modello di *augmented (in store) customer experience*, a cui sia connesso anche un'efficiente attività di *monitoring* in ordine alla selezione dei dati addizionali proposti sulla base del contesto di consumo (situazione, occasione, ecc).

Ancora, occorre interrogarsi sulla applicabilità delle soluzioni di AR alle diverse tipologie di acquisto in maniera più o meno generalizzata. Intuitivamente si può affermare che la AR risulta maggiormente efficace in ipotesi di acquisti problematici, complessi, mentre per le altre tipologie- p.e. acquisti d'impulso- sarebbe forse da preferire una "realtà diminuita" (*Diminished Reality, DR*)?

Oltretutto vanno considerate anche le implicazioni delle soluzioni aumentate/diminuite sulla identificazione dei segmenti/target di mercato in funzione della ricerca di contenuti informativi alti/bassi (figura 5):

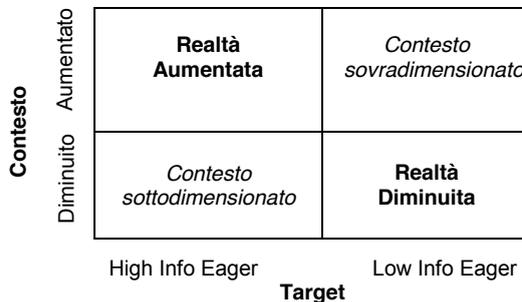


Fig. 5 – Segmenti di mercato in contesti AR/DR

Infine, la AR costituisce comunque "uno strumento" all'interno di un più ampio e complesso *brand's overall customer communication ecosystem*, in cui creatività, tecnologia e contenuti devono far convivere tradizionali e innovative forme, e relativi strumenti, per generare efficaci *customer engagement* (Schultz e Block, 2011). In tale ambito, pertanto, è opportuno che la ricerca scientifica proceda verso la validazione empirica dei modelli concettuali proposti (nel presente come in altri studi) per poter offrire una valida base teorica a piani e processi di *augmented marketing*.

Bibliografia

- Addis, M., L'esperienza di consumo. Analisi e prospettive di marketing, Pearson-Prentice Hall, Milano, 2005.
- Arvidsson, A., Giordano, A., Societing reloaded. Pubblici produttivi e innovazione sociale, Egea, Milano, 2013.
- Caudell, T.P., Mitzell, D.W., Augmented reality: an application of heads-up display technology to manual manufacturing processes. System Sciences, 2, 1992, 659-69.
- Cova, B., Il marketing tribale. Legame, comunità e autenticità come valori del Marketing Mediterraneo, Il Sole 24 Ore, Milano, 2003.
- Cova, B., Marketing tribale e altre vie non convenzionali: quali ricadute per la ricerca di mercato? Micro & Macro Marketing, 3, 2008, 437-447.

Cuomo, M.T., Metallo, G., Tortora, D., Opportunità, limiti e criticità dei social network. *Esperienze d'impresa*, 2, 2011, 25-48.

Cuomo, M.T., Metallo, G., Tortora, D., Testa, M., Kitchen, P.J., Building brand equity: the genetic coding of Mediterranean brands. *Euromed Journal of Business*, 4, 3, 2009, 237-253.

De Felice, L, Marketing conversazionale: dialogare con i clienti attraverso i social media e il Real-Time Web di Twitter, FriendFeed, Facebook, Foursquare, *Il Sole 24 Ore*, Milano, 2011.

Editoriale, Realtà Aumentata e QR code a confronto. www.arnews.it, 25 gennaio, 2013b.

Editoriale, Realtà aumentata: suo impiego nel mondo retail. www.puntodivendita.info, 11 febbraio, 2013a.

Filardo, V., Messina, M., Bortolussi, S., Marino, S., Megna, L. (a cura di), Augmented reality. Nuove applicazioni, nuove soluzioni, in *La comunicazione multimediale*, 2012, <http://www.arproject.altervista.org/intro.html>.

Frà, C., Lamorte, L., Martini, G., Dall'augmented reality al check-in. *Notiziario tecnico Telecom Italia*, 3, 2011, 20-31.

Kim, H.W., Lee, H-L., Son, J.E., An exploratory study on the determinants of smartphone app purchase, in *Proc. Of the 11th International DSI and the 16th APDSI Joint Meeting*, Taipei, Taiwan, July 12 – 16, 2011, 1-10.

Mandelli, A., Accoto, C., *Social Mobile Marketing. L'innovazione dell'ubiquitous marketing con device mobili, social media e realtà aumentata*, Egea, Milano, 2012.

Mandelli, A., *Processes of Value Creation in Markets as Mediated Conversations*. Working paper, 2011.

Mardegan, P., Riva, G., Pettiti, M., *Mobile Marketing: la pubblicità in tasca*, Lupetti Editore, Bologna, 2012.

Pantano, E., Servidio, R., Modeling innovative points of sales through virtual and immersive technologies. *Journal of Retailing and Consumer Services*, 19, 2012, 279-286.

Pine, J.B. II, Gilmore, J.H., *L'economia delle esperienze. Oltre il servizio*, Etas, Milano, 2000.

Prunesti, A., Lalli, F., *Geolocalizzazione e mobile marketing. Fare business con le App e i social game*, Franco Angeli, Milano, 2011.

Riva, G., *Digital Marketing 2.0. Multicanale, Sociale, Esperienziale, Mobile*. *Micro & Macro marketing*, 2, 2012, 213-218.

Riva, G., *I social network*, Il Mulino, Bologna, 2010.

Rohm, A.J., Gao T., Sultan, F., Pagani, M., Brand in the hand: A cross-market investigation of consumer acceptance of mobile marketing. *Business Horizons*, 55, 2012, 485-493.

Scatena, S., Mardegan, P., *Mobile Marketing: stato dell'arte e applicazioni pratiche*. *Micro & Macro marketing*, 2, 2012, 219-235.

Schultz, D.E., Block, M.P., Understanding customer brand engagement behaviors in today's interactive marketplace. *Micro & Macro marketing*, 2, 2011, 227-244.

Tortora, D., *Experience marketing e creazione di valore. Relazioni e interazioni tra consumatore, offerta e contesto*, Giappichelli editore, Torino, 2007.

Trubow, M., *Augmented reality marketing strategies: the how to guide for marketers*. Hidden Creative Ltd, www.hiddenltd.com, 22 marzo, 2011b, 1-18.

Trubow, M., *Sales technology: selling with augmented reality*. Hidden Creative Ltd, www.hiddenltd.com, 5 settembre, 2011a, 1-14.

Varnali, K., Toker, A., *Mobile marketing research: The-state-of-the-art*. *International Journal of Information Management*, 30, 2010, 144-151.

A Multi-Criteria Decision Aid for Agent-Based Decision Makers

Alba Amato, Beniamino Di Martino, Salvatore Venticinque
Dipartimento di Ingegneria Industriale e dell'Informazione, Seconda Università di Napoli
Via Roma, 29, Aversa (CE)
{alba.amato, beniamino.dimartino, salvatore.venticinque}@unina2.it

Abstract. *The multi-criteria decision-aid is intended to provide to agent-based decision-makers some tools, which help them to choose among alternatives, with many and contradictory points of view to be taken simultaneously in consideration. In this kind of problems such decision process deals with measuring the relevance of decision criteria and with combining them in order to evaluate and choose among the available alternatives. Decision support techniques are needed to solve those problems automatically. Specific rules can be derived from requirements in order to evaluate the each criteria in terms of increasing user's utility. In this paper we present how these techniques have been developed to help agents to take their decisions in three different application scenarios. Correspondent multi agents software solutions are described.*

Keywords: Multi-criteria decision, Semantic Discovery, Brokering

1. Introduzione

L'analisi multicriterio, o multiple criteria decision-making (MCMD) o multiple criteria decision-aid (MCDA), costituisce un valido strumento nei processi di supporto alla decisione negli ambiti in cui si deve affrontare la risoluzione di un problema di scelta in cui ci sono molteplici fattori da prendere in considerazione, e molteplici alternative spesso talmente diverse l'una dall'altra che risulta estremamente difficile indicare una preferenza (incomparabilità). In generale in questi casi non esiste nessuna decisione (soluzione, azione) che sia contemporaneamente la migliore rispetto a tutti i punti di vista coinvolti. Inoltre ci interessano quei contesti in cui ci sono differenti attori (stakeholder) che danno un valore differente ai fattori da prendere in considerazione, questi valori diversi rappresentano i criteri di selezione. Pertanto il problema risiede nel come misurare l'importanza di tali criteri e come combinare questi valori. Di solito non è possibile trovare soluzioni che perseguono contemporaneamente tutti gli obiettivi. Il problema di decisione viene risolto cercando la soluzione più soddisfacente, o "più coerente" con la logica dei decisori. La scelta deve quindi

essere eseguita tra quelle soluzioni che realizzino un certo livello di soddisfazione dei vari obiettivi, tale che non è possibile migliorare uno di loro senza degradarne altri. In particolare, in questo lavoro, l'analisi multicriterio è stata utilizzata per modellare una soluzione generale per questo tipo di problemi, ed è stato applicato in tre diversi scenari. In ognuno dei contesti applicativi presi in considerazione, il modello è stato implementato al fine di consentire ad agenti software di prendere decisioni circa le azioni da intraprendere per raggiungere i propri obiettivi.

2. Modellizzazione del problema

La metodologia di supporto alle decisioni qui proposta è mostrata in Fig. 1. E' necessario innanzitutto definire un modello di conoscenza e la sua rappresentazione formale che servirà anche a rappresentare contenuti e requisiti e, successivamente, a valutarli e confrontarli.

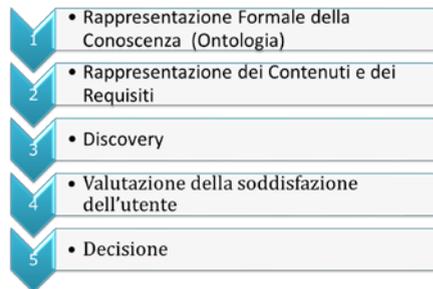


Fig.1 - Metodologia di supporto alle decisioni

Inoltre si deve fornire un servizio di discovery per raccogliere le informazioni inerenti le scelte disponibili. In generale, il processo di discovery consiste nel recupero di un insieme di risorse che sono rilevanti per la query dell'utente, espressa da un insieme di requisiti o da un profilo utente.

Il processo di discovery viene eseguito in modi differenti a seconda del problema. E' possibile definire obiettivi diversi, e in relazione a questi è necessario fornire sia la caratterizzazione quantitativa e qualitativa delle informazioni disponibili. In realtà i problemi decisionali sono regolati da una serie di soggetti interessati con i loro obiettivi e priorità. Infine è necessaria una strategia per valutare la decisione ottimale. Il modello di processo decisionale viene qui formalmente definito dalle seguenti voci:

- Un insieme di Believes $B = \{b_1, \dots, b_{nb}\}$, usati per descrivere la conoscenza del contesto applicativo, utilizzando un'ontologia di dominio definita per il caso specifico;
- La conoscenza dell'agente i al tempo t circa l'ambiente e le sue preferenze, descritta come un insieme di concetti ed individui appartenenti all'ontologia. Tale conoscenza viene utilizzata anche nella costruzione delle

query per recuperare le opzioni disponibili da valutare poi nel processo decisionale. Per ogni utente i è disponibile un profilo $P_i \subset B$ ed un insieme di vincoli $C(B)$;

- Un servizio di discovery per ricercare tutte le alternative disponibili che sono rilevanti per il profilo P_i ;

- Un insieme di obiettivi $G = \{g_1, \dots, g_{ng}\}$, che, se raggiunti, incrementano la soddisfazione dell'utente;

- Un insieme di azioni $A = \{a_1, \dots, a_{na}\}$ ottenute mediante il servizio di discovery. E' necessario definire una metrica (utilità) $U: A \rightarrow u_1(A), \dots, u_{nw}(A)$ per calcolare la misura del raggiungimento dell'obiettivo associato ad ogni azione. U rappresenta un insieme di funzioni e definisce il criterio di valutazione in base al quale sono valutate le azioni alternative.

- Un decisore (o un gruppo di decisori) provvederanno a selezionare il miglior insieme di azioni $A' \subset A$ che ottimizzano l'utilità dell'utente rispettandone i vincoli. L'insieme ottimo di azioni da proporre all'utente sarà pertanto $U(G(A')) > U(G(A' \cup a_j)) \forall j: a_j \in A - A'$. Ciò significa che ogni altra azione che non contribuisce ad aumentare l'utilità dell'utente è da escludere. Ovviamente la soluzione può anche non esistere oppure ve ne possono essere di equivalenti.

La realizzazione di tale modello è stata progettata utilizzando tecniche ad agenti basati su obiettivi il cui funzionamento può essere descritto utilizzando il modello BDI (Belief, Desires, Intentions), secondo cui un agente possiede: delle credenze (beliefs), ossia assunzioni relative al proprio stato interno e all'ambiente in un determinato istante; dei desideri (desires), cioè degli stati del mondo che vorrebbe che si producessero; delle intenzioni (intentions), ossia quei desideri che sono diventati obiettivi da raggiungere, attraverso la selezione di opportune azioni. Le intenzioni rappresentano lo stato deliberativo dell'agente. Nel modello descritto:

- le percezioni forniscono informazioni riguardo il contesto di decisione,

- il processo di revisione della conoscenza consente l'aggiornamento del profilo dell'utente e quindi dei criteri per la decisione,

- nella fase di discovery l'obiettivo determina il piano da seguire per arrivare alla decisione.

3. Scenari Applicativi

Il primo caso di contesto applicativo è MARA (Mobile Augmented Reality in Archeologia). MARA [Amato et al, 2012b] è un progetto multidisciplinare, con finalità sia culturali che tecnologiche, il cui scopo è la progettazione e realizzazione di una piattaforma context-aware per assistere gli utenti durante la visita nei siti archeologici. Servizi innovativi forniscono contenuti multimediali per i visitatori attraverso i loro dispositivi mobili, al fine di orientare il loro tour e migliorare la loro percezione della realtà. In questo caso il problema della decisione riguarda la scelta del migliore insieme di contenuti multimediali, che ottimizzano la soddisfazione dell'utente, tenendo conto del profilo dell'utente, le capacità del dispositivo, i vincoli di tempo e di spazio disponibile, ecc. In questo caso abbiamo un unico obiettivo da ottimizzare, che è l'affinità semantica tra

l'interesse dell'utente e la serie di contenuti multimediali. La base di conoscenza è costituita da un'ontologia OWL, dalle annotazioni semantiche dei contenuti forniti e dal profilo dell'utente. La base di conoscenza è stata definita da esperti del dominio applicativo del Dipartimento di Studio delle Componenti Culturali del Territorio della Seconda Università degli Studi di Napoli. Il modello di spazio vettoriale è stato utilizzato per rappresentare le annotazioni e profilo con i concetti e gli individui dell'ontologia.

Pervasive Smart Market, presentato in [Amato et al, 2012a], rappresenta il secondo scenario, in cui l'obiettivo è quello di fornire un servizio di e-commerce, che profili i clienti, e supporti i loro acquisti, consigliando prodotti e negozi. In questo caso la il problema di decisione consiste nello scegliere i prodotti da raccomandare in modo che massimizzino la soddisfazione degli utenti, tenendo conto delle loro esigenze, delle loro preferenze, del bilancio, di vincoli di tempo, ecc. Per costruire la base di conoscenza anche in questo caso è stata utilizzata un'ontologia, ma attingendo alla letteratura. In questo caso vengono definiti molteplici obiettivi da ottimizzare quali il tempo per l'acquisto, il budget, le preferenze, la distanza percorsa.

L'ultimo caso risolve un problema di brokering del migliore servizio Cloud che soddisfa i requisiti di utente. La scelta del fornitore di servizi Cloud le cui offerte siano in grado di soddisfare le esigenze di una particolare applicazione, è una problematica complessa a causa dei diversi modelli di business che sono associati a tali sistemi di calcolo, per l'eterogeneità delle tecnologie utilizzate dai providers, per le diverse funzionalità realizzate e per i differenti livelli di servizio garantiti. In questo caso la decisione consiste nel scegliere la proposta che meglio soddisfa le esigenze dell'utente tra quelle disponibili sul mercato. Questo problema è stato modellato come un problema di ottimizzazione multi-obiettivo completo. Per rappresentare le esigenze degli utenti e la proposta del fornitore abbiamo adottato uno schema XML standard per la rappresentazione formale dei Service Level Agreement (SLA). Ogni proposta consiste in un insieme di elementi, ognuno descritto da una serie di parametri standard. Una politica basata su regole consente la valutazione di tutti gli obiettivi per ogni proposta. L'applicazione di brokering è implementata come componente di Cloud Agency [Venticinque, 2012], un sistema multi agente per il provisioning e la gestione di infrastrutture cloud eterogenee.

4. Ambienti pervasivi e Internet of Things

Nei primi due casi di studio è stato utilizzato lo stesso modello decisionale e, per semplicità, si è ipotizzato che gli obiettivi da raggiungere nello smart market sono tutti indipendenti realizzando decisori paralleli, ognuno dei quali propone il suggerimento ottimo per il conseguimento dell'obiettivo di competenza.

4.1 Il framework ad agenti MARA

In Fig. 2 è mostrata la soluzione architeturale del framework MARA realizzata che realizza il supporto alle decisioni nello scenario di valorizzazione del patrimonio archeologico. Il framework si compone di diversi

strumenti e applicativi. Vengono forniti agli esperti del dominio dei Beni Culturali strumenti per arricchire il sito archeologico mediante l'inserimento di una serie di contenuti multimediali. Essi comprendono un editor di mappe, un annotatore semantico e un gestore di contenuti. Gli esperti del dominio applicativo definiscono l'ontologia per il caso di studio specifico. In particolare il loro compito consiste nel progettare una mappa del sito, nella quale vengono indicati Punti di Interesse (POI) per geo-referenziare i contenuti multimediali realizzati o per georeferenziare concetti dell'ontologia stessa.. Lo strumento di annotazione consente di collegare contenuti quali immagini e testi, o parti di essi ai concetti o agli individui della ontologia.

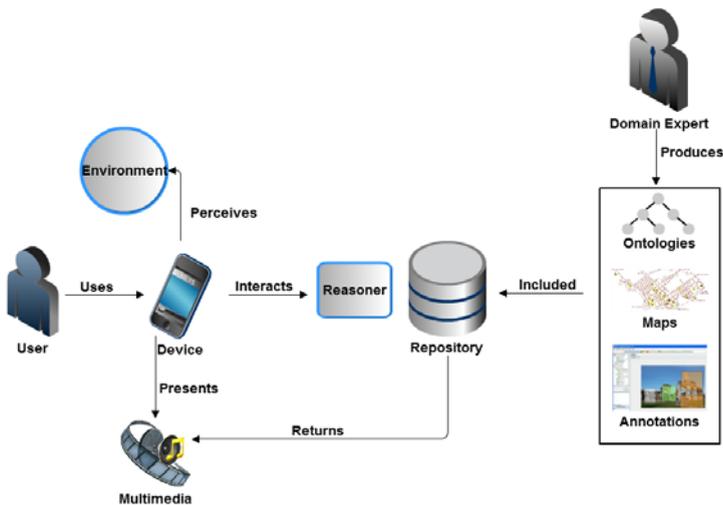


Fig.2 – Architettura del framework MARA

Un insieme di servizi context-aware sono a disposizione per il discovery intelligente e la visualizzazione di contenuti multimediali. Tali servizi utilizzano i dispositivi degli utenti per raccogliere percezioni dall'ambiente attraverso sensori e per la proposizione e la visualizzazione di contenuti. Una vera e propria guida turistica supporta l'utente durante la visita del sito archeologico. Nella Fig. 2, l'utente utilizza il suo dispositivo che ospita un agente in grado di percepire informazioni relative all'ambiente utilizzando sensori pervasivi, o reti di sensori come è stato descritto in [Amato et al, 2012b]. Agenti software eseguono autonomamente e proattivamente sul dispositivo e in remoto supportando l'utente all'interno dell'ambiente in cui si muove. Utilizzano le percezioni per aggiornare la rappresentazione del profilo dell'utente, reagiscono utilizzando la base di conoscenza per organizzare e proporre i contenuti disponibili mediante un'interfaccia interattiva. Se la connessione è disponibile il dispositivo può accedere ai servizi remoti, che utilizzano una

maggiore capacità di ragionamento a al fine di cercare contenuti e applicazioni aggiuntive.

4.2 Pervasive Smart Market

Il secondo caso di studio proposto è un servizio di e-commerce che profila i clienti e li supporta durante gli acquisti, raccomandandogli prodotti e negozi [Amato et al, 2012a]. Il servizio viene alimentato dai venditori, che inseriscono le informazioni circa le loro offerte, i propri negozi e una descrizione degli oggetti in vendita.

Immaginiamo che utenti di diverse età e abitudini, e con scopi diversi, si trovino a dover effettuare degli acquisti. Ognuno di loro ha uno smartphone con un'applicazione intelligente che percepisce l'ambiente, fornisce delle interfacce semplici da utilizzare per accedere ai servizi. Gli utenti utilizzando tale applicazione le consentono di aggiornare il loro profilo nel sistema. Informazioni personali quali nome, sesso, età, paese di origine, indirizzo e-mail possono essere impostate dagli utenti, altre possono essere inferite dall'applicazione stessa. E' possibile inoltre impostare alcune preferenze su quali informazioni il dispositivo può comunicare ai servizi remoti, quali la posizione, gli oggetti vicini, il comportamento dell'utente. I clienti utilizzano il loro smartphone per aggiornare la loro lista dei prodotti da acquistare scegliendo in una vetrina di categorie o di prodotti. I venditori possono aggiungere o rimuovere i loro prodotti in vetrina, e aggiornare la loro descrizione nella base di conoscenza. Per semplificare il problema, si è supposto che i clienti non possano aggiungere, descrivere e consigliare prodotti per conto proprio. Utilizzando le capacità di apprendimento e di inferenza del servizio, l'obiettivo del sistema è quello di proporre la migliore serie di raccomandazioni, che può aiutare l'utente a migliorare la propria utilità, sfruttando le informazioni disponibili sui prodotti, l'ambiente all'interno del quale i clienti si stanno muovendo e riguardo le preferenze e le abitudini dei clienti stessi.

4.3 Il modello decisionale

Per ragioni di semplicità, si suppone che le percezioni notificate dal dispositivo ai servizi remoti si limitano alla posizione dell'utente, al tempo a disposizione per gli acquisti o per la visita, ed ai concetti dell'ontologia che descrivono il suo profilo. Ogni volta che una nuova percezione è stata ricevuta, il profilo dell'utente viene valutato al fine di eventualmente attivare il servizio di discovery che aggiorna il set ottimale di contenuti da proporre. Il servizio di discovery semantico restituisce un insieme di riferimenti ad oggetti digitali. Ogni oggetto è annotato per mezzo di concetti dell'ontologia e può essere recuperato mediante una query SPARQL (SPARQL Protocol and RDF Query Language), un linguaggio di interrogazione che consente di estrarre informazioni dalle basi di conoscenza. Il risultato della query è un insieme di N istanze di oggetti digitali cui rilevanza per l'utente viene ottenuta calcolando una metrica di distanza di ciascun contenuto dal profilo p , che è la rappresentazione semantica del contesto, degli interessi dell'utente, della posizione, ecc. L'annotazione e il

profilo dell'utente vengono rappresentati utilizzando il Vector Space Model (VSM). VSM è un modello di rappresentazione semantica come vettori di oggetti definito da G. Salton in [Salton e Lesk, 1968]. Nel nostro caso le componenti del vettore sono i concetti $c_{i,j}$ dell'ontologia di dominio.

$$\vec{a}_j = \langle \{c_{1,j}, o_{1,j}\}, \{c_{2,j}, o_{2,j}\}, \dots, \{c_{l,j}, o_{l,j}\} \rangle \quad \forall j = 1 \dots N$$

$$\vec{p} = \langle c_1, c_2, \dots, c_m \rangle$$

I valori l ed m rappresentano il numero di concetti \vec{a}_j differenti che appaiono nell'annotazione e nel profilo utente. Se un termine compare nell'annotazione, $o_{k,j}$ rappresenta il numero di occorrenze di quel termine nell'annotazione. Abbiamo poi definito un punteggio $A(\vec{a}_j, \vec{p})$ per misurare la rilevanza di un annotazione \vec{a}_j rispetto al profilo \vec{p} mediante la seguente formula:

$$w_j = w(\vec{a}_j, \vec{p}) = \sum_{k=1}^l r_k \quad \text{dove} \quad r_k = o_{k,j} * \sum_{i=1}^m \frac{1}{d_{k,i} + 1}$$

dove $d_{k,i}$ è il numero minimo di archi che connettono il nodo che rappresenta il concetto $c_{k,j}$ dell'annotazione al concetto c_i del profilo. Per ogni elemento $c_{k,j}$ del vettore \vec{a}_j la rilevanza rispetto al profilo \vec{p} è calcolata aggiungendo la rilevanza di quel concetto rispetto ad ogni concetto del profilo, e moltiplicando ognuno di questi contributi per il numero di occorrenze $o_{k,j}$. La rilevanza tra due concetti è calcolata dividendo 1 per il numero di archi dell'ontologia che connettono il nodo che rappresenta il concetto c_k del profilo a $c_{k,j} + 1$. Come risultato abbiamo un punteggio per ogni contenuto annotato associato ad un POI in modo tale da poter ordinare contenuti e POI in base alle preferenze dell'utente. Tuttavia, un utente ha alcuni vincoli aggiuntivi come il tempo disponibile, la distanza massima che può percorrere, la capacità del dispositivo di riprodurre alcuni tipi di contenuto, oppure il prezzo e la frequenza di acquisto. Per questo motivo è necessario selezionare i contenuti validi escludendo, per esempio, quelli che non possono essere erogati dal dispositivo dell'utente ed in generale quelli che non rispettano i vincoli dell'utente. Quindi, il miglior insieme di contenuti da erogare, ovvero l'obiettivo, dovrebbe: massimizzare il punteggio, essere compatibile con i vincoli dell'utente. I vincoli sono espressi dalle seguenti disequazioni:

$$\sum_{i=1}^n b_{i,j} x_i \leq a_j \quad \forall j \in \{1 \dots m\}$$

che definiscono lo spazio di soluzione del problema. Nel nostro caso si ha:

$$w_i \geq 0, \quad \forall j = 1 \dots N \quad W = \{\vec{b}_1, \dots, \vec{b}_m\}$$

dove w_i rappresenta il punteggio e B un insieme di vincoli \vec{b}_i ognuno composto da $N + 1$ interi. E' necessario quindi calcolare:

$$\max \sum_{k=1}^N w_i x_i \quad \text{tale che} \quad \sum_{i=1}^N b_{i,j} x_i \leq b_{j,N+1} \quad \text{con } x_i \in \{0, 1\} \quad \forall i = 1 \dots N$$

Il vettore \vec{x} rappresenta una possibile soluzione. Le sue componenti x_i sono 1 se l'oggetto corrispondente è incluso nell'insieme dei contenuti migliori, 0 altrimenti.

5. Cloud Broker

Il problema del Brokering di risorse Cloud [Amato e Venticinque, 2013], illustrato in Fig. 3, consiste nello scegliere la migliore proposta tra il numero di offerte ricevute da differenti provider che forniscono offerte differenti ed eterogenee sia dal punto di vista delle tecnologie, sia dal punto di vista dei termini e delle condizioni. Pertanto spesso non è possibile stabilire metriche di comparazione. I requisiti richiesti dall'utente sono espressi in un documento chiamato Call for Proposal (CFP) che si compone di:

- *SLA template*: è un documento che esprime la configurazione delle risorse che sono necessarie all'utente e consiste di un insieme di servizi che possono includere risorse di calcolo, rete e storage, le loro caratteristiche di interesse, livelli e termini di servizio. Viene realizzato usando come modello di rappresentazione `sla@soi` (`sla-at-soi.eu`);

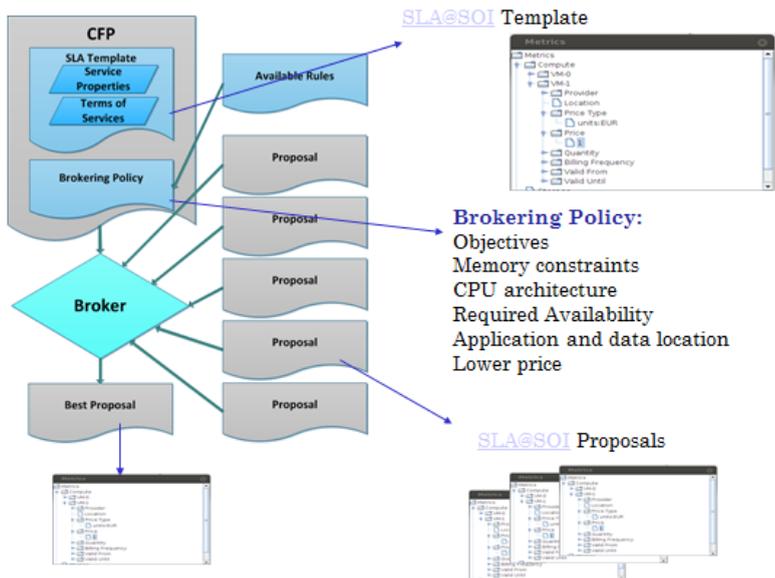


Fig.3 – Brokering

-*Brokering policy*: un documento che contiene un insieme di regole per la scelta della migliore proposta (requisiti) come ad esempio il miglior prezzo, il maggior numero di core, il miglior provider accreditato l'availability minima accettata.

L'utente quindi seleziona le caratteristiche del servizio richiesto, i livelli di servizio in termini di prestazioni, disponibilità, costo, ecc. Per ciascuna caratteristica l'utente definirà un vincolo ,specificando se esso debba essere obbligatorio o facoltativo. Definerà inoltre per uno o più parametri del servizio una funzione obiettivo da ottimizzare. In risposta all'SLA template le proposte vengono generate dai Cloud Vendor complementando i termini dello SLA template $T=\{t_1, \dots, t_n\}$ con i corrispondenti valori offerti: $P_j=\{(t_1, v_{j1}), \dots, (t_n, v_{jn})\}$. Per risolvere il problema del brokering abbiamo definito un modello multicriterio in cui, per ogni proposta ricevuta il broker valuterà se questa può essere presa in considerazione (cioè se non viola nessun vincolo obbligatorio), ed i vincoli facoltativi da essa soddisfatti.

Nel dettaglio il broker utilizza un insieme di regole $R = C \cup O$ che possono rappresentare un insieme di vincoli C con $cr \in C$ e/o un insieme di obiettivi O con $or \in O$. Vincoli e obiettivi sono definiti da:

$$cr : (t_i, c_i, m_i) \rightarrow [0,1] \quad or : (t_i, f_i, o_i) \rightarrow [0,1]$$

Dove t_i è un termine dello SLA, c_i è un'espressione booleana, m_i specifica se il vincolo è obbligatorio (0) oppure facoltativo (1), f_i è una funzione che definisce per ogni valore t_i il corrispondente punteggio e o_i specifica se l'obiettivo è esplicito ovvero da negoziare separatamente (1) oppure implicito (0). Per ogni proposta ricevuta il broker esclude quelle che non rispettano i vincoli obbligatori ovvero:

$$M_j = \prod_{i=1}^n (c_i \vee \bar{m}_i) \quad \forall j = 1, \dots, m$$

Il broker cerca in un insieme di proposte $P=\{p_1, \dots, p_n\}$ quelle che massimizzano gli obiettivi :

$$\max_{j=0}^m (or_j) : o_i = true \quad \wedge \quad M_j = 1 \quad \forall i = 1, n \quad \forall j = 1, m$$

massimizzano la valutazione complessiva

$$\max_{j=0}^m (V_j) : M_j = 1 \quad \forall j = 1, m \quad V_j = \sum ((-o_i) * f_i(v_{j,i})) \quad \forall j = 1, \dots, m$$

e massimizzano il numero di vincoli preferiti:

$$\max_{j=0}^m (Opt_j) : M_j = 1 \quad Opt_j = \sum_{i=1}^n (\neg m_i * (cr_i)) \quad \forall j = 1, \dots, m$$

6. Conclusioni

In questo lavoro è stata presentata una metodologia per la progettazione e lo sviluppo di sistemi multi-agente per la risoluzione di problemi decisionali, la cui complessità è caratterizzata da diversi criteri che devono essere valutati e confrontati in modo indipendente e in base alle preferenze dell'utente/agente. Abbiamo presentato i risultati dell'applicazione della metodologia proposta in tre diversi casi di studio di crescente complessità. Come si è visto l'obiettivo finale della metodologia proposta è quello di supportare lo sviluppatore nella realizzazione di un modello di ragionamento che consente agli agenti di prendere decisioni migliori in presenza di incertezza e di abbondanza di

alternative al fine di orientare il processo decisionale. In tutti i casi di studio proposti la decisione è stata modellata da una serie di vincoli e obiettivi utilizzati per valutare i differenti elementi appartenenti allo spazio di decisione.

Successivamente è stato calcolato un punteggio per ciascun obiettivo per fornire informazioni che consentono ai decisori di scegliere tra le azioni rilevanti che sono a disposizione per perseguire gli obiettivi di interesse. Diverse tecnologie sono state utilizzate sia per la rappresentazione e la valutazione delle alternative disponibili nell'ambito di scenari appartenenti al Cloud e all'Internet of Things (IoT). Attività sperimentali sono in fase di studio per verificare il comportamento del sistema, all'aumentare del numero dei soggetti interessati, del numero di criteri e della dimensione del processo decisionale.

Bibliografia

[Amato et al, 2012a] Amato A, Di Martino B, Venticinque S., A Semantic Framework for Delivery of Context-Aware Ubiquitous Services in Pervasive Environments, Proc. of the 4th International Conference on Intelligent Networking and Collaborative Systems (INCoS), 2012, 412-419.

[Amato et al, 2012b] Amato A, Di Martino B., Venticinque S., Semantically augmented exploitation of pervasive environments by intelligent agents. Proc. of the 10th IEEE International Symposium on Parallel and Distributed Processing with Applications, 2012, 807-814.

[Amato e Venticinque, 2013] Amato, A., Venticinque, S.: Multi-objective decision support for brokering of cloud sla. Proc. of the 27th IEEE International Conference on Advanced Information Networking and Applications (AINA), 2013, 1241-1246.

[Salton e Lesk, 1968] G. Salton and M. E. Lesk, "Computer evaluation of indexing and text processing," J. ACM, vol. 15, no. 1, 1968, 8–36.

[Venticinque, 2012] Venticinque S., Agent Based Services for Negotiation, Monitoring and Reconfiguration of Cloud Resources, European Research Activities in Cloud Computing, Cambridge Scholars, 2012, 178-202.

Optimizing a static greedy algorithm for influence maximization

Leonardo Capone, Nicola Di Mauro, Floriana Esposito
Università degli Studi di Bari "Aldo Moro"
Via Orabona, 4, 70125 Bari

leonardocapone89@gmail.com, nicola.dimauro@uniba.it, floriana.esposito@uniba.it

One of the main problem in social networks and viral marketing is that of finding a set of nodes maximizing the spread of influence. Corresponding algorithms solving this problem are required to have both guaranteed accuracy and high scalability. Greedy algorithms are able to find accurate solutions but fail in efficiency. This paper presents a modification of an existing greedy algorithm to solve the influence maximization problem by integrating a memoization technique. Experimental results with a first prototypical implementation on real-world social networks proved the validity of the proposed technique.

Keywords: *influence maximization, memoization, social networks*

1. Introduction

Nowadays people are connected by heterogeneous social relationships in large-scale online social networks that provide a platform for information dissemination and marketing. The success of viral marketing is rooted in the interpersonal influence empirically studied in [Richardson et al., 2002; Huang et al., 2012]. With the advent of social networks the information spreads in the form of “word-of-mouth” communications, and it is noticeable to observe how much they affect our daily life style.

Influence maximization is a fundamental problem for viral marketing and it has been originally formulated as an optimization problem in [Kempe et al., 2003]. It consists in finding a set of seed nodes which maximize the *influence spread* in a social network computed as the expected number of nodes influenced by the seed nodes. In [Kempe et al., 2003] has been proved the NP-completeness of the influence maximization problem and it has been provided a greedy approximation algorithm that yields an influence spread solution that is no less than a given bound of the optimal value.

Given a seed set of nodes there is no exact algorithm that gives the corresponding influence spread. Usually it is approximated using a large

number of Monte Carlo simulations. However, this reduces the scalability of the general greedy algorithm proposed in [Kempe et al., 2003] since it requires too many Monte Carlo simulations. To overcome this problem one can reduce the times of influence spread estimations [Leskovec et al., 2007, Cheng et al., 2012], or proposes various heuristics to use more efficient methods for influence spread estimation [Chen et al., 2010]. The hot interest in scalable and accurate methods to solve the influence maximization problem is confirmed by a lot of recent works [Chen et al., 2009; Goyal et al., 2011; Leskovec et al., 2007; Jiang et al., 2011; Kimura et al., 2010].

Two widely used information diffusion models to solve the *influence maximization problem* [Kempe et al., 2003] are the *independent cascade* (IC) [Kempe et al., 2003] and the *linear threshold* (LT) [Watts, 2002] models. The IC model is sender-centered and each active node independently influences its inactive neighbors with given diffusion probabilities, while the LT model is receiver-centered and a node is influenced by its active neighbors if their total weight exceeds the threshold for the node.

In this paper, we provide an optimization of an existing greedy algorithm [Cheng et al., 2012] with its corresponding evaluation adopting the IC model on two real world social networks.

2. Preliminaries

In order to mathematically model the information diffusion in a social network, we firstly recall the IC and LT models according to [Kempe et al., 2003].

Let $G = (V, E)$ be a directed network (graph) where V is a set of nodes and $E \subset V \times V$ is a set of directed links. For each node $v \in V$, let $F(v)$ be the set of all the nodes that have links from v (its *children nodes*), i.e., $F(v) = \{u \in V; (v, u) \in E\}$, and $B(v)$ be the set of all the nodes that have links to v (its *parent nodes*), i.e., $B(v) = \{u \in V; (u, v) \in E\}$. We say a node is *active* if it has been influenced with the information. In the IC and LT models the diffusion processes unfold in discrete time steps, and it is assumed that nodes can switch their state only from inactive to active.

2.1 Independent Cascade Model

In the IC model it is necessary to define in advance a probability $p_{u,v}$, called *propagation probability*, for each directed link. Given an initial set A of active nodes the diffusion process works as follows. If a node u is active at step t , it has a chance to activate each currently inactive node $v \in F(u)$ with probability $p_{u,v}$. Whether or not u succeeds in activating v , it cannot make any further attempts to activate v in subsequent time steps and the overall process terminates if no more activations are possible.

2.2 Linear Threshold Model

In the LT model, for any node $v \in V$, it is necessary to specify a weight $w_{u,v} > 0$ for each of its parent node u , such that their sum is lesser or equal to 1.

When an initial set A of active nodes is given, and a threshold θ_v for each node v is set to be uniformly distributed in the interval $[0,1]$, the diffusion process works as follows. An inactive node v at time step t is influenced by each active parent node u according to the weight $w_{u,v}$. If the total weight from active parents is greater than θ_v , then v will be active at time step $t+1$. The overall process terminates if no more activations are possible.

Given an initial active set A , let $\sigma(A)$ denote the expected number of active nodes at the end of the random process in the IC or in the LT model. We call $\sigma(A)$ the *influence degree* of the set A .

2.3 Influence Maximization Problem

The influence maximization problem is defined as follows. Given a positive integer k , find a set A of k nodes such that, for each set B of k nodes, $\sigma(A) \geq \sigma(B)$.

A greedy algorithm that approximately solves this problem has been proposed in [Kempe et al., 2003], and it is sketched in the following.

1. Set $A \leftarrow \emptyset$
2. **for** $i=1$ to k **do**
3. choose a node $v_i \in F$ maximizing $\sigma(A \cup \{v_i\})$
4. set $A \leftarrow A \cup \{v_i\}$
5. **endfor**

In [Kempe et al., 2003], this algorithm has been proved to obtain an approximate solution whose value is at least $(1 - 1/e) \sigma(A^*)$, where A^* is the optimal solution. This factor has been obtained by proving that the influence function $\sigma(\cdot)$ is *submodular* and using a result obtained in [Nemhauser et al., 1978]. Formally, a submodular function satisfies:

$$f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T),$$

for all elements v and all pairs of sets $S \subseteq T$.

2.4 Approximate Influence Degree

However, how to exactly computing $\sigma(A)$ is not known and it is approximated using a large amount of Monte Carlo simulations [Kempe et al., 2003], thus degrading the efficiency of the algorithm.

In particular, the Monte Carlo approach used to approximate the influence degree $\sigma(A)$ in the IC model works as follows. Let A_t be the set of nodes that are activated in the time step t , and $A_0 = S$. For any link $(u,v) \in E$ such that $u \in A_t$ and v is not yet activated, then v is activated by u in the time step $t + 1$ with the propagation probability $p_{u,v}$. This process is repeated until A_{t+1} is empty.

The random process to estimate the influence degree in the LT model is quite similar to that used for the IC model but taking into account that the probability of u activating v is usually not the same as the probability of v activating u , thus requiring a slightly modification.

3.The new optimized algorithm

A recent interesting approach to overcome the problem to use a large amount of Monte Carlo simulations has been proposed in [Cheng et al., 2012]. The authors proved that the submodularity is not guaranteed in existing implementations of greedy algorithm, caused by the independence among Monte Carlo simulations executed in different iterations of the greedy algorithm. They proposed a static greedy algorithm to strictly guarantee the submodularity property, by reusing the results of Monte Carlo simulations during the whole process of greedy algorithm. The results is to dramatically reduce the random simulations thus effectively improving the scalability of the greedy approach.

In particular the authors of [Cheng et al., 2012] introduced the concept of *snapshot* obtained a priori according to the characteristic of the IC model. A snapshot is a graph G' obtained from the original graph G , where an edge (u, v) is removed with probability $1-p_{u,v}$. Then, for each snapshot G' , the influence spread of a set of nodes S is the number of nodes reachable from S . Hence, the influence degree $\sigma(S)$ can be obtained by averaging over many snapshots.

The process of the corresponding static greedy algorithm is the following:

- 4.randomly sampling R snapshots from the underlying social network G ;
- 5.start from an empty seed set S , then iteratively add one node a time into S such that the node provides the largest marginal gain of $\sigma(S)$, which is estimated on the R snapshots.

In particular, the static greedy algorithm, named Static, is formalized as follows, where $R(S)$ is a function returning the nodes reachable from the nodes in S :

1. initialize $S = \emptyset$
2. **for** $i = 1$ to R **do**
3. generate G'_i by removing each edge (u, v) from G with probability $1-p_{u,v}$
4. **endfor**
5. **for** $j = 1$ to k **do**
6. set $sv = 0$ for all $v \in V \setminus S$
7. **for** $j = 1$ to R **do**
8. for all $v \in V \setminus S$ **do**
9. $sv += |R(S \cup \{v\})|$
10. **endfor**
11. **endfor**
12. $S = S \cup \{\text{argmax}_{v \in V \setminus S} \{ sv / R \}\}$
13. **endfor**
- 14 output S

In this paper, we tried to improve this static greedy algorithm by adopting a memoization approach. In particular, at time step t , we have to compute $R(S \cup \{v\})$ for each node $v \in V \setminus S$. However we can notice that the nodes reachable from the set $\{S \cup \{v\}\}$ corresponds to the nodes reachable from S plus the nodes reachable from v and not already reached from S . More

formally, let S be equal to the nodes $\{s_1, s_2, \dots, s_n\}$, then:

$$\begin{aligned} |R(S \cup \{v\})| &= |R(S) \cup R(\{v\})| = \\ &|R(\{s_1, s_2, \dots, s_n\}) \cup R(\{v\})| = \\ &|R(\{s_1\}) \cup R(\{s_2\}) \dots \cup \dots R(\{s_n\}) \cup R(\{v\})|. \end{aligned}$$

Hence, at each time step t , instead of computing $R(S \cup \{v\})$ we can use its decomposition and take advantage of the previous computations. Indeed, for each snapshot, before to start the iterative process, we can compute the sets $R(\{v\})$ for each node of the snapshot. Then in the iterative process we can exploit this sets and the basic union operations to compute the influence spread. The static greedy algorithm exploiting memoization, named StaticM, becomes:

1. initialize $S = \emptyset, Q = \emptyset$
2. **for** $i = 1$ to R **do**
3. generate G_i' by removing each edge (u, v) from G with probability $1-p_{u,v}$
4. **endfor**
5. **for** $i = 1$ to k **do**
6. set $sv = 0$ for all $v \in V \setminus S$
7. **for** $j = 1$ to R **do**
8. for all $v \in V \setminus S$ **do**
9. $sv += |Q \cup R(\{v\})|$
10. **endfor**
11. **endfor**
12. $S = S \cup \{\operatorname{argmax}_{v \in V \setminus S} \{sv / R\}\}$
13. $Q = Q \cup R(\{v\})$
13. **endfor**
- 14 output S

The set Q in the algorithm represents the set of nodes reachable from the best chosen k nodes at the previous time step. It is enlarged, at each iteration, with the nodes $R(\{v\})$ for that node v such that the cardinality of $R(S \cup \{v\})$ is maximized.

4. Experimental evaluation

In this section we use two real social network datasets, such as Epinions and Slashdot datasets available at <http://http://snap.stanford.edu/data/>, in order to evaluate the efficacy of our proposed approach improving the static greedy algorithm.

The Epinions dataset [Richardson et al., 2003] concerns a *who-trust-whom* online social network of a a general consumer review site Epinions.com. Members of the site can decide whether to “trust” each other. All the trust relationships interact and form the Web of Trust which is then combined with review ratings to determine which reviews are shown to the user. The corresponding graph consists of 75879 nodes and 508837 edges.

Slashdot is a technology-related news website know for its specific user community. The website features user-submitted and editor-evaluated current

primarily technology oriented news. In 2002 Slashdot introduced the Slashdot Zoo feature which allows users to tag each other as friends or foes. The Slashdot dataset [Leskovec et al., 2009] network contains friend/foe links between the users of Slashdot. The network was obtained in February 2009 and consists of 82168 nodes and 948464 edges.

For both the datasets we used the IC model as a diffusion model and we set the propagation probability p to 0.5.

We compared the new StaticM algorithm against the Static and Random algorithm. The Random algorithm simply selects k seed nodes among the possible and computes its corresponding influence spread.

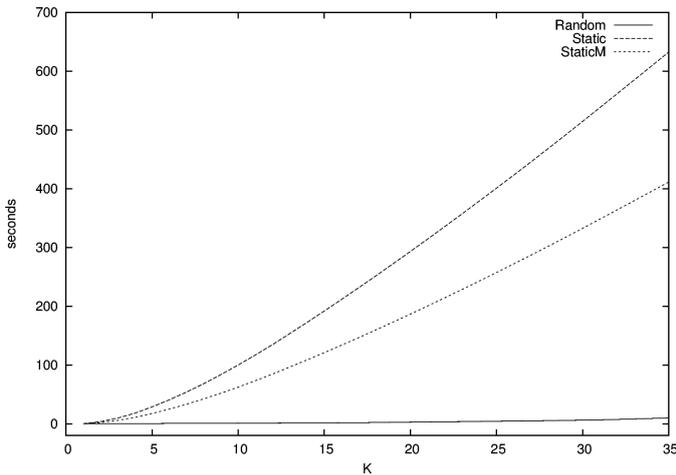


Fig.1 – Time in seconds on the Epinions dataset for Random, Static and StaticM.

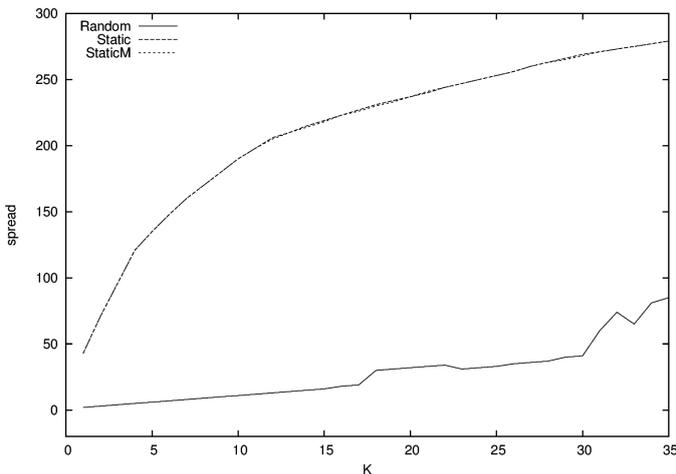


Fig.2 – Influence spread on the Epinions dataset for Random, Static and StaticM.

As we can see from Figure 1 StaticM requires a time lesser than that required by Static to compute the influence spread on the Epinions dataset. Random has the best performances since it just randomly picks nodes from the available ones. Figure 2 plots the influence spread for k ranging from 1 to 35 obtained by the algorithms. Both Static and StaticM reach, for each k , the same value and both obviously find solutions better than those obtained with Random. Similar results are obtained on the Slashdot dataset as reported in Figure 3 and Figure 4.

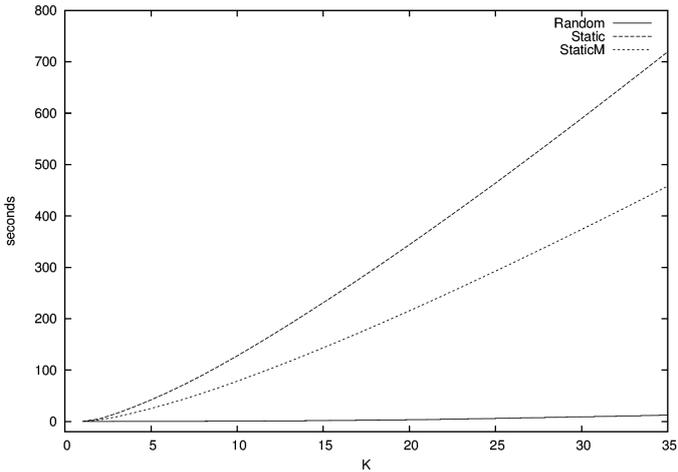


Fig.3 – Time in seconds on the Slashdot dataset for Random, Static and StaticM.

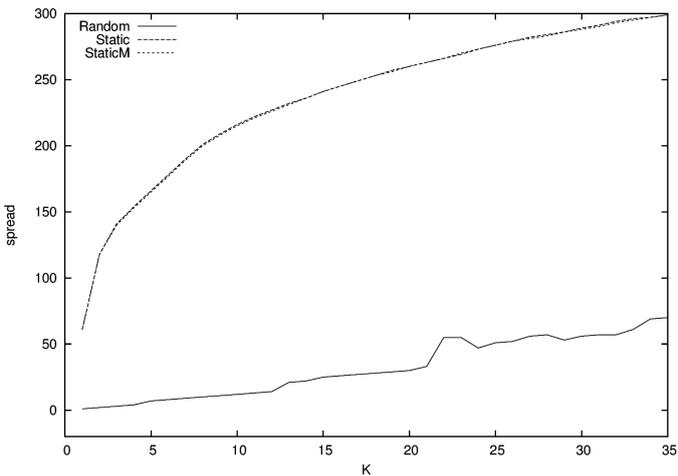


Fig.4 – Influence spread on the Slashdot dataset for Random, Static and StaticM.

5. Conclusions

One of the main problem in social networks and viral marketing is that of finding a set of nodes maximizing the spread of influence. Algorithms for solving this problem are required to have both guaranteed accuracy and high scalability. The proposed greedy algorithms are able to find accurate solutions but fail in efficiency. In this paper we have proposed a modification of a static greedy algorithm by integrating a memoization technique. Experimental results on the Epinions and Slashdot real world datasets exploited with a prototypical implementation proved the validity of the proposed technique.

Acknowledgements

This work fulfills the research objectives of the PON02_00563_3489339 project "PUGLIA@SERVICE - L'Ingegneria dei Servizi Internet-Based per lo sviluppo strutturale di un territorio "intelligente" funded by the Italian Ministry of University and Research (MIUR).

Bibliography

[Chen et al., 2009] Chen W., Wang Y., Yang S., Efficient influence maximization in social networks, in Proceedings of the 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2009, 199-207.

[Chen et al., 2010] Chen W., Wang C., Wang Y., Scalable influence maximization for prevalent viral marketing in large-scale social networks, in Proceedings of the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2010, 1029-1038.

[Cheng et al., 2012] Cheng S., Shen H., Huang J., Zhang G., Cheng X., Static greedy: solving the apparent scalability-accuracy dilemma in influence maximization. CoRR abs/1212.4779, 2012.

[Goyal et al., 2011] Goyal A., Lu W., Lakshmanan L.V.S., CELF++: optimizing the greedy algorithm for influence maximization in social networks, in Proceedings of the 20th International Conference on World Wide Web, 2011, 47-48.

[Huang et al., 2012] Huang J., Cheng X.Q., Shen H.W., Zhou T. and Jin X., Exploring social influence via posterior effect of word-of-mouth recommendations, in Proceedings of the 5th ACM International Conference on Web Search and Data Mining, 2012, 573-582.

[Jiang et al., 2011] Jiang Q., Song G., Cong G., Wang Y., Si W., Xie K., Simulated annealing based influence maximization in social networks, in Proceedings of the 25th AAAI Conference on Artificial Intelligence, 2011, 127-132.

[Kempe et al., 2003] Kempe D., Kleinberg J., Tardos E., Maximizing the spread of influence through a social network, in proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining, 2003, 137-146.

[Kimura et al., 2010] Kimura M., Saito K., Nakano R., Motoda H., Extracting influential nodes on a social network for information diffusion, Data Mining and Knowledge Discovery, 20(1), 2010, 70-97.

[Leskovec et al., 2007] Leskovec J., Krause A., Guestrin C., Faloutsos C., VanBriesen J., Glance N.S., Cost-effective outbreak detection in networks, in Proceedings of the 13th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2007, 420-429.

[Leskovec et al., 2009] Leskovec J., Lang K., Dasgupta A., Mahoney M., Community Structure in Large Networks: Natural Cluster Sizes and the Absence of Large Well-Defined Clusters, Internet Mathematics 6(1), 2009, 29-123.

[Nemhauser et al., 1978] Nemhauser G., Wolsey L., Fisher M., An analysis of the approximations for maximizing submodular set functions. Mathematical Programming, 14, 1978, 265–294.

[Richardson et al., 2002] Richardson, M. and Domingos, P., Mining knowledge-sharing sites for viral marketing, Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, 2002, 61-70.

[Richardson et al., 2003] Richardson M., Agrawal R., Domingos P., Trust Management for the Semantic Web, The Semantic Web - ISWC 2003, 2003, 351-368.

[Watts, 2002] Watts D.J., A simple model of global cascades on random networks, PNAS 99, 2002, 5766–5771.

Social and Smart Mobility for Future Cities: the S²-Move project

Davide De Pasquale¹, Pietro Marchetta², Eduard Natale²
Alessandro Salvi², Antonio Tirri², Manuela Tufo¹

¹University of Sannio
Piazza Guerrazzi 1, 82100 Benevento (Italy)
davide.depasquale@studenti.unisannio.it,
manuela.tufo@unisannio.it

²University of Napoli Federico II
Via Claudio 21, 80125 Napoli (Italy)
{pietro.marchetta, alessandro.salvi}@unina.it
{ed.natale, an.tirri}@studenti.unina.it

Abstract. *The main idea behind S2-Move, a social innovation project financed by the Italian Ministry of Education, Research and University, is to supply soft real-time information exchange among citizens, public administrations and transportation systems by exploiting Information and Communication Technologies. In this paper, we describe the technological S2-Move infrastructure and the mobility services used to demonstrate the potentialities of the project.*

Keywords: smart mobility, fleet control, traffic monitoring.

1. The project S²-Move

The main idea behind S2-Move [Marchetta et al., 2012], a project financed by the Italian Ministry of Education, Research and University, is to supply soft-realtime information exchange among citizens, public administrations and transportation systems. As shown in literature [Schaffers et al., 2011], to this aim ICT plays an essential role and, currently, several projects investigate the potentialities of similar approaches ([Bakıcı et al., 2012] and references therein). In our view, the urban mobility represents a complex and ever-evolving ecosystem that can be just partially measured through a widely and well distributed set of probes. In this scenario, it is essential to exploit at best the forcedly limited information collected from the urban environment: a set of different types of elaboration may allow to gather accurate knowledge on the current status of the urban mobility and supply citizens with a set of different services. In order to maximize the quality of experience of the citizens as well as the impact of the project, we identified the mechanism of customized maps as the most user-friendly and intuitive approach to supply urban mobility services. In order to meet this underlying philosophy, we designed an architecture consisting of two main components, described as follows. The urban probes represent a heterogeneous set of devices/sensors deployed in the urban environment. Within the project, we also designed and developed a prototype of On Board Unit (hereafter simply OBU): a smart electronic device, installed into a vehicle, that is able to collect and send information, to process

data by applying filtering techniques and communicate with the other components of the architecture. The OBU is responsible for both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. In particular, this device is able, through the CAN (Controller Area Network) bus, to take a set of information from the different electronic control units in vehicles. By analysing the data collected from the CAN bus, it is possible to monitor the vehicle and, indirectly, the urban environment by crossing the data provided by multiple sources of information. Finally, we are currently implementing mobile apps to encourage citizens to actively collaborate toward a more effective and efficient urban mobility by reporting mobility-related events to the central system. The Central Processing System (hereafter simply CPS) represents the core of our architecture. This component has three main responsibilities: (a.) it permanently stores the information collected by the urban probes (Data layer); (b.) it hosts all the data-processing algorithms part of the services at the basis of the project (Core layer); (c.) it implements all the mechanisms for the interaction with the final users through customized maps (Presentation layer). In our current prototype, we adopted the Play! Java Framework [Reelsen, 2011] for the implementation of most modules, Geoserver [Deoliveria, 2008] to generate annotated maps, and PostgreSQL [Stonebraker,1986] extended with PostGIS to persistently store the collected data. The data cycle of the project can be described as follows. A generic urban probe collects an information about a parameter of interest (e.g. a car park has been released), this bit of information is pre-processed by the urban probe and sent to the CPS via the Internet, to be permanently stored by the data layer. From now on, this information is available to all the mobility services to be exploited for totally different purposes, such as updating the information about a specific parking lot. Few moments later, when the map manager refreshes the map, the number of free parks is updated.

2. S²-Move Applications

While the platform at the basis of the project is purposely defined to be easily extended with always new and more sophisticated services, we considered a set of applications as a valuable proof of concept of the potentialities of our approach. In particular, in this work we describe Traffic Monitoring and Fleet Control.

Traffic monitoring aims at determining an accurate and complete knowledge on the traffic in the city by exploiting the limited information collected from the urban environment. Since traditional traffic monitoring systems based on heterogeneous traffic detectors (inductive loops, video camera, etc.) are particularly expensive, the S²-Move traffic monitoring service mainly relies on the information provided by OBU devices and collected thanks to the GPS (Global Positioning System), an approach already proposed in literature [Barbagli et al., 2011] [Hadachi, 2012] [Yang et al., 2005]. The idea is to infer the traffic condition by observing the speed of the monitored vehicles along the urban routes. To this end, a data collection module hosted on the OBU device samples and filters kinematics information coming from the CAN bus. On the CPS, two main modules are responsible for the data processing phase. The

Map Matching Module associates data and coordinates to the involved streets. Currently the algorithms used in project are based on geometric relations between the GPS coordinates and the position of the closest street. The Speed Estimation Module extracts a representative value for each street from the speed occurrences provided by the monitored vehicles travelling across the city.

Fleet management can be divided in two main action: Fleet Monitoring and Fleet Control. Fleet monitoring aims at tracking group of vehicles moving in the urban environment. Fleet Control, instead, allows to manage vehicle fleets such that vehicles travel in the same direction, one behind another; moreover vehicles are equipped with a longitudinal control law in order to guarantee that vehicles travel with a common velocity and a predefined intra-vehicular distance [Hedrick et al., 1994]. The platform includes the Communication Module, letting the on-board collected data to be shared between all the vehicles in the fleet (i.e. GPS, target position and speed) and includes both V2I [Belanovic et al., 2010] and V2V [Sichitiu, 2008]) communications. In this context, heterogeneous wireless communication technologies and their performance and quality of service must be carefully taken into account [Bernaschi et al., 2005],[Iannello et al., 2004],[Karres et al., 2006],[Botta et al., 2008]. The Coordination Module coordinates operations between adjacent vehicles in order to guarantee the safety for drivers. The Regulation Module collects and processes information coming from the on-board sensors in order to compute the right control effort on the vehicle actuators (i.e. throttle, braking and steering control). A more detailed description of the control law used in the simulated scenario can be found in [Marchetta et al., 2012]. A possible advancement of this control strategy is under investigation [Di Bernardo et al., 2010] (recently used to control mechanical automotive and power electronic systems in canonical forms [Di Bernardo et al., 2008],[Salvi et al., 2013]) to guarantee robustness with respect to noise and non modeled dynamics as well as rapid adaptation to changes of road conditions. Finally, the Vehicle Dynamics Module is the physical layer providing information about the vehicle motion.

3. Conclusion

The social innovation projects financed by the Italian Ministry of Education demonstrate a particular attention to the quality of life in the urban environment. In this context, S2-Move aims to contribute towards a more effective and efficient urban mobility thanks to an ICT-powered architecture able to collect and process soft-real time data. By using the collected information and applying the most innovative data analysis techniques, it is possible to derive new information that may help citizens and public administrations to develop new management strategies. The modular architecture behind the project is designed to allow the integration of new mobility services that, using the information gathered and made available at the storage level, would be able to determine new information for the end user. Accordingly, the services described in the work aim at demonstrating the potentialities of this approach.

Acknowledgement. The activities described in this paper are funded by MIUR (PON04a3_00058). Authors thank the University of Napoli Federico II for lodging some S2-Move activities, and to D. Di Nocera and A. S. Valente for their work.

References

- [Bakıcı et al., 2012] Bakıcı, T., Almirall, E., and Wareham, J. (2012). "A smart city initiative: the case of Barcelona". *Journal of the Knowledge Economy*, 1-14.
- [Barbagli et al., 2011] Barbagli B., Bencini L., Magrini I., Manes G. and Manes A, "A real-time traffic monitoring based on wireless sensor network technologies", IWCMC 2011.
- [Belanovic et al., 2010] Belanovic, P., Valerio, D., Paier, A., Zemen, T., Ricciato, F., and Mecklenbrauker, C. F., "On wireless links for vehicle-to-infrastructure communications", IEEE TVT 2010.
- [Bernaschi et al., 2005] Bernaschi M., Cacace F., Pescapè A., and Za S., "Analysis and Experimentation over Heterogeneous Wireless Networks", TRIDENTCOM 2005.
- [Botta et al., 2008] Botta A., Pescapè A., Ventre G., 'Quality of Service Statistics over Heterogeneous Networks: Analysis and Applications', Performance Evaluation of QoS-aware Heterogeneous Systems, Elsevier, 2008.
- [Deoliveira, 2008] Deoliveira, J. "GeoServer: uniting the GeoWeb and spatial data infrastructures." GSDI. 2008.
- [Di Bernardo et al., 2008] Di Bernardo, M., Montanaro, U., Santini, S., "Novel hybrid MRAC-LQ control schemes: Synthesis, analysis and applications" *International Journal of Control* 81 (6) , pp. 940-961, 2008.
- [Di Bernardo et al., 2008a] Di Bernardo, M., Montanaro, U., Santini, S., "Novel switched model reference adaptive control for continuous piecewise affine systems", IEEE CDC, 2008.
- [Di Bernardo et al., 2010] Di Bernardo, M., Di Gaeta, A., Montanaro, U., Santini, S. "Synthesis and experimental validation of the novel LQ-NEMCSI adaptive strategy on an electronic throttle valve", IEEE TCST, 2010.
- [Hadachi, 2012] Hadachi A., "Travel time estimation using sparsely sampled probe GPS data in urban road network", Doctorate Thesis, 2012.
- [Hedrick et al., 1994] Hedrick, J. K., Tomizuka M., and Varaiya P., "Control issues in automated highway systems." *Control Systems*, IEEE 14.6 (1994).
- [Iannello et al., 2004] Iannello G., Pescapè A., Ventre G., Vollero L., *Experimental Analysis of Heterogeneous Wireless Networks*. WWIC 2004.
- [Karres et al., 2006] Karrer R., Matyasovszki I., Botta A., Pescapè A., 'Experimental evaluation and characterization of the magnets wireless backbone'. WINTeCH 2006.
- [Marchetta et al., 2012] Marchetta, P., Salvi, A., Natale, E., Tirri, A., Tufo, M., and De Pasquale, D., "S2-MOVE: Smart and Social Move", IEEE GIIS 2012.
- [Reelsen, 2011] Reelsen, A. *Play Framework Cookbook*. 2011.
- [Salvi et al., 2013] Salvi A., Santini S., Biel D., Olm J. M. and Di Bernardo M. "Model reference adaptive control of a full-bridge buck inverter with minimal controller synthesis", IEEE CDC, 2013.
- [Schaffers et al., 2011] Schaffers, H., Komninos, N., Pallot, M., Trousse, B., Nilsson, M., and Oliveira, A. "Smart cities and the future internet: towards cooperation frameworks for open innovation" *The future internet* 2011.
- [Sichitiu, 2008] Sichitiu, M. L., and Kihl, M. (2008). "Inter-vehicle communication systems: a survey". *IEEE Comm. Surveys and Tutorials*.
- [Stonebraker, 1986] Stonebraker, M., and Rowe L.A., "The design of Postgres", Vol. 15. No. 2. ACM, 1986.
- [Yang et al., 2005] Yang J., Kang S. and Chon K., "The map matching algorithm of GPS data with relatively long polling time intervals", EASTS, vol. 6, 2005.

"Dis(ease)Ability Game"

Quando applicare una teoria pedagogica innovativa é un gioco da ragazzi, anzi...una App.

Raffaella Conversano
Media Educator – La Sapienza
Docente Specializzata
Istituto Comprensivo “A. Moro”
Via Alfieri 14, 74021 Carosino (TA)
raffaellaconversano@tin.it

After a complex research and experimentation on the definition of program guidelines, methodological and didactic application of my innovative pedagogical theory, based on re-reading application of Media Education in schools about the strategic use of new technologies in the definition of educational plans, to custom various difficulties - pathological and disabling of socio/cultural - I propose its instrumental application: an App that allows teachers and students, finally, to participate together with the construction of knowledge and how to learn it rather, to play video. An evolutionary ambitious project, which has pioneered the adoption of SMART information technologies in teaching and the use of technologically advanced software platforms for the integration of pupils with various disabilities. The research landscape to propose national educational seals the willingness to share, not only the success of my educational theory, now recognized worldwide, and titled: "Dis(ease)Ability", but also and above all, its new application.

Keywords: Disability, New Technology, App Game, New Theory.

1. Spunti critici

Frutto di un laborioso percorso di sperimentazione, questo lavoro è scaturito dal desiderio di implementare idee innovative circa l'uso di alcune tecnologie, utilizzate come strumenti strategiciolutivi nella comunicazione con determinate patologie disabilitanti gravissime. Tali applicazioni hanno condotto successivamente ad individuare nel fare quotidiano una linea teorica di riferimento, i cui risultati ottenuti altro sono stati che la sperimentazione della validità applicativa della stessa, circa le modalità di uso delle tecnologie con tutti, quindi anche per le patologie disabilitanti. L'idea è stata quella di proporre

Congresso Nazionale AICA 2013

e diffondere modelli originali di utilizzo didattico non solo delle più comuni strumentazioni tecnologiche (postazioni pc e sistemi connessi) ma, anche e soprattutto, uno strumento in particolare - il cellulare o sistema mobile - proponendo, proprio nel suo approccioolutivo di comunicazione e veicolazione delle informazioni (apprendimento funzionale alla competenza), le direttive ispiratrici della teoria pedagogica che ho chiamato:"Dis(agi)Abilità".

Difatti, si parla molto e a largo spettro circa l'uso delle tecnologie che inondano le stanze e i depositi delle scuole, senza tenere conto che ciò che manca realmente é:

- La predisposizione naturale alla PEDAGOGIA come identità di pensiero filosofico di riferimento.
- La padronanza DIDATTICA e METODOLOGICA al fine di indirizzare il processo di insegnamento alla soluzione dei problemi di comunicazione con tutti, per facilitare l'apprendimento generando competenze che, reinterpretate in modo critico dagli allievi, possano diventare nuove nella spendibilità quotidiana per l'autonomia personale.
- La padronanza peculiare della spendibilità intrinseca - come soluzione ipotizzabile ed apprezzabile non esclusivamente per disabili ma, accessibile comodamente "anche" da loro - di tutte le potenzialità che gli strumenti tecnologici possiedono, primo fra tutti il cellulare, la cui abilità di consumo risiede soprattutto nel saper gestire la sintesi funzionale dell'enorme panorama comunicativo/virtuale di cui é portatore.
- L'intenzionalità docente, intesa come regia dell'apprendimento attraverso una formazione consapevole al consumo delle potenzialità tecnologiche che gli strumenti per la comunicazione possiedono, per favorire l'autonomia di vita e di approccio ad essa per tutti.
- La necessità di ghetizzare chi ancora fa "audience" delle proprie incompetenze tecnologiche demonizzandone l'uso didattico, strategicamente valido, invece, per supportare la metodologia quotidiana che, attivando piani adeguati di condivisione formativa sociale, favorisce la possibilità di nuovi spunti creativi.
- La necessità di invertire la tendenza che da un lato inonda di informatizzazione e strumentazioni le scuole e, dall'altro, ne delega l'uso limitandolo solo a pochi eletti, senza considerare che "nuovo" non significa "buttare" ma "evolvere" soprattutto nel modo di pensare il contesto concettuale del termine "scuola".

2. Le risorse

Punto di partenza sono state le varie realtà presenti all'interno delle classi - patologie disabilitanti, etnie diverse con presenza anche di disabilità, difficoltà generalizzate di apprendimenti, difficoltà empatiche di comunicazione - dove la presenza di alunni con stili normali di apprendimento erano solo il caso limite per attivare percorsi di studio "normali". Il nodo da risolvere era quello di mettere tutti nella stessa condizione di immagazzinare, rielaborare e ricomunicare i dati culturali attraverso strumentazioni peculiarmente utilizzabili

da tutti. Mi sono guardata intorno e rovistato tra tutto ciò che era a nostra disposizione e tirato fuori, oltre i PC e le strumentazioni ad esso correlate, tutti i cellulari in possesso degli studenti confrontando i vari modelli e trovando i percorsi comuni e strumentali offerti dalle smart tech fino alle attuali Apps. Ci sono riuscita poiché a monte ero fermamente consapevole del mio ruolo intenzionale di essere docente, inteso come regia assoluta ed indiscussa dell'apprendimento, attraverso una formazione consapevole al consumo delle potenzialità tecnologiche che gli strumenti per la comunicazione possiedono, per favorire l'autonomia di vita e di approccio ad essa per tutti. Affrontando l'analisi delle reali difficoltà riscontrate all'interno delle classi dove, la presenza di alunni affetti da patologie più o meno gravi di approccio alla vivibilità della vita comunemente normale erano solo il caso limite da risolvere, ho dapprima analizzato le loro funzionalità residue abili e le modalità personali di comunicazione per individuare successivamente: il loro "funzionamento" inteso come approccio alle dinamiche comunicative per l'autonomia personale; le loro attitudini personali poiché tutti le hanno e vanno individuate, coltivate e potenziate; il grado di stress che la presenza del deficit scatenava non solo all'interno della classe, nelle relazioni con i compagni, ma anche tra i docenti della classe - stato di "ana-empatia" intesa come impedimento nel cogliere comunque una persona, in grado di comprendere con modalità comunicative diverse nella sua interezza - attivando quella barriera psicologica e comportamentale che porta ad etichettare l'individuo con termini vari quali "handicap, disabile....ecc."; gli obiettivi generali di tipo educativo/istruttivo che la scuola aveva programmato di perseguire; le strumentazioni a disposizione, quelle disponibili, quelle realmente funzionanti e sempre accessibili e quelle in nostro possesso; pensato in modo metodologicamente valido sotto il profilo della didattica al fine di organizzare percorsi comunicativi nuovi dei contenuti che, progettati insieme al gruppo classe come processo nuovo per "Imparare ad Imparare", fosse tecnologicamente valido ma, soprattutto, accessibile da tutti e non solo da chi aveva delle difficoltà visibili da arginare; superato i preconcetti legati all'uso e consumo della tecnologia mobile.

La sperimentazione di tali percorsi, progettati scientificamente, ha previsto una organizzazione funzionale ai risultati previsti, intesa come:

- Igiene professionale definita come team di ricerca/azione: darsi linee pedagogiche, didattiche e metodologiche di riferimento la dove il panorama della Media Education ne risultava sprovvisto al di là di una semplice analisi sociologica di riferimento circa l'impatto di ricaduta dell'uso che le tecnologie avrebbero avuto nel tempo (M. McLhuan) e di una indicazione blanda sulla necessità di "... far studiare i media in un modo serio e come una disciplina. oltre i semplici esercizi (...)" (Len Masterman); del resto, tutti parlano, ipotizzando una stesura condivisa, dell'avvento futuro di una pedagogia dei media, senza cogliere che serve una pedagogia che modifichi l'approccio comunicativo con tutti, favorendo così il nuovo modo di intendere strategico della tecnologia in funzioneolutiva.
- Riscontrare i successi ottenuti per ricalibrarli in funzione di evoluzione continua, adeguandoli alla evoluzione della tecnologia stessa.

- Organizzare, in funzione statistica, i risultati conseguiti, il materiale costruito e gli spunti conseguibili.
- Credere fermamente nella solidità delle linee teoriche delineate e condivise al fine di confrontarle, costruttivamente e condivisibilmente, in contesti di ricerca accademici nazionali, internazionali e mondiali per la divulgazione della stessa.

Non basta creare ottime tecnologie se poi manca la teoria pedagogica con la sua didattica e metodologia applicativa di riferimento, per poterle gestire in campo istruttivo; difatti, con la "Dis(agi)Abilità" ho puntato a modificare l'approccio comunicativo con tutti favorendo il nuovo modo strategico di intendere la tecnologia partendo dal presupposto che: non sono i "disabili" che devono comunicare con noi in modo univoco e standardizzato ma noi interpretare, attraverso le tante modalità che le varie tipologie cliniche e patologiche disabilitanti pongono in essere, ciò che costoro vogliono dirci e/o comunicarci.

3. Come ho progettato

In prima istanza mi sono chiesta: come i media col loro supporto tecnologico potessero effettivamente penetrare nella quotidianità della didattica e se, tra Media Education e lavoro didattico, potesse instaurarsi una organica e durevole relazione a favore di una migliore e reale qualità degli apprendimenti. Attraverso la ricerca/azione ho cercato di individuare i punti di sinergia che normalmente restavano offuscati dalla retorica tecnologica, oltre che dalla scarsa obiettività della documentazione educativa che spesso accompagna l'innovazione. Sono partita dal considerare in quale misura dovevo ricorrere alle tecnologie e come riscoprirle, rispetto a quelle più tradizionali, per favorire apprendimenti per lo più basati sulla manualità e corporeità. Il problema non è stato da poco, poiché andavano riconfigurati spazi e stili nuovi di pensiero, suscettibili nel tempo di articolazioni e consolidamenti culturali. Ho stabilito dei criteri per essere in grado di effettuare delle scelte su: quale ruolo attribuire ai media; a quali fini e contesti impiegarli; funzionali a quali finalità risolutive dei problemi cognitivi e di apprendimento che il progetto scolastico poneva in essere. La FINALITA' della sperimentazione è stata definita dall'analisi filosofica della teoria pedagogica (impostazione culturale, finalità, obiettivi), intese come linee direttive del percorso da intraprendere compresi i criteri di individuazione delle classi, le caratteristiche territoriali del bacino di utenza e di quelli diversamente coinvolti nel tentativo di indagare, anche indirettamente, come la stessa didattica, attivata con criteri analoghi, potesse variare in contesti socio-culturalmente differenti. Ho funzionalizzato la previsione operativa su gruppi di lavoro differenti allo scopo di declinare lo stesso percorso formativo/educativo a diverse fasce di apprendimento, definendo percorsi appropriati, con la individualizzazione di una metodologia condivisa in un sistema laboratoriale di rete. Entrando nel campo più specifico degli apprendimenti, ho considerato gli effetti che la ricaduta dell'uso delle tecnologie avrebbe inciso sulla didattica focalizzando in itinere: quali tecnologie utilizzare; quali ambienti esplorare o ricostruire; quali linguaggi

ricodificare per sviluppare capacità logiche al fine di favorire gli apprendimenti di contenuti specifici o per superare determinate difficoltà comunicative legate a patologie cliniche o deficit disabilitanti. **PRINCIPI ISPIRATORI:** Sperimentare: l'uso di codici verbali e non verbali superando l'uso di prodotti mediali come meri ed esclusivi sussidi didattici; Coinvolgere: in linea di continuità, vari livelli scolastici partendo dalla situazione di coinvolgimento con collegamenti di macrosperimentazioni in contesti classi tra i segmenti istruttivi di mia pertinenza; Promuovere e delineare: attraverso percorsi didattici specifici, l'apprendimento della cultura e dei linguaggi mediali col macro obiettivo di formazione della persona; Attivare: un confronto operativo su nuove applicazioni progettuali con dinamiche applicative di rete. **RICADUTA OPERATIVA sulla FORMAZIONE PROFESSIONALE:** promozione della ricerca/azione nelle aree tematiche di applicazione; sollecitazione attiva nello sperimentare una didattica e sua ricaduta metodologica nuova in classe; sostegno reciproco alla collaborazione/scambio in rete; promozione: di un atteggiamento di osservazione/ascolto volto a favorire l'efficacia degli interventi didattici attivati, la libera espressione professionale e la rielaborazione creativa degli alunni. **CONDIVISIONE di:** competenze professionali specifiche: specializzazione in Didattica Speciale e Media Educator; articolazione dell'impianto metodologico riferito allo specifico segmento scolastico individuato; attivazione di strategie conoscitive e progettazione di percorsi didattici specifici nei vari livelli educativi; adeguamento delle modalità auto valutative e verifica del feedback; previsione delle direzioni di sviluppo dell'attività didattica nelle classi individuate e di coinvolgimento in quelle successivamente coinvolte. La **FINALITÀ GENERALE** del lavoro ha ruotato intorno alla: consapevole importanza di una proposta culturale condivisa in modo sistemico; implementazione di una Didattica ed una Metodologia nuova nella scuola di tutti; promozione, in modo rinnovato, di "sapere e saper fare" metodologico e didattico affinché fosse funzionale alla formazione di nuove competenze e consapevolezze rispetto alla fruizione delle forme più svariate di comunicazione ed espressione audio e visiva. Con **FINALITÀ** da soddisfare **FUNZIONALI** alla teoria pedagogica di riferimento quali: incentivare la gestione autonoma e consapevole del rapporto con la sfera dell'autonomia, dell'informazione e della comunicazione; contribuire a coltivare le potenzialità immaginative, espressive e creative, elevandone il gusto estetico degli alunni coinvolti; promuovere il riconoscimento della specificità del linguaggio utilizzato e del loro approccio operativo, quale visione culturalmente determinata e linguisticamente strutturata; sostenere l'assunzione di un atteggiamento analitico e sistemico che consenta ulteriori generalizzazioni, approfondimenti e contestualizzazioni; sollecitare l'accostamento ai nuovi linguaggi in funzione del ruolo determinante che questi assumono nel modo degli alunni di affrontare, conoscere ed interpretare la realtà in un quadro definito di **FINALITÀ FORMATIVE** per: formare ad apprendere e a predisporre la progettazione di attività specifiche inerenti la Media Education in ambito scolastico; consentire la disponibilità, il monitoraggio e la valutazione delle singole azioni intraprese. In questa interazione, tra processo educativo e spettacolarità comunicativa, è nei media che ho trovato la chiave di lettura per uno sviluppo applicativo

interessante, una formula che, definita all'interno della mia teoria – la Dis(agi)Abilità -, ha consentito il passaggio definitivo ad un nuovo concetto di uso delle tecnologie nei processi educativi classici. Questo nuovo modello pedagogico, evolvendosi verso un uso integrato dei mezzi di comunicazione di massa nei percorsi scolastici, ha consentito di rimodellarli in base alle diverse esigenze di tipologie (difficoltà cognitive, psico/fisiche di accesso, disagio socio-culturale) al servizio anche della conseguente metodologia, alleanza forte e allo stesso tempo necessaria, per la sopravvivenza di tutti gli attori coinvolti nel processo di formazione ed educazione. La sfida, nella ricerca applicativa di una didattica ed una metodologia di tipo educativo, che si ponesse, secondo modalità costruttiviste al centro, tra ormai obsoleti impianti manieristici di cognitivismo e comportamentismo, è stata quella di coniugare la rigorosa esigenza formativa del mondo della scuola con quella di intrattenimento tipica dei media in tutte le loro espressioni, allo scopo di raccordare complessità strumentali e loro uso in ambito istruttivo/comunicativo.

4. Dai casi di studio alla progettazione della AppGame

Partendo dai risultati di autonomia personale ottenuti dallo studio dei vari casi di patologie disabilitanti, che mi hanno consentito di sperimentare con successo l'esito applicativo di quanto via via teorizzavo, e ad altri ancora in corso di studio, ho lavorato ad un'idea più ampia di apprendimento e costruzione dello stesso come processo di crescita cognitiva ed affettiva, poiché ogni vicenda ha la sua storia, le sue dinamiche, le sue potenzialità e genialità di azione e dove il vero limite è proprio il docente, con le sue ansie e paure di fronte alle diagnosi e deficit vari. Ho finalizzato il tutto verso un processo della conoscenza della persona costruito sul suo progetto di vita, contribuendo a dare loro un'immagine della scuola che fosse meno angusta e meno denigratoria di quanto spessissimo accade, diventando per loro - gli alunni "con e in" difficoltà - non un semplice punto di riferimento ma la disponibilità spaziale dell'incontro. Per i docenti, invece, ho puntato a superare la "difficoltà" di vivibilità della classe proponendo la presenza della patologia, quindi del deficit, come elemento di allineamento agli standard comuni della classe, restituendo loro il timone del potere decisionale e riorganizzativo della didattica. La scuola, difatti, è il luogo dove ci si forma, un percorso sinergico che nasce costruendolo insieme giorno per giorno; un luogo dove educare alla libera espressione creativa del pensiero personale di tutti, cercando soluzioni moderne, all'avanguardia per rendere dinamica e innovativa la propria didattica. Ecco che ciò che mancava era lo strumento che consentisse di interagire con gli allievi della propria classe, per conoscerli al meglio, nel profondo, al di là di competenze ed abilità già definite, attraverso cui reimpostare la metodologia. Una didattica fluida per nuova impostazione, funzionalizzata su obiettivi concreti da raggiungere attraverso una adattabilità degli argomenti/lezioni gestiti insieme - docente/allievo - in un "game", un gioco virtuale interattivo dove, inserendo insieme gli elementi del sapere, questi portassero ad una competenza condivisa: vincere insieme, inteso come vittoria della competenza degli elementi costruttivi la lezione

videogioco. Lo spunto più importante è arrivato proprio dal caso di Davide, un ragazzo con Sindrome di Asperger e autismo gravissimo che ho seguito quest'anno; data la sua sfrenata passione per i videogiochi e la profonda competenza in campo informatico, si muoveva all'interno del contesto scolastico proprio con la logica dei virtualgames: lui non veniva a scuola ma ogni giorno entrava in un circuito di cui lui decideva le sorti. È stato sufficiente cercare quello che lo appassionava di più "giocando" con lui dal vivo, prevedendo azioni e reazioni, per definire i suoi punti di criticità e debolezze per condurlo a "studiare" ciò che rifiutava. Gli è stato costruito, tramite un primo prototipo di App, un circuito game dove, in una sorta di percorso di guerra, ho fatto inserire gli esercizi proposti dalle varie discipline, soprattutto ambito scientifico e logico, come percorsi obbligati da risolvere per proseguire nel gioco, proprio come nel film "Rambo", dove il protagonista si rifugia nella giungla perché lì dentro si era allenato e li sapeva gestirsi. Il rifiuto di Davide verso il metodo classico di insegnamento è stato l'imput cruciale per andare alla ricerca di modalità diverse ma vere, per farlo comprendere e apprendere partendo dalla strada che lui stesso indicava tramite il suo funzionamento con lettura dinamica in ICF-CY. Partendo da questo prototipo ho progettato e definito la "Dis(ease)Ability Game" App, uno strumento pedagogico per i docenti, che consente di interpretare con metodo e didattica elaborata, la teoria innovativa della "Dis(agi)Abilità" come strumento operativo quotidiano, divenendo al contempo un virtual game per gli alunni, con cui interagire per costruire insieme la lezione che conduce al traguardo, inteso come acquisizione delle abilità e competenze di riferimento. Si parte dalla presentazione dell'allievo che inserisce i primi dati e foto personale, successivamente viene sottoposto un test psicoattitudinale a carattere personale; al termine del test la App rielabora i dati della foto trasformandola in un avatar che entra nel virtual game. È qui che entra in gioco il docente che propone l'argomento e l'obiettivo educativo/cognitivo cui la App, elaborando metodologia e strumenti, costruisce l'aula virtuale dando inizio al gioco.

5. Innovare è pensare in modo pedagogico nuovo

L'utilizzo delle nuove tecnologie favorisce significative modifiche del contesto educativo poiché, se da un lato accresce la motivazione degli alunni, dall'altro ne incrementa l'autostima favorendo una capacità di coinvolgimento ed un livello partecipativo maggiore, facilitando tutti i processi di relazione. Ero consapevole che proprio all'interno della didattica (pratica dell'insegnamento) e della metodologia (percorso per la realizzazione di un fine educativo) dovevo individuare partecipazioni dinamiche operative, nuove rispetto ai processi standardizzati, anche se più elevati di astrazione, ho riflettuto, così, sulla funzione che le tecnologie dovevano avere per favorire i processi di apprendimento, considerando come primaria la necessità di valorizzare lo studente non solo come attore ma anche come protagonista del percorso di conoscenza e di apprendimento evolutivo co-costruito. Ho desunto che la metodologia da individuare doveva adeguare l'insegnamento alle personali caratteristiche degli alunni – ai loro ritmi e stili di apprendimento, alle capacità

linguistico/comunicative ed ai loro prerequisiti cognitivi -, per far conseguire individualmente obiettivi di apprendimento comuni al resto della classe; difatti, se la dinamica del processo di insegnamento-apprendimento non pone i contenuti scolastici al centro ma li riporta al loro giusto ruolo di stimolo percepibile ed utilizzabile da tutti, per l'alunno in difficoltà questa diventa una occasione per ripensare la formazione scolastica come strumento di successo formativo per tutti. Per realizzare ciò, mi sono distaccata dal pensiero standardizzato dell'insegnamento basato unicamente sulle discipline per orientarmi verso l'operatività dell'insegnamento centrato sulle persone che apprendono, superando a prescindere tutte le barriere dinamico/operative, socio/culturali, clinico/patologiche di tipo anche disabilitante di interazione comunicazionale con cui ho operato. Ho delineato una strategia DIDATTICA innovativa, che fosse nuova per analisi ed impostazione, orientata in tre principali aree: Motivazionale: alimentare capacità creative e comunicative favorendo il superamento di ansie ed inibizioni connesse alla comunicazione dei saperi; Percettivo/Cognitiva: sviluppare abilità percettivo-motorie e specifiche abilità cognitive e metacognitive; Disciplinare/Culturale: espandere segmenti di apprendimento disciplinare/inter-disciplinare, favorendo esperienze multiculturali, criticizzando l'apprendimento attraverso la presentazione di tematiche secondo varie angolature e chiavi di lettura promuovendo FINALITÀ METODOLOGICO-DIDATTICHE olistiche per: acquisire la capacità di definire obiettivi specifici di apprendimento connessi a piani di lavoro attinenti la Media Education; prevedere il feedback finale dei risultati attesi dal percorso sperimentale progettato, attraverso l'analisi territoriale scolastica di riferimento; funzionalizzare la metodologia ai reali bisogni formativi degli allievi, organizzati in una concreta situazione laboratoriale per: favorire l'acquisizione di specifiche competenze anche trasversali ed incrementare l'autostima. In pratica, ho individuato le linee direttive di una didattica valida secondo la reinterpretazione della Media Education - quale approccio tecnologico e intenzionalità docente sugli sviluppi interattivi tra regia dell'apprendimento e competenze qualitative dei discenti - per puntare ad una metodologia che fosse una valida tabulazione a priori di tutte le informazioni che occorrono al docente per implementare e adeguare la didattica e le strumentazioni tecnologiche alle peculiarità di ciascun discente, affinché i contenuti e la loro scelta fossero il più adeguati possibile agli obiettivi standardizzati da raggiungere. Mi sono messa all'opera per creare uno strumento tecnologico innovativo, dinamicamente strutturato, open for all e non solo per casi isolati, da fornire ai docenti per dare ai discenti la possibilità di "dire", presentare chi sono veramente, come operano ed elaborano le informazioni, quali strumenti utilizzano e come li vorrebbero per esprimersi al meglio dando così loro la possibilità di disegnare anche la tecnologia del futuro ma costruendola insieme. Questo strumento, concretizzatosi in una App, tecnologicamente impostata su device mobile, consente di dare a tutti, e non solo allo studente con diversa abilità, l'opportunità di comunicare in modo alternativo le proprie capacità e i propri interessi. Grazie a questa App il docente acquisisce i dettagli che riguardano tutte le aree di studio ed interesse oltre che multimediale, consentendogli così di capire la diversità funzionale di uso delle varie abilità nei confronti della sua interazione con l'esterno, al fine di

calibrare le strumentalità più confacenti ad aumentare tale processo di autonomia; in funzione di ciò, si potrà decidere la strategia di intervento didattico più opportuna puntando sulle capacità e gli interessi evidenziati, motivando al massimo il discente.

5. Conclusioni

Insegnare é un lavoro impegnativo cui nessuno ha garantito l'esito: richiede preparazione, passione, flessibilità, energia e grande professionalità. Con grande disponibilità di intenti, mi sono rimessa in discussione per comunicare in modo efficace, per "inventare" e "reinventare" lezioni interessanti con modi e tecniche di trasmissione dei saperi che fossero al contempo attraenti e stimolanti per l'apprendimento degli studenti, tutti nessuno escluso, indipendentemente dalle loro modalità di approccio alle tecniche di interazione. Storie quotidiane, diverse, sempre stuzzicanti e sempre colme di sorprese e piccole grandi conquiste di autonomia, dove in gioco ho messo le mie idee: efficaci, buone e spesso eccellenti anche in situazioni apparentemente difficili da gestire. Non era il momento di pensare a ciò che loro non avevano, piuttosto di pensare a ciò che potevo fare con quello che io avevo per loro: le mie idee. Mi sono posta come "lavoratore della conoscenza" facendo della trasmissione e della costruzione dei saperi il mio grande progetto professionale, realizzando uno spazio, virtuale e solido, dove inserire la connessione di due mondi differenti e paralleli al contempo. Mi sono calata con immagini in un progetto innovativo per efficacia di soluzione, discusso idee e analizzato vicende per: "raccontare, confrontarmi e comunicare - da persona a persona - nella scuola e per la scuola", perché sono le persone con l'eclettismo della loro individualità, la specificità della loro esperienza e l'originalità delle loro idee che fanno la differenza. Non ho fatto della scuola il mio valore ma, poiché sono una persona che ci lavora, ho voluto costruire il "valore della scuola" per come lo sentivo pulsare in me prima di ogni altra componente. Ci credevo da sempre, fondere le mie energie e perseguire successi nei risultati mi ha confortata, ed ora più che mai mi stimola a crederci e ad "...essere ancora più affamata e folle...!" (S. J.)

Bibliografia

- [1] CONVERSANO R., *Interagire per crescere – Interazione tra Media e Formazione*. EDIZIONI PUGLIESI, 2005
- [2] CONVERSANO R., *Progetto di sperimentazione in M.E. – La Media Education nella scuola elementare*. EDIZIONI PUGLIESI, 2005
- [3] CONVERSANO R., BINACCHI M., *The Enchanted Maze*, in "Challenges in International Communication" Edited by ATINER 2012
- [4] CONVERSANO R., MANZULLI G., BINACCHI M., *Work in Progress*, Didamatica2012 – Atti del Convegno

[5] CONVERSANO R., MANZULLI G., **Work in Progress**, in “Lecture Notes in Electrical Engineering” - volume 2010 - Information Technology by Wei Lu, Guoqiang Cai, Weibin Liu, Weiwei Xing Editor SPRINGER 2012

[6] MANZULLI G., SALENTINO A., **Mediateca Virtuale**, 1° premio in Didamatica 2008 – Atti del Congresso Editore Laterza Giuseppe Edizioni, 2008

[7] MASTERMAN L., **A scuola di media. Educazione, media e democrazia**, La Scuola Brescia 1997

[8] McLUHAN M., **Gli Strumenti del Comunicare**, Il Saggiatore, 2008

[9] VYGOTSKIJ L., **Il processo cognitivo - Raccolta di scritti** a cura di Michael Cole, Sylvia Scribner, Vera John-Steiner, Ellen Souberman, Ed. Bollati Boringhieri, 1987-2002

[10] SCHIANCHI M., **La terza Nazione del Mondo. I disabili tra pregiudizio e realtà**, Editore Feltrinelli, Serie Bianca Feltrinelli Milano 2009

[11] A cura di DE POLO G., PRADAL M., BORTOLOTTI S., **ICF-CY Servizi per la Disabilità**, FrancoAngeli Milano 2011

[12] PRENSKY M., **Digital Game-Based Learning**, 2001

Link utili:

RAI RADIO 1: “Area di servizio” - Intervista a Raffaella Conversano e Gaetano Manzulli
www.rai.tv/.../ContentItem-2d54f9c3-fe29-4e53-99af-ae8c7ce1eb22

RAI RADIO 1: “Area di servizio” - Intervista a Gaetano Manzulli
www.rai.tv/.../ContentItem-a13fe48c-891d-4dd8-ae4a-688db20f2b1

RAI RADIO 1: “Diversi da chi?” – Intervista a Raffaella Conversano
www.radio.rai.it/radio1/diversidachi/view.cfm?Q_EV_ID=321903

HANDIMATICA 2010 – Seminario “Migranti e disabili: tecnologia delle mediazioni e dei mediatori”
<http://www.youtube.com/watch?v=DCU0A74cwfo>

Il progetto “Sempre Giovani 2.0”

Domenico Consoli
Istituto Tecnico Commerciale “C. Battisti”
Viale XII Settembre n. 3, 61032 Fano
domenico.consoli@istruzione.it

Abstract. *The paper describes the project "Young 2.0 Forever", which was implemented in two editions at a High Education Institute. The goal of the project was to compare the two generations of Digital Natives (young) and Paper Natives (elderly). Nowadays young people use technological tools with a greater familiarity and speed and can transmit this technological skill to elderly. Older people can use these interactive tools in their private life to communicate with others, relatives and friends or to interact with public administrations (e-democracy). The project could be framed in a context of overcoming the Digital Divide to stimulate a "digital inclusion" and an active citizenship. In this way, it is possible to reduce the technological gap of senior citizens towards the new technologies.*

Keywords: *digital divide, digital natives, adult education, web 2.0, social networks.*

1. Introduzione

Il progetto “Sempre Giovani 2.0” si è attuato in due edizioni presso un Istituto Scolastico Superiore, con il patrocinio degli Assessori alle Politiche Sociali e alle Innovazioni Tecnologiche e il supporto di una Fondazione di Solidarietà del territorio. Il progetto ha avuto come obiettivo quello di far spiegare a studenti, nella veste di docenti, l'utilizzo dei social network (Facebook, Youtube, Twitter,...) ad un gruppo di corsisti anziani. In questo modo si sono messe a contatto le due generazioni: i nativi digitali (giovani) e i nativi cartacei (anziani).

Oggi le tecnologie interattive e collaborative del web 2.0, social web o web partecipativo, che si basano sui concetti di condivisione e collaborazione, stanno apportando dei profondi cambiamenti nel modo di relazionarsi con la gente e nella società. Il web 2.0 permette alla singola persona di esprimere un parere/suggerimento su un particolare prodotto/servizio in modo che l'azienda/ente possa apportare dei miglioramenti [McAfee, 2006].

Anche le persone adulte, per non essere escluse, devono saper utilizzare e applicare, nel modo corretto, queste nuove tecnologie. Questi strumenti, che danno maggiore enfasi alla comunicazione, alla trasmissione e scambio di informazioni, devono essere utilizzati da tutti: giovani, adulti e anziani. In questo modo ognuno può apportare un contributo alla creazione di conoscenza e

intelligenza collettiva e allo scambio di suggerimenti, proposte e idee su come migliorare le condizioni di vita sociale, culturale e politica di una comunità.

In questo progetto, oltre all'aspetto tecnologico, è stato importante anche il ruolo di docente svolto dagli studenti. Loro generalmente sono abituati a recepire in maniera, più o meno passiva, le lezioni impartite dai professori. Questa volta sono stati loro a "salire" in cattedra e ad essere responsabili dei contenuti didattici che impartivano ai corsisti.

L'articolo è così strutturato: nel paragrafo 2 si analizza la letteratura di riferimento e nel terzo si profilano le figure di questi giovani internauti e dei nativi cartacei. Si passa quindi, nel paragrafo 4, alla descrizione del progetto e all'analisi delle risposte al questionario somministrato agli studenti-docenti e ai corsisti anziani (paragrafo 5). Infine nel paragrafo 6 si trae qualche conclusione e si delineano dei probabili sviluppi futuri.

2. Letteratura di riferimento

Il termine web 2.0, che descrive lo stato corrente di Internet, è stato coniato per la prima volta da O'Reilly [2005] e denota la transizione dalla concezione del Web come contenitore a Web come "fornitore" di servizi reali, dove gli utenti, attivamente, possono esprimere opinioni/consigli e interagire tra di loro. Nel web 2.0, più che l'aspetto tecnologico, prevale l'aspetto sociale che si basa su concetti fondamentali come la relazione, lo scambio di informazioni, la partecipazione attiva, la collaborazione e l'intelligenza collettiva [Levy, 1997; De Kerchove, 1997]. Nel web 2.0 è importante la persona, come soggetto pensante, che si relaziona con gli altri e come generatore di contenuti. Grossman [2006], assegna a questo soggetto il titolo della copertina del settimanale Time "La persona dell'anno sei tu".

Il social web ha favorito la nascita e lo sviluppo di ambienti virtuali interattivi e collaborativi quali i social media. Secondo Kaplan e Haenlein [2010] i social media rappresentano quel gruppo di applicazioni Internet basate sui presupposti ideologici e tecnologici del web 2.0 che consentono la creazione e lo scambio di contenuti generati dagli utenti.

Oggi questi strumenti sociali sono molto utilizzati dai nativi digitali [Prensky, 2001a; Prensky, 2001b] che sono più portati rispetto ai loro insegnanti e alle persone adulte ad utilizzare queste nuove tecnologie. Don Tapscott [2000, 2008], analizza dettagliatamente la Net Generation che ha raggiunto la maturità nell'era digitale e tutte le implicazioni e i cambiamenti che si stanno verificando sia nella famiglia che nella scuola e nella società.

I nativi digitali non sono più semplici e passivi consumatori ma sono a tutti gli effetti dei "prosumer" [Toffler, 1980] e cioè consumano ma soprattutto producono informazioni e partecipano, attivamente e in maniera costante, alla creazione di beni e servizi. Si scambiano commenti, opinioni, feedback e collaborano tutti insieme contribuendo alla creazione di una intelligenza collettiva. La produzione, in maniera trasparente, di contenuti collaborativi sta cambiando il modo di comunicare e partecipare nel dibattito e nella vita pubblica.

3. I nativi digitali e i nativi cartacei

Oggi i giovani utilizzano i nuovi strumenti digitali con maggiore dimestichezza, velocità rispetto agli adulti [Consoli, 2013]. Loro sono i nativi digitali, nati con il web, i canali virtuali e i multi-schermo (cellulari, TV, computer, iPad, iPod). Sono orientati al multi-tasking, passano da un media ad un altro e fanno più cose contemporaneamente: studiano, ascoltano musica, rispondono a qualche SMS, vedono un filmato su Youtube. Per capire come funziona un dispositivo elettronico provano, esplorano, tentano, sbagliano, correggono e si consultano tra di loro. Controllano gli spazi virtuali e si esprimono liberamente sul web tramite il blog, le chat, i forum che sono diventati oramai oggetti integranti della loro vita e delle loro relazioni. Non conducono una vita reale e una virtuale ma, essendo sempre connessi, ne vivono una unica che si integra e che tende sempre di più a convergere grazie all'utilizzo dei social media.

La generazione dei nativi digitali o Net Generation [Tapscott, 2000] è una generazione, socialmente responsabile, in grado di interagire, collaborare e di socializzare nel villaggio globale [McLuhan, 1964] a prescindere dalle distanze fisiche geografiche. E' una generazione che guarda meno la televisione; più che altro l'ascolta mentre chatta con gli amici e naviga su Internet condividendo idee ed emozioni.

L'utilizzo intensivo delle nuove tecnologie interattive e dei social network da parte dei nativi digitali, oltre ad evidenti vantaggi, comporta anche degli svantaggi come la frammentarietà della conoscenza e la debolezza di un pensiero critico costruttivo che si forma con una lettura e analisi approfondita dei testi e non con lo sfogliare o navigare velocemente tra una mole di informazioni proveniente da molti canali virtuali. Queste loro azioni veloci e diversificate sono le cause di un loro linguaggio frammentato e sintetico [Contri, 2013]. I nativi digitali raccolgono e collezionano frammenti di vari media e non si focalizzano, in profondità, su qualcosa. Essendo soggetti ad un overload informativo non riescono a ricordare tutto ma solo dei frammenti e non riescono neanche a distinguere le informazioni utili dal "rumore" aggiuntivo. Con questi frammenti sono portati ad elaborare un pensiero destrutturato del tutto privo di senso critico.

Gli anziani, nativi cartacei, a differenza dei giovani, sono mono-schermo (principalmente guardano il televisore), sono abituati a lavorare in sequenza e fare una cosa alla volta, controllano gli spazi fisici (la scuola, la fabbrica, l'ufficio,...) e non quelli virtuali e nell'installazione di un dispositivo meccanico-elettronico seguono il manuale e non procedono per tentativi dando più spazio alla logica e non all'istinto e all'esplorazione.

Le persone anziane non usano molto la tecnologia ma hanno tanta esperienza sia nel settore sociale che politico.

Bisogna superare il Digital Divide tra giovani e adulti. Gli adulti non devono diventare esperti ma devono lasciarsi coinvolgere dai cambiamenti, aprirsi al nuovo e capire le logiche di funzionamento di questi nuovi strumenti per poter comunicare e condividere pensieri, parole e idee con tutti gli altri.

Se gli adulti utilizzano le nuove tecnologie possono partecipare attivamente al dibattito sui canali virtuali ed esprimere le loro opinioni su argomenti che

conoscono bene. Con gli strumenti digitali si possono attivare progetti di e-democracy; ognuno può esprimere le proprie idee e opinioni e può implementare azioni di democrazia partecipativa. Gli adulti possono essere coinvolti nella partecipazione e discussione di norme che rispettano il territorio, l'ambiente e la società e possono quindi giocare un ruolo essenziale nello sviluppo della comunità locale. In tal senso è importante intensificare il rapporto generazionale tra adulti e giovani. L'integrazione giovane-adulto può essere interessante per lo scambio di esperienze e competenze, tecnologiche da una parte e socio-politiche e culturali dall'altra.

La Comunità Europea, ha emanato diverse direttive per promuovere l'educazione, l'apprendimento, l'aggiornamento e l'impiegabilità lungo l'intero corso di vita (lifelong learning)[Sloep et al., 2011]. Un cittadino ben formato contribuisce allo sviluppo della comunità. Se tutte le persone, anche gli adulti, hanno certe competenze tecnologiche, la Pubblica Amministrazione può offrire migliori servizi online e avere dei feedback interessanti per il miglioramento delle condizioni sociali della comunità.

4. Il progetto

Il progetto "Sempre Giovani 2.0" si è svolto in due edizioni. Complessivamente nelle due edizioni hanno partecipato 26 anziani (13 per corso) e 26 studenti, 6 dei quali hanno partecipato ad entrambi le edizioni.

Il corso consisteva in n. 8 lezioni su Internet e i Social Network e in particolare in un incontro settimanale dalle 14.00 alle 16.00 presso il laboratorio di Informatica dell' Istituto.

Ogni studente affiancava un corsista anziano nella veste di docente. Lo scrivente espletava solo la funzione di supervisore. Prima di iniziare il corso si concordava la lezione da svolgere: Internet e motori di ricerca, Facebook, Twitter, YouTube,... e dopo ogni studente, in piena autonomia, portava avanti un percorso didattico individualizzato in base alle esigenze dell'anziano. Su Facebook si è creato un gruppo "Sempre Giovani 2.0" e tutti i partecipanti, studenti e anziani, si sono scambiati le amicizie tra di loro.

Gli studenti partecipanti avevano un'età compresa tra i 15 e i 17 anni e frequentavano le classi II e III. In totale sono stati coinvolti 8 ragazzi e 12 ragazze.

Per quanto riguarda gli adulti (15 femmine e 11 maschi) erano tutti pensionati e l'età variava dai 62 anni ai 72 anni. Qualcuno era alla prima esperienza di utilizzo del computer, qualcun altro era autodidatta e la maggior parte aveva seguito un primo corso di introduzione all'uso del computer organizzato da una Fondazione Solidale del territorio.

Alla fine di ogni corso ho provveduto a somministrare dei questionari sia agli studenti che ai corsisti anziani. Il questionario e le risposte saranno analizzati nel paragrafo successivo relativo all'analisi dei dati.

Alla fine di ogni ciclo di lezioni si è proceduto alla consegna degli attestati presso la sala del Comune in presenza degli Assessori alle Innovazioni Tecnologiche e alle Politiche Sociali.

5. Analisi dei dati

Sia agli studenti che alle persone adulte, a fine percorso, sono state somministrate dei questionari.

5.1 Analisi delle risposte degli anziani

Di seguito si riportano le domande e su appositi riquadri le risposte più interessanti di entrambi i due corsi.

D1. Domande generiche rivolte agli anziani

<u>Ha utilizzato un gioco elettronico?</u>		<u>Cosa le piace fare di più su Internet?</u>	
mai	69%	leggere i giornali	42%
playstation	15%	cercare immagini	54%
nintendo	8%	cercare informazioni	81%
altro	8%	comprare qualcosa online	15%
<u>Continua a giocare ancora?</u>		prenotare	23%
si	12%	<u>Prima di comprare confronta i prezzi su Internet?</u>	
no	88%	mai	31%
<u>Prima di questo corso era mai entrato su Facebook?</u>		qualche volta	69%
si	23%	<u>Utilizza ebay per comprare?</u>	
no	77%	no	88%
<u>Le piace Facebook?</u>		si	12%
si	92%	<u>Quante volte utilizza l'e-mail?</u>	
no	8%	spesso	23%
<u>Prima di questo corso aveva mai usato Twitter?</u>		qualche volta	69%
si	4%	mai	8%
no	96%	<u>Preferisci mandare messaggi tramite:</u>	
<u>Le piace Twitter?</u>		posta elettronica	42%
si	38%	sms	58%
no	62%	<u>Le piace leggere i contenuti in:</u>	
<u>Cosa le piace condividere su Facebook:</u>		cartaceo	58%
commenti	42%	digitale	12%
messaggi	62%	<u>16. Quali dei seguenti media preferisce?</u>	
foto	50%	televisione	42%
filmato	31%	Internet	23%
mi piace	23%	cinema	15%
		computer	8%
		altro	12%

Come si può osservare dal riquadro relativo alla domanda D1, la maggior parte degli anziani non aveva mai giocato con un gioco elettronico (69%) e solo in pochi continuano a farlo (12%). Pochi usavano Facebook (23%) e non conoscevano Twitter (96%). La maggior parte preferisce condividere su Facebook messaggi (62%), foto (50%) e commenti (42%). Su Internet piace di più cercare informazioni (81%) e immagini (54%). Gli anziani non usano ebay (88%) e preferiscono mandare più SMS (58%) che messaggi di posta elettronica (42%). In media preferiscono di più la televisione (42%) e leggere i contenuti in formato cartaceo (58%).

Nelle risposte alle domande successive si riportano solo quelle quelle più interessanti ai fini della nostra analisi.

D2. Cosa ne pensa dei ragazzi di oggi?

- sono molto svegli, intelligenti e sanno riflettere. Usano la tecnologia in maniera intuitiva e riescono ad avere informazioni utili in tempo reale;
- sono molto bravi per il digitale ma un pò poveri nello scrivere e nel leggere;
- hanno delle competenze che oggi sono indispensabili ma non bisogna dimenticare che la macchina è una propaggine dell' uomo;
- riescono a far un uso disinvolto delle nuove tecnologie e nella maggior parte delle volte, agiscono in maniera costruttiva e non violenta, alle iniziative legislative che li riguardano come studenti e cittadini;
- stanno perdendo i veri valori: il confrontarsi con gli altri e guardarsi in viso mentre si parla;
- faticano a sognare e immaginare;
- un pò di manualità e fantasia non fa sicuramente male;
- sono bravi ma c'è il rischio che, a forza di presidiare i canali virtuali, perdano il gusto della compagnia. La solitudine sarà il male del futuro;
- sono delle brave persone ma i media li bloccano intellettualmente. Più che usare Internet dovrebbero riflettere sull'essere umano e sulla sua interiorità per scoprire cose interessanti e emozionanti;
- bravi nell'uso delle tecnologie ma insicuri nelle amicizie e nelle sfide che la vita comporta.

D3. Quale consiglio darebbe ad un ragazzo per rafforzare quei valori a cui noi adulti crediamo?

- di essere sempre se stessi e in modo sincero;
- di comportarsi sempre secondo gli schemi dei veri valori e riflettere sui valori del passato;
- di ascoltare di più gli adulti, i loro consigli e di apprezzare i genitori;
- cercare di approfondire le cose e non fermarsi alla superficialità;
- leggere molto, ascoltare un pò di più gli anziani e fare riferimento alle loro esperienze vissute;
- non seguire mai il gregge ma ragionare e decidere con la propria testa;
- di non basarsi esclusivamente su quanto apprendono dai contenuti dei social network poiché spesso le notizie possono essere inquinate mentre i suggerimenti degli adulti ed in particolare dei genitori e dei nonni sono i più sinceri;
- credere nei valori della famiglia e imparare anche a sporcarsi le mani;
- essere più altruisti e non isolarsi;
- aiutarsi tra di loro a superare le situazioni problematiche;
- avere come regola i comandamenti di Dio;
- confrontarsi sempre con gli altri, essere consapevoli con umiltà dei propri limiti. Porsi sempre degli obiettivi, piccoli o grandi, da raggiungere.

D4. Le chiedo di esprimere un giudizio sullo studente che l'ha affiancato in questo corso?

- è molto bravo, educato, paziente, disponibile e gentile;
- è molto preparato dal punto di vista tecnologico;
- sono diventato nonno in questi giorni e sinceramente mi auguro che mio nipote possa assomigliare a questo studente;
- è molto bravo ma è troppo preso dal suo chattare;
- ha faticato a trasmettere la sua "manualità tecnologica" a me che sono lenta. Un 10 per la fatica di sopportarmi;
- si impegna molto ad insegnarmi;
- oltre ad essere un'insegnante più che valido è anche un angelo custode.

D5. Per quale motivo noi adulti dobbiamo usare di più le tecnologie?

- per saper rispondere ed essere sempre più vicini ai nostri figli e nipoti;
- per rimanere sempre aggiornati e competitivi;

- per cercare di stare al passo con i tempi visto che la società si evolve continuamente;
- per risparmiare e fare risparmiare lo Stato;
- ce lo impone la globalizzazione;
- se mio padre non avesse preso la patente sarebbe stato escluso da molte occasioni. Se noi non apprendiamo l'uso delle nuove tecnologie, in futuro, saremo esclusi da diverse iniziative;
- per soddisfare i nostri bisogni fisici senza però eccedere ed esserne succubi;
- per una ginnastica mentale e sentirsi attivi nella società.

5.2 Analisi delle risposte degli studenti

Da domande generiche fatte agli studenti, tranne due, gli altri sono convinti che i corsisti che affiancavano hanno appreso tutto quello che gli hanno spiegato.

Chi ha insegnato in entrambi i due corsi, nel secondo ha migliorato il modo di spiegare, di porsi, di andare più in dettaglio negli argomenti. C'è chi invece ha tenuto la stessa linea di condotta.

Tutti sono rimasti entusiasti dell'esperienza svolta e pensano di aver spiegato gli argomenti nel modo giusto. Interessante la risposta di diversi studenti: "Le 2 ore del corso non erano poi tanto lunghe; stando in loro compagnia, è stato un pò come stare con un vero e proprio nonno".

D6. Cosa ne pensi della persona anziana che seguivi?

- brava, preparata, perspicace, volenterosa e veloce nell'apprendere;
- faceva fatica ad apprendere e non era molto disponibile a imparare;
- simpatica, gentile, affettuosa e divertente; capiva al volo quello che gli spiegavo;
- aveva difficoltà nel ricordare determinati passaggi ma con pazienza e concentrazione è riuscita sempre a portare a termine tutti i passaggi;
- entusiasta nell'apprendere questi argomenti;
- aveva tanto voglia di imparare e tanto da insegnarmi sull'aspetto sociale;
- è una persona colta e istruita in molti settori e aveva una generica conoscenza di Internet. Ha seguito sempre le spiegazioni e messo in pratica tutto quello che le veniva spiegato;
- è stato bello interagire con lui e scambiare informazioni;
- era un'anziana molto simpatica anche se non gli bastava una volta per capire le cose. Alla fine è riuscita ad imparare l'utilizzo dei vari strumenti e adesso mi scrive anche su Facebook.

D7. Esprimi un giudizio sulle persone anziane?

- solitamente lente nell'apprendimento di nuovi argomenti; bisogna ripetere le cose molte volte;
- hanno molta voglia di imparare e sono rispettosi nei nostri confronti;
- sono molto intelligenti, cordiali e gentili;
- si impegnano molto a capire la nostra generazione;
- sono molto più attive di quanto noi giovani pensiamo;
- qualche volta li spaventa l'assenza totale di privacy dei canali virtuali come Facebook, Twitter...;
- persone colte con tanti principi;
- sono più interessate e impegnate dei giovani nel voler apprendere le cose che non sanno;
- si ritengono sempre superiori e non stanno a sentire quando un ragazzo gli spiega qualcosa;
- le persone anziane sono persone che bisogna rispettare perché hanno fatto tanti sacrifici e hanno vissuto una vita intensa;
- alcuni possono pensare che sono molto retrogradi e che non gli interessa entrare nel "nostro mondo" ma non è così; basta un pò di pazienza per fargli raggiungere il nostro livello.

D8. Cosa ti ha trasmesso e ti ha fatto capire il tuo corsista anziano?

- tanti valori a cui oggi noi giovani non crediamo. Loro li hanno vissuti in prima persona;
- che non dobbiamo perdere i valori primari della vita stando sempre davanti al computer o al cellulare;
- le differenze di età e di conoscenze sono indici di aspetti di vita differenti che ognuno può completare con l'altro;
- che bisogna fare degli sforzi per raggiungere i propri obiettivi, anche i più banali;
- noi giovani abbiamo perso qualche valore della vita e non siamo più quelli di una volta; la nostra generazione è molto diversa dalla loro;
- il valore della famiglia, del mantenere i rapporti con persone lontane, il valore dell'amicizia e questo mi ha fatto capire che basta anche un piccolo gesto per fare felice qualcuno;
- il bello degli anziani è che sono molto amichevoli; mi ha fatto vedere i suoi figli su Facebook, la sua idea politica su Twitter;
- mi parlava della sua vita, della famiglia e dei figli (dove hanno studiato con chi sono sposati, i vari nipoti,...);
- che tutto è possibile; basta l'impegno e la buona volontà;
- bisogna aver sempre voglia di imparare;
- bisogna mettere impegno in qualsiasi cosa si faccia;
- si può vivere felice anche senza Internet;
- non è mai troppo tardi per imparare;
- è bello che i giovani, che potrebbero essere loro nipoti, diventano i loro insegnanti;
- si può insegnare tutto a tutti e a qualsiasi età.

D9. Cosa pensi di aver trasmesso al tuo corsista?

- velocità e capacità di apprendere le nuove tecnologie;
- l' utilizzo e il modo di gestire la tecnologia avanzata di oggi;
- tutte le cose brutte e gli imbrogli ma anche le cose belle che si possono fare con i social network;
- la manualità nell'uso del computer;
- come oggi sia importante l'uso del computer;
- credo di aver "modernizzato" il suo stile di vita con qualcosa in più da condividere anche con i nipoti;
- l'amore per i vari social network che conoscevano solo per sentito dire;
- come nell'utilizzare Internet bisogna basarsi sul ragionamento e avanzare per passi successivi;
- gli ho insegnato ad essere paziente;
- l' uso più approfondito di strumenti di comunicazione come i social network.

D10. Cosa ti è piaciuta di più di questa esperienza?

- ridere e scherzare insieme;
- trovarmi in cattedra;
- aver scoperto che le nuove tecnologie interessano anche alle persone anziane che hanno ancora voglia di imparare;
- la partecipazione attiva tra giovani e anziani e la complicità tra le due parti;
- interagire con una generazione di cui conoscevo ben poco;
- il rapporto che, lezione dopo lezione, si viene a creare con un nuovo nonno;
- la comunicazione che si è instaurata tra due generazioni molto differenti tra di loro;
- il fatto che anche ad una certa età le persone sono volenterose ad imparare nuove cose;
- la simpatia degli anziani e la condivisione e collaborazione tra noi studenti; se non sapevamo fare una cosa ci aiutavamo tra di noi;
- il fatto di aver provato l'esperienza di insegnante;
- il rapporto che si crea con l'anziano che segui;
- poter insegnare delle cose che sapevo ad altre persone;
- il mettermi in gioco come insegnante.

6. Discussione e conclusioni

Dalle risposte alle domande sintetiche fatte ai miei studenti, che per la prima volta svolgevano il ruolo di docenti e dalla risposta dei corsisti anziani, si nota subito che questo progetto è stato interessante per tutti i partecipanti. Gli studenti hanno trasmesso agli anziani la loro manualità tecnologica, il loro amore per questi social network a cui dedicano parte della loro vita. Gli adulti, da parte loro, hanno trasmesso agli studenti dei messaggi valoriali, l'impegno, la costanza, la saggezza, la memoria storica e la cultura. Io come supervisore o tutor del corso, non ho fatto altro che apprezzare la complicità che si è venuta a creare tra queste due generazioni che conversavano continuamente tra di loro scambiandosi consigli, idee e suggerimenti.

E' bello aiutarsi a vicenda ognuno dà all'altro quello che possiede o è in grado di dare. Questo momento di vita comune li ha portati gli uni a scoprire qualcosa degli altri: i ragazzi non pensavano che gli anziani avessero questa voglia di imparare e gli adulti non pensavano che i ragazzi potessero trovare e dedicargli del tempo nello spiegare, con amore, questi nuovi strumenti digitali. Gli anziani sono rimasti molto contenti di apprendere l'utilizzo di queste nuove tecnologie ma non si sono sottratti a dare ai giovani dei consigli quali quelli di avere più fantasia, capacità di ascolto, più voglia di studiare, di scrivere e leggere e di non isolarsi trascorrendo molto tempo su questi canali virtuali.

In questo progetto la tecnologia è stata un collante validissimo per le due generazioni. Insieme giovani e adulti possono contribuire ad abbattere il Digital Divide, ad acquisire una maggiore cittadinanza attiva e a migliorare la nostra comunità/società con lo scambio di idee, suggerimenti e proposte. L'affermarsi di una e-democracy globale, con l'uso delle tecnologie e dei canali virtuali, può favorire la ricerca di soluzioni innovative a problemi ambientali e sociali.

I dati ottenuti da questa prima analisi possono essere utilizzati come base di partenza per un'ulteriore indagine di tipo quantitativo che coinvolga un campione più grande di giovani e corsisti anziani. In questo modo si possono validare alcune ipotesi e ottenere dei risultati più generalizzabili. Non si deve però alimentare il pregiudizio che i ragazzi e gli anziani di oggi sono diversi da quelli di ieri ma bisogna sottolineare il fatto che oggi le nuove tecnologie interattive e collaborative si evolvono in maniera più veloce rispetto al passato e che hanno delle implicazioni maggiori nel settore sociale, politico e culturale. Oggi i computer,

soprattutto con l'avvento del web 2.0, hanno perso la loro valenza di elaboratori di calcoli complessi per dare maggiore spazio all'aspetto sociale e allo scambio di informazioni contribuendo alla creazione di un'intelligenza collettiva utile al benessere della società.

Bibliografia

Consoli D. Gli studenti nativi digitali e i social network. Atti di conferenza 27^a Didattica 2013 Tecnologie e metodi per la didattica del futuro, Pisa 7-9 maggio 2013, pp. 433 – 442.

Contri A., I nativi digitali: salviamoli prima che si friggono il cervello. Disponibile online in data 9/2/2013 su:<http://www.ilsussidiario.net/News/Educazione/2013/1/19/SCUOLA-Contri-i-nativi-digitali-Salviamoli-prima-che-si-friggano-il-cervello/355658>.

De Kerchove D., *Connected Intelligence, the Arrival of the Web Society*, Toronto, Somerville House, 1997.

Grossman L., Time's Person of the Year: You, Time Magazine. December 13, 2006, Available online: <http://www.imli.com/imlog/archivi/001051.html>.

Kaplan A. M., Haenlein Michael. Users of the world, unite! The challenges and opportunities of social media, *Business Horizons*, Vol. 53, Issue 1, 2010, pp. 59-68.

Levy P., *Collective Intelligence: Mankind's Emerging World in Cyberspace*, Perseus, Cambridge, 1997.

McAfee A.P., *Enterprise 2.0: The Dawn of Emergent Collaboration*. *Sloan Management Review*, Spring, Vol. 47, Issue 3, 2006, pp. 21-28.

McLuhan M., *Understanding Media: The Extensions of Man*; 1st Ed. McGraw Hill, NY, 1964.

O'Reilly T., What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, *International Journal of Digital Economics*, 65, 2007, pp.17-37.

Prensky M., Digital natives, digital immigrants, *On the Horizon*, 9, 5, 2001a, pp. 1-6.

Prensky M., Digital Natives, Digital Immigrants, Part II: Do They Really Think Differently?, *On the Horizon*, 9, 6, 2001b, pp. 1-9.

Sloep P., Boon J., Cornu B., Klebl M., Lefrere P., Naeve A., Scott P., and Tinoca L., A European research agenda for lifelong learning. *Int. J. Technol. Enhanc. Learn.* 3, 2 2011, pp. 204-228.

Tapscott D. *Growing up digital. The rise of the Net Generation*. Harvard Business Press, 2000.

Tapscott D. *Grown Up Digital: How the Net Generation is Changing Your World*, McGraw-Hill Publ.Comp, 2008.

Toffler A., *The Third Wave*, Bantam Books, US, 1980.

Informatica Solidale – Web per ogni età'

Claudio Tancini, Antonio Savarese
Informatica Solidale

Via Miramare, 9, Milano
info@informatica-solidale.org

Abstract. Technology Innovation is a major opportunity for social improvement, but not all the citizens can get full benefits from it. Elder people, even in well developed Countries, in many cases suffer about economical or cultural gaps that prevent them from internet services access. This document presents a program driven by the Italian NGO Informatica Solidale, who's target is to enable the digital inclusion of aged population.

Keywords:: Digital Divide, Digital inclusion, Elder

1. Introduzione

La Onlus **Informatica Solidale** nasce dall'impegno di professionisti ICT e del Terzo Settore per promuovere lo sviluppo sociale e ridurre le situazioni di disagio attraverso l'utilizzo delle tecnologie informatiche ed il volontariato.

I progetti sviluppati si articolano in diverse aree di intervento per affrontare situazioni di disagio relative a:

- Individui soggetti a disabilità
- Anziani
- Persone in difficoltà per problemi economici o di inserimento nella società e nel mondo del lavoro
- Soggetti aiutati da altre iniziative di altre organizzazioni NonProfit
- Soggetti aiutati da organizzazioni di Cooperazione Internazionale

Una delle aree di intervento di Informatica Solidale è l'impulso alla diffusione di tecnologia informatica per la riduzione del Digital Divide, una moderna forma di esclusione sociale e culturale, dovuta non solo alla mancanza di copertura di rete a larga banda, ma anche alla difficoltà nell'uso delle nuove tecnologie.

2. Il progetto Web per ogni Età

Il nostro obiettivo, per questo Programma, è la riduzione del Digital Divide culturale della popolazione, agendo in modo specifico verso persone anziane e/o persone in difficoltà economica.

Il programma, oltre alla organizzazione di corsi di alfabetizzazione informatica, prevede un **Servizio di Assistenza post Corso** attraverso laboratorio informatico permanente (che chiameremo “palestra informatica”).

A supporto del programma di assistenza sarà implementato un sistema di **E-Learning** con la funzione di rendere disponibile a tutor ed utenti un sistema semplice di recupero di documentazione e di lezioni on-line.

Queste ultime potranno essere seguite a casa o presso la palestra (limitando l'impiego di trainer e garantendo una maggiore omogeneità nelle lezioni)

Nel 2013 il progetto si focalizzerà sulla città di Milano, per avviare 3 palestre informatiche in 3 circoscrizioni già identificate, ma il programma proseguirà, con l'obiettivo di arrivare ad una palestra per circoscrizione e per allargare l'esperienza ad altre città nel 2014.

2.1 Il modello operativo

L'approccio tradizionale dei corsi di alfabetizzazione prevede solo alcuni giorni di training in aula, ed è invece carente nel seguire l'evoluzione dell'apprendimento e supportare i normali problemi durante il concreto utilizzo del personal computer e della rete a casa propria.

Soltanto un supporto continuativo post corso può permettere il raggiungimento di una reale autonomia ed evitare la dispersione del lavoro di preparazione.

Per questo l'impegno maggiore che viene dedicato dai nostri volontari è indirizzato alla organizzazione delle cosiddette “palestre informatiche”, luoghi attrezzati dove le persone possono recarsi per approfondire alcuni argomenti o superare alcune specifiche difficoltà.

Le palestre avranno una apertura fissa, attraverso un sistema di prenotazione potranno recarvisi tutte le persone over-60 che avranno aderito al progetto.

2.2 Il sistema di E-Learning

Per rendere più efficace il modello operativo stiamo sviluppando alcune partnership che consentiranno la disponibilità di un sistema di e-learning, che permetta la registrazione della documentazione, delle lezioni e, dove opportuno costruire specifici percorsi formativi.

L'obbiettivo del sistema di e-learning è duplice, il primo è quello di supportare le persone che partecipano ai corsi e alle palestre, permettendo l'accesso a documentazione aggiornata in modo semplice e di poter seguire lezioni di approfondimento in totale autonomia e anche da casa.

Il secondo obiettivo è invece di efficienza per l'organizzazione delle lezioni, infatti la disponibilità di docenti preparati è uno dei vincoli maggiori del progetto, la visualizzazione di lezioni preimpostate permetterà di poter organizzare i corsi prevedendo in aula solo volontari pronti per il supporto, e non necessariamente personale specializzato nella docenza.

2.3 La organizzazione delle sedi operative

Essendo una iniziativa NonProfit una particolare attenzione deve essere dedicata alla sostenibilità e alla economicità del programma.

Prima di avviare ogni sede operativa è necessario soddisfare alcuni prerequisiti:

- ⌚ disponibilità di **locali idonei**, spesso in partnership con altre associazioni, cooperative o con centri di supporto della terza età
- ⌚ disponibilità di **volontari** per organizzare i corsi e il supporto
- ⌚ disponibilità di **materiale informatico** per attrezzare la sede

Una volta soddisfatti i prerequisiti inizia la organizzazione dei corsi e del successivo supporto, e ogni centro operativo vive di vita propria pur potendo disporre, attraverso il sistema di e-learning, di tutto il materiale generato dalle diverse iniziative

3.3.Conclusioni

Informatica Solidale, attraverso il programma “**Web per ogni età**”, vuole portare **un contributo concreto alla Digital Inclusion**.

Il modello che stiamo sperimentando, per le sue caratteristiche di alta modularità, di minimo impatto economico e di grande coinvolgimento delle persone coinvolte, potrà essere facilmente replicato e quindi incidere in modo significativo al superamento degli ostacoli che oggi limitano in molti casi l'accesso alla rete ai ai servizi telematici alle persone della terza età.

Noi siamo disponibili per aiutare altre organizzazioni che vorranno seguire il nostro approccio, maggiori informazioni sono reperibili sul portale della associazione : **www.informatica-solidale.org**



Confronto tra Metodologie Didattiche per l'Apprendimento delle Discipline Giuridiche

Teresa Consiglio¹, Elvira Vitiello¹,
¹Università Telematica "Leonardo da Vinci"
Piazza S. Rocco, 2 - 66010 Torrecchia Teatina (CH)
t.consiglio@unidav.it
e.vitiello@unidav.it

Abstract. *Knowledge transmission nowadays is adapted to technological evolution. Computer has become an indispensable didactic tool. It allows you to create a variety of resources and activities very helpful in order to draw the attention of students and to foster the research. It is a teacher's task making the use of technology in didactic is perceived by students as an integrated instrument like the blackboard or commented codes. In higher education the deliver of online lectures besides the face-to face lectures is more and more widespread. In this paper we compare this two teaching methods delivered in two Italian Universities in master programs of law field. In particular, we analyze a module of Philosophy of Law attended either in presence or in e-learning.*

Keywords: e-Learning, higher education, didactic of Law.

1. Introduzione

Nel corso degli anni il progresso tecnologico, strettamente legato a quello della comunicazione, ha inevitabilmente avuto come conseguenza un'evoluzione dei sistemi di trasmissione della conoscenza in generale e della didattica in particolare [Bacceli, 2002].

Molte Università, ormai, si avvalgono della formazione in rete, di sperimentazioni attraverso l'utilizzo di mezzi informatici, di internet e della modalità e-learning. L'istruzione acquista non solo un canale di comunicazione in più, ma si ha una elaborazione del concetto stesso di insegnamento a distanza [Capineri, 2003].

L'avvento delle reti nell'istruzione telematica evidenzia il modo nuovo in cui è concepita l'attività didattica. L'elemento fondamentale che costituisce il tratto differenziale della tipologia comunicativa consiste nel fatto che con l'avvento delle reti si passa da una tipologia comunicativa del tipo uno a molti ad una tipologia comunicativa molti a molti, in cui non solo il discente può interagire più rapidamente con il docente ma può anche stabilire interazioni e rapporti cooperativi con tutti i partecipanti [Ardizzone e Rivoltella, 2008, Cornoldi, 1995].

Congresso Nazionale AICA 2013

Mentre negli interventi tradizionali a distanza l'aspetto comunicativo è sempre stato considerato meno importante rispetto alla fruizione individuale dei materiali, nella formazione in rete è proprio la collaborazione tra i partecipanti il meccanismo che stimola e facilita l'apprendimento: il docente cambia il suo ruolo. Egli non soltanto trasmette le conoscenze, ma diventa il tutor, il facilitatore, l'animatore in rete. Il processo formativo assume sempre più il carattere di attività di negoziazione cooperativa, peer tutoring [Trentin, 1998].

Gli studi di Bourdieu [Bourdieu e Passeron, 1971] e degli altri esponenti del funzionalismo pedagogico hanno evidenziato che la principale attenzione degli insegnanti è sempre stata quella di favorire la trasmissione culturale e con essa l'acquisizione di quei comportamenti e quei valori sui quali certa cultura si reggeva. Qualcosa di non dissimile accade normalmente nel campo della media education [Thompson, 1995]. Si pensi alla funzione dell'analisi (dei messaggi, dei testi) e alle modalità attraverso le quali normalmente essa viene condotta: sono le modalità dell'analisi guidata, dell'analisi esemplare in cui l'insegnante mostra alla classe come gli strumenti debbano essere applicati fornendo allo stesso tempo una analisi-modello del testo su cui si sta lavorando.

Ciascun docente sarà sollecitato ad interventi orientati a riconsiderare il tema dello scambio, a ripensare il proprio stile comunicativo, a pensare a soluzioni innovative di comunicazione della comunicazione, ad introdurre attività cooperative.

2. scopo del lavoro

In questo lavoro si vogliono confrontare due metodologie didattiche adottate presso due Atenei italiani in corsi di laurea di area giuridica. In particolare si vuole comparare la lezione di Filosofia del Diritto in presenza con quella in modalità e-learning.

Due sono le ragioni principali quanto alla scelta degli oggetti didattici: l'unicità del docente in entrambi i corsi di laurea e la diretta esperienza di lavoro di uno degli autori in entrambi gli Atenei.

2.1 Il linguaggio giuridico e la didattica del diritto

Il linguaggio giuridico, ossia il linguaggio con cui sono formulate le norme giuridiche, è un linguaggio *tecnicizzato* in quanto, insieme a parole di uso comune, utilizza termini che appartengono in modo specifico al mondo giuridico.

“Come altri linguaggi specialistici, il linguaggio giuridico è un linguaggio non formalizzato, che adotta la sintassi e la semantica del linguaggio naturale. Inoltre il linguaggio giuridico, a differenza di altri linguaggi specialistici, non costituisce un sottoinsieme rigidamente delimitato: il diritto regola tutti gli ambiti dell'attività umana e prende a prestito, di volta in volta, il vocabolario della materia da disciplinare.

Il linguaggio giuridico “naturale”, così come ci è consegnato dalla tradizione, costituisce il mezzo linguistico di impiego generale più adeguato per la trattazione dei problemi giuridici. Infatti, solo il linguaggio naturale è

sufficientemente ricco e versatile da consentire di descrivere i molteplici oggetti, comportamenti, ambiti sociali regolati dal diritto. Solo uno strumento umano e sociale come il linguaggio naturale è in grado di riflettere gli aspetti umani e sociali presenti in ogni problema giuridico; solo la sua flessibilità consente di rappresentare i casi giuridici dando espressione all'individualità, alle particolari sfaccettature di ciascuno di essi. Quanto detto non esclude che, rispetto a problemi e scopi particolari, il linguaggio giuridico in uso possa rivelarsi insufficiente e incompleto, e che taluni problemi giuridici possano essere meglio affrontati mediante strumenti concettuali nuovi e specialistici, alcuni dei quali possono essere tratti da discipline formali (matematica, logica, informatica, teoria dei giochi, statistica, ecc.) o svilupparsi sul modello di queste" [Sartor, 1998].

Il lessico giuridico è necessario e bisogna appropriarsene ma la cosa peggiore che si possa fare all'inizio di un percorso formativo è iniziare fornendo delle definizioni stipulative nella convinzione che, in tal modo, l'allievo si formi il lessico necessario per introdursi alla disciplina. Occorre al contrario, una disamina, di situazioni tratte dalla vita quotidiana mediante la quale far emergere la dimensione giuridica e quella economica presente in tali fatti. Solo così l'allievo si formerà degli schemi mentali grazie ai quali rappresentare a se stesso una molteplicità di fenomeni interconnessi tra loro e il linguaggio giuridico entrerà in tali schemi con logicità e significato e solo così la scuola formerà delle "menti complete" [Bruner, 1988] e non sarà "cinghia di trasmissione della cultura dominante" [Bacceli, 1998].

La didattica del diritto, come in generale la didattica di qualsiasi disciplina, non potrà porsi obiettivi avulsi dal contesto in cui l'insegnamento del diritto è preposto. Perché si insegna diritto e perché del diritto si insegnano determinati contenuti? La risposta più semplice è quella di formare un cittadino responsabile. Per fare ciò l'insegnante deve possedere le competenze per formulare degli obiettivi di insegnamento disciplinari scientificamente corretti, insegnabili perché apprendibili e utilizzabili, perché funzionali a precisi bisogni formativi.

Lo studio effettuato nel presente lavoro, ha come oggetto l'insegnamento di Filosofia del Diritto.

I casi che sono utilizzati come spunti essenziali per riflettere quanto all'ipotesi sperimentale della ricerca sono la lezione in presenza del corso di laurea in Servizi Giuridici per l'Impresa (Ateneo "G. d'Annunzio") e la lezione in modalità e-learning del corso di laurea in Giurisprudenza (Ateneo telematico "Leonardo da Vinci"). In particolare vogliamo comparare la lezione di Filosofia del Diritto in presenza con quella in modalità e-learning.

Due sono le ragioni principali quanto alla scelta degli oggetti didattici: l'unicità del docente in entrambi i Corsi di Laurea e la diretta esperienza di lavoro di uno degli autori in entrambi gli Atenei.

L'indagine esplorativa vuole mettere in luce i seguenti aspetti

- In quali rapporti sono gli studenti con le Tecnologie dell'Informazione e della Comunicazione?

- Come sono predisposti verso la lezione on-line?

- Quali effetti hanno le due differenti modalità di erogazione delle lezioni sul livello di preparazione degli studenti?

3. Materiali e Metodi

Il corso di laurea in Servizi Giuridici per l'Impresa dell'Università degli Studi "G. d'Annunzio" di Chieti-Pescara prevede, nel piano di studi del I anno, l'insegnamento di Filosofia del Diritto per complessivi nove CFU. L'ipotesi sperimentale di questo lavoro si è sviluppata da un'analisi di tale insegnamento e, nello specifico, il modulo relativo all' Interpretazione giuridica.

Su base volontaria, abbiamo suddiviso in due gruppi i settantatré studenti partecipanti, con un'età media di ventuno anni e non in possesso di un titolo di istruzione universitaria.

Un gruppo ha seguito il modulo in modalità tradizionale e l'altro gruppo in modalità e-Learning.

L'esperienza si è sviluppata in tre fasi. Nella prima fase ai due gruppi sono stati somministrati i questionari di inizio modulo; nella seconda fase gli studenti hanno seguito le lezioni secondo la modalità didattica prescelta; nella terza sono stati somministrati i questionari di fine modulo. Tutti i partecipanti hanno sostenuto l'esame finale con lo stesso docente, indipendentemente dalla modalità prescelta.

I questionari somministrati ex ante ci hanno permesso di raccogliere informazioni di carattere generale riguardanti l'età e la formazione scolastica, ma, soprattutto, di capire il tipo di rapporto con i mezzi informatici e con l'utilizzo di Internet. I questionari ex post ci hanno permesso di rilevare il gradimento delle lezioni e le considerazioni degli studenti in merito alla fruizione in e-learning.

Organizzazione del corso in presenza

Il modulo di *Interpretazione giuridica* si compone di cinque lezioni.

Le lezioni si sono svolte in cinque giorni successivi con una durata di due ore ciascuna.

Il docente distribuisce in aula il materiale didattico consistente in dispense cartacee relative all'argomento delle lezioni.

L'interazione tra docente e studenti è sincrona, ossia si realizza durante lo svolgimento delle lezioni in aula e durante l'orario di ricevimento stabilito dal docente; e asincrona mediante l'utilizzo della posta elettronica. L'interazione tra studenti, è sincrona durante le lezioni in aula ed, eventualmente, asincrona al di fuori della classe.

Organizzazione del corso in modalità e-learning

Il corso si caratterizza per due figure di riferimento, il docente, autore delle lezioni fruibili in ambiente digitale e il tutor di contenuto e di metodo. La frequenza si è svolta senza vincoli di giorni e orari.

Il modulo si compone di cinque lezioni audio/video della durata di circa venti minuti ciascuna, di dispense stampabili e di un test di autovalutazione che permette allo studente di saggiare l'apprendimento. Caratteristica importante di questo test è la possibilità di eseguirlo più volte e, laddove non è superato con

successo, di capire dove e perché si è sbagliato visionando le risposte esatte mediante i feedback del docente presenti sulla piattaforma. L'interazione tra docente, tutor e studenti è asincrona mediante l'utilizzo della posta elettronica e del forum didattico e sincrona attraverso l'uso della chat e della videoconferenza. Sono stati creati due forum, un forum "C@fè" all'interno del quale gli studenti sono stati invitati a socializzare e un forum disciplinare relativo a questioni di didattica e di contenuti. Il moderatore per entrambi i forum è stato il tutor.

4. Risultati

Dall'elaborazione dei dati raccolti abbiamo ottenuto i seguenti risultati.

Relativamente alle domande riguardanti il possesso di un PC e l'utilizzo di Internet, tutti hanno risposto affermativamente.

Per ciò che concerne altre esperienze formative, molti dichiarano l'assenza di esperienze formative ulteriori, pochissimi dichiarano di aver frequentato corsi di informatica.

Come sopra specificato la suddivisione degli studenti in due gruppi è di tipo volontario, pertanto l'ultima domanda dei questionari è diversa a seconda della scelta effettuata.

La maggior parte di coloro che hanno scelto la fruizione del modulo in modalità e-learning con la risposta "credo possa giovarmi all'apprendimento di questo modulo" ha presumibilmente eseguito un ragionamento di tipo utilitaristico, con la finalità di apprendere meglio il contenuto delle lezioni al fine di un miglior risultato in sede di esame finale.

La maggior parte dei membri del gruppo che non ha scelto la modalità e-learning, ha dichiarato di preferire la tradizionale lezione in presenza. Per una minima parte di studenti, invece, è stata una scelta obbligata, legata al non possesso a casa di un'adeguata connessione a internet.

Agli studenti che hanno deciso di seguire il corso in modalità e-learning sono state somministrate domande specifiche relative alla misura dei vantaggi attesi.

Tra le risposte si evidenziano: "l'organizzazione dei momenti di formazione secondo i tempi dello studente", "la conciliazione dello studio con altre attività" e "la comunicazione più veloce e diretta con il docente".

Il supporto di un tutor, anche se considerato positivamente, non è tra i vantaggi ritenuti basilari.

Elaborazione dei dati finali (gruppo e-learning)

È stato interessante scoprire come gli studenti siano stati positivamente colpiti dalla nuova modalità didattica loro prospettata: in particolare si rileva che nessuno studente ha espresso un giudizio negativo sul gradimento di tali lezioni. Più della metà ha espresso un giudizio "abbastanza positivo" e la restante parte un giudizio "molto positivo"; buono anche il giudizio sul materiale didattico messo a disposizione.

Anche per le domande sulla qualità complessiva del corso e la disponibilità del docente non vi sono risposte negative.

Si noti che alla domanda “*solo se credi che il corso on-line sia utile, scegli due motivazioni tra le seguenti*”. Tra le risposte più frequenti: il 47% valuta utile il fatto che un corso in modalità e-learning permette allo studente di ascoltare più volte la lezione del docente; il 33% valuta utile il fatto che ciò che viene proposto con le slide e spiegato a voce è poi possibile stamparlo, in luogo degli appunti.

Lo strumento forum non è stato considerato utile al fine della preparazione di questo modulo.

Elaborazione dei dati finali (gruppo in presenza)

Diversamente da quanto espresso dal gruppo e-learning, alcuni studenti che hanno frequentato le lezioni in presenza hanno riscontrato delle difficoltà nell'elaborazione personale dei contenuti trattati dal docente, nonostante il materiale didattico di supporto. Una possibile spiegazione potrebbe essere che, dato l'alto livello di astrazione dei contenuti proposti, probabilmente l'impegno e la concentrazione in aula non sono stati sufficienti e gli studenti avrebbero beneficiato della possibilità di riascoltare più volte le lezioni per arrivare a un'attiva costruzione del sapere.

5. Discussione

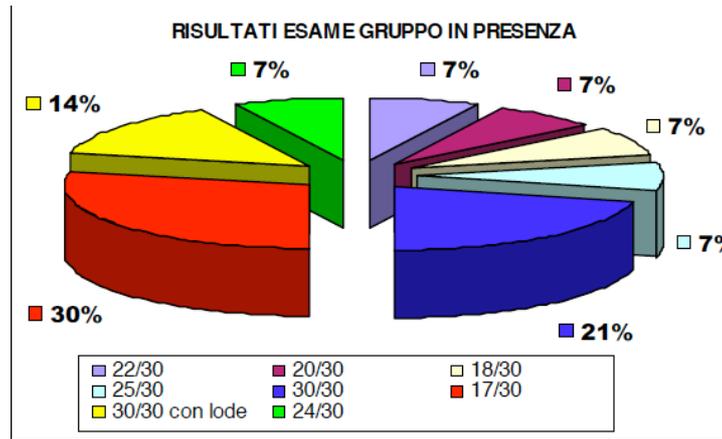
Per entrambi i gruppi l'esame si è svolto in presenza nello stesso giorno.

Il modulo oggetto di sperimentazione, si ricorda, riguarda le lezioni di *Interpretazione giuridica* nell'ambito del corso di *Filosofia del Diritto*.

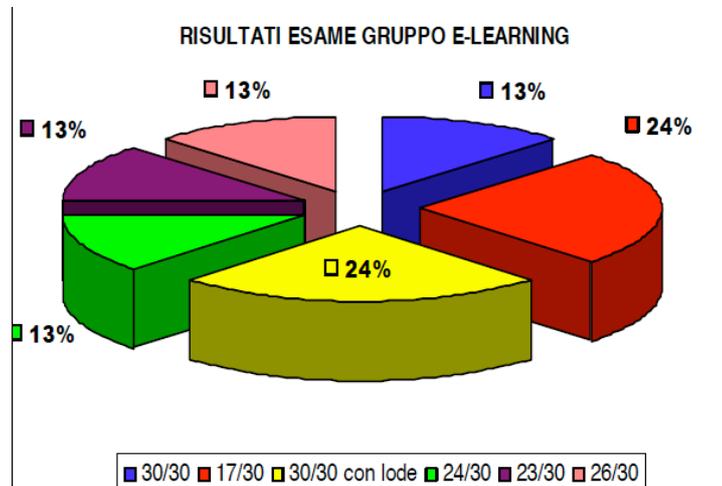
Il corso di filosofia del diritto è stato suddiviso dal docente in tre moduli: I modulo (Kelsen), II modulo (Interpretazione), III modulo (Carcattera). Per i singoli moduli è stato previsto un esame di fine modulo con attribuzione di votazione finale.

Dal risultato delle valutazioni riguardanti l'esame di fine modulo, si evince in prima battuta che i voti degli studenti che hanno frequentato in presenza (Fig.1) hanno una maggiore dispersione rispetto a quelli attribuiti ai frequentanti in modalità e-learning (Fig. 2).

La votazione media degli studenti frequentanti on-line è leggermente migliore di quella in modalità tradizionale. In particolare: gli studenti frequentanti la lezione frontale hanno ottenuto una media voti di 23,3p (Fig.3) rispetto alla media voti di 24,75p, riportata dai frequentanti in modalità e-learning (Fig.4).



**Fig. 1 -
votazione degli studenti che hanno frequentato in presenza**



**Fig. 2 -
votazione degli studenti che hanno frequentato online**

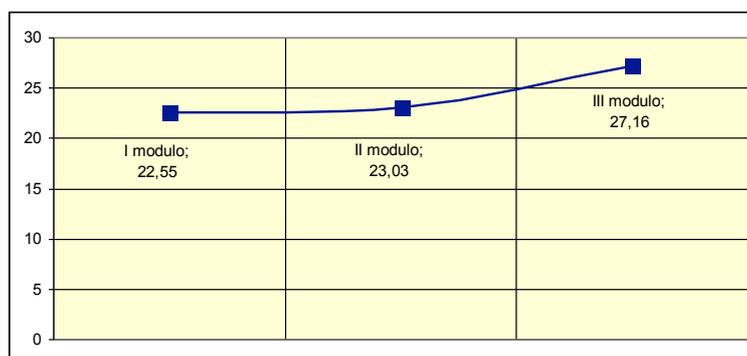


Fig. 3 - votazione media degli studenti frequentanti in presenza

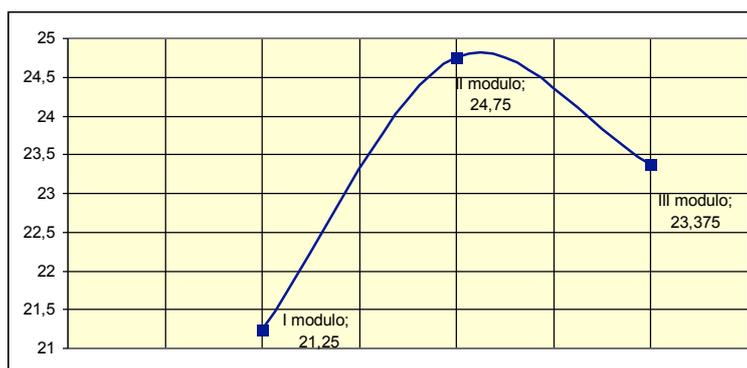


Fig. 3 - votazione media degli studenti frequentanti online

Dall'esame dei grafici, si evince che: gli studenti che hanno frequentato le lezioni frontali hanno avuto un miglioramento progressivo dal I al III modulo in termini di votazione e quindi di apprendimento riscontrato dal docente.

Gli studenti che hanno seguito parte dei moduli in presenza e parte in e-learning non evidenziano un miglioramento costante ma la curva mostra un picco in salita in corrispondenza del II modulo frequentato in modalità on-line. Ci siamo spiegate questo risultato con il fatto che, a seguito di interviste effettuate ex post, gli studenti ci hanno detto che la maggior parte di loro era fuori sede o studente lavoratore. Probabilmente il tempo risparmiato per gli spostamenti ha influito positivamente sulla qualità dello studio.

6. Conclusioni

Anche negli anni accademici 2010 e 2011, si è deciso di riproporre l'esperimento nel medesimo corso di laurea ma pochissimi studenti hanno aderito e quindi non si è potuto procedere con la comparazione.

Tuttavia si è deciso di somministrare i questionari per raccogliere dati di carattere generale.

Dall'esito negativo delle ultime due esperienze si evince che gli studenti non sono ben predisposti verso un tipo di lezione on-line a scapito di quella tradizionale, anche se la maggior parte di loro si è detta ben disposta a sperimentare altri insegnamenti in modalità e-learning.

In generale la scelta di un percorso formativo in e-learning non è pregiudicata dalla non conoscenza di questa modalità, ma emergono altri fattori quali il non contatto umano in primis forse, a nostro avviso, legato alla giovane età degli intervistati.

Sotto il profilo dell'apprendimento, valutato con esame finale, si rileva che gli studenti che hanno seguito le lezioni in e-learning, seppur di poco, hanno ottenuto una votazione media più alta. Situazione, questa ultima, rilevata e osservata, ma sulla quale non vengono tratte conclusioni definitive considerato il numero ridotto di studenti coinvolti.

Bibliografia

Ardizzone P., Rivoltella P.C., "Media e tecnologie per la didattica", Vita e Pensiero - Milano, 2008

Bacelli G., "Un'area da esplorare: la didattica dell'economia", Pisa - Roma, 1998, p.230

Bacelli G., "Didattica del Diritto", Elemond Scuola & Azienda - Milano, 2002

Bruner J., "La mente a più dimensioni", Roma - Bari, 1988, p.104

Capineri C., "Reti di comunicazione nell'era dell'informazione", Storia e Futuro, N° 2, 2003, p. 1-2

Carcattera G., "Corso di Filosofia del Diritto", Bulzoni - Roma, 1996, p.72ss.

Cornoldi C., "Metacognizione e apprendimento", Il Mulino - Bologna, 1995

Trentin G., "Insegnare e apprendere in rete", Zanichelli - Bologna, 1998

Bourdieu P., Passeron J.C., "Les Heritres: les étudiants et la culture", Trad. Ital. I Delfini: gli studenti e la cultura Guaraldi - Firenze, 1971

Thompson T.B., "The Media and Modernity: A Social Theory of the Media" Stanford University Press, 1995.

Sartor G., "I linguaggi (e i sistemi) informatici e linguaggio giuridico". In: Proceedings of the conference: Il diritto nella società dell'Informazione, Florence: IDG, 1998

A survey on the math test at the A-level scientific high school-leaving exam in Italy

[Indagine Nazionale sulla Prova di Matematica agli Esami di Stato nei Licei Scientifici Italiani]

Marco Scialdone, Salvatore Venticinque
Dipartimento di Ingegneria Industriale e dell'Informazione
Seconda Università degli Studi di Napoli
Via Roma 29, 81031 - Aversa (CE)

marcoscialdone@hotmail.com, salvatore.venticinque@unina2.it,

Abstract. In this paper we present an analysis of results about a national survey, that was carried out in July 2012, on the math test of school-leaving exam in Italian scientific high schools. The Ministry of Education invited all committees to answer an on-line questionnaire. The first part of the survey is about the students' results. The second part concerns teaching and assessment characteristics. In the end committee members provide a critical evaluation of the math test.

Keywords: mathematic, high school leaving, statistical analysis.

Introduzione

La valutazione e la certificazione delle competenze, dei risultati scientifici e didattici attraverso indici misurabili è oggetto di interesse in diversi contesti, nazionali e internazionali, e in particolare in diversi livelli del sistema di istruzione. Tra le iniziative più note si ricordano l'indagine Invalsi [Lasorsa e Amici, 2010] per la valutazione della scuola primaria, la più recente valutazione delle Università condotta dall'ANVUR e i nuovi criteri, bibliometrici e non, definiti per il conseguimento delle abilitazioni per il ruolo di professore di prima e seconda fascia nelle sistema universitario italiano. La facile raccolta delle informazioni per mezzo delle tecnologie web e la loro condivisione attraverso formati aperti consente oggi di disporre di una ricchezza di informazioni, accessibile e processabile da tutti, per ogni eventuale riflessione.

Dal 2001 la prova scritta di matematica agli Esami di Stato nei licei scientifici italiani è strutturata in problemi e quesiti. I candidati sono invitati a svolgere un problema a scelta tra i due proposti e cinque quesiti tra dieci. Le prove di matematica sono differenti a seconda degli indirizzi di studio. Dallo stesso anno 2001 è realizzata un'indagine nazionale sui risultati della prova, curata da

ispettori tecnici. Obiettivo dell'indagine è capire l'oggetto e i metodi di studio della matematica nei licei scientifici, raccogliendo le opinioni dei commissari che sono chiamati a rispondere a una serie di domande circa la complessità e la coerenza della prova rispetto ai programmi svolti, le difficoltà incontrate dagli studenti e le motivazioni, e infine i risultati ottenuti. Si accompagna, a tale iniziativa, la proposta di adozione di un metodo uniforme di valutazione sul territorio nazionale, pubblicando una griglia che definisce i criteri di correzione, e che viene puntualmente completata con i pesi assegnati ad ogni problema e quesito al termine dello svolgimento della prova. Dal 2001 il sito www.matmedia.it, che ospita l'indagine, si è posto come riferimento per docenti ed alunni, offrendo riflessioni, critiche e spunti. La Facoltà di Ingegneria della Seconda Università degli Studi di Napoli, dal 2009, il Dipartimento di Ingegneria Industriale e dell'Informazione (www.diii.it), dal 2013, offrono un supporto tecnico alla conduzione dell'indagine e all'analisi dei risultati [Venticinque et al, 2010]. In questo lavoro si presentano i risultati della sessione 2012.

Partecipazione all'Indagine

La prova di matematica dell'Esame di Stato nell'anno 2012 è stata sostenuta da 109.495 studenti, raggruppati in 5103 classi. Il 12 giugno 2012 la Direzione Generale per gli ordinamenti scolastici e per l'autonomia scolastica, attraverso una circolare inviata a tutti i licei italiani, dà notizia dello svolgimento dell'indagine e invita tutti i commissari a partecipare, rispondendo al questionario proposto attraverso il sito www.matmedia.it. E' data la possibilità di rispondere ai questionari dal 25 giugno al 14 luglio. Il database delle risposte viene, come ogni anno, reso disponibile a tutti, consentendone sia la navigazione web sia il download. Ogni docente è in grado di confrontare le informazioni riguardanti la propria classe con i risultati delle altre classi nella propria scuola, delle scuole nel proprio territorio o nell'intero contesto nazionale. I commenti liberi dei commissari, e quindi le opinioni di altri docenti, rappresentano inoltre un'ulteriore base di riflessione. Al termine della fase di raccolta delle informazioni, risultano inviati 2850 questionari su un totale di 5103 classi, per un totale di 58.920 studenti pari al 53,8% dei candidati. Si vuole evidenziare come tale percentuale non rappresenti un piccolo campione statistico, ma un'enorme quantità di informazioni che fornisce una fotografia delle classi italiane. La Fig. 1 mostra la partecipazione al questionario dal 2009 al 2012.

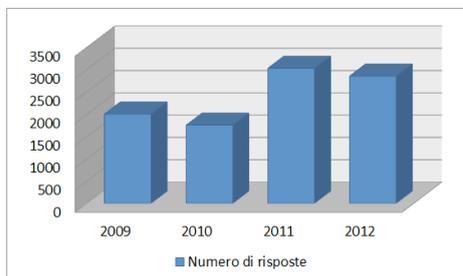


Fig. 1 – Numero di questionari pervenuti dal 2009 al 2012

Si evince che nel 2012 il numero di partecipanti è leggermente calato rispetto al 2011. La causa di questo calo può essere attribuita al fatto che il numero delle classi è diminuito di 800 unità; inoltre, a causa del terremoto che ha colpito l'Emilia Romagna, molte classi sono state esonerate dal sostenere la prova scritta di matematica. Infine una diminuzione si nota ad anni alterni, ogni qual volta il membro interno della commissione di esame è il professore di matematica.

La Fig. 2 mostra l'adesione percentuale per ogni singola regione. La percentuale è calcolata attraverso il rapporto tra numero di questionari inviati e numero di classi che hanno sostenuto la prova nella corrispondente regione. Si nota che il picco massimo, 100%, si è avuto per l'Abruzzo e il minimo, 23.7%, per il Trentino Alto Adige.

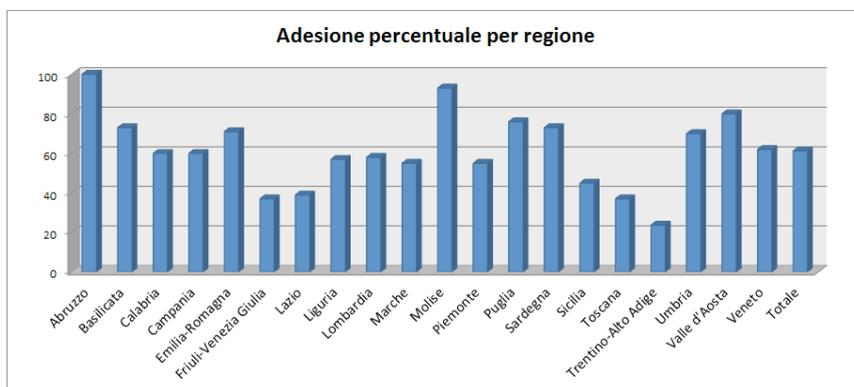


Fig. 2 – Adesione percentuale per regione

I licei scientifici sono classificati secondo 4 indirizzi di studio. Il grafico in Fig. 3 mostra la distribuzione dei candidati per ogni singola regione, suddivisi in base all'indirizzo di studi.

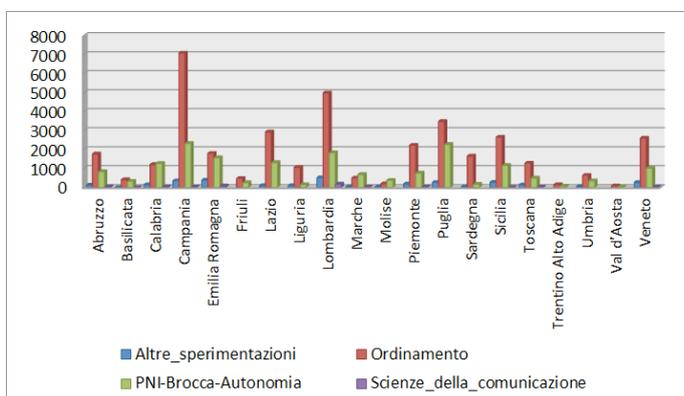


Fig. 3 – Numero di candidati suddivisi per indirizzo e regione

E' facile notare che il numero maggiore di studenti si ha nell'Ordinamento e nel PNI. In particolare, sempre dal grafico in Fig. 3, si evince che, tranne in Calabria, Marche e Molise, il numero di candidati dell'indirizzo Ordinamento è maggiore rispetto al PNI e a qualsiasi altro indirizzo. Per questo motivo l'attenzione sarà focalizzata su questi due indirizzi la cui principale differenza è un maggior numero di ore dedicate alla matematica nelle sperimentazioni PNI (3 per l'indirizzo di Ordinamento e 5 per il PNI).

Svolgimento della prova

La prima analisi riguarda la scelta che hanno effettuato i candidati verso una traccia piuttosto che l'altra. Il problema 1 per l'indirizzo di Ordinamento proponeva lo studio di due funzioni piuttosto standard, mentre il secondo problema proponeva la costruzione e lo studio di un luogo geometrico. Entrambi, come ultimo quesito, richiedevano il calcolo del volume di un solido di rotazione. Per l'indirizzo PNI, invece, il primo problema invitava i candidati allo studio di una funzione di cui però non si proponeva l'equazione, ma una rappresentazione grafica della derivata più altre informazioni. Il secondo problema, formulato in maniera più standard, proponeva due funzioni e le rispettive equazioni, ne chiedeva lo studio, il calcolo di un'area e un volume di rotazione. Per motivi di spazio si rimanda alle tracce complete sul sito www.matmedia.it. Dalla Fig. 4 è possibile notare un forte sbilanciamento nella scelta in entrambi gli indirizzi. Nel caso dell'Ordinamento si può notare che la maggior parte degli studenti ha scelto il problema 1, trovandolo, evidentemente, più semplice rispetto al 2. La scelta si inverte nel caso del PNI anche se lo sbilanciamento è minore rispetto all'Ordinamento. In quest'ultimo caso, infatti, il problema 1 è stato scelto da oltre 30000 studenti mentre il 2 da circa 3000. Nel caso del PNI il problema 1 è stato svolto da circa 4000 candidati e il problema 2 da poco più di 12000.

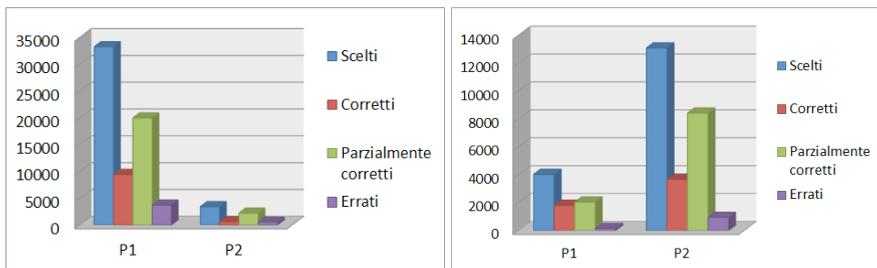


Fig. 4 – Numero di candidati che hanno scelto il primo o il secondo problema (a sinistra Ordinamento, a destra PNI)

E' interessante notare (fig. 5), come a differenza degli altri anni, nel caso del PNI è stato svolto maggiormente un compito con elementi che nel passato non venivano scelti. Questo ha portato ad una distribuzione degli insuccessi nello svolgimento dei problemi diversa rispetto agli anni precedenti. Mentre in passato l'insuccesso maggiore (in percentuale) si aveva nel problema meno scelto, perché evidentemente più complesso, quest'anno nel caso del PNI

l'insuccesso maggiore si è avuto in quello più scelto e quindi in quello che i candidati ritenevano più semplice. Infatti, il maggior insuccesso, sia nel caso dell'Ordinamento che nel caso del PNI, si è avuto nel problema 2 (fig. 5), meno scelto dagli studenti dell'Ordinamento e maggiormente scelto nel PNI (fig. 4).

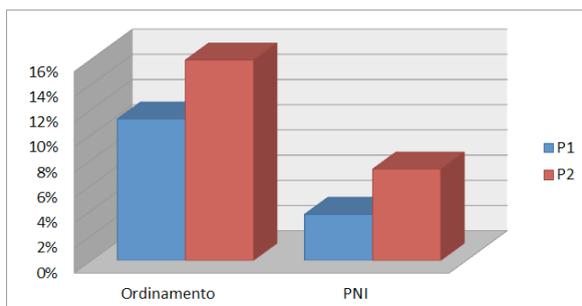


Fig. 5 – Insuccesso nello svolgimento dei problemi

A questo punto è parso interessante calcolare la polarizzazione della scelta verso una stessa traccia. Nel grafico in fig. 6 è riportato il numero di classi in cui una certa percentuale dei candidati ha scelto una stessa traccia. Oltre a notare l'elevato numero di classi in cui tutti gli studenti scelgono la stessa traccia, è stato interessante verificare come la scelta degli studenti si distribuisce per l'Ordinamento e per il PNI. I candidati dell'Ordinamento hanno scelto praticamente tutti la traccia che hanno ritenuto più semplice, mentre gli studenti del PNI che hanno scelto tutti la stessa traccia si sono orientati verso entrambi i problemi.

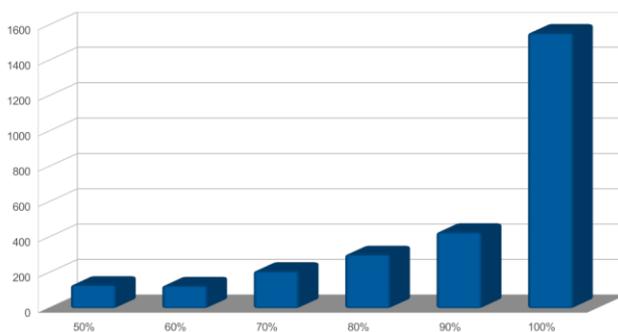


Fig. 6 - Polarizzazione degli studenti verso una stessa traccia

Per quanto riguarda i quesiti, la Fig. 7 mostra le scelte effettuate dai candidati dell'Ordinamento. I meno scelti risultano essere quelli in comune col PNI. Infatti, i diversi insiemi di 10 quesiti somministrati nei due diversi indirizzi presentavano tre quesiti comuni per l'esattezza il 5, il 7 e il 9. I quesiti proponevano, rispettivamente, un calcolo combinatorio, il calcolo di un angolo di un tetraedro e la risoluzione del problema di Erone (citando il periodo storico del matematico e probabilmente disorientando in questo modo gli studenti).

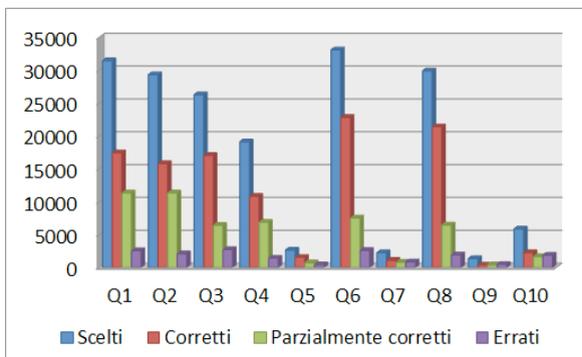


Fig. 7 – Numero di scelte per i quesiti di Ordinamento

Per l'indirizzo Ordinamento si ha che proprio i quesiti meno scelti risultano quelli con le maggiori percentuali di errore che vanno dal 35% a oltre il 40% (fig.8). Una piccola eccezione la fanno il quesito numero 5 che, nonostante sia stato poco scelto, è stato svolto abbastanza correttamente, e il numero 10 che invece è stato scelto da molti candidati, ma che ha una percentuale di errore molto alta. Esso chiedeva quale, tra 4 funzioni, era sempre positiva per ogni x.

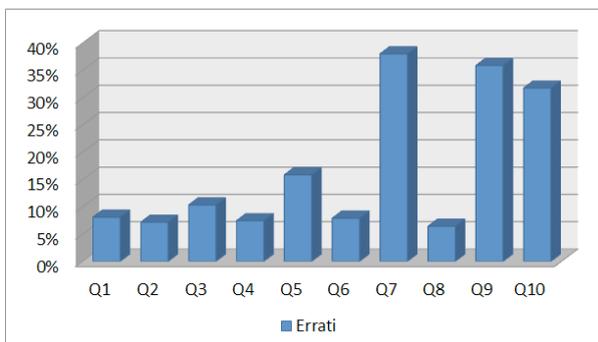


Fig. 8 – Percentuale di errore per i quesiti di indirizzo Ordinamento

Per quanto riguarda il PNI, i candidati si distribuiscono in modo più uniforme rispetto agli studenti dell'ordinamento, dimostrando difficoltà minori e una più completa preparazione rispetto agli argomenti proposti. I quesiti sbagliati maggiormente, anche in questo caso, sono risultati quelli meno scelti, ad eccezione del 2 che è stato scelto da un sostanzioso numero di studenti. Tale quesito chiedeva di riportare, commentandoli e senza calcolarli, gli integrali di due solidi di rotazione. Un'altra eccezione la fa il quesito numero 9, comune all'Ordinamento, che seppur poco scelto, ha un percentuale di errore non troppo elevata. La Fig. 9 mostra, rispettivamente, la percentuale di scelta e quella di insuccesso relativa ai quesiti in comune ai due indirizzi.

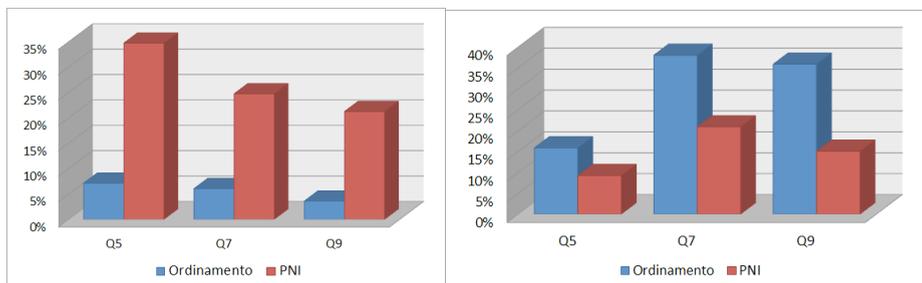


Fig. 9 – Percentuali di scelte (sinistra) e di insuccessi (destra) dei quesiti comuni

Dai grafici si evince che, come accade di solito, questi quesiti sono scelti maggiormente dai candidati del PNI, e vengono sbagliati maggiormente dai candidati dell'Ordinamento.

Utilizzo della griglia e valutazione della prova

Dal sito Matmedia era possibile scaricare una griglia di valutazione elaborata nel corso di seminari nazionali da una rappresentanza di docenti provenienti dalle diverse regioni. La griglia è il frutto di un ampio lavoro di studio e di ricerca che, se condivisa ed adottata su larga scala, può costituire lo strumento per conseguire una maggiore uniformità nella valutazione della prova, contribuire a renderne comparabili i risultati e costituire altresì l'occasione per un'azione decisamente utile sul piano scientifico e della crescita della cultura della valutazione [www.matmedia.it]. La griglia definiva gli indicatori e i punteggi massimi da assegnare ad ogni criterio per i diversi problemi e quesiti.

Dall'analisi dei risultati si nota che la griglia è stata adoperata praticamente dalla metà delle commissioni, il 49% per l'esattezza. La Fig. 10 mostra la percentuale di utilizzo della griglia per ogni singola regione. L'utilizzo maggiore si è avuto nel Molise dove si è arrivati al 95% circa. Il minor utilizzo, invece, si è avuto in Val D'Aosta dove non è stato raggiunto nemmeno il 15%.



Fig. 10 – Percentuale di utilizzo della griglia per ogni regione

Nella compilazione del questionario era, inoltre, possibile indicare i motivi per cui si era scelto di non utilizzare la griglia. E' risultato che quasi il 100% di coloro che non hanno scelto la griglia, non condividevano i pesi proposti. Un'altra larga fetta, circa il 90%, non ha scelto di utilizzare la griglia per la tabella che riportava il punteggio totale conseguito su una scala da 0 a 15. Un'ultima fascia, circa il 45%, non ha scelto la griglia perché non ne condivideva gli indicatori.

Per quanto riguarda la valutazione delle prove, il grafico in Fig. 11 mostra la percentuale dei voti per i vari indirizzi, suddivisi in ottimo, sufficiente e insufficiente. Si sottolinea che un giudizio insufficiente non equivale alla bocciatura del candidato, ma a un punteggio minore di 10. Gli ottimi, invece, rappresentano le prove valutate con un punteggio massimo pari a 15.

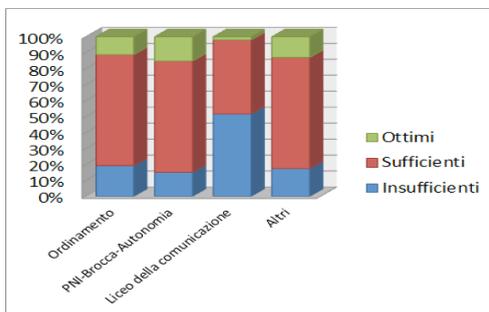


Fig. 11 – Distribuzione dei giudizi nei vari indirizzi

Si nota immediatamente che gli studenti dei Licei della Comunicazione hanno incontrato molte difficoltà nella risoluzione della prova. La percentuale delle insufficienze, infatti, supera il 50%. Per ciò che riguarda l'Ordinamento e il PNI si evince, invece, che nel primo ci sono più insufficienti e nel secondo più ottimi, anche se la differenza non è così netta. La Fig. 12 mostra la distribuzione dei giudizi nelle varie regioni italiane. Si può notare che la percentuale di sufficienze supera il 50% in tutte le regioni, con punta massima dell'80% in Sicilia; la percentuale di insufficienze è tra il 10% e il 30%; infine, la percentuale di ottimi si assesta intorno al 10% in tutte le regioni.

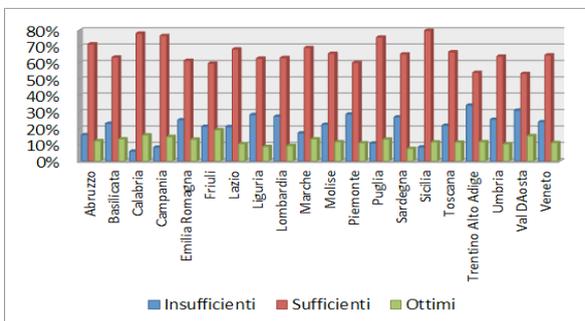


Fig. 12 – Distribuzione dei giudizi nelle varie regioni

E' parso interessante confrontare le valutazioni effettuate con l'utilizzo della griglia, con quelle in cui la stessa non è stata utilizzata, calcolando le media e deviazione standard di insufficienze, sufficienti e ottimi. Non sono state notate differenze sostanziali nelle valutazioni e, nella pratica, le valutazioni restano abbastanza uniformi. Ciò che invece è interessante analizzare è un confronto delle valutazioni con gli anni precedenti. Sono state prese in considerazione gli anni 2010, 2011 e 2012, calcolando medie e deviazioni standard per le singole valutazioni.

<i>Media/DevStnd</i>	Insufficienti	Sufficienti	Ottimi
2010	19.70/7.65	65.95/5.34	14.35/3.09
2011	35.24/6.83	56.68/6.24	6.08/1.03
2012	21.93/7.91	66.19/7.29	12.42/2.65

Tab. 1– Media e deviazione standard percentuale delle valutazioni

Dalla tabella 1, e ricordando che nel 2010 e nel 2012 il membro di matematica era un docente interno, si nota come, in questi due anni, la percentuale degli ottimi è più del doppio del 2011 e che la deviazione standard in presenza del membro esterno è praticamente 1.

Giudizio sulla prova e analisi delle difficoltà

L'ultima parte del questionario è volta a verificare come i presidenti di commissione hanno giudicato le tracce d'esame. La Fig. 13 mostra come i docenti hanno ritenuto la corrispondenza tra traccia e programmi per i due problemi e per i dieci quesiti, per l'Ordinamento e il PNI. Era possibile fornire tre risposte: *completa*, *solo in parte*, *per niente*.

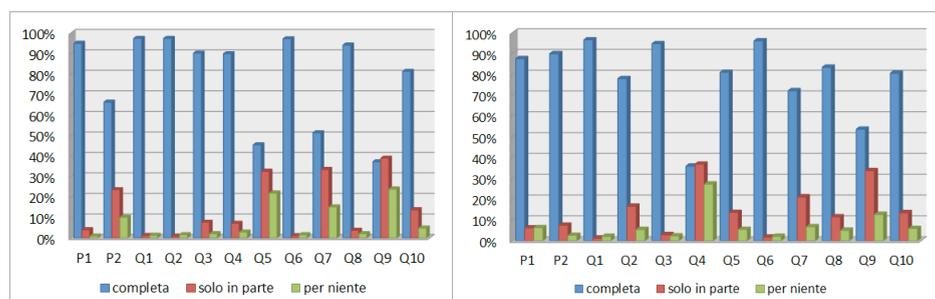


Fig. 13 – Corrispondenza tra traccia e programmi (a sinistra Ordinamento, a destra PNI)

E' risultato che, tranne per il quesito numero 9, nel caso dell'Ordinamento, e del quesito numero 4, nel caso del PNI, tutti i commissari hanno ritenuto che la prova di matematica corrispondesse effettivamente ai programmi svolti. Si richiedeva, infine, di indicare se la prova fornita per l'anno 2012 era in continuità rispetto alle prove degli ultimi cinque anni o meno. Anche in questo caso era possibile scegliere tra tre diverse opzioni: Sì, Solo in parte, No.

Paradossalmente, coloro che hanno ritenuto maggiormente che la traccia fosse in continuità con gli ultimi cinque anni, sono stati i commissari del Liceo della Comunicazione, che, analizzando le domande successive, hanno però anche ritenuto la prova più complessa rispetto a tutti gli altri indirizzi. Praticamente i commissari di tutti gli indirizzi hanno ritenuto la traccia abbastanza chiara. Per ciò che riguarda le percezioni dei commissari circa le difficoltà incontrate dagli studenti nello svolgimento della prova, i pareri dei commissari sono alquanto distribuiti. In particolare, il giudizio dei commissari del Liceo della Comunicazione circa la complessità della prova è confermato dal fatto che, come visto in precedenza, è proprio nel Liceo della Comunicazione che si ha avuto la maggiore percentuale di insufficienze.

Conclusioni

La facilità con cui le tecnologie informatiche rendono possibile oggi comunicare e condividere informazioni ha favorito il moltiplicarsi di iniziative finalizzate alla valutazione dei risultati scientifici e didattici del sistema educativo nazionale. Esse rappresentano la base di conoscenza necessaria alla costruzione di un sistema di certificazione delle competenze acquisite conseguendo i diversi titoli di studio rilasciati nel nostro paese e in ambito europeo. Nel presente lavoro è stata illustrata un'iniziativa che da anni si occupa di censire e valutare i risultati di apprendimento della matematica nei licei scientifici italiani attraverso una indagine sulla seconda prova scritta agli Esami di Stato. La dimensione dell'indagine e la ricchezza delle informazioni è cresciuta negli anni e ha innescato un processo di riflessione promosso dal Ministero dell'Istruzione e della Ricerca attraverso progetti e iniziative, supportate dal sito www.matmedia.it.

Ringraziamenti

Si ringraziano tutti i Commissari che hanno inviato i numerosi questionari, i Referenti Regionali e i docenti che hanno partecipato alle attività di predisposizione e di conduzione dell'indagine, e l'Ispettore Emilio Ambrisi della ripartizione per il suo coordinamento, nonché la società italiana di scienze matematiche e fisiche Mathesis (www.mathesisnazionale.it).

Bibliografia

[Lasorsa e Amici, 2010] C. Lasorsa, M. Amici, La struttura organizzativa del Servizio Nazionale di Valutazione. *L'Educatore*, 57, 12, 2010, 41-44

[Venticinque et al., 2010] S. Venticinque et al., Indagine Matmedia 2009, La Prova di Matematica nei Licei Scientifici Italiani. *Periodico Di Matematica*, Numero 1, Volume 2 , Serie XI, Anno CXX, Gen-Apr 2010

Can Mathematics Education Take Advantage of Massively Multiplayers Online Competitions?

Mario A. Bochicchio, Antonella Longo, Antonio Bernardo
DIDALab – Università del Salento
via per Monteroni, Complesso Ecotekne, 73100 Lecce (Le)
{mario.bochicchio, antonella.longo, antonio.bernardo}@unisalento.it

Abstract. Collaborative learning can provide remarkable support to the efficacy of teaching Mathematics. Collaboration, indeed, is essential to reinforce the interaction among learners and to foster active learning and comprehension construction in the student. In the social setting of a collaborative learning experience the learner interacts directly with the instructor and with the other actors, and these interactions can guide interpretation and construction of concepts. To test these assumptions on the field, in 2008 we created a platform and a format for Massive Multiplayer Online Competitions (MMOCs), able to support large groups of students (25 to 250 teams with 500 to 10.000 players at a time). In the paper we describe the proposed format and the main results of two national editions of *MatematiCup*. *MatematiCup* is a large students' completion with the primary objective to encourage young people's interest in mathematics, to experience how to participate in the network.

Keywords: Collaborative Environments & Tools, Online Experimentation, Serious Games, Innovation in mathematics education, Labs, Contents and Tools.

1. Introduction

In the last decade, the Creative Commons philosophy of freely sharing information and the pervasiveness of the Internet have created many new opportunities for teaching and learning. Meanwhile the use of learning management systems (LMSs), such as the proprietary Blackboard or the open-source Moodle software, has become ubiquitous. One of the results of combination of these phenomena are David Cormier's MOOCs (Massively Online Open Courses). They integrate the connectivity of social networking, the facilitation of an acknowledged expert in a field of study, and a collection of freely accessible online resources. Perhaps most importantly, however, a MOOC builds on the active engagement of several hundred to several thousand

“students” who self-organize their participation according to learning goals, prior knowledge and skills, and common interests. “Although it may share in some of the conventions of an ordinary course, such as a predefined timeline and weekly topics for consideration, a MOOC generally carries no fees, no prerequisites other than Internet access and interest, no predefined expectations for participation, and no formal accreditation” [17].

A MOOC is typically offered through an online social network. A central web address may be used to consolidate a registration process, outline the suggested course schedule, and provide a nexus for support and communication. Apart from this, however, individuals may continue to use the central site to consolidate their participation or they may spin it off into their own blogs and develop and maintain ties through other technologies such as Twitter. They can negotiate and define collaborative topics, working networks, and goals with others who share common interests and concerns. The results of a MOOC collaboration may extend far beyond the MOOC itself: the network negotiated is just as important as the topic covered, if not more so. Participation in a MOOC is emergent, fragmented, diffuse, and diverse.

The MOOCs phenomenon of online social network is parallel of another common likewise pervasive one: Massive Multiplayer Online Games (MMOGs): online computer games that draw together players from all over the world have come to dominate the digital entertainment industry and have increasingly become the dominant form of entertainment, particularly for children and adolescents ([18]).

Scientific investigation on educational use of MMOGs is not new. They concern both pedagogical and technical aspects. Researchers have concentrated on how to use MMOGs to teach either languages ([19]) or Science, Technologies, Engineering and Mathematics (STEM) [20] in a collaborative environment.

In this scenario we implemented MatematiCup, an online mathematical competition for large groups of Italian students, with the intent to investigate how recreational and competitive approaches can engage and interest in STEM subjects at an early age – in a way that deepens their knowledge while also developing 21st-century skills.

From the same IT platform and competition format EcologicaCup, for supporting ecology and biodiversity, has been derived. The scientific intent of these competitions are to encourage young people's interest in mathematics and sciences, to experience how to participate in the network, to propose new solutions to the logistical problems associated to large students' competitions, to spread the use of new technologies and to analyze a new form of collaborative and participatory Internet. In this paper we present the format and the results of MatematiCup competition.

1.1 Mathematical competitions in Italy

In Italy there are several mathematical competitions. Among them the most important ones are:

- Mathematical Olympiad, is an individual competition, for secondary school students. It is articulated on different increasing levels. At the first stage the

Games of Archimedes are held in individual schools, the best students go further, to the provincial stage (second step). The national phase (third step) takes place in Cesenatico. The 20 best students attend a week long pre-Olympic training session in Pisa (fourth step) from which six participants are selected to represent Italy to the International Olympiad (last step)[4].

- Mathematical Games, organized by the Pristem Eleusis center of Bocconi University, in Milan, involve students from primary school to university in four phases named "Autumn Games", the national phase of the "International Championship of Mathematical Games", the "Race Team" and the "Spring Games";
- *Mathematica Senza Frontiere*, the Italian edition of *Mathématiques Sans Frontières*. It involves 23 countries and it is characterized by a collaborative approach, as a class is divided into groups that contribute to the solution of any assigned problems. It is also characterized by its attention to the languages. The first question, in fact, is always formulated in a foreign language selected by the student, and the solution must be described in that language. The better classes are selected to proceed to the national phase [5].
- Kangaroo Games, started in Australia in '78, they are mainly addressed to students of primary school. Participants include thirty countries. They are based on mass participation and do not encourage competition between countries. The competition game consists of a single session, without further steps, and consists of a questionnaire with 30 multiple-choice questions of increasing difficulty [6].
- Mathematical Games of Sicily is an annual initiative of the Association of Teachers and Students of Mathematics (AICM). It includes primary and secondary schools from Sicily [7].
- Rally Matematico Transalpino is a math contest for pupils from third grade (primary schools) to second year of high schools. It started in '92, in Switzerland; it was then extended to other countries.

Recent investigations from OECD-PISA (Organization for Economic Cooperation and Development - Programme for International Student Assessment) [8][9] have highlighted the difficulties of Italian students in solving math problems in practical situations. PISA experts define the mathematical literacy as: "an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgments and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen." [10]. Therefore, according to PISA Programme, mathematical competence deals with a functional use of mathematics. In this context, several authors have investigated the role of mathematical games as a tool to enhance learning and encourage motivation. Fengfeng Ke [11], for example, reported the findings of a research carried out by J. Randel et al. [12], which had analyzed 68 studies about the performance of students who attended and didn't attend mathematical games. He states that in 56% of cases no difference is detected, in 32% of cases a clear improvement has been detected in the performance of students who have participated in one or more competitions and in the 5% of cases the

performance has been better in students who followed conventional approaches with no competitions. Ke also reports the findings of other researchs on relational dynamics and on student motivation in the classroom. According to Ke, several scientists agree on splitting students' approaches in three categories: cooperative, competitive and individualistic. In the first category, the student perceives that he is working together with others to achieve a common prize. Conversely competitive student perceives that he is working with others to get an award that comes from the positive comparison between his and others' work. Finally, the individualistic student perceives that he is working alone on the basis of forms of self-gratification. The results reported by Ke, referring to 122 studies, show that the collaborative approach is more effective in terms of learning and productivity than the other [13]. In the context outlined, the authors believe that the competition-based problems can be usefully employed in the secondary level to stimulate young people's interest in mathematics. Moreover, we believe that existing competitions leave wide space to novel online contests to stimulate the collaborative approach.

On this basis, in 2007 the authors have organized the first edition of *MatematiCup*, which was attended by about 5000 students from all Italian regions. The specificity of *MatematiCup* in relation to other competitions is discussed in section II. Section III describes the main organizational aspects of *MatematiCup*. The results of five years of operations are discussed in Section IV. Section V sets out the conclusions and the main future developments.

2. MATEMATICUP: REQUIREMENTS AND OBJECTIVES

The requirements of *MatematiCup* have been derived from a detailed analysis of the above mentioned math competitions, from the unique experience of conducting the *Matematicamente.it* community, with more than 50.000 subscribed users, for over 12 years, from the research on collaborative Web applications carried out at the DIDA-Lab of the University of Salento about the activities developed in collaboration with several Italian schools.

2.1 Objectives

O1. The primary objective of *MatematiCup* is to **encourage young people's interest in mathematics** and to **reinforce math skills** that meet the demands of young people as persons exercising a constructive role in society.

O2. The competition aims to **test virtual participation**, based on the Web and to spread out the use of new technologies, with specific attention to collaborative web applications and participatory Internet

2.2 Requirements

R1. Web-based Collaborative learning: in our opinion "teaching activities" must go beyond the interaction between students and teachers in classrooms. New technologies, for example, allow multiple interactions between multiple groups of people, and this can help to expand the "teaching experience". Furthermore, the use of the Internet, if correctly addressed by the teachers in class, and analyzed with the help of research organizations like universities, can

finalize the eLearning experience making it systematic and not accidental, even for primary and secondary schools.

R2. Open Enrollment: registration for the competition shall be free to allow schools and teachers to involve not only the best students, but **all** those who have pleasure and desire to compete on this discipline.

R3. Collaboration between classes: in tournament student teams must be composed by students from different classes. The student must discuss and “defend” his/her opinion before answering questions. In this way there is a form of collaborative learning among students from different classes rarely practiced in the usual teaching, and it is known that this type of collaboration can be very effective.

R4. Situated learning: according to the framework defined in OECD PISA surveys, and on the research on situated learning, questions and problems proposed during the competition must be focused on the critical knowledge and skills for solving mathematical problems drawn from real-life contexts.

R5. Measurability of the learning outcomes: quizzes must be built to be automatically evaluated, so that performance of participating teams can be ranked in real time for all ranking categories. This allows the active involvement of online spectators (e.g. families). However, multiple-choice answers should be avoided, because they lead students to ‘guess’ the answers rather than to compute it. Simple open answers like numbers or letters can be useful to discourage those who just ‘guess’ the result rather than arguing on it.

R6. Attention to the school-level: Italian primary and secondary schools are often not supported by qualified technical personnel, and not equipped with advanced computers and high speed connections. Online competition can stimulate schools to compare each other’s and to upgrade laboratories and Internet connections, bridging the technological gap with the high schools.

R7. Attention to gender issues. The competition should help push girls to deal with mathematics, exceeding the idea of the social fund for which they are not inclined to scientific studies.

R8. Support orientation. The competition allows students a comparison with peers from different schools and regions and supporting guidance in the choice of upper secondary level according to th personal interest and / or capacity.

R9. Students training. The preparation for tournament should be during the school year without placing many organizational constraints to teachers and promoting the achievement of curriculum objectives.

3. MatematiCup: Format and Technical Aspects

MatematiCup is made up of three phases:

- registration and organization of the teams, from February to May;
- training, from March to May;
- Online competition: first half of May.

Each secondary school can participate with one or more teams, of 12- 30 students per team, plus a teacher as a coach. Each team is split into three groups (one per class, requirement R3), each of the three groups must decides

specific tasks for each student: a captain who organizes the work, a computer operator, who is responsible for entering the answers and two or more “brains” who contribute, together with the captain and the computer operator, to the resolution of the questions (requirement R1). The coach helps pupils in the early stages of organization, online subscription and training. Schools are also encouraged to form teams of “all-girls”: there are, in fact, awards for the best all-female teams (R7). The teams can work out at any time, both at school and at home, from February to May (R9). Practice sessions are based on problems and questions similar to those of the competition and scores are calculated in the same way. The final competition, lasting about 3 hours, takes place in May, in the morning, in curricular time, at the same time for all teams. Each team works with three computers and on each of them only appear quizzes that must be solved by the sub-team working on that computer. The rankings are updated in real time for all the teams, they are available to participants and to families and friends connected to the Web (R5). The coach is responsible for the educational value of the initiative towards his/her students and toward the organizers of the competition, therefore he/she will discourage any unsportsmanlike behavior of students. During the competition the coach is enabled to remove any technical problem or computer malfunctioning. He can also connect to another computer to see in real time the overall performance of his teams and to encourage them to do their best. During the competition each student works with the team to try and solve as correctly and as quickly as possible the proposed quizzes. Participants can use calculators, software, books and anything deemed necessary. On the other side, they can't disclose quizzes to others and cannot ask for coach's help. Each answer is weighted and assessed according to both the difficulty level of the question and the time taken to solve it. For each quiz the difficulty is graphically expressed by a 5 star indicator. The questions marked with a green circle in the column “Giocato” (i.e. Played) are not still solved, while those marked with a red circle have been already answered. Questions marked in yellow have been read but not yet solved, and they must be answered as soon as possible, as its score decreases over time. Since a quiz is displayed, students have had two minutes to read it and to discuss it with the teammates. During this time the computer will not accept replies. This requires students to carefully read the text and to discuss it with peers before answering. After this time a stopwatch starts and runs until the computer operator enters and confirms the answer. Figure 2 shows an example of the question screenshot. The score assigned to each answer is based on the elapsed time evaluated on the client, and not on the server, so that it is independent on the connection speed (a number of schools are still connected via analog modems, ISDN or other slow or high-latency connections not suitable for server-side operations). The overall team's score is the sum of the quizzes' scores solved by each sub-teams. The ranking is evaluated and published in near-real time and simultaneously for all the teams. Moreover, it can be displayed at any time during the game. To avoid distraction, the ranking is normally displayed only to the coach on another computer. According to R3 and R4, the problems and questions for the competition and for the practice sessions are built according to the guidelines of the the OECD-PISA

Programme, especially for those skills that allow to understand real life problems and critical knowledge.

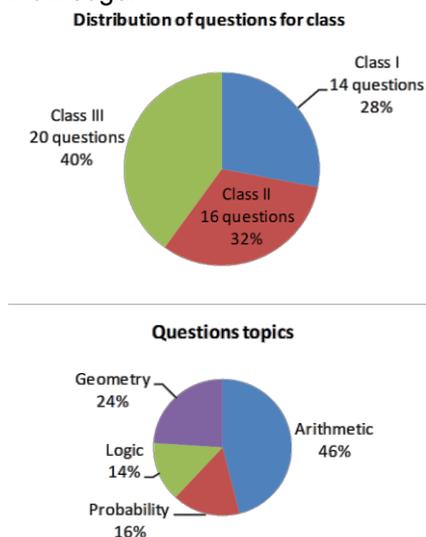


Figure 1: Distribution of questions per topic and class

We paid attention to equally distribute questions among classes and topics, as shown in Figure 1.

Each team must have:

- a classroom or a laboratory reserved for the competition;
- at least three computers connected to the Internet (one for each sub-team).

The questions are usually structured as a short explanatory text followed by a picture, as shown in Figure 1. Given the large number and variety of users we decided not to use video, animations or other bandwidth-demanding media to avoid delays and network congestions. Ranking pages are updated in near-real time by "forcing" the reload of the corresponding web page at each minute.

4. Results and discussion

In its first year, MatematiCup summed up about 5,000 students, divided into 250 teams, belonging to about 150 Italian secondary schools. The distribution of institutions by region is shown in Figure 2. The graph shows a good level of penetration and a fairly even distribution throughout the national territory. Equality of participants by gender, shown in Figure 3, demonstrates the fulfillment of the requirement R7.

The data on enrollment for the second year shows that participants increased of about 30%. Confirmation of participation by the majority of participating schools of the first year is indicative of substantial approval from teachers and pupils.

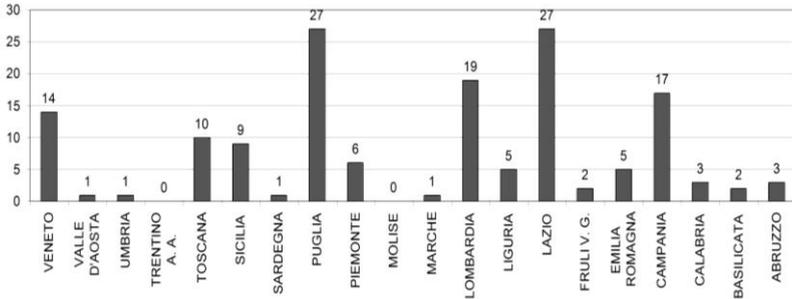


Figure 2. Regional distribution of participants at the first edition of MatematiCup

A review of the effectiveness of the initiative was conducted through informal interviews, by phone and in person, to two groups:

- The first, composed by headmasters and teachers of the 15 winning teams of the first MatematiCup, stressed the importance of the initiative as complementary to the other extracurricular activities (eg, afternoon workshops in mathematics) organized by the schools on a local scale. Some executives have clearly appreciated the opportunity provided by the competition, to acquire an additional element of comparison between the performance of teams belonging to their institution and teams of others institutions. Many of the participants have asked for a formal certification of achievement, essential to complete their personal curricula and/or portfolios of activities;
- The second group (about 5% of the participants), consisting of headmasters, teachers and students, pointed out, among the positive aspects, the quality of the questions (for concreteness and for their connection to the experiential world of students). The existence of a ranking and a prize for the training phase has prompted many students to a constant commitment in the period from February to May, much appreciated by teachers and by families. Some teachers have asked for the availability of the questions as educational material for holidays.

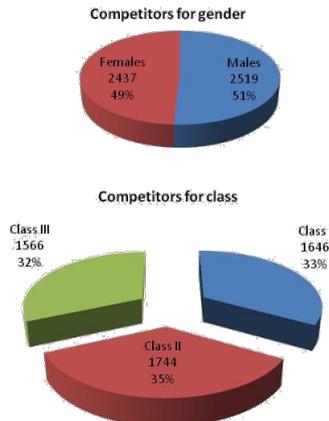


Figure 3 Enrolled students per genre and class

During the final competition the beginning stage was critical for several teams: in fact the difficulties of connection of some schools contributed to delay the start of the competition and it flooded us of phone calls, with requests for explanations, which further delayed the start of the competition. To overcome this issue, from the second edition we created a messaging area, visible in every pages, to communicate with all participants. A fifth of respondents expressed doubts and misgivings about the impartiality of teachers / coaches of the winning teams. Others would have liked a longer duration or more tests.

5. Conclusions and future works

The mathematical competitions are a useful motivational incentive to improve the learning of mathematics in secondary schools, but they require careful organization and appropriate use of tools (especially the Internet). In this context, the authors have presented the motivations, organizational aspects and the main results of MatematiCup.

The competition, conducted by the University of Salento in close collaboration with the online community Matematicamente.it, highlighted interesting areas for improvement and research in comparison to other national competitions. In particular, teachers and students have appreciated the style adopted for the questions, the opportunity to compete online, in real time, with other Italian schools, and to have rankings updated in real time and accessible to friends and families.

The positive results achieved by MatematiCup in its two editions convinced us to apply the same organizational model and the same software platform to EcoLogicaCup (<http://www.ecologicacup.unisalento.it/>), organized by the Observatory of Mediterranean Ecosystems Ecology and Health in collaboration with the Italian Society of Ecology, the WWF and other organizations). In this case, some extensions and modifications have been necessary to take into account the nature of extra-curricular initiative with regard to ministerial programs of Italian secondary schools.

Further developments are about the definition of more sophisticated and media-rich environment, to approach the behavior of Massively Multiplayer Online Games already used by students in secondary schools.

6. Acknowledgment

The research work is partially founded by Engineering Ingegneria Informatica and Essentia Spa.

Bibliografia

[1] Michel Criton, "Le origini: I giochi matematici fino al Medioevo", in Lettera Matematica Pristem, n.54, 2005, p.54.

[2] Alcuino di York, Giochi matematici alla corte di Carlomagno, Problemi per rendere acuta la mente dei giovani, a cura di R. Franci, Edizioni ETS, 2005.

[3] Nando Geronimi, Giochi matematici del medioevo, I conigli di Fibonacci e altri rompicapo liberamente tratti dal Liber Abaci, Bruno Mondatori, 2006.

[4] Progetto Olimpiadi della Matematica, <http://olimpiadi.ing.unipi.it/>

[5] Matematica Senza Frontiere, <http://www.matematicasenzafrontiere.it/>

[6] Kangourou Italia, <http://www.kangourou.it/>

[7] Associazione A.I.C.M., <http://dipmat.math.unipa.it/~grim/aicm/>

[8] La sezione italiana di OCSE-PISA 2006 presso il sito dell'Invalsi, l'Istituto incarica di svolgere l'indagine in Italia <http://www.invalsi.it/ric-int/Pisa2006/sito/>

[9] OCSE- PISA international site. http://www.oecd.org/pages/0,3417,en_32252351_32235731_1_1_1_1_1,00.html

[10] AA. VV., Valutare le competenze in scienze, lettura e matematica, Quadro di riferimento di PISA 2006, Armando Editore, 2007, p. 17

[11] Fengfeng Ke, Classroom goal structures for educational math game application, "Proceedings of the 7th international conference on Learning sciences", June 2006, International Society of the Learning Sciences.

[12] Randel, J.et Al. (1992) The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming*, 23(3), 261-276.

[13] Davidson, N.E. Cooperative learning in mathematics: a handbook for teachers, Addison-Wesley, 1990.

[14] S. Locatello, G. Meloni, Apprendimento collaborativo in matematica, Pitagora, 2003.

[15] L. Camizzi, E. Mosa, F. Rossi, Un'esperienza di "collaborative learning" in matematica, Contenuti innovativi + apprendimento collaborativo + ambienti on line = la formula ideale per l'apprendimento e l'insegnamento della matematica!, INDIRE, <http://www.bdp.it/content/index.php?action=read&id=1439>.

[16] W. J. Clancey, "A tutorial on situated learning", in Proceedings of the International Conference on Computers and Education (Taiwan) Self, J. (Ed.) Charlottesville, VA: AACE. 49-70, 1995; <http://cogprints.org/323/0/139.htm>

[17] McAuley, A., Stewart, B., Siemens, G., and Cormier D. The MOOC Model for Digital Practice. 2010;http://www.elearnspace.org/Articles/MOOC_Final.pdf.

[18] F. Paraskeva, S. Mysirlaki, A. Papagianni (2010), Multiplayer online games as educational tools: Facing new challenges in learning. *Comput. Educ.*54, 2 (February 2010), 498-505. DOI=10.1016/j.compedu.2009.09.001 <http://dx.doi.org/10.1016/j.compedu.2009.09.001>

[19] Rankin, Y.; Morrison, D.; Shute, M., "Utilizing Massively Multiplayer Online Games to foster collaboration and learning," Science and Innovation Policy, 2009 Atlanta Conference on , vol., no., pp.1,10, 2-3 Oct. 2009, doi: 10.1109/ACSIP.2009.5367811

[20] R. Araya, A. Jiménez, M. Bahamondez, P. Dartnell, J. Soto-Andrade, P. González, P. Calfucura, "Strategies Used by Students on a Massively Multiplayer Online Mathematics Game", in Advances in Web-Based Learning - ICWL 2011, Lecture Notes in Computer Science Volume 7048, 2011, pp 1-10, DOI: 10.1007/978-3-642-25813-8_1

Competenza di lettura e didattica on line:

Eugenio, tutor digitale

Piera Schiavone
I.I.S.S. "R. Canudo" (liceo scientifico)
Via Aldo Moro, sn 70023 Gioia del Colle (Ba)
piera.schiavone2@gmail.com

Abstract. *Homework, pens, white/blackboards, books : we do hope they will never die. They represent the traditions Eugenio, the first Italian digital tutor , relies on . But Eugenio goes further ,it is curious and amazed: it looks at the new school, understands and anticipates its needs. Eugenio responds to specific teaching /educational needs, getting closer to students' needs and involving them with the tools and methods that are part of their world and are the way of learning things. This work describes the potentialities of the digital tutor Eugenio used for the reading, comprehension and analysis of literary texts. Eugenio is a digital tutor who, with rigor and flexibility, can help students with their homework at home, adapting himself himself to the demands of the teacher and the educational level of the class. Eugenio is an innovative attempt to integrate teaching and new technologies, combining the classical tradition with the most advanced ICT resources and network. Eugenio is also an invitation to read.*

Keywords: digital tutor, cloudschooling, Maieutical labs

1. Introduzione

Dall'anno scolastico 2010/2011 anche per le classi seconde del biennio superiore a maggio sono previste le prove Invalsi. Quali sono i problemi di questa novità? Sono molti e non sempre di facile soluzione, dal momento che l'impostazione didattica, spesso ancora troppo tradizionale, non prepara in modo adeguato i ragazzi e, nella maggior parte dei casi, alunni "modello"...di fronte alle prove della tipologia invalsi...crollano! Come risolvere il problema?Potenziando il procedimento di lettura e comprensione e avvalendosi anche della metodologia invalsi e del sistema di valutazione invalsi, che prevede un preciso quadro di competenze. Certo non è sufficiente proporre solo domande a risposte chiusa, perché in tal caso non si potrebbe insegnare al

Congresso Nazionale AICA 2013

ragazzo a potenziare l'aspetto argomentativo, tuttavia la capacità di sintesi e la precisione nella risposta sono elementi a cui il ragazzo deve abituarsi

Eugenio riesce a conciliare queste esigenze didattiche:potenzia il procedimento di lettura, affianca quello di comprensione e di analisi, definisce quello di rielaborazione, guidando a piccoli passi il giovane lettore. Il fatto che proponga esercitazioni sia per la terza media che per il biennio delle scuole superiori inoltre crea un elemento di continuità fra i due ordini di scuola e di stretto collegamento fra le diverse fasi dell'apprendimento, preoccupandosi del potenziamento delle conoscenze e dell'acquisizione delle competenze, puntando sul miglioramento delle abilità di lettura e comprensione testuale. Per avvalersi di questo tutor digitale è necessario il seguente setting tecnologico: l'aula deve disporre di LIM con collegamento Wi-fi; gli alunni devono utilizzare a casa il computer.

2. Situazione di partenza

Guidare l'alunno nella lettura, comprensione e analisi testuale: questa l'esigenza didattica di partenza. Come gestire la situazione? Accanto alla didattica tradizionale, l'uso del tutor digitale *Eugenio* può essere un utile completamento per il recupero e il potenziamento delle abilità e per avvicinarsi all'acquisizione di precise competenze. *Eugenio* è un tutor digitale di nuova concezione che offre agli studenti una guida mirata alla comprensione del testo; un sistema di interazione adattiva, poiché quando pone una domanda, *Eugenio* si ricorda il livello di competenza dello studente e formula la domanda successiva ad un livello appena più difficile di quello esplicitato dallo studente nella risposta corretta; infine *Eugenio* offre una classe virtuale semplice da usare e consultare.

In quali classi utilizzare *Eugenio*? Io insegno in un liceo scientifico ed ho provato ad usare *Eugenio* con ragazzi di prima classe e di terza classe.

2.1 Fasi dell'esperienza

Durante l'anno scolastico appena concluso ho utilizzato *Eugenio* con alcuni alunni in una prima classe e in una terza classe di liceo scientifico.

Nel caso della prima classe, proporre questo tutor digitale è stato un modo per *potenziare* le tecniche di lettura e comprensione ed abituare gradatamente gli alunni alla metodologia invalsi che sarà proposta nel prossimo anno scolastico (dato che è prevista la prova nazionale invalsi per la seconda classe del primo biennio).

Nel caso della terza classe, proporre questo tutor digitale è stato un modo per *recuperare* carenze e aiutare gli alunni, in itinere, ad affrontare le analisi testuali di ambito letterario e, nello specifico, poetico.

La piattaforma è stata usata in aula attraverso l'impiego della LIM.

1ª fase: innanzi tutto è stato spiegato ai ragazzi in classe il concetto di piattaforma, la sua funzionalità e contestualmente si è provveduto ad assegnare a ciascun ragazzo un accesso con credenziali. Si è creata la classe virtuale. Io ho ricoperto il ruolo di tutor in questa classe virtuale.

2ª fase : sono stati da me scelti e assegnati on line esercizi specifici, inviati in alcuni casi a tutti gli alunni, in altri casi solo ad alcuni, per potenziare o per recuperare alcuni argomenti.

3ª fase : gli allievi hanno svolto esercizi on line da me assegnati per potenziare o recuperare specifici contenuti di carattere grammaticale-sintattico (a seguito di errori negli esercizi, la piattaforma ha fornito al singolo alunno sintesi per il ripasso dell'argomento).

4ª fase: è importante consolidare/potenziare la competenze di lettura secondo i descrittori PISA (Programme for International Student Assessment-indagine internazionale promossa dall'OCSE). La competenza di lettura è considerata come un insieme di conoscenze, capacità e strategie in continua evoluzione, che gli individui sviluppano nel corso della vita, attraverso le situazioni e i rapporti in cui sono coinvolti e che serve loro per continuare ad apprendere. PISA definisce la competenza di lettura come la capacità di comprendere e utilizzare testi scritti e riflettere sui loro contenuti al fine di raggiungere i propri obiettivi, sviluppare le proprie conoscenze e potenzialità e svolgere un ruolo attivo nella società. Partendo dall'acquisizione della tecnica di lettura e comprensione, in prospettiva si intende ampliare tale attività attraverso il coinvolgimento dei docenti appartenenti ai diversi assi culturali, proponendo l'analisi non solo di testi letterari, ma anche di ambito scientifico, tecnico, in lingua straniera, testi "autentici", legati cioè alla concreta quotidianità, anche attraverso la metodologia CLIL. Alla conclusione di tali attività l'alunno dovrebbe acquisire competenze sia disciplinari che trasversali. Peraltro l'attività potrebbe, in linea con le indicazioni previste dalla Nota Miur del 16/1/2013 recante le norme transitorie per *l'insegnamento di discipline non linguistiche (DNL) in lingua straniera secondo la metodologia CLIL nei Licei Linguistici* essere una valida modalità operativa per avviare, nell'ambito del consiglio di classe, la metodologie didattiche CLIL. Come recita la nota ministeriale infatti *si auspica la costituzione di veri e propri team CLIL (docente di DNL, docente di lingua straniera, conversatore di lingua straniera, eventuale assistente linguistico), finalizzati allo scambio e al rafforzamento delle reciproche competenze.*

3. Riflessioni sull'uso di *Eugenio*

Per sperimentare l'uso di *Eugenio* è necessario avere conoscenze di piattaforme (per es. della tipologia moodle) dove creare classi virtuali per avvalersi della possibilità di differenziare e individualizzare le tipologie di esercizio, oltre che i contenuti, per ogni singolo alunno. È inoltre necessario essere in grado di gestire come tutor la piattaforma,risolvendo anche inevitabili inconvenienti tecnici e rispondendo ai ragazzi nel raro caso in cui si dovessero trovare in difficoltà (per esempio nella procedura di iscrizione).Gli alunni imparano facilmente ad entrare in piattaforma e ad utilizzarla. Dal punto di vista pedagogico-didattico l'uso di tale piattaforma consente di creare, pur nella eterogeneità della classe, quasi una didattica individualizzata, molto utile per gestire situazioni di recupero, potenziamento, ma anche di eccellenza.

L'uso della piattaforma Eugenio appare utile sia nella modalità collettiva che in quella individuale. I *testi bussola* sono un ottimo strumento per sostenere e guidare la lezione in classe (con la LIM); per fornire, a casa, un'introduzione ad argomenti che verranno trattati la volta successiva a scuola, consentendo al docente, ancor prima di entrare in classe, una quadro esaustivo delle principali difficoltà incontrare dai suoi studenti sul testo-bussola; infine sono un ottimo strumento di recupero nel passaggio al successivo ciclo scolastico, quando è particolarmente importante tener conto, nella programmazione didattica, di eventuali lacune creatisi nei cicli precedenti. Nel dettaglio i *testi bussola* sono vere e proprie lezioni introduttive o di ripasso/potenziamento, che partono da un esempio concreto per ricostruirne la struttura, l'organizzazione e la finalità testuale. Tutti i testi bussola sono adattivi: *Eugenio* si ricorda il livello dello studente sulla competenza associata e formula la domanda ad un livello appena più difficile di quello che ha ricavato da una risposta corretta. Di seguito si indicano alcune caratteristiche specifiche delle singole tipologie. Nella tipologia di *Analisi e comprensione di un testo letterario*, il tutor guida alla scoperta dei metodi di analisi e dei principali generi testuali. Il docente può scegliere fra i testi bussola sulla base dell'insieme delle *abilità*, delle *conoscenze* che vengono introdotte dalla lezione, del *genere*, del *titolo* dell'*autore* del testo-bussola. I testi bussola non presuppongono la padronanza delle conoscenze, ma le introducono progressivamente, approfondendo, in caso di errore, il quadro teorico in cui tali conoscenze vengono tradizionalmente sistemate a livello disciplinare. Al tempo stesso, testano il grado di padronanza delle *competenze* effettivamente attivate dallo studente nel corso dell'analisi, fungendo da placement-test per impostare una corretta programmazione didattica anche e soprattutto a livello individuale. Nella tipologia *Analisi e comprensione di un testo espositivo, descrittivo, discontinuo*, il tutor guida alla scoperta degli elementi della descrizione, la loro collocazione nello spazio e il punto di vista dell'osservatore; stimola lo studente a individuare informazioni implicite ed esplicite date in un testo espositivo. Nel caso dei testi discontinui, il tutor suggerisce, all'occorrenza, esempi grafici che aiutano lo studente ad interpretare correttamente le informazioni visive e le relazioni semantiche tra le parti discontinue del testo. Nella tipologia *Analisi e comprensione di un testo argomentativo*, il tutor guida all'individuazione della *tesi centrale* del testo e degli *argomenti* a sostegno, attraverso una serie di domande sul testo. Ogni qual volta lo studente giunge a un determinato *snodo* argomentativo, il tutor avverte lo studente o lo guida a considerare la posizione dell'argomento nel complesso dell'argomentazione. Nella tipologia *Analisi e comprensione di un testo multimediale*, il tutor mostra (in esecuzione) un testo multimediale (es. video, notiziari, spot pubblicitari, presentazioni) e guida alla scoperta della sua organizzazione e della sua struttura. Nel dialogo fra interlocutori diversi, guida ad individuare contesto, scopo, destinatario della comunicazione e registro utilizzato. Ognuna delle suddette tipologie, pone particolare attenzione alla grammatica esplicita e alla riflessione sugli usi della lingua, ma ciò avviene sempre *in atto*, cioè nella concreta dimensione del testo che si sta analizzando. A differenza dei testi bussola, da intendersi come verifiche *formative*, le simulazioni mettono

l'accento sulla verifica sommativa di abilità e competenze. Tali esercizi simulano una prova standardizzata (es. Prova Invalsi), rispettandone fedelmente il numero e la distribuzione delle tipologie testuali e delle domande. Possono anche essere *adattive*, e in tal caso avvicinano lo studente alla struttura, alla metodologia e alle finalità del test, uscendo dalla dimensione di mero "teaching to the test". Eugenio fornisce al momento diverse tipologie di esercizi nel quadro della Competenza chiave europea "Comunicazione nella madrelingua" ripresa dalle Indicazioni Nazionali per il curricolo della scuola dell'infanzia e del primo ciclo di istruzione, relativamente alle competenze specifiche: 1) "*Leggere, comprendere e interpretare testi scritti di vario tipo*" e 2) "*Riflettere sulla lingua e sulle sue regole di funzionamento*". A partire dall'a.s. 2014 saranno disponibili esercizi che toccano anche la terza competenza specifica 3) "*Produrre testi di vario tipo in relazione ai differenti scopi comunicativi*". Il docente può scegliere gli esercizi free o quelli legati al libro di testo in adozione (della casa editrice Loescher). Nello svolgimento degli esercizi lo studente può contare su aiuti (direttamente forniti dalla piattaforma in base all'analisi dell'errore) e sul paratesto (cioè note contenenti glossario e/o definizioni); per un'autovalutazione può essere molto utile la cronologia in cui vengono indicati i tentativi effettuati prima di rispondere in modo corretto al quesito .

4. Conclusioni

Quali vantaggi comporta l'uso in classe e a casa di *Eugenio*? Innanzi tutto offre agli alunni l'immediata correzione dell'errore; una guida pratica alla comprensione del testo (è come avere un docente accanto che spiega precisamente come è opportuno muoversi di fronte al testo). L'esercizio quindi viene inviato dall'alunno al docente: a questo punto il docente dispone di un set di parametri che gli consente di confrontare velocemente tutte le esercitazioni svolte dai singoli alunni della classe e di capire quali sono le maggiori debolezze di ogni studente. Il tutor Eugenio ha archiviato finora circa 10.000 analisi del testo ed oltre 250.000 feedback, ciascuno dei quali taggato con una competenza e ove possibile, un'abilità riconducibile al Quadro della Prova Nazionale Invalsi, integrato, ove necessario, in tema di competenze, dalle Indicazioni Nazionali per il curricolo 2012, pubblicate nella primavera del 2013.

L'archiviazione dei dati, che vengono restituiti volta per volta al docente e allo studente, prevede un secondo step di aggregazione a livello regionale e nazionale, finalizzato allo studio dei cosiddetti "errori tipici", fondamentali per operare una revisione periodica dei contenuti e dei feedback in chiave adattiva.

Con l'uso di *Eugenio* si è notato maggiore interesse per l'analisi dei testi. Non va trascurata tuttavia una criticità: i ragazzi già dotati di un adeguato metodo di lettura e comprensione evidenziano che l'esercizio svolto in piattaforma comporta tempi più lunghi rispetto alla metodologia tradizionale. Come affrontare il problema? La possibilità di scegliere diversi brani d'autore con caratteristiche stilistiche differenti viene in aiuto del docente: in questo la piattaforma *Eugenio* evidenzia tutte le potenzialità della metodologia e learning e i relativi vantaggi. Nell'aula reale è molto difficile differenziare la didattica e rispondere alle esigenze conoscitive dei singoli discenti, nell'aula virtuale invece

tutto questo è possibile: per ogni alunno il docente potrà scegliere l'iter didattico più adeguato. Unica vera difficoltà? L'impegno del docente si moltiplica, anche se in modo direttamente proporzionale ai risultati: il profitto dei ragazzi migliora... e questo non è un risultato da trascurare!

Attualmente, gli iscritti a Cloudschooling, la classe virtuale che gestisce Eugenio Tutor di Italiano, Cicero Latin Tutor e le Palestre di Cloudschooling sono circa 27000, con un incremento medio mensile di circa 2.000 utenti, distribuiti su secondaria di primo e di secondo grado. Nel 2014 è prevista l'uscita all'interno dello stesso LMS (Learning Management System) di 8 tutor di altrettante discipline (inglese, tedesco, francese, spagnolo, storia, filosofia, greco, storia dell'arte), con contenuti per ora limitati alla secondaria di secondo grado.

A breve Eugenio sarà ancora più ricco di contenuti, sarà accompagnato da *Eugenio - tutor di scrittura* (la versione in grado di valutare le competenze di scrittura degli studenti), da *Eugenio per la LIM* (che gestisce le squadre e le classifiche, per trasformare in esperienze ludiche i ripassi collettivi in classe). Cloudschooling diventerà un social network per docenti sincronizzato via cloud e una nuova generazione di libri interattivi sviluppati da Maieutical Labs che consentirà di scambiarsi domande, commenti, collegamenti direttamente *sul testo*, realizzando l'aspetto collaborativo che è previsto anche dalle Indicazioni nazionali per il curriculum.

Bibliografia

<http://www.eugeniotutoritaliano.it/>

http://archivio.pubblica.istruzione.it/riforma_superiori/nuovesuperiori/index.html

Nota MIUR AOODGOS prot. n. 240/R.U./U del 16/1/2013 e relativi allegati
Elia G., Murgia G. (a cura di, 2008), *Collaborative E-learning* Franco Angeli, Milano, 2008.

Kaye A. (1994), *“Apprendimento collaborativo basato sul computer”*, in *Tecnologie Didattiche N.4*, Istituto per le Tecnologie Didattiche Consiglio Nazionale delle Ricerche

http://www.sapete.org/site/images/stories/glossario_ada.pdf

Liscia (a cura di, 2007), *E-learning – strategie per lo sviluppo delle competenze*; Osservatorio AITech-Assinform, Apogeo, Milano.

S. Penge, M. Terraschi (2004), *Ambienti digitali per l'apprendimento*, Anicia, Roma.

Internet Social Network e DSA

¹PASTENA N. (I) ²NAZZARO L. (I)

¹ Department of Human, Philosophical and Educational Sciences

² University of Salerno

Abstract. *L'impressionante velocità con cui Internet e i Social network (Facebook, Twitter, You Tube.....) dilagano nel mondo dell'Informazione Globale è la dimostrazione evidente di un'evoluzione dei meccanismi comunicativo-relazionali che, se da un lato, ha facilitato il contatto in tempo reale con il resto del mondo in ogni luogo e in ogni tempo, dall'altro lato, ma ha generato ed alimentato il senso di solitudine, di disagio e di precarietà dell'uomo moderno.*

È fatto evidente che, in un mondo di nativi digitali, l'uso corretto e consapevole delle ICT in ambito scolastico (in particolar modo per i bambini affetti da disturbi specifici dell'apprendimento) diventa attività privilegiata attraverso la quale scoprire e conoscere il mondo; condizione per imparare a gestire curiosità, emotività, creatività, edificando il proprio sé attraverso la comprensione dell'altro da sé e divenire, nel contempo, contesto per acquisire conoscenze e produrre nuove forme di pensiero.

Tradotto, dunque, in processo di insegnamento-apprendimento, all'interno di percorsi formativi in ambienti formali e non, esso rappresenta il nucleo centrale di una teleologia professionale atta a garantire ad ognuno il diritto ad un percorso formativo-educativo in grado di consentire ad ognuno di esprimere, al massimo livello possibile, le proprie capacità e divenire cittadino planetario.

In termini di educazione/educabilità, questo vuol dire che in uno spazio sociale e tecnologico (eticamente e valorialmente orientato) ogni bambino ha diritto a condizioni di rispetto, di tutela e di massima espressione della propria diversità e specificità.

Abstract. *The impressive speed of propagation of the Internet and Social Networking (Facebook, Twitter, You Tube) in the global world of information is the clear demonstration of the evolution of communicative-relational processes that, on the one hand, facilitated the actual contact with the rest of the world (in any place and at any time) and, on the other hand, has created and nurtured a sense of loneliness, discomfort and insecurity of modern man.*

It is an obvious fact that, in a world of digital natives, the correct and conscious use of ICT in schools (especially for children with Specific Learning Disorders) becomes privileged activity through which to discover the world, to learn how to condition manage curiosity, emotion, creativity, building themselves through an understanding of themselves and becoming, at the same time, the context for attaining knowledge and produce new ways of thinking.

Therefore, translated in the process of teaching and learning in training programs of formal and non-formal contexts, it represents the core of a teleology professional such as to guarantee the right of everyone to a educational training in a position to allow everyone to express, at the highest level possible, their skills and become a global citizens.

In terms of education/educability, it means that in a social and technological space (ethically and in values oriented) every child has the right to conditions of respect, protection and maximum expression of his diversity and specificity.

Keywords: Social Network, DSA, Internet, Digital Education.

1. Introduzione

“Un buon uso della media nei processi di insegnamento-apprendimento va ricercato, non in una sterile quanto mai inutile visione istruzionista, bensì in un approccio costruttivista, in cui bambini e adolescenti svolgono un ruolo attivo, predominante e centrale, al fine di sviluppare conoscenza ma, soprattutto, maturare competenze spendibili nel loro percorso di vita. Di qui l'attualità di Comenio e della sua idea di rendere il sapere spendibile nella vita di ogni giorno: punto fermo e di indiscutibile fondamentale importanza per chi fa didattica ad ogni livello districandosi nel caos universale dei media” (Pastena 2011)

“Malizia definisce i media **diffusori di oggetti culturali**, confermando quel ruolo pervasivo ad essi riconosciuto da molti studiosi e tendente ad influenzare l'agire sociale nelle sue forme di **gatekeeping** e **newsmaking**. Questi termini, mutuati dalla terminologia giornalistica, secondo l'Autore sono adattabili a qualsiasi contesto socio-culturale e ad ogni tipologia di media. Sia i new media che gli old media investono –

egli afferma – in maniera totalitaria il nostro essere nel mondo, costruendone fondamentalmente l'immagine in maniera tale da poter essere definiti **sense-making**. È su questi principi fondanti che si colloca l'esigenza generalizzata, e da più parti evidenziata, di essere educati ed alfabetizzati ai nuovi linguaggi dei media. Questo era già chiaro a James Hallorn, a Marsha Jones, a Len Masterman, pionieri delle pedagogie dei media del XX secolo. (Pastena 2011)

L'intento di queste note è fondamentalmente quello di indagare in che modo i nuovi bisogni educativi contribuiscano alla ri-definizione delle modalità di esplicazione della comunicazione educativa, in generale, e dei processi di insegnamento-apprendimento riferiti all'ambito dei DSA, in particolare. Nello specifico, in che modo i **new media** intervengono nei processi di apprendimento e nella crescita cognitiva e sociale dei soggetti in età evolutiva? In quale misura allargano il panorama delle strategie educatico-didattiche nei processi di insegnamento-apprendimento e quanto influiscono nell'articolazione concettuale e nella strutturazione di percorsi educativi che conducano verso un progetto di vita e verso un orizzonte di senso?

2. ICT e DSA

Nel panorama educativo-istituzionale italiano si respira, ancora a tutt'oggi, nonostante le innumerevoli iniziative messe in atto per fronteggiare il problema, un'aria di scarsa conoscenza e di scarsa consapevolezza dei problemi sociali e culturali che caratterizzano i soggetti con **Disturbi Specifici di Apprendimento**. Ciò comporta inevitabilmente gravi ritardi e innumerevoli difficoltà nella reale comprensione del fenomeno, con conseguente inadeguatezza delle risposte e degli interventi effettuati per garantire il successo formativo di questa tipologia di alunni (dotati di normale intelligenza e di notevoli potenzialità).

L'uso delle tecnologie, ed in particolare dei social network, può rappresentare, in tal senso, uno strumento compensativo di enorme impatto non solo per il controllo delle emozioni ma, in maniera prevalente, per garantire la possibilità di creare comunità a distanza attraverso le quali socializzare esperienze, condividere risorse, pianificare strategie. Di fatto, l'alunno con DSA che frequenta la scuola vive all'interno di un preciso contesto territoriale nel quale manifesta le sue esigenze formative, i suoi disagi, le sue difficoltà.

Può, dunque, nell'epoca dell'**umanesimo tecnocratico**, un **soggetto-persona** esprimere in modo completo il proprio diritto alla cittadinanza senza il possesso degli **strumenti** e dei **linguaggi** che caratterizzano, in maniera globale, la **Società dell'Informazione e della Comunicazione**?

Essere educati all'uso consapevole dei **media** e dei **social network** può consentire di beneficiare a pieno titolo di strategie e di strumenti tecnologici a sostegno dei processi di apprendimento?

Gli studenti con **DSA** necessitano di recuperare competenze di base non ancora consolidate attraverso strategie in grado di rispondere in maniera ottimale alle loro peculiari caratteristiche cognitive e ai loro particolari bisogni educativi speciali (**BES**).

L'uso consapevole delle tecnologie può sicuramente garantire il raggiungimento di numerosi obiettivi come:

- Garantire la valorizzazione di linguaggi comunicativi diversi rispetto al codice scritto (linguaggio iconografico, linguaggio parlato) e l'uso di mediatori didattici.
- Incentivare la promozione di inferenze, integrazioni e collegamenti tra le conoscenze e le discipline.
- Attivare il sostegno dell'apprendimento attraverso percorsi metacognitivi e attraverso la didattica laboratoriale.
- Sollecitare la promozione dell'apprendimento collaborativo, il tutoraggio tra pari, l'autocontrollo e la gestione delle proprie azioni.

L'attività di gestione di un percorso didattico basato sull'utilizzo di **Wiki** implica, ad esempio, l'impiego dell'ambiente di apprendimento per attivare azioni collaborative di **editing**, di **immissione di dati/informazioni**, di **cancellazione**, di **creazione di hyperlink** e di **nuove pagine** attraverso **link di rimando**.

È necessario creare **ambienti di apprendimento** in grado di privilegiare attività di collaborazione del gruppo-classe senza preoccupazioni rispetto alla possibilità di sbagliare; il gruppo, in tal senso, assume la funzione di supporto nell'individuare e/o correggere eventuali errori.

Un'attività che aspiri all'acquisizione di una competenza comunicativo-relazionale, interagendo con la più ampia comunità civile e sociale, contribuisce allo sviluppo del **senso di responsabilità**, alla **valorizzazione delle qualità personali**, all'**accrescimento dell'autostima** e alla **concreta realizzazione di un obiettivo pienamente condiviso**.

Questi elementi, non solo mettono in risalto la forte valenza dell'uso appropriato delle tecnologie ma, contemporaneamente, sottolineano la possibilità di **costruire curricoli** seguendo uno **schema reticolare** non rigido e lineare ma aperto, flessibile e suscettibile di continue revisioni.

Lo studente con DSA ha pieno diritto all'acquisizione di specifiche competenze, declinabili in comportamenti attivi e responsabili durante tutto il corso della sua vita. (Costituzione italiana, art.3: principio di

uguaglianza e rimozione degli ostacoli; art.32 diritto alla salute; art.34: diritto all'istruzione e formazione - Legge 104/92 e D.lgs 297/94: integrazione scolastica per lo sviluppo delle potenzialità delle persone con handicap).

La scuola, responsabile del conseguimento del successo formativo di tutti e di ciascuno, deve applicare e rendere concreta la normativa vigente riuscendo a far perseguire anche agli alunni con DSA, le seguenti finalità:

- a) garantire il diritto all'istruzione;
- b) favorire il successo scolastico, anche attraverso misure didattiche di supporto, garantire una formazione adeguata e promuovere lo sviluppo delle potenzialità;
- c) ridurre i disagi relazionali ed emozionali;
- d) adottare forme di verifica e di valutazione adeguate alle necessità formative degli studenti;
- e) preparare gli insegnanti e sensibilizzare i genitori nei confronti delle problematiche legate ai DSA;
- f) favorire la diagnosi precoce e percorsi didattici riabilitativi;
- g) incrementare la comunicazione e la collaborazione tra famiglia, scuola e servizi sanitari durante il percorso di istruzione e di formazione;
- h) assicurare eguali opportunità di sviluppo delle capacità in ambito sociale e professionale.

3. Educare ai “Media”... con i “Media”

Quanti sono i bambini già in possesso di un telefono cellulare di ultima generazione? Come e quando lo usano? Oggi i bambini si muovono nel mondo dell'alta tecnologia con estrema disinvoltura, spesso alienandosi dalla realtà sociale e in un clima di totale immersione nella realtà virtuale dove ogni atto è consentito, dove ogni abuso è possibile e dove ogni violenza assume le sembianze del reale.

In questo contesto le pratiche educative tradizionali risultano obsolete, non più sufficienti a fronteggiare le insidie del mondo moderno e a fornire le competenze necessarie per preparare gli studenti alle sfide culturali del mondo della **globalizzazione/universalizzazione/mondializzazione dell'informazione**.

La scuola, **attrezzata di tecnologie**, deve prevedere almeno due momenti d'azione educativo-didattica: da un lato, l'esercizio all'uso strumentale delle tecnologie e, dall'altro lato, l'esercizio all'uso di una didattica metacognitiva.

È risaputo che i bambini e i ragazzi arrivano a scuola già con un solido background tecnologico perché **digital natives** (Prensky 2010); posseggono una buona manualità, una grande abilità nell'esecuzione di numerose azioni (multitasking), ma è la scuola (intesa come sottosistema sociale e incaricata dalla stessa società di curare l'istruzione e l'educazione delle giovani generazioni) a dare significato e senso all'azione, a specificare i giusti canali d'accesso direzionando il **saper fare** dell'alunno verso un suo più consapevole **saper essere**. Essere competenti oggi non significa minimizzare i tempi d'esecuzione di un compito bensì essere in grado di individuare le giuste connessioni, le idonee relazioni, le interdipendenze tra i vari livelli della conoscenza. Vuol dire, altresì, **saper comunicare, saper interpretare con senso critico** le informazioni, le insidie e i pericoli nascosti nell'affascinante mondo delle ICT, dei social network e di quel **pianeta misterioso** qual è spesso **Internet**.

“Gli strumenti digitali già estendono e arricchiscono le nostre capacità cognitive in molti modi. La tecnologia digitale migliora la memoria, per esempio attraverso gli strumenti di acquisizione, archiviazione e restituzione dei dati. La raccolta digitale di dati e gli strumenti di supporto alle decisioni migliorano la capacità di giudizio consentendoci di raccogliere più dati di quanti ne riusciremmo a mettere insieme da soli, aiutandoci a condurre analisi più complesse e aumentando il nostro potere di chiedere “*cosa succederebbe se...?*” e verificare tutte le implicazioni derivanti da quella domanda. Il **potenziamento digitale** in ambito cognitivo, reso possibile da laptop, data-base on-line, simulazioni tridimensionali virtuali, strumenti collaborativi on-line, palmari e da una serie di altri strumenti specifici per diversi contesti, è una realtà in molte professioni, anche in campi non tecnici come la giurisprudenza e le discipline umanistiche” (Prensky2010).

Un utilizzo responsabile delle tecnologie implica la capacità di controllare con lucidità gli eventi che si avvicinano in questi ambienti, giungendo a riconoscere e gestire consapevolmente le proprie emozioni, evitando di subire il fascino di un incontro in rete, o di sentirsi offesi per il comportamento di qualche amico, o essere turbati dalla visione di certe immagini o essere succubi dell'influenza che possono produrre talune informazioni (così come evidenziano gli ultimi tristi casi di giovani, tra i quali citiamo l'attivista informatico del web Aaron Swartz, che, con gesti estremi, hanno deciso di porre fine alla loro vita per essersi sentiti violati nella dignità di persone).

Scattare una fotografia o girare un piccolo video per poi pubblicarlo in Internet è oggi estremamente facile, con il risultato di non poter più parlare di produzione cinematografica rigidamente separata dal pubblico. La posizione degli adulti, in particolar modo degli insegnanti, gioca qui un ruolo cruciale, diventa essenziale che l'alunno sia in grado di analizzare consapevolmente sia gli effetti degli atti da lui prodotti, sia le cause che

comportamenti inconsapevoli possono generare (così come diventa centrale capire il perché di determinati comportamenti e ancor più essere in grado di prevenirli).

Ci sono comportamenti, infatti, che possono risultare inspiegabili se si considera il soggetto come singolo, ma che acquistano immediatamente senso se collocati all'interno di un contesto più ampio (come la scuola) in cui il soggetto normalmente vive e si relaziona con gli altri.

In questo scenario trova sicuramente spazio una scuola che fonda la sua **mission** sulla **qualità delle comunicazioni e relazioni** (Bateson) favorendo un clima sereno e collaborativo grazie al quale si alimenti la motivazione ad apprendere in modo cooperativo, si espanda il desiderio di cambiamento-innovazione, si innalzino i livelli di autostima, si prevengano atteggiamenti di illegalità, di bullismo e di conflittualità: sintomi primari di mancanza di comunicazione, che troviamo ampiamente diffusi tra gli studenti.

A tal proposito, già da tempo molte realtà scolastiche si sono orientate verso percorsi di **media education** e non soltanto di **educational through media**.

E' solo aiutando non solo a conoscere realmente gli strumenti che abitualmente si usano, ma a valutarne gli esiti in maniera consapevole, che si potrà "garantire" un'autentica capacità di costruire e di sviluppare in libertà e armonia una **propria identità** e un **proprio futuro migliore**.

In ultima analisi, il **digital native** necessita di un'educazione **ai media** oltre che **con i media**; l'uso delle tecnologie dovrebbe essere pienamente integrato nella pratica educativa per garantire un apprendimento personalizzato e flessibile ma sempre eticamente e valorialmente orientato.

Tutto ciò richiede la necessità di muoversi in una dimensione educativa che tenga prevalentemente conto dei bisogni affettivi e sociali di riferimento, del riconoscimento dei diritti di ognuno, primo fra tutti quello alla piena e consapevole partecipazione ai sistemi di convivenza e di co-esistenza sociale che generano ineludibilmente il destino di ogni essere umano.

Bibliografia

BATESON, G. (2008). *Mente e natura. Un'unità necessaria*. (trad. it. di G. Longo). Milano: Adelphi. [1979]

CALVANI A. (2011) *Principi dell'Istruzione e strategie per insegnare. Criteri per una didattica efficace*. Roma: Carocci.

MALIZIA P. (2002). *La seconda realtà. Riflessione sulla realtà "dai" media*. In "Sociologia". Rivista Quadrimestrale di Scienze Storiche e Sociali. Gangemi Editore, Roma, 2002, Anno XXXVI, n. 1, p. 141

PASTENA N. (2011). *Giochi di parole per giochi di pensiero nel labirinto dell'integrazione*. Napoli: De Nicola EditorEE, p.48

PRENSKY M. (2010). *H. Sapiens Digitale: dagli Immigrati digitali e nativi digitali alla saggezza digitale*. TD-Tecnologie Didattiche, 50, pp. 17-24

Digital Literacy nella società della conoscenza: alcune riflessioni educativo-didattiche per superare il Digital Divide

Alessandro Barca, Mariella Tripaldi¹

Docente c/o XVI C.D. "Europa" di Taranto e Docente a contr. c/o l'Univ. Lumsa di TA, la Facoltà Teologica Pugliese di Bari e Molfetta e l' I.S.S.R. "R. Guardini" di TA

E-mail: alessandro.barca@istruzione.it

¹Docente presso I.C. "A. Manzoni" di Lizzano (TA)

E-mail: mariella.tripaldi@istruzione.it

Technological innovation, which permeates more and more quickly and deeply into the knowledge society makes it imperative continuous adaptation of our knowledge, which are increasingly subject to rapid obsolescence. The school's task is therefore to define new strategies for education, instruction and training to enable, along with traditional transmission of knowledge and know-how, also the acquisition of skills needed for easy access to information in a wise and critical mastering all the various communication tools and make the individual able to cope with change and complexity that characterize contemporary society and avoiding the risk of social exclusion.

Keywords: Digital Literacy 1, Knowledge Society 2, Digital Divide 3, Lifelong Learning 4.

1. La Knowledge Society e il Digital Divide

Nell'attuale società poliedrica e "liquida" [Bauman, 2008] del terzo millennio, dove si vivono contesti che si modificano ancor prima che il suo agire riesca a consolidarsi in procedure, routine e abitudini, sorge più che mai l'esigenza non solo di "sapere" e "saper fare", ma principalmente di imparare a imparare per saper e voler agire e reagire con efficacia e efficienza in situazioni e condizioni imprevedibili, complesse, precarie, incerte, "liquide", trovando il giusto equilibrio nel movimento. Per descrivere la società contemporanea è stato frequentemente utilizzato il concetto di "società dell'informazione e della comunicazione" o «villaggio globale dei media» [Alberici, 2002] ma in questi ultimi anni è nata l'esigenza di definirla "società della conoscenza" [Spinelli, 2009]. Con "knowledge society" o società della conoscenza ciò che si vuole mettere in evidenza è che gli individui, nella vita personale come nel lavoro, sono tenuti continuamente a cercare, acquisire ed elaborare il sapere e la conoscenza come un nuovo capitale e come fondamento strutturale dell'economia e dello sviluppo sociale e culturale. Occorre, di conseguenza, favorire «un'interazione sinergica tra innovazione tecnologica e valori umani» [Castells, 2004] e definire delle nuove strategie di educazione, istruzione e formazione per consentire, insieme alla tradizionale trasmissione dei saperi e allo scambio e alla condivisione di buone pratiche [Tornerò et al., 2010] anche un'acquisizione di competenze che rendano l'individuo in grado di affrontare e

Congresso Nazionale AICA 2013

rispondere ai cambiamenti e alla complessità che caratterizzano la società contemporanea e accedere facilmente alle informazioni permettendo a tutti, e a tutte le età, di essere inclusi nella società della conoscenza per contrastare e, successivamente “sradicare” del tutto il Digital Divide, ossia ogni forma di disuguaglianza, barriera, limitazione esistente nell’accesso e nell’utilizzo delle nuove tecnologie dell’informazione e della comunicazione da parte di categorie di persone o di intere aree geografiche, alla cui base esistono disparità di tipo economico, sociale, culturale e politico.

2. Digital Literacy nella scuola e nella società

In tutto questo turbinio di cambiamenti e trasformazioni sta mutando profondamente ed inesorabilmente anche la scuola. Se infatti la società cambia la scuola deve necessariamente cambiare ma se la scuola non si evolve mentre la società cambia, è inevitabile che la scuola entri in una crisi radicale in quanto prepara i ragazzi a una società che non c’è già più.

La scuola anche nell’attuale contesto a cui si aggiunge la sempre più rapida obsolescenza delle conoscenze, non può sottrarsi al compito di «sviluppare l’attitudine a contestualizzare e globalizzare saperi e competenze, stimolare la capacità di fare collegamenti fra conoscenze, abilità, esperienze», [Morin, 2000] fornire ai giovani una metodologia di apprendimento di tipo critico e riflessivo e soprattutto le competenze necessarie per vivere al meglio nel mondo attuale e ancor più in quello ancora in buona parte imprevedibile che li attende; ma per far ciò occorre essenzialmente una nuova formazione iniziale del corpo docente (partendo già dall’università) e un aggiornamento costante e continuo in modo da renderlo consapevole del nuovo processo di insegnamento/apprendimento e pronto a fronteggiare con competenza le sfide educative che di giorno in giorno gli si pongono innanzi. Assieme ai numerosi altri obiettivi che già persegue, non potrà, quindi, non tener conto delle ICT e le sue implicazioni nei vari ambiti; dovrà pertanto: valutare meglio le potenzialità e i valori aggiunti delle ICT (questo anche grazie alla formazione e all’aggiornamento continuo dei docenti) mettendo a punto tipologie di impiego e metodologie educative che in aggiunta ai consueti mezzi didattici siano in grado di stimolare l’acquisizione di conoscenze e meta-competenze e di sviluppare le capacità di auto-apprendimento critico da parte degli allievi; adottare le pratiche necessarie affinché gli studenti di oggi e i cittadini di domani possano essere utilizzatori attivi e consapevoli di ogni strumento e forma comunicativa; valutare quali siano le competenze specifiche necessarie per un uso efficiente e creativo delle ICT per poter attivare gli insegnamenti ritenuti più idonei e favorire tutte le competenze necessarie, che possano essere aggiornate continuamente e ad ogni età, affinché si possa arginare il digital divide.

Nelle Raccomandazioni del Parlamento e del Consiglio Europeo del 18 dicembre 2006 viene definito requisito indispensabile per l’esercizio di una cittadinanza attiva, la competenza digitale o Digital literacy: «essa consiste nel saper utilizzare con dimestichezza e spirito critico le tecnologie della società dell’informazione per il lavoro, il tempo libero e la comunicazione. Essa deve essere supportata da abilità di base nelle ICT: l’uso del computer per reperire,

valutare, produrre, presentare e scambiare informazioni, nonché per comunicare e partecipare a reti collaborative tramite Internet».

Il tema della digital literacy, è andato affermandosi negli ultimi decenni parallelamente al processo di digitalizzazione della società contemporanea come nodo cruciale nella riflessione educativa del nuovo millennio [Calvani, et al., 2010]. Al di là delle definizioni date degli autori più rappresentativi, di cui si rimanda per maggiori approfondimenti, come Martin, Midoro, Jenkins, Tornero, che sottolineano come la digital literacy sia la risultante di una combinazione stratificata e complessa di capacità, abilità e conoscenze, Antonio Calvani sottolinea che «la competenza digitale non riguarda solo l'expertise tecnica, ma è densa di implicazioni cognitive e relazionali. E oggi questo concetto si sposta sempre più ad accogliere ulteriori dimensioni etiche e sociali come e-citizenship, e-participation ed e-inclusion»; [Calvani et al., 2012]

Ma allora oltre al cosa, come si apprende nella società della conoscenza? La knowledge society, infatti, così come si caratterizza e come si evolve, impone un nuovo modo di pensare e approcciarsi all'educazione e alla formazione e, di conseguenza alla digital literacy. Quello che cambia anzitutto è la finalità stessa della formazione; alla base vi è il concetto di empowerment ossia l'apprendimento visto come potenzialità individuale. Non basta avere accesso e acquisire un numero sempre maggiore di informazioni ma occorre sviluppare quelle metacompetenze, vale a dire quella abilità e quelle capacità di appropriarsi strategicamente delle informazioni e delle conoscenze, ma anche quei saperi "taciti", formali e informali, che permettono agli individui di essere capaci di affrontare i continui cambiamenti della società odierna, di costruire sempre nuove forme di sapere condiviso, apprendendo lungo tutto l'arco della vita (lifelong learning). Questo inoltre fa sì che i luoghi dell'apprendimento non siano soltanto quelli istituzionalmente deputati ad esso, ma una pluralità di situazioni. L'apprendimento diviene un continuo processo di interazione tra variabili interne ed esterne e ciò lo rende un fenomeno interdisciplinare inteso come pluralità di approcci disciplinari e metodologici, come diversificazione dei luoghi e delle situazioni di apprendimento. [Montedoro, 2004] Le competenze richieste saranno quindi meno rivolte allo "svolgimento" di compiti e azioni o di una mansione secondo procedure determinate a priori e molto di più legate al ragionamento e alla capacità di imparare ad imparare. A nostro parere ma con le dovute accortezze, utili a questo nuovo approccio possono essere tutte e tre le prospettive educative monocognitive, metacognitive e fantacognitive indicate da Guerra. [Guerra, 2002]

Le nuove tecnologie dovranno essere utilizzate pertanto come strumenti per attivare esperienze che mettano il soggetto in contatto con il mondo e stimolino la capacità di esprimersi e di scoprire in modo originale, creativo e orientato a comprendere anche l'altro mettendo nell'ottica tipica dell'empowerment, il soggetto nella condizione di poter diventare/ridiventare realmente protagonista e responsabile della propria vita, delle proprie scelte, del mondo che lo circonda.

2. Conclusioni

Le ricerche in ambito educativo e didattico degli ultimi tempi mettono sempre più in evidenza che l'uso delle ICT nei sistemi educativi favorisce il miglioramento dell'apprendimento e soprattutto il processo dell'imparare a imparare, abilità basata sulla capacità di adattarsi ai cambiamenti che si verificano a livello relazionale e ambientale nel sistema di cui l'individuo fa parte attraverso la costruzione e la trasformazione dei suoi modelli di conoscenza e di azione. Risulta necessario a tal proposito fornire i mezzi per collegare una pluralità di conoscenze, evitare i pericoli nascosti nel cyberspazio e per favorire la capacità di gestione dell'imprevisto, la comunicazione, e garantire la valutazione critica e l'autonomia dell'individuo; utile appare una nuova tecnologia dell'educazione intesa come un complesso sistema di artefatti che possono potenziare la comunicazione didattica e come tale intervenire nei processi di insegnamento/apprendimento. [Falcinelli, 2012] Dalla riflessione sulle trasformazioni della società e della scuola emerge l'esigenza di un nuovo tipo di formazione che parte già dagli alunni della scuola primaria per giungere, con l'aggiornamento continuo, fino all'anziano; una formazione che pone il focus sulle competenze strategiche e trasversali, connesse e interdipendenti con le competenze "di base" e quelle specialistiche; una nuova formazione "alta", quindi universitaria e post-universitaria e una formazione e/o aggiornamento che pone il proprio focus sugli adulti/anziani per ridurre il digital divide e favorire l'e-governance e l'e-participation.

Bibliografia

- Alberici, A., *Imparare sempre nella società della conoscenza*, Bruno Mondadori, Milano, 2002
- Bauman, Z., *Vita liquida*, Laterza, Roma-Bari, 2008
- Calvani, A., Fini, A., Ranieri, M., *La competenza digitale nella scuola. Modelli e strumenti per valutarla e svilupparla*, Erickson, Trento, 2010
- Calvani, A., et al., *Are young generations in secondary school digitally competent? A study on Italian teenagers*, in "Computers & Education", vol. 58, 2012, 797-807
- Castells, M., *La città delle reti*, Marsilio Editore, Venezia, 2004, 23
- Falcinelli, F., *Tecnologie dell'educazione in L'Agire Didattico, Manuale dell'insegnante*, a cura di Rivoltella, P.C., Rossi, P.G., La Scuola, Brescia, 2012
- Guerra, L., *Tecniche e tecnologie per la mediazione didattica* in L. Guerra (a cura di), *Educazione e tecnologie. I nuovi strumenti della mediazione didattica*, Junior, Bergamo, 2002
- Guerra, L., Zanetti F., *Digital divide: analisi del fenomeno e prospettive di superamento*, Stampa inedita, Bologna, 2005
- Montedoro, C., in ISFOL, *Apprendimento di competenze strategiche. L'innovazione nei processi formativi nella società della conoscenza*, Franco Angeli, Milano, 2004
- Morin, E., *La testa ben fatta. Riforma dell'insegnamento e riforma del pensiero*, Raffaello Cortina, Milano, 2000
- Spinelli, A., *Un'officina di uomini. La scuola del costruttivismo*, Liguori Editore, Napoli, 2009, 28
- Tornero J.M.P., Paredes O. and Simelio N., *Media literacy in Europe. From promoting digital literacy to the audiovisual media services directive*, URL: <http://formare.erickson.it/wordpress/en>, Form@re, n. 70 September, 2010.

Sampling issues and management solutions in internet-based market researches

Maria V. Ciasullo, Giuseppe Festa

Università degli Studi di Salerno, Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology)
Via Giovanni Paolo II – 132, 84084 Fisciano (SA)
miciasullo@unisa.it, g festa@unisa.it

Questo lavoro è frutto di un impegno comune. Tuttavia, i paragrafi 1, 2 e 6 sono da attribuirsi a Maria V. Ciasullo, mentre i paragrafi 3, 4 e 5 sono da attribuirsi a Giuseppe Festa.

Abstract. *In the wide context of internet-based market researches, this paper focuses on those ones carried out on net users. An analysis of the advantages, but also disadvantages, related to online market researches shows that a crucial problem is the 'vagueness' of the representativeness of sample surveys, up to affecting the credibility of the online survey. Therefore, the authors have developed a theoretical / practical framework that integrates contributions from marketing, statistics and informatics, following a logical-methodological path oriented to give minor uncertainty degrees to online surveys. This model recovers useful considerations from inferential statistics and proposes also some solutions for managing internet samples as virtual communities. The conceptual paper aims to contribute to an advancement in the field of online market researches, shifting the focus about the validity / reliability of the investigation from the mere software or statistical tool to a wider analytical internet marketing strategy.*

Keywords: *e-research, analytical internet marketing, online sampling.*

1. Introduzione

L'orientamento *customer based* [Valdani e Busacca, 2000], requisito inderogabile per la sopravvivenza e il successo delle organizzazioni, caratterizza in modo sempre più preponderante l'agire delle imprese dei nostri giorni. In tale logica, diviene fondamentale un'approfondita conoscenza delle esigenze di chi acquista, del suo modo di interpretare la realtà e dei meccanismi attraverso i quali operano i processi di scelta. Parallelamente, un'economia sempre più globalizzata, caratterizzata da una rapida trasferibilità delle informazioni, ha contribuito a delineare non solo un cliente più attento e complesso, ma anche nuovi stili di consumo basati sulla capacità dei beni di esprimere la cultura, il gusto, lo stile di ciascun individuo.

I continui cambiamenti associati a una maturazione dei mercati e delle relative dinamiche innovative di acquisto impongono alle imprese un costante impegno nell'analisi e nella comprensione dei fenomeni anzidetti. Diviene pertanto comprensibile come, nell'ambito degli studi di marketing management, sia stata posta particolare attenzione alla funzione analitica del marketing, tesa ad analizzare la struttura del mercato e l'evoluzione della domanda, al fine di individuare i fattori più significativi che possano influenzare i comportamenti d'acquisto.

Le ricerche di mercato, generalmente intese quali indagini sistematiche e obiettive finalizzate alla raccolta di elementi rilevanti per successivamente elaborare e implementare strategie di interazione con i clienti e con gli altri attori del sistema del valore dell'impresa, assolvono alla finalità di osservare mercati e delineare scenari. È quindi evidente come le stesse costituiscano *driver* fondamentali nell'agire dell'impresa, laddove la competitività di quest'ultima si basi sulla capacità di conquistare e alimentare costantemente il sostegno a una domanda matura, esigente e innovativa.

Nella complessità di tale contesto s'inserisce la valenza delle tecnologie ad alta intensità connettiva (prevalentemente, se non esclusivamente, *internet-based*), capace di contribuire in modo significativo alla ricerca sociale e, nello specifico, alle ricerche di mercato. Internet si presenta infatti quale strumento "naturalmente" idoneo a raccogliere informazioni in maniera mirata, approfondita e metodologicamente corretta sulla domanda [Kiang et al., 2000], sia nei termini delle ricerche di mercato sia ancor prima tramite opportune operazioni di *marketing intelligence* [Kotler, 2001]. In tale prospettiva, si sottolinea un'importante proprietà della rete, ossia la sua capacità di rendere i processi comunicativi sostanzialmente indipendenti dai vincoli spazio-temporali: è evidente, infatti, come tale proprietà si dimostri estremamente valorizzante per la ricerca online, in quanto consente il superamento di una serie di barriere fisiche che possono esistere tra il ricercatore e i soggetti sui quali viene condotta la rilevazione.

Inoltre, come evidenziato in letteratura [Cantone et al., 2006; Vescovi, 2007], le nuove tecnologie dell'informazione e della comunicazione, in particolare quelle che si basano su standard aperti e universali, assumono un ruolo fondamentale nel miglioramento della gestione del sistema delle relazioni intra e inter-impresa e dei sottostanti processi di creazione di valore, facilitando altresì la fase di avvicinamento a culture e mercati internazionali. Per tutte queste ragioni, è evidente come l'efficacia delle ricerche di mercato richieda un continuo aggiornamento che si combina inevitabilmente all'evoluzione *internet-based* dei sistemi informativi di marketing e quindi, nella prospettiva di questo lavoro, alle strategie di internet marketing analitico in senso lato.

2. Le ricerche di mercato internet-based: natura, tassonomia e problematica

La combinazione tra marketing e statistica è tradizionalmente alla base del sistema di competenze che regola il mondo delle ricerche di mercato [Bassi, 2008]; a voler essere più precisi, in particolare, il marketing, in questo particolare contesto, si serve della statistica inferenziale prevalentemente nella fase del campionamento, allo scopo di estendere alla popolazione le osservazioni sul campione (*sample*). Nel tempo, a questo binomio di competenze ha finito per accodarsi una terza disciplina, ossia l'informatica, che, in ragione dei vantaggi conseguibili prevalentemente in termini di tempi e costi, ha contribuito a potenziare notevolmente l'interazione tra le due competenze anzidette, al punto che oggi, quasi senza timore di smentita, il sistema delle ricerche di mercato è per l'appunto governato dalla sintesi di queste tre discipline (marketing, statistica e informatica).

L'avvento dell'internet, inoltre, non ha soltanto generato una nuova serie di strumenti informatici a disposizione del ricercatore, ma, in ragione del suo successo "sociale", ha finito per consegnare al mondo delle ricerche di mercato anche un nuovo polmone d'indagine. In tal modo, pertanto, diventa possibile distinguere tra le ricerche di mercato svolte grazie alle nuove soluzioni in rete (e parleremo propriamente di ricerche di mercato "con" internet) e le ricerche di mercato svolte presso gli utenti della rete (e parleremo propriamente di ricerche di mercato "su" internet): le due categorie insieme, evidentemente, costituiscono la famiglia allargata delle *ricerche di mercato internet-based*, che, in altre parole, in un modo o in un altro hanno a che fare con l'internet.

Nell'ambito del presente lavoro, in particolare, non ci soffermeremo sulla prima categoria ("con"), in primo luogo perché esclusivamente strumentale e in secondo luogo perché già analizzata in un precedente lavoro [Festa, 2001b]. In questa sede, invece, si vuole ragionare della seconda categoria ("su"), identificandone la principale problematica di riferimento (ossia la rappresentatività campionaria) e proponendone, se possibile, un'analisi teorica, che tenga in conto il contributo integrato proprio del marketing, della statistica e dell'informatica.

Si ritiene, anzitutto, che un corretto approccio metodologico alla tematica delle ricerche di mercato *internet-based* non possa prescindere dall'impostazione delle ricerche di mercato tradizionali, verificando se contestualmente la relativa "cassetta degli attrezzi" (in termini di tecniche e strumenti) sia inapplicabile (perché non adattabile all'ambiente internet), applicabile con indispensabile adattamento (dovendo necessariamente operare in modalità informatica) o, infine, applicabile in chiave innovativa (sfruttando le notevoli possibilità dell'ambiente internet). A muovere questa classificazione, che parte da un mero ragionamento epistemologico [Festa, 2001a], sono prevalentemente due fattori: le caratteristiche "tecnologiche" della piattaforma internet e le caratteristiche "sociali" della popolazione internet, che per brevità si analizzano congiuntamente grazie a una rapida panoramica dei vantaggi e degli svantaggi collegati alle ricerche di mercato online (senza voler ovviamente approfondire la tematica, che di per sé sarebbe vastissima).

Per quanto riguarda alcuni dei principali vantaggi, si pensa, prevalentemente, all'efficienza e all'economicità, in termini di risparmi sui tempi (data la velocità di elaborazione delle attuali attrezzature informatiche, hardware e software) e sui costi (non in senso assoluto, ma ovviamente se paragonati ai costi normali di una ricerca di mercato tradizionale), arrivando a una sorta di consacrazione dell'internet quale canale più efficiente per la raccolta dei dati [Weible e Wallace, 1998]. Tuttavia, esistono anche vantaggi più significativamente qualitativi: le potenzialità dell'elaborazione informatica, infatti, non sono da considerare soltanto in riferimento ai dati immagazzinati, ma anche e soprattutto in riferimento alle modalità di immagazzinamento, capaci di realizzare un'esperienza più attraente per l'intervistato (grazie alla multimedialità), una maggiore indipendenza dei tempi e degli spazi di rilevazione (e, quindi, maggiore comodità e consequenziale precisione di rilevazione, sia per l'intervistato sia per l'intervistatore), una maggiore indipendenza dal contesto (personale e/o situazionale: si pensi a disabili, ammalati, reclusi, zone di guerra, zone di epidemia, ecc.), un abbattimento delle barriere psicologiche e socioculturali (facendo leva sulla sensazione di assenza della personalità fisica dell'interlocutore) e così via [Schmitz, 2004; Wright, 2005].

Il complessivo maggior senso di comfort, inoltre, produce in genere risultati più ampi e profondi, soprattutto nelle ricerche qualitative, che, fin d'ora e come si suggerirà per altre motivazioni anche in seguito, sembrano costituire uno scenario di approfondimento altamente interessante per le ricerche di mercato *internet-based* [Andreani e Conchon, 2001b], richiedendo contestualmente nuove competenze al ricercatore (si pensi, per esempio, alle abilità richieste al moderatore di un gruppo online). In generale, infatti, la ricerca qualitativa aspira a ottenere una profonda comprensione del comportamento del consumatore, attraverso un'analisi approfondita delle aspirazioni e delle sensazioni associate per esempio a una categoria, un brand, un'immagine [Mariampolski, 2001; Marbach, 2010]. Al di là della rappresentatività statistica (indispensabile per le tecniche quantitative), le tecniche qualitative presentano il vantaggio di essere applicabili anche nei casi in cui l'informazione sia difficile da rilevare, complessa da analizzare o di problematica interpretazione [Andreani, 1998; Tissier-Desbordes, 1998; Andreani e Conchon, 2001a] proprio perché conferiscono maggiore libertà espressiva ai soggetti intervistati [Prandelli e Verona, 2006]. In questo modo, grazie a internet, l'intervistato diventa completamente "padrone" dell'intervista, senza che l'intervistatore possa giocare da "manipolatore". Se così non fosse e dovesse accorgersene, infatti, all'intervistato basterebbe un clic per andarsene o, peggio ancora, per imbrogliare l'intervistatore.

Al progettista della ricerca qualitativa su internet e al relativo intervistatore, pertanto, serviranno competenze tecnologiche, metodologiche e soprattutto socio-relazionali perfettamente in linea con l'ambiente internet. Anche per questi motivi, in particolare, le ricerche di mercato più interessanti su internet sono a nostro avviso proprio le ricerche qualitative, perché possono far emergere informazioni che in altri contesti difficilmente si può essere

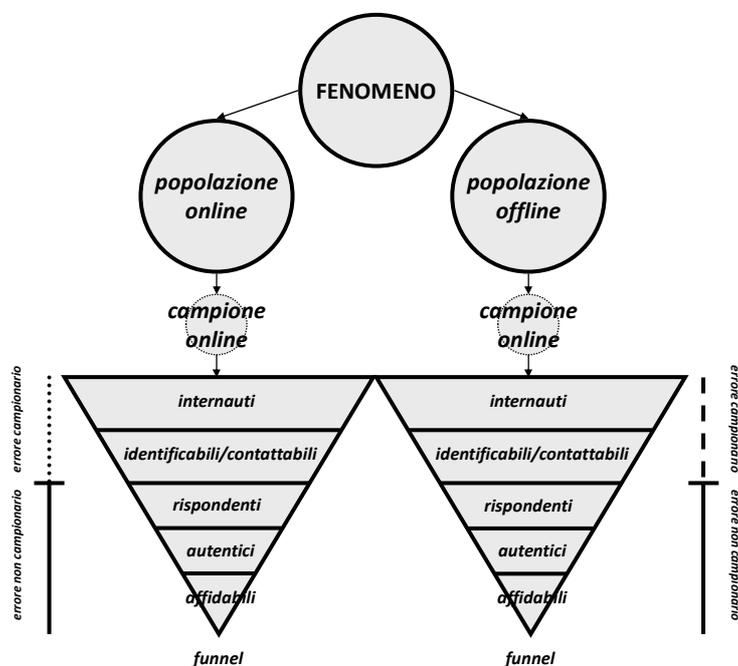
disposti a confessare (anche ricorrendo alla distorsione complessivamente generabile con il linguaggio para-verbale e non-verbale).

Esistono naturalmente anche diversi svantaggi collegati alle ricerche di mercato online, che sembrano, in ogni caso, ampiamente compensati dai relativi vantaggi, già commentati in precedenza: si fa riferimento, per esempio, alla maggiore o minore familiarità dell'utente con il "linguaggio" della tastiera e del mouse (si pensi al monitoraggio di una chat), alla perdita d'informazione sul contesto (si pensi a un questionario online a cui si risponde in ufficio oppure da casa, dovendo evidentemente "sottostare" a regole di comportamento diverse in ambienti diversi) e all'impossibilità di raccogliere segnali non-verbali e para-verbali (e forse anche verbali: si pensi, per esempio, a modi di esprimersi fisiologicamente diversi nella scrittura elettronica rispetto all'uso della penna o ancor di più della voce, ecc.). Tuttavia, come evidenziato da molteplici studi [Howard et al., 2001; Andrews et al., 2003; McDonald e Stewart, 2003; Di Fraia, 2004; Wright, 2005] la principale mancanza delle ricerche su internet è legata alla "vaghezza" della rappresentatività dell'indagine campionaria, che finisce per diventare un problema di notevole importanza, al punto da condizionare la stessa credibilità dell'indagine online.

Avendo ampiamente verificato in letteratura la consapevolezza di tale lacuna, la principale finalità di questo paper di ricerca concettuale è quella di elaborare un *framework* teorico / pratico che possa contribuire a meglio delimitare il perimetro del problema della rappresentatività campionaria su internet, proponendo altresì specifiche soluzioni gestionali, derivanti da una complessiva strategia di internet marketing analitico. In tale logica, si ritiene di poter far emergere alcuni specifici spunti di riflessione che potranno essere oggetto di ulteriori ricerche in materia.

3. Alcune caratteristiche delle indagini campionarie su internet

L'analisi della problematica in oggetto sembra rappresentabile per successive "differenze": rispetto a un fenomeno da osservare, infatti, non tutta la popolazione di riferimento potrebbe essere attiva su internet (*internauti*); degli utenti internet che appartengano (anche) alla popolazione di riferimento, non tutti potrebbero essere *identificabili / contattabili* dal ricercatore ai fini del campionamento; anche tra quelli contattati, non tutti potrebbero decidere di rispondere (*rispondenti*); e tra questi, infine, non tutti potrebbero rispondere con sincerità (*autentici*) o con cura (*affidabili*), in ragione della distanza spazio-temporale "fisica" dell'indagine e del senso di "irresponsabilità" socialmente diffuso in rete. Si pensi, per esempio, a quanti utenti internet si avvalgano di servizi gratuiti (e-mail, video sharing, freeware, ecc.) alla cui registrazione, obbligatoriamente richiesta dal provider a fini prevalentemente commerciali, abbiano fornito dati completamente inattendibili o persino inventati, pur di superare l'ostacolo inevitabile della registrazione e usufruire del servizio internet. Questi ostacoli permettono di costruire un vero e proprio "imbuto" (*funnel*), che associamo graficamente a tale problematica, come rappresentato in Figura 1.



(Figura 1) Errori campionari e non campionari nell'impiego di campioni online

Dalla rappresentazione grafica si evince facilmente che le ricerche di mercato su internet, riferite a popolazioni online oppure offline, presentano ovviamente identiche problematiche per quanto riguarda l'errore non campionario (ossia errori dell'intervistatore, dell'intervistato, del ricercatore, ecc.), in cui bisogna gestire i soggetti rispondenti, autentici e affidabili, ma soltanto "a valle della base campionaria" (ossia l'insieme dei soggetti identificabili /

contattabili su internet). In precedenza, invece, si pone quell'evidente questione di rappresentatività, che è alla base del ragionamento sviluppato nel presente lavoro, ossia la possibile (pressoché certa) discrepanza tra la popolazione offline e il campione online, circostanza che ovviamente contribuisce a rendere più complesso l'errore campionario (graficamente rappresentato con un tratteggio più intenso).

Bisogna infatti specificare che l'errore campionario in senso stretto è in linea di principio identico per entrambi i campioni (e questa osservazione costituirà proprio una delle basi per la soluzione metodologica di seguito proposta), perché deriva dall'applicazione del teorema del limite centrale, secondo il quale all'aumentare della numerosità campionaria aumenta la rappresentabilità del fenomeno (o, più precisamente, all'aumentare della numerosità campionaria la distribuzione campionaria del fenomeno tende alla distribuzione normale standardizzata, consentendo stime attendibili dopo aver definito errore e rappresentatività) [Natale, 2004]. Tuttavia, è chiaro che un campione online è "fisiologicamente" ereditario di una popolazione online, mentre nel caso di un campione online impiegato per rappresentare una popolazione offline tale ereditarietà è inevitabilmente spuria.

Poiché proprio queste sembrano le fondamentali problematiche riguardanti le ricerche online, a nostro avviso non soltanto non è possibile arrivare a una tassonomia esaustiva dei dispositivi impiegabili nelle ricerche di mercato *internet-based*, perché gli strumenti internet sono in continua evoluzione (tecnologica e sociale) e soprattutto è in continua evoluzione la loro combinazione, ma, dal punto di vista del marketing, tale sforzo potrebbe essere tutto sommato nemmeno dirimente. Piuttosto, si può tentare una rappresentazione della problematica campionaria e della relativa soluzione seguendo uno schema concettuale, cominciando dall'impostazione logico-metodologica della ricerca di mercato e proseguendo con una visione sinottica delle strumentazioni e, soprattutto, delle tecniche disponibili e/o attuabili.

In tal senso, nel voler seguire rigorosamente l'impostazione classica delle ricerche di mercato, ferma restando la necessaria elasticità rispetto alle diverse situazioni professionali, è doveroso attingere in primo luogo a una solida base concettuale su cui innescare successivi approfondimenti, legati alla natura dell'ambiente internet. Pertanto, un approccio corretto (o, almeno, metodologicamente non scorretto) alle ricerche di mercato *internet-based* è proprio identificabile nel tradizionale processo della ricerca di mercato, articolabile in otto fasi [Barile e Metallo, 2002]:

- a) definizione del problema di marketing;
- b) identificazione della finalità della ricerca;
- c) formulazione degli obiettivi dell'indagine;
- d) progettazione dell'intervento (in termini di qualità, tempi e costi);
- e) valutazione economica delle azioni da intraprendere;
- f) raccolta dei dati (primari e/o secondari, da fonti interne e/o esterne);
- g) elaborazione e analisi dei dati (con eventuali proposte di applicazione);
- h) relazione finale (corredata, come si vedrà successivamente, delle "credenziali").

Si ribadisce come questa impostazione non sia chiaramente obbligatoria, ma è ovviamente facile apprenderne i vantaggi in termini di schematicità. Proseguendo nella strutturazione, la fase più delicata del processo, ossia la "progettazione dell'intervento", può essere a sua volta articolata in ulteriori otto sotto-fasi, come può evincersi dal seguente percorso [ibidem]:

- 1) formulazione degli obiettivi operativi della ricerca;
- 2) progettazione dell'indagine in termini concreti e fattuali;
- 3) progettazione del campione;
- 4) selezione e strutturazione degli strumenti di rilevazione;
- 5) gestione delle risorse umane impegnate nell'indagine;
- 6) programmazione temporale delle attività da svolgere;
- 7) pianificazione delle tabulazioni e delle codificazioni dei risultati;
- 8) stima dei costi operativi.

Si evidenzia come questa sotto-articolazione sia perfettamente inserita nel processo logico-metodologico della ricerca di mercato. Si noti, infatti, come la prima (*formulazione degli obiettivi operativi della ricerca*) e l'ultima sotto-fase (*stima dei costi operativi*) della "progettazione dell'intervento" siano fisiologicamente collegate alla fase precedente ("formulazione degli obiettivi dell'indagine") e a quella successiva ("valutazione economica delle azioni da intraprendere").

Anche soltanto con una rapida scorsa delle fasi e sotto-fasi così esposte, risulta ben chiara la presenza di diverse "zone" concettuali caratterizzate da identica problematicità, sia nelle ricerche *earth-based* sia in quelle *internet-based* (si pensi, per esempio, all'analisi del problema di marketing). Altre "zone", invece, presentano situazioni da contestualizzare in termini di strumentazioni (si pensi, per esempio, al questionario online e non più cartaceo), ma capaci di realizzare i vantaggi più evidenti collegati alle ricerche di mercato online. Esiste, infine, una "zona" caratterizzata da specifiche problematiche attinenti le ricerche di mercato *internet-based*, tra le quali la più rilevante è senza dubbio la rappresentatività, come già si anticipava in precedenza e come si è in questo modo verificato anche metodologicamente.

Osservazione (grazie al *funnel*) e metodologia (grazie alla *literature review*) permettono quindi di determinare come, in definitiva, la principale problematica relativa alle ricerche *internet-based* risieda nell'ancora limitata rappresentatività dell'indagine. Per cercare di superare tale limite, nella visione proposta in questo lavoro si ritiene indispensabile un supporto da parte del marketing alla statistica inferenziale, contrariamente a quanto avviene tradizionalmente nelle ricerche di mercato, in cui il contributo statistico è funzionale all'analisi ed eventualmente alla soluzione del problema di marketing [Molteni e Troilo, 2003].

Infatti, la matematica e la statistica, che costituiscono le discipline maggiormente contribuenti allo studio della rappresentatività del campione, non potranno che correttamente discendere da regole GIGO ("garbage in, garbage out") [Seglin, 1994] e, in un certo senso, disinteressarsi della qualità del dato di partenza: proprio su tale questione, invece, dovrà soffermarsi il ricercatore, il quale ben intuisce le potenzialità dell'internet e si adopera per il superamento dei contestuali limiti. Al momento, infatti, la problematica in oggetto si riferisce alla rappresentatività del campione online rispetto a fenomeni (anche) offline, mentre è evidente che un'indagine su un fenomeno online potrebbe avere seri problemi di rappresentatività se svolta offline, ossia presso intervistati non normalmente fruitori dei servizi internet oggetto dell'indagine.

Paradossalmente, pertanto, si dimostra che le ricerche online, pur patendo profondamente il "funnel", non possono essere considerate inattendibili a priori, perché, per esempio, è epistemologicamente sbagliato indagare mediante tecniche tradizionali un utente, anche soltanto in sede di *concept test*, in merito all'esperienza di navigazione in un *virtual store* multimediale, magari anche tridimensionale. Alcune proposte in tal senso sono sviluppate nella parte finale del paper, in cui si ragiona sulla metodologia di raccolta, all'interno di *internet sample* concepiti come comunità virtuali [Hagel e Armstrong, 1997], del dato "autentico" e "affidabile" rispetto al problema di marketing in questione.

4. Riflessioni sulle possibili soluzioni al problema della rappresentatività online

Si proverà, di seguito, a ragionare sulle principali caratteristiche del campione adeguatamente rappresentativo, esponendo genericamente alcune motivazioni statistiche e successivamente soffermandosi su approfondimenti gestionali, che, ovviamente, più sono in sintonia con lo spirito di questo lavoro.

In generale, la rappresentatività del campione (n) rispetto alla popolazione (N) dovrebbe essere conseguita, supposta la correttezza nel calcolo della numerosità campionaria, tramite il campionamento probabilistico, in primo luogo tramite il campionamento casuale semplice [Molteni e Troilo, 2003]: infatti, soltanto un campione casuale semplice (ossia con probabilità di campionamento nota e uguale per tutti i membri della base campionaria) garantisce l'assenza dell'errore di selezione (che certamente figura, invece, nel campionamento non probabilistico, ossia per convenienza, giudizio e quota). Proprio l'azzeramento dell'errore di selezione, ricavabile dal campionamento casuale semplice, dovrebbe assicurare, in concomitanza come si diceva di un'adeguata numerosità, anche assenza di distorsione.

Nel campionamento non probabilistico, peraltro, il ricercatore non è ignaro della possibile distorsione, ma anzi la usa a proprio vantaggio, ritenendo che quel consapevole errore di selezione gli consenta una rappresentatività contestualmente meno infelice, anche se, ovviamente, non può saperlo *ex ante*. Dall'estrazione casuale semplice, infatti, potrebbe venir fuori proprio il campione non probabilistico, che sarebbe più rappresentativo, perché casuale, rispetto allo stesso (identico) campione non probabilistico: in tale possibilità risiede il paradosso del campionamento casuale semplice, che trova la propria forza non nel risultato, ma nel procedimento: del resto, anche un campione casuale (forse soprattutto un campione casuale) può produrre un valore estremo.

Il campionamento, come si anticipava in precedenza, si basa sul teorema del limite centrale, secondo il quale l'aumento della numerosità campionaria rende la distribuzione campionaria sempre più simile a una distribuzione normale standardizzata, consentendo pertanto di calcolare la numerosità campionaria necessaria per sopportare un dato margine di errore (E_{s_x}) e un dato intervallo di confidenza (z). Infatti, possiamo ricorrere alle seguenti formule (cfr. Tabella 1) per il calcolo della numerosità campionaria (è infinita, per convenzione statistica, la popolazione superiore a 100.000 elementi; è binomiale il fenomeno rappresentabile in due sole modalità, mentre diversamente è continuo) [Barile e Metallo, 2002].

	Popolazione finita	Popolazione infinita
Fenomeno continuo	$n = N z^2 s^2 / [(N - 1) E^2 s_x + z^2 s^2]$	$n = (z s / E_{s_x})^2$
Fenomeno binomiale	$n = N z^2 p (1 - p) / [(N - 1) E^2 s_x + z^2 p (1 - p)]$	$n = p (1 - p) (z / E_{s_x})^2$

(Tabella 1) Calcolo della numerosità campionaria (n) rispetto alla numerosità della popolazione (N)

È del tutto evidente, pertanto, che per la precisione della stima non conta tanto la frazione di campionamento ($f = n / N$) quanto la numerosità campionaria, proprio in forza del teorema del limite centrale. Queste considerazioni, essenziali nella teoria dei campioni, permettono di affermare che la questione della rappresentatività su internet, per certi versi, ispirati più alla pratica dell'indagine che alla scientificità della metodologia, possa trascurare un primo stadio del *funnel*, ossia la non coincidenza della popolazione osservata con l'utenza internet, purché si riesca a costruire un campione online sufficientemente numeroso e ovviamente aggiornato rispetto alla mortalità statistica.

In pratica, il problema più rilevante della rappresentatività online si riferisce a quel numero di elementi della base campionaria che concorra a formare l'adeguata numerosità campionaria calcolata, risultando una caratteristica aggiuntiva e non penalizzante a priori il fatto che tali elementi della popolazione offline siano anche utenti attivi online. Invece, reclutare esclusivamente utenti attivi online per un campione di una popolazione offline potrebbe per certi versi essere assimilato a un campionamento non probabilistico [Andreani e Conchon, 2001a], che avvenga per giudizio o per convenienza (eventualmente, anche per quota: si pensi, per esempio, a un'indagine sulla soddisfazione dei clienti di una banca che sia attiva con filiali tradizionali e con servizi internet).

Nel campionamento non probabilistico, infatti, non ha alcun valore il concetto di errore standard, alla base, assieme alla confidenza, del calcolo della numerosità campionaria. A servirsi del campionamento non probabilistico sono prevalentemente le ricerche di mercato qualitative, che, anche per questo motivo, sembrano essere le indagini meglio praticabili tramite ricerche di mercato online, non avendo per natura pretesa di descrizione, ma soltanto di esplorazione.

Si è ragionato, pertanto, del come la rappresentatività del campione online non debba essere collegata tanto alla grandezza dell'utenza internet rispetto alla popolazione, quanto piuttosto alla numerosità dell'utenza internet indagata. In tal senso, di conseguenza, potrebbe non essere necessario estrarre un campione online da una popolazione online, corrispondente in tutto o in parte alla popolazione offline, perché potrebbe bastare il positivo riscontro delle caratteristiche della popolazione presso il campione, purché adeguatamente numeroso. Si seguirebbe, in tal modo, un procedimento non di estrazione (dall'alto verso il basso), ma di astrazione (dal basso verso l'alto).

È fondamentale, quindi, mitigare le impurità collegate all'indagine con la trasparenza delle relative credenziali: caratteristiche della popolazione, tipologia di campionamento, numerosità campionaria, percentuale di errore ammissibile e confidenza della stima. A queste indicazioni statistiche bisognerebbe correttamente aggiungere tutte le altre informazioni ritenute di una qualche influenza nell'estrapolazione del risultato (una sorta di nota integrativa alla ricerca): considerazioni sulla formazione degli intervistatori, sulla formulazione delle domande, sulla catalogazione delle risposte e così via, facendo anche attenzione alla corretta implementazione di sistemi per l'eliminazione o almeno la mitigazione dell'errore non campionario online, con particolare riferimento agli stadi del "funnel" relativi a *rispondenti, autentici e affidabili*.

Tali credenziali rappresentano non soltanto la "cartina di tornasole" per l'interpretazione dei risultati dell'indagine, ma anche e soprattutto la corretta chiave di lettura per la valutazione delle informazioni ottenute, diventando le stesse credenziali, pertanto, informazioni per il processo decisionale. Infatti, la naturale distorsione statistica del campionamento, dovuta alla fisiologica distanza tra universo e campione, non può essere messa in discussione: come vale per le ricerche *earth-based*, vale certamente anche per quelle *internet-based* (la natura socio-economica è perennemente cambiante e anche gli utenti internet, ovviamente, sono persone).

Pertanto, la problematica relativa alla rappresentatività della popolazione internet rispetto ai fenomeni offline, che costituisce il principale *focus* di questo studio, sembra poter essere risolta facendo tesoro delle riflessioni logico-metodologiche espresse in precedenza (anche se ovviamente con l'adeguata considerazione delle caratteristiche volta per volta contestuali del fenomeno). Tramite una complessiva strategia di internet marketing analitico, finalizzata a costruire, monitorare e gestire gli *internet sample* come se fossero comunità virtuali, diventa inoltre possibile ragionare sui meccanismi utili a garantire, a valle, la "pulizia" del dato in ingresso, risolvendo in questo modo gli ulteriori limiti rappresentati dal *funnel*, con una possibile proposta operativa che viene di seguito analizzata.

5. Alcune proposte operative per la gestione di un internet sample

In base al ragionamento esposto nel presente lavoro, sembra possibile delineare alcune fondamentali caratteristiche, soprattutto metodologiche (*strategiche*) e applicative (*operative*), non soltanto per la corretta progettazione, ma anche per la corretta gestione di un *internet sample* per le ricerche online. In tal senso, sembra particolarmente interessante pensare a una sorta di comunità virtuale, da costruire con abnegazione e attenzione, spostando la prospettiva di riferimento dalle ricerche di mercato in senso stretto a una più complessiva strategia di internet marketing, basata (anche quando analitico) sul binomio visibilità e attrattività [Metallo e Festa, 2003].

In altre parole, sembra paradossalmente non dirimente porsi problemi eccessivi sulla validità e sull'attendibilità delle ricerche di mercato online in quanto tali, perché la principale problematica di riferimento sembra riguardare invece la numerosità della popolazione online, da cui estrapolare la base campionaria e successivamente il campione per l'indagine. Quest'ultimo, se adeguatamente numeroso, potrebbe essere sufficiente in termini di rappresentatività almeno da un punto di vista empirico: da un punto di vista metodologico, invece, tale possibilità potrebbe anche non essere mai praticabile, non essendo sostanzialmente possibile procedere su internet a un campionamento casuale semplice che sia certamente rappresentativo di una popolazione offline [Schmitz, 2004].

In questa direzione, il *funnel* deve essere continuamente aggiustato e mitigato, nel senso che bisogna insistere costantemente non soltanto sulla progettazione statistica del campione, ma anche e soprattutto sulla complessiva correttezza operativa nella gestione del campione, attivando una serie di meccanismi di continua autenticazione dell'intervistato, orientata a garantire identità e sincerità nella rilevazione.

Questo continuo imbrigliamento, al contempo, potrebbe essere causa di eccezioni sulla complessiva spontaneità e naturalezza dell'intervistato nonché sulla sua predisposizione all'intervista (la buona lena nel rispondere, per dire, potrebbe essere fiaccata da continue autenticazioni), ma appare per le motivazioni finora esposte una delle poche strade al momento perseguibili per incrementare la complessiva attendibilità dell'indagine online. I passaggi di autenticazione dovrebbero essere molteplici e potrebbero essere schematizzati come segue:

- *iscrizione*: l'utente deve in primo luogo iscriversi alla comunità virtuale (il ricercatore, quindi, deve fin dall'inizio dell'indagine immaginare il campione online di una qualsiasi ricerca online come una comunità virtuale, potendone ricavare anche un senso di maggior "appartenenza" dell'intervistato al buon esito della ricerca), compilando un *form* di registrazione, in base al quale il ricercatore dovrebbe desumere una prima serie di dati per procedere alla classificazione dell'utente in base ai principali criteri di segmentazione dell'intera comunità. Gli strati in questione non sono statici e nemmeno dinamici: sono semplicemente "virtuali" (allo stesso modo in cui, in un database, le *query* possono essere considerate delle tabelle virtuali), potendo sfruttare la capacità del sistema informatico sottostante alla comunità virtuale di ricavare, grazie a opportune operazioni di filtraggio, la segmentazione più utile in quel preciso momento (tramite comuni applicazioni OLAP). A tale registrazione il sistema assegna, su scelta dell'utente e conformemente alle politiche della comunità, un *account*, costituito da *username* e *password*, che d'ora in avanti saranno la chiave di accesso al sistema, con relativa autenticazione, autorizzazione e responsabilità;

- *accreditamento*: in sede di svolgimento della singola rilevazione, all'utente dovrebbe essere richiesta non soltanto l'autenticazione (tramite *account*), ma anche un successivo livello di autorizzazione, identificabile p.e. in un *captcha* (sulla pagina web di accesso alla rilevazione) e/o in una *password* (nell'e-mail di "convocazione all'indagine" spedita all'utente). *Captcha* è l'acronimo dell'espressione "completely automated public Turing test to tell computers and humans apart" ("test di Turing pubblico e completamente automatico per distinguere tra macchine e umani"): per confermare l'uso del servizio internet da parte di un individuo e non di una macchina, viene proposta a video un'immagine con dei caratteri, che l'utente è chiamato a digitare in un'apposita casella di testo per la verifica. I caratteri nell'immagine sono artatamente distorti nell'apparenza, ma perfettamente comprensibili a occhio nudo e, pertanto, sono interpretabili da un individuo e non da una macchina (aumentando, si spera, il livello di attenzione dell'intervistato). In questo modo, il sistema si garantisce rispetto alla possibilità che l'utente (o lo stesso sistema informatico sottostante alla comunità virtuale) sia inconsapevolmente infettato da un *malware*, che potrebbe quindi accedere alla comunità virtuale, ma non alla rilevazione, in cui potrebbe generare risposte inutili;

- *verifica*: infine, per avere l'ulteriore certezza che a essere oggetto della rilevazione sia proprio l'utente, diventa inevitabile costringere quest'ultimo a fornire adeguate risposte a determinate domande, alle quali egli soltanto può rispondere. Tali domande dovrebbero essere presentate durante la rilevazione in modalità *random*, così da "costringere" quello specifico utente a essere "fisicamente" presente nell'interazione con il sistema. È da evidenziare che l'autenticazione degli utenti rappresenta un elemento centrale delle infrastrutture di sicurezza informatica e le tecniche basate sulla conoscenza sono attualmente il metodo usato più frequentemente per l'autenticazione dell'utente [Ellison et al., 2000; Dhamija e Perrig, 2000]. Nello specifico, attraverso la *knowledge based authentication* (KBA) l'utente, per poter essere riconosciuto come tale, deve rispondere correttamente a una domanda (standard) alla quale ha già risposto in precedenza e della quale, ovviamente, il sistema informatico conserva la memoria della relativa risposta (generando, probabilmente, anche una maggior empatia con l'utente, il quale dovrebbe / potrebbe vedersi non come un "numero" della batteria degli intervistati, ma come un soggetto al quale è richiesto di interagire in maniera univocamente individuale). Tale tecnica è ampiamente usata per ripristinare accessi a servizi internet nel caso in cui l'utente abbia dimenticato la password: nel caso di risposta corretta alla domanda standard, il sistema invia un'e-mail all'indirizzo precedentemente lasciato dall'utente, ricordandogli username e password (o generandone una nuova). Alcune domande, infine, potrebbero essere concepite anche come vere e proprie "domande di controllo", allo scopo di recuperare la maggiore affidabilità possibile in merito alle risposte dell'intervistato [De Luca, 2006].

Come si accennava in precedenza, questo sistema di continua autenticazione potrebbe rallentare la rilevazione e renderla persino scomoda per l'utente, ma al momento queste sembrano le soluzioni gestionali più incisive nell'efficacia e diffuse nella pratica. Tali riflessioni, naturalmente, riguardano anche altri meccanismi di accesso, basati non su quello che l'utente "sa" (come analizzato finora), ma su quello che l'utente "ha" (per esempio, una smart card, un token, ecc.) e/o su quello che l'utente "è" (scansione biometrica) [Teti e Festa, 2009]: anche in questi casi, tuttavia, nulla vieta di sospettare che l'utente possa accedere al sistema una prima volta con le proprie credenziali, per lasciare svolgere a qualcun altro la rilevazione in un momento successivo.

Con la continua tecnica dell'autenticazione (*iscrizione-accreditamento-verifica*) si fa quindi riferimento a una pratica di costante verifica della correttezza comportamentale dell'utente (eliminando o mitigando le problematiche relative agli stadi relativi a *rispondenti, autentici e affidabili* del "funnel"); come in tutte le attività di controllo, si finirà per essere meno produttivi da un punto di vista operativo, assicurandosi però una migliore qualità dell'indagine. Potrebbe anche accadere, a dire il vero, che parte degli utenti selezionati abbandoni l'indagine o persino l'*internet sample* gestito come comunità virtuale: a questo punto, bisognerebbe chiedersi a quale compromesso si possa arrivare tra quantità e qualità del campione. Anche per tale motivo sembrano maggiormente adeguate all'ambiente internet le ricerche qualitative, meno sensibili alla problematica del campionamento e alla stessa numerosità campionaria, perché principalmente finalizzate, come già ampiamente evidenziato, a meccanismi di esplorazione e non di descrizione.

6. Risultati, implicazioni e conclusioni

L'evoluzione dell'informatica e, in particolare, l'introduzione, la diffusione e lo sviluppo dell'internet hanno avuto, continuano ad avere e avranno sempre di più in futuro uno straordinario impatto sui quotidiani processi economici e sociali. Queste innovazioni hanno naturalmente interessato anche il mondo dei sistemi informativi aziendali (anzi, soprattutto questo mondo), provocando radicali trasformazioni nella progettazione, nel funzionamento e nello sviluppo delle procedure fisiche (processi organizzativi) e logiche (flussi informativi) chiamate a svolgere specifiche attività d'impresa.

Una delle discipline maggiormente coinvolte in tale cambiamento è quella delle ricerche di mercato, che si trova a poter usare internet come strumento aggiuntivo rispetto agli "attrezzi" tradizionali (ricerche "con" internet) o persino come bacino aggiuntivo / sostitutivo / innovativo di soggetti da sottoporre a indagine (ricerche "su" internet). Tra queste due possibilità, oltre all'inevitabile esigenza di procedere a un'adeguata contestualizzazione delle nuove strumentazioni informatiche, le maggiori problematiche di studio derivano dalle ricerche "su" internet, soprattutto in considerazione della distanza (sociale, culturale, economica, ecc.) che potrebbe esserci (che, anzi, costantemente esiste) tra popolazione offline e popolazione online, con ovvie ricadute sull'apprezzamento della rappresentatività del campione online rispetto alla popolazione offline.

In questo paper, chiaramente orientato a uno studio di natura concettuale, si è pertanto proceduto in primo luogo a una rappresentazione logica delle principali problematiche di campionamento nelle ricerche online, arrivando a sviluppare un *framework* teorico / pratico ("funnel") che riesca a tenere complessivamente conto degli errori, campionari e non campionari, che potrebbero metodologicamente emergere nel collegamento di rappresentatività (da una parte) tra popolazione online e campione online e (dall'altra) tra popolazione offline e campione online. Grazie a tale schematizzazione, ricorrendo a continue riflessioni in contaminazione tra marketing, statistica e informatica, è stato possibile proporre alcune soluzioni empiriche per la gestione della rappresentatività online, più di statistica "sostanziale" nel caso dell'errore campionario e più di marketing "applicato" nel caso dell'errore non campionario.

Tali soluzioni, in ogni caso, sono utilmente proponibili soltanto se perseguite nella prospettiva dell'*internet sample* concepito come comunità virtuale, in cui i soggetti sottoposti alla rilevazione siano considerati dal ricercatore (e soprattutto da sé stessi) membri di uno specifico gruppo attivo online, anche se soltanto per la durata dell'indagine. Diventa fondamentale, pertanto, immaginare una strategia di internet marketing analitico per alimentare e mantenere il campione / comunità, facendo ricorso alla cangiante combinazione tra visibilità e attrattività.

Le implicazioni scientifiche del presente lavoro sembrano riguardare soprattutto le possibili evoluzioni dell'*humus* alla base della disciplina delle ricerche di mercato, ossia la tradizionale combinazione tra marketing, statistica e (ormai) informatica, individuando quindi in tale mix di competenze un probabile oggetto di ulteriori studi, sia in termini metodologici che applicativi. In ragione della pervasività dell'internet, infatti, gli approfondimenti sulle ricerche di mercato dovranno sempre di più confrontarsi con un'evoluzione sociale della Rete che già espone (e sempre di più esporrà in futuro) alla necessità / opportunità di trovare soluzioni ragionevoli (per l'oggi) e/o lungimiranti (per il futuro), come nello spirito delle riflessioni e delle proposte sviluppate in questo studio, che sembrano pertanto fornire interessanti spunti applicativi soprattutto in termini di implicazioni operative (e quindi per chi sia chiamato da subito a svolgere ricerche online).

Non bisogna infatti dimenticare che il mondo delle ricerche di mercato online è ancora ai suoi primordi (come lo sono in generale tutte le ambientazioni *internet-based*) per quanto riguarda i valori, i principi e soprattutto le regole di funzionamento (si pensi alla netiquette, alla privacy, ai social network, ecc.): tale caratteristica, assieme alla vertiginosa velocità di sviluppo delle strumentazioni internet e delle relative possibilità d'uso, che comprime esponenzialmente la durata dell'*internet year*, rende già obsoleta o almeno rischiosa qualsiasi soluzione, soprattutto se empirica (ed è questo un probabile limite della ricerca concettuale che si è presentata). Per tale motivo, è sempre fondamentale recuperare, contestualizzare e sviluppare un rigoroso impianto logico-metodologico, com'è avvenuto in questo studio, che ha attinto dalla "tradizionale" metodologia delle ricerche di mercato, successivamente proponendo, prima in merito alla prospettiva dell'internet e successivamente in merito alla prospettiva del campionamento su internet, soluzioni sostanzialmente gestionali (se si vuole, di "buon senso", come normalmente avviene nell'economia d'impresa), non limitandosi alla sola analisi dell'utilità del singolo strumento statistico o informatico, ma mescolando continuamente osservazioni e rappresentando costantemente applicazioni rispondenti a una complessiva strategia di internet marketing analitico.

Bibliografia

- Andreani J.C., Conchon F., Gli studi qualitativi in Internet, *Micro & Macro Marketing*, 1, 2001a, 65-74.
- Andreani J.C., Conchon F., *Les études qualitatives en marketing*, Paris: ESCP-EAP, Les Cahiers de Recherche, 2001b, 01-150.
- Andreani J.C., L'interview qualitative en marketing, *Revue Française du Marketing*, 1988, 3-4.
- Andrews D., Nonnecke B., Preece J., Electronic survey methodology: A case study in reaching hard-to-involve Internet users, *International Journal of Human-Computer Interaction*, 16, 2, 2003, 185-210.
- Barile S., Metallo G., *Le ricerche di mercato. Aspetti metodologici e applicativi*, Giappichelli, Torino, 2002.
- Bassi F., *Analisi di mercato. Strumenti statistici per le decisioni di marketing*, Carocci, Roma, 2008.
- Cantone L., Calvosa P., Testa P., Tecnologie ad alta intensità connettiva, relazioni di marketing e processi di creazione di valore per i clienti nei mercati BtB, *Mercati & Competitività*, 2006, 3, 1.38.
- De Luca A., *Le ricerche di mercato. Guida pratica e metodologica*, Angeli, Milano, 2006.
- Dhamija R., Perrig A., D'ej`a Vu: A User Study Using Images for Authentication, *Proceedings of the 9th USENIX Security Symposium Denver, Colorado, USA, August 14-17, 2000*, 1-15.
- Di Fraia G., Validità e attendibilità delle ricerche on line, in Di Fraia G. (a cura di), *e-Research. Internet per la ricerca sociale e di mercato*, Laterza, Roma-Bari, 2004.
- Ellison C., Hall C., Milbert R., Schneier B., Protecting secret keys with personal entropy, *Future Generation Computer Systems*, 16, 2000, pp. 311-318.
- Festa G., Il ruolo dei nuovi strumenti informatici nel sistema informativo di marketing, *Esperienze d'impresa*, 2, 2001b, pp. 67-114.
- Festa G., *Digital Business English - Glossario ragionato di linguistica d'impresa per la new economy*, Edisud, 2001a.
- Hagel J., Armstrong A., *Net gain*, Harvard Business School Press, Boston, 1997.
- Howard P., Rainie L., Jones S., Days and nights on the Internet: The impact of a diffusing technology, *American Behavioral Scientist*, 45, 3, 2001, 383-404.
- Kiang M.Y., Raghu T.S., Huei-Min Shang K., Marketing on the Internet — who can benefit from an online marketing approach?, *Decision Support Systems*, 27, 2000, 383-393.
- Kotler P., *Marketing Management*, Pearson Education Italia, Milano, 2001.
- Marbach G., *Ricerche per il Marketing*, Utet, Torino, 2010.
- Mariampolski H., *Qualitative Market Research*, Sage, London, 2001.
- Martone D., Furlan R., *Online Market Research Tecniche e metodologia delle ricerche di mercato tramite Internet*, Franco Angeli, Milano, 2007.
- McDonald H., Stewart A., A comparison of online and postal data collection methods, *Marketing Intelligence & Planning*, 21, 2, 2003, 85-95.
- Melody Y. Kiang M.Y., Raghu T.S., Huei-Min Shang K., Marketing on the Internet — who can benefit from an online marketing approach?, *Decision Support Systems*, 2000, 27, 383-393.
- Metallo G., Festa G., La progettazione dei portali web nell'ottica del customer service, in Barile S., Metallo G. (a cura di) *Soluzioni problematiche d'impresa. Riflessioni e modalità risolutive*, Edizioni Culturali Internazionali, Roma, 2003.
- Molteni L., Troilo G., *Ricerche di marketing*, McGraw-Hill, Milano, 2003.
- Natale P., *Il sondaggio*, Laterza, Bari, 2004.
- Prandelli E., Verona G., *Marketing in rete, oltre Internet verso il nuovo Marketing*, McGraw Hill, Milano, 2006.
- Schmitz N., Rappresentatività e campionamenti nelle survey on line, in Di Fraia G. (a cura di), *e-Research. Internet per la ricerca sociale e di mercato*, Laterza, Roma-Bari, 2004.
- Seglin J.L., *Marketing (la guida al)*, McGraw-Hill, Milano, 1994, pag. 37.
- Teti A., Festa G., *Sistemi informativi per la sanità*, Apogeo, Milano, 2009.
- Tissier-Desbordes E., *Les études qualitatives dans un monde post-moderne*, *Revue Française du Marketing*, 1998, 3-4.
- Valdani E., Busacca B., Customer Based View: dai principi alle azioni, *Proc. de Convegno Le tendenze del Marketing in Europa*, Università Ca' Foscari Venezia, 2000, 1-24.
- Vescovi T., *Il Marketing e la rete. La gestione integrata del web nel business. Comunicazione, e-commerce, sales management, business to business*. Il Sole 24Ore, Milano 2007.

Weible R., Wallace J., Cyber research: the impact of the Internet on data collection, *Market Research*, 1998, 10, 3, 19-31.

Wright K.B., Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services, *Journal of Computer-Mediated Communication*, 10, 3, 2005.

Document Image Retrieval by Layout Analysis

G. Pirlo⁽¹⁾, M. Chimienti⁽²⁾, M. Dassisti⁽³⁾, D. Impedovo⁽⁴⁾, A. Galiano⁽⁴⁾

⁽¹⁾ Dipartimento di Informatica, Università degli Studi di Bari "A. Moro", via Orabona 4,
70125-Bari, Italy

⁽²⁾ Laboratorio Kad3, C.da Baione, 70043 Monopoli (BA), Italy

⁽³⁾ Dip. Meccanica, Management e Matematica, Politecnico di Bari, viale Japigia 182,
70126 - Bari, Italy

⁽⁴⁾ Dyrecta Lab, Via V. Semplicio 45, 70014 Conversano (BA) - Italy

(corresponding author: giuseppe.pirlo@uniba.it)

A new layout-based document image retrieval system is presented in this paper. The system is specifically designed for commercial form retrieval and uses mathematical morphology to extract structural components from the document image. Document layout description is performed by the Radon Transform whereas Dynamic Time Warping is used for matching. The experimental results have been carried out on both real and simulated data sets. They demonstrate the effectiveness of the proposed approach and their robustness with respect to different classes of commercial forms and shifted/rotated document images.

Index Terms — Document management, Document Image Retrieval, Mathematic Morphology, Radon Transform, Dynamic Time Warping.

1. Introduction

The increasing number of documents available in databases and digital libraries makes document retrieval a critical task of current document management systems. Traditional document retrieval systems – based on set-theoretic, algebraic and probabilistic models - require a document to be present in text form and the querying method is based on a specific textual content in the document [Doermann, 1998; Manning et al., 2009]. Whatever the model used, text-based document retrieval systems require a document in text form, since the search for similar documents is based on comparing the textual contents. As a consequence, a preliminary stage of image to text conversion by an Optical Character Recognizer (OCR) is required when a document is in image form. OCR is a time-consuming error-prone process, specifically in the

case of multi-lingual/multi-font documents and poor-quality document images [Marukawa et al., 1997; Taghva et al., 1996; Lorpesti, 1996], as discussed in comprehensive surveys on this topic [Doermann, 1998; Mitra and Chaudhuri, 2000].

Along with the spreading of multimedia documents, it is useful to search a document on the basis of its structure and not only on the basis of its textual content. In such cases, methods adopted for document retrieval use feature vectors in which each feature is extracted from a specific region of the document image. For instance, some researchers used a static zoning strategy for document image decomposition to extract a fixed-size feature vector from the document image. In this approach, a regular grid is superimposed to the document image in order to extract regional characteristics [Tzacheva et al., 2002]. In another approach, a hierarchical zoning strategy was proposed to overcome the problem of optimal grid selection, in order to face with the treatment of set of documents of different characteristics [Duygulu and Atalay, 2002]. A system that extracts text lines and describes the layout by means of relationships between pairs of these lines was also discussed in the literature [Huang et al., 2005], whereas some researchers used Brick Wall Coding Features (BWC) features to represent bounding boxes of the words [Erol et al., 2008]. Although the features are scale invariant and robust to slight perspective distortion, the accuracy of their system is very low. In addition the method does not work correctly when documents are written in languages such as Japanese and Chinese, in which words are not separated. Several approaches can also be combined to identify a document, such as barcode, micro optical patterns, encoding hidden information, paper fingerprint, character recognition, local features and RFID. Owing to utilize SIFT. Unfortunately, the retrieval process is time consuming and requires special equipment [Liu and Liao, 2011].

In this paper a Layout-based Document Image Retrieval (LDR) system is presented, that is specifically devoted to commercial form processing, such as invoices, waybills, receipts, etc., in which layout is strongly characterized by a grid-structure. In fact, in these particular cases, traditional document-image approaches are not effective since they are not able to describe documents on the basis of the grid-based structure. In the first step the system uses a technique based on mathematical morphology for removing textual components from the document image and for extracting the grid-based structure in the document layout. Subsequently Radon Transform is used to obtain the feature vector characterizing the specific grid structure of the document. Dynamic Time Warping (DTW) is finally adopted to perform document matching.

The paper is organized as follow. The architecture of the system is presented in Section 2. Section 3 describes the preprocessing phase, which uses operators of mathematical morphology. The feature extraction phase is presented in Section 4. In Section 5 the matching process based on DTW is discussed while the decision combination process is illustrated in Section 6. The experimental results are reported and discussed in Section 7. Section 8

presents the conclusion of the work and highlights some directions for further research.

2. The Radon Transform for Layout-based Document Image Retrieval

The LDR system presented in this paper is based on three main phases: Acquisition and Preprocessing; Feature Extraction; Matching. After document image acquisition, the document is preprocessed and transformed by Radon Transform. The features extracted are then stored in the reference database in the enrollment stage. In the running stage, an unknown document is first scanned and preprocessed, successively the features are extracted compared to the those stored into the database. The matching module performs matching by Dynamic Time Warping (DTW) and outputs the ranked list of similar documents. More precisely, the input document is acquired as a standard 256 gray-level – 100dpi PDF file. Figures 1 shows an input document concerning a real invoice.

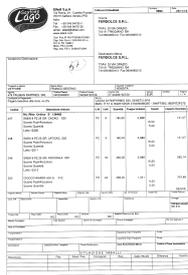


Figure 1. Input document image $I=l(x,y)$.

Successively, after noise removal, document is resampled to 100 dpi and grid-based structure is extracted by mathematical morphology [Serra, 1982]. More precisely, let $I=l(x,y)$ be the document image ($1 \leq x \leq X$, $1 \leq y \leq Y$) and let be

- B_{hor} the horizontal structure element defined as (see Figure 2a):
 $B=\{(-s,0), \dots, (-1,0), (0,0), (1,0), \dots, (s,0)\};$
- B_{ver} the horizontal structure element defined as (see Figure 2b):
 $B=\{(0,-s), \dots, (0,-1), (0,0), (0,1), \dots, (0,s)\};$

being s a small positive integer which determine the size of the structure element.

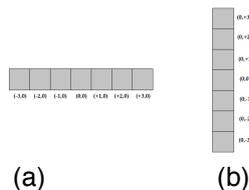


Figure 2. Structure elements ($s=3$): (a) B_{hor} , (b) B_{ver} .

In the preprocessing phase from the image $I(x,y)$ two filtered images $I_{hor}=I_{hor}(x,y)$ and $I_{ver}=I_{ver}(x,y)$, which contains respectively horizontal and vertical segments, are obtained by a closure operator as follows (see Figure 3):

$$I_{hor} = I \bullet B_{hor} = (I \oplus B_{hor}) \ominus B_{hor} \quad (1a)$$

$$I_{ver} = I \bullet B_{ver} = (I \oplus B_{ver}) \ominus B_{ver} \quad (1b)$$

being “ \bullet ” the closure operator, while “ \oplus ” and “ \ominus ” indicate respectively Minkowski sum and difference.

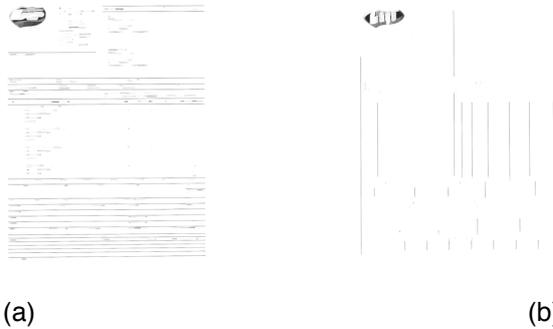


Figure 3. Example of filtered images: (a) I_{hor} , (b) I_{ver} .

Finally, $I_{hor}(x,y)$ and $I_{ver}(x,y)$ are combined to reconstruct the preprocessed image I^* according to XOR operator:

$$I^* = I_{hor} \text{ XOR } I_{ver} \quad (2)$$

Figure 4 shows an example of document image after preprocessing.



Figure 4. The preprocessed image I^* .

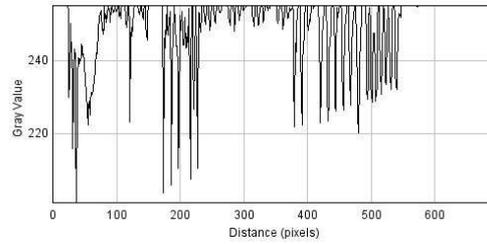
In the feature extraction step, in order to extract grid-based layout document images, the Radon Transform was considered. It is worth noting that the Radon Transform was extensively used in image analysis and has a number of important applications, like those related to MRI and computed tomography [Cormack, 1983; Deans, 1983]. The complete description of the Radon Transform is beyond the scope of this paper (see further details in [Jafari-Khouzani and Soltanian-Zadeh, 2005; Seo et al., 2004]). For the aim of this paper

we only remind that the Radon Transform computes projection sum of the image intensity along a oriented at line $(\rho - x \cos \vartheta - y \sin \vartheta) = 0$, for each ϑ and ρ . More precisely the Radon Transform of a function $I(x,y)$ in an Euclidean space is defined by [Hjoui and Kammler, 2008]:

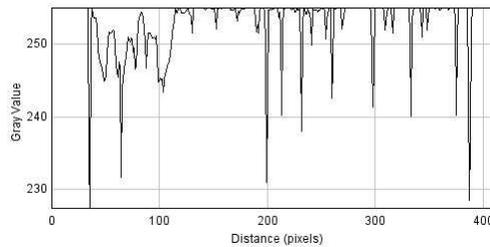
$$S_{\vartheta,\rho} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I^*(x,y) \cdot \delta(\rho - x \cos \vartheta - y \sin \vartheta) dx dy \quad (3)$$

where the $\delta(r)$ is Dirac function, which is infinite for argument zero and zero for all other arguments (it integrates to one).

Therefore, computing the Radon Transform of a two dimensional image intensity function $I(x,y)$ results in its projections across the image at arbitrary orientations ϑ and offsets ρ . Figure 5 presents the results of the Radon Transform applied to the preprocessed image I for the parameter values related to horizontal ($\vartheta=0$, $\rho=0$) and vertical ($\vartheta=\pi/2$, $\rho=0$) projections.



(a) Horizontal Projection



(b) Vertical Projection

Figure 5. Feature extraction by Radon Transform.

Dynamic Time Warping (DTW) is used for matching the feature vectors extracted by the radon transform from two document images. More precisely, let be F^r , S^t the feature vectors of M elements extracted from the document images I^r and I^t , a warping function between S^r and S^t is any sequence of couples of indexes identifying points of S^r and S^t to be joined [Salvador and Chan, 2004; Lemire, 2009]:

$$W(S^r, S^t) = c_1, c_2, \dots, c_K \quad (4)$$

where $c_k = (i_k, j_k)$ (k, i_k, j_k integers, $1 \leq k \leq K$, $1 \leq i_k \leq M$, $1 \leq j_k \leq M$). Now, if we consider a distance measure $d(c_k) = d(z^r(i_k), z^t(j_k))$ between elements of S^r and S^t , we can associate to $W(S^r, S^t)$ the dissimilarity measure

$$D_{w(S^r, S^t)} = \sum_{k=1}^K d(c_k) \quad (5)$$

The DTW detects the warping function $W^*(S^r, S^t) = c^*_1, c^*_2, \dots, c^*_{K^*}$ which satisfies the condition of [Salvador and Chan, 2004]:

• Monotonicity (i.e. $i_{k-1} \leq i_k$, $j_{k-1} \leq j_k$ for $k=2, \dots, K$) (6a)

• Continuity (i.e. $i_k - i_{k-1} \leq 1$, $j_k - j_{k-1} \leq 1$ for $k=2, \dots, K$) (6b)

• Boundary (i.e. $i_1 = 1$, $j_1 = 1$ and $i_K = M$, $j_K = M$) (6c)

and which provides the distance value between S^r and S^t defined as [Salvador and Chan, 2004; Lemire, 2009]:

$$D_{w^*(S^r, S^t)} = \min_{W(S^r, S^t)} D_{w(S^r, S^t)} \quad (7)$$

The value in eq. (7) represents the similarity between the document images I^r and I^t . Therefore, given a document image as input, the matching module will output the ranked list of the k top similar document images retrieved from the database.

The matching procedure provides two distance-based ranked lists of documents, obtained respectively from horizontal and vertical projections. The decision making process obtains the final decision combining the two ranked lists using the Borda-count strategy [Kittler et al., 1998; Xu et al., 1992]. According to this strategy, let $D = \{D_1, D_2, \dots, D_k, \dots, D_K\}$ be the set of K documents enrolled into the system for reference and D^* the unknown input document. Furthermore, let be:

- $L^h : \langle D^h_1, D^h_2, \dots, D^h_k, \dots, D^h_K \rangle$ the ranked list of documents obtained from the match of the horizontal projection ($D^h_k \in D$, for $k=1, 2, \dots, K$ and $D^h_{k1} \neq D^h_{k2}$ for $k_1 \neq k_2$);
- $L^v : \langle D^v_1, D^v_2, \dots, D^v_k, \dots, D^v_K \rangle$ the ranked list of documents obtained from the match of the vertical projection ($D^v_k \in D$, for $k=1, 2, \dots, K$ and $D^v_{k1} \neq D^v_{k2}$ for $k_1 \neq k_2$);

The Borda-count approach assigns to each reference document D_k a confidence score $S(D_k)$ defined as [Ho et al., 1994]:

$$S(D_k) = S^h(D_k) + S^v(D_k) \quad (8)$$

being $S^h(D_k) = K - i$, if $D_k = D^h_i$; $S^v(D_k) = K - j$, if $D_k = D^v_j$.

Hence, the final list of ranked documents is

$$L^* : \langle D^*_1, D^*_2, \dots, D^*_k, \dots, D^*_K \rangle \quad (9)$$

so that D^*_{k1} precedes D^*_{k2} in L^* if and only is $S(D_{k1}) \geq S(D_{k2})$, and – of course – D^*_1 is the top candidate document [Ho et al., 1994].

3. Experimental Results

The experimental test was carried out on a dataset of 33 commercial forms belonging to 16 different categories. Figure 6 shows some examples of commercial forms in the dataset. In this case they belong to the category n. 3.

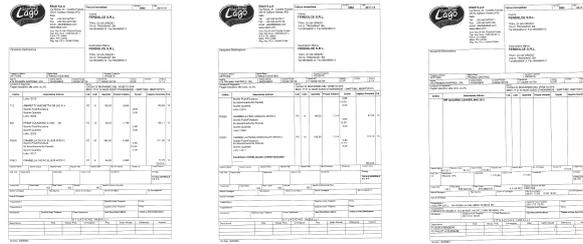


Figure 6. Examples of commercial forms of the same category.

Documents were scanned (100dpi , 256 gray-level) and preprocessed. Finally they were stored into a database along with the values of the Radon Transform concerning the horizontal ($S_{0,0}$) and vertical ($S_{0,\pi/2}$) projection. Table 1 reports the number of forms for each category.

Table 1. Dataset

Category	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Number of Documents	9	4	3	3	2	2	1	1	1	1	1	1	1	1	1	1

In the testing phase the leave-one-out method was considered to verify the effectiveness of the system. In order to estimate the quality of the ranked list provided by the system for a given query, the Average Normalized Rank (ANR) was adopted, defined as follows [Huang et al., 2005]:

$$ANR = \frac{1}{N \cdot N_w} \cdot \sum_{i=1}^{N_w} \left(R_i - \frac{N_w + 1}{2} \right) \tag{10}$$

being

- N the number of documents in the set,
- N_w the number of relevant documents (for the given query) in the set,
- R_i is the rank of each relevant document in the set.

It is worth noting that ANR ranges in $[0, 1]$:

- ANR=0 means that relevant documents are at the top of the ranked list (right position);
- ANR=1 means that relevant documents are at the bottom the ranked list (wrong position).

Figure 7 shows the experimental results. They demonstrate that the proposed approach is very robust with respect to different categories of documents. On average the value of ANR is equal to 0.08. Furthermore, 26 cases out of 33 the ANR is less than 10%, whereas only in one case it is greater than 0.5.

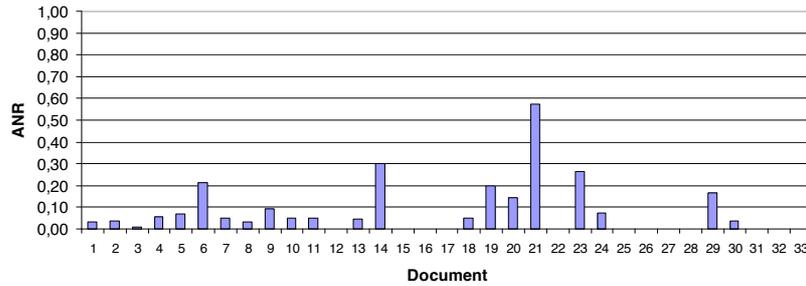


Figure 7. Experimental Results.

In order to estimate the robustness of the new approach two additional tests have been carried out using shifted and rotated document images as input.

When shifted documents are fed into the system the experimental results are shown in Figure 8. In this case a shift of 5 pixel is considered in the four main directions and the average result is computed. Also in this case the value of ANR is equal to 0.08, on average.

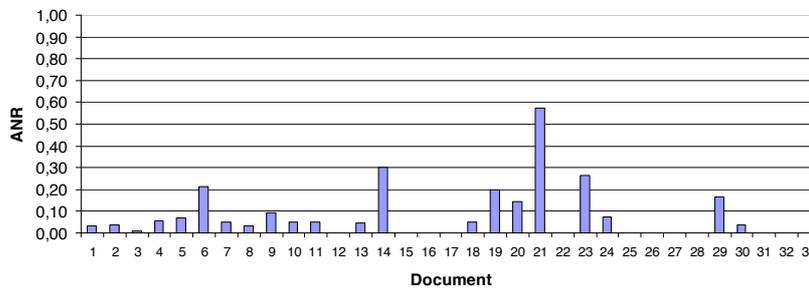


Figure 8. Experimental Results: shifted documents

Conversely, Figure 9 shows the results when rotated document images are fed into the system. In this case a rotation of 2° clockwise and anticlockwise is considered and the average result is computed. In this case the value of ANR is equal to 0.14 on average.

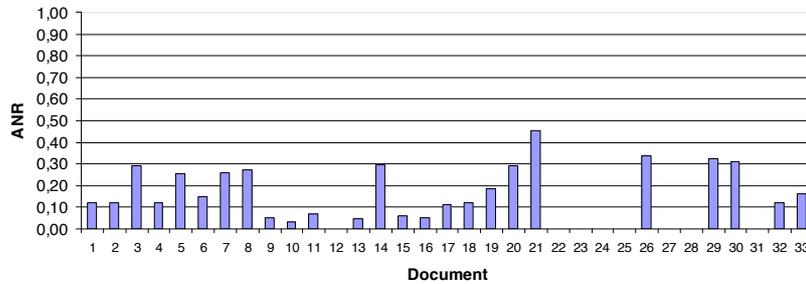


Figure 9. Experimental Results: rotated documents

Figure 10 shows the results when document images are shifted and rotated before to be fed into the system. In this case a shift of 5 pixels and a rotation of 2° clockwise and anticlockwise are considered. In this case the value of ANR is equal to 0.15 on average.

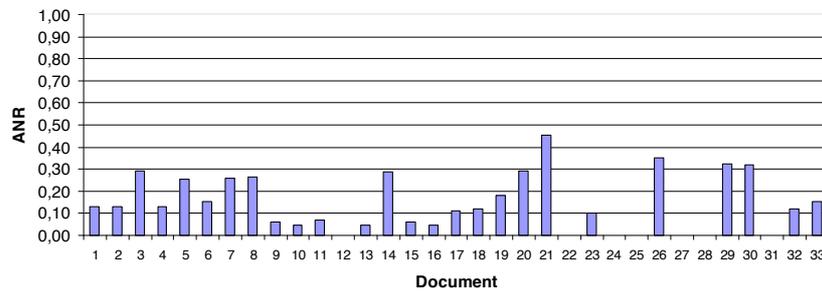


Figure 10. Experimental Results: shifted and rotated documents

4. Experimental Results

This paper presented a new system for layout-based document image retrieval. The system was specifically designed for retrieval of commercial forms as invoices, waybills and receipts, to optimize document management and sustainability. It used a morphologic filtering technique and the Radon Transform to obtain multiple document image descriptions. Document matching was then performed by Dynamic Time Warping whereas the Borda-count decision combination strategy was used to combine multiple decisions.

The experimental results, carried out on a dataset of real commercial documents, demonstrate the effectiveness of the proposed solutions and the robustness also with respect to shifted and rotated input document images.

5. References

- Lopresti, D., Robust Retrieval of noisy text, *Proc. of ADL'96*, 1996, 76–85.
- Cormack, A. M., Computed tomography: Some history and recent developments, *Proc. Symposia in Applied Mathematics*, 27, 1983, 35–42.
- Deans, S. R., *The Radon Transform and Some of Its Applications*, New York: Wiley, 1983.
- Doermann, D. , The Indexing and Retrieval of Document Images: A Survey, *Computer Vision and Image Understanding*, 70, 3, 1998, 287–298.
- Duygulu, P., Atalay, V., A Hierarchical Representation of Form Documents for Identification and Retrieval, *International Journal on Document Analysis and Recognition*, 5, 1, 2002, 17–27.
- Erol, B., Ant´unez, E., Hull, J. J., Hotpaper: multimedia interaction with paper using mobile phones, *Proceeding of the 16th ACM international conference on Multimedia*, 2008, 399–408.

Hjouj, F., Kammler, D. W., Identification of Reflected, Scaled, Translated, and Rotated Objects From Their Radon Projections, *IEEE Trans. Image Processing*, 17, 3, 2008, 301-310.

Ho, T.K., Hull, J.J., Srihari, S.N., Decision combination in multiple classifier systems, *IEEE Trans. Pattern Anal. Mach. Intell.*, 16, 1, 1994, 66–75.

Huang, M., Dementhon, D., Doermann, D., Golebiowski, L., Document ranking by layout relevance, *Proc. 8th ICDA*, 2005, 362–366.

Jafari-Khouzani, K., Soltanian-Zadeh, H., Radon Transform orientation estimation for rotation invariant texture analysis, *IEEE Trans. Pattern Anal. Mach. Intell.*, 27, 6, 2005, 1004–1008.

Kittler, J., Hatef, M., Duin, R.P.W., Matias, J., On combining classifiers, *IEEE Trans. on Pattern Analysis Machine Intelligence*, 20, 3, 1998, 226-239.

Lemire, D., Faster Retrieval with a Two-Pass Dynamic-Time-Warping Lower Bound, *Pattern Recognition*, 42, 9, 2009, 2169-2180.

Liu, Q., Liao, C., PaperUI, Proceeding of the 4th International Workshop on Camera-Based Document Analysis and Recognition, 2011, 3–10.

Manning, C. D. , Raghavan, P. , Schütze, H., An Introduction to Information Retrieval, Cambridge Press, 2009.

Marukawa, K., Hu, T., Fujisawa, H., Shima, Y., Document retrieval tolerating character recognition errors - Evaluation and application, *Pattern Recognition*, 30, 8, 1997, 1361-1371.

Mitra, M., Chaudhuri, B., Information retrieval from documents: A Survey, *Information Retrieval*, 2, 2/3, 2000, 141–163.

Salvador, S., Chan, P., Fast DTW: Toward Accurate Dynamic Time Warping in Linear Time and Space, *Proc. KDD Workshop on Mining Temporal and Sequential Data*, 2004, 70-80.

Seo, S., Haitzma, J., Kalker, T., Yoo, C. D., A robust image fingerprinting system using the Radon transforms, *Signal Process.: Image Commun.*, 19, 4, 2004, 325–339.

Serra, J., *Image Analysis and Mathematical Morphology*, Academic Press, 1982.

Taghva, K., Borsack, J., Condit, A., “Evaluation of model-based retrieval effectiveness with OCR text”, *ACM TOIS*, 14, 1, 1996, 64–93.

Tzacheva, A., El-Sonbaty, Y., El-Kwae, A., Document Image Matching Using a Maximal Grid Approach, *Proceedings of the SPIE Document Recognition and Retrieval IX*, 2002, 121-128.

Xu, L., Krzyzak, A., Suen, C. Y., Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition, *IEEE Transaction on Systems, Man and Cybernetics*, 22, 3, 1992, 418-435.

Acquisitions, what will happen to my business? Predicting the acquired firm status

Bice Della Piana, Filomena Ferrucci, Gerardina Flammia, Carmine Gravino
Dipartimento di Studi e Ricerche Aziendali – Management and Information Technology
Università degli Studi Salerno
Via Giovanni Paolo II, 132 - 84084 - Fisciano (SA)
{bdellapiana, fferucci, gravino}@unisa.it

Abstract. *Mergers and acquisitions are of great practical importance in strategic, monetary, and social terms. They have attracted the research attention of several management disciplines. However, no studies have used prediction techniques for analyzing these challenging and high risk activities, particularly if we consider the viewpoint of the acquired firm. Guided by the firms' needs on a support to their strategic decisions, we analyzed the stability of the acquired firm's status. With the aim of predicting the evolution of these cooperative arrangements, we carried out an empirical study meant to verify the ability of some Machine Learning techniques to predict the acquired firm's status after two years from the acquisitions. In our analysis we employed 107 firms that were involved in acquisition, and applied a 10-fold cross validation. The preliminary results suggest that two of the employed techniques can be profitably exploited to provide accurate estimates of the acquired firm's status.*

Keywords: Acquisition, Machine Learning, Empirical Study

1. Introduction

Inter-firm cooperation is today a crucial part of management (Contractor & Lorange, 2002). The most common types of cooperative arrangements are short-term contracts, relational contracting, licensing, logistical supply-chain relationships, equity joint ventures and the complete merger or acquisition of two or more organizations (Gulati & Singh, 1998). Among these, mergers and acquisitions (M&A) are characterized by a high mutual commitment between partners, a high expected longevity of the cooperation and high risks. For these reasons, M&A are popular and controversial (Hayward, 2002) and have attracted the interest of a broad range of management disciplines encompassing the financial, strategic, behavioral, operational and cross-cultural aspects (Cartwright & Schoenberg, 2006) (Gomes et al., 2013).

The stability of the acquired firm's status in the post-acquisition stage (intended here as the status of the acquired firm that does not change over time) represents an important factor of this arrangement with relevant managerial implication. Considering the firms' needs on a support to their strategic decisions and the complexity of the multifaceted phenomenon of the acquisitions, it is useful to verify the possibility to predict the status of the firm acquired in the post acquisition stage. To this aim, we carried out an empirical study aiming to assess the prediction capability of some Machine Learning (ML) techniques. The application of ML techniques is not new in management sciences (for a review see Hakimpour et al. 2011); indeed several problems have been addressed in different subject areas. Nevertheless, at the best of our knowledge, the prediction of the acquired firm status has not been investigated.

For the empirical analysis we employed as data set information about 107 acquisitions and analyzed the use of a subset of the classification techniques available in Weka (Hall et al., 2009). As for the validation method, we employed a 10-fold cross validation, while Precision, Recall, F-measures, Accuracy, and K-statistic were exploited as evaluation criteria.

The paper is organized as follows. Section 1 introduces the main research problem upon which this study is based and justifies the research methodology adopted. Section 2 describes the sources of data; Section 3 illustrates the design of the empirical analysis. Section 4 summarizes the findings of our investigation and highlights the future directions of the research.

2. Data set

The choice of the starting point for the identification of the firms being analyzed has not been easy to determine. There are many researches that utilize large alliance databases to explore issues related to firm collaboration (e.g., Folta and Miller, 2002; Hagedoorn, 2002; Lavie and Rosenkopf 2006; Vanhaverbeke, Duysters, and Noorderhaven, 2002; Villalonga and McGahan, 2005). Securities Data Company (SDC), MERIT-CATI, CORE, Bioscan, or Recombinant Capital (RECAP) are the alliance databases most commonly used in empirical studies published in top strategy journals (Shilling, 2009). Considering that our focus was the Italian context, these alliance databases were not useful for the purposes of the research. On the other hand the KPMG Report 2012 is very useful to understand the phenomenon of acquisitions in Italy since it aims to offer an across-the-board investigation of development trends in Italian business, by observing consolidation, internationalization, and competitive processes from a global standpoint. Nevertheless, this report was not helpful to our research goal since it does not provide the following information for both the acquired and the acquiring firms (name of the variables among brackets):

- general information about the firms: status as active or inactive (FirmsAcquiredStatus and FirmsAcquiringStatus, respectively), year in which the activity of the firms started, location of the headquarter (HLA1 and HLA2, respectively), legal capital (LegalcapitalofAcquired and LegalcapitalofAcquiring, respectively), value of production (ValueofProductionAcquired and

Value of Production Acquiring, respectively), number of employees (Number of Employees Acquired and Number of Employees Acquiring, respectively);

- specific information regarding the acquisitions: year of acquisition, specific constraints between firms.

As there is no database with similar features, it has been created ad hoc supported by the basic idea that it was necessary to have a data set with a certain degree of homogeneity among firms. To this aim, the most correct solution seemed to be to search for all Italian acquired firms which share the characteristic of being subject to the direction and coordination of other businesses (acquiring firms). In practice, the latter coordinates the strategic and operational decisions of the acquired firms. According to the art. 2497-bis of the Italian Civil Code, it has been specifically asked to select all the Italian companies that declared to the Chambers of Commerce to be involved in that type of acquisitions. Thus, the degree of homogeneity among firms is expressed by the communication of acquired firms of their situation of subjection to the management of the acquiring firms.

The research partnership between the Department of Management & Information Technology (University of Salerno) and the Chamber of Commerce allowed us to obtain information on the acquisitions carried out in Italy up to 2011. We obtained information on 26881 acquisitions in terms of the attributes/variables described above. We performed a preliminary analysis to remove those entries having null fields, obtaining as result 5.562 acquisitions. Moreover, before to process the data to address our research question, we carried out a data analysis with the aim of obtaining a more homogeneous sample and new attributes/variables needed for our prediction purposes. In particular, we introduced two further variables: Pre-acquisition Years of the Acquiring firm (PYA1) and Pre-acquisition Years of the Acquired firm (PYA2), where the PYA1 and PYA2 were obtained by taking into account the differences between the year in which the activity of the firms started and the year of acquisition (i.e., of being subject to the direction and coordination to other businesses - acquiring firms) communicated to the Chambers of Commerce.

From the selected 5562 acquisitions we could identify 4967 acquired firms having the status set to active, while the remaining 595 were acquired firms having status set to inactive. On the other hand, the analysis of the 5562 acquisitions allowed us to identify 5183 acquiring firms having status set to active, while the remaining 379 have status set to inactive. Moreover, we observed that more than 400 firms were characterized by PYA1 e PYA2 less than 7. The main part of the acquired firms was from regions in the north of Italy (i.e., Lombardia 1461, Veneto 774, Emilia Romagna 722) as well as the acquiring firms (i.e., Lombardia 1672, Veneto 753 Emilia Romagna 703), thus revealing that acquisitions are mainly generated among firms locating in the same region or in neighboring regions.

We also exploited the Principal Component Analysis (PCA) to perform a pre-analysis on the data to understand the relation among variables/attributes. PCA is widely used to reveal the internal structure of the data in a way that best explains the variance in the data (Abdi and Williams, 2010). The PCA results

showed that the variables positively correlated were: FirmsAcquiredStatus, FirmAcquiringStatus, LegalCapitalofAcquired, LegalCapitalofAcquiring, ValueofProductionAcquired, PYA1, PYA2, and NumberofEmployeesAcquired. On the other hand, FirmsAcquiredStatus and FirmAcquiringStatus were negatively correlated with NumberofEmployeesAcquiring and ValueofProductionAcquiring. Due to these results, we decided to exploit all these variables to address our research question. FirmsAcquiredStatus was the dependent variable, while the remaining variables were considered as independent.

Since we were interested in predicting the status of the acquired firms after two years from the acquisitions we needed the information on firms' status up to 2013 for all the firms obtained from the Chamber of Commerce (i.e., 5562 after preliminary analysis). This information can be achieved by manually querying the database made available by the Chamber of Commerce. Thus, we decided to randomly (according to the IP identification number ID) extract a number of firms (around 100) involved in acquisitions trying to obtain a balanced data set taking into account the information on the status. The result of our selection was: 57 firms tagged with an active status and 50 firms as inactive. Thus, we employed 107 observations (i.e., firms) in the empirical analysis, whose design is reported in the next section.

3. Design of the empirical analysis

We investigated some ML techniques to assess if it is possible to predict the status (active/inactive) of the firm involved in the acquisition as acquired. To this end, we formulated the following research question:

RQ: *Are the considered ML techniques able to predict the stability of an acquisition by estimating the status of the acquired firm?*

In the rest of this section we present the estimation techniques and the validation and evaluation criteria we employed in our empirical analysis. The results and their discussion are presented in Section 4.

3.1 Classification techniques

In this preliminary analysis we applied a subset of classification techniques available in Weka (Hall et al., 2009), which is a software “workbench” incorporating several classification techniques. It is worth noting that for all the employed techniques, no parameter setting was performed and we executed the default implementation provided by Weka (see Table 1). In the following we provide a brief description of these techniques. The reader interested to further details can refer to (Weka, 2013) and the references provided for the techniques.

Zero rule (ZeroR). It is the simplest of classifiers proposed by Weka, which considers no rules in its analysis and always outputs the most commonly found value for the dependent variable. For example, it predicts the mean if the type of the target attribute is numeric or the mode for an attribute whose type is nominal (Weka, 2013). It can be used as a lower bound on performance, i.e., a sort of benchmark, to assess the effectiveness of a proposed estimation technique.

Naive Bayesian Classifier (NBF). It is a simple but important probabilistic model (Caruana and Niculescu-Mizil, 2006), that uses Bayes's Theorem, i.e., assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. It can be considered very efficient in a supervised learning setting. In many practical applications, parameter estimation uses the method of maximum likelihood.

Support Vector Machines (SVMs). SVMs are a maximum-margin linear classification technique (Vapnik, 1995), since they find the optimal hyperplane between two classes defined by a number of support vectors (Vapnik, 1995). This technique has good generalization ability, thanks to the introduction of a penalty factor, named C, which allows us to cope with the effects of outliers, permitting a certain amount of misclassification errors. Moreover, the use of a kernel function to map the data points into a higher-dimensional feature space allows SVMs to handle also non-linear problems (Vapnik, 1995). The parameter *epsilon* determines the level of accuracy of the approximation. Its value relies on the target values in the training set and the parameter C determines the trade-off between the occurrences of errors larger than epsilon in the training set and the flatness of the function. We chose to apply the Polynomial and the RBF kernels since they are widely used for prediction problems in different contexts (Scholkopf et al., 1997, Corazza et al., 2011, 2013).

MultiLayer Perceptrons (MLP). MLP is a feedforward artificial neural network that is trained with the backpropagation learning algorithm, a standard algorithm for any supervised-learning pattern recognition process (Miller et al., 2011). MLP uses three or more layers of neurons (nodes) with nonlinear activation functions and each layer fully connected to the next one. In our work we employed three layers, where the input layer is connected with the input data; the intermediate layer or hidden layer receives data from the input layer and sends them to the layer of output neurons; the last level or output layer receives data from neurons intermediate layer, and it interfaces with the output. As for the parameter settings (see Table 1), *HiddenLayers* represents the number of hidden neurons and there are some wildcards for this value as $a=(attribute+classes)/2$. *LearningRate* and *Momentum* are hyper parameters that determine the behavior of the algorithm, while *seed* is used for the selection of the starting point and for the shuffling of data. *TrainingTime* and *ValidationSetSize* represent the maximum number of iterations to execute and the number of patterns for training set that will not be used to find the weights. *ValidationThreshold* is a stopping criterion.

j48. It is the implementation of decision tree C4.5, a method of supervised classification structured in non-terminal and terminal nodes. The terminal nodes represent the results of the decision, while the non-terminal nodes indicate the test on one or more attributes (Zhao and Zhang, 2011). The algorithm tries to recursively partition the data set into subsets by evaluating the normalized information gain (difference in entropy) resulting from choosing a descriptor for splitting the data. The training process stops when the resulting nodes contain instances of single classes or if no descriptor can be found that would result to the information gain. *ConfidenceFactor* is used for pruning while *NumberObjects* allows setting the minimum number of instances before to fix a

leaf in the tree (set as reported in Table 1). *Seed* is used for randomizing the data when reduced error pruning is employed.

Simple Classification And Regression Trees (SCART). SCART is an algorithm to construct decision trees using historical data (Breiman et al., 1984). Through a pruning process, SCART uses cross-validation to split the input data into the two subgroups that are the most different with respect to the outcome. As for parameter setting (see Table 1), *MinNumObjects* represents the minimal number of observations at the terminal nodes while *NumFoldsPruning* is the number of folds in the internal cross-validation. *Seed* is the random number seed to be used and *SizePer* denotes the percentage of the training set size (in the range [0-1], where 0 means not included).

Table 1. The parameters employed for classifying

C	epsilon	Kernel,
1.0	1.0E-12	RBF with gamma=0.01
1.0	1.0E-12	Polynomial with c0=250007 and Exponent=1.0

(a) SVM

Hidden Layers	Random Seed	Training Time	Learning Rate	Momentum	Validation SetSize	Validation Threshold
a	16	1000	0.3	0.2	20	20

(b) MultiLayer Perceptrons

Confidence Factor	Number Objects	Seed
0.25	2	16

(c) J48

MinNumObjects	Random Seed	NumFoldsPruning	SizePer
2	16	10	1.0

(d) SCART

3.2 Validation method and evaluation criteria

We carried out a validation to verify whether or not ML techniques can provide useful predictions of the status of the acquired firm. To this end, we applied a 10-fold cross validation, randomly partitioning the original data set into 10 equal size subsets where a single subset is used as the validation set for assessing the predictions, and the remaining 9 subsets are used as training set to obtain the predictions. Thus, the validation process is repeated 10 times to use each of the 10 subsets exactly once as the validation set.

Regarding the evaluation criteria, we used several accuracy measures to assess the achieved predictions. In particular, we employed four performance measures, i.e., Accuracy, Precision, Recall, and F-measure (Witten I., Frank, 2005), which leverage on the concepts reported in the confusion matrix of Table 2 and are defined as follows. *Accuracy* is the ratio between the number of components correctly predicted (i.e., classified as TP and TN) and the total number of components (i.e., the sum of TP, TN, FP, FN). *Precision* is the ratio between the number of components classified as TP and the number of those classified as TP or FP. *Recall* is the ratio between the number of components classified as TP and the number of those classified as TP or FN

Table 2. The confusion matrix

		Predicted	
		Inactive	Active
Actual	Inactive	True Positive (TP)	False Negative (FN)
	Active	False Positive (TP)	True Negative (TN)

These definitions suggest that Precision concerns with the correctness of the responses provided by a method, while Recall allows us to measure the completeness of the responses. A measure that provides an indication of a balance between correctness and completeness is the harmonic mean of Precision and Recall (or *F-measure*), defined as follow:

$$F - measure = 2 * \frac{(Precision * Recall)}{(Precision + Recall)}$$

Moreover, we employed the K-statistic which is a chance-corrected measure of agreement between the classifications and the true classes (Weka, 2013). K-statistic is defined as:

$$K - statistic = \frac{P(A) - P(E)}{1 - P(E)}$$

where P(A) is the proportion of times the k raters agree and P(E) is the proportion of times the k raters are expected to agree by chance alone. K=1 means complete agreement while K=0 denotes lack of agreement (i.e. purely random coincidences of rates).

3.3 Validity evaluation

The validity of our empirical studies can be biased by different factors. First of all, threats to construct validity can be due to the way the variables have been selected and collected. We tried to mitigate such a threat employing publicly available data sets from the Chambers of Commerce. Threats to internal validity can be due to the bias introduced by the default choices considered in the application of the estimation techniques. However, our focus was not on the use of a specific technique but on the variables employed in combination with machine learning approaches to produce the predictions of the stability of an acquisition after the formation stage, by detecting the status (active/inactive) of the acquired firm. As for the external validity, it can be affected by the specific data set employed for the empirical analysis. However, this threat is mitigated by the fact that we randomly selected firms from the original data set provided by the Chambers of Commerce.

4. Results and discussion

Table 3 summarizes the results achieved in our empirical analysis in terms of K-statistics, Accuracy, Precision, Recall, and F-measure. We can observe that the decision tree based algorithms (i.e., NBF and SCART) provided predictions with the highest values for the considered performance measures

and with a high K-statistic value. In particular, the predictions on the status of the acquired firm in an acquisition with a Precision of about 0.87, thus with an incorrect estimates of about 13%. The predictions are also good in terms of completeness since the Recall value is 0.86 for NBF and SCART. Even if SVM_poly, MLP, and J48 achieved performance measure values close to the ones of NBF and SCART the K-statistic values suggest that this result was obtained by chance c.a. 50%. Thus, the performed analysis seems to reveal that the considered variables in combination with NBF and SCART are useful to predict the status of the acquired firms. More complex estimation techniques, such as SVMs, need accurate parameter settings “to give their best”. Using the default setting of WEKA did not allow us to get useful predictions (with the RBF kernel the results are equal to the ones provided by ZeroR).

The results of this preliminary analysis are encouraging and motivate us to further investigate ML techniques in this context. We think that the error percentage characterizing the predictions can be reduced by optimizing the parameters setting.

Thus, we can positively answer our research question: *two of the employed ML techniques (i.e., NBF and SCART) are able to predict the stability of an acquisition by estimating the status of the acquired firm.*

Table 3. Results for each technique

Technique	K statistic	Performance Results			
		Accuracy	Precision	Recall	F-measure
ZeroR	0	0.533	0.284	0.533	0.370
NBF	0.716	0.860	0.865	0.860	0.859
SVM_Poly	0.436	0.729	0.820	0.729	0.701
SVM_RBF	0	0.533	0.284	0.533	0.370
MLP	0.51	0.757	0.757	0.757	0.756
J48	0.586	0.794	0.794	0.794	0.794
SCART	0.716	0.860	0.865	0.860	0.859

From a managerial perspective it is possible to highlight some of the strengths and weaknesses of the research. Referring to the objective of the research, knowing in advance the evolution of the firm acquired status could have relevant managerial implication. Strength of our work lies in the originality in terms of the specific condition (status) and specific phase (post acquisition) analyzed. None so far has bothered to understand what happens in the post-acquisition phase by analyzing the stability of the status of the acquired firm. The information about the stability of the status is particularly relevant. The prediction of the stability could be taken as an ex ante measure of the effectiveness of the strategy formulated by both the acquiring and acquired firms. Effectiveness intended as a measure of the chance to tie the actions, which in this case refers to the strategic choice of the acquisition, preferences and objectives of the companies involved. A weak point of our work involves the paucity of information on firms under investigation. It would be wise to have more information for both acquired and acquiring firms, as for example the change in sales and added value. Future developments will cover the use of prediction in order to analyze the coordination of activities between acquiring

and acquired firm whose status was predicted as active. As Agarwal et al. (2012) suggested, when one firm acquires another, there are numerous challenges in achieving coordination of activities between the acquired and acquiring entities.

As for future work, we intend to replicate the performed analysis by considering different and larger data sets and further estimation techniques. Furthermore, we are aware that the choice of parameters can have a strong impact on the application of the considered techniques, such as SVMs. In the presented preliminary study we applied the techniques with the default setting provided by Weka. In the future, we will search for the best parameter settings using approaches that allowed us to explore a wide range of variables' values.

References

- [1] Abdi H., Williams L., Principal component analysis, in Wiley Interdisciplinary Reviews: Computational Statistics, 2, 4, July/August 2010, 433–459.
- [2] Agarwal R., Croson R., Mahoney J. The role of incentives and communication in strategic alliances: an experimental investigation. *Strategic Management Journal*, 31(4), 2010, 413–437.
- [3] Agarwal R., Anand J., Bercovitz J., Croson R., Spillovers across organizational architectures: the role of prior resource allocation and communication in post-acquisition coordination outcomes. *Strategic Management Journal*, 33, 2012, 710–733.
- [4] Breiman L., Friedman J., Stone C., Olshen R., *Classification and Regression Trees*, 1984, Ed. Chapman and Hall/CRC.
- [5] Cartwright S., Schoenberg R., Thirty years of mergers and acquisitions research: recent advances and future opportunities. *British Journal of Management*, 17(S1), 2006, S1-S5.
- [6] Caruana R., Niculescu-Mizil A., An empirical comparison of supervised learning algorithms, *Proceedings of the International Conference on Machine Learning*, 2006.
- [7] Conover W., *Practical Nonparametric Statistics*, 3rd edition. Wiley, 1998.
- [8] Contractor F.J., Lorange P., The growth of alliances in the knowledge-based economy. *International Business Review*, 11, 2002, 485–502.
- [9] Corazza A., Di Martino S., Ferrucci F., Gravino C., Mendes E., Investigating the use of support vector regression for web effort estimation. *Empirical Software Engineering*, 16, 2, 2011, 211–243.
- [10] Corazza A., Di Martino S., Ferrucci F., Gravino C., Sarro F., Mendes E., Using Tabu Search to configure support vector regression for effort estimation. *Empirical Software Engineering*, 8, 3, 2013, 506-546.
- [11] Folta T.B., Miller K.D. Real options in equity partnerships. *Strategic Management Journal* 23(1), 2002, 77–88.
- [12] Gulati R, Nickerson J. Interorganizational trust, governance choice, and exchange performance. *Organization Science*, 19(5), 2008, 688–708. Sarkar M, Aulakh PS, Madhok A. Process capabilities and value generation in alliance portfolios. *Organization Science*, 20(3), 2009, 583–600.

[13] Gulati, R., & Singh, H. The architecture of cooperation: Managing coordination costs and Appropriation Concerns in Strategic Alliances. *Administrative Science Quarterly*, 1998, 43(4), 781–814.

[14] Hakimpoor H., Arshad K.A.B., Tat H.H., Khani N., Rahmandoust M., Artificial Neural Networks' application in management, *World Applied Sciences Journal*, 14(7), 2011, 1008-1019.

[15] Hall M., Eibe F., Holmes G., Pfahringer B., Reutemann P., Witten I., The WEKA Data Mining Software: An Update; *SIGKDD Explorations*, 11, 1, 2009.

[16] Hagedoorn J., Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. *Research Policy*, 2002, 31, 477–492.

[17] Hayward M.L.A., When do firms learn from their acquisition experience? Evidence from 1990-1995. *Strategic Management Journal*, 23, 2002, 21-39.

[18] Kitchenham B., Mendes E., Travassos G., Cross versus Within-Company Cost Estimation Studies: A Systematic Review. *IEEE Trans. Software Eng.* 33,5, 2007, 316-329.

[19] Lavie D., Rosenkopf L., Balancing exploration and exploitation in alliance formation. *Academy of Management Journal* 49, 2006, 797–818.

[20] Miller F., Vandome A., McBrewster J., Multilayer Perceptron: Artificial neural network, Activation function, Backpropagation, Linear separability, Perceptron, Supervised learning, Action potential, 2011, Ed.: Alphascript Publishing, ISBN-13: 978-6134149709

[21] Scholkopf B., Sung K., Burges C., Girosi F., Niyogi P., Poggio T., Vapnik V., Comparing support vector machines with Gaussian Kernels to radial basis function classifiers. *IEEE Trans Signal Process* 45, 11, 1997, 2758–2765.

[22] Vanhaverbeke W., Duysters G., Noorderhaven N., External technology sourcing through alliances or acquisitions: an analysis of the application-specific integrated circuits industry. *Organization Science* 13, 2002, 714–733.

[23] Vapnik V., *The Nature of Statistical Learning Theory*. Springer-Verlag, 1995.

[24] Weka Documentation, <http://www.cs.waikato.ac.nz/ml/weka/documentation.html>. Last check July 2013.

[25] Villalonga B., McGahan A.M., The choice among acquisitions, alliances, and divestitures. *Strategic Management Journal* 26(13), 2005, 1183–1208.

[26] Witten I., Frank E., *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2nd Ed., San Francisco, 2005.

[27] Zhao Y., Zhang Y., Comparison of decision tree methods for finding active objects, *Advances in Space Research*.

Il ruolo delle start-up ICT-based nel cambiamento del paradigma di business: il caso delle smart grid¹

Roberto Parente¹, Rosangela Feola², Valter Rassega³

¹Università di Salerno
Via Giovanni Paolo II, 132 - Fisciano 84084 (SA)
rparente@unisa.it

²Università di Salerno
Via Giovanni Paolo II, 132 - Fisciano 84084 (SA)
rfeola@unisa.it

³Regione Basilicata
Ufficio Società dell'Informazione
Via V. Verrastro, 4 - 85100 Potenza (PZ)
walter@rassega.it

Abstract. *Business prospects related to the development of green technologies and energy saving systems have attracted the interest of ICT-based firms that have the crucial task of making more and more "intelligent" the system of production, distribution and use of the energy produced. In this process of radical redefinition of the energy business, an important role can be played by start-up that come from the world of research. The paper starts from a definition of the concept of smart grid as a trigger for radical change of the paradigm of the energy business. From a technological point of view we will examine in particular the role of ICT as enabling technologies in the development of smart grid. At the end we will analyze the role of ICT-based start-up in the development of smart grid, adding further evidence to the mainstream literature available on the role of new ventures in the development of radically new technologies.*

Keywords: Start-up; Technology Innovation; Smart grid

¹ Pur essendo il lavoro frutto della comune attività di ricerca degli Autori, il paragrafo 1 è attribuibile a Valter Rassega, il paragrafo 2 a Rosangela Feola, i paragrafi 3 e 4 a Roberto Parente. L'introduzione è attribuibile a tutti gli Autori. Gli Autori ringraziano la dott.ssa Valentina Cucino e il dott. Ezio Marinato per il supporto nella ricerca sugli spin-off italiani. Tuttavia, ogni responsabilità resta imputabile agli Autori.

Introduzione

Il paradigma tradizionale del business elettrico sta profondamente cambiando spostando l'attenzione dalle tradizionali fonti energetiche non rinnovabili verso fonti energetiche rinnovabili. Il sistema energetico è sempre più orientato al perseguimento di tre obiettivi: l'integrazione ottimale delle fonti rinnovabili nella rete di distribuzione elettrica, l'aumento dell'efficienza, della sicurezza e dell'affidabilità del servizio e la creazione di un mercato competitivo dell'energia. Il passaggio che ormai si impone è quello verso la **smart grid**, ovvero "una rete di trasporto e distribuzione di energia elettrica di nuova generazione in grado di integrare intelligentemente il comportamento e le azioni di tutti gli utenti ad essa connessi – produttori, consumatori e produttori/consumatori – allo scopo di assicurare un approvvigionamento elettrico efficiente, sostenibile, economico e sicuro" [European Technology Platform «Smartgrids», 2006]. Nello sviluppo delle smart grid, le tecnologie ICT rappresentano un elemento essenziale: le tecnologie informatiche rappresentano infatti una tecnologia abilitante in grado di conferire al sistema le caratteristiche di "intelligenza" necessarie a svolgere le funzioni "smart" della rete. Nel processo di ridefinizione del business dell'energia che si caratterizza per l'interrelazione che viene a crearsi tra il settore elettrico ed il settore ICT, diviene fondamentale indagare il ruolo che le imprese, vecchie e nuove, possono giocare nel processo di cambiamento del settore. In questo contesto, il paper si ripropone di mettere in luce il ruolo delle start-up ICT provenienti dal mondo della ricerca nello sviluppo delle smart grid. Dal punto di vista teorico il lavoro può essere inquadrato nell'ambito degli studi che analizzano i processi di trasformazione dei settori economici, ed in particolare di quelli che si sono focalizzati sull'analisi dei fattori e degli attori che guidano il cambiamento [Geels, 2002; Geels, 2004; Geels e Schot, 2007; Geels e Raven, 2006; Dolata, 2009; Erlinghagen e Markard, 2012].

1. Smart grid: verso un nuovo paradigma nel settore energetico

L'attuale evoluzione del mercato energetico si caratterizza per la regolazione delle strutture di mercato per l'energia elettrica che devono integrare agevolmente nella rete un numero crescente di produttori di energia elettrica, anche di piccola taglia, proveniente da fonti rinnovabili: vengono introdotti, in questo modo, nuovi concetti, come la gestione e modulazione della domanda di energia elettrica da parte dell'utilizzatore finale del servizio. Non a caso, una delle funzioni fondamentali della smart grid è quella di integrare nel sistema elettrico la micro generazione di energia elettrica operata da privati che, con i loro impianti fotovoltaici o eolici, potendo conferire in rete il loro surplus energetico, assumono la doppia veste di consumatore e di produttore: per l'appunto, i c.d. "prosumers". Per quanto limitata in termini di potenza prodotta per singolo impianto, la prevedibile diffusione su vasta scala di impianti di generazione di tipo casalingo segna l'avvento di un nuovo paradigma economico: il passaggio da un mercato di monopolio naturale integrato verticalmente, ad un mercato orizzontale dove potenzialmente tutti possono consumare e produrre senza limitazioni tecniche o normative. Quando produrre e vendere elettricità diventa più semplice, si attivano scambi e si crea uno

Il ruolo delle start-up ICT-based nel cambiamento del paradigma di business: il caso delle smart grid

spazio di business per nuovi soggetti economici che offrono servizi di aggregazione. L'aggregatore è una delle figure emergenti nel nuovo panorama del settore energetico, e svolge la funzione di mediatore tra i consumatori, ai quali vende le flessibilità di carico ovvero la disponibilità a modulare i propri consumi energetici, e il mercato. In tal modo, l'aggregatore risponde alle richieste provenienti dal mercato e dal gestore del sistema di distribuzione, cercando di soddisfarle, utilizzando e aggregando la capacità di poter variare gli assorbimenti e, parallelamente, valutando le disponibilità energetiche fornite da un aggregato di prosumers. Dunque, l'attuale mercato dell'energia si configura come un ambito di interazione che vede un insieme composito di portatori di interesse. In chiave economica, deve essere analizzata tale interazione applicata al modello di tecnologia di rete della smart grid. Per gli economisti le tecnologie di rete sono quelle infrastrutture tecnologiche che connettono più agenti in rete coinvolti in un'interazione, e per le quali esistono delle esternalità positive. Detto in altri termini, ciò significa che il beneficio per il singolo agente non dipende solo dalla sua adozione individuale della tecnologia, ma è proporzionale al numero di altri agenti che entrano nel gioco. La smart grid sarà utile ai vari stakeholders solo se tutti i players di mercato parteciperanno alla loro implementazione, e l'utilità sarà tanto maggiore quanto più il modello diventerà diffuso e pervasivo. La transizione verso il nuovo paradigma sembra ormai avviata, ma sono subito emerse alcune barriere che è necessario superare. Una fra queste è quella delle infrastrutture di trasporto e delle connessioni di rete, che devono modificare le proprie caratteristiche e modalità operative per poter gestire flussi di energia multi-direzionali ed integrare in rete i sempre più numerosi impianti di generazione di energia elettrica. Per rispondere a queste esigenze, è necessario che i dispositivi di controllo e di gestione di tutti gli impianti di produzione e distribuzione elettrica, possano dialogare, sincronizzarsi e coordinarsi automaticamente fra loro per provvedere al trasporto e alla distribuzione dell'energia elettrica con la massima efficienza ed affidabilità. Una rete intelligente dunque fortemente caratterizzata dal ruolo fondamentale svolto dalle tecnologie ICT, grazie alle quali i dispositivi e gli apparati di controllo e gestione della rete elettrica, gli utilizzatori, i produttori e ogni altra figura coinvolta nel network possono scambiarsi informazioni in tempo reale, determinando lo sconvolgimento dei tradizionali paradigmi di settore e le conseguenti modificazioni degli scenari economici e tecnologici nel campo energetico. Non a caso, nella prospettiva economica, la smart grid può essere considerata un agente "transattivo" che permette di effettuare transazioni di natura finanziaria e informativa tra tutti gli attori del mercato energetico.

Per ritenersi "Smart", la grid di nuova generazione deve necessariamente avere alcune caratteristiche fondamentali: deve essere innanzitutto un diffuso e potente sistema di elaborazione computerizzato, connesso con i dispositivi di misura e le apparecchiature di controllo della rete, in grado di effettuare misure, calcoli e comunicare in tempo reale con gli altri sistemi di elaborazione costituenti l'arcipelago tecnologico della rete stessa. La tecnologia necessaria è già disponibile e alcuni degli apparati di gestione della grid sono già installati e in esercizio sulle reti di distribuzione ad alta e media tensione e molti produttori stanno rilasciando nuovi dispositivi elettronici di potenza per il controllo del

flusso energetico a livello di distribuzione a bassa tensione. Tuttavia, va rilevato che tali dispositivi, per concorrere realmente alla realizzazione di una rete intelligente, devono integrarsi perfettamente e, anche se molto è già stato fatto, ad oggi si devono ancora definire alcuni standard necessari a supportare un livello di interoperabilità hardware e software di tipo "plug-and-play" tra tutti i dispositivi elettronici utilizzabili nella smart grid. In termini operativi, la smart grid è in grado di conoscere e profilare le nostre abitudini nell'ambito dei consumi elettrici e il fabbisogno degli elettrodomestici e degli eventuali veicoli elettrici posseduti. Quando nella rete la richiesta di elettricità è al massimo, la Smart grid si adatta automaticamente alla domanda raccogliendo l'energia in eccesso disponibile da altre fonti connesse in rete, evitando problemi di sovraccarico. Oltretutto, il sistema elettrico Smart supporta e gestisce comunicazioni in tempo reale tra i consumatori e i gestori della distribuzione e ciò fornisce automaticamente agli utenti le informazioni e gli strumenti per il controllo e l'esercizio delle opzioni che rendono possibile adattare dinamicamente il proprio profilo di consumo di energia in funzione dell'andamento del prezzo o, nel caso di utenti che dispongono di un proprio impianto di generazione eolica o fotovoltaica, il sistema consente agli stessi di intervenire attivamente nel mercato elettrico immettendo in rete il proprio surplus di energia prodotta, ad un prezzo di vendita determinato dinamicamente dal sistema. Il percorso verso l'effettiva e completa realizzazione della smart grid appare solo agli inizi e la strada da compiere è ancora lunga. Il mercato delle apparecchiature e dei servizi per la costruzione di una rete sempre più intelligente è in forte crescita ed evidenzia un trend positivo che non è soltanto italiano ma che è diffuso anche a livello internazionale.

2. Il ruolo delle tecnologie ICT nello sviluppo delle smart grid

Nel passaggio dal vecchio al nuovo paradigma energetico il ruolo delle tecnologie ICT risulta sempre più determinante. Le tecnologie ICT rappresentano infatti il fattore abilitante di un processo di trasformazione strutturale di ogni fase del ciclo energetico, dalla generazione all'accumulo, al trasporto, alla distribuzione, alla vendita e, soprattutto, al consumo intelligente di energia. Le tecnologie ICT rappresentano l'elemento in grado di conferire al sistema le caratteristiche di "intelligenza" necessarie a svolgere le funzioni "smart" della rete e di poter operare il controllo automatico e adattivo delle infrastrutture in funzione delle condizioni operative del momento. Il connubio tra smart grid e tecnologie ICT è implicito nell'essenza stessa del processo di realizzazione della smart grid: realizzare una smart grid non significa infatti rifare da zero le infrastrutture della rete elettrica ma significa arricchire di intelligenza quelle esistenti. Si potrebbe parlare di un Internet of Energy in cui una rete di informazione si affianca ad una rete di distribuzione fisica dell'energia consentendone una gestione intelligente. Le tecnologie ICT intervengono nello sviluppo della smart grid a due livelli [Bellifemine et al, 2009]: forniscono gli strumenti, le piattaforme e gli algoritmi di controllo distribuito, necessari ad ottimizzare l'efficienza di tutti i sistemi coinvolti; divengono il fattore abilitante per lo sviluppo di una serie di servizi legati all'energia.

Nello sviluppo di una rete energetica intelligente, le tecnologie ICT consentono di rispondere a tre esigenze fondamentali: favorire l'integrazione

Il ruolo delle start-up ICT-based nel cambiamento del paradigma di business: il caso delle smart grid

della generazione distribuita; facilitare la gestione dei flussi multi direzionali; favorire il coinvolgimento del cliente.

Va tuttavia precisato che molte delle tecnologie ICT necessarie per le smart grid sono oggi disponibili come elementi separati, sviluppati a diversi gradi di maturità. Si rendono necessari ulteriori investimenti in R&S finalizzati a sviluppare le tecnologie funzionali alla smart grid portandole a un livello di maturità adeguato a poterle installare su vasta scala. Il potenziale del mercato ICT collegato agli investimenti nelle reti intelligenti è estremamente elevato come dimostra il crescente coinvolgimento delle imprese ICT nei progetti europei relativi alle smart grid. Se si analizza infatti l'andamento della partecipazione delle imprese ICT ai progetti europei, si riscontra, dal 2004 ad oggi una tendenza crescente: da una percentuale di circa il 20% di progetti partecipati da almeno un'impresa ICT nel 2004, si arriva ad una percentuale di oltre il 60% nel 2012 [EU, 2013]. La consapevolezza del ruolo fondamentale delle tecnologie informatiche è testimoniata, sul fronte istituzionale, dall'avvio di numerose iniziative europee (quali il progetto SEESGEN ed il progetto Finsey) che hanno l'obiettivo di potenziare il ruolo dell'information technology e delle telecomunicazioni per migliorare l'efficienza energetica nella produzione e distribuzione di energia. Le prospettive di business legate alle applicazioni ICT nel campo delle smart grid hanno richiamato l'interesse dei grandi operatori del settore (Telecom, Cisco, Ibm, Microsoft e Google) che hanno effettuato notevoli investimenti nella direzione smart e nella direzione dello sviluppo di dispositivi e tecnologie informatiche necessarie alla implementazione di una rete energetica intelligente. Tuttavia, il processo è appena agli inizi e la strada da percorrere verso la completa implementazione di una rete energetica intelligente è ancora lunga: il progetto "European Smart Grids Technology Platform" della Commissione Europea stima che siano ancora necessari investimenti per 750 miliardi di Euro nei prossimi trent'anni, di cui 100 miliardi nella trasmissione, 300 miliardi nella distribuzione e 350 miliardi nella generazione.

Il tema del ruolo dell'ICT nello sviluppo delle smart grid, può essere inquadrato nell'ambito degli studi che analizzano i processi di trasformazione dei settori economici, ed in particolare di quelli che si sono focalizzati sull'analisi dei fattori e degli attori che guidano il cambiamento [Geels, 2002; Geels, 2004; Geels e Schot, 2007; Geels e Raven, 2006; Dolata, 2009]. A questo proposito è stato riconosciuto che un ruolo determinante nella trasformazione di un settore può essere giocato da imprese consolidate operanti in settori magari anche tecnologicamente distanti da quello di riferimento ma portatori di competenze e tecnologie in grado di innescare processi di cambiamento [Erlinghagen e Markard, 2012]. Tali incumbents infatti potrebbero decidere di entrare in un nuovo settore con l'obiettivo di sfruttare al meglio la già esistente base di conoscenze e competenze [Nicholls-Nixon e Jasinski, 1995]. Il ruolo di tali incumbents diventa particolarmente rilevante quando le tecnologie sviluppate all'esterno del settore di riferimento (quelle che potremmo definire come outside technologies) diventano abilitanti per la trasformazione del settore stesso [Dolata, 2009]. In questi casi infatti, settori precedentemente lontani e completamente separati, divengono strettamente interconnessi e dipendenti gli uni dagli altri [Erlinghagen e Markard, 2012]. L'interrelazione tra settori che

viene a determinarsi quando la forza del cambiamento è esterna al settore di riferimento, ha importanti conseguenze sulle dinamiche competitive del settore stesso. In questi contesti infatti, il gioco competitivo coinvolge non solo gli incumbents del settore ed i nuovi entranti, ma anche gli incumbents dei settori adiacenti, che potrebbero decidere di entrare nel focal sector per sfruttare le opportunità che in esso si presentano. L'ingresso degli incumbents adiacenti è in grado di alterare il "naturale gioco competitivo" che come dimostra la letteratura sull'argomento, generalmente vede gli incumbents del settore soccombere di fronte alla forza innovatrice dei new-comers [Hannan e Freeman, 1977]. Di fronte alla minaccia derivante dagli incumbents adiacenti, le imprese consolidate nel settore focale, piuttosto che resistere al cambiamento, tendono a reagire adottando strategie offensive con l'obiettivo di acquisire, ad esempio attraverso acquisizioni di nuove imprese, le competenze e le tecnologie che non possiedono al proprio interno [Christensen, 1997].

Il settore delle smart grid è perfettamente rappresentativo di tali dinamiche competitive. Lo sviluppo di una rete intelligente infatti, implica una profonda trasformazione del settore energetico derivante dall'applicazione al suo interno delle conoscenze e delle tecnologie derivanti dal settore ICT. Lo sviluppo delle smart grid dipenderà quindi dai comportamenti e dalle strategie che saranno adottate dagli incumbents elettrici, dagli operatori incumbents del settore ICT e da nuovi entranti sia nel settore elettrico che nel settore ICT.

La nostra analisi si focalizzerà in particolare sul ruolo che la ricerca e le start-up ICT possono svolgere nel cambiamento del paradigma energetico.

3. Il contributo della ricerca e delle start-up nello sviluppo del nuovo paradigma energetico

Shock esogeni al sistema produttivo e nuove acquisizioni scientifiche e tecnologiche rappresentano le driving forces dei cambiamenti radicali nella struttura dei settori industriali [Geels, 2004]. Ovviamente, pur essendo possibile che nuove scoperte scientifiche e nuove tecnologie nascano sulla scia di una ricerca curiosity driven, è ancor più verosimile che siano shock esogeni al sistema produttivo, com'è nel caso del cambiamento climatico in corso, ad agire quale elemento catalizzatore nello sviluppo di nuove tecnologie destinate a cambiare i connotati dei processi produttivi tradizionali [Schaltegger e Wagner, 2011]. La direzione di questo cambiamento, ed i tempi necessari per portarlo a compimento, dipendono da una molteplicità di attori e dai comportamenti che essi mettono in campo, dando vita ad una complessa rete di influenze e di condizionamenti reciproci. Le imprese rappresentano gli attori chiave del cambiamento, giacché in ultima analisi sono queste peculiari strutture organizzative che, alla ricerca delle più idonee condizioni di sviluppo e di profittabilità, sono chiamate a dover elaborare una sintesi fra bisogni del mercato e tecnologie disponibili, attraverso una loro peculiare offerta di prodotti/servizi. In un sistema competitivo aperto è pacifico che tipologie diverse di imprese possano avere atteggiamenti, capacità, valutazioni diverse (anche fortemente divergenti) sia rispetto alle strategie da adottare che ai tempi della loro implementazione [Denrell et al, 2003]. A questo proposito, un tema ben noto in letteratura è legato al diverso atteggiamento nei confronti delle

Il ruolo delle start-up ICT-based nel cambiamento del paradigma di business: il caso delle smart grid

innovazioni radicali espresso dalle aziende già consolidate (i cosiddetti incumbents) rispetto alle aziende di nuova formazione (le cosiddette start-up). Per spiegare tale fenomeno si è fatto ricorso al concetto di nicchia ecologica e di inerzia organizzativa [Hannan e Freeman, 1977]. Le imprese, cioè, che hanno avuto successo in un determinato ambiente tecnologico difficilmente hanno la forza e la capacità di sostituire/modificare radicalmente le proprie competenze tecnologiche e finiscono così per subire un processo di “adverse selection” ovvero ad essere sostituite da coorti di nuove imprese che nascono fin dall’inizio cavalcando le nuove tecnologie [Hill e Rothaermel, 2003]. Dosi ed altri [2008], analizzando l’impatto delle nuove tecnologie ICT sulla struttura competitiva dei settori economici hanno verificato che la rivoluzione tecnologica delle ICT ha modificato solo marginalmente la struttura dei settori industriali. I confini verticali ed orizzontali delle imprese sono cambiati, c’è stato un certo turnover nel club delle grandi imprese, ma non si osserva una crisi delle grandi imprese a favore di quelle più piccole, e più specializzate, che verosimilmente coincidono in buona parte con la categoria delle imprese più giovani. Una delle spiegazioni plausibili di questa apparente contraddizione risiede nella capacità che hanno mostrato le grandi imprese che hanno abbracciato la filosofia dell’Open Innovation [Chesbrough, 2008], di saper integrare al loro interno asset tecnologici e competenze sviluppate al loro esterno, magari da giovani start-up. Un indizio in questo senso viene portato da Hockerts e Wustenhagen [2010], che, analizzando in particolare i cambiamenti in corso nel settore delle energie rinnovabili, hanno evidenziato che nell’ambito di queste nuove tecnologie svolgono un ruolo portante sia gli “Emerging Davids” (ovvero le start-up tecnologiche), sia i “Greening Goliath” (ovvero le vecchie imprese capaci di innovarsi). I rapporti fra nuove e vecchie imprese sono quindi più complessi e vanno oltre il paradigma sostituito/sostituto, con le nuove imprese technology-based che possono giocare un ruolo più indiretto di agenti di cambiamento e di ringiovanimento tecnologico del sistema produttivo preesistente [Autio e Yli-Renko, 1998]. Il ruolo di ringiovanimento dei Big Incumbents è documentato anche nell’ambito specifico delle tecnologie smart grid [Erlinghagen e Markard, 2012]. Questi autori evidenziano una triangolazione fra Grandi imprese del settore ICT, nuove start-up ICT e Grandi imprese delle vecchie Utilities. Ad un ingresso nelle tecnologie smart grid delle grandi imprese ICT, le Grandi utilities hanno infatti risposto accelerando i processi di acquisizione di nuove start-up specializzate nelle tecnologie ICT. Questo processo prosegue tuttora e trae linfa anche da nuovi e sempre più sofisticati sistemi di scouting dei progetti di nuove start-up tecnologiche. Incubatori, Acceleratori, Start cup Competition, Seed Capital sono soltanto alcuni dei filoni di questo nuovo attivismo dei Big Players per attrarre giovani start-up. Per fermarci al contesto italiano Enel Lab e Working Capital promosso da Telecom Italia sono alcuni degli esempi più interessanti in questo campo.

Quando si parla, però, di innovazioni radicali che nascono quindi sui confini della ricerca scientifica entra in gioco un ulteriore attore che diventa fondamentale, ovvero il sistema della Ricerca pubblica rappresentato dalle Università e dai Centri di ricerca pubblici. Il loro impegno nella ricerca scientifica sulle aree di frontiera e nel trasferimento tecnologico diventa determinante.

L'impegno di questi soggetti si può misurare attraverso una serie di indicatori, quali il numero di pubblicazioni scientifiche in un determinato ambito scientifico/tecnologico, ma soprattutto il numero di brevetti e di spin-off attivati a valle del lavoro svolto dai gruppi di ricerca. Per ciò che concerne le pubblicazioni scientifiche, il rapporto pubblicato dall'Istituto per la Competitività per il settore delle Smart Grid rileva per il 2012, 199 pubblicazioni a livello mondiale, a fronte di 86 pubblicazioni nel 2011 e di sole 18 pubblicazioni nel 2010 [I-Com, 2013]. L'Italia, nell'anno 2012, con le sue 8 pubblicazioni, si attesta nella media in questo campo, in forte crescita rispetto al dato del 2011 (solo 3 pubblicazioni). Da notare che all'interno delle smart grid, la ricerca italiana si è focalizzata sull'importante settore dello *Smart Metering*: l'Italia nel 2012 con 6 pubblicazioni (erano 2 nel 2012) si attesta come primo Paese seguita da Germania e Cina. Analizzando invece il numero di brevetti, l'Italia accusa un significativo ritardo. Nel 2012 risultano complessivamente 171 brevetti a fronte di 96 nel 2010, 34 nel 2009 e solo 10 brevetti nel 2000. L'Italia nel 2012 presenta solo 2 brevetti che sono di provenienza lombarda e registrati da aziende. Per quanto riguarda il ruolo degli spin-off accademici ICT-based, attraverso una ricerca on desk condotta principalmente mediante consultazione dei siti web di 967 spin-off universitari, sul totale di 1.082 censiti dal Netval nel 2012 [Netval, 2013] si è osservato in primo luogo che circa il 30% di tali spin-off hanno matrice di provenienza ICT. Su questo gruppo di spin-off si è proceduto ad una analisi più approfondita per verificare se operano in almeno una delle aree classificate dall'OECD come aree di applicazione delle tecnologie ICT al settore Smart Grid. La classificazione OECD distingue 5 macroaree: Generation, Transport, Storage, Retail, Consumption [OECD, 2011]. Sulla base di questa catalogazione sono state censite complessivamente 44 imprese spin-off con tecnologie e competenze applicabili ad almeno una di queste macroaree. All'interno di questo gruppo, 27 spin-off pur possedendo le potenzialità per operare nel settore smart grid al momento sono attive solo in settori affini. 17 imprese invece dichiarano di essere già operative in aree applicative delle tecnologie ICT alle smart grid. Le aree specifiche di attività dei 17 spin-off ICT già impegnati nelle smart grid sono riportate nella Fig. 2.

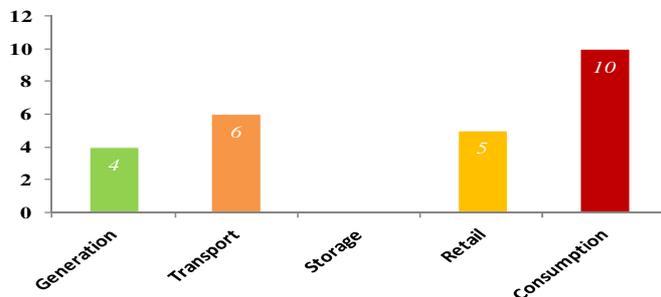


Fig. 2 – Aree di attività degli spin-off nelle smart grid

Come si evince dalla figura, l'area più presidiata è quella "Consumption", con 10 aziende presenti. Da rilevare che diversi spin-off operano contemporaneamente in più aree di business.

4. Conclusioni

Il mondo della ricerca è dunque fortemente coinvolto nella produzione di nuove conoscenze scientifiche e tecnologiche da applicare al settore delle smart grid, come testimonia il volume crescente di pubblicazioni in questo campo, ma sembrerebbe, almeno in Italia, ancora poco interessato al tema della valorizzazione economica considerata la scarsa propensione alla brevettazione. Se ci si fermasse a questi dati ne verrebbe però fuori un'immagine distorta e poco realistica. Molto spesso, infatti, la valorizzazione delle nuove conoscenze scientifiche e tecnologiche prodotte dai gruppi accademici viene perseguita direttamente tramite la costituzione di spin-off accademici che poi evidentemente nel corso della loro operatività procedono alla eventuale brevettazione del proprio know how. Il numero di spin-off ICT-based in Italia operanti nel settore smart grid è, infatti, significativo in termini assoluti e se comparati con il volume di pubblicazioni scientifiche e, soprattutto di brevetti Universitari prodotti. Il fiorire, anche in Italia, di spin-off che si propongono di applicare tecnologie ICT-based al settore smart grid si presta ad una chiave di lettura che fa riferimento alle difficoltà legate al trasferimento tecnologico diretto verso i cosiddetti "incumbents" [Cerrato et al, 2011]. Questa chiave di lettura mette in luce la difficoltà degli incumbents a gestire cambi di paradigmi tecnologici che sono "competence destroying" e che minano quindi alle fondamenta, gli assetti organizzativi e tecnologici che sono stati alla base del successo nel passato, e per converso la maggiore capacità delle new ventures ad adattarsi ai cambiamenti repentini che si verificano all'interno di rugged competitive landscapes [Levinthal, 1997] quale è, appunto, il caso delle smart grid technologies. Vi sono rilevanti indizi, però, per ritenere che le start-up ed in particolare gli spin-off accademici possano rappresentare per gli incumbents più "smart" delle opportunità per ampliare/modificare la propria base di competenze e conoscenze utili a dominare il nuovo paradigma tecnologico che sta emergendo con la smart grid.

Bibliografia

Autio, E., Yli-Renko, H., New technology-based as agents of technological rejuvenation. *Entrepreneurship & Regional development*, 10, 1998, 71-92.

Bellifemine, F. L., Borean, C., De Bonis, R., Smart grids: Energia e ICT. *Notiziario Tecnico Telecom Italia*, Anno 18, Num. 3, 2009.

Cerrato, D., Parente, R., Petrone, M., La collaborazione tra università e industria: un'indagine sui brevetti co-generati in Italia. *L'Industria*. Vol. A. XXXIII, N.2, 2012, 255-282.

Chesbrough H., *Open. Modelli di business per l'innovazione*, a cura di A. Di Minin, tradotto da R. Merlini, Egea, Milano, 2008.

Christensen C.M., *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business School Press, Boston, MA, 1997.

Denrell, J., Fang, C., Winter, S., The economics of strategic opportunity. *Strategic Management Journal*, 24, 2003.

Congresso Nazionale AICA 2013

Dolata, U., Technological innovations and sectoral change: transformative capacity, adaptability, patterns of change: an analytical framework. *Research Policy*, 38, 2009, 1066–1076.

Dosi, G., Gambardella, A., Grazzi, M., Orsenigo, L., Technological revolutions and the evolution of industrial structures. *Capitalism and society*, 3, 2008.

Erlinghagen, S., Markard, J., Smart grid and the transformation of the electricity sector: ICT firms as potential catalyst for sectoral change. *Energy Policy*, 51, 2012, 895–906.

EU-European Commission, Smart Grid projects in Europe: Lessons learned and current developments, Report by the Joint Research Centre of European Commission, 2013.

European Technology Platform (ETP) Smartgrids, Vision and Strategy for Europe's Electricity Networks of the Future, Directorate-General for Research, Bruxelles, 2006.

Geels, F.W., From sectoral systems of innovation to socio-technical systems: insights about dynamics and change from sociology and institutional theory. *Research Policy*, 33, 2004, 897–920.

Geels, F.W., Raven, R., Non-linearity and expectations in niche-development trajectories: ups and downs in Dutch biogas development (1973–2003). *Technology Analysis & Strategic Management*, 18, 2006, 375–392.

Geels, F.W., Schot, J., Typology of socio technical transition pathways. *Research Policy*, 36, 2007, 399–417.

Geels, F.W., Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study, *Research Policy*, 31, 2002, 1257–1274.

Hannan, M., J. Freeman, The population ecology of organizations. *American Journal of Sociology*, 82, 5, 1977, 929-964.

Hill, C.W.L., Rothaermel F.T., The performance of incumbent firms in the face of radical technological innovation. *Academy of Management Review*, 28, 2, 2003, 257–274.

Hockerts, K., Wüstenhagen R., Greening Goliaths versus emerging David. Theorizing about the role of incumbents and new entrants in sustainable entrepreneurship. *Journal of Business Venturing*, 25, 5, 2010, 481-492.

Istituto per la competitività, Rapporto I-Com 2013 sull'innovazione energetica, 2013.

Levinthal, D., Adaptation on rugged landscapes. *Management Science*, 7, 1997.

Netval, X Rapporto Netval sulla valorizzazione della ricerca Pubblica Italiana, Seminiamo Ricerca per Raccogliere Innovazione, 2013.

Nicholls-Nixon, C., Jasinski, D., The blurring of industry boundaries-an explanatory model applied to telecommunications. *Industrial and Corporate Change*, 4, 1995, 755–768.

OECD, ICT Applications for the Smart Grid: Opportunities and Policy Implications, OECD Digital Economy Papers, N.190, OECD Publishing. <http://dx.doi.org/10.1787/5k9h2q8v9bnl-en>, 2012.

Schaltegger, S., Wagner, M., Sustainable entrepreneurship and sustainability innovation. *Business Strategy and Environment*, 20, 2011, 222-237.

Investigating the Effect of Social Media on Trust Building in Customer-Supplier Relationships

Fabio Calefato, Filippo Lanubile, Nicole Novielli
Università degli Studi di Bari Aldo Moro,
Dipartimento di Informatica, Via Orabona 4, 70125 Bari (BA)
{fabio.calefato, filippo.lanubile, nicole.novielli}@uniba.it

Abstract. *Trust is a concept that has been widely studied in e-commerce since it represents a key issue in building successful customer-supplier relationships. In this sense, social software represents a powerful channel for establishing a direct communication with customers. As a consequence, companies are now investing in social media for building their social digital brand and strengthening relationships with their customers. In this paper, we present the preliminary results of an ongoing research aimed at investigating the role of social media in the process of trust building, with particular attention to the case of small-medium enterprises. Our findings show that social media contribute to build more affective trust than traditional websites, suggesting that social media have the potential to enhance the business of SMEs other than large companies, by fostering the affective commitment of customers.*

Keywords: Trust Building, Social Media, Empirical Studies, Human Factors, Social Computing.

1. Introduzione

Negli ultimi anni, l'uso del Web ha influenzato significativamente la comunicazione interpersonale, grazie alla diffusione capillare dei social media. Twitter, Facebook e LinkedIn sono solo alcune tra le più famose piattaforme di social networking online e sempre più numerose sono le aziende che investono oggi in strategie di marketing, sfruttando i social media per diffondere l'immagine associata al proprio marchio nonché per costruire e rafforzare le relazioni di fiducia con i clienti nuovi e abituali [Blanchard, 2011]. Ciò vale anche per le piccole-medie imprese (PMI), che possono trarre beneficio dalla viralità della comunicazione su social media come versione contemporanea del tradizionale passaparola [Tvesovat e Kouznetsov, 2011].

Costruire rapporti di fiducia con i propri clienti e fornire un'immagine affidabile della propria azienda sono entrambe questioni cruciali in ambito commerciale [Büttner e Göritz, 2008], soprattutto nel campo dell'e-commerce, in

Congresso Nazionale AICA 2013

cui la dimensione della comunicazione faccia a faccia è inesistente. Il motivo del successo dei social media come luogo per l'implementazione di strategie di marketing ha origine dal successo dei classici paradigmi di marketing. Le piattaforme di social networking, infatti, restituiscono alle aziende operanti sul Web la possibilità di gestire i rapporti con i clienti in modo personale e personalizzato. Questo è fondamentale poiché la ricerca ha dimostrato come spesso la fiducia sia stabilita nei confronti di uno specifico venditore, piuttosto che rispetto al marchio [Doney e Cannon, 1995]. In questo senso, il personale gioca un ruolo chiave nel relazionarsi con i clienti quando è in grado di fare ricorso a strategie di persuasione basate su argomentazione sia razionale sia emotiva [Petty e Cacioppo, 1986]. I social media consentono ai fornitori di realizzare questo comportamento in un ambiente virtuale in modo da fornire ai clienti la percezione di un'azienda più raggiungibile e preoccupata delle loro necessità [Blanchard, 2011]. Una PMI che miri ad attrarre nuovi clienti fornendo un'immagine affidabile di sé e a fidelizzare i propri clienti di lunga data, dovrebbe prendere in considerazione la possibilità di sfruttare il canale affettivo come strategia chiave di persuasione. In tal senso, i social media offrono la possibilità di implementare forme di interazione che simulano i tradizionali scambi faccia a faccia con i clienti, migliorando l'immagine dell'azienda in termini di affidabilità e benevolenza [Blanchard, 2011].

In questo articolo descriviamo i risultati preliminari di uno studio in corso volto a comprendere il ruolo dei social media nella costruzione di un rapporto basato sulla fiducia tra PMI e consumatori. In particolare, focalizziamo l'attenzione sull'impatto delle diverse modalità web (sito web tradizionale vs *fanpage* su social media) sulla fiducia affettiva, ossia la fiducia costruita sulla base della valutazione di argomenti e informazioni che fanno leva sulla sfera emotiva.

2. Stato dell'arte e motivazione

La fiducia è stata oggetto di studio in numerosi ambiti applicativi [Rusman et al, 2010], dalle scienze cognitive [Castelfranchi e Falcone, 2000] all'economia [Doney e Cannon, 1995], e più recentemente nell'ingegneria del software [Al-Ani e Redmiles, 2009; Schumann et al, 2012]. Secondo Hung et al. [2004], il riporre fiducia in qualcuno implica la convinzione che il fiduciario si comporterà secondo le nostre attese. Per quanto riguarda il marketing, diverse sono le definizioni di fiducia elaborate dai ricercatori. In particolare, il processo di costruzione della fiducia si sviluppa principalmente lungo diverse dimensioni che possono essere identificati come antecedenti della fiducia o *'trust antecedent'* [Rusman et al, 2010], vale a dire, le proprietà del fiduciario che innescano la valutazione cognitiva finalizzata a misurarne l'affidabilità. Coerentemente con il punto di vista dei ricercatori nel campo dell'e-commerce [Büttner e Göritz, 2008] [Dooney e Cannon, 1995], il modello adottato in questo studio è un adattamento del *Tripod Model* [Mayer et al, 1995] che definisce l'attendibilità di una persona o di un'organizzazione in termini di *abilità, benevolenza e integrità*.

L'*abilità* consiste nella capacità del fiduciario di adempiere ad un compito o un'obbligazione, rispondere a una richiesta, fornire una risposta competente. È valutata in funzione delle capacità professionali dell'altro, delle sue conoscenze e delle sue competenze specifiche. Ne sono un esempio, la descrizione dell'esperienza lavorativa di un professionista nel suo curriculum o le informazioni sulla storia professionale dell'azienda e del suo staff, solitamente fornite a scopo informativo sul sito web della società. La *benevolenza* si riferisce a caratteristiche al livello di cortesia, atteggiamento positivo, disponibilità, intenzione di condividere le informazioni e/o le risorse, disponibilità ad aiutare, gentilezza e ricettività. Una persona o azienda che soddisfi questi requisiti è solitamente percepita come incline ad accontentare le esigenze della propria clientela e a fornire supporto in caso di necessità. Per *integrità* si intende un insieme di norme morali e le caratteristiche del fiduciario considerate auspicabili dalla morale comune come, per esempio, integrità, onestà, correttezza, lealtà e discrezione.

Per quanto riguarda il dominio commerciale, McKnight et al. [1998] estendono questo modello includendo la quarta dimensione di *prevedibilità* del comportamento del fiduciario. La prevedibilità è un concetto legato a quello di responsabilità (*accountability*) enunciato da Rusman et al [2010], ossia la capacità di una persona (l'azienda, in questo caso) di soddisfare le aspettative dell'acquirente, nel dominio di riferimento, in termini di affidabilità e consistenza rispetto agli standard dichiarati. Questa quarta dimensione è rilevante in ambito e-commerce [Büttner e Göritz, 2008] e dunque considerata nel presente studio.

Rispetto al processo di costruzione della fiducia, è possibile evidenziare la differenza tra fiducia cognitiva (*cognitive trust*) e fiducia affettiva (*affective trust*) [McAllister, 1995]. La fiducia cognitiva si basa su meccanismi di valutazione razionali delle caratteristiche del fiduciario oggettivamente riconducibili al dominio di riferimento, come ad esempio le competenze professionali. Tale valutazione viene solitamente combinata con il processo di ponderazione dei vantaggi e dei rischi derivanti dal fidarsi dell'altro. Al contrario, la fiducia affettiva si manifesta in presenza di legami affettivi e di sincera preoccupazione per il benessere altrui [Hung et al, 2004].

Schumann et al. [2012] definiscono un'operazionalizzazione di questa distinzione mappando gli antecedenti della fiducia rispetto alle due dimensioni, cognitiva e affettiva. Manterremo tale corrispondenza in questo articolo, in riferimento a tutti e quattro gli antecedenti discussi finora, come rappresentato in Fig. 1. In particolare, l'abilità e la prevedibilità sono valutate per mezzo di elaborazione cognitiva delle informazioni personali e professionali del fiduciario, rilevanti rispetto al dominio di riferimento. Al tempo stesso, la valutazione lungo la dimensione affettiva porta alla costruzione della fiducia in presenza di elementi che supportano la percezione dell'altrui benevolenza e integrità.

In questo studio analizziamo come la valutazione di informazioni a disposizione sul fiduciario sia influenzata dai diversi mezzi di comunicazione online, con riferimento alle dimensioni cognitiva e affettiva. In particolare, ci riferiamo ai tradizionali siti web come *information oriented*, ossia come orientati principalmente a veicolare contenuti informativi. Al contrario, i social media permettono di osservare il comportamento di un'azienda mentre interagisce con

i propri clienti. Accedendo a un profilo aziendale su social network, i clienti possono valutare come lo staff risponde a critiche o commenti negativi, verificare che si prenda attivamente cura della relazione con i clienti, e valutare quanto il personale e i titolari si dimostrino aperti e disponibili a essere contattati. Come evidenziato da ricerche sul marketing [Blanchard, 2011], i social media possono essere utilizzati per simulare con successo una relazione uno-a-uno e amplificare la percezione di ricevere cure e attenzioni personali dedicate, come tipicamente avviene nei mercati tradizionali in cui la componente di interazione faccia a faccia è ancora presente.

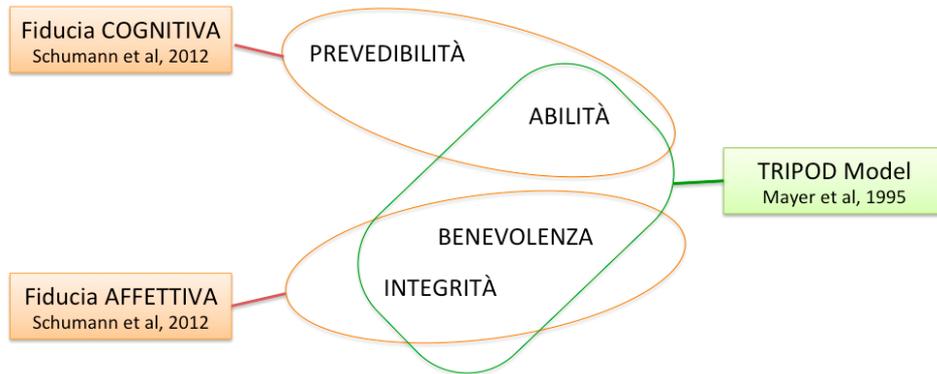


Fig.1 –Dimensione cognitiva e affettiva nel modello di costruzione della fiducia.

3.Obiettivi

3.1 Dominio

Come evidenziato da recenti ricerche, la volontà di fidarsi è inversamente proporzionale al rischio intrinseco associato al dominio di riferimento. Un esempio tipico è l'acquisto di prodotti tecnologici [de Ruyter et al, 2001] o farmaceutici [Büttner et al, 2006]. La decisione di fidarsi di un'azienda dipende anche dal rischio percepito rispetto alla classe di prodotto specifico, oltre che dalle caratteristiche del venditore e dallo specifico contesto di acquisto [Büttner e Göritz, 2008]. Nel nostro esperimento, abbiamo selezionato due piccoli ristoranti poiché per l'utente comune è possibile valutare l'affidabilità di simili aziende in conformità a semplici criteri basati buon senso senza che la valutazione richieda alcuna particolare abilità o conoscenza specifica. Le aziende A e B sono, rispettivamente, una pasticceria e un ristorante biologico. Entrambe hanno iniziato la loro attività circa un anno fa e attuano strategie di marketing basate su una forte presenza su web e social media.

3.2 Ipotesi

Nel nostro modello, assumiamo che la fiducia cognitiva sia determinata dalla valutazione positiva delle informazioni professionali, solitamente acquisite attraverso il sito web dell'azienda. Al contrario, i social media possono contribuire allo sviluppo di un'interazione con il cliente più informale e amichevole, rispondendo ai commenti e alle domande degli utenti sulle community online (ad esempio, sulla fanpage dell'azienda su Facebook). Tale potenziale dei social media può essere impiegato a favore dall'aumento della fiducia lungo la dimensione affettiva. Ipotizziamo, dunque, che la possibilità di monitorare e rispondere adeguatamente e in tempo reale al feedback della *community* online offra all'azienda la possibilità di migliorare la propria immagine in termini di benevolenza e integrità. In altre parole, ipotizziamo che i siti web tradizionali, orientati a comunicare contenuti informativi, e i social media, strutturati in modo da favorire l'interazione diretta con i clienti, possano avere un impatto diverso, rispettivamente, sulla fiducia cognitiva e affettiva.

Di conseguenza, formuliamo due ipotesi che, rispetto alla creazione di rapporti di fiducia tra azienda e nuovi potenziali clienti sulla base della prima impressione ricavata dall'osservazione del profilo aziendale online:

H_{aff} – *I social media favoriscono la fiducia affettiva più dei siti web tradizionali.*

H_{cog} – *I siti web tradizionali favoriscono la fiducia cognitiva più dei social media.*

4. Lo studio

Lo studio è progettato secondo uno schema 2 x 2 (Tab. 1). La modalità Web (*Sito web tradizionale vs Profilo su social media*) costituisce l'unica variabile indipendente, mentre l'azienda (*A vs B*) è considerata semplicemente un fattore di blocco poiché quello che intendiamo indagare è l'effetto della modalità web sulla fiducia. L'esperimento ha coinvolto 19 studenti di laurea specialistica in informatica (16 maschi, età media = 25 anni). Su una scala Likert a 4 punti, tutti i partecipanti tranne uno hanno dichiarato di utilizzare Facebook su base giornaliera (media = 3.74). La maggior parte di loro ha inoltre dichiarato di utilizzare spesso Internet per acquisti (media = 2.58). E' dunque ragionevole supporre che il campione presenti un livello omogeneo di familiarità con il Web e con il meccanismo di valutazione dell'affidabilità di un'azienda attraverso la sua presenza online tipico dell'e-commerce. L'ordine di visualizzazione di sito web e pagina Facebook per ognuno delle due aziende è stato randomizzato per evitare condizionamenti imputabili alla sequenza di presentazione.

Ogni partecipante ha valutato l'affidabilità percepita di entrambe le aziende rispetto a una delle due possibili combinazioni (Tab. 1). La fiducia è valutata da ciascuno dei partecipanti rispetto a tre degli antecedenti nel modello: abilità, prevedibilità e benevolenza. Coerentemente con la letteratura [Schumann et al, 2012], la valutazione dell'integrità è esclusa poiché è un antecedente peculiare di relazioni a lungo termine. Nel nostro studio, invece, i

partecipanti non conoscevano le aziende prima dell'esperimento: sarebbe impossibile fornire una valutazione dell'integrità in un simile contesto in cui la non conoscenza delle aziende costituisce un prerequisito per la partecipazione all'esperimento.

Per la valutazione, è stato definito un questionario [Calefato et al, 2013] ottenuto integrando le linee guida e gli elementi inclusi nel questionario di studi sulla fiducia basata sulla prima impressione [Büttner e Göritz, 2008; Rusman et al 2010, de Ruyter et al, 2001]. Per ogni antecedente è stato definito un insieme specifico di domande: 7 per abilità, 3 per prevedibilità, 11 per benevolenza (questionario basato su scala Likert da 1 a 5). La valutazione è associata alle due dimensioni cognitiva e affettiva, secondo quanto discusso nella Sezione 2.

Modalità web x Azienda	<i>A</i>	<i>B</i>
<i>Sito web tradizionale</i>	Gruppo 1	Gruppo 2
<i>Profilo su social media</i>	Gruppo 2	Gruppo 1

Tab.1 – Progettazione dello studio

4.4 Procedimento

L'esperimento si è svolto in ambiente controllato: tutti i partecipanti sono stati coinvolti simultaneamente sotto la supervisione dello stesso sperimentatore. Lo sperimentatore ha introdotto e spiegato l'esperimento ai partecipanti contemporaneamente, illustrando lo scenario e fornendo istruzioni dettagliate sulla sequenza di esecuzione delle attività. Lo sperimentatore è rimasto in laboratorio durante lo svolgimento dell'esperimento per garantire che i partecipanti non interagissero né scambiassero opinioni sulle due aziende, nonché per rispondere alle loro domande durante l'esperimento. Ogni studente ha lavorato in modo indipendente, accedendo alla procedura sperimentale attraverso un'apposita interfaccia guidata sul web.

All'inizio dell'esperimento, i partecipanti hanno risposto a un breve questionario introduttivo volto a valutare il loro grado di familiarità con le tecnologie web, i social network e l'e-commerce. In seguito, è stato loro introdotto il task dell'esperimento ossia la scelta di una delle due aziende proposte come fornitore per l'allestimento di un buffet. Lo scenario della ristorazione è stato scelto poiché presenta un fattore di rischio più elevato rispetto alla semplice scelta di un ristorante per un pranzo individuale, poiché nello scenario riguardante il catering è in gioco una somma di denaro più consistente e che l'eventuale insuccesso di un buffet presenta necessariamente anche delle implicazioni sociali rispetto agli invitati all'evento.

Una volta presentato lo scenario, i partecipanti sono stati invitati a visualizzare e valutare i due profili aziendali attraverso il sito web dell'azienda o il suo profilo di Facebook, secondo l'ordine di presentazione casuale descritto in precedenza. I partecipanti hanno esplorato il profilo aziendale per non più di cinque minuti. Di conseguenza, ciascuna delle due fasi di valutazione è durata al massimo quindici minuti, includendo visualizzazione del profilo online e compilazione del questionario. Pre-condizione necessaria per la partecipazione

all'esperienza era la non conoscenza da parte dello studente né del sito web né del profilo Facebook di nessuna delle due aziende. Al termine dell'esperienza, i partecipanti sono stati sollecitati a esprimere pareri e opinioni in modo non strutturato nel corso di una discussione collettiva con lo sperimentatore.

5. Risultati

Le due ipotesi H_{aff} and H_{cog} sono state valutate eseguendo un'analisi multivariata della varianza (MANOVA) sulle due variabili dipendenti, ossia la fiducia affettiva e la fiducia cognitiva. Il test ha rivelato che l'interazione tra la modalità web e l'azienda ha un effetto significativo sul livello di fiducia affettiva stabilita sulla base della prima impressione ricevuta dai partecipanti osservando i profili online delle aziende ($F = 4.528$, $p = .041$), come mostrato dal grafico di interazione in Fig. 2a. Al contrario, nessun effetto significativo è stato osservato sulla fiducia cognitiva assumendo un livello di significatività pari a .05, né per l'interazione tra modalità web e azienda ($F = .035$, $p = .853$) né per gli effetti principali di modalità web ($F = 1.590$, $p = .216$) e azienda ($F = .977$, $p = .330$) (Fig. 2b).

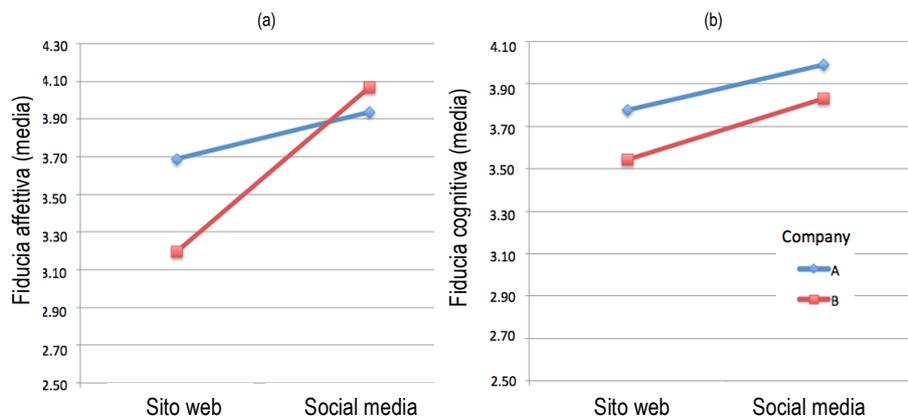


Fig. 2. Effetto dei fattori su trust affettivo (a) e cognitivo (b)

Per meglio comprendere la causa dell'interazione tra modalità web e azienda sulla fiducia affettiva, abbiamo eseguito due ANOVA separate per ognuna dei due ristoranti coinvolti nell'esperienza. I risultati sono riportati in Tab. 2 e mostrano come l'uso di social media abbia un impatto significativo sulla fiducia affettiva solo per l'azienda B ($F = 22.152$, $p = 0.000$), per la quale osserviamo un valore medio di fiducia affettiva pari a 3.20 (Dev. Std. = .42) nella modalità sito web, per poi salire a 4.07 (Dev. Std. = .38) nella modalità social media. Al contrario, per l'azienda A non riscontriamo nessuna differenza statisticamente significativa tra fiducia affettiva osservata a fronte della visualizzazione di sito web (media = 3.69, Deve. Std. = .60) e profilo Facebook (media = 3.94, Deve. Std. = .35).

Azienda	Effetto della modalità web	
	F	P
A	1.211	0.287
B	22.152	0.000

Tab. 2 ANOVA sulla fiducia affettiva per le due aziende

5.1 Discussione

Abbiamo testato due ipotesi, rispettivamente H_{aff} (*i social media esercitano un maggiore impatto sulla fiducia affettiva*) e H_{cog} (*i siti web tradizionali esercitano un maggiore impatto sulla fiducia cognitiva*). Rispetto alla fiducia affettiva, i risultati supportano parzialmente l'ipotesi H_{aff} suggerendo che effettivamente il ricorso al web marketing attraverso i social media contribuisce a creare un'immagine più aperta e benevolente dell'azienda presso i potenziali nuovi consumatori che ne visitano il profilo online, già sulla base del primo impatto. L'analisi della varianza ha evidenziato che tale impatto è statisticamente significativo solo per una delle due aziende considerate nello studio. L'interazione tra la modalità web (variabile indipendente) l'azienda (fattore di blocco) suggerisce la necessità di controllare, nelle future repliche dell'esperimento, l'effettiva equivalenza sia delle informazioni fornite sia della strategia di comunicazione attuata dalle due aziende.

Rispetto alla fiducia cognitiva, i risultati non supportano l'ipotesi H_{cog} poiché non rileviamo nessun impatto significativo da parte dei siti web delle due aziende sulla costruzione della fiducia rispetto alla dimensione cognitiva. Tuttavia, interessanti spunti di riflessione sono stati forniti dai partecipanti nel corso del colloquio collettivo e informale con lo sperimentatore, a conclusione dell'esperimento controllato. La maggior parte dei partecipanti ha infatti dichiarato che il ricorso a fanpage e community su social media, per quanto concepite per favorire l'interazione informale con gli utenti, può favorire una migliore valutazione dell'affidabilità dell'azienda anche rispetto alla dimensione cognitiva. Infatti, i partecipanti dichiarano di valutare come più competenti le aziende che deliberatamente si espongono direttamente al feedback dei clienti sui social media, poiché tale comportamento è percepito come indice di maggiore credibilità nonché come asserzione di maggiore qualità nei beni e servizi forniti.

Inoltre, la maggior parte degli studenti coinvolti ha dichiarato di individuare nella disponibilità di foto (del ristorante, dei prodotti e dello staff) un elemento fondamentale per la valutazione della competenza e professionalità dell'azienda. Ciò è coerente con la letteratura sul rapporto tra fotografie e fiducia [Riegelsberger et al, 2013] [Steinbrueck et al, 2002] e sull'effetto che l'accesso agli elementi di informazione normalmente associati alla dimensione affettiva (ad es. foto del personale) può avere anche lungo la dimensione cognitiva [Schumann et al, 2012].

6. Conclusioni e sviluppi futuri

Nel presente studio abbiamo analizzato il ruolo dei social media nella costruzione della fiducia lungo la dimensione affettiva, in un contesto in cui potenziali clienti visualizzano per la prima volta le informazioni dell'azienda visitandone i profili disponibili sul web. Abbiamo inoltre indagato il potenziale del social software per le piccole aziende che vogliono costruirsi un'immagine online basata su affidabilità e competenza.

Abbiamo osservato come, consistentemente con le ipotesi iniziali, i social media detengano il potenziale necessario per incrementare la percezione di disponibilità e apertura rispetto alle specifiche esigenze dei clienti. Includendo il ricorso ai social media nella propria strategia di web marketing, un'azienda può simulare interazioni personali e individuali attraverso gli strumenti offerti dalle più comuni piattaforme di social networking online.

I risultati di questo studio sono da considerarsi preliminari e richiedono successiva validazione attraverso future repliche dell'esperimento che coinvolgano un campione di partecipanti più ampio e vario e introducano un maggiore controllo degli elementi informativi e dell'effettiva equivalenza delle strategie di comunicazione attuate dalle aziende selezionate.

Ringraziamenti

La presente ricerca è finanziata dal progetto Intersocial, nell'ambito del Programma di Cooperazione Transfrontaliera Grecia-Italia 2007-2013. Ringraziamo tutti gli studenti che hanno partecipato allo studio.

Bibliografia

Al-Ani, B. and Redmiles, D. In Strangers We Trust? Findings of an Empirical Study of Distributed Teams. Proc. 4th Int'l Conf. on Global Software Engineering (ICGSE '09), 2009

Blanchard, O. Social Media ROI, 2011. Pearson Education.

Büttner, O. B., and Göritz, A. S. Perceived trustworthiness of online shops. Journal of Consumer Behaviour 7, 2008, 35-50

Büttner, O. B., Schulz, S. and Silberer, G. Perceived Risk and Deliberation in Retailer Choice: Consumer Behavior towards Online Pharmacies. Advances in Consumer Research, 33, 2006, 197-202.

Calefato, F., Lanubile, F., Novielli, N. A Preliminary Investigation of the Effect of Social Media on Affective Trust in Customer-Supplier Relationships. Proc. of Affective Computing and Intelligent Interaction (ACII 2013), to appear.

Castelfranchi, C. and Falcone, R. Trust is more than subjective probability: mental components and sources of trust. Proc. of the 33rd Hawaii International Conference on System Sciences, 2000.

Doney, P. M., and Cannon J. P. An Examination of the Nature of Trust in Buyer-Seller Relationships. Journal of Marketing 61, 1995, 35-51.

Hung, Y.C., Dennis, A.R., and Robert, L. Trust in Virtual Teams: Towards an Integrative Model of Trust Formation. Proc. of the 37th Hawaii International Conference on System Sciences. HICSS '37. IEEE Computer Society, Washington DC, 2004

Mayer, R. C., Davis, J. H., and Schoorman, F. D. An integrative model of organizational trust. *Academy of Management Review* 20, 1995, 709-734.

McAllister, D.J. Affect and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal* 38, 1, 1995, 24-59.

McKnight D.H., Cummings, L.L., Chervany, N.L. Initial trust formation in new organizational relationships. *Academy of Management Review*, 23 (3), 1998, 473-490.

Petty, R. E., and Cacioppo, J. T. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, Springer, N, 1986.

Riegelsberger, J., Sasse, M.A., McCarthy, J.D. Shiny Happy People Building Trust? Photos on e-Commerce Websites and Consumer Trust. Proc. of CHI 2003, 2003.

Rusman, E., van Bruggen, J., Sloep P., and Koper R. Fostering trust in virtual project teams: Towards a design framework grounded in a TrustWorthiness Antecedents (TWAN) schema. *International Journal of Human-Computer Studies* 68, 2010, 834-850.

de Ruyter, K., Moorman, L. and Lemmink, J. Antecedents of Commitment and Trust in Customer-Supplier Relationship in High Technology Markets. *Industrial Marketing Management*, 30, Elsevier Science, 2001, 271-286.

Schumann, J., Shih, P., Redmiles, D., and Horton, G. Supporting Initial Trust in Distributed Idea Generation and Evaluation. Proc. International ACM SIGGROUP Conference on Supporting Group Work (GROUP 2012) ACM, 2012, 199-208.

Steinbrueck, U., Schaumburg, H., Duda, S., and Krueger, T., A Picture Says More Than A Thousand Words - Photographs As Trust Builders In E-Commerce Websites. In. *Proceedings of CHI2002: Extended Abstracts*, 2002, 748-749.

Tvesovat, M. and Kouznetsov, A. *Social Network Analysis for Startups*. O'Reilly, 2011.

Smart Cities, Urban Sensing and Big Data: Mining Geo-location in Social Networks

Daniele Sacco¹, Gianmario Motta², Linlin You³, Nicola Bertolazzo⁴, Chen Chen⁵

Università di Pavia

Via Ferrata 1, 27100, Pavia (PV)

¹ daniele.sacco01@ateneopv.it

² motta05@unipv.it

³ linlin.you01@ateneopv.it

⁴ nicola.bertolazzo01@ateneopv.it

⁵ chen.chen01@ateneopv.it

Location based social networks offer spatio-temporal information which can be accessed through public Application Programming Interfaces (APIs) and drew the interest of researchers with diverse scientific backgrounds. This availability of data enables a potential use of geo-located content as an additional, low cost and infrastructure-less source of information for urban sensing in Smart Cities. All these aspects bounded with the need of real-time analytics for urban sensing takes to Big Data management and its related issues. A systematic literature review outlines related works and gaps in current research. We propose a reference model to exploit Big Data and Open Data for urban sensing and we validate it by a case study. Finally, we give recommendations for future research about location and mobility mining of social network data.

Keywords: big data, urban sensing, smart city, location-based social network, twitter, data mining, open data

1. Introduction

Last years have seen a paradigm shift of the role of the user on the web, from consumer to producer of information. Due to last technological developments and their world-wide dissemination, Web 2.0 changed users' approach to information creation and exploitation. The paradigm shift has been also a social shift. Web has become an essential need for people, a sort of commodity, that is used to communicate, interact, share information and even maintain relationships thanks to social networks.

The last advent of smartphones, equipped with GPS sensors allowing users to geo-locate themselves, can take to the next shift, from a social and collaborative Web 2.0 to a local and mobile Web 3.0. One among the first achievements has been the integration of Geographic Information Systems

(GISs) and social networks resulting in new location-based capabilities. Social networks that include location information into their contents are called Location Based Social Networks (LBSNs). So, LBSNs offer spatio-temporal information which can be accessed through public Application Programming Interfaces (APIs) and draw the interest of researchers with diverse scientific backgrounds. This availability of data enables a potential use of geo-located content as an additional, low cost and infrastructure-less source of information for urban sensing in Smart Cities.

All these aspects bounded with the need of real-time analytics for urban sensing takes to Big Data management and its related issues. Real-time urban sensing use citizens as active and passive sensors and can reveal important insights of human behavior in the city. Diverse use scenarios can enable new perspectives in society level, e.g. community healthcare, public safety, city resource management and transportation management.

2. Systematic Literature Review

Systematic Literature Review (SLR) can be regarded as “explicitly formulated, reproducible and up-to-date summary” [Egger et al, 2008] that includes and extends the statistical results of a meta-analysis methodology. As opposed to narrative reviews, it is based on a structured method that is always explicitly specified at the beginning of the review.

Our objective is to identify initiatives, experiences and viewpoints on location and mobility mining of social network data. So, our research question is “How to exploit geo-located data from social networks and what level of maturity has reached its application?”. So, the expected outcomes of our SLR are: (a) a complete overview of the “state of the art”, (b) the identification of gaps in current research, solutions, trends and future research and suggestions to the community of researchers and practitioners, and (c) recommendations about best practices for location and mobility mining of social network data.

Extracted information contains techniques, issues, models and any other kind of topic useful to provide an accurate snapshot of the current state of location and mobility mining of social network data. 87 out of 109 articles have been selected and classified by year of publication, geographical area, research method and publication channel.

There are no significant articles before 2009 because Location Based Social Networks (LBSNs) raised in the same year. Afterwards researchers could start to consider data provided by LBSNs. The articles considered in our research range from 2009 to second quarter of 2013. Considering that publications in 2013 refer only to the first months of the year, the number of publications tends to have an exponential growth (2^x): 2 publications in 2009, 4 in 2010, 14 in 2011, 38 in 2012. Case studies represent the majority of the study types. The number of case studies and instrument development publications (87.3% in total) reflects the experimental approach to the issue. It is also motivated by the low number of theoretical papers (5.7%) and by the lack of position papers.

3. Discussion

We have identified 3 main domains: (a) data sources, (b) technologies, (c) use scenarios. Let us consider the main contributions to each domain.

4.1 Data Sources

The data sources that can be inferred for urban sensing are heterogeneous. Three innovative data sources exist: (a) mobile sensor data about the individual devices, (b) infrastructure sensor data about the context, and (c) social data from social network and other internet services [Zhang et al, 2011]. Data sources can be used independently but the combination of the three kinds can provide a comprehensive understanding of human behavior and its context. Here we focus only on data sources for geo-location from the social network tier and how they have been used in literature.

Twitter is a widely-used platform for the real-time social sharing of short text-based messages called tweets. Twitter and smartphone usage reflected same growth, indeed Twitter users interact frequently on mobile devices. As Twitter is easy to use and interactions are short, many users post tweets despite they are engaged in other activities. This gives Twitter data good spatial and temporal coverage because tweets can be automatically geo-located [Mai et Hranac, 2013]. Twitter provides a free real-time streaming API through which a sample of all tweets can be retrieved. The API provides filters that can be set on these data streams to capture tweets within a geographic area or only those containing certain terms. However, data stream is limited to 1% of total tweet volume. So, only a subset of the total tweets can be used.

Foursquare is a location-based social network where users can “check in” to different locations and share them with friends both on Foursquare itself and other social networks. Users can upload pictures at a venue or leave ‘tips’ on the venue page (e.g. a user may ‘check-in’ to a hotel and leave a ‘tip’ about how bad the service is) [Cheng et al, 2011]. Foursquare check-in data is not directly accessible: however, users typically decide to share their check-ins publicly on Twitter, so they can be retrieved via Twitter streaming API.

Several other papers use data sets published for research purposes, because no API is available or the social network stopped its service. For example, Gowalla was a location-based social network created in 2009. The concept behind the service was to advertise your exact location to all your friends in real-time [Scellato et al, 2010]. Also BrightKite was a social networking website where users could share their location, to post notes and to upload photos. By making “check-ins” at places, users could see people who is nearby. Now, only already collected data sets are available [Li et Chen, 2009].

Other available services are Momo and Flickr. However, a question could rise: why not to use Facebook, the most popular service? The main reason is that Facebook API can be used to retrieve data from those users who accepted to publish their posts to your application or system, so it is not publicly available.

4.2 Technologies

We discuss here the main technologies used for urban sensing. The SLR shows that related works focus on machine learning techniques.

K-means can be used to reveal clusters of common behaviors across land segments. The land use of each cluster can be derived by analyzing the activity vectors of the regions within the cluster. K-means depends on the initial random selected seeds and it needs to specify the number of clusters k (land uses) to identify [Frias-Martinex et al, 2012]. [Lee et al, 2011] used a similar approach, but they also formed a Voronoi diagram using the center points (latitude, longitude) of the K-means results and regarded the formed regions as a set of region of interests, to identify the occurrence of local events.

A Self-Organizing Map (SOM) is an unsupervised neural network that reduces the input data dimensionality to be able to represent its distribution as a map. So, SOM forms a map where similar samples are mapped close together. [Frias-Martinex et al, 2012] used SOM to build a map that segments the urban land into geographical areas with different concentrations of tweets in the time period under study.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) has specific characteristics: (a) it is based on the concept of density reachability, producing satisfying results identifying arbitrarily shaped clusters, (b) the number of clusters is not given a priori, and (c) the algorithm tolerates noise, allowing for some data points not to be assigned to any cluster [Villatoro et al, 2013]. The advantage of density based clustering algorithms is that clusters are defined by the density of data points and not by spatial size and form of cluster.

Spectral methods for data clustering are popular because of the quality of the clusters that are produced and the simplicity of implementation. Spectral clustering is able to find arbitrarily shaped clusters and does not pose any constraints on them (in contrast to the k -means, for example, which assumes cluster to be convex). It requires anyway the parameter k to define the number of desired clusters [Roslrer et Liebig, 2013].

Mean-shift is a non-parametric clustering technique that detects the modes of an underlying probability distribution from a set of discrete samples. So, mean-shift can be used both as an algorithm to detect local maxima (modes) as well as a clustering technique (areas associated to the modes). [Frias-Martinex et al, 2012] assume that there exists an unobservable underlying probability distribution of where people tweet from. The modes of that distribution are determined to represent urban landmarks or points of interest in the city.

4.3 Use scenarios

The identification of use scenarios aims to enable new perspectives in society level, e.g. community healthcare, public safety, city resource management and transportation management..

We here propose a classification for visualization objectives in urban sensing, representing an evolution of existing Business Intelligence (BI) solutions towards Geographic Information Systems (GISs) and Big Data visualization [Stodder, 2013], as shown in Tab. 1.

Class	Description	BI similarities	Objectives
Urban characterization	The results can be stored for users as a “snapshot” of a certain point in time. Users examine snapshots to identify changes in data over time, so they must be provisioned and presented consistently so that trends and comparisons are valid	It recalls dashboards. The view is static and it is previously defined by data analysts.	To visualize and predict social ties and urban structure
Spatial discovery	It enables users to interact with data through analytical processes. Visual functionality for filtering, comparing, slicing and dicing, drilling down, and correlating data can then be integrated with the users’ analytical application functions for forecasting, modeling, and statistical, what-if, and predictive analytics.	It recalls On-Line Analytical Processing (OLAP). The view is dynamic and it allows users to navigate data.	To analyze behaviors “on-line”, in space and time
Exception alerting	It notifies users of particularly important changes in the data or when situations arise that demand immediate attention. Alerts mean that something important in real-time data or event streams is happening.	It recalls event processing in modern Business Activity Monitoring (BAM) solutions, that detect and warn of problems or exceptions in real-time.	To detect events or exceptions to standard behaviors, e.g. disasters, diseases, unexpected crowds

Tab. 1 – Visualization objectives classification

The three visualization classes reflect papers analyzed by our SLR. Respectively, 30 papers deal with urban characterization, 18 with spatial discovery, 12 with exception alerting. The rest of papers use social networks geo-located data mainly to build recommender systems or user profiling. These papers are single user oriented, so we do not take them into account in following paragraphs because they do not give a broad view of the city.

4.3.1 Urban characterization

[Rosler et Liebig, 2013] provide insights on the activity profiles in urban environments. Clusters identified by Foursquare check-ins help to describe the socio-dynamics of urban districts in different times of the day. [Ferrari et al, 2011] extend the work on activity profiles by providing also mobility patterns that occur in an urban environment and understanding of social commonalities between people. Traditional municipal organizational units such as neighborhoods are studied by Livehoods project from [Cranshaw et al, 2012], who shows that their boundaries do not always reflect the character of life in these areas. [Joseph et al, 2012] mine check-ins to identify groups of people which are of different types (e.g. tourists), communities (e.g. users tightly clustered in space) and interests, and how they use urban space.

By processing Twitter data, [Wakamiya et al, 2011] are able to examine the relation between regions of common crowd activity patterns and major categories of local facilities. [Wakamiya et al, 2012] extend their work and correlate psychological and geospatial proximity of urban areas by borrowing crowd's experiences from geo-tagged tweets, in order to demonstrate that people often rely on geospatial cognition to the real space than the exact physical distance in the real world.

4.3.2 Spatial discovery

[Silva et al, 2012] study social behaviors by monitoring check-ins in Foursquare, [Sagl et al, 2012] by Twitter and Flickr data. They are able to analyze city dynamics spatially and temporally and to identify seasonality in human behaviors. [Cheng et al, 2011] investigate 22 million check-ins across 220,000 users and report a quantitative assessment of human mobility patterns by analyzing the spatial, temporal, social, and textual aspects associated with these footprints.

4.3.3 Exception alerting

[Boettcher et Lee, 2012] introduce EventRadar, a detection system that finds local events such as release parties, musicians in a park, or art exhibitions. [Watanabe et al, 2011] provide a similar system to detect events. [Baldwin et al, 2012] also try to predict events impact. [Mai et Hranac, 2013] assert that tweets can be matched to traffic incidents by examining the content of the tweets for key words and comparing locations of the tweets and incidents. Results are confirmed for areas with sufficient density of Twitter usage.

4. A Big Data-based approach

Big data is a popular term used to refer to the exponential growth, availability and use of information. However, as Gartner states (<http://www.gartner.com/it-glossary/big-data/>), Big Data doesn't focus only on the high volumes of information, but also on "data velocity", that involves streams of data that are produced and processed in real-time, and on "data variety", that refers to more types of information to analyze, e.g. social media, context aware data, documents, images, videos, audio, etc.

All papers in our review obviously consider "data variety" because location-based social networks' data is geo-located and unstructured (typically JSON files), however the challenge results from the expansion of all three properties given by Gartner, rather than just "data volume" or "data velocity" alone.

We here propose a big data-based approach to mining of social network geo-located data. Our approach intends to consider the 3 properties by applying technologies that fit Big Data management. Our literature review found out no papers take into account big data-based system architectures to process "variety", "velocity" and "volume" of social network data. Here we propose a simple architecture to process this kind of data and possible extensions to support more analytics.

In order to validate our approach, we report a case study about public transport monitoring. First we developed a simulator that transforms timetables from open transport data in Torino (Italy) to a visualization tool that shows planned position of each bus in a specific time (each moving point in Fig. 1). We decided to enrich this visualization and leverage urban sensing. Our specific question was: “how to correlate transport planning and urban activity areas?”. Exploitation of location-based data from social network is a viable way.

To define activity areas in a city we built a tool for spatial discovery that could allow us to model crowds by density clusters within the city and drill down them according to different scales. In order to keep the case study simple, our approach has been to compute density on a fixed grid applied over the map (rectangles with same dimension). However, what kind of data could be used? We exploited Twitter streaming API, collected geo-located data in Torino area in real-time and clustered them on the map for each specific time range (1 hour).

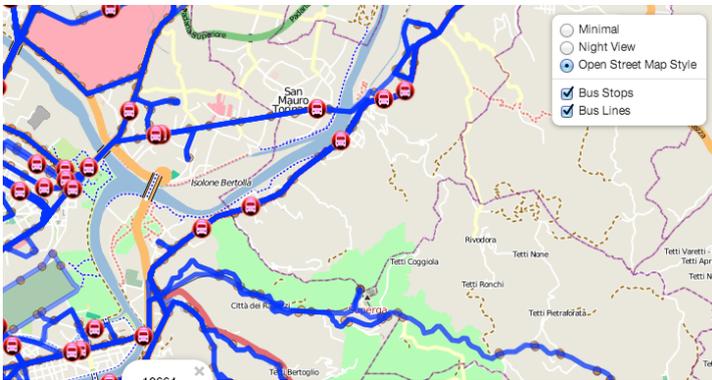


Fig. 1 – Public transport simulator

To retrieve Twitter data we implemented FluenTD agents by Node.js on a virtual machine in Amazon Elastic BeanStalk. FluenTD (<http://fluentd.org/>) is an open-source log collector that enables to have a logging architecture with more than an hundred types of systems, by treating logs as JSON. Node.js (<http://nodejs.org/>) is a platform built on JavaScript for easily building fast, scalable network applications. It uses an event-driven, non-blocking I/O model that makes it lightweight, efficient, and oriented to data-intensive real-time applications that run across distributed devices. FluenTD and Node.js integration allows to build real-time collection and filtering of geo-located tweets, and their storage in TreasureData (<http://www.treasure-data.com/>). TreasureData is a “Big Data as a Service” cloud solution that offers a time series, columnar, Hadoop-based, schema-free data warehouse stored on Amazon S3. It allows you to access data using Hive query language (<http://hive.apache.org/>) by JDBC. An Extraction-Transformation-Loading (ETL) process access data in TreasureData, processes it and uploads it in CartoDB every 5 minutes. CartoDB is a “database as a service” cloud solution, based on a PostgreSQL database with GIS extension. As PostgreSQL databases can use

extensions to run data mining algorithms, we decided to apply density-based clustering online, whenever the user changes the zoom scale in the map. Fig. 2 shows the final architecture for our prototype.

As this solution intends to implement Service Oriented Architecture (SOA), it has two main advantages: (a) it can be easily extended, and (b) a component can be replaced by others because of decoupling. For example, the data stored in the data mart can be used to perform different analytics or to build an alerting system. As SAP Hana has a Predictive Analysis Library that offers native support to DBSCAN, we could use SAP Hana to implement it as a supplementary layer between TreasureData and CartoDB. If we want to easily implement Latent Dirichlet Allocation we could replace TreasureData with Mahout on a Hadoop cluster. The ETL process, that may represent a bottleneck in terms of performances, can be replaced by ad hoc solutions, e.g. interfaces developed in Node.js to provide data filtering and storage in the data mart.

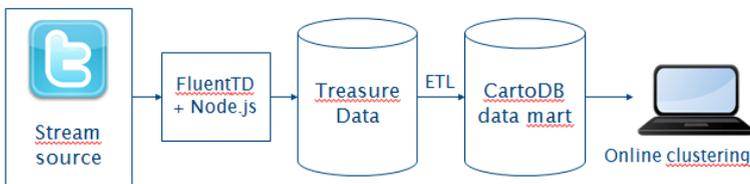


Fig. 2 – A Service Oriented Architecture for Big Data management

Fig. 4 shows early results of our prototype. It considers tweets collected from 8 AM to 9 AM on July 10th, 2013. Color density of rectangles change according to tweet density in the same area.

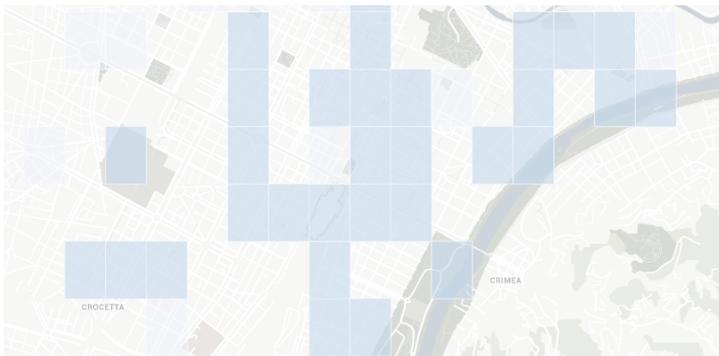


Fig. 4 - Tweet density in Torino

Final integration between the crowd modeling tool and the public transport simulator allows spatial discovery of urban areas not covered by transport service during peaks of presence of people who may need to move from one place to another within the city.

5. Conclusion

Our SLR demonstrated the exponential growth in number of papers about this trending topic. The high number of case studies and instrument development publications reflects the experimental approach to the issue and it may also lead developers and freelancers to provide smart solutions to contextual problems in the city.

Most of related works focus on validation of data mining techniques, rather than their application in real use scenarios. Our prototype demonstrated that the integration of urban sensing and open data can help municipalities to reveal enhanced insights about their services to citizens.

So, future research should move from validation of techniques to their real application in Smart Cities by exploiting Big Data technologies, thus providing real time analytics to municipalities and end users.

Our future works intend to integrate different data sources, such as sensors, and more social media, in order to provide deeper insights for urban sensing. Furthermore, our system needs to scale up and validate performances against user needs. Next step will be the development of a service orchestration layer to provide complete validation of our Service Oriented Architecture.

References

- [Baldwin et al, 2012] Baldwin, T., Cook, P., Han, B., Harwood, A., Karunasekera, S., & Moshtaghi, M. (2012, April). A support platform for event detection using social intelligence. 13th Conference of the European Chapter of the Association for Computational Linguistics (pp. 69-72).
- [Boettcher et Lee, 2012] Boettcher, A., & Lee, D. (2012, November). EventRadar: A Real-Time Local Event Detection Scheme Using Twitter Stream. In Green Computing and Communications (GreenCom), IEEE International Conference on (pp. 358-367).
- [Chen et al, 2013] Chen, T., Kaafar, M. A., & Boreli, R. (2013). The Where and When of Finding New Friends: Analysis of a Location-based Social Discovery Network. International Conference On Weblogs And Social Media
- [Cheng et al, 2011] Cheng, Z., Caverlee, J., Lee, K., & Sui, D. Z. (2011). Exploring Millions of Footprints in Location Sharing Services. ICWSM, 2011, 81-88.
- [Cranshaw et al, 2012] Cranshaw, J., Schwartz, R., Hong, J. I., & Sadeh, N. M. (2012, June). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. In ICWSM.
- [Egger et al, 2008] Egger, M., Smith, G. D., & Altman, D. (Eds.). (2008). Systematic reviews in health care: meta-analysis in context. Wiley.
- [Ferrari et al, 2011] Ferrari, L., Rosi, A., Mamei, M., & Zambonelli, F. (2011, November). Extracting urban patterns from location-based social networks. 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks (pp. 9-16)
- [Frias-Martinez et al, 2012] Frias-Martinez, V., Soto, V., Hohwald, H., & Frias-Martinez, E. (2012, September). Characterizing Urban Landscapes using Geolocated Tweets. 2012 IEEE International Conference on Social Computing (pp. 239-248).

[Joseph et al, 2012] Joseph, K., Tan, C. H., & Carley, K. M. (2012, September). Beyond local, categories and friends: clustering foursquare users with latent topics. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 919-926)

[Lee, 2012] Lee, C. H. (2012). Mining spatio-temporal information on microblogging streams using a density-based online clustering method. *Expert Systems with Applications*, 39(10), 9623-964

[Lee et al, 2011] Lee, R., Wakamiya, S., & Sumiya, K. (2011). Discovery of unusual regional social activities using geo-tagged microblogs. *World Wide Web*, 14(4), 321-349.

[Li et Chen, 2009] Li, N., & Chen, G. (2009, August). Analysis of a location-based social network. *IEEE International Conference on CSE'09* (Vol. 4, pp. 263-270)..

[Mai et Hranac, 2013] Mai, E., & Hranac, R. (2013, January). Twitter Interactions as a Data Source for Transportation Incidents. In *Transportation Research Board 92nd Annual Meeting* (No. 13-1636).

[Rosler et Liebig, 2013] Rösler, R., & Liebig, T. (2013). Using Data from Location Based Social Networks for Urban Activity Clustering. In *Geographic Information Science at the Heart of Europe* (pp. 55-72). Springer International Publishing.

[Sagl et al, 2012] Sagl, G., Resch, B., Hawelka, B., & Beinat, E. (2012). From Social Sensor Data to Collective Human Behaviour Patterns: Analysing Spatio-Temporal Dynamics in Urban Environments. *GI-Forum: Geovisualization, Society and Learning*.

[Scellato et al, 2010] Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010, June). Distance matters: geo-social metrics for online social networks. In Proceedings of the 3rd conference on Online social networks (pp. 8-8). USENIX Association.

[Silva et al, 2012] Silva, T. H., Melo, P. O., Almeida, J. M., Salles, J., & Loureiro, A. A. (2012, November). Visualizing the invisible image of cities. In *Green Computing and Communications (GreenCom), 2012 IEEE International Conference on* (pp. 382-389)

[Stodder, 2013] Stodder, D. (2013). *Data Visualization and Discovery for Better Business Decisions*. TDWI Best Practices Report. TDWI Research

[Villatoro et al, 2013] Villatoro, D., Serna, J., Rodríguez, V., & Torrent-Moreno, M. (2013). The TweetBeat of the City: Microblogging for Discovering Behavioural Patterns during the MWC2012. *Citizen in Sensor Networks* (pp. 43-56). Springer Berlin

[Wakamiya et al, 2011] Wakamiya, S., Lee, R., & Sumiya, K. (2011, November). Crowd-based urban characterization: extracting crowd behavioral patterns in urban areas from twitter. *ACM SIGSPATIAL Workshop on LBSNs* (pp. 77-84)

[Wakamiya et al, 2012] Wakamiya, S., Lee, R., & Sumiya, K. Measuring Crowd-sourced Cognitive Distance between Urban Clusters with Twitter for Socio-cognitive Map Generation. *International Conference On Emerging Databases*

[Watanabe et al, 2011] Watanabe, K., Ochi, M., Okabe, M., & Onai, R. (2011, October). Jasmine: a real-time local-event detection system based on geolocation information of microblogs. *ACM conference on Information and knowledge management*

[Zhang et al, 2011] Zhang, D., Guo, B., & Yu, Z. (2011). The emergence of social and community intelligence. *Computer*, 44(7), 21-28.

Thinking BigData: Motivation, Results and a Few Recipes for a Balanced Growth of High Performance Computing in Academia

P. Margara, N. Nepote, E. Piccolo, C. G. Demartini and P. Montuschi
Department of Control and Computer Engineering - Politecnico di Torino
C.so Duca Degli Abruzzi 24, 10129 Torino (TO)

[paolo.margara, nicolo.nepote, elio.piccolo, claudio.demartini, paolo.montuschi]@polito.it

Big Data is today both an emerging research area and a real present and future demand. High Performance Computing (HPC) Centers cannot neglect this fact and have to be reshaped to fulfill this need. In this paper we share our experience of upgrading a HPC Center at Politecnico di Torino, originally designed and deployed in 2010. We believe that this issue could be common to some other existing "general purpose" HPC centers where, at least in the short term, the possibility to start from scratch a new Big Data HPC center cannot be afforded but a balanced upgrade of the existing system has to be preferred.

Keywords: Big Data, High Performance Computing, HPC, MapReduce, Hadoop, Parallel Systems, Academia.

1. Introduction

A possible solution to survive the economic crisis and keep the global competition in all fields, is to develop large-scale projects. From crash simulation to disaster and climate modeling, from ethical clinical trials to drug discovery, from [Graphene](#) to [The Brain Project](#), from nano-technologies to smart cities. All these studies share commonalities: the demand for fast, complex computation and the need to analyze huge datasets.

So far, these two aspects have been treated separately, leading both to the development of parallel algorithms running on supercomputers, and to the birth of data mining which exploits large DBMS. Today we need instead a greater integration between the tools for processing massive amounts of data and the supporting underlying computation architecture. The key to this breakthrough has a name: Big Data. Differences are more and more emerging between large amounts of data and Big Data, both in terms of correlation with physical or natural events that generate them, and also for the semantics that are assigned to them: it is a complete reversal of perspective. In other words, the first has a background model, while the second is assumed to be able to derive different models, even in many different fields [Kindratenko & Trancoso, 2011].

Dealing with Big Data has therefore become a stimulating challenge for many small-to-medium sized supercomputing centers, also in Academia. For three years our High Performance Computing (HPC) center at [Politecnico di Torino](#) has provided support to research groups by means of a standard cluster widely described in [Della Croce et al. 2011] [Nepote et al. 2013]. So far, the demands focused on rough computing power rather than Big Data analysis, but now the time has come to take up the challenge.

We have taken this new demand both as a challenge and as an opportunity, and 6+ months ago have started a project of upgrading our HPC center to efficiently run also tasks massively using Big Data.

At the beginning we figured out that this request could be addressed by installing some new packages followed by proper tuning of some parameters. The results of this first experimental phase have been very negative but have stimulated us to explore further and think at more structural solutions, by entering each specific problem with the aim to find feasible solutions besides the trivial one to start from scratch a completely new Big Data HPC center.

Today we can proudly say that, given the small resources available, we have not only met our goals, but also obtained a set of specifications for future balanced upgrades of our HPC. As we believe that this issue could be common especially nowadays as Big Data is emerging as a new and real requirement, we have decided to share our choices and experience with other Readers both to get their feedback and to transmit what we have learned during this enhancement of our HPC system.

In the following sections we outline the steps taken to make our HPC system able to run MapReduce [Dean & Ghemawat, 2008] tasks using [Apache Hadoop](#), the results obtained, and the considerations that have followed.

2. The CASPER cluster at PoliTO

With [HPC@POLITO](#) we refer to the initiative boosted in 2010 to start a “supercomputing” center at Politecnico di Torino, and aimed at providing computational resources and technical support to both academic research and university teaching. Over the years, the computing center has set up a general-purpose and campus-wide available and dynamically evolving cluster called CASPER (Cluster Appliance for Scientific Parallel Execution and Rendering), now serving several research groups operating in different areas [Della Croce et al. 2011] [Nepote et al. 2013].

CASPER is a Linux-based MIMD distributed shared memory InfiniBand heterogeneous cluster, reaching 1.3 TFLOPS with its 136 cores and 632 GB of overall main memory. As the majority of clusters, in 2010 CASPER was designed to focus more on the need to perform several computations on relatively small amount of data possibly stored together with the Operating System, inside a small local hard disk of each computational node of the architecture. The role of shared resource to store all experimental data, especially when it was too much large for the local disks, was played by a Network Attached Storage (NAS).

Over the years, this initial configuration (Figure 1) has mostly evolved towards the increasing of the local node computational power, as the applications used were more computation-intensive than data-intensive, i.e. they were very rarely were using big amount of data.

In Figure 1 it is depicted the current structure of CASPER when observed from the viewpoint of the general job scheduler. It can be noted that the frontend node provides all the storage, as local disks are only used to maintain cache and temporary files. In CASPER communications occur under the umbrella of [InfiniBand](#) connection, which is fast and with low latency linking all the nodes in the cluster as well as coexisting with the Ethernet network. To run custom code or third-party software, including also the widely used [Message Passing Interface](#) (MPI) libraries, CASPER is normally operated through the industry standard job scheduler and resource manager [Oracle GridEngine](#) (formerly known as Sun GridEngine).

Several research activities have been supported by HPC@POLITO, where, during the past three years 30+ research projects and 20+ scientific papers have benefited of CASPER computational resources.

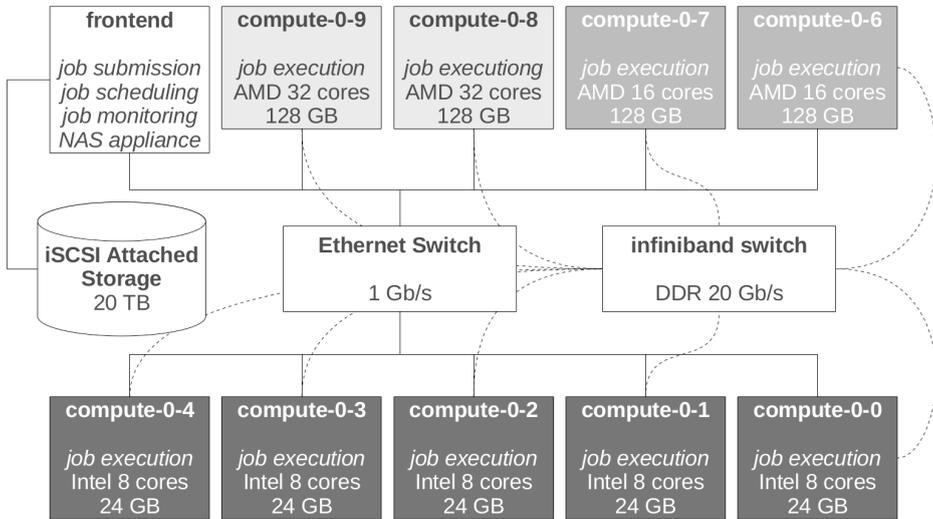


Figure 1: The CASPER cluster configuration (mid 2013) from SGE's point of view.

3. Addressing the BigData problem

3.1 Motivation and our choice

During the last academic year, we have experienced the starting of a new need: the capability to perform analysis of large amounts of data has been more and more often required by CASPER users. Due to budget restrictions, we have immediately discarded the option of a complete rejuvenation and/or replacement of our cluster with another one specifically designed for Big Data. Therefore, we have decided to study the feasibility of “upgrading” the current configuration to tackle also Big Data. We have started by verifying if an implementation of a framework for Big Data processing would perform well inside a general purpose HPC system like CASPER and, if so, what changes to the general architecture would have been necessary to achieve better performances. We believe that this issue could be common and therefore we are here to share our choices and experience with other Readers both to get their feedback and to transmit what we have learned.

After long discussions, some pre-tests and a deep analysis of advantages/drawbacks/tradeoffs/costs, we have converged to a first-experimental solution adopting the Apache Hadoop open source framework, which exploits the well known *MapReduce* [Dean & Ghemawat, 2008] paradigm for the efficient implementation of data-intensive distributed applications. Besides these characteristics, Hadoop is also well known to: run very well on large clusters of commodity hardware, provide a framework possibly minimizing the impact of hardware failures and exploit local computation and storage.

3.2 General Remarks

Before starting the deployment, we spent some more time to analyze the current configuration of the cluster to identify and possibly quickly address critical issues of a straight implementation of Hadoop on CASPER. We focused our attention on the main components, i.e. storage, CPU & Memory, and Network. In the following we have listed our main findings and issues.

- **Storage**

Fact: Hadoop accesses memory by using its own distributed file system, called HDFS. To prevent storage from becoming the bottleneck, it is necessary to properly harmonize the configuration of the physical storage and the way it is accessed by the software.

Problem: We observed that, being a distributed file system, HDFS uses the local storage resources of the individual computational nodes. However, usually HPC general purpose clusters (like CASPER), are based on centralized storage systems (such as a NAS). In such cases, the computational nodes of the cluster are equipped with low performance very basic local hard disks, very poorly suitable to efficiently store and handle any big data. In Figure 2 we observe that in CASPER the amount of local disk space available for each CPU has even decreased during the years. Other HPC cluster configurations have even diskless computational nodes bootstrapping from the network and using a small ramdisk for local operations.

Our Solution: An easy way to workaround the problem before installing Hadoop, is to add more local hard disks on each computational node for the exclusive use of HDFS.

- **CPU and memory**

Fact: Hadoop exploits CPU and memory especially during the Map phase of the MapReduce process. At this stage, many tasks are executed in parallel by all the Mappers processes running on each computational node. For correct operation, all CPUs in the cluster should have similar performances. If not, the slower ones would represent a bottleneck. In other words the cluster should be as much symmetric (or homogeneous) as possible.

Problem: The last nodes (from compute-0-6 to compute-0-9 in Figure 1) added to the initial configuration of CASPER were acquired as part of "fellowship" agreements [Della Croce et al. 2011] with some research groups. The needs of these groups have therefore driven the growth of CASPER, basically towards three different types of computational nodes (Figures 1 and 3), while at the same time aiming the RAM/CPU ratio to the initially target value of 4 GB (Figure 2). The outcome of this growth process is that the four most recent nodes are characterized by more CPUs operating at lower frequencies than these of the first five nodes, thus resulting into a current configuration of CASPER which is highly heterogeneous.

Our Solution: This problem is difficult to solve without radically changing the current configuration of the cluster. A quick, but limiting, workaround could be to exclude the slower computational nodes from the pool used by Hadoop. We will discuss more effective solutions in our "Recipe list" in section 3.5.

- **Network**

- *Fact:* HDFS commonly relies on IP for node-to-node communication. Much attention should be paid while choosing the type of Layer 2 network on which IP traffic flows, in order to ensure high performances of HDFS when transferring blocks of data between nodes.

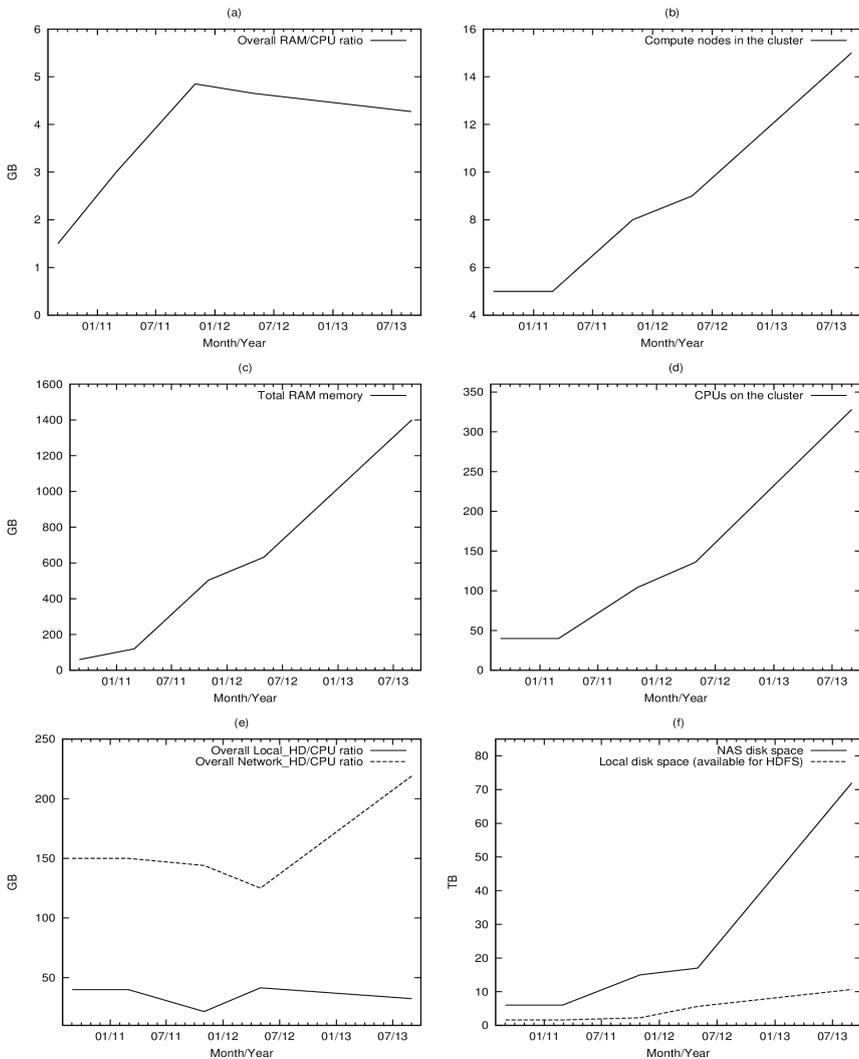


Figure 2: Temporal evolution of some hardware features of the CASPER cluster

Problem: CASPER has a total of three Layer 2 networks carrying different information. A InfiniBand DDR 20 Gb/s network carries all the MPI node-to-node data transfer during standard jobs (i.e. not using Hadoop) execution. A dedicated 4x1 Gigabit Ethernet aggregate link connects the central NAS storage to the nodes via NFS. Finally, a Gigabit Ethernet (Gbe) network still remains as a supporting layer 2 for carrying the remaining IP traffic, usually related to system management communications or data generated by monitoring tools. Unfortunately HDFS traffic falls under the slower Gbe network.

Our Solution: The only possible solution that we have found to be effective in our environment without doing a complete redesign of the cluster, is to implement an aggregated link also on each slave node.

3.3 Hadoop deployment and tuning

The installation of Apache Hadoop [Garza et al. 2013] was made by trying to harmonize its needs to those of our specific HPC system. From the operating system viewpoint, CASPER is an installation of ROCKS Cluster Distribution 5.4.31. Therefore, to install Hadoop 1.0.4 we used the packages from the EPEL repository for [CentOS](#), as derived by ROCKS. As it can be observed from Figure 3, one of our main choices has been to use the frontend node of the cluster as master node of Hadoop and the computation nodes as slave nodes.

Targeting to performances improvements HDFS was configured with permissions disabled, data block size set to 256MB and number of replicas per data block set to 2. On the first five nodes of the cluster (from *Slave 0* to *Slave 4* in Figure 3), we added one single 1TB 7200 rpm SATA-3 local hard disks for the exclusive use of HDFS. On the remaining four nodes (the most recent) this was not possible because of lack of physical space in the case and therefore the available storage for HDFS is only 160GB.

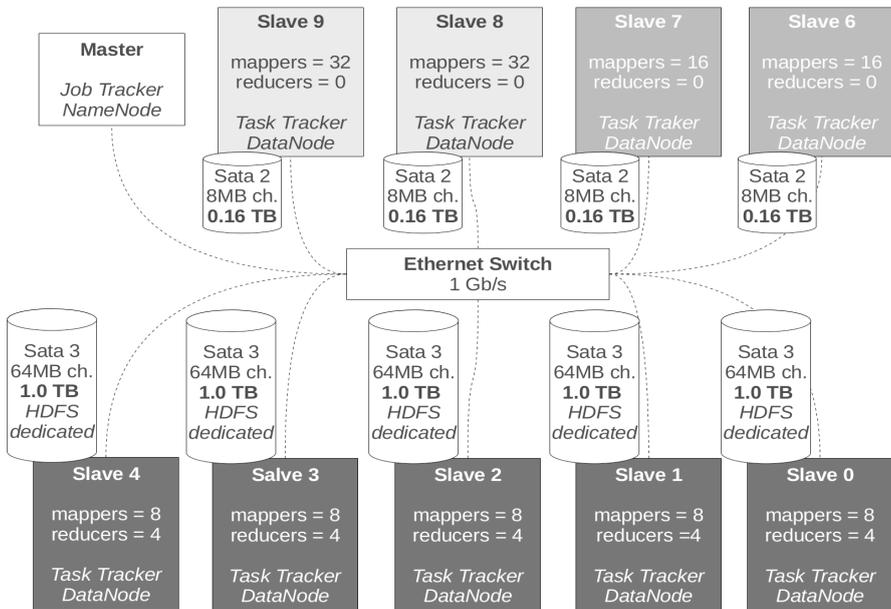


Figure 3: The CASPER cluster configuration (mid 2013) from Hadoop's point of view.

We then configured Hadoop so as to meet the heterogeneous nature requirements of CASPER. Specifically, the maximum number of *mappers* for each node was set equal to the number of cores, while the maximum number of *reducers* was set to 4 for the five nodes with a large and efficient 1TB local disk, and to 0 for the other four nodes, in order to comply with potential decreases of the overall file system performances due to the slow

efficiency and small size and their hard disks.

3.4 Benchmarking the system

We run a set of experiments based on Hadoop to evaluate its performances and scalability on CASPER, in order to possibly identify which upgrades could be useful for CASPER to become an efficient system for Big Data. In particular, we run one of the most widely used and available benchmarking tests to analyze efficiency and scalability on Hadoop, i.e. the Hadoop-based implementation of [TeraSort](#). For the purpose of running TeraSort in different boundary conditions, we used two datasets of sizes 100GB and 200GB respectively.

As it is well known that for Hadoop, asymmetry equals inefficiency, we decided to evaluate TeraSort this two different configurations of CASPER:

- The **SomeNodes** test configuration is based solely on the 5 first nodes of the cluster, labeled in Figure 3 as *Slave 0* to *Slave 4*. It is the only possible homogeneous configuration that we can build in CASPER using more than two nodes. It relies on 5 Intel 3.2GHz nodes with 1TB of secondary memory per node, accounting a total of 40 CPUs. The 1TB disk was added before deploying Hadoop to possibly enhance HDFS performances.
- The **AllCluster** test configuration is basically an extension (or an enhanced version) of *SomeNodes* exploiting all the resources of the cluster. It is designed to squeeze every drop of power from CASPER, but it is also extremely heterogeneous. It is based on 136 CPUs with different frequencies (3.2 GHz and 2.4 GHz) and local disks with size ranging from 160GB to 1TB. This configuration has +240% more CPUs than *SomeNodes*, therefore the expected execution time decrease when running TeraSort should be -58%.

The results reported in Table 2 show that TeraSort can process in less than 4 hours a 200GB dataset and that, even on a cluster that was not specifically designed for Hadoop, it is possible to process files larger than those manageable by traditional parallel frameworks like MPI [Garza et al. 2013].

However, the time scored by the *SomeNodes* configuration is only slightly worse than the one scored by the *AllCluster* configuration. The *AllCluster* configuration decreases the execution time by only -13% when the file size is 200GB and -31% when the file size is 100GB, thus showing that the extra 4 non-homogeneous nodes appear not to contribute in a relevant way to the overall computations.

Name	CPUs	Tot. Memory	Tot. HDFS size	Mappers	Reducers
SomeNodes	40	120 GB	1.6 TB	40	20
AllCluster	40+96	632 GB	5.64 TB	136	20

Table 1: Test configurations chosen for TeraSort

Dataset	Test Configuration	Execution Time
100 GB	SomeNodes	1h23m2s
	AllCluster	57m26s
200 GB	SomeNodes	3h52m57s
	AllCluster	3h21m49s

Table 2: TeraSort execution time on CASPER

We have identified some possible reasons for this result on the *AllCluster* configuration:

- Having set to zero the number of *reducers* on the most recent slaves without additional 1TB disk, has allowed us to not degrade the performance of HDFS, but at the same time has forced the same slaves to transfer all the data to the first 5 slaves having at least 4 *reducers* (see Figure 3) . This generates a quantity of traffic on the Gigabit Ethernet network so high that the advantage given by having more than double the mappers is almost completely zeroed.
- Although *AllCluster* involves several additional CPUs, their operating frequency is too low: 2.4 Ghz versus 3.2 Ghz of the nodes of *SomeNodes*. This has the effect of slowing the overall *map()* phase of the MapReduce paradigm, thus showing that this configuration is probably too much heterogeneous.

3.5 Some recipes for a better Hadoop implementation on general purpose HPC

Based on the results that we have obtained, it can be observed that a slightly “upgraded” CASPER, and in general also possibly any similar general-purpose architecture, can potentially be used to run also the new Hadoop-based applications. The key is what and how to implement the slightly upgrading specially targeted to improve performances on large datasets.

While integrating Hadoop within a general purpose HPC cluster as CASPER we identified and set up some possible recommendations to achieve integration with good performance at the same avoiding the need to change the architecture of the cluster itself. In particular, we have observed that the main ingredients are, as expected, **Network**, **Distributed Storage** and **Computational Power**. Clearly, their combination, interoperability and configuration can really make the difference. Here below our “recipes”, i.e. how to weight, size and combine the different ingredients.

- **Network:** Communication in Hadoop suggest a Gigabit Ethernet connection for the slave nodes, and a faster connection for the master node, such as:
 - An aggregate of several 1 Gigabit Ethernet links. Incidentally, this has been proven to be working well with the current configuration of CASPER.
 - The combination of a faster 10 Gigabit Ethernet link on the master node and some aggregate links on the slaves.
 - An InfiniBand link, if already available in the cluster, implementing the IP-over-InfiniBand stack, which unfortunately implies a complete redesign of the cluster.
- **Distributed storage:** The distributed storage in a Hadoop based cluster is managed through HDFS, which has a master/slave architecture. In HDFS, a cluster consists of: at least one *NameNode*, a master server that manages the file system namespace and controls access to files by clients, and several *DataNodes*, usually one per node in the cluster, each one managing their local node storage. Under this premise, our thoughts are:
 - We strongly advise to setup a RAID on the *NameNode*, since it contains all the metadata needed by HDFS to ensure data consistency. On the contrary, no RAID configuration is required on *DataNodes*, as HDFS already operates

with intrinsic internal data redundancy.

- Regarding the number of disks on the slave nodes, we would like to recommend a configuration with 2 disks, one for the exclusive use of HDFS and another one for storing temporary files needed for processing the tasks that will be assigned to the computational nodes. As in a general purpose HPC cluster the computational nodes could also be diskless, it could be likely that often it is not possible to add more than 1 or 2 additional disks per node due to physical limitations of the individual nodes. In general, in a cluster specifically designed for Hadoop the number of disks recommended for node usually ranges from 2 to a number depending by the number of cores available on the node and the physical space needed/available. In our case we have tested CASPER to exhibit good performances with one additional 1 TeraByte disk, as we did not have more physical space to install more disks.
- As the frequency of communication between *NameNode* and *DataNodes* increases together with the fragmentation of the data (i.e. when several small files are often exchanged), in order to reduce the negative impact of a possibly slow *NameNode*, it is highly recommended to equip the master node with both enough RAM memory and computing power. A fine tuning on “how much”, depends on number and size of files that will be likely stored into HDFS, and can be done by running some practical experiments.
- **Computational power:** Into a Hadoop cluster the jobs are submitted to the MapReduce framework consisting of a single master *JobTracker* and one slave *TaskTracker* per slave node. The master is responsible for: scheduling the tasks of each job on the slaves, monitoring them and re-executing the failed tasks.
 - For the master node we have not identified additional relevant constraints besides these outlined in the previous points.
 - On the other hand, slave nodes require a sufficiently large memory size to keep the processor busy without frequently swapping. Although this is highly dependent on the type of job that will run on the cluster, we have found that satisfactory performances for basic configuration are obtained with slave nodes equipped with 8 cores and 24GB of RAM.
 - As we have seen in the previous sections, good performances require a homogeneous system with at least the same choice of hardware for the slave nodes both in the number of cores and in the frequencies of the processors. In CASPER we have experienced that any asymmetry between the individual nodes can cause re-schedule of tasks and long waiting times between the operations of *map()* and *reduce()*, due to the different execution times of similar tasks scheduled on differently equipped nodes. Unfortunately, in our case these issues should have been addressed both in the design phase of the initial configuration (i.e. when CASPER was designed in 2010) and in its following upgrades, as later (i.e. today) there is no way to make such changes without full replacing at least some of the computational nodes.

Additional points that, based on our experience, we recommend to observe are:

- Future upgrades of the system should comply with scalability and keep the system as most homogeneous as possible.
- Based on our tests, Hadoop seems to achieve better results when the nodes have sufficiently large and efficient local disks. In other words, the choice of the disk's

size and type will be in a trade-off between performances and total size of the distributed file system, based on the footprint that experimental data will have at runtime (i.e. considering intermediate writes).

- One way to avoid a problem of cluster-heterogeneity, like in CASPER, is to exploit the features provided by the general scheduler (see Figure 1). All popular schedulers for HPC systems, in fact, provide the ability to create different execution queues including homogeneous subsets of nodes, therefore creating many “smaller”, more efficient Hadoop clusters.

4. Conclusions and Future Developments

Big Data is an emerging issue and the need to rethink at High Performance Computing to be able also to manage tasks with massive Big Data is not just a research topic, but a real need. At Politecnico di Torino we have taken this new issue both as a challenge and as an opportunity and have started a project of identifying the requirements for a balanced upgrading of our HPC towards Big Data. In this paper we have shared our experience gained while designing and deploying this growth. In particular, we have briefly outlined the steps that we have taken to make our HPC system able to run MapReduce tasks using Hadoop, the results obtained, and shared a set of possibly useful general recommendations.

The plans for the close future of our HPC center include the expansion of CASPER up to 296 CPUs, 1.3 TB of memory and 60 TB of central storage by September 2013. Following the recipes in this paper, we have adopted some measures:

- The system will be kept as homogeneous as possible by adding computational nodes similar to the current ones.
- Each new node will have 2 additional disks for the exclusive use of Hadoop.
- The master node will be connected with a 10 Gigabit Ethernet link.

Finally, although the primary goal is to have a stable and reasonably efficient system in every field of HPC, it is our intention to follow the current trend of Big Data: not only “large data-bases”, but also new algorithms, computational models and applications, starting from NoSQL DBMS like [Apache HBase](#). This will bring the computing center to face new stimulating challenges such as: strengthening the Ethernet network with multiple switches, optimizing the sharing of local resources between HBase and SGE, redounding the master node to make the whole system more fault tolerant.

References

- [Nepote et al. 2013] N. Nepote, E. Piccolo, C. G. Demartini, P. Montuschi, “[Why and How Using HPC in University Teaching? A Case Study at PoliTo](#)”. Proc. of the DIDAMATICA 2013 conference, Pisa, Italy, 2013.
- [Garza et al. 2013] P. Garza, P. Margara, N. Nepote, L. Grimaudo, E. Piccolo, “Hadoop on a Low-Budget General Purpose HPC Cluster in Academia”. Proc. of the ADBIS 2013 Conference - Special Session on Big Data (in press), Geneva, Italy, 2013.
- [Kindratenko & Trancoso, 2011] V. Kindratenko, P. Trancoso, “[Trends in High-Performance Computing](#)”. In: IEEE Computing in Science & Engineering, vol. 13, n. 3, pp. 92-95, Sept. 2011.
- [Della Croce et al. 2011] F. Della Croce, N. Nepote, E. Piccolo, “[A Terascale Cost-Effective Open Solution for Academic Computing: Early Experience of the DAUIN HPC Initiative](#)”. Proc. of the 49th AICA national conference on smart tech and smart innovation, Turin, Italy, 2011.
- [Dean & Ghemawat, 2008] J. Dean and S. Ghemawat, “[MapReduce: simplified data processing on large clusters](#)”. In: Commun. ACM, vol. 51, n. 1, pp. 107–113, 2008.

Integrating Statistical Data with the Semantic Web: The Istat Experience

Raffaella Aracri, Stefano De Francisci, Andrea Pagano,
Monica Scannapieco, Laura Tosco, Luca Valentino
Istituto Nazionale di Statistica – Istat
Via Balbo 16, 00185, Roma, Italia
E-mail:(aracri,defranci,pagano,scannapi,tosco,luvalent)@istat.it

Abstract. *The paper describes an experience about the dissemination of Istat (Italian National Institute of Statistics) data according to Semantic Web technologies, starting from standards adopted in the statistical domain. More specifically, the paper shows the design and implementation issues related to the development of a translator from the SDMX (Statistical Data and Metadata eXchange) data model to the RDF Data Cube vocabulary data model. The effectiveness and efficiency of the translator are validated against real Istat dataset.*

Keywords: Semantic Web, Statistical data, Linked data

1. Introduction

“Linked Data” (<http://linkeddata.org/>) permits to create and interlink arbitrary volumes of structured data across the Web. The Linked Data initiative is made possible by the widespread adoption of Web standards for publishing data according to the Resource Description Framework (RDF) model. RDF allows to uniquely identifying resources on the Web, by means of a specific URI (Uniform Resource Identifier). This feature has several advantages, including (i) the possibility of a direct access to resources via a query language and (ii) the ability to link data together in order to access them in an integrated way (with the clear positive side-effect of higher quality, more information more easily accessed, and so on).

In this paper, we present the first results of the IS-LOD (Istat - Linked Open Data) project. The project started in September 2012, with the general objective of investigating how to integrate the huge amount of data disseminated as Official Statistics with the Semantic Web.

National Statistical Institutes have a consolidated language for representing the statistical data they publish, namely SDMX (Statistical Data and Metadata eXchange) [SDMX-2013]. The focus of the contribution of this paper is on the design and implementation issues related to the development of a “translator” from the SDMX data model to the RDF Data Cube vocabulary data model

[RDF-QB-2013], this latter being the current standardization initiative for statistical data publication based on Semantic Web standards. An extensive validation of the translator against real Istat datasets is a further presented contribution.

2. Background: SDMX Framework and RDF Data Cube Vocabulary

SDMX [SDMX-2013] is an ISO standard, based on XML, designed for the exchange of statistical data. The SDMX Information Model defines the following structures:

- Key Family: ordered list of dimensions, measures and attributes that define the structure of the data set.
- Concepts: descriptor of dimensions, measures and attributes (and optional group key) used in the data set.
- Code Lists: defining the valid content of each of the concepts used in the dataset.
- Dataset: containing data (observations) and related metadata whose content conforms to the specification of the Key Family definition.

The first three structures are grouped into a single data flow named as Data Structure Definition (DSD).

In the following, we will refer to SDMX Standards Version 2.0 which is the version currently used in Istat dissemination architecture.

The RDF Data Cube Vocabulary (RDF-QB) [RDF-QB-2013] is a W3C candidate recommendation (as of 25 June 2013), and is based on SDMX. More specifically, it is built on the SDMX Information Model and on other existing vocabularies listed in Table 1. Like in the other vocabularies, the names of the entities are URIs expressed with a compact notation *prefix:localname* where the prefix identify a URI namespace and its concatenation with the localname gives the complete URI. The prefix to be used in RDF-QB is *qb*.

3. The IS-LOD (Istat- Linked Open Data) Project

Istat is the principal provider of open data among all the Italian public administrations, according to an analysis performed by Formez [Marras-2013]. In particular, Istat provides more than 600 datasets as open data. Such datasets are mainly published on the I.stat web warehouse (<http://dati.istat.it>), from which it is possible to download datasets in CSV, Excel and SDMX formats. The machine-to-machine communication is managed by a dedicated service named Single Exit Point (SEP), that allows immediate access and reusability of the disseminated data. The SEP system is based on the Eurostat SDMX Reference Infrastructure, hence using SDMX for data and metadata descriptions.

However, Istat open data are at level 3, of Tim-Berners Lee five-star stack [Five-Star-2013]. In order to improve Istat open data dissemination towards level 4 (RDF) and level 5 (LOD), in September 2012 we started the IS-LOD project.

The first objective of this project was to investigate how to publish Istat data taking advantage from the standardization work carried out for the publication of

Integrating Statistical Data with the Semantic Web: The Istat Experience
 data according to SDMX model and format. It is notable that data expressed in SDMX come with a large amount of metadata that, if reused, will allow us to publish Linked Data of high quality.

Vocabulary	Namespace (prefix)	Description
SKOS – Simple Knowledge Organization	skos	SKOS gives specifications and standards to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading systems and taxonomies within the framework of the Semantic Web.
Void – Vocabulary of Interlinked Dataset	void	RDF schema vocabulary for expressing metadata about RDF datasets. It is a bridge between publisher and user of RDF data, with application ranging from data discovery to cataloging and archiving of datasets.
FOAF – Friend Of A Friend	foaf	RDF-OWL vocabulary for describing persons, their activities and their relations to other people and objects.
Dublin Core	dc	Basic vocabulary to describe documents.
ORG	org	Core ontology for organizational structures, aimed at supporting linked-data publishing of organizational information across a number of domain.
RDF Vocabulary	rdf, rdfs	A defined set of predicates that can be used in an application

Table 1: RDF-QB Vocabularies

3.1 Translator: from SDMX to RDF Data Cube Vocabulary

As mentioned, our objective was to understand how much and how costly was to approach the semantic Web starting from SDMX. Hence, we decided to design and implement an automated translator from the SDMX model and format to RDF Data Cube Vocabulary.

As a first step, we performed an evaluation of the technological environments that could support the implementation of the Translator. Three different solutions were identified and compared, namely:

- R package RSDMX developed at FAO (http://r-forge.r-project.org/R/?group_id=1298). The package could serve the purpose of supporting reading/writing operations for manipulating SDMX files. However, the package is not complete and not actively maintained.
- Apache Jena (<http://jena.apache.org/>). Jena is a Java framework for reading, processing and writing data in RDF, it also provides a query engine compliant with the latest SPARQL specification. While useful

for RDF manipulation, the framework does not cover our further requirement related to manipulation of SDMX documents.

- XSLT (Extensible Stylesheet Language Transformations) technology: useful for specifying transformations between XML files. We found a successful usage of XSLT transformation within the MIMAS Linked data Project [MIMAS-2011], with an objective similar to ours, i.e. mapping data expressed in SDMX format to RDF-QB format.

This XSLT solution was hence selected for our Translator.

As a second step, we designed the mapping transformations between the elements of the SDMX data model and the elements of the RDF-QB data model, as depicted in Fig.1.

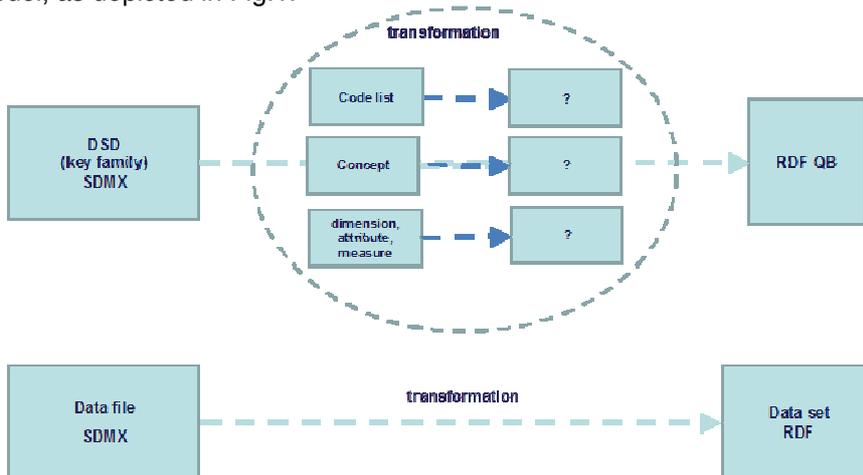


Fig. 1:Translator’s mapping rules

In the following, we show an example of a mapping rule expressed via XSLT.

Let us consider the component REF_AREA stated in the key family section of the DSD related to "Separate collection indicators" of the Istat data on Environment and Energy Waste:

```

<registry:KeyFamilies>
  <structure:KeyFamily id="SEP_AMBIENTE_RIF_INDRACDIFF" version="1.0" agencyID="IT1" isFinal="false">
    <structure:Name xml:lang="en">Environment and energy - Waste</structure:Name>
    <structure:Description xml:lang="en">Separate collection indicators for the provincial capitals
      Separate collection indicators
      Separate collection kilogram per capita by type of waste</structure:Description>
    <structure:Components>
      <structure:Dimension conceptRef="REF_AREA" conceptSchemeRef="SEP_AMBIENTE" conceptSchemeAgency="IT1"
        conceptSchemeVersion="1.0" codelist="CL_REFAREA" codelistVersion="1.3" codelistAgency="IT1" />
    </structure:Components>
  </structure:KeyFamily>
</registry:KeyFamilies>
    
```

The component REF_AREA stated in the key family refers to the CL_REF_AREA code list:

```
<structure:CodeList id="CL_REFAREA" version="1.3" agencyID="IT1" isFinal="false">
  <structure:Name xml:lang="en">Territory</structure:Name>
```

REF_AREA is defined as follows:

```
<structure:Concept id="REF_AREA">
  <structure:Name xml:lang="en">Geographical reference area</structure:Name>
</structure:Concept>
```

The element under consideration, when described in RDF-QB is more verbose as it must be declared not only as a dimension of the Data Cube Vocabulary but also as an SDMX DimensionProperty and as an SDMX CodedProperty as it refers to a specific code list (in this case "territory" code list):

```
<rdf:Description rdf:nodeID="REF_AREA">
  <qb:dimension rdf:resource="http://dati.istat.it/output/code/territory"/>
  <dc:language>en</dc:language>
  <rdf:type rdf:resource="http://purl.org/linked-data/sdmx#DimensionProperty"/>
  <rdf:type rdf:resource="http://purl.org/linked-data/sdmx#CodedProperty"/>
  <sdmx:codeList rdf:resource="http://dati.istat.it/output/code/territory"/>
  <rdfs:range rdf:resource="http://dati.istat.it/output/code/territory"/>
  <rdfs:label xml:lang="en">Geographical reference area</rdfs:label>
</rdf:Description>
```

XSLT and XPATH statements, included in the portion of the stylesheet showed in the following, write the tags rdf:Description and qb:dimension which refer respectively to the Core RDF vocabulary (<http://www.w3.org/1999/02/22-rdf-syntax-ns#>) and to the Data Cube Vocabulary (<http://purl.org/linked-data/cube#>):

```
<xsl:for-each
  select="doc($dsd)//*[local-name()='KeyFamilies']/structure:KeyFamily/structure:Components/*[@conceptRef]">
  <xsl:variable name="conceptRef" select="./@conceptRef" />
  <xsl:variable name="codeList" select="./@codeList" />
  <xsl:variable name="codeListName"
    select="doc($dsd)//*[name()='registry:CodeLists']/*[@id=$codeList]/structure:Name[@xml:lang=$language]" />
  <xsl:element name="rdf:Description" inherit-namespaces="no">
    <xsl:attribute name="rdf:nodeID" select="$conceptRef" />
    <xsl:element name="qb:dimension">
      <xsl:attribute name="rdf:resource" select="concat($istatRoot, '/code/', $codeListName)" />
    </xsl:element>
```

XPATH Instruction to detect all key family's components contained in SDMX DSD file

XPATH Instruction to detect detect the name of the specific component.

XSLT Instruction to write a tag element

3.2 Technological Solution

As described in the previous Section, we used XSLT to convert data expressed in SDMX to RDF-QB. XSLT is a very powerful language xml-based proposed by W3C to apply transformations to XML documents. It is designed to be used as part of XSL, which is a stylesheet language for XML.

In more details, XSL consists of three elements:

- XSL Transformations (XSLT): a language for transforming XML;
- XML Path Language (XPath): an expression language used by XSLT (and many other languages) to access or refer to parts of an XML document;
- XSL Formatting Objects (XSL-FO): an XML vocabulary for specifying formatting semantics.

As XSLT Processor we chose Saxon (<http://saxon.sourceforge.net/>). It provides implementations of XSLT 2.0, XQuery 1.0, and XPath 2.0 at the basic level of conformance defined by W3C. It is available both for Java and .NET. We used Saxon Java Home Edition, the open source version of the tool. This tool version can be used from the operating system command line or can be integrated into other tools and applications.

For the Translator, an XSLT file, even if generalized, needs the following parameters to be set at runtime for each transformation to be executed:

- Language: indicating the reference language, that is the language in which SDMX code lists, DSDs and data are expressed.
- Alternative language: indicating an alternative language in case no language is specified in SDMX files.
- DSD: containing the path and the name of the DSD file.

We developed a standalone Java application that, before carrying out the transformation, performs a preprocessing of the stylesheet setting the necessary parameters. For accessing, manipulating, and outputting stylesheets we have used JDOM, a Java-based "document object model" for XML files (<http://www.jdom.org/>).

Finally, RDF can be serialized in different formats, namely: RDF/XML, Turtle [Turtle-2013], Notation-3 [N3-2013] and N-triples [N-triples-2013]. Since the source SDMX data and DSD file are both in XML format, we chose to use the XML serialization of RDF as our primary output. By integrating a Java library provided by any23.org (<http://any23.apache.org/>) in our Translator, we can also obtain Turtle format as an alternative output.

In Fig. 2 the overall technological architecture of our Translator is shown.

4. Validation and Experiments

The validation of our Translator was carried out according to three steps:

1. syntactic validation,
2. semantic validation,
3. experimental validation.

The validation process was carried out on real Istat datasets, represented in Table 2. The datasets were selected in order to maximize their diversity with respect to both dataset number of observations and number of dimensions.

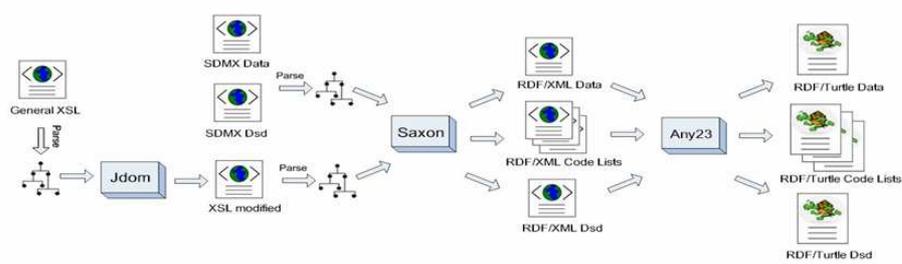


Fig. 2: Technological architecture of the translator

In the *syntactic validation* step, the obtained output files, expressed in RDF/XML format, were syntactically validated by creating the RDF triples and their graph representation. To this purpose, we used a free service provided by W3C called Validator (<http://www.w3.org/RDF/Validator/>) that takes as input RDF/XML document and returns a 3-tuple representation of the corresponding data model and an optional, but very explanatory, graphical visualization of the data model. Fig.3 shows the graphical representation of a simplified DSD containing only the component REF_AREA.

The syntactic validation of our output does not imply that our RDF files are "valid" with respect to the RDF-QB model. Therefore, we undertook a second step of *semantic validation*. To this purpose, we used the RDF Data Platform: Virtuoso Open Link Software (<http://virtuoso.openlinksw.com/>), and in particular we used the Virtuoso SPARQL entry point. Once uploaded the RDF/XML files (code lists, DSD and data), we examined our data by performing a set of focused SPARQL queries. For example, we issued a SPARQL query extracting the "Percentage of separate collection of waste in Rome", as shown in Fig. 4:

Dataset Name	Size (KB)	Dataset Description	
		# Dimensions	# Observations
Households Economic Conditions and Disparities - Poverty	17	8	160
Households Economic Conditions and Disparities - Housing conditions	60	17	532
Gross earnings per full time equivalent unit index - quarterly data	81	10	1.071
Households Economic Conditions and Disparities - Income	159	8	2.152
Services – Tourism	753	9	12.663
Households Economic Conditions and Disparities - Consumer confidence	764	13	10.550
Services - Air transport, rail transport, maritime transport, road transport	791	10	13.104
Environment and energy – Waste	1.338	8	14.302
Consumer price index for the whole nation (NIC - from 2011 onwards)	4.341	8	71.050
Annual national accounts + Quarterly national accounts + Territorial Economic Accounts and Aggregates + Annual sector accounts	4.937	12	63.576
Employment	16.367	19	148.400
Households Economic Conditions and Disparities - Consumptions	18.560	12	205.980

Table 2: Datasets used for validation

Graph of the data model

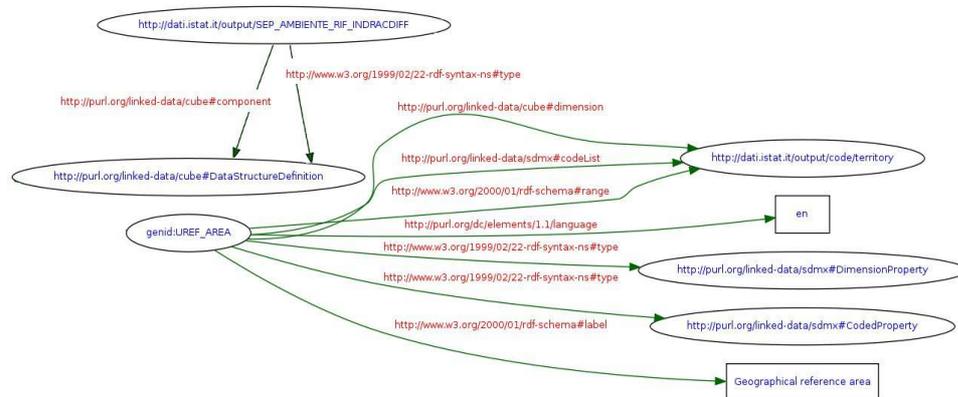


Fig. 3: Syntactic validation: RDF triples graph representation of DSD

The screenshot shows a SPARQL query interface with the following query:

```

SELECT * WHERE {
?obs <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://purl.org/linked-data/cube#Observation> .
?obs <http://dati.istat.it/property#TIPO_RIFIUTO> .
<http://dati.istat.it/output/code/typeOfWaste#9> .
?obs <http://dati.istat.it/property#IND_TYPER>
<http://dati.istat.it/output/code/environmentAndEnergyIndicators#RD_PERC> .
?obs <http://dati.istat.it/property#REF_AREA> ?territory .
?obs <http://purl.org/linked-data/sdmx/2009/measure#obsValue> ?
perc_sep .
?obs <http://purl.org/linked-data/sdmx/2009/dimension#timePeriod> ?
year .
OPTIONAL (?territory skos:prefLabel ?territorylabel) FILTER (?
territorylabel="Roma"@en) } ORDER BY ASC(?year) LIMIT 90

```

The results table below shows data for the year 2000 to 2009 for the territory of Rome:

obs	territory	perc_sep	year	territorylabel
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2000-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	5.7399798644	2000	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2001-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	6.1174714389	2001	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2002-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	6.9238965952	2002	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2003-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	10.504471576	2003	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2004-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	13.628205659	2004	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2005-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	16.001796488	2005	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2006-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	16.23427636	2006	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2007-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	17.099267384	2007	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2008-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	19.512051202	2008	"Roma"@en
http://dati.istat.it/output/datasets/environmentAndEnergy-Waste#58091-RD_PERC-9-A-2009-1-PURE_NUMB	http://dati.istat.it/output/code/territory#58091	20.673779347	2009	"Roma"@en

Fig. 4: Semantic validation: example query returning the percentage of separate collection of waste in Rome

With respect to the experimental validation, three different versions of the Translator were compared, namely:

- **Base:** This version was actually a generalization of the stylesheet used by the MIMAS project. Such a stylesheet had been really tailored on a specific case study, the World Development Indicators from the World Bank's Databank service. Hence, the translator performances suffered from such an ad-hoc development of the XSLT code.
- **Optimized_1:** We proceeded to rewrite the stylesheet of the Base version, to avoid redundancy, and unnecessary cycles minimizing repeated access to the information contained in the DSD file.
- **Optimized_2:** In this version, a considerable improvement was obtained using a technique that allows to load in memory, as an XSLT variable, an XML structure that contains information of the DSD avoiding multiple accesses to the file. This version reduced drastically the execution time.

Fig. 5 shows the Translator execution times in relation to data flow sizes, for the different Translator versions.

The experiments show that the performance gain from the Base to the Optimized_2 version is of two order of magnitude.

5. Conclusions and Future Work

The actual results of the IS-LOD project tested the viability of disseminating statistical data expressed in RDF-QB starting from SDMX data. Inevitably, the resulting data are actually documents, and, as such, not directly accessible via

related URIs. In future work, the project will investigate how integrating the current Istat dissemination architecture with an architecture fully supporting LOD publication (levels 4 and 5 of Tim-Berners Lee five-star stack).

Test Case	Data Input Size (KB)	Execution Time (min,sec)
Base	1.338	6m 7,90s
Base	2.674	12m 32,86s
Base	4.009	18m 42,25s
Base	5.345	23m 49,13s
Optimized_1	1.338	2m 17,71s
Optimized_1	2.674	4m 36,93s
Optimized_1	4.009	7m 28,95s
Optimized_1	5.345	9m 12,63s
Optimized_2	1.338	5,60s
Optimized_2	2.674	7,87s
Optimized_2	4.009	11,00s
Optimized_2	5.345	13,97s

Fig.5 Translator’s execution times for different versions

Bibliography

[SDMX-2013]: SDMX: <http://sdmx.org/>

[RDF-QB-2013] RDF Data Cube Vocabulary: <http://www.w3.org/TR/2013/CR-vocab-data-cube-20130625/>

[Marras-2013]: Marras S.: Open data: Lo scenario di riferimento, XI Conferenza Italiana di Statistica, 2013

[Five-star-2013]: Five star Open Data, <http://5stardata.info/>, 2013

[MIMAS 2011]: MIMAS Project : <http://mimasld.wordpress.com/2011/11/25/final-product-post-esds-the-sdmx-to-rdf-process/>

[Turtle-2013] W3C, “Turtle - Terse RDF Triple Language”, <http://www.w3.org/TeamSubmission/turtle/>, 2013

[N3-2013] W3C, “Notation3 (N3): A readable RDF syntax”, <http://www.w3.org/TeamSubmission/n3/>, 2013

[N-Triples-2013] W3C, “N-Triples”, <http://www.w3.org/TR/n-triples/>, 2013.

Big Data: limitazioni e opportunità geopolitiche e geoeconomiche

Carlo Muzzi

AICA - Associazione Italiana per l'Informatica ed il Calcolo Automatico
muzzi@acm.org

Abstract. *The idea that Big Data can constitute a new frontier for innovation, competition and productivity in the global economy is now generally accepted; but there are many geo-political and geo-economic conditions that influence them in the global context. These conditions will be analyzed in this study, along with the opportunities offered.*

Keywords: Big Data, conditions, opportunities, geo-political, geo-economic.

1. Introduzione

L'idea che i Big Data possano costituire una nuova frontiera per l'innovazione, la competizione e la produttività dell'economia globale [Manyika et al, 2011] è oramai generalmente condivisa. Lo stesso termine Big Data, inizialmente circoscritto alla letteratura scientifica o specialistica, è ora rilevabile anche nella più generale comunicazione pubblica; in effetti la generica ricerca (eseguita l'8 giugno 2013) dell'occorrenza *big data* nel motore di ricerca Google ha restituito circa 1.870.000.000 voci, di cui 520 introdotte nelle precedenti 24 ore. Anche se l'ambito di applicazione del concetto dei Big Data sembrerebbe estremamente vario e distribuito, scopo del presente studio sarà dimostrare come alcune condizioni geopolitiche e geoeconomiche oggi esistenti nello scenario planetario lo influenzeranno fortemente, inducendo conseguenti vincoli sulle possibilità di costituire realmente uno dei volani economici sui quali agire per uscire dalla crisi economica ancora in atto.

2. Cosa sono i Big Data

Il termine Big Data [Manyika et al, 2011] è utilizzato per riferirsi a quegli insieme di dati dal volume talmente grande da non poter più essere catturati, conservati, gestiti e analizzati con gli strumenti tipici dei database tradizionali; questa definizione è di tipo soggettivo poiché, non utilizzando metriche precise per specificare quanti dati sono necessari per poterli definire "grandi", preserva dal rischio di stabilire soglie che verrebbero certamente rese obsolete dal costante incremento delle informazioni digitali.

Nonostante il termine sia entrato in un uso più ampio solo di recente i Big Data sono già ampiamente diffusi nella vita di tutti [White, 2012] e nei contesti più diversi: nella capacità della rete Internet di aggregare e rendere fruibili documenti, testi, indici di ricerca, immagini, filmati, ecc.; nell'accumulazione senza precedenti di dati personali derivanti dalla social data revolution

Congresso Nazionale AICA 2013

[Weigend, 2009] innescatasi dalla diffusione dei social networks; dal continuo incremento delle informazioni digitalizzate su prodotti, merci, servizi, ecc. indotta dalla crescita del mercato dell'e-commerce; dalle raccolte di grandi masse di dati compiute da diverse discipline scientifiche (astronomia, chimica, biologia, meteorologia, genomica, ecc.); dai nuovi dati creati dalla sempre più pervasiva diffusione delle reti di sensori (come gli RFID).

È opportuno evidenziare come l'equazione "Big Data = Grandi Quantità di Dati" sia fuorviante; in effetti, quando ci si riferisce ai Big Data si pone più attenzione alla complessità [Rindler, 2009] piuttosto che alla quantità dei dati da trattare: ed è proprio da tale complessità che derivano le difficoltà ad effettuare analisi utilizzando i software statistici standard [Snijders et al, 2012].

2.1 Le 3 dimensioni dei Big Data

Fin dal 2001 [Laney, 2001] è stato posto in evidenza che la capacità di gestire la progressiva crescita di grandi quantità di informazioni produce ritorni significativi degli investimenti effettuati anche in ambito business. In questo contesto il termine Big Data è stato descritto anche attraverso il cosiddetto modello "3Vs": Data Volume + Data Velocity + Data Variety; un modello, successivamente rielaborato ed esteso in altri studi [Bayer e Laney, 2012], che espone 3 dimensioni significative per la crescita e l'affermazione dei Big Data: la *Volume*, che rappresenta la quantità delle informazioni disponibili (sempre in crescita); la *Velocità*, che esprime la rapidità di movimento dei dati (sia in input che in output); la *Varietà*, che sottolinea l'eterogeneità delle tipologie e delle sorgenti delle informazioni.

3 Leve di analisi per i Big Data

Dopo aver brevemente trattato del significato dei Big Data, analizzeremo sinteticamente alcune leve che, si ritiene, possano influenzare significativamente questa disciplina agendo come vettori di stimolo o di freno; in particolare valuteremo sia gli impatti negativi riconducibili ad alcuni limiti geopolitici e geoeconomici che le opportunità potenzialmente esistenti.

3.1 Storage

Non si possono conservare informazioni senza disporre degli opportuni strumenti tecnologici di memorizzazione. Questi strumenti variano in funzione dei soggetti che accumulano i Big Data: i singoli individui (che conservano video, foto, musica, messaggistica, comunicazione sociale, ecc.), le organizzazioni più o meno grandi (che nei propri database conservano dati strutturati e destrutturati di gestione aziendale, produzione, clientela, ecc.), gli stessi governi (che dispongono di banche dati anagrafiche, fiscali, welfare, sicurezza e polizia, ecc.) hanno esigenze di storage specifiche, essenzialmente riconducibili agli investimenti che accettano di sostenere.

Se analizziamo il solo mercato degli hard disk, che costituiscono l'elemento atomico per la costruzione dei vari livelli di storage tradizionale (SAN, NAS,

DAS, ecc.), si rileva come la progressione nelle politiche di acquisizione delle diverse corporate internazionali [Taylor, 2012] ha ristretto i produttori mondiali a soli 3 soggetti: Seagate (USA), Western Digital (USA) e Toshiba (Giappone).

Anche l'analisi [Santa Clara Consulting Group Executive Summary, 2012] dei produttori leader nel mercato dello storage USB, che ovviamente costituisce il riferimento prevalente per il mondo consumer, vede i seguenti brand assorbire circa il 71% del mercato globale: SanDisk (USA), Kingston (USA), Transcend (Taiwan), Verbatim (Giappone), HP (USA), Sony (Giappone), Buffalo (Giappone), Imation (USA) e PNY (USA).

La diffusione del modello del cloud ha anche aperto la via al mercato dello storage cloud; limitandosi, per semplicità d'analisi, alla comparazione [Ruth e Chandrasekaran, 2012] di 10 soluzioni rispetto a 6 capacità ritenute critiche per il mercato corporate o professionale troviamo: Amazon S3 (USA), AT&T Synaptic Storage as a Service (USA), Google Cloud Storage (USA), HP Cloud Object Storage (USA), IBM SmartCloud Enterprise Object Storage (USA), Internap AgileFILES (USA), Microsoft Windows Azure Blob Storage (USA), Nirvanix Public Cloud Storage (USA), Rackspace Cloud Files (USA) e SoftLayer Object Storage (USA).

Per le soluzioni sopra esposte è stata riportata l'indicazione dello stato nel quale ha sede prevalente il quartier generale di ciascuna società; risulta evidente che tali società, pur essendo in gran parte organizzazioni operanti su scala internazionale, sono prevalentemente localizzate nel Nord America e, solo in minor misura, nell'area asiatica. Si è dunque in presenza di un mercato fortemente guidato dalla leadership USA, quindi basato su una dipendenza strategicamente rilevante che sbilancia fortemente la leva dello storage su un novero estremamente limitato di attori prevalentemente riconducibili alla sfera di controllo statunitense. È verosimile che opportunità per operatori esterni, come quelli italiani, potrebbero aprirsi nel mercato dello storage USB (dove l'aspetto del design, così tipico del nostro paese, permetterebbe anche alle FabLab di operare in settori di nicchia) o dei servizi ad alto valore aggiunto nello storage cloud (soluzioni appetibili anche da chi ha minori capacità di investimento e competenze ICT, come nel caso delle piccole imprese italiane).

3.2 Internet

Il modello "3Vs" ha enfatizzato nella velocità di trasferimento dei dati uno dei fattori qualificanti per la diffusione dei Big Data; è ovvio che oggi non possa che avvenire attraverso internet. L'ampia disponibilità di connettività pubblica adeguata, veloce e pervasiva, diviene quindi un elemento infrastrutturale essenziale per il progresso dei Big Data. Dal punto di vista della disponibilità della rete recenti studi [Internet World Stats, 2012] hanno evidenziato come, pur in presenza di sensibili divari a livello continentale, la rete sia oggi utilizzata dal 34,3% della popolazione mondiale (stimata in poco più di 7 miliardi di individui): un tasso ampio che verosimilmente aumenterà tenendo conto che era cresciuto del 566,4% nei 12 anni precedenti. Anche l'accesso veloce alla rete è una realtà per una significativa parte dei 2,4 miliardi di individui già oggi connessi; analisi effettuate sui 34 stati aderenti

all'Organizzazione per la Co-operazione e lo Sviluppo Economico (OECD\OCSE) evidenziano l'esistenza di 321.317.973 accessi sottoscritti [OECD Broadband statistics, 2012a] alla banda larga su rete fissa e 698.656.918 su rete wireless [OECD Broadband statistics, 2012b].

Rank	Country	Fixed	Rank	Country	Wireless
34	Iceland	109.359	34	Iceland	206.359
33	Luxembourg	166.630	33	Luxembourg	334.800
32	Estonia	334.655	32	Slovenia	566.793
31	Slovenia	495.953	31	Estonia	850.601
30	Slovak Republic	771.638	30	Hungary	1.342.196
29	Ireland	1.011.730	29	Slovak Republic	1.972.551
28	New Zealand	1.239.569	28	Belgium	2.844.187
27	Finland	1.600.400	27	Ireland	2.845.357
26	Czech Republic	1.715.700	26	New Zealand	2.942.301
25	Norway	1.789.279	25	Israel	3.831.000
24	Israel	1.910.000	24	Chile	3.866.878
23	Austria	2.063.612	23	Portugal	3.960.947
22	Chile	2.101.524	22	Norway	3.987.776
21	Hungary	2.114.832	21	Switzerland	4.245.550
20	Denmark	2.131.167	20	Austria	4.334.468
19	Portugal	2.309.702	19	Greece	4.480.265
18	Greece	2.558.830	18	Denmark	5.125.855
17	Sweden	3.053.000	17	Finland	5.160.000
16	Switzerland	3.276.700	16	Czech Republic	6.134.301
15	Belgium	3.586.116	15	Sweden	9.616.000
14	Australia	5.601.000	14	Netherlands	9.760.000
13	Poland	5.602.280	13	Turkey	10.661.171
12	Netherlands	6.573.000	12	Mexico	10.725.623
11	Turkey	7.666.609	11	Canada	14.279.850
10	Canada	11.013.711	10	Italy	20.991.956
9	Spain	11.384.588	9	Poland	22.049.733
8	Mexico	12.108.478	8	Australia	22.178.000
7	Italy	13.444.129	7	Spain	22.277.983
6	Korea	18.010.256	6	France	30.562.000
5	United Kingdom	21.066.624	5	Germany	31.666.400
4	France	23.167.000	4	United Kingdom	37.629.000
3	Germany	27.677.770	3	Korea	51.894.381
2	Japan	35.142.132	2	Japan	107.508.636
1	United States	88.520.000	1	United States	237.824.000
	OECD	321.317.973		OECD	698.656.918

Fig. 1 – OECD broadband statistics: numero totale di sottoscrizioni fixed (via cavo) e wireless alla banda larga per paese, dati in milioni a giugno 2012

Una corretta analisi sui dati esposti (vedi Fig. 1) deve comparare il numero di accessi a banda larga in rapporto agli abitanti dei singoli stati. Ad esempio l'Islanda, pur disponendo del numero più basso di accessi sottoscritti su rete fissa e wireless, dispone di una copertura da ritenersi altissima rispetto alla propria popolazione stimata, nello stesso periodo, in circa 320.000 unità; in effetti questo successo è stato perseguito grazie al forte impegno della nazione islandese che, nel 2005, ha regolamentato il mercato delle telecomunicazioni [eGovernmentFactsheets, 2011] per garantire a tutti i cittadini, particolarmente quelli residenti in aree rurali o svantaggiate, l'accesso alla larga banda.

Ad oggi la disponibilità di connettività adeguata è una opportunità limitata solo ad una parte, anche se significativa, della popolazione mondiale; la

disparità diviene macroscopica in alcune regioni: ad esempio nel 2011 l'utente medio africano disponeva di una banda media di 2.000 bps contro i 90.000 bps del corrispondente utente europeo [ICT Data and Statistics Division, 2011]. Purtroppo l'esigenza di diffondere le tecnologie della comunicazione ha creato un boom del mercato mobile in Africa e di conseguenza una diffusione della connettività di tipo Broadband Wireless Access (BWA). Anche se questi limiti infrastrutturali sulla connettività incidono negativamente sulla diffusione dei Big Data, non bisogna però dimenticare che le opportunità offerte dall'evoluzione del web negli ultimi anni permettono, anche a coloro che utilizzano la rete a velocità più basse, di utilizzare comunque grandi quantità di informazioni; un solo esempio: grazie alla diffusione del modello Ajax [Garret, 2005] per lo sviluppo di applicazioni web interattive si sono ridotti sensibilmente i flussi dati http scambiati tra i nodi client e server in una comunicazione internet.

Naturalmente in uno scenario in cui la disponibilità di connettività ad alta capacità costituisce un elemento di vantaggio strategico, risulta evidente che taluni vincoli strutturali che ancora gravano sull'Italia rischiano di incrementare il divario con i paesi concorrenti. Per combatterlo dal 2008 lo stato italiano ha lanciato il Piano Nazionale Banda Larga con l'obiettivo di offrire ampia connettività anche agli 8 milioni di cittadini residenti in aree ancora scoperte dalle normale offerta del mercato commerciale; siccome gli ultimi investimenti definiti [Piano Nazionale Banda Larga, 2013] prevedono il raggiungimento di tale obiettivo nel breve periodo è verosimile che, nel momento in cui internet diverrà effettivamente un servizio universale per tutta la popolazione italiana, nuove opportunità per i Big Data potranno aprirsi anche nel nostro paese.

3.3 I grandi accumulatori di informazioni

Coloro che accumulano grandi quantità di informazioni diventano, o possono diventare, soggetti naturalmente avvantaggiati nel mondo dei Big Data. Tra questi vi sono ovviamente i soggetti che istituzionalmente dispongono di proprie banche dati anagrafiche, fiscali, assistenziali, sanitarie, bancarie, assicurative, militari, ambientali, scientifiche ecc. Queste informazioni sono di elevata importanza ma le analisi che possono realizzarsi sulle stesse sono fortemente circoscritte all'ambito in cui sono raccolte; difficilmente analisi correlate potranno avvenire senza interventi coercitivi e regolatori da parte delle competenti autorità sovrane. Diverso è invece il contesto internet dove i dati divengono pubblici, talvolta senza neanche la consapevolezza dei legittimi titolari.

In questo ambito diversi sono i grandi soggetti accumulatori di informazioni: i *motori di ricerca*, che per propria natura aggregano e correlano informazioni, occupano un ruolo rilevante. L'analisi dei dati esposti in Fig. 2 [Market Share Statistics for Internet Technologies, 2013] evidenzia il predominio del motore Google (USA) negli accessi da desktop (altre statistiche confermano il predominio anche per gli accessi da mobile, tablet e console). In generale gli attori residui più significativi sono Yahoo (USA), Bing (USA), Baidu (Cina), AOL (USA), ASK (USA), Lycos (USA). Dal punto di vista dei vincoli geopolitici e

geoeconomici anche in questo settore si rileva una prevalenza dell'area statunitense; il solo competitore alternativo, il cinese Baidu, pur disponendo del

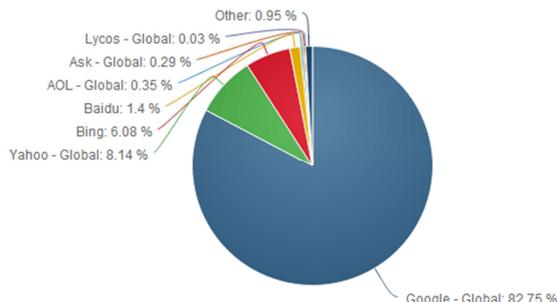


Fig. 2 – <http://marketshare.hitslink.com> market share dei motori di ricerca per accesso da desktop (29 giugno 2013)

vantaggio di operare nella florida economia industriale cinese, realizza comunque delle percentuali minime. È quindi naturale che in tale settore le opportunità per altri paesi siano estremamente ridotte (basta ricordare l'esperienza del progetto franco-tedesco per la realizzazione del motore di ricerca Quaero avviato nel 2005 e poi abbandonato dalla Germania), le sole possibilità

ipotizzabili potrebbero aprirsi in ambiti di nicchia; ad esempio per l'Italia, come già discusso in un precedente congresso AICA [Muzzi, 2009], nell'area del patrimonio storico-artistico.

La social data revolution ha fatto nascere nuovi soggetti che sono divenuti naturali aggregatori di informazioni direttamente provenienti dagli utenti registrati; la comparazione ([HTTP://SOCIAL-NETWORKING.FINDTHEBEST.COM](http://social-networking.findthebest.com), 2 luglio 2013, [FindTheBest, 2013]) sui *social network* più frequentemente visitati nel mese vede le prime 10 posizioni così occupate: Facebook (USA) 960.000.000; Twitter (USA), 170.000.000; Tumblr (USA), 120.000.000; LinkedIn (USA), 97.000.000; Mobikade (Gran Bretagna), 70.200.000; Multipl (USA), 56.000.000; Badoo (Gran Bretagna), 56.000.000; Flickr (USA), 51.000.000; Orkut (USA), 45.000.000; MySpace (USA), 31.000.000. Generalizziamo ora la ricerca prendendo in considerazione i 20 *siti internet* più visitati al mondo in termini assoluti, la classifica che si ottiene è: Facebook (USA), Google (USA), YouTube (USA), Yahoo! (USA), Amazon (USA), Baidu (Cina), Wikipedia (USA), Windows Live (USA), QQ.COM (Cina), Taobao (Cina), Google India (USA), Twitter (USA), Blogspot (USA), Yahoo! Japan (USA), LinkedIn (USA), Bing (USA), 新浪新闻中心/ sina.com.cn (Cina), Яндекс yandex.ru (Russia), ВКонтакте/vk.com (Russia) e Ask (USA). Riepilogando anche i grandi accumulatori sono prevalentemente allocati negli Stati Uniti: che quindi dispongono di una sovranità giuridica e di fatto sul mondo dei Big Data.

3.4 I giganti del calcolo

Grandi quantità di informazioni necessitano di grandi capacità di calcolo in grado di poterle trattare: un obiettivo che è certo più semplice raggiungere disponendo di adeguati supercomputer.

La FIG. 3 [TOP500 Authors, 2013] riporta la distribuzione geografica dei 500 computer più potenti del mondo. Emerge con evidenza che oltre il 50% si trovano nei soli USA; mentre la Cina, che pur recentemente ha reso operativo presso la National University of Defense Technology il calcolatore più potente

del mondo (il *Tianhe-2 MilkyWay-2*), ne ha circa $\frac{1}{4}$; il primo supercomputer presente in Italia occupa il 12° posto (è il *Fermi* dislocato presso il consorzio interuniversitario Cineca). In generale emerge la rapida crescita dell'Asia che insidia l'Europa; un dato negativo è l'assoluta assenza del continente africano che, ancora oggi, non dispone di nessun super-computer.

Stati / Continenti	Africa	Americhe	Asia	Europa	Oceania	Totale
USA		252				252
Cina			66			66
Giappone			30			30
Gran Bretagna				29		29
Francia				23		23
Germania				19		19
India			11			11
Canada		9				9
Russia				8		8
Svezia				7		7
Italia				6		6
Australia					5	5
Corea del Sud			4			4
Arabia Saudita			4			4
Svizzera				4		4
Brasile		3				3
Norvegia				3		3
Polonia				3		3
Spagna				3		3
Finlandia				2		2
Israele			2			2
Olanda				2		2
Austria				1		1
Belgio				1		1
Danimarca				1		1
Hong Kong			1			1
Taiwan			1			1
Totale	0	264	119	112	5	500

Fig. 3 – [http://www.top500.org/list/2013/06,top500 supecomputer sites](http://www.top500.org/list/2013/06,top500%20supecomputer%20sites)
(3 luglio 2013)

3.5 Le influenze dei contesti giuridici

Gli operatori dei Big Data sono fortemente influenzati dai contesti giuridici all'interno dei quali operano; talvolta tale influenza condiziona in misura ben maggiore che aspetti tecnologici o infrastrutturali. Consideriamo ad esempio la grande quantità di informazioni raccolte dai dispositivi che circondano la nostra esistenza: apparati mobili, sensori di rilevamento, log software, telecamere, microfoni, ecc. e tutti quegli strumenti che sono destinati a moltiplicarsi e disseminarsi grazie anche all'esponentiale diffusione dell'*Internet of Things*. Particolarmente sui dispositivi che comunicano attraverso sistemi wireless, come i dispositivi elettronici autonomi connessi nelle reti WSN (Wireless Sensor Network) o i tag RFID utilizzati nelle tecnologie d'identificazione a radio

frequenza, hanno impatto le normative di compatibilità elettromagnetica che mirano a preservare da disturbi su quanto si trova nell'ambiente circostante; ma ulteriori riflessi giuridici derivano anche dalla crescente attenzione sul tema controverso dell'elettrosensibilità [Knowledge note, 2012] manifestata verso campi magnetici, elettrici o elettromagnetici e sui sintomi fisici e/o psicologici che sembrano accusare taluni soggetti esposti ai medesimi. Quanto questi ed altri aspetti normativi inducano vantaggi o svantaggi nella competizione internazionale è da sempre oggetto di dibattito; è stato talvolta evidenziato [Talone e Russo, 2008] come lo svantaggio competitivo dell'Europa rispetto agli USA derivasse da un quadro giuridico più sfavorevole. L'azione svolta dai paesi europei, anche a livello comunitario, ha permesso di recuperare in parte tale svantaggio ma, nel frattempo, anche l'attore cinese è entrato in gioco: tuttavia il ruolo degli USA rimane prevalente [de Panizza et al, 2010].

Impatto giuridico rilevante hanno poi le regolamentazioni connesse alla privacy e più in generale alla data protection; una tematica ampia che, nell'ottica di tutelare i diritti dell'individuo, si riflette sulle possibilità di poter effettuare profilazioni e schedature di massa. Pur ricordando [Movius e Krup, 2009] che l'art. 12 della Dichiarazione Universale dei Diritti dell'Uomo recita che *"nessun individuo potrà essere sottoposto ad interferenze arbitrarie nella sua vita privata..."*, i meccanismi che implementano tali garanzie differiscono fortemente da nazione a nazione: creando così aree geografiche all'interno delle quali possono operare soggetti, pubblici o privati, che possono acquisire poteri politici ed economici eccezionali che, grazie ad internet, non si limitano solo al proprio ambito di sovranità nazionale ma sconfinano a livello mondiale: occorre qui ricordare che, nei Big Data, il potere derivante dalla semplice disponibilità di immense aggregazioni di informazioni si amplifica enormemente quando queste vengono elaborate con sofisticati strumenti di analisi predittiva che ne enfatizzano la valenza strategica, socio-politica e patrimoniale [Mantelero, 2012]. Anche il contesto giuridico può dunque spingere gli attori di questo settore a trasferire o implementare i propri sistemi (data center, cloud computer, ecc.) nelle nazioni ritenute più favorevoli. Non essendo questo il luogo deputato a trattare di cyber crime escludiamo dalla nostra analisi approfondimenti su quelle nazioni che, per minor sviluppo o per qualche forma di complicità, dispongono addirittura di un quadro legislativo favorevole verso talune azioni illecite; rivolgiamo invece la nostra attenzione sulla chiarezza del quadro normativo: emerge che negli USA, al vantaggio della preminenza del settore dell'ICT, si aggiunge la leva competitiva di una normativa più armoniosa rispetto a quella ad esempio esistente nei diversi stati dell'Unione Europea [Mantelero, 2012]. La recente proposta della Commissione europea per un nuovo regolamento continentale che costituisca un quadro giuridico omogeneo in materia di protezione dei dati, potrebbe, se approvata, fornire maggiore competitività anche ai player europei [Commissione europea, 2013].

4 Conclusioni

L'esame delle leve sopra esposte evidenzia due aspetti: la grande importanza strategica del tema dei Big Data e la prevalente superiorità

dell'attore USA a livello mondiale e, in subordine, una significativa posizione degli attori asiatici. In generale l'Europa e il nostro paese subiscono più che competere in questo mercato; per il nostro continente questa situazione non costituisce solo uno svantaggio economico, ma anche elemento di dipendenza strategica per certi versi assimilabile a quella esistente verso i paesi produttori di petrolio (lo scandalo del Datagate costituisce un chiaro esempio di tale subordinazione e dei rischi connessi [La Stampa, 2013]). Un'azione comunitaria più omogenea potrebbe costituire lo strumento per recuperare lo svantaggio esistente; un'azione che l'Italia dovrebbe promuovere arricchendola di azioni specifiche sul proprio territorio. Specificatamente promuovendo una maggiore sinergia tra 3 soggetti: i distretti industriali, che tipicamente costituiscono l'anima produttiva del nostro paese, le università e i centri di ricerca di cui ancora si dispone e la leva pubblica (governo e regioni) per sfruttare i vantaggi competitivi sopra esposti.

Bibliografia

[Bayer e Laney, 2012] Bayer M.A., Laney D., 2012, *The Importance of 'Big Data': A Definition*, Gartner.

[Commissione europea, 2013] Commissione europea, 2013, *Salvaguardare la privacy in un mondo interconnesso. Un quadro europeo della protezione dei dati per il XXI secolo*, Bruxelles, 25.1.2012, COM(2012) 9 final.

[de Panizza et al, 2010] de Panizza A., Lindmark S., Rotter P., 2010, *RFID: Prospects for Europe, item-level tagging and public transportation, fondamenti di una tecnologia silenziosamente pervasiva*, EUR 24416 EN, JRC European Commission.

[eGovernmentFactsheets, 2011] eGovernmentFactsheets, 2011, *eGovernment in Iceland*, European Commission.

[FindTheBest, 2013] FindTheBest unbiased, data-driven comparisons, 2013, *Compare social networking, monthly visitors*, FindTheBest.

[Garret, 2005] Garret J.J., 2005, *Ajax: A New Approach to Web Applications*, Adaptive Path.

[ICT Data and Statistics Division, 2011] ICT Data and Statistics Division-Telecommunication Development Bureau, 2011, *The World in 2011 – ICT Facts and Figures*, International Telecommunication Union.

[Internet World Stats, 2012] Internet World Stats, 2012, *Internet users on June 30, 2012*, Miniwatts Marketing Group.

[Knowledge note, 2012] Knowledge note 2012-001, 2012, *Electrosensitivity. Health complaints near sources of electromagnetic fields*, Dutch Knowledge Platform on Electromagnetic Fields and Health.

[Laney, 2001] Laney D., 2001, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, META Group.

[La Stampa, 2013] La Stampa, 2013, *Datagate, i giganti del Web aderiscono al programma Usa*, LASTAMPA.IT (Esteri, 8 giugno 2013).

[Manyika et al, 2011] Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Hung Byers A., 2011, *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute.

[Mantelero, 2012] Mantelero A., 2012, *Big Data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, Il diritto dell'informazione e dell'informatica (Fasc. 1, 2012).

[Market Share Statistics for Internet Technologies, 2013] Market Share Statistics for Internet Technologies, 2013, *Device Types Search Engine Market Share*, Netmarketshare.

[Muzzi, 2009] Muzzi C., 2009, *Ridondanza dei sistemi nel mondo reale, relazione al Congresso Nazionale AICA 2009*, Università La Sapienza, Roma.

[Movius e Krup, 2009] Movius L. B., Krup N., 2009, *U.S. and EU Privacy Policy: Comparison of Regulatory Approaches*, International Journal of Communication (3, 2009).

[OECD Broadband statistics, 2012a] OECD Broadband statistics, 2012, *Total number of fixed (wired) broadband subscriptions, by country, millions, June 2012*, Organisation for Economic Co-operation and Development (OECD/OCSE).

[OECD Broadband statistics, 2012b] OECD Broadband statistics, 2012, *Total number of wireless broadband subscriptions, by country, millions, June 2012*, Organisation for Economic Co-operation and Development (OECD/OCSE).

[Piano Nazionale Banda Larga, 2013] Piano Nazionale Banda Larga, 2013, *Al via bandi per 900 mln di euro per azzerare digital divide e banda ultralarga*, Agenda Digitale Italiana.

[Rindler, 2009] Rindler A., 2012, *Big Data Definition*, MIKE2.0 The open source standard for Information Management.

[Ruth e Chandrasekaran, 2012] Ruth G., Chandrasekaran A., 2012, *Critical Capabilities for Public Cloud Storage Services*, Gartner.

[Santa Clara Consulting Group Executive Summary, 2012] Santa Clara Consulting Group, Executive Summary–Q4/Year-End 2012, 2012, *USB Flash Tracker™ - Trends for USB Flash Drive Markets*, Santa Clara Consulting Group (Vol. 9 No. 4).

[Snijders et al, 2012] Snijders C., Matzat U., Reip U-D., 2012, *"Big Data": Big Gaps of Knowledge in the Field of Internet Science*, International Journal of Internet Science (2012, 7).

[Talone e Russo, 2008] Talone P., Russo G., 2008, *RFId, fondamenti di una tecnologia silenziosamente pervasiva*, Fondazione Ugo Bordoni.

[Taylor, 2012] Taylor P., 2012, *HDD business to become Mexican standoff*, TechEYE.net.

[TOP500 Authors, 2013] TOP500 Authors, 2013, *Top500 List - June 2013*, TOP500 supercomputer sites.

[Weigend, 2009] Weigend A., 2009, *The Social Data Revolution(s)*, Harvard Business Review.

[White, 2012] White T., 2012, *Hadoop: The Definitive Guide*, O'Reilly Media.

Business Transformation with Big Data

Raffaele Stifani
Executive Architect, IBM Italia
Circonvallazione Idroscalo, 20090 Segrate (MI)
raffaele_stifani at it.ibm.com

Organizations today are collecting tremendous volumes of data, generated by a wide variety of sources, often at extreme velocities. As the amount of data available to enterprises and other organizations dramatically increases, more and more companies are looking to turn this data into actionable information and intelligence in real time. Addressing these requirements requires applications that are able to analyze potentially enormous volumes and varieties of continuous data streams to provide decision makers with critical information almost instantaneously. Leading organizations in financial services, telecommunications, retail, healthcare, digital media, insurance and other industries are adopting advanced technologies to generate new actionable insights from Big Data that can help them dramatically reduce financial risks, increase operational efficiencies, enhance customer loyalty and improve healthcare outcomes. These organizations are tapping into Big Data to transform not only their businesses but also their industries.

1. Introduction

The term “Big Data” is a bit of misnomer since this implies that pre-existing data is somehow small (and it isn't) or that the only challenge is the size of data (size is only one of them). The term Big Data applies to information that can't be processed or analyzed using traditional processes or tools. Today organizations are increasingly facing more and more Big Data challenges. We live in a climate where companies have the ability to store anything generating data like never before in history. Today's business has more access to potential insight than ever before; as this potential gold mine of data piles up, the percentage of data the business can process is going down fast.

The Big Data era is in full force today because the world is changing. Through instrumentations, we are able to sense more things, and if we can sense it, we tend to store it (or at least some of it). Through advances in communications technology, people and things are becoming increasingly interconnected. Usually referred to as machine-to-machine (M2M), interconnectivity is responsible for double-digit year over year data growth rates.

Finally, because small integrated circuits are now so inexpensive, we are able to add intelligence to almost everything.

Three characteristics define Big Data: volume, variety and velocity. Big Data have created the need for a new class of capabilities to augment the way things are done today to provide better control over the existing knowledge domains and the ability to act on them. A robust and effective Big Data platform gives the unique opportunity to extract insight from an immense volume, variety and velocity of data, in context, beyond what was previously possible.

1.1 Volume

The volume of data being stored today is exploding. In the year 2000, about 800,000 petabytes (PB) of data were stored in the world. We expect this number to reach about 35 zettabytes (ZB) by 2020. Twitter alone generates more than 7 terabytes (TB) of data every day, Facebook more than 10 terabytes (TB). Organizations are facing massive volume of data. Organizations that don't know how to manage this data are overwhelmed by it. Companies have the opportunity, using the right technological platform, to analyze almost all of the data (or at least more of it by identifying the data useful to them) to gain a better understanding of the business, the customers and the marketplace.

As the amount of data available to the enterprise is on the rise, the percent of data it can process, understand and analyze is on the decline; thereby it is creating a blind zone. What is in the blind zone? We don't know: it might be something great or may be nothing at all, but the "don't know" is the problem (or the opportunity, depending how we look at it).

1.2 Variety

With the explosion of sensors and smart devices as well as social collaboration technologies, data in an enterprise has become complex because it includes not only traditional relational data but also semi-structured and unstructured data from web pages, log files (including streaming data), search indexes, social media forums, e-mails, documents, sensors data from active and passive systems, and so on.

Variety represents all type of data: it is a fundamental shift in analysis requirements from traditional structured data to include raw, semi-structured and unstructured data as part of the decision-making and insight process. Today database represents about 20 percent of the data; 80 percent of the world data is unstructured or semi-structured at best. To capitalize on Big Data opportunity, enterprises must be able to analyze all types of data, both relational and non-relational: text, sensor data, audio, video, transactional and more.

1.3 Velocity

Just as the volume and variety of data we collect and store has changed, so too has the velocity at which it is generated and needs to be handled. In a lot of use cases there is the need to analyze the data before it is stored. This data that is continuously flowing across interconnected communication channels is considered streaming data or *data in motion*.

Today's enterprises are dealing with petabytes of data instead of terabytes and the increase in RFID sensors and other information streams has led to a constant flow of data at a pace that has made it impossible for traditional systems to handle.

In addition more and more of the data being produced today has a very short shelf-life, so organizations must be able to analyze this data in near real time if they hope to find insights in this data.

In traditional processing, companies can think of running queries against relative static data; with streams computing, companies need to execute a process similar to a continuous query with continuously updated results (for example, because location information from GPS data is refreshed in real time). Dealing effectively with Big Data requires performing analytics against the volume and the variety of data while it is still in motion, not just after it is at rest.

2. What's new in Big Data

In the past, organizations used technology solutions to analyze historical data and identify broad trends based on a limited collection of information housed in structured databases. Today cutting-edge technologies enable organizations to analyze much more data from an extensive array of sources at incredible speeds. Now leading organizations can:

- conduct real-time analysis of customer behaviors to produce tailored experiences and targeted promotions;
- measure the effectiveness of online advertising to fine-tune campaigns while they are in progress;
- adopt advanced content analytics solutions to mine social media posts and call-center logs in order to assess customer sentiment and avoid churn;
- analyze data continuously streaming in from operational sensors to increase service uptime, facilitate better planning and anticipate risks;
- implement predictive analytics solutions to anticipate future customer behaviors, avoid risks and identify potential outcomes.

With the right tools, organizations are capitalizing on the wealth of opportunities that Big Data presents.

3. Serve customers as individuals

According to the IBM Global CEO Study 2012, forward-thinking CEOs identify customer insight as the most important area for new investment. These leaders want to capitalize on the vast potential of Big Data to provide deeper insight into customer preferences, needs and trends.

For most organizations, there is no shortage of customer data available. Banking customers provide financial information on credit applications and discuss banking problems during customer service calls; telecommunications subscribers continuously generate smartphone usage data; retail customers enter information for online transactions and register their "likes" on social media networks. The challenge lies in handling the volume and velocity:

businesses must efficiently analyze that data in order to generate timely insights that can help them enhance the customer experience.

Many organizations already use data management solutions to integrate customer information from multiple sources and create a single, holistic view of each customer. New solutions designed for understanding Big Data can now push those capabilities further. Organizations can use advanced analytics to provide near-real-time trend analysis and anticipate future outcomes. These solutions can produce insights that help organizations create targeted marketing promotions, optimize ad campaigns, avoid churn and improve cross-sell and up-sell opportunities.

In retail, leading organizations are investing in advanced analytics solutions to explore Big Data, detect patterns and reveal new insights that are helping them better understand and engage with individual customers. They are capturing customer sentiment and discovering new insights by examining a variety of structured and unstructured data, ranging from information they have already collected to sentiments expressed through social media.

Predictive analytics capabilities enable retailers to conduct precise segmentation, down to the individual level, by gauging future customer behaviors. Marketing teams can be more precise in identifying prospects, managing their marketing budgets to maximize their marketing return on investment, and developing targeted and relevant offers across all channels that deliver a richer, more personalized shopping experience. At the same time, retailers can optimize merchandising decisions for pricing, assortment, inventory and demand forecast.

Some forward-thinking telecommunications companies are capitalizing on Big Data to transform call centers from cost centers to revenue drivers. They use solutions to analyze previous customer interactions, integrate information with existing customer information and present real-time results to call-center agents so they can provide timely cross-sell and up-sell offers. Drawing on micro-segment, location and search-history information from across multiple channels (smartphone, TV, Internet services, etc.) enables companies to deliver highly targeted product bundles.

Communication service providers are also leveraging predictive analytics to reduce customer churn. By gaining insight about customers with a high propensity to change services or move to competitors, they can proactively engage and service those customers' individual needs and retain their business.

Banks are using analytics solutions to mine Big Data for insights that help them create more customer-focused enterprises that foster lasting relationships. They are moving beyond customer surveys and the review of customer service logs to analyze a variety of structured and unstructured information. For example, leading banks are analyzing Internet feedback and social media posts to address negative comments and build on positive ones to improve their reputation and retain customers. Building stronger, lasting relationships with customers is having a direct, positive impact on revenues.

4. Operational Efficiency

While many organizations have experienced the benefits that analytics and business intelligence solutions can provide for specific back-office functions, advanced solutions for in-depth analysis of Big Data are providing important new opportunities to change the way businesses operate. Analyzing streaming data from instrumented operational systems, deeply analyzing data from inventory and supply-chain operations, and analyzing data streaming from financial systems can help organizations significantly increase operational efficiency, boost revenues and ensure service availability.

For example, leading telecommunications organizations are performing real-time, root-cause analysis on data streaming in from a variety of sensors as well as retrospective analysis on massive volumes of call detail records and network event data. Instead of struggling to address customer problems and unanticipated outages, organizations are using solutions for analyzing Big Data to help rapidly respond to and prevent connectivity and bandwidth problems while optimizing performance and improving capacity planning. Business users can perform network quality-of-experience (QoE) analysis, traffic engineering and data analysis to identify and address network bottlenecks faster.

Leading investment firms are using real-time analytics solutions for Big Data to improve financial decision making. For example, stock market traders are conducting real-time analysis of streaming market data and incorporating contextual awareness (such as global news events and weather) into trading decisions. Consuming, analyzing and acting on real-time market data helps traders maximize gains.

5. Product innovation

In many industries, product innovation is critical for success, but the process of researching, developing, testing, reporting, adjusting and retesting new offerings can be long and resource-intensive. Big Data is a common factor in product development efforts, and leading organizations are using it to drive innovation by employing analytics solutions to explore large, complex data sets and generate new insights.

Analytics solutions also enable organizations to test “what-if” scenarios and anticipate the performance of new products and services, fueling experimentation and guiding research investments. Retailers, banks and telecommunications companies can use analytics to collect valuable feedback on current offerings and identify emerging market trends. These organizations can draw insights from unstructured customer data generated through social media posts, call-center interactions and online chat sessions, and they can use these insights as the basis for developing new products and services.

Predictive analytics solutions enable organizations to anticipate how new products and services will be received in the marketplace. By helping to gauge future customer behaviors, predictive analytics solutions enable organizations not only to create products that meet customer needs but also to scale production appropriately.

In addition, analytics solutions help organizations in all industries efficiently and successfully market innovative products. Using cluster analytics, organizations can identify shared customer attributes that may not have been obvious to analysts. Teams can then create marketing campaigns or promotions tailored to precise customer segments, allowing them to focus their resources on the customers most likely to embrace new products.

Statistical analytics solutions also help companies test the effectiveness of websites, direct email campaigns or other marketing collateral. Organizations can determine which outreach efforts generate the most qualified leads and ultimately yield the strongest customers.

6. Manage frauds and risks

Effective analysis of Big Data provides tremendous potential for enhancing risk management and avoiding costly losses. With analytics solutions for Big Data, organizations gain instant awareness of risks and can generate insights that help enhance investment decisions, improve lending decisions and increase fraud detection.

Leading financial services organizations are using Big Data to minimize credit risks and make smarter investments. Data management solutions help to integrate market, credit, operational and regulatory data to create a comprehensive view of enterprise risk exposure. Advanced analytics capabilities enable these organizations to analyze years' worth of structured and unstructured identity, behavior and financial transaction data to make more informed decisions. Implementing new solutions for analyzing Big Data is delivering significant returns on the investment: with better risk analysis, these organizations can reduce write-offs and help minimize losses.

For savvy insurance companies, analytics solutions for Big Data are helping to prevent and detect fraud which can account for a significant portion of an insurance company's losses. Predictive analysis technologies assess the future fraud potential of policy applicants by scrutinizing the past history of applicants and other people associated with them, giving organizations another tool to help prevent fraud attempts. Content analytics technologies deliver a more complete view of information than service bureaus or existing solutions can provide. As a result, organizations are able to correlate information from multiple departments and data sources: for example, they can monitor and analyze social media sources for rumors, deliberate misinformation and fraudulent impersonation of employees. Better fraud prediction and detection helps these insurers significantly reduce costs.

7. Exploit instrumented assets

From RFID tags and smart utility meters to building security systems and railroad trackside sensors, today's world is more instrumented than ever before. Data collection devices offer organizations access to a tremendous volume of information that streams in at high velocity. Leading organizations are capitalizing on this availability by employing analytics solutions for Big Data to

identify problems in real time, improve asset management, enhance operational efficiencies and provide real-time feedback to customers.

Many telecommunications companies, for example, are using solutions for analyzing Big Data to improve service quality and availability by analyzing data that streams in from a wide range of sources. With analytics and reporting solutions for Big Data, they can conduct operational and failure analysis from device, sensor and GPS inputs to solve existing problems as well as prevent future ones.

In healthcare, top providers are integrating large volumes of data from multiple sources so that doctors can access a full range of individual patient information right away. They are using analytics solutions for streaming data to improve real-time decision making. Incorporating content analytics capabilities enables healthcare organizations to find valuable information in unstructured content (such as doctors' notes or medical journal articles) to treat current patients, identify important trends and improve treatment regimens over the long term.

8. Big Data Platform

Big Data requires a new platform with a collection of best-of-breed technologies and services that help organizations integrate data from disparate sources, analyze Big Data in real time, help anticipate future outcomes and rapidly generate insights for capitalizing on new opportunities. Platform components should include:

- discovery and navigation software that provides real-time access and fusion of Big Data with rich and varied data from enterprise applications for greater insight and ROI;
- an enterprise-ready Apache Hadoop-based system with sophisticated text analytics, visualization, performance, security and administrative features for managing and analyzing massive volumes of structured and unstructured data;
- in-motion streaming analytics software that enables continuous analysis of massive volumes of streaming data with sub-millisecond response times, helping to improve organization's level of insight and decision making, as well as promoting real-time response to events as they happen;
- high-performance data warehouse systems able to make advanced analytics on exploding data volumes and to use advanced analytics to deliver deep insights in minutes on petabyte-scale volumes of relational data;
- comprehensive data warehouse software platform that delivers access to structured and unstructured information in real time; able to support operational analytics and applications with up-to-the-minute insights;
- a complete suite of data integration and data quality capabilities that help ensure delivery of trusted information; enables organizations to understand, cleanse, transform and deliver trusted information to critical business initiatives by integrating Big Data across enterprise IT systems;

- a Master Data Management software to create trusted views of master data about customers, products and more, and provides a centralized data source that promotes accuracy and data quality to help improve applications and business processes.

Powering analytics for Big Data requires an infrastructure with the potential to reap the benefits and drive business outcomes. It requires systems designed to access the latest information, regardless of type or location, by allocating the right resources at the right time for analysis on demand. The systems have to provide:

- scalable capabilities that helps optimize analytic workload performance by using all available data and information;
- high-performance parallel technologies that optimize complex decision making by spotting trends and anomalies to predict outcomes;
- resilient architectures, either on-premise or in the cloud, that help organizations deploy analytics throughout their business, and with customers and suppliers.

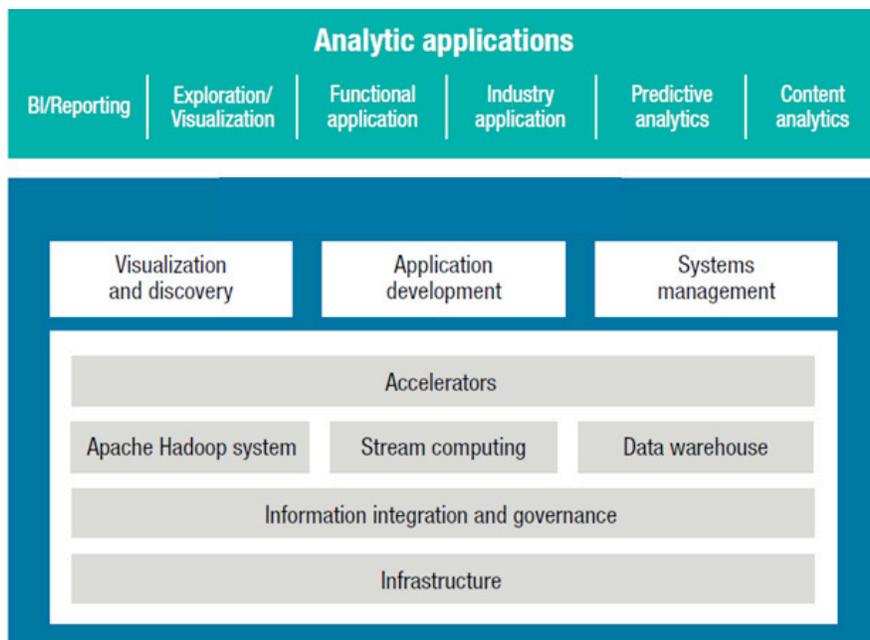


Figure 1 – Big Data Platform

9. Readiness Assessment

Big Data offers organizations new ways to use enterprise data to deepen business insights and extend competitive advantage but today many companies are unsure of where to start. Which Big Data projects will deliver the biggest

and fastest returns with the smallest risk? How will Big Data fit into an existing technical environment?

A *Readiness Assessment for Big Data* answers these questions so companies can begin using Big Data and its associated technologies in use cases that have a direct, positive business impact within a reasonable time frame. Scope of the assessment is to provide findings and recommendations on one or more high-value Big Data use cases to consider for deploying. The assessment report should contain estimated benefits, costs, a high-level architecture and a deployment model associated with each Big Data use case. The main activities of a readiness assessment are:

- **Business Opportunity Identification:** an interactive discussion with experts to explore the many possible use cases for Big Data within the organization; the list is also prioritized according to business value and importance for the organization;
- **Business Justification:** analysis for developing the business case and quantifying the potential Return On Investment (ROI) from each use case by identifying the key performance indicators (KPI) and the business and technical processes that can be improved through deployment of that use case;
- **Solution Definition and Architecture:** an interactive session to explore and document the architecture of the proposed use cases as mapped against the organization's existing IT environment, enterprise data assets, business processes and systems.

10. Conclusions

Analytics solutions for Big Data offer tremendous possibilities for transforming businesses and their industries. Leading organizations in retail, financial services, telecommunications, healthcare and other industries are already capitalizing on new opportunities.

In the past the barriers to capitalizing on Big Data appeared formidable. Today the platforms for Big Data offer a comprehensive, integrated collection of state-of-the-art technologies for data management and analytics that can help organizations address the challenges of Big Data and realize its potential for transformation.

Organizations can assess their current standing and future plans to work with Big Data and evaluate their maturity. Based on an objective industry technique, this maturity model helps organizations evaluate their place on the Big Data adoption pattern and benchmark their status against emerging best practices around Big Data. The maturity model incorporates data and findings from the latest Big Data studies and research, and helps organizations assess their maturity level across Big Data focused areas such as business strategy and value, technologies, architectures and components, approaches, governance and data types - key inputs for defining a Big Data roadmap.

References

- [1] Barry D. and others, Big Data Comes of Age, An Enterprise Management Associates and 9sight Consulting Research Report, November 2012
- [2] Burghard C., Big Data and Analytics Key to Accountable Care Success, IDC Health Insights #HI237529, October 2012
- [3] Ferguson M., Architecting A Big Data Platform for Analytics, Intelligent Business Strategies, October 2012
- [4] IBM Global CEO Study 2012 at www.ibm.com/services/us/en/c-suite/ceostudy2012
- [5] IBM Institute for Business Value, Analytics: The real-world use of big data, IBM Institute for Business Value in collaboration with Saïd Business School at the University of Oxford Business, November 2012
- [6] IBM Institute for Business Value, Analytics: The real-world use of big data in financial services, IBM Institute for Business Value in collaboration with Saïd Business School at the University of Oxford Business, May 2013
- [7] Kiron D. and others, Analytics: The widening divide - How companies are achieving competitive advantage through analytics, IBM Institute for Business Value In collaboration with MIT Sloan Management Review, October 2011
- [8] TechAmerica Foundation, Demystifying Big Data: A Practical Guide To Transforming The Business of Government, Prepared by TechAmerica Foundation's Federal Big Data Commission, 2012
- [9] Zikopoulos P. and others, Understanding Big Data: Analytics for Enterprise Class, Hadoop and Streaming Data, McGraw Hill, 2012

Riconoscimento Automatico di Design Pattern da documentazione UML di Software Artifacts

Beniamino Di Martino¹, Antonio Esposito²
Dipartimento di Ingegneria Industriale e dell'Informazione
Seconda Università di Napoli
Aversa (CE), Italy
¹beniamino.dimartino@unina.it
²antonio.esposito7@studenti.unina2.it

This paper describes a procedure and a prototype implementation for automatic recognition of Software Design Patterns from documentation of Software artifacts design and implementation, provided in machine readable form, namely UML diagrams. The procedure is based on a semantic representation of the patterns to be recognized, provided by the ODOL language developed at the Massey University (New Zealand), which we have augmented with an OWL-S based representation of the dynamic behavior of the pattern. The UML set of diagrams which are included in the design and implementation documentation of the Software artifact are automatically analyzed and translated into a first-order logic representation. The ODOL+OWL-S Design patterns representation is likewise translated in a Prolog based representation. A set of first order logic rules have been defined for describing the heuristics and features which trigger the recognition, based on the Prolog description of the patterns to be recognized, and applied to the base of prolog facts which represent the UML documentation. Such set of rules is independent from the specific pattern to be recognized.

Keywords: Design Patterns, UML, OWL, OWL-S, ODOL

1. Introduzione

Il ruolo dei design pattern nell'Ingegneria del Software è piuttosto noto e nel tempo sono stati proposti diversi approcci e tool per supportare il progetto di processi software attraverso pattern. Anche l'uso dei design pattern nel reverse engineering è stato analizzato e studiato. In [Nija e Olsson, 2003] si afferma che, riconoscendo design pattern da codice sorgente, è possibile comprendere meglio programmi complessi e riconoscere l'architettura complessiva dei sistemi. I pattern possono essere utili anche per stimare la qualità del progetto di un sistema software [Beyer e Lewerentz, 2003] in quanto, insieme alla

definizione e al riconoscimento dei cosiddetti anti-pattern, possono essere usati come metriche di qualità. La motivazione principale per il riconoscimento di pattern da codice è rappresentata dalla possibilità di supportare il refactoring dei componenti degli artefatti software, che possono essere identificati come specifici elementi del pattern identificato. In letteratura sono state analizzate diverse tecniche per il riconoscimento automatico dei design pattern a partire da codice sorgente o da documentazione software. Nel presente lavoro proponiamo una procedura e una implementazione di un prototipo per il riconoscimento di Software Design Pattern dalla documentazione relativa al progetto e alla implementazione di artefatti software. La rappresentazione dei pattern si basa su una descrizione semantica che sfrutta un modello OWL chiamato ODOL, definito e sviluppato presso l'Università di Massey in Nuova Zelanda e descritto ampiamente in [Dietrich e Elgar,2007]. In letteratura sono stati proposti diversi linguaggi formali per rappresentare Software Pattern, sia da un punto di vista statico che dinamico, con risultati anche molto differenti [Tahibi, 2007]. Lo stesso UML è stato sfruttato per rappresentare i diversi aspetti dei Pattern. Le rappresentazioni semantiche si sono dimostrate particolarmente utili per descrivere i Pattern, almeno per quanto riguarda la loro struttura. In particolare, la rappresentazione ontologica ODOL si è mostrata sufficientemente completa, flessibile e facilmente estendibile, fornendo un solido punto di partenza per la descrizione dei pattern. Il modello proposto da ODOL è stato analizzato ed esteso per migliorarne l'espressività, dopodiché è stato affiancato da descrizioni basate su OWL-S per la rappresentazione del comportamento dinamico dei pattern utilizzando annotazioni specifiche per i differenti pattern esaminati. La documentazione relativa agli artefatti software è invece rappresentata da diagrammi UML, in particolare diagrammi delle classi e delle interazioni. I diagrammi UML sono spesso usati per documentare artefatti Software (o possono essere estratti in maniera automatica dal software quando non direttamente disponibili). Esistono in letteratura delle rappresentazioni machine readable dei modelli UML, in particolare XML Metadata Interchange (XMI), che permettono di analizzare i documenti UML in maniera automatica. Per comparare la rappresentazione XMI della documentazione UML e la descrizione ODOL dei pattern è necessario sfruttare un modello di rappresentazione intermedio comune. Per i nostri scopi, il modello che più si è mostrato utile è rappresentato da una descrizione in logica del primo ordine che utilizza il motore Prolog. XMI può essere rappresentato automaticamente attraverso fatti prolog sfruttando appositi tool, mentre la rappresentazione semantica offerta da ODOL può essere analizzata e convertita in prolog sfruttando il framework Thea [Thea,2013]. A partire da una base di conoscenza comune espressa in Prolog, è possibile definire regole per descrivere le euristiche che definiscono il riconoscimento dei pattern, producendo uno o più possibili mapping tra i diagrammi UML e la rappresentazione semantica, se esistono. La procedura descritta è basata su un precedente lavoro sul riconoscimento automatico di pattern algoritmici [Di Martino e Kessler, 2000]. Il nostro scopo è quello di applicare tale tecnica al supporto del refactoring di software legacy e al porting verso piattaforme di Cloud Computing [Di Martino e Cretella, 2012].

2. Background e stato dell'arte

Di seguito forniamo una panoramica generale sui tool e sulle procedure utilizzati in letteratura per riconoscere i pattern da codice sorgente. Forniamo inoltre alcuni dettagli sui linguaggi e sui modelli semantici che utilizziamo nella nostra procedura.

2.1 Analisi dei tool esistenti

In letteratura sono descritti diversi tool e tecniche di riconoscimento di pattern. Il "Web of Pattern Project" [WoP,2013] usa la rappresentazione base di ODOL per rappresentare un set limitato di pattern, utilizzando la Java reflection e l'analisi dell'Abstract Syntax Tree (AST) per riconoscere pattern in progetti Java Eclipse. Le limitazioni imposte dalla mancanza di una descrizione dinamica dei pattern ci ha portato ad estendere l'ontologia base di ODOL e a definire un nuovo tool, meno legato ad un linguaggio di programmazione specifico o ad un framework. MARPLE-DPD [Marple, 2013] usa un linguaggio per i pattern interno chiamato DiSCO, descritto in [Tahibi, 2007]. DiSCO può essere usato per descrivere sia la struttura sia la dinamica dei pattern, ma esso risulta molto meno flessibile ed estendibile di altri formalismi basati sulla semantica. Inoltre, mentre i linguaggi semantici sono supportati da numerosi tool, DiSCO è sfruttato solo in framework interni. Pinot [Pinot,2012] usa il linguaggio logico SPINE, basato su Prolog, che sfrutta il motore inferenziale Hedgehog [Tahibi,2007]. SPINE è in grado di descrivere unicamente gli aspetti strutturali di un pattern, come confermato dagli stessi autori, per cui esso è in grado di supportare la definizione di un loro numero limitato. Altri tool sono presentati in [Essere, 2013], insieme ad una valutazione delle loro caratteristiche e delle loro abilità di riconoscimento.

2.2 Descrizione dei linguaggi e dei modelli semantici utilizzati

Il Web Ontology Language (OWL) [Owl, 2013] è un linguaggio per il Semantic Web progettato per rappresentare conoscenza riguardo oggetti, gruppi di entità e relazioni esistenti tra di essi. OWL può essere elaborato attraverso programmi e tool per verificare la consistenza di basi di conoscenza preesistenti o per inferire nuova conoscenza. I documenti OWL, noti come ontologie, possono essere pubblicati sul Web e possono fare riferimento o essere referenziati da altre ontologie. OWL-S [Owl-S,2013] è a sua volta una ontologia, all'interno del framework basato su OWL per il Semantic Web, che descrive Web Services semantici. Esso permette agli utenti di scoprire, invocare, comporre e monitorare risorse web che forniscono servizi, sotto vincoli specifici. Può essere facilmente utilizzato per descrivere workflow di servizi, organizzando il loro ordine di esecuzione attraverso strutture per il controllo di flusso e delle decisioni, rendendo inoltre possibile annotare i servizi stessi rispetto ad ontologie OWL di riferimento.

ODOL [Dietrich e Elgar, 2007] è un modello semantico per la rappresentazione di pattern, in grado di descrivere i suoi partecipanti, le sue

proprietà e le relazioni esistenti tra di essi, utilizzando una ontologia OWL. L'ontologia può essere facilmente utilizzata per descrivere la struttura di un pattern possibilmente, ma non necessariamente, derivandola da un diagramma delle classi UML. L'ontologia definisce classi specifiche capaci di descrivere nei dettagli i partecipanti di una classe, insieme ai loro metodi e attributi interni, come pure le differenti relazioni esistenti tra di essi.

3. Descrizione della procedura di riconoscimento dei pattern

Di seguito descriviamo nei dettagli la procedura proposta (vedi Fig 1.). Possiamo definire la procedura come un processo in tre passi: (1) rappresentazione interna dei pattern, (2) rappresentazione interna della documentazione UML relativa ai software artifact analizzati, (3) analisi delle rappresentazioni e matching.

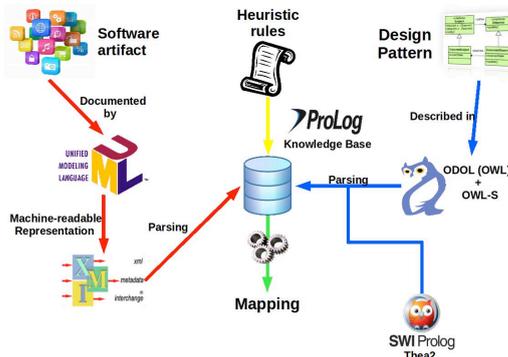


Fig. 1: Schematizzazione della procedura

Durante il **primo passo**, abbiamo bisogno di descrivere formalmente i pattern utilizzando un linguaggio adatto, per poi rappresentarli (internamente) attraverso fatti prolog. In questa fase utilizziamo ODOL per descrivere gli aspetti strutturali dei pattern, semplicemente istanziando classi, proprietà e relazioni OWL definiti nel modello. Usando la rappresentazione di ODOL è anche possibile costruire un catalogo online di pattern, che può essere consultato in ogni momento per ricercare descrizioni di pattern che siano compatibili con OWL. Inoltre, poiché la documentazione OWL standard (e così ODOL) non è in grado di definire workflow, utilizziamo OWL-S per descrivere gli aspetti dinamici dei pattern: sfruttando le strutture di controllo di flusso messe a disposizione per gestire l'ordine delle chiamate tra i partecipanti del pattern, ne formalizziamo il flusso di elaborazione.

Le descrizioni OWL-S sono generalmente utilizzate per annotare semanticamente Web Service, per cui occorre cercare delle corrispondenze tra i metodi dei partecipanti e i servizi OWL-S. Ciò può essere ottenuto definendo una corrispondenza uno ad uno tra metodi e servizi, esplicitando l'associazione tra condizioni di esecuzione e parametri di ingresso e uscita nelle due diverse rappresentazioni. Inoltre, occorre definire un processo OWL-S composto per descrivere l'ordine di esecuzione di ciascun servizio semplice identificato. Una

volta che abbiamo definito il pattern utilizzando una combinazione di ODOL e OWL-S, dobbiamo riempire la base di conoscenza con in fatti appropriati: ciò può essere ottenuto utilizzando il framework Thea, che è in grado di ricavare fatti prolog a partire da un qualunque documento OWL. Poiché Thea non supporta OWL-S, può risultare necessario utilizzare librerie esterne.

Nel **secondo passo**, a partire da artefatti software, dobbiamo utilizzare la documentazione UML esistente, come diagrammi delle classi e delle interazioni, oppure dobbiamo creare nuova documentazione, manualmente oppure automaticamente sfruttando dei tool esterni di analisi del codice. Qui possiamo assumere che la documentazione UML sia già disponibile. A partire dalla documentazione in possesso, dobbiamo derivarne una rappresentazione XMI: la maggior parte dei tool UML esistenti sul mercato è in grado di fornire questa funzionalità. D'altro canto, esistono diverse versioni di XMI e non tutti i tool sono in grado di supportare. Per esaminare i documenti XMI abbiamo definito un parser che analizza l'albero XMI e genera fatti prolog, utilizzando un modello interno in grado di descrivere diverse tipologie di diagrammi UML in logica del primo ordine. Il parser è in grado di distinguere le differenti versioni di XMI e di esaminare i corrispondenti documenti in maniera concorde. Unendo i fatti prodotti dal parser XMI e dal framework Thea, siamo in grado di riempire la base di conoscenza comune.

Nel **terzo passo**, utilizzando regole prolog che implementano le euristiche e catturano le caratteristiche dietro il processo di riconoscimento, siamo in grado di trovare le corrispondenze tra diagrammi UML e descrizioni di pattern se, ovviamente, tali corrispondenze esistono.

La nostra rappresentazione interna di UML e quella Thea di OWL sono basate su modelli piuttosto differenti, per cui occorre definire delle regole prolog aggiuntive che permettano alle due descrizioni di comunicare, così da definire più facilmente l'algoritmo di riconoscimento. Queste regole di supporto non sono direttamente responsabili per l'implementazione dell'algoritmo, ma sono usate per semplificare la sua descrizione e per slegarlo dai diversi possibili modelli di rappresentazione di pattern e artefatti software.

3.1 Approcci per l'implementazione dell'algoritmo di riconoscimento

Di seguito descriviamo nello specifico il funzionamento dell'algoritmo di riconoscimento. Per definire il meccanismo da utilizzare per riconoscere i pattern, abbiamo analizzato due approcci possibili.

Possiamo rappresentare entità e relazioni individuate nel diagramma UML e nel pattern utilizzando dei grafi orientati, per poi applicare degli algoritmi di matching fra di essi, basati sulle regole che definiscono uguaglianza tra classi e partecipanti (visti come nodi) e relazioni tra di essi (visti come archi). Il problema che andiamo a definire in questo modo può essere ricondotto ad un problema di isomorfismo tra sotto-grafi, che sappiamo essere NP-completo. Sebbene per alcune classi di grafi il problema possa essere risolto in tempo polinomiale, poiché non possiamo fare ipotesi sulla natura dei diagrammi UML su cui andiamo ad operare, non possiamo avvalerci di queste semplificazioni.

Un'alternativa consiste nel considerare delle liste di mapping accettabili, analizzando preventivamente tutte le classi compatibili con ciascun partecipante e solo le relazioni effettivamente valide ai fini del pattern, come indicato precedentemente, per poi operare per tentativi. Ad esempio potremmo considerare una possibile configurazione di mapping tra partecipanti e classi per poi verificare se le regole di corrispondenza tra relazioni nelle due rappresentazioni sono compatibili con la struttura del pattern. In seguito potremmo modificare la configurazione iniziale, se ciò risultasse possibile, ed effettuare nuove considerazioni in base alle relazioni. Rispetto alla tecnica precedente si ha il beneficio di non dover esaminare tutte le classi del diagramma UML, ma solo quelle che sono già state identificate come compatibili, con l'ulteriore vantaggio di non dover considerare un elevato numero di configurazioni possibili in quanto, per la maggior parte dei pattern, le caratteristiche dei partecipanti sono piuttosto specifiche e difficilmente compatibili con un elevato numero di classi. Anche se, in linea teorica, dovendo considerare tutte le combinazioni possibili di associazioni partecipanti/classi, la complessità computazionale risulta comunque esponenziale, in pratica il problema risulterà molto più semplificato proprio grazie alle considerazioni preliminari fatte sulla natura delle entità coinvolte.

Sfruttando le caratteristiche di Prolog, in particolare il meccanismo di back-tracking che lo contraddistingue e la capacità nativa di elaborare liste di predicati e fatti, possiamo facilmente implementare l'algoritmo seguendo il secondo approccio illustrato.

3.1 Descrizione dell'algoritmo e criteri di mapping

L'algoritmo è basato su un processo di riconoscimento in due fasi.

Durante la prima fase sono individuati tutti i possibili mapping tra relazioni UML e associazioni che coinvolgono partecipanti del pattern: ciò avviene semplicemente confrontando tipologia e direzionalità di tali elementi. Sfruttando il meccanismo di back-tracking di Prolog e la gestione delle liste, siamo in grado di analizzare in breve tempo tutte le possibili configurazioni di mapping tra relazioni UML e associazioni tra partecipanti dei pattern.

Nella seconda fase, sfruttando le corrispondenze individuate al punto precedente, non siamo obbligati ad esaminare tutte le possibili analogie tra classi UML e partecipanti del pattern, in quanto dobbiamo solo verificare se le corrispondenze imposte dai mapping individuati durante la prima fase sono accettabili. Ogni volta che uno di questi mapping tra classi e partecipanti non risulta accettabile, l'intera configurazione che l'ha generato viene scartata e si passa ad analizzare quella successiva. Onde evitare di controllare lo stesso mapping più volte, vengono generati dinamicamente opportuni fatti prolog.

Generalmente una classe UML e un partecipante di un pattern sono considerati compatibili se la prima contiene almeno gli stessi attributi e metodi del secondo. Le relazioni tra classi e partecipanti non sono considerati durante questa fase, poiché le conoscenze relative sono già state sfruttate durante la prima fase. Ovviamente non possiamo prendere in considerazione per il matching i nomi degli attributi e dei metodi, ma solo la tipologia dei campi e dei

parametri di ingresso e uscita. Alcuni pattern descrivono partecipanti la cui struttura, in termini di attributi e metodi, dipendono dall'esistenza di particolari elementi nell'implementazione reale: ciò vuol dire che, accanto alle regole di riconoscimento generale, è possibile definire condizioni di matching più specifiche che dipenderanno dal particolare pattern esaminato.

Tutti i mapping trovati usando questo algoritmo basato su regole dipendono dalle similarità esistenti tra struttura del pattern e diagramma UML di una sua implementazione. Utilizzando le conoscenze derivanti dall'analisi dei diagrammi delle interazioni UML è possibile eliminare uno o più mapping, raffinando il processo di riconoscimento. Un mapping può essere accettato se, usando le corrispondenze tra classi, partecipanti e relazioni nei diversi modelli di riferimento, individuate dall'algoritmo, è possibile ricostruire il workflow del pattern così come definito dal documento OWL-S corrispondente.

4. Un esempio pratico

In questo esempio utilizzeremo una implementazione del pattern di concorrenza Active Object, definito in [Lavender e Schmidt, 1996] (vedi Fig. 2). Active Object è stato creato per disaccoppiare l'invocazione dall'esecuzione di un metodo, così da migliorare la concorrenza e semplificare gli accessi sincronizzati ad oggetti risiedenti in thread di controllo diversi. L'implementazione a cui facciamo riferimento è descritta nei dettagli in [Huston, 2003]. Il diagramma UML (vedi Fig 3.) corrispondente è alquanto diretto e semplice da analizzare: confrontandolo con la definizione di Active Object in [Lavender e Schmidt, 1996] o con una rappresentazione grafica conforme, un utente umano sarebbe facilmente in grado di riconoscere il pattern.

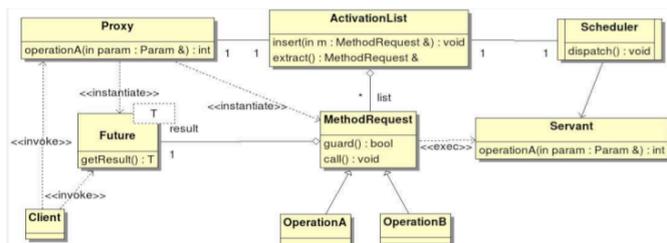


Fig. 2 Diagramma delle classi che definisce una implementazione di Active-Object

Come ci aspettiamo da un artefatto software reale, non tutte le classi presenti nel diagramma sono legate al pattern esaminato: le classi “ACE_Task_Base” e “CompletionCallBack” non hanno corrispondenze nell’Active Object. Inoltre, “CompletionCallBack” viene indicato come una possibile implementazione del pattern “Observer” [Gamma et al, 1995].

Per rendere il processo di riconoscimento automatico, scriviamo prima la rappresentazione XMI del diagramma delle classi e poi, usando il nostro parser, definiamo un set di fatti prolog che sia compatibile con il nostro modello interno. Forniamo, come esempio, alcuni dei fatti prodotti dal parser.

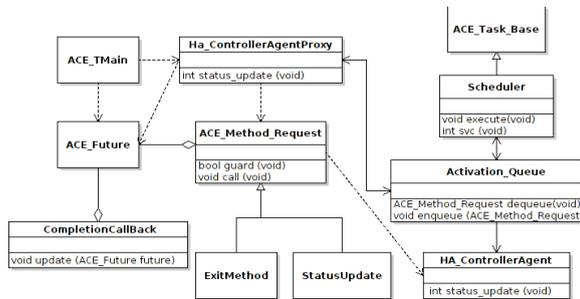


Fig. 3 Diagramma delle classi del software usato nell'esempio

```

- uml_class("','ACE_Method_Request',[],[[guard:Boolean,[]],[call:Void,[]])

```

Il predicato descrive una classe UML chiamata "ACE_Method_Request", per cui non è definito un package (il primo campo è un ""), senza attributi ([] nel terzo campo indica una lista di attributi vuota), e con due metodi ("guard" ha un tipo di ritorno booleano, mentre per "call" + void; nessuno dei due ha dei parametri di ingresso).

```

- uml_dependency("','ACE_Method_Request','HA_ControllerAgent)

```

Questo predicato descrive una dipendenza anonima, (il primo campo " rappresenta il nome della dipendenza), tra una classe client "ACE_Method_Request" e la classe server "HA_ControllerAgent", entrambe localizzate in package root (indicate dal secondo e terzo campo ").

Una volta che il parsing della sorgente XMI è stato completato e che tutti i fatti prolog generati sono stati inseriti nella base di conoscenza, possiamo lanciare il processo di riconoscimento vero e proprio. Assumiamo che tutti i fatti che descrivono i pattern e le regole di matching necessarie siano già stati impostati nella base.

Come abbiamo affermato prima, è possibile creare dei semplici predicati di appoggio per supportare il processo di interrogazione della base di conoscenza, utili anche perché la rappresentazione di Thea è piuttosto difficile da navigare. Per esempio, immaginiamo di voler conoscere le associazioni tra partecipanti nel pattern, un predicato semplice può essere definito nel modo seguente:

```

- participantAssociation(Pattern,Association,Client,Supplier,Directed):-
classAssertion('AO.owl#AssociationTemplate',Association),
objectPropertyAssertion(' AO.owl#client',Assoc,Client),
objectPropertyAssertion('AO.owl#supplier',Assoc,Supplier),
dataPropertyAssertion('AO.owl#isDirected',Assoc,Directed).

```

In questo modo possiamo trovare tutte le associazioni in un pattern, insieme ai partecipanti coinvolti, i loro rispettivi ruoli e la direzionalità della relazione. Una volta definita tale regola, possiamo evitare di interrogare direttamente il modello Thea, in quanto è sufficiente utilizzare la testa della regola e dimenticarci del suo corpo. Ciò significa anche che, se cambiassimo modello di riferimento, l'algoritmo potrebbe funzionare correttamente purché aggiornassimo i predicati di supporto.

Il prossimo aspetto è rappresentato dalle regole di matching che sono usate dall'algoritmo di riconoscimento. Abbiamo definite un set di regole che comparano la rappresentazione UML con la descrizione del pattern. Per esempio, consideriamo la seguenti regole.

– *compliantAssociation(UML_Relation,Pattern,Pattern_Relation):-
participantAssociation(Pattern,Pattern_Relation,Client,Supplier,Directed),
uml_association(UML_Relation,_,AssCL,_,_,AssServ,_,Directions),
rightDirection(Directed,Directions).*

La regola è usata per individuare associazioni compatibili nel diagramma UML e nella descrizione del pattern, interrogando la base di conoscenza. Il predicato di supporto “rightDirection” confronta la direzionalità delle associazioni.

– *compliant(Package,Class,Pattern,Participant):-
attributeCompliant(Package,Class,Pattern,Participant),
methodCompliant(Package,Class,Pattern,Participant).*

La regola precedente definisce classi UML e partecipanti di Pattern compatibili esaminando i loro attributi e metodi, secondo le regole di riconoscimento introdotte precedentemente.

Nessuna delle regole definite precedentemente si riferisce in maniera specifica al pattern esaminato: esse possono essere usate per ogni altra rappresentazione di pattern senza alcuna differenza. Se esaminiamo il pattern Active-Object, possiamo verificare che il partecipante “Proxy” ha una caratteristica peculiare: tutti i suoi metodi corrispondono a sottoclassi del partecipante “Method_Request”. Dunque, possiamo introdurre una regola specifica:

– *isAOProxy(Proxy,Method_Request):-
setof(Method,uml_class(_,Proxy,_,[_:Method,_]),Methods),
setof(SubClass,uml_generalization(_,_,Request,_,Method_Request),SubClasses),
subList(Methods,Requests).*

“Proxy” e “Method_Request” sono anche classi UML che si candidano ad essere mappate negli omonimi partecipanti del pattern. Il modello prolog di UML viene interrogato per trovare tutti i metodi appartenenti alla classe Proxy e tutte le sottoclassi di Method_Request. In seguito, usando il predicato subList, esso verifica se la lista di metodi individuati contiene almeno gli elementi corrispondenti alle sottoclassi elencate.

Il riconoscimento del pattern può essere lanciato, una volta che sono state caricate nel database prolog tutte le informazioni relative ad esso e al software artifact analizzato, semplicemente invocando una regola prolog.

– *matching(PatternRelations,UmlRelations,ClassMapping):
participantRelationList(_,PatternRelations),
relationMatching(_,PatternRelations,[],UmlRelations,ClassMapping).*

5. Conclusioni

La procedura di riconoscimento automatico dei pattern e la relativa implementazione proposte consentono di analizzare e identificare i diversi

componenti di un Software Artifact, fornendo un utile strumento di supporto al refactoring del software. La genericità delle regole di riconoscimento descritte, insieme alla netta separazione imposta tra definizione dell'algoritmo e modelli di rappresentazione usati, rende la procedura facilmente adattabile, sia nel caso di modifiche agli stessi modelli sia in seguito a modifiche dell'algoritmo utilizzato.

Ringraziamenti

Questo lavoro è stato supportato dal Progetto di Rilevante Interesse Scientifico e Tecnologico anno 2009 dal titolo "Fruizione assistita e context aware di siti archeologici complessi mediante dispositivi mobili", finanziato dalla Seconda Università degli Studi di Napoli.

Bibliografia

- [Huston, 2003]Huston et al, "ACE Programmer's Guide, The: Practical Design Patterns for Network and Systems", 2003
- [Dietrich e Elgar,2007] Dietrich, Elgar, "An Ontology Based Representation of Software Design Patterns", capitolo XIII di [Tahibi, 2007]
- [Lavender e Schmidt, 1996] Lavender, Schmidt, "Active Object: An Object Behavioral Pattern for Concurrent Programming", 1996
- [Beyer e Lewerentz, 2003] Beyer, Lewerentz. "CrocoPat: Efficient Pattern Analysis in Object Oriented Programs". IWPC 2003.
- [Gamma et al, 1995] Gamma et al, "Design patterns: Elements of reusable object-oriented software",1995.
- [Tahibi, 2007] Toufik Tahibi, "Design Patterns Formalization Techniques", IGI Publishing, 2007
- [Nija e Olsson, 2003] Nija Shi, Ronald A. Olsson, "Reverse Engineering of Design Patterns for High Performance Computing", University of California, U.S.A
- [Di Martino e Cretella, 2012] Beniamino Di Martino, Giuseppina Cretella, "Towards a Semantic Engine for Cloud Applications Development Support", Proc. of CISIS-2012: The Sixth International Conference on Complex, Intelligent, and Software Intensive Systems, July 4-6th 2012, Palermo, Italy, pp. 198-203, IEEE CS Press, 2012
- [Di Martino e Kessler, 2000] B. Di Martino and C.W. Kessler, "Two Program Comprehension Tools for Automatic Parallelization", IEEE Concurrency, Vol. 8, n. 1, pp. 37-47, Jan-Mar. 2000.

Sitografia

- [Essere, 2013] Essere: Evoluzione dei Sistemi Software e Reverse Engineering, <http://essere.disco.unimib.it/>, verificato Luglio 2013
- [OWL,2013] Web Ontology Language (OWL), www.w3.org/TR/owl-features/, Luglio 2013
- [OWL-S,2013] OWL-S:Semantic Markup for Web Services, www.w3.org/Submission/OWL-S/, verificato Luglio 2013
- [Thea, 2013] Thea, an OWL2 Prolog library, www.semanticweb.gr/thea/index.html, verificato Luglio 2013
- [WOP, 2013] The Web of Patterns Project, <http://www-ist.massey.ac.nz/wop/>, verificato Luglio 2013
- [Pinot, 2013] Pinot:Pattern Inference and Recovery Tool, <http://www.cs.ucdavis.edu/~shini/research/pinot>, verificato Luglio 2013

The Knowledge Experience Base as a support for the global development of software

Pasquale Ardimento, Danilo Caivano¹, Nicola Convertini², Giuseppe Visaggio³

¹ Università degli Studi di Bari "Aldo Moro" Department of Informatics
Via Orabona, 4 70125 Bari

pasquale.ardimento@uniba.it, danilo.caivano@uniba.it, giuseppe.visaggio@uniba.it

² Ser&Practices s.r.l. Department of Informatics
Via Orabona, 4 70125 Bari
nicolaconvertini@gmail.com

Abstract. *As it is known in the production process of software, a lot of knowledge and experience are created and remain silent; therefore, it cannot be reused to improve the effectiveness and the efficiency of these processes and can not be transferred between practitioners improving the capabilities of each one. The capitalization and transfer of knowledge and experience ensures the compliance of the different teams of the company and ensures the ability to continuously improve the quality of processes and products and the enhancement of skills of all developers the same company. To support the capitalization and the transfer of knowledge and experience we propose a Knowledge Experience Base (KEB) with Knowledge and Experience Package (KEP) that are enriched and improved incrementally with the use of content thanks to lessons learnt. The paper not only describes the structure and the content of the KEP but also shows an empirical research that currently validates its effectiveness in knowledge and experience transfer and the resulting lessons learned.*

Keywords: Knowledge Management, Knowledge Management Practices, Software Development

1. Introduction

Over the last decade software engineering, industry has started to use progressively nationally and globally distributed way to carry out software development projects (Herbsleb and Moitra, 2001).

Knowledge Management (KM) is a very important skill of an organization and implicates human resource, enterprise organization and culture, so as the information technology, methods and tools that support and enable it. For managing knowledge into the organization, a proper KM Strategy needs to be

Congresso Nazionale AICA 2013

developed. KM is a strategy that has coordinated and defined plan of actions to enable, in core business processes, the use of KM techniques. Nonaka model (Nonaka and Konno, 1998) for knowledge creation and sharing; knowledge repositories for reuse; document management systems; yellow pages for competence management are well suited in DSD environments. Software companies have large quantity of knowledge resided in their methodologies, processes, culture, people and working environment. There is a need to spread knowledge in/across organizations in order to understand, manage, and identify customer and business needs, know-how of culture and co-ordinate throughout the SDLC effectively (Rus and Lindvall, 2002).

This paper is structured as follows:

In section II are presented the related works; in section III is presented PROMETHEUS, a KEB (Knowledge Experience Base) made up of KEPs (Knowledge Experience Packages) to support knowledge delivery for software development; in section IV is presented a controlled experiment to measure whether the knowledge transferred by a KEP, and delivered by PROMETHEUS, from one development site to another one make more understandable than the knowledge transferred by traditional documentation.

2. Related works

Software development is a very multifaceted, knowledge intensive and rapidly changing activity with different people involved in it (Nicholson and Sahay, 2004). Likewise, as time evolves technology changes, processes changes, development methodology changes and people interest changes; as a consequence there is a need to retain this knowledge properly in company for reuse in future. In DSD, the teams have lot of inter-dependency of work. To ensure no rework, learning from experiences, not repeating the mistakes and utilizing the resources properly, KM is essential. KM strategies can help manage the knowledge effectively and will abate the problems faced in DSD. In this section, an overview to the discipline of KM is given to understand what KM is and the theories that can support to achieve the study's goal to comprehend the importance of KM in SD. It has been argued that knowledge has become a key asset to an organization that drives organizational survival and success (Drucker, 1992). KM has its roots in a number of disciplines like Information Systems, Human Resource Management and in Sociology, Anthropology, Economics, Philosophy, Management Science and Computer Science (Jashapara, 2004). So the strength and challenge of KM comes from its interdisciplinary approach. The most common notion of knowledge in the current KM literature is Tacit and Explicit Knowledge (Nonaka, 1994). These notions have its origins in the ideas of logical behaviorism of Ryle (Ryle, 1949) and Polanyi (Humphrey, 1997). In the think of this authors, knowledge exists along a continuum between tacit (know-how) and explicit (know-that) knowledge. Nonaka (Nonaka, 1994) marks a distinction between tacit and explicit knowledge based on Polanyi's original categories as:

- Explicit knowledge – It is the knowledge of sequential and rationality knowledge.

- Tacit knowledge – It is the knowledge about experience—simultaneously knowledge and practice.
- Locus of Knowledge: Knowledge exists in individuals, as well as there is collective knowledge.
- Individual Knowledge – Individual knowledge is ‘driven’ by an individual (Ricardo, 2008). It is the knowledge possessed by individual in minds.
- Collective Knowledge – Collective knowledge is shared and distributed between members of a group.

3. Model for creation and sharing

A KEP is an organized set of: knowledge content, teaching units on the use of the demonstration prototypes or tools and all other information that may strengthen the package’s ability to achieve the proposed goal. The KEP, in figure 1, must be usable independently of its author or authors, and the content must then have a specific structure: distance education and training must be available through an e-learning system. A user can access one of the package components and then navigate along all the components of the same package according to her/his needs.

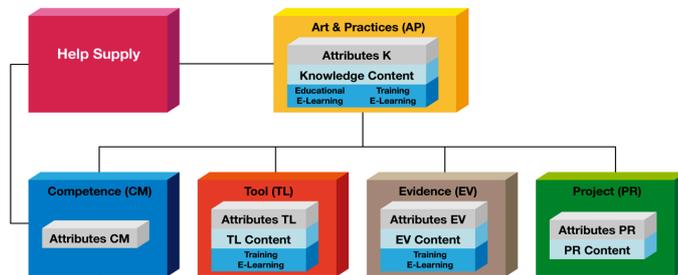


Figure 1: KEP Structure

The KEP does not contain the conceptual basis of the subject, because it is considered to be the background of the user's knowledge, and can be found in traditional sources of knowledge such as technical reports, papers and books. When users need some of the basic concepts for understanding the contents of KEP they can use an educational e-learning course, and if users should need more information, they can use the "attachments" regarding reports, papers and books concerning the basic topics of KEP. When a package also has support tools, rather than only demonstration prototypes, Knowledge Content (KC) links the user to the available tool. The tools are collected in the Tools Component (TL). A KEP is generally based on conjectures, hypotheses and principles and when it is “mature” its content must all become principle-based. The transformation of a statement from conjecture through hypothesis to principle must be based on experimentation showing evidence of validity. The experimentation, details of its execution and relative results, are collected in the Evidence component (EV). In the case the user needs support from those who have knowledge of the contents of KEP, a list of resources is provided as a

reference. The list is collected in the Competence component (CM). The following sections describe attributes, and further details are available in the following publication (Ardimento and Cimitile, 2008) (Ardimento et al, 2011a) (Ardimento et al, 2011b).

3.2. Attributes

A search within the package starting from any of its components is facilitated by the component's. Each component in the knowledge package has its own attribute structure. In the case of all the components, the attributes allow a rapid selection of the relative elements in the knowledge base. To facilitate the research, we used a set of selection classifiers and a set of descriptors which summarize the contents. The summary descriptors include: a brief summary of the content and a history of the events occurring during the life cycle of the package, The history may also include information telling the reader that the content of all or some parts of the package are currently undergoing improvements. The classifiers include: the adoption threats of the technological innovation for which it is a provider; the risk mitigation initiatives that ensure a better performance of the KEP in the solution of the main problems; the impact that the KEP will have on the active processes of the production lines to which it will be applied, supposing that the difficulties to be solved correspond with those in the KEP; a forecast of the Return of Investment that the new introduction will have in the company. Both the economical impact of the KEP and its impact on the value chain are therefore specified; the acquisition plan of the KEP methods; the history of the KEP, i.e. the set of practices that have required its use and the results obtained from their application in order to ensure a higher perception of the KEP's reuse.

4. Empirical investigation

In this section the main details of the experiment will be detailed. The experiment material is available at the Prometheus website (Prometheus, 2011). The main goal of this experiment is to assess whether the knowledge, about software documentation, transferred by a KEP, and delivered by PROMETHEUS, is more comprehensible than the knowledge transferred by traditional documentation in a distributed software development scenario. We used the GQM template to define the goals of our experiment. This is defined as follows: "Analyze the comprehension of KEPs for the purpose of comparing it with the comprehension of traditional documentation with regard to the effectiveness and efficiency from the viewpoint of IT software engineers in the context of professional IT stakeholders from the University of Castilla-La Mancha".

4.2 Planning

Various issues related to the planning of the experiment are introduced in this section.

Subjects Selection. The experiment was carried by two groups located in different sites: one in Italy and the other one in Spain. The Italian team located in Bari was made up of 4 IT professionals contracted in projects led by the Software Engineering Research LABORatory (SERLAB) group at the University of Bari Aldo Moro (Italy). The Spanish team was made up of 14 IT professionals contracted in projects led by the Alarcos research group at the University of Castilla-La Mancha (Spain), all of whom have degrees in computer science. The KEP selected, referred to basic software engineering knowledge such as: software maintenance, the software development lifecycle and so on. The topic was therefore comprehensible to each single experimental subject..

Sample assignment. To compare the comprehension of a KEP with the comprehension of traditional documents, it was necessary to select more than one experimental group: When there is more than one group in an empirical experimentation it is important that the groups are balanced as regards a set of specific characteristics whose values could affect the empirical results. If the groups are not equal, then it is not possible to know whether the independent variable has caused changes in the subjects or whether this has occurred because of the inequality. In our investigation, we identified two control variables: the level of instruction, and knowledge of the English language. The matched sample was applied by asking the subjects to fill in a background questionnaire. The results obtained were used to select two groups whose standard deviation was the lowest with regard to the two control variables.

Independent and dependent variables. In accordance with certain suggestions concerning the measurement of comprehension (Juristo and Moreno, 2000), we have used the Cognitive Theory of Multimedia Learning (CTML) (Mayer and Moreno, 1998). This choice was made for several reasons. First, it focuses on words and graphics, which are the elements used in papers and in KEPs. Second, it provides principles for the design of effective multimedia presentations which can be tested empirically. Third, it has evolved through years of work and development of experimental instruments and methods related to model comprehension. By following CTML, the comprehension has been defined through three dependent variables:

- Semantic Comprehension: the ability to comprehend the semantic material being presented.
- Retention: the comprehension of the material being presented, and the ability to retain knowledge from it.
- Transfer: the ability to use the knowledge gained from the material to solve related problems which are not directly answerable from it.

The quality model, defined according to the GQM approach presented in Fig. 2 which also indicates the experimental hypothesis defined

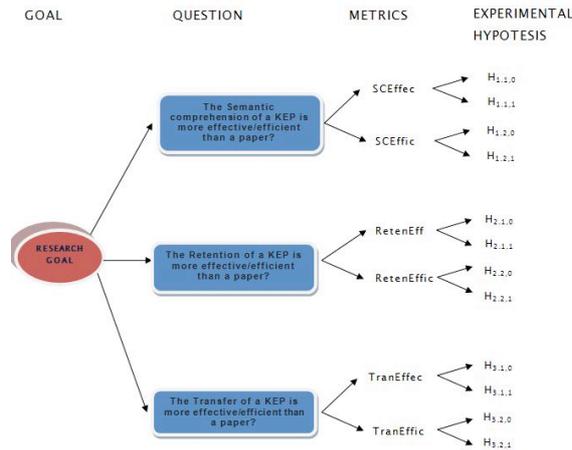


Figure 2: Quality Model and Experimental Hypotheses

Each dependent variable was measured in terms of :Effectiveness: the proportion of correct answers provided in each test (number of correct answers / number of questions). This measure reflects the ability to understand the material presented correctly. It is important to note that in our analysis, the unanswered questions are considered to be wrong answers and Efficiency: the proportion of correct answers divided by time (Effectiveness / Time).

In this work, we therefore defined the following measures: SCEffec/SCEffic (Semantic Comprehension Effectiveness/ Semantic Comprehension Efficiency), TranEffec/TranEffic (Transfer Effectiveness/ Transfer Efficiency), and finally, RetenEffec/RetenEffic (Retention Effectiveness/ Retention Efficiency).

Design of experiment. The design used was a "Between-subject design" in which each group is subjected to one treatment in one run. The knowledge transferred refers to the topic of Reiterative reengineering. All the subjects were assigned to 2 groups (A and B). The experiment consisted of one round. Two different diagrams were presented to every subject in each group. The design is shown in Table 1.

Topic	Treatment	
	KEP	Traditional documentation
Iterative reengineering	Group B	Group A

Table 1: Between-Subjects Design

Experiment objects. There were two experimental objects, related to the two values of the independent variable (KEP, traditional documentation). One consisted of a set of four documents on the "Iterative Reengineering" topic, whilst the other was a KEP made up of one AP component related to one EV

component. Both the AP and the EV components were built by analyzing the 4 previously mentioned documents.

All the documents represent the formalization of the interactive reengineering process. The documents have been written by the development pool of Bari and sent to the Spanish development pool to be used within their own production activity. Each single document was written by one different author to aderire il più possibile a scenari reali di sviluppo distribuito del software. It is important to note that we selected the topic so that no experimental subject had any specific knowledge of it.

Hypotheses. For Semantic Comprehension we wished to test the following hypothesis:

- Null hypotheses: H1.1,0: KEPs do not improve SCEffec with regard to those subjects using papers; H1.2,0: KEPs do not improve SCEffic with regard to those subject using papers.
- Alternative hypotheses: H1.1,1 = \neg H1.1,0 and H1.2,1 = \neg H1.2,0

We analogously formulated a set of hypothesis H2 for Retention measures (RetenEffec and RetenEffic), and another set, H3, related to the Transfer measures (TransfEffec and TransEffic) (see Fig. 1).

Instrumentation. In the experiment we used three questionnaires for both the treatments, one for each dependent variable. For each questionnaire the subjects were asked to write down the starting time and the finishing time.

4.3. Post-experiment Survey

The goal of this survey (see Table 2) was to obtain feedback about the subjects' perceptions of the execution of the experiment. The 5-point Likert item ranges from strongly agree (1) to strongly disagree (5), The items concerned the perceived level of difficulty of: each single questionnaire, whose corresponding items are A1, A2 and A3, as shown in Table 2; the adequateness of the training received, A4; the clearness of the material provided during the experiment execution, A5; and the clearness of all the questions provided during the experimental execution, A6.

Id	Assertion	values
<i>A1</i>	During the experimental run question. 1 was easy	(1-5)
<i>A2</i>	During the experimental run question. 2 was easy	(1-5)
<i>A3</i>	During the experimental run,question. 3 was easy	(1-5)
<i>A4</i>	The training received in the two-hour lecture which took place the day before the experiment was sufficient to execute the experimental run	(1-5)
<i>A5</i>	During the experimental run, all the materials provided were clear	(1-5)
<i>A6</i>	During the experimental run, all the questions were clear	(1-5)
Five-level Likert scale: 1 = strongly agree; 2 = agree; 3 neutral; 4 = disagree; 5 = strongly disagree		

Table 2: Post Experiment Survey

4.4. Experiment Execution

The experiment took place in a two-hour session. In the first 30 minutes we explained how to perform the experiment and assigned the subjects to each group using the matched sample assignment based on the data collected from the background questionnaire. The experiment was conducted in a laboratory, where the experimental subjects were supervised, and no communication among them was allowed. The experiment was executed as follows: the subjects in group A downloaded the four selected papers. After logging onto PROMETHEUS, the subjects in group B selected the KEP concerning the Iterative reengineering topic. Immediately after all the subjects had finished these operations, the supervisor gave each one of them questionnaire 1. After a subject stated that s/he had finished questionnaire 1, s/he received questionnaire 2. After a subject stated that s/he had finished questionnaire 2, s/he received questionnaire 3. Only in this experimental case were the subjects unable to use the KEP or Papers, previously received, to answer the questions. After a subject stated that s/he had finished questionnaire 3, s/he received the post-experiment survey. Each subject immediately handed in each questionnaire to the supervisor upon its completion.

5. Presentation and analysis

In this section, we present, for each dependent variable the results of hypothesis testing and the analysis of the data collected from the post-experiment survey.

The Mann-Whitney U Test was used to test the formulated hypotheses because all its assumptions were satisfied :

- each dependent variable measured at least on an ordinal measurement scale; the independent variable had only two levels; a between-subjects design was used; the subjects were not matched across conditions.

Measures of dependent variables	Rank Sum KEP	Rank Sum papers	U	p-value	Statistical difference
<i>SCEffec</i>	31,50	23,50	8,5	0,4033	NO
<i>SCEffic</i>	37,00	18,00	3	0.0472	YES
<i>TranEffec</i>	29,50	25,50	10,5	0,6761	NO
<i>TranEffic</i>	32,00	23,00	8	0.3472	NO
<i>RetenEffec</i>	38,50	16,50	1,5	0,0215	YES
<i>RetenEffic</i>	39,00	16,00	1	0,0162	YES

Table 3: Mann Whitney analysis

A statistically significant difference exists between the Semantic Comprehension Efficiency and Retention Efficiency variables. The efficiency, which represents the proportion of correct answers divided by time, of Semantic Comprehension and Retention shows a statistically significant difference between the KEP and the traditional documents. This result is probably owing to

the fact that the traditional documents contain more and redundant information, in addition to the fact that they are not as structured as KEPs. Nevertheless, Transfer Efficiency, measured by the questionnaire made up of open questions, has no statistical difference. This measure was collected using a questionnaire made up of open questions for which we chose questions that were as simple as possible owing to the time available: only about one hour and forty-five minutes to answers all the questionnaires. This consideration is also supported by the post-experiment analyses in which the subjects agreed that this questionnaire was easy. With regard to the Effectiveness, there is no statistically significant difference between KEP and the papers. This result is probably owing to the fact that the selected documents were written by the same authors and did not therefore contain incoherent information. In general, traditional documents written by different authors could contain different and incoherent information despite dealing with the same topic.

6. CONCLUSIONS

In this work, blend of KM practices and Carmel's centripetal forces as proposed solution to the problems faced in DSD are given. This paper continues our research into supporting knowledge transfer between different stakeholders. More precisely, in this paper we have analyzed whether the knowledge formalized and transferred in the form of KEPs is more comprehensible than the same knowledge when formalized and transferred in the form of traditional documents. We have therefore carried out a controlled experiment, from which the following findings have been obtained:

- KEPs are more effective in retaining knowledge than traditional documents. It is important to remember that this measure was collected using the answers provided for questionnaire 3 without the help of any material, KEP or traditional documents. This result was probably affected by the reduced size of a KEP in comparison to the size of traditional documents.
- With regard to semantic comprehension and the transfer of knowledge, the KEP are not more effective than the traditional documents. This result was probably affected by the documents selected, since they were all written by the same authors and the information and concepts expressed were therefore homogenous and coherent between each paper. Moreover, only four papers were selected owing to the time available to carry out the experimentation. If there had been much more than four documents the difference would, in our opinion, have been greater.
- KEPs are more efficient at transferring knowledge than the traditional documents. In other words, it is faster to find the concepts, the empirical evaluations and the empirical data expressed in a KEP than in the traditional documents. This result was probably obtained because the KEP structure provides a set of attributes, defined for each specific component, to allow a rapid selection of knowledge content.

These results have encouraged us to make further studies in order to: analyze other contexts, such as industrial or educational contexts; these contexts are not equal because, for instance, the industrial context has more pressures than the educational context while the latter requires more attention to educational aspects. For completeness, we should state that we are aware that the structure of KEPs requires a supplementary effort in comparison to the papers because it is produced from papers.

References

Ardimento, P., Cimitile, M. An Empirical Study on Software Engineering Knowledge/Experience Packages. In A. Jedlitschka y O. Salo (Eds.), *Product-Focused Software Process Improvement* (Vol. 5089, pp. 289-303): Springer 2008

Ardimento, P., Boffoli V.N, Visaggio, G., "PROMETHEUS: a web platform for supporting knowledge management in an environment based on experience factory", ICDS 2011,, 23-28 February 2011, pp. 182-186, ISBN: 978-1-61208-003-1,

Ardimento, P., Boffoli V.N, Visaggio, G., "PROMETHEUS: A Web Platform To Support Open Innovation", *Information Systems: People, Organizations, Institutions, and Technologies ItAIS: The Italian Association for Information Systems 2011*, Rome 7-8 October 2011

Ashapara, A. (2004) *Knowledge Management: An Integrated Approach*. England: PearsonEducation Limited.

Drucker, P. (1992) 'The new society of organizations', *Harvard Business Review*, September/October, pp.95-105

Herbsleb, J. and Moitra, D. (2001) *Global Software Development*, *IEEE Software*, 18(2), pp. 16-20.

Juristo, N., Moren, A.,. *Basics of Software Engineering Experimentation*, Kluwer Academic Publishers, 2001

Mayer, R. E.; R. Moreno (1998). "A Cognitive Theory of Multimedia Learning: Implications for Design Principles". <http://www.unm.edu/~moreno/PDFS/chi.pdf>. - See more at: <http://www.learning-theories.com/cognitive-theory-of-multimedia-learning-mayer.html#sthash.4RgojPRm.dpuf>.

Nicholson, B. and Sahay, S. (2004) *Embedded Knowledge and offshore software development*, *Information and Organization*, Elsevier, 14(4), pp. 329-365.

Nonaka, I. (1994) *A Dynamic Theory of Organizational Knowledge Creation*, *Organizations Science*, 5(1), pp. 14-37.

Nonaka, I. and Konno, N. (1998) *The Concept of Ba: Building A Foundation For Knowledge Creation*, *California Management Review*, 40(35), pp. 40-56.

Prometheus.com available at <http://prometheus.serandp.com/en/content/documentation-experiment-spain-december-2011>

Ricardo, F. (2008) *Cyberculture and New Media*. Netherlands: Rodopi

Rus, I and Lindvall, M. (2002) *Knowledge management in software engineering*, *IEEE Software*, 19(3), pp.26-38.

Ryle, G. (1949) *The Concept of Mind*. London: Hutcheson.

Textual Analysis and Software Quality: Challenges and Opportunities

Gabriele Bavota¹, Andrea De Lucia², Rocco Oliveto³,
Fabio Palomba², Annibale Panichella²

¹Department of Engineering, University of Sannio
Palazzo ex Poste, Via Traiano, 82100 Benevento, Italy
gbavota@unisannio.it

²Department of Management and Information Technology, University of Salerno
Via Giovanni Paolo II, 84084 Fisciano, Italy
adelucia@unisa.it, fabio.palomba.89@gmail.com, apanichella@unisa.it

³Department of Bioscience and Territory, University of Molise
c.da Fonte Lappone, 86090 Pesche (IS), Italy
rocco.oliveto@unimol.it

Abstract. Source code lexicon (identifier names and comments) has been used – as an alternative or as a complement to source code structure – to perform various kinds of analyses (e.g., traceability recovery). All these successful applications increased in the recent years the interest in using textual analysis for improving and assessing the quality of a software system. In particular, textual analysis could be used to identify refactoring opportunities or ambiguous identifiers that may increase the program comprehension burden by creating a mismatch between the developers' cognitive model and the intended meaning of the term, thus ultimately increasing the risk of fault proneness. In addition, when used “on-line” during software development, textual analysis could guide the programmers to select better identifiers aiming at improving the quality of the source code lexicon. In this paper, we overview research in text analysis for the assessment and the improvement of software quality and discuss our achievements to date, the challenges, and the opportunities for the future.

Keywords: Software quality, Textual analysis, Survey.

1. Introduction

During software development and evolution a variety of software artifacts are created, such as, requirements, bug descriptions, documentation, source code, test cases, etc. These artifacts have different representations and contain different types of information, i.e., structural (e.g., control and data flow), dynamic (e.g., execution traces), process (e.g., CVS logs), and textual (e.g., identifiers and comments in source code, documentation). The textual

Congresso Nazionale AICA 2013

information captures knowledge about the problem and solution domain, about developer's intentions, client demands, etc. and it is the most common type of information present in software. Text is also the common form to represent information among various artifacts at different abstraction levels. Among other things, developers use textual information to understand what a specific piece of code implements and make decisions during their daily tasks. For very small software systems, developers could read all the text found in software artifacts and extract and use only the information that is useful for their current task. However, as the size and complexity of the system increases, tools are required to extract, analyze, and retrieve this information to the developers.

For these reasons, in recent and past years, textual analysis (TA) has been successfully applied to leverage the textual information and help developers in several software engineering tasks, such as traceability recovery [Antoniol et al, 2002], impact analysis [Canfora and Cerulo, 2005], clone detection [Marcus and Maletic, 2001], feature location [Poshyvanyk et al, 2007]. The use of TA in software engineering has demonstrated to be effective for various reasons:

- it is *lightweight* and to some extent *independent on the programming language*, as it does not require a full source code parsing, but only its tokenization and (for some applications) lexical analysis;
- it *provides information complementary to what structural or dynamic analysis can provide* [Marcus et al, 2008];
- it models software artifacts as textual documents, thus *it can be applied to different kinds of artifacts* (i.e., it is not limited to the source code) and, above all, can be used to perform combined analysis of different kinds of artifacts (e.g., requirements and source code), as in the case of traceability recovery.

All these successful applications increased in the recent years the interest in using TA for improving and assessing the quality of software systems. This paper offers an overview of the process that is usually followed when using TA techniques to support software engineering tasks, focusing the attention on activities aimed at improving software quality. Other than discussing the state of the art, the paper also presents challenges and opportunities for the future. Thus, it represents a useful roadmap for both practitioners, who want to know how to use TA in their working environment, and researchers, who want to get closer and doing research on this topic.

Paper structure. Section 2 provides background information on how to extract and manage textual information. Sections 3 and 4 describe how textual information can be used to measure quality aspects, i.e., cohesion and coupling, and identifying refactoring opportunities, respectively. Finally, Section 6 concludes the paper highlighting challenges and new horizons.

2. Background

TA has been proposed for a variety of software engineering tasks, and making use of different text retrieval (TR) techniques. No matter the particular task or

retrieval technique used, an approach based on TR generally follows the same process:

- *extracting text documents from software artifacts (the corpus)*;
- *indexing the corpus*;
- *computing similarity between documents*.

Each step is explained in detail in the following subsections.

2.1 Extracting the corpus

The first step in using TR techniques is to define a collection of text documents, also known as *corpus*, which are extracted from the software artifacts. Documents can be extracted at different granularities from an artifact. For example, in the case of source code, a document could be represented by structural elements of the code, such as a class. In the case of textual software documentation, sentences, paragraphs, sections, or chapters could represent the documents. Thus, a software artifact may be represented by one or more documents in the corpus. The document granularity needs to be decided up front according to the needs of the task at hand and can influence greatly the results of text retrieval.

Once the corpus is extracted, a few optional, corpus normalization steps can be performed before the documents are indexed by the text retrieval technique [Baeza-Yates and Ribeiro-Neto, 1999]:

- *term extraction*, aimed at extracting words from the artifacts and removing anything useless (e.g., punctuation or programming language operators);
- *identifier splitting*, aimed at splitting composite identifiers. This step is important to align source code and documentation vocabulary, since identifiers are often composed of several concatenated dictionary words. The simplest approaches for identifier splitting are based on common conventions for separating words in identifiers, such as using camel case, underscore, numbers and symbols as separators. For example, `SETpointer`, `set_pointer`, `setPointer` would be all split to *set* and *pointer*. More advanced techniques make use of dictionaries and abbreviation lists to identify words in the cases where common naming conventions are not used, e.g., the identifier `setptr` would be split into *set* and *pointer* based on these techniques [Guerrouj et al, 2012];
- *term filtering*, aimed at removing common terms, referred to as “stop words” (e.g., articles, prepositions, common use verbs, or programming language keywords). Words shorter than a given length (e.g., shorter than three characters) are removed as well.

Morphological analysis of the extracted words is often performed to bring back words to the same root (e.g., by removing plurals to nouns, or verb conjugations). The simplest way to do morphological analysis is by using a stemmer, e.g., the Porter stemmer [Porter, 1980]. Other stemmers used by

researchers in software engineering are WordNet's morphstr function¹ and the Snowball stemmer².

2.2 Indexing the corpus

The extracted information is stored in a $m \times n$ matrix (called *term-by-document matrix*), where m is the number of terms occurring in all artifacts, and n is the total number of artifacts in the repository. A generic entry w_{ij} of this matrix denotes a measure of the weight (i.e., relevance) of the i^{th} term in the j^{th} document [Baeza-Yates and Ribeiro-Neto, 1999]. Such weight, independently by the used technique, is based on two criteria: how well they describe the current document (*local weight*) and how they relate to the entire corpus (*global weight*). A widely used measure is the *tf-idf* (*term frequency-inverse document frequency*), which gives more importance to words having a high frequency in a document (*tf*) and appearing in a small number of documents, thus having a high discriminating power (high *idf*).

2.3 Computing similarity between documents

Based on the *term-by-document matrix* representation, different Information Retrieval (IR) methods can be used to compute similarity between documents aiming at deriving latent patterns between them. A survey of available research papers reveals that probabilistic models, Vector Space Model (VSM), its extension Latent Semantic Indexing (LSI), and Latent Dirichlet Allocation (LDA) are the four most frequently used IR methods in software engineering. In the following we describe in details VSM and LSI since these two techniques have been used to assess software quality. However, the interested reader can find more details on IR methods the book by Baeza-Yates and Ribeiro-Neto [1999].

In the VSM, a document is represented by a vector of terms, i.e., column of the *term-by-document matrix*. Since any document contains a limited set of terms, while the vocabulary (all the terms in the documents) can be millions of terms, most document vectors are very sparse and they generally operate in a positive quadrant of the vector space, i.e., no term is assigned a negative value. In VSM the angle between two vectors is used as a measure of divergence between the vectors, and the cosine of the angle is used as the numeric similarity between the corresponding documents. The cosine has a property indicating 1.0 for identical vectors (very similar documents) and 0.0 for orthogonal vectors (completely different document). A common criticism of VSM is that it does not take into account relations between terms [Deerwester et al, 1990], e.g., having “automobile” in one document and “car” in another document does not contribute to the similarity measure between these two documents.

LSI [Deerwester et al, 1990] was developed to overcome the synonymy and polysemy problems, which occur with the VSM model. In LSI the dependencies between terms and documents, in addition to the associations between terms and documents, are explicitly taken into account. LSI assumes that there is an

¹ <http://wordnet.princeton.edu/>

² <http://snowball.tartarus.org/>

underlying or “latent structure” in word usage that is partially obscured by variability in word choice, and uses statistical techniques to estimate this latent structure. For example, both “car” and “automobile” are likely to co-occur in different documents with related terms, such as “motor”, “wheel”, etc. LSI exploits information about co-occurrence of terms (i.e., latent structure) to automatically discover synonymy between different terms.

Specifically, LSI defines a term-by-document matrix as well as VSM. Then it applies the Singular Value Decomposition (SVD) [Deerwester et al, 1990] to project the original term-by-document matrix into a reduced space of concepts. The size of this space is k , that is much lower than n , i.e., number of terms. The cosine of the angle between two vectors in the k -space represents the similarity of the two documents (terms, respectively) with respect to the concepts they share. In this way, SVD captures the underlying structure in the association of terms and documents. Terms that occur in similar documents, for example, will be near each other in the space of concepts, even if they never co-occur in the same document. This also means that some documents that do not share any word, but share similar words may nonetheless be near in the space of concepts. The choice of k is critical and the proper way to make such a choice is an open issue in the factor analysis literature [Deerwester et al, 1990].

The obtained similarity measure is used to support different software engineering tasks. For instance, given a software artifact used as query (e.g., requirement), the similarity measure can be used to find similar source code classes in order to identify traceability links between requirements and source code [Antoniol et al, 2002] or to perform impact analysis [Canfora and Cerulo, 2005]. In the next sections we show how textual similarity can be used to monitor and improve the internal quality of software systems.

3. Capturing cohesion and coupling

Cohesion is a desirable property of software as it positively impacts understanding, reuse, and maintenance. There are several metrics to measure the cohesion of a software component. These metrics are generally based on structural information extracted from the source code, such as attribute references [Chidamber and Kemerer, 1998]. However, information contained in identifiers and comments could be worthwhile to measure the cohesiveness of software component. Specifically, two components are conceptually related if their (domain) semantics are similar, i.e., they perform conceptually similar actions. To this aim, Marcus *et al.* [2008] proposed a semantic measure, called Conceptual Cohesion of Classes (C3), to capture the cohesion of a class inspired by the mechanisms used to measure textual coherence in cognitive psychology and computational linguistics. In order to capture the semantic cohesion of a class, LSI is used to represent each method as a real-valued vector that spans a space defined by the vocabulary extracted from the code. The conceptual similarity between two methods (CSM) is then calculated as the cosine of the angle between their corresponding vectors. Thus, the higher the value of CSM the higher the similarity between two methods. The average value between the CSM of all passable pairs of methods of a class represents the

Conceptual Cohesion of the class. In short, the measure captures relationships between the comments, identifiers, and other text present in the methods, based on word usages in the entire code. It is clear that C3 depends on the consistency of naming conventions used in the source code as well as on the comments contained in it. An empirical study conducted on three open-source systems has been performed to compare the novel metric with a set of existing metrics based on structural information, e.g., Lack of COhesion in Methods (LCOM) [Chidamber and Kemerer, 1998] and Coh [Briand et al, 1998]. The results achieved indicated that C3 captures different aspects of class cohesion compared to any of the existing cohesion measures.

Following the same idea, Poshyvanyk *et al.* [2009] proposed the Conceptual Coupling Between Classes (CCBC). Specifically, they used LSI to measure the similarity between methods of two classes. The average similarity of all possible pairs of the two classes provides an indication of the conceptual coupling between the two classes. The higher the similarity, the higher the coupling between them, i.e., the methods of the two classes perform conceptually similar actions. An empirical study indicated that CCBC provides orthogonal information as compared to coupling metrics based on structural information, e.g., Coupling between Objects (CBO) [Chidamber and Kemerer, 1998].

3.1 Improving defect prediction

Empirical studies have indicated that semantic metrics capture different aspects of cohesion and coupling as compared to structural metrics. This suggests that the combination of semantic and structural metrics is a more complete cohesion indicator than any combination of structural metrics. In order to verify such a conjecture, Marcus *et al.* [2008] used the C3 metric to improve defect prediction models³ based only on structural cohesion metrics. They performed a study where logistic regression models were used to predict the defects of classes of version 1.6 and 1.7 of Mozilla. The considered predictors were the metrics from the CK suite [Chidamber and Kemerer, 1998] and the C3. Results indicated that the three models with the best precision in prediction defect-prone classes are C3+LCOM3, C3+LCOM1, and C3+Coh. Thus, C3 is a useful indicator of an external property of classes in OO systems, that is, the defect proneness of classes. More importantly, the results support the initial conjecture that the combination of C3 with other cohesion metrics allows to build superior models for detecting defect prone classes.

3.2 Capturing developer perception of quality

From the analysis of the literature we can derive that semantic metrics complement structural metrics and are useful to measure the quality of software components. However, little is known about how developers actually perceive quality and if existing measures align with this perception.

³ One of the most interesting applications of cohesion metrics in software engineering is to predict defects in classes.

In order to bridge this gap, Bavota *et al.* [2013a] empirically investigated how a specific quality attribute, namely class coupling – as captured by structural, dynamic, semantic, and logical coupling measures – aligns with developers' perception of coupling. The study has been conducted on three Java open-source systems, i.e., ArgoUML, JHotDraw and jEdit, and involved 64 students, academics, and industrial practitioners from around the world, as well as 12 active developers of these three systems. In the context of the study, the authors asked participants to assess the coupling between pairs of classes exhibiting high (low) coupling as indicated by the four different types of coupling and provide their ratings and some rationale.

The results indicate that the peculiarity of the semantic coupling measure allows it to better estimate the mental model of developers than the other coupling measures. In other words, when the semantic measure indicates high (low) coupling between two classes, also the developers feel the same. This is because, in several cases, the interactions between classes are encapsulated in the source code vocabulary, and cannot be easily derived by only looking at structural relationships, such as method calls.

4. Identifying refactoring opportunities

During software evolution the internal structure of the system undergoes continuous modifications. These continuous changes push away the source code from its original design, often reducing its quality, including class cohesion [Fowler, 1999]. In such scenarios a refactoring of the system is recommended. Refactoring has been defined as “*the process of changing a software system in such a way that it does not alter the external behavior of the code yet improves its internal structure*” [Fowler, 1999].

Typical advantages of refactoring include improved readability and reduced complexity of source code, a more expressive internal architecture and better software extensibility [Fowler, 1999]. For these reasons, refactoring is advocated as a good programming practice to be continuously performed during software development and maintenance. However, despite its advantages, performing refactoring in non-trivial software systems might be very challenging. First, the identification of refactoring opportunities in large systems is very difficult, due to the fact that the design flaws are not always easy to identify. Second, when a design problem has been identified, it is not always easy to apply the correct refactoring operation to solve it. All these observations highlight the need for (semi)automatic approaches supporting the software engineer in (i) identifying refactoring opportunities (i.e., design flaws) and (ii) designing and applying a refactoring solution.

Existing approaches for suggesting refactoring opportunities generally use metrics, e.g. cohesion metrics, which capture structural relationships between the members of a class, e.g., method-to-attribute references. However, semantic metrics have been proved to complement structural cohesion metrics. Based on this observation, Bavota *et al.* [2013c] proposed a method for automating the Extract Class refactoring. The proposed approach analyzes (structural and semantic) relationships between the methods in a class to

identify chains of strongly related methods. The identified method chains are used to define new classes with higher cohesion than the original class, while preserving the overall coupling between the new classes and the classes interacting with the original class. The proposed approach has been empirically evaluated on open source systems in order to assess how good and useful the proposed refactoring solutions are considered by software engineers and how well the proposed refactorings approximate the refactorings done by the original developers. Results indicate that (i) the semantic measure plays a crucial role in identifying meaningful refactorings; and (ii) the proposed solutions are useful in guiding refactorings. Following the same idea, Bavota *et al.* [2013d] also proposed an approach for suggesting Extract Package refactoring aiming at improving cohesion in packages.

Semantic measures were also used to identify Move Class refactoring [Bavota et al, 2013c]. The proposed approach, called R3 (Rational Refactoring via RTM), analyzes underlying latent topics in source code as well as structural dependencies to recommend (and explain) refactoring operations aiming at moving a class to a more suitable package. R3 has been evaluated in two empirical studies. The results of the first study conducted on nine software systems indicate that R3 provides a coupling reduction from 10% to 30% among the software modules. The second study with 62 developers indicates that more than 70% of the recommendations (and explanations) were considered meaningful from a functional point of view.

5. Conclusion and directions for future work

Textual analysis has been widely and successfully applied in software engineering for supporting several tasks. However, they also have some weaknesses, and poses challenges for researchers:

–**C₁**: *Quality of source code lexicon*. TA depends on the quality of the lexicon: a bad lexicon often means inaccurate – if not completely wrong – results. There are two common problems in the TA of software artifacts. The first is represented by the presence of inconsistent terms in related documents (e.g., requirements express some concepts using certain words, whereas the source code uses synonyms or abbreviations). The second problem is related to the presence of “noise” in software artifacts, for example due to recurring terms that do not bring information relevant for the analysis task, e.g. programming language keywords and terms that are part of a specific document template, such as a test case specification, a use case, or a bug report.

–**C₂**: *Setting of textual analysis technique*. TA techniques require configuring different components and their respective parameters, such as type of pre-processors (e.g., splitting identifiers, term weighting schema, stemmers). Despite this overwhelming popularity of IR methods in SE research, most of the proposed approaches are based on ad-hoc methods to configure these solutions, components, and their configurations, thus resulting oftentimes in suboptimal performance of such promising analysis methods. In other

words, existing methods for designing IR-based solutions for SE tasks is currently based on art, rather than science. This also makes the practical use of IR-based processes quite difficult and undermines the technology transfer to software industry.

Clearly, these challenges represent on one hand an obstacle for the technology transfer in industry, on the other hand a fertile ground for the definition of new solutions aimed at advancing the state of the art. We briefly outline some opportunities for future work aimed at mitigating the weaknesses of TA for software engineering:

–**O₁**: *Inducing developers to improve the quality of source code lexicon.* In 1990 Chikofsky and Cross [1990] defined the continuous reverse engineering, i.e., “reverse engineering, used with evolving software development technologies, will provide significant incremental enhancements to our productivity”. Inspired by continuous reverse engineering, we conjecture that *continuous textual analysis* could be one of the solutions for inducing developers to improve the quality of source code lexicon. Preliminary study indicated that developers are induced to improve the source code lexicon if the software development environment provides information about the textual similarity between the source code under development and the related high-level artifacts [De Lucia et al, 2011].

–**O₂**: *Automatic setting on textual analysis process.* Approaches for solving the problem of assembling IR-based solutions for a given SE task and accompanying dataset are required. Our underlying assumption, which is supported by a large body of empirical research in the field, is that it is not possible to build a set of guidelines for assembling IR-based solutions for a given set of tasks as some of these solutions are likely to underperform on previously unseen datasets. Thus, automatic approaches need to be defined to find an optimal setting of the TA technique given a specific dataset.

–**O₃**: *Empirical studies.* Existing approaches based on textual retrieval have been evaluated off-line, i.e., they have been applied on ended software projects and the benefits have been evaluated considering some metrics. However, such approaches are intended to be used by software engineers. Thus, they need to be validated on-line, i.e., when used by real users during the development of software systems. Only in this way it is possible to measure the actual benefits of TA in supporting software engineering tasks.

References

[Antoniol et al, 2002] G. Antoniol, G. Canfora, G. Casazza, A. De Lucia, and E. Merlo, Recovering traceability links between code and documentation, IEEE TSE, 28, 10, 970–983, 2002.

[Baeza-Yates and Ribeiro-Neto, 1999] Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. Addison-Wesley, 1999.

[Bavota et al, 2013a] G. Bavota, B. Dit, R. Oliveto, M. Di Penta, D. Shybyanyk, A. De Lucia, An Empirical Study on the Developers Perception of Software Coupling, Proc. of ICSE, San Francisco, USA, 2013, 692-701.

[Bavota et al, 2013b] G. Bavota, M. Gethers, R. Oliveto, D. Shybyanyk, A. De Lucia. Improving software modularization via automated analysis of latent topics and dependencies. TOSEM, 2013.

[Bavota et al, 2013c] G. Bavota, A. De Lucia, A. Marcus, R. Oliveto. Automating extract class refactoring: an improved method and its evaluation. EMSE, 2013.

[Bavota et al, 2013d] G. Bavota, A. De Lucia, A. Marcus, R. Oliveto. Using structural and semantic measures to improve software modularization. EMSE, 2013.

[Briand et al, 1998] L. C. Briand, J. W. Daly, and J. Wüst, A unified framework for cohesion measurement in object-oriented systems, EMSE, 3, 1, 65-117, 1998.

[Canfora and Cerulo, 2005] G. Canfora and L. Cerulo, Impact analysis by mining software and change request repositories, in Proc. of METRICS, Como, Italy, IEEE CS Press, 2005, 20–29.

[Chidamber and Kemerer, 1998] S. R. Chidamber, C. F. Kemerer. A metrics suite for object-oriented design. IEEE TSE, 20,6, 476-493, 1994.

[Chikofsky and Cross, 1990] E. J. Chikofsky, James H. Cross II: Reverse Engineering and Design Recovery: A Taxonomy. IEEE Software 7(1): 13-17 (1990).

[Deerwester et al, 1990] S. Deerwester, S. T. Dumais, G. W. Furnas, T.K. Landauer, R. Harshman. Indexing by latent semantic analysis. JASIST, 41, 6, 391–407, 1990.

[De Lucia et al, 2011] A. De Lucia, M. Di Penta, R. Oliveto: Improving Source Code Lexicon via Traceability and Information Retrieval. IEEE TSE, 37,2, 205-227, 2011.

[Fowler, 1999] M. Fowler, Refactoring: improving the design of existing code. Addison-Wesley, 1999.

[Guerrouj et al., 2012] L. Guerrouj, P. Galinier, Y.-G. Guéhéneuc, G. Antoniol, M. Di Penta: TRIS: A fast and accurate identifiers splitting and expansion algorithm. Proc. of WCRE, Kingston, Canada, IEEE Press, 2012, 103-112.

[Marcus and Maletic, 2001] A. Marcus and J. I. Maletic, Identification of high-level concept clones in source code, in Proc. of ASE, San Diego, California, USA, IEEE CS Press, 2001, 107–114.

[Marcus et al, 2008] A. Marcus, D. Shybyanyk, and R. Ferenc, Using the conceptual cohesion of classes for fault prediction in object-oriented systems, IEEE TSE, 34, 2, 287–300, 2008.

[Porter, 1980] Porter, M.F.: An algorithm for suffix stripping. Program, 14, 3, 130–137, 1980.

[Shybyanyk et al, 2007] D. Shybyanyk, Y. Gael-Gueheneuc, A. Marcus, G. Antoniol, and V. Rajlich, Feature location using probabilistic ranking of methods based on execution scenarios and information retrieval, IEEE TSE, 33, 6, 420–432, 2007.

[Shybyanyk et al, 2009] D. Shybyanyk, A. Marcus, R. Ferenc, T. Gyimóthy. Using information retrieval based coupling measures for impact analysis, EMSE, 14, 1, 5–32, 2009.

Building Trust through Social Awareness: The SocialCDE Project

Fabio Calefato, Filippo Lanubile
Università degli Studi di Bari Aldo Moro
Dipartimento di Informatica, Via Orabona 4, 70125 Bari (BA)
{fabio.calefato, filippo.lanubile}@uniba.it

Abstract. *Trust is paramount in distributed software development to prevent geographically distributed sites to feel and act like distinct, distant teams. Nevertheless, how to build trust among developers with few or no chances to meet is an open issue. To overcome such a challenge we hypothesize that increased social awareness may foster trust building in global software teams. Here, we first present SocialCDE, a tool that aims at augmenting Application Lifecycle Management (ALM) platforms with social awareness to facilitate the establishment of interpersonal connections by disclosing developers' personal interests and contextual information. Then, we present two different empirical studies, specifically designed to test our hypothesis.*

Keywords: Trust Building, Distributed Software Development, Social Awareness, Empirical Studies, Human Factors, Social Computing.

1. Introduzione

Benché importante per qualsiasi tipo di team, il trust (fiducia) è un fattore che contribuisce notevolmente al successo o al fallimento di progetti su scala globale [Herbsleb et al., 2005]. Infatti, in tali contesti distribuiti, il trust è fondamentale per evitare che la distanza fisica porti anche a quella psicologica. E' stato osservato che un basso livello di fiducia (a) aggrava la percezione di essere team diversi e con obiettivi conflittuali, (b) diminuisce la disponibilità a condividere informazioni e cooperare per risolvere i problemi, e infine (c) influenza negativamente la benevolenza verso gli altri in caso di obiezioni e disaccordi [Al-Ani et al., 2011]. Il trust tra i componenti di un team si sviluppa di solito attraverso una stretta interazione e comunicazione faccia-a-faccia (F2F) poiché essa rappresenta il modo più efficace per stabilire relazioni con gli altri e acquisire consapevolezza (awareness). Purtroppo, l'interazione F2F è proprio l'attività maggiormente penalizzata della distanza nel global software development. Altresì, lo sviluppo del software è condotto sempre più in modo distribuito con il passare degli anni, nonostante il fatto che nessuna risposta sia stata fornita alla seguente domanda: *Come si può stabilire o rafforzare il trust*

tra i componenti di team distribuiti su scala globale quando essi hanno scarsa o nessuna possibilità di incontrarsi?

Nel tentativo di fornire una risposta, abbiamo ipotizzato che la condivisione degli interessi personali dei componenti di un team all'interno di uno spazio di lavoro comune possa facilitare la creazione di connessioni interpersonali, aumentare la probabilità di interazioni di successo, e contribuire a stabilire rapporti di fiducia tra i componenti di un global software team. In altre parole, ci aspettiamo che la possibilità di avere accesso alle informazioni condivise sui social media, così come la possibilità di monitorare l'attività dei propri collaboratori sui social network possano funzionare come un surrogato della interazione sociale che avviene naturalmente durante incontri informali faccia a faccia, aumentando così la fiducia reciproca attraverso l'incremento di consapevolezza (awareness). Il concetto di awareness (consapevolezza) proviene dal campo del CSCW e ha avuto una notevole influenza nella campo della ricerca dell'ingegneria del software in quanto fornisce meccanismi per coordinare le attività di gruppo [Mistrík et al., 2010].

Pertanto, abbiamo sviluppato SocialCDE, un tool che estende le piattaforme di Application Lifecycle Management (ALM), aggiungendo il supporto alla social awareness. Nella parte restante di questo articolo, prima descriviamo lo strumento; dopodiché, illustriamo due studi empirici preliminari, al fine di rispondere alla precedente domanda di ricerca.

2. Stato dell'Arte

2.1 Trust

Il trust è un fenomeno complesso da studiare in quanto coinvolge sia i rapporti interpersonali, ossia i problemi culturali tra fiduciario (trustor) e fiduciante (trustee), sia le sfaccettature del comportamento umano, ossia, i tratti personali del carattere. A oggi sono state date diverse definizioni di trust. Una definizione ampiamente utilizzata e concisa è fornita da Jarvenpaa et al. [Jarvenpaa et al., 1998], i quali definiscono il trust come le aspettative del trustor che il trustee si comporterà come previsto. In altre parole, la fiducia positiva emerge quando le azioni degli altri incontrano la nostra aspettativa; in caso contrario, si genera sfiducia. Altre definizioni di fiducia distinguono tra dimensione *cognitiva* (*cognitive trust*) e *affettiva* (*affective trust*). Per esempio, Wilson et al. [Wilson et al., 2006] definiscono la fiducia cognitiva in termini di aspettative circa la competenza e l'affidabilità degli altri nello svolgimento di azioni importanti che il trustor non può monitorare. Al contrario, la fiducia affettiva si riferisce ai legami emotivi reciproci, le preoccupazioni e l'affetto tra il trustor e il trustee, che spinge quest'ultimo a fare qualcosa per il primo per dovere morale.

Diversi sono stati gli approcci proposti per favorire la crescita della fiducia. L'idea di fondo è che il processo di costruzione della fiducia si sviluppa lungo diverse dimensioni chiamate antecedenti (*trust antecedents*) [Jarvenpaa et al., 1998], cioè le proprietà del trustee che scatenano la valutazione da parte

del trustor nel valutare l'affidabilità degli altri. Secondo Jarvenpaa et al. [Jarvenpaa et al., 1998], l'*abilità* (e.g., competenze, conoscenze), la *benevolenza* (e.g., la cortesia, la disponibilità), l'*integrità* (e.g., la fedeltà, l'adesione alle norme morali), e la *prevedibilità* (e.g., l'affidabilità, i comportamenti coerenti) sono le caratteristiche personali di un trustee che facilitano la costituzione del rapporto di fiducia con un trustor. In particolare, le dimensioni di *capacità* e la *prevedibilità* sono valutate per mezzo di un'elaborazione cognitiva delle informazioni personali e professionali. Allo stesso tempo, la valutazione affettiva porta alla costruzione della fiducia lungo le dimensioni di benevolenza e integrità. Infine, per un trustor, è invece la *propensione a fidarsi* il tratto personale del carattere che conta.

2.2 Awareness

La consapevolezza di gruppo, o *group awareness*, è stata definita da Dourish e Bellotti, come "*la comprensione delle attività degli altri, per fornire un contesto alla propria attività*" [Dourish e Bellotti, 1992].

Secondo Gutwin et al. [Gutwin et al., 1996], oltre che riguardo ai propri collaboratori, tipicamente i componenti di un gruppo cercano informazioni anche sulle attività e gli artefatti. Di conseguenza, sono stati identificati quattro tipi di group awareness finora, vale a dire: la *informal o presence awareness* (cioè, chi è online e la loro disponibilità), la *group-structural awareness* (vale a dire, i ruoli di ciascun componente e la struttura interna dei team), la *workspace awareness* (cioè, chi ha modificato un artefatto condiviso e quando), e la *social awareness* (cioè, l'informazione e la comprensione che compagni hanno circa loro connessioni sociali all'interno di un gruppo [Omoronya et al., 2010]).

3. Modello di Ricerca e Ipotesi

Jarvenpaa e Leidner hanno osservato l'evoluzione del trust nei team distribuiti su scala globale in cui l'interazione avviene esclusivamente attraverso la comunicazione mediata da computer [Jarvenpaa e Leidner, 1999]. La loro analisi ha indicato che i team con un basso livello di trust iniziale erano quelli carenti a livello di comunicazione sociale all'inizio del progetto. Al contrario, i team che manifestavano un elevato livello di fiducia al termine del progetto, si focalizzavano inizialmente sulla comunicazione "sociale", che in seguito diminuiva per dare spazio all'interazione focalizzata sulle attività da compiere e le procedure da seguire. Coerentemente con questa osservazione, noi ipotizziamo che la divulgazione delle informazioni personali, contestualizzate nello spazio di lavoro condiviso, possa aumentare la percezione di similarità e vicinanza tra i componenti di team distribuiti, favorendo in tal modo la quantità di comunicazione sociale scambiata, la manifestazione di un numero più elevato di interazioni di successo, e nuovi legami più forti stabiliti tra individui distanti. Infine, la ricerca condotta da Jarvenpaa et al. [Jarvenpaa et al., 1998] ha dimostrato come la percezione dell'integrità altrui, la benevolenza, e la propensione a fidarsi percepita siano gli antecedenti rilevanti per favorire la costruzione di fiducia affettiva. Pertanto, ipotizziamo che (vedi Figura 1):

H1 - Esiste una correlazione positiva tra il livello di social awareness acquisita attraverso i social media e il livello di fiducia affettiva stabilito reciprocamente fra team distribuiti.

Per verificare questa ipotesi attraverso un esperimento, abbiamo bisogno di misurare i livelli di fiducia affettiva e cognitiva. Dal momento che la fiducia è una *sensazione percepita* dagli individui, è necessario fare affidamento su dati riportati dai partecipanti all'esperimento (ossia, utilizzare questionari e/o interviste).

Treinen e Miller-Frost [Treinen e Miller-Frost, 2006] hanno osservato che lo sviluppo di fiducia reciproca tra siti distanti all'inizio di un progetto è fondamentale per la sua riuscita. Infatti, essi hanno osservato che, durante la fase iniziale di un progetto, la conoscenza personale tra i componenti di un team e la fiducia reciproca sono più importanti della risoluzione dei primi problemi tecnici, dal momento che la fiducia permette di risolvere tali problemi in seguito e a distanza (e.g., attraverso conference call), ottenendo così una maggiore efficienza complessiva per l'intero progetto. Pertanto, ipotizziamo che (Figura 1):

H2 - Esiste una correlazione positiva tra il livello di fiducia affettiva reciprocamente stabilita tra team distribuiti e la performance generale di progetto.

Per quanto riguarda il test di questa ipotesi, riconosciamo che stabilire un rapporto di causa / effetto tra la fiducia e le prestazioni di un progetto sia un compito arduo, giacché molti altri confounding factor – per esempio, il tipo di progetto, le competenze individuali – possono interferire lungo il processo sperimentale.

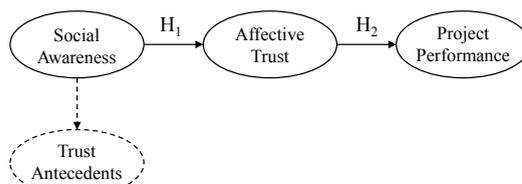


Fig.1 – Il modello di ricerca proposto.

4. Il Progetto SocialCDE

Per Application Lifecycle Management (ALM) si intende un processo continuo di gestione del ciclo di vita di un progetto attraverso piattaforme che forniscono uno spazio di lavoro condiviso e integrato con un set di strumenti di supporto a tutte le attività di sviluppo software, come per esempio la gestione dei requisiti, la progettazione, la codifica, il testing e gestione del rilascio [Calefato e Lanubile, 2012]. Le piattaforme di ALM più popolari, note anche come ambienti di sviluppo collaborativo (CDE), non supportano affatto la social awareness, oppure, quando disponibile, come nel caso di Jazz e GitHub, il livello di supporto offerto non è sufficiente [Lanubile et al., 2013].

Per far fronte a questa limitazione, abbiamo sviluppato SocialCDE [Calefato e Lanubile, 2013], uno strumento che estende le piattaforme ALM

Microsoft Team Foundation Server (TFS) e GitHub rendendo disponibili le informazioni raccolte dalla maggior parte dei più diffusi social network oggi disponibili (ad esempio, Twitter, LinkedIn). Dal lato client, due sono i plugin disponibili, uno estende Visual Studio, l'altro supporta qualsiasi IDE basato su Eclipse. La Figura 2 mostra i componenti principali del progetto SocialCDE: i plugin gestiscono la visualizzazione di tutti i contenuti sociali raccolte presso i servizi abilitati dall'utente (Figura 3).

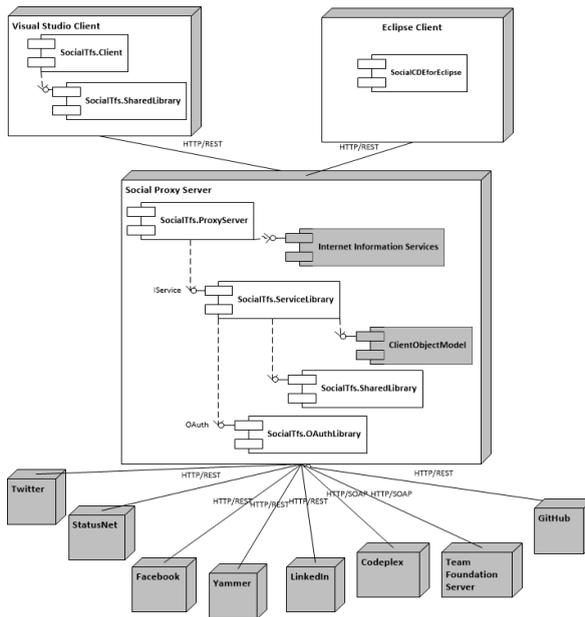


Fig. 2 – Diagramma di deployment delle componenti del progetto SocialCDE.

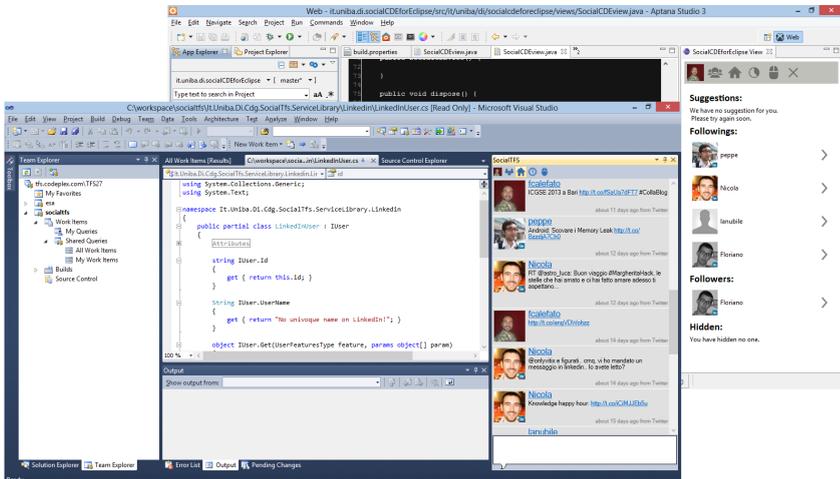


Fig.3 – I plugin di SocialCDE per Visual Studio (sx) e Aptana (dx).

L'altra componente principale, il Social Proxy Server, è un aggregatore che sfrutta le API di microblog aziendali e siti di social networking. Essendo un proxy, tale componente interagisce attraverso protocolli RESTful sia con i plugin lato client sia con le piattaforme ALM. I suoi compiti principali sono il recupero di informazioni riguardo agli account utenti sui servizi di social network (SNS) e ai progetti software dalle piattaforme ALM. Il proxy server ottiene i permessi di accesso attraverso il protocollo OAuth ed effettua il caching dei post, consentendo all'utente di personalizzare le informazioni che è possibile recuperare dai propri account. Inoltre, il proxy può gestire e memorizzare i dati di connessione a GitHub, TFS e CodePlex.

Come accennato prima, il contenuto recuperato dai SNS è caricato e memorizzato nella cache del Social Proxy Server. Quindi, tutte le informazioni richieste da un client sono presentate all'utente finale attraverso una vista all'interno dell'IDE. In particolare, tali informazioni sono mostrate attraverso tre prospettive diverse (dette *timeline*), ossia la *home*, la *iteration*, e la *interactive*. La home timeline replica le timeline disponibili in siti di microblogging come Twitter o Yammer, in quanto essa è popolata dai post dell'utente e di coloro che lo "seguono" (follower) in ordine cronologico. Noi definiamo *statico* questo tipo di follower, perché per ricevere i messaggi di un utente è richiesta esplicitamente l'azione di "follow", cioè di aggiunta di un contatto alla propria *rete di consapevolezza* (o *awareness network*) – l'insieme di colleghi le cui azioni sono monitorate da un utente e ai quali sono mostrate le sue stesse azioni. Tuttavia, le *awareness network* sono fluide [de Souza e Redmiles, 2011]: le persone le cui azioni si dovrebbero monitorare e a cui le proprie azioni dovrebbero essere mostrate cambiano nel tempo, a seconda (a) delle attività correnti o (b) della fase di sviluppo in cui si trova il progetto. Pertanto, oltre la visualizzazione del flusso dei follower statici mostrati nella home timeline, abbiamo anche progettato un tipo dinamico di follower e altre due timeline. A differenza di quelli statici, i follower dinamici non richiedono alcuna azione di follow esplicito e, pertanto, sono aggiunti e rimossi automaticamente dalla *awareness network* quando: i) si verificano dei cambiamenti riguardo ai task di un utente nella iterazione corrente; ii) un artefatto nel repository del progetto è aperto nell'editor dell'IDE. Nel primo caso, la *iteration timeline* aggiunge dinamicamente alla *awareness network* di un utente ogni componente del team che ha riportato o anche commentato un workitem assegnatogli/le. Nel secondo caso, la *interactive timeline* visualizza dinamicamente i messaggi di chi ha contribuito modifiche alla risorsa aperta e visualizzata al momento nell'editor di testo (e.g., un file sorgente).

5. Gli Studi

In questa sezione presentiamo il design di due studi empirici progettati per condurre un'indagine preliminare della nostra ipotesi di ricerca. Il primo (Studio A) è un field study attualmente in corso che coinvolge un solo team di

sviluppo. Il secondo (studio B) è un esperimento controllato, che si prevede abbia inizio entro la fine dell'estate.

5.1 Studio A

Questo studio coinvolge un piccolo team di sviluppo di 7 componenti che lavorano ad "I Speak Again", un progetto promosso dalla onlus Informatici Senza Frontiere (ISF). Abbiamo selezionato questo progetto perché verifica i seguenti requisiti socio-tecnici. In primo luogo, il team è distribuito e non costituito da molto tempo. In particolare, i tre responsabili del progetto sono gli unici che si conoscono da uno o più anni e hanno la possibilità di incontrarsi di persona. Invece, i restanti quattro membri sono completamente distribuiti e hanno iniziato a lavorare per il progetto proprio quando lo studio ha avuto inizio. In secondo luogo, il progetto utilizza GitHub come piattaforma ALM e Aptana, un IDE basato su Eclipse per lo sviluppo di applicazioni per il Web. Infine, i membri del team hanno accettato di connettere a SocialCDE uno o più dei loro account di social network.

Poiché alcuni dei soggetti sperimentali coinvolti hanno già lavorato insieme per qualche tempo prima dell'inizio dell'esperimento, abbiamo adottato il seguente design sperimentale di tipo *A-B-A*. Come mostrato in Figura 4, la fase iniziale *A* è quella in cui si trovava il team prima di iniziare l'esperimento, ossia quando i componenti sono stati addestrati all'uso del plugin. Inoltre, uno dei team leader ha partecipato a una sessione di demo per dare supporto tecnico agli altri sviluppatori. Prima di iniziare l'esperimento, tutti i soggetti hanno risposto a un questionario preliminare, che si proponeva di raccogliere alcune informazioni personali (e.g., età, sesso, esperienza di lavoro, istruzione), così come misurare la loro propensione a fidarsi degli altri e la quantità di trust già stabilita con gli altri componenti del team. Tali misure sono calcolate indirettamente, attraverso la valutazione degli antecedenti della fiducia.

Al momento, l'esperimento è in fase *B*, ossia, il team ha iniziato le pratiche di sviluppo agile e le attività collaborative, utilizzando l'IDE Aptana esteso con il plugin di SocialCDE. Questa fase durerà da quattro a sei settimane (anche se imporre scadenze rigide non è fattibile a causa della natura volontaria del progetto). Durante questa fase, i membri del team condurranno regolarmente le loro attività lavorative, usando in più il plugin per leggere e condividere contenuti. I dati di utilizzo saranno automaticamente raccolti dal componente proxy (e.g., il numero di posti condivisi, variazioni nell'awareness network). Alla fine della fase *B*, un secondo questionario sarà somministrato ai partecipanti, con l'obiettivo di individuare differenze nei livelli di fiducia percepita rispetto agli altri. Infine, durante la terza e ultima fase *A*, la squadra tornerà a lavorare con l'IDE senza il plugin, come nella fase iniziale. Dopo una settimana di lavoro senza il supporto del nostro tool, uno dei ricercatori condurrà delle interviste semi-strutturate individuali (tramite Skype o faccia a faccia), che saranno registrate e trascritte. Inoltre, sarà anche organizzata una riunione finale con i tre team leader.

Data la natura di field study, saremo in grado di ottenere da questo studio solo alcuni dati per testare l'ipotesi H_1 sulla relazione tra social awareness e fiducia affettiva.

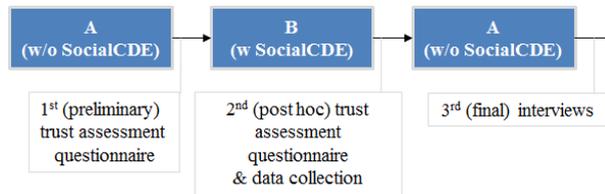


Fig. 3 – Design sperimentale del field study “Studio A”.

5.2 Studio B

E' stato realizzato anche il design di un esperimento controllato che, rispetto allo studio precedente, confronta due progetti differenti, realizzati da due diversi team distribuiti di 5 o 6 membri. Entrambi i team applicano la metodologia di sviluppo agile SCRUM e lavorano a un vero prodotto software per un cliente reale. La composizione dei team è la seguente: 3 sviluppatori (stagisti per una grande multinazionale presso PUCRS, Università di Porto Alegre, Brasile), 1 team leader (docente), e 1 o 2 product owner (clienti). Come nel caso dello Studio A, anche questo esperimento richiede che i team siano distribuiti e formati di recente. Tuttavia, in questo caso sarà utilizzato TFS come piattaforma ALM e Visual Studio come ambiente di sviluppo.

Poiché sono disponibili due diversi progetti, abbiamo deciso di utilizzare il progetto A come gruppo di controllo e il progetto B come gruppo di trattamento (Figura 5). In particolare, durante la prima fase, ossia, prima di iniziare la fase di sviluppo, il team del progetto B assisterà a una lezione introduttiva che illustrerà come utilizzare il plugin. La lezione di formazione si concluderà con la somministrazione dello stesso questionario pre-esperimento usato nell'altro studio. Dopodiché, durante la seconda fase, il team di progetto A utilizzerà Visual Studio così com'è, mentre il team di progetto B utilizzerà l'IDE esteso con il plugin di SocialCDE per tutta la durata dell'attività di sviluppo (circa 6-8 settimane). Durante questa fase di sviluppo, i membri del team di entrambi i progetti dovranno svolgere le loro attività di lavoro regolarmente. Le uniche attività extra svolte dal team di progetto B attraverso l'uso del plugin riguarderanno la lettura e la condivisione di contenuto sociale con gli altri membri del team. Un secondo questionario post-hoc sarà somministrato al termine di questa fase, ma solo per il gruppo di trattamento. Infine, durante la terza ed ultima fase, al termine della attività di sviluppo per entrambi i progetti, saranno condotte e trascritte interviste semi-strutturate al fine di eseguire analisi qualitative.

Per questo studio controllato saremo in grado di valutare gli effetti di SocialCDE durante l'intera fase di sviluppo del progetto di test, per poi confrontarli con il progetto di controllo. Pertanto, otterremo dati utili per testare sia H_1 sia H_2 . Tuttavia, rispetto alla seconda ipotesi, considerato il numero limitato di soggetti sperimentali coinvolti, sarà necessario condurre ulteriori repliche dell'esperimento per confermare eventuali differenze riscontrate in termini di performance tra i due progetti.

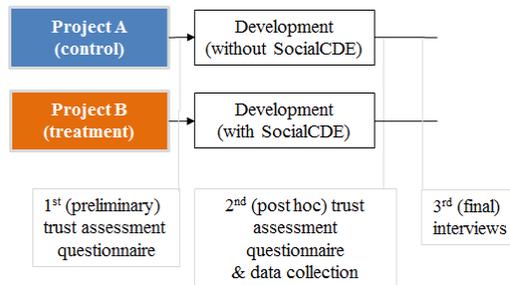


Fig. 5 – Design sperimentale dell’esperimento controllato “Studio B”.

6. Conclusioni e Sviluppi Futuri

In questo articolo abbiamo descritto SocialCDE, un tool che ha l’obiettivo di estendere le piattaforme ALM attraverso l’aggregazione di contenuto da diversi social media e la condivisione all’interno di uno spazio di lavoro condiviso. SocialCDE è stato progettato per supportare la nostra ipotesi che le informazioni condivise sulle reti sociali possano costituire un surrogato della social awareness ottenuta durante le interazione informali, utile per generare fiducia tra i componenti di software team distribuiti. Infine, abbiamo anche presentato il design sperimentale di due studi preliminari, volti a valutare sperimentalmente la nostra ipotesi.

Ringraziamenti

Si ringraziano Nicola Sanitate, Giuseppe Santoro, e Floriano Fauzzi per il loro contributo allo sviluppo del progetto. SocialCDE è parzialmente finanziato dal progetto Intersocial, nell’ambito del Programma di Cooperazione Transfrontaliera Grecia-Italia 2007-2013, e da Microsoft Research attraverso il premio Software Engineering Innovation Foundation (SEIF) Award 2011.

Bibliografia

Al-Ani, B., Wilensky, H., Redmiles, D., and Simmons, E. An Understanding of the Role of Trust in Knowledge Seeking and Acceptance Practices in Distributed Development Teams, Proc. 6th Int’l Conf. Global Soft. Eng. (ICGSE ’11), Helsinki, Finland, Aug. 15-18, 2011, pp. 25-34.

Calefato F. and Lanubile F. Practice: Collaborative Development Environments, in Global Software and IT: A Guide to Distributed Development, Outsourcing, and Supplier Management (C. Ebert ed.), IEEE, Wiley, 2012.

Calefato F. and F. Lanubile F. SocialCDE: A Social Awareness Tool for Global Software Teams. ESEC/FSE 2013, St. Petersburg, Russia.,18-26 Aug. 2013.

de Souza C.R.B. and Redmiles, D.F. The Awareness Network, To Whom Should I Display My Actions? And, Whose Actions Should I Monitor? IEEE Trans. on Sw Eng, 37, 3, pp. 325-340, 2011.

Dourish P. and Bellotti, V. Awareness and Coordination in Shared Workspaces. Proc. Int'l Conf. on Computer Supported Cooperative Work (CSCW '92), Toronto, Canada, Nov 1-4, 1992, pp. 107-114.

Gutwin, C., Greenberg, S., and Roseman, M. Workspace Awareness in Real-Time Distributed Groupware: Framework, Widgets, and Evaluation. Proc. HCI (HCI'96), London, Aug. 20-23, 1996, pp. 281-298.

Herbsleb, J.D. et al. Global Software Development at Siemens: Experience from Nine Projects. Int'l Conf. Sw Eng (ICSE'05), St Louis, USA, May 15-21, 2005, pp. 524-533.

Jarvenpaa, S.L., Knoll, K., and Leidner, D.E. Is anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems*, 14(4), pp. 29-64, 1998.

Jarvenpa S.L. and Leidner D.E. Communication and Trust in Global Virtual Teams. *Journal of Organization Science*, 10(6), pp. 791-815, Nov—Dec. 1999.

Lanubile, F., Calefato, F. and C. Ebert, C. Group Awareness in Global Software Eng., *IEEE Softw.*, Mar. 2013, pp.18-23.

Mistrić, I., Grundy, J., van der Hoek, A., and Whitehead, J. Collaborative Software Engineering: Challenges and Prospects. *Collaborative Software Eng.*, 2010, pp. 389–403.

Omoronyia, I., Ferguson, J., Roper, M., and Wood M. A review of awareness in distributed collaborative software engineering. *Software: Practice & Experience*, 40, pp. 1107–1133, 2010.

Treinen J.J. and Miller-Frost S.L. Following the sun: case studies in global software development, *IBM Syst. J.*, 45(4), Oct. 2006, pp. 773-783.

Wilson, J.M., Strausb, S.G, and McEvily, B. All in due time: The development of trust in computer-mediated and face-to-face teams, *Organizational Behavior and Human Decision Processes*, 99(1), Jan. 2006, pp. 16–33.

Analisi comparativa di linguaggi formali per la definizione di Design Pattern

Beniamino Di Martino¹, Antonio Esposito²
Department of Industrial and Information Engineering,
Second University of Naples
Via Roma 29, Aversa
[1beniamino.dimartino@unina.it](mailto:beniamino.dimartino@unina.it)
[2antonio.esposito7@studenti.unina2.it](mailto:antonio.esposito7@studenti.unina2.it)

This paper describes, analyses and compares different formal languages which have been used in literature to define Design Patterns. For every language we provide a brief description of its elements and discuss their ability to completely define a Design Pattern, both through a graphical and a machine readable formalism, also in relation to the other languages examined, in order to identify their weak and strong points.

Keywords: Design Patterns, Lepus, BPSL, ODOL, RBML, UML, OWL, First Order Logic

1. Introduzione

L'uso dei Design Pattern nell'Ingegneria del Software è diffuso e supportato da un numero sempre più consistente di tool. La capacità dei Pattern di sostenere il riuso del software, la riduzione di errori e bug che deriva dal loro utilizzo e il notevole risparmio, in termini di tempo, che si ottiene sfruttando le soluzioni da essi proposte, hanno portato allo sviluppo di una vasta letteratura che, ancora oggi, si arricchisce di continuo con nuove definizioni e soluzioni.

Tuttavia, i cataloghi esistenti non classificano i pattern in maniera formale, e gli stessi pattern non sono pubblicati seguendo una struttura o un formalismo condivisi: la tendenza generale è quella di sfruttare un mix di linguaggio naturale e diagrammi UML, seguendo uno schema deducibile dalle descrizioni presenti in [Gamma et al, 1995]. Diversi tentativi di formalizzazione sono stati proposti [Taibi, 2007], ma nessun linguaggio è riuscito ad imporsi definitivamente sugli altri come standard. In questo paper descriviamo alcuni di questi linguaggi formali, fornendo una comparazione tra i formalismi esaminati. Il documento è organizzato come segue: nel paragrafo 2 sono introdotti i Design Pattern; nel paragrafo 3 sono descritti i linguaggi formali esaminati; infine indichiamo le conclusioni tratte dalla nostra analisi.

2. I Design Pattern

Un design pattern viene definito come una soluzione generale e riutilizzabile a un problema comune all'interno di un dato contesto nella progettazione software. L'obiettivo dei design pattern è quello di fornire indicazioni sulla soluzione di un problema già affrontato precedentemente in maniera efficiente ed efficace, evitando nuove implementazioni inclini a errori e bug, sia di progettazione che di programmazione. Un design pattern non è un progetto finito che può essere trasformato direttamente in codice sorgente o macchina: essi possono essere visti come un insieme di best-practice che il programmatore deve implementare in un'applicazione. Le definizioni dei principali Design Pattern sono fornite in [Gamma et al, 1995], l'opera che ha dato impulso alla ricerca in tale ambito. I pattern si concentrano essenzialmente su due aspetti fondamentali dei problemi affrontati e delle soluzioni proposte: l'aspetto **strutturale**, consistente nella descrizione dei componenti attivi (detti partecipanti) nella definizione delle funzionalità offerte dal pattern e delle relazioni tra essi esistenti; l'aspetto **dinamico** del pattern, cioè le modalità di interazione tra i partecipanti. L'analisi dei linguaggi formali si basa proprio sulla relativa capacità di descrivere questi aspetti dei diversi pattern.

3. Linguaggi formali: analisi comparativa

La maggior parte dei linguaggi formali proposti in letteratura si sofferma unicamente sull'aspetto strutturale dei pattern: è il caso del **Language for Pattern Uniform Specification (LePUS)** [Eden et al, 1998]. LePUS è un linguaggio descrittivo visuale per la progettazione orientata agli oggetti, in grado di descrivere la struttura dei pattern attraverso un formalismo grafico rappresentato da diagrammi chiamati **Codechart**, affiancati da rappresentazioni codificate dette schema **Class-Z**. I Codechart permettono di identificare i differenti partecipanti del pattern descritto, descrivendone la tipologia e i prototipi dei metodi posseduti: in questo modo la struttura del pattern è completamente definita. LePUS può facilmente supportare progettazione e sviluppo degli aspetti architeturali del software, permettendone la descrizione completa dei componenti e delle reciproche relazioni attraverso pattern. Inoltre, eventuali modifiche strutturali al software sono facilmente rappresentabili nei **Codechart**, anche attraverso il supporto di tool automatici che sfruttino la rappresentazione codificata messa a disposizione. Nonostante tali potenzialità, la mancanza di una descrizione comportamentale dei pattern rappresenta un punto debole importante di LePUS, in quanto risulta necessario ricorrere a rappresentazioni di supporto (come diagrammi UML) per non perdere informazioni importanti relativamente ai componenti descritti.

Il **Balanced Pattern Specification Language (BPSL)** [Taibi, 2003] descrive pattern sfruttando un linguaggio basato sulla logica del primo ordine per la definizione della struttura, e sulla temporal logic of actions (TLA) per le specifiche comportamentali. Sfruttando le relazioni primarie fornite dal linguaggio stesso la descrizione dei pattern risulta piuttosto semplice ed agevole. Tale descrizione è suddivisa in sezioni: le prime definiscono i diversi

elementi che partecipano al pattern e le relazioni esistenti tra di essi; le ultime descrivono gli stati interni dei vari partecipanti e riproducono le diverse possibili interazioni tra di essi. BPSL risulta essere il linguaggio basato su predicati logici più completo per la descrizione dei pattern, in quanto definizioni di struttura e comportamento fornite sono complete e semplici da realizzare. Se la possibilità di descrivere agevolmente anche il workflow dei pattern, senza dover ricorrere ad altri linguaggi di supporto, permette di superare alcune delle problematiche presenti in LePUS, la totale assenza di un formalismo grafico limita molto le potenzialità di BPSL: la documentazione relativa al design e allo sviluppo software dovrebbe essere redatta manualmente senza l'ausilio di tool grafici, rendendo altamente probabili errori umani; inoltre, tali descrizioni hanno uno scarso impatto visivo, risultando poco immediate e intuitive.

Di natura diametralmente opposta è il **Role Based Modeling Language** [France, Robert B. et al., 2004]. RBML estende UML per descrivere in maniera specifica struttura e comportamento di design pattern, definendo un insieme di ruoli che identificano sia classi che associazioni, i quali possono essere poi mappati verso diagrammi UML veri e propri: in questo modo è possibile sia identificare un pattern a partire da un diagramma UML già esistente, sia creare un nuovo diagramma seguendo le indicazioni fornite dal pattern. Poiché sfrutta appieno le caratteristiche di UML, RBML ne eredita l'espressività: gli **Static Pattern Specification** descrivono la struttura di un pattern in maniera simile ad un diagramma delle classi, mentre gli **Interaction** e gli **State Machine Pattern Specification** definiscono i suoi aspetti comportamentali, sia da un punto di vista globale sia specifico di un determinato componente. Tuttavia, nonostante la notevole espressività, RBML possiede alcune importanti lacune: innanzitutto, esso sfrutta una rappresentazione machine readable basata su XML che, a causa della scarsa ricezione dello standard e delle ulteriori estensioni aggiunte da RBML, rende i grafi poco inclini alla analisi e alla elaborazione automatiche; in secondo luogo, come accade per l'UML puro e con praticamente tutti i linguaggi dedicati ai pattern, tra cui anche LePUS e BPSL, RBML non riesce a descrivere aspetti del pattern quali i suoi intenti, motivazioni e contesto applicativo, che risultano importanti per meglio comprendere ruoli e funzionalità dei componenti descritti nei pattern.

La definizione formale dei Design pattern, dei loro partecipanti, delle proprietà e delle relazioni esistenti tra i partecipanti descritta da **Object-Oriented Design Ontology (ODOL)** [Dietrich et al, 2005] è basata su una ontologia OWL, cioè su un sistema di classi OWL, dotate di proprietà e relazioni ben definite, che sono istanziate in definizioni di Design pattern. L'ontologia alla base di ODOL contiene classi come **ClassTemplate** e **MethodTemplate**, e le relazioni associate, per descrivere i partecipanti al pattern e i rispettivi metodi, così come una classe **Pattern** che rappresenta i pattern stessi. Ciò comporta una estrema facilità nella descrizione della struttura del pattern, in quanto essa è completamente rispecchiata dall'intreccio di relazioni esistenti tra le classi istanziate. Essendo OWL uno standard condiviso, è possibile sfruttare per la sua rappresentazione grafica e per la manipolazione dell'ontologia di base uno dei qualsiasi tool esistenti sul mercato. Inoltre, ODOL prevede la possibilità di associare all'ontologia una serie di meta-dati che forniscono informazioni

relative a scopi, motivazioni e intenti del pattern, elemento assente negli altri linguaggi qui presentati. Il problema è rappresentato dalla definizione della parte comportamentale, non molto accurata, soprattutto a causa delle restrizioni imposte da OWL stesso.

Conclusioni

Nel presentare i diversi linguaggi presenti in letteratura abbiamo analizzato le tecniche con cui ognuno di essi esprime gli aspetti fondamentali dei pattern e abbiamo potuto osservare come alcuni di essi siano incapaci di descrivere l'aspetto comportamentale (LePUS, ODOL) o si rifacciano semplicemente ad UML (RBML) per la formulazione di comportamento e struttura dei pattern, rendendone difficile una rappresentazione codificata. L'unico linguaggio effettivamente completo è il BPSL: tuttavia, esso è penalizzato dall'inesistenza di un formalismo grafico, oltre alla sua incapacità di esprimere concetti relativi ai pattern quali scopi, intenti e conseguenze del loro utilizzo, utili ad una catalogazione e una successiva ricerca, anche con l'obiettivo di automatizzare la produzione di codice. L'uso di una ontologia per la descrizione di design pattern sembra essere, nonostante le difficoltà emerse analizzando ODOL, una buona metodologia per la loro rappresentazione, soprattutto in conseguenza della possibilità di estendere i concetti statici esprimibili in OWL con definizioni comportamentali dinamiche facilmente formulabili con Owl-S. Studi in merito sono riportati in [Yu et al,2010] e sono stati ripresi e sfruttati in alcuni dei nostri lavori relativi al riconoscimento di Design Pattern da documentazione UML.

Ringraziamenti

Questo lavoro è stato supportato dal Progetto di Rilevante Interesse Scientifico e Tecnologico anno 2009 dal titolo "Fruizione assistita e context aware di siti archeologici complessi mediante dispositivi mobili", finanziato dalla Seconda Università degli Studi di Napoli.

Bibliografia

- [Dietrich et al, 2005] Dietrich Jens and Chris Elgar. "A formal description of design patterns using OWL." *Software Engineering Conference, 2005. Australian*. IEEE.
- [France, Robert B. et al, 2004] France, Robert B., et al. "A UML-based pattern specification technique." *Software Engineering, IEEE Transactions on* 30.3(2004):193-206.
- [Gamma et al, 1995] Gamma et al, "Design patterns: Elements of reusable object-oriented software",1995.
- [Taibi, 2007] Toufik Taibi, "Design Patterns Formalization Techniques", IGI Publishing, 2007
- [Taibi, 2003] Taibi, Toufik, and David Chek Ling Ngo. "Formal Specification of Design Patterns - A Balanced Approach." *Journal of Object Technology* 2.4 (2003): 127-140.
- [Eden et al, 1998] Eden, Amnon H., Yoram Hirshfeld, and Amiram Yehudai. "LePUS- a declarative pattern specification language." (1998).
- [Yu et al, 2010] Yu, Kui, et al. "Using OWL-S for Formalizing Workflow Patterns." 2nd International Conference on Information and Multimedia Technology (ICIMT 2010)