

DEVE L'UNIVERSITÀ OCCUPARSI DI TRASFERIMENTO TECNOLOGICO?

Le università e i centri di ricerca possono divenire partner efficaci nel supportare la competitività del sistema industriale. Solo negli ultimi anni è iniziata, infatti, una nuova considerazione del ruolo dell'università nello sviluppo economico e si è assistito a un progressivo interesse di quest'ultima alle attività di innovazione tecnologica affiancando, ai tradizionali interventi nel campo sia della formazione professionale sia della ricerca, un intervento centrato sulla collaborazione stretta tra università e industrie, nonché sullo sviluppo economico.

1. INTRODUZIONE

Se si osservano alcuni indicatori dei livelli di innovazione delle nazioni, per esempio il TAI (*Technology Achievements Index*), basato sul numero di brevetti effettuati, sul livello di scolarità, sulla creazione e diffusione di tecnologie nuove e mature, si scopre che alcune nazioni sono valutate come più disposte alla diffusione dell'innovazione tecnologica rispetto ad altre: la Finlandia più che la Francia, Hong Kong più del resto della Cina, gli USA più che Israele; negli USA, inoltre, la capacità di un Paese di produrre e commercializzare tecnologie innovative è stata misurata tramite il *National Innovative Capacity*, che offre una valutazione dinamica attualizzata del potenziale di innovazione di una singola nazione, analizzato con una prospettiva a medio-lungo termine [5]. Cambiano le posizioni ai vertici, ma indipendentemente dal parametro di misura, senza sorprese l'Italia, nell'ambito dell'innovazione, della ricerca e sviluppo o nella creazione d'impresa, non tiene certo il passo degli altri Paesi avanzati. Se la domanda è "come è possibile favorire

l'innovazione in Italia?", la risposta richiama regole che sono ormai note a tutti: un effetto di portata nazionale e di consistente durata, si ottiene solo con la congiuntura tra la diffusione della cultura¹ dell'innovazione e la presenza di un terreno reso fertile dall'offerta di strumenti disponibili e agevoli sul piano politico, finanziario, di sviluppo e di accesso alle competenze e alle tecnologie.

Il processo di innovazione può seguire traiettorie differenti, a volte fondate sulle nuove tecnologie o metodologie, altre volte sulle volontà imprenditoriali. Non sempre innovazione e trasferimento tecnologico coincidono²: è anche possibile, infatti, trasferire le metodologie

¹ Nella lingua inglese si usa un solo termine, *culture*, per indicare sia l'insieme di sapere e *modi di vivere*, sia un luogo reale o virtuale dove esistano le condizioni per sviluppare un sistema biologico.

² Il telefono cellulare è un esempio di nuova tecnologia che ha determinato l'innovazione nel modo di comunicare, mentre i nuovi *concept* di prodotto, tipici del *design* strategico, portano all'innovazione attraverso un nuovo modo di percepire tecnologie note.



Sergio Campodall'Orto
Paolo Vercesi

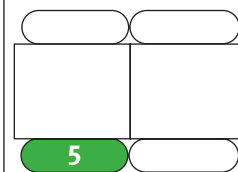
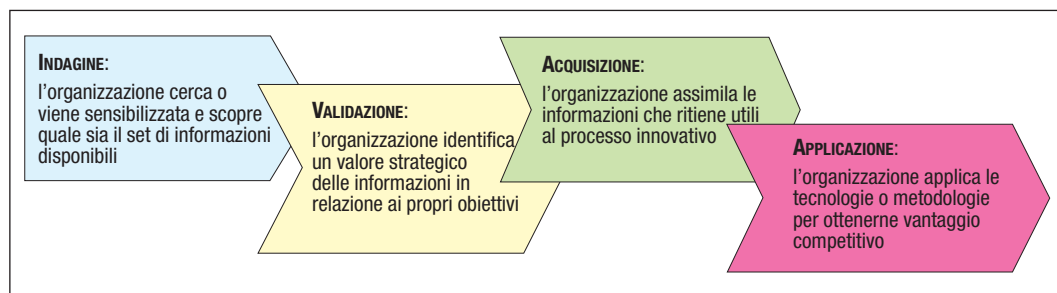


FIGURA 1
Il processo
di trasferimento
tecnologico



di approccio alle funzioni dei processi, dei prodotti o dei sistemi³, che consentono di vederli come se avessero proprietà nuove modificando la posizione spaziale o temporale delle parti che li compongono. Questo aspetto rappresenta un beneficio per chi lavora nelle università rispetto alle divisioni di R&D (*Research & Development*) tecnologica delle PMI (*Piccole e Medie Imprese*). Dall'approccio aziendale è invece importante mutuare una più puntuale focalizzazione verso gli obiettivi e, in particolare, la permanenza della possibilità di sfruttare economicamente l'innovazione.

Per capire in che modo le università o i centri di ricerca possono divenire *partner* efficaci nel supportare la competitività del sistema industriale, si deve considerare il fatto che nelle imprese il processo di innovazione si genera, nella maggioranza dei casi, dalla percezione di un differenziale di competitività sul mercato, che richiederà interventi nell'azienda per mantenere una posizione di vantaggio o per colmare il distacco dai propri competitori. Se questo fattore deriva da aspetti di tipo commerciale l'innovazione andrà, con ogni probabilità, a influire sui processi, sull'organizzazione e sulle reti di vendita. A volte, invece, il fattore determinante è di natura tecnologica e richiede alle imprese di intraprendere un percorso per generare o acquisire le conoscenze o le tecnologie e per renderle profittevoli. Questo percorso, almeno in parte, può essere guidato. Tecnologie, metodologie e il loro trasferimento rappresentano, spesso, la struttura portante

di questi processi. Alcune di esse, possono guidare l'innovazione disponendosi lungo le traiettorie del processo, altre possono abilitare questo processo distribuendosi trasversalmente a più processi⁴. In sintesi, il processo di trasferimento tecnologico può essere schematizzato [6] come illustrato nella figura 1.

2. LE CONNESSIONI INDUSTRIA-UNIVERSITÀ IN ITALIA

Negli ultimi anni, si è assistito a un progressivo interesse dell'università italiana alle attività di innovazione tecnologica affiancando, ai tradizionali interventi nel campo sia della formazione professionale sia della ricerca, un intervento centrato sulla collaborazione stretta tra università e industrie, nonché sullo sviluppo economico.

Questo fenomeno è conosciuto come "Seconda Rivoluzione Accademica"⁵ e comprende una pluralità di attività che coinvolgono le strutture accademiche e che spaziano dalle ricerche finalizzate, al sostegno al trasferimento tecnologico, alla valorizzazione dei brevetti generati dalla ricerca, alla nascita di *spin-off* universitari, fino al più recente coinvolgimento nello sviluppo economico locale. La schematizzazione delle differenti tipologie di collaborazione università-industria è rappresentata nella tabella 1, nella quale vengono sottolineati gli aspetti principali che la caratterizzano.

Il connubio università-industria, in Italia, ha

³ Un esempio di metodologia di innovazione a supporto della pianificazione strategica di sistema è data dal foresight tecnologico.

⁴ È perciò possibile classificarle come "infratecnologie" o "cross-cutting technologies" [4].

⁵ La Prima Rivoluzione Accademica si riferisce all'introduzione della ricerca in ambito universitario e si colloca tra la fine del XIX secolo e l'inizio del XX. La Seconda, invece, ha il suo *incipit* nella seconda decade del 1900 ed è tuttora in corso.

LE CONNESSIONI INDUSTRIA - UNIVERSITÀ

Connessioni	Obiettivi economici	Strategie di sviluppo
Connessioni nel campo della ricerca e sviluppo	Generazione e applicazione della conoscenza scientifico-tecnologica	- Potenziare la ricerca di base e la ricerca applicata nelle università - Aumentare l'interazione tra ricerca di base, ricerca applicata e processi di sviluppo a livello dell'università e dell'industria
Connessioni per lo sviluppo imprenditoriale	Efficienza e innovazione nelle industrie nuove e nelle industrie esistenti	- Supportare il meccanismo tecnologico per aumentare l'innovazione a livello delle grandi e delle piccole imprese - Incoraggiare l'offerta dei servizi per aiutare gli imprenditori a creare e sviluppare nuove aziende
Connessioni per lo sviluppo delle risorse umane	Formazione di personale a tutti i livelli nell'area scientifico-tecnologica	- Potenziare i programmi formativi di rango universitario in campo scientifico, ingegneristico, aziendale e tecnico-professionale - Potenziare il <i>training</i> e <i>retraining</i> nelle <i>skill</i> tecnologiche - Miglioramento delle informazioni sul mercato del lavoro e le previsioni occupazionali

TABELLA 1

Le connessioni industria-università.
Fonte: [8]

vissuto una fase di florido sviluppo nei periodi che hanno preceduto e seguito la seconda guerra mondiale: esempi eclatanti di tale rapporto sono stati i risultati conseguiti dal Centro di studi sulla gomma sintetica, diretto dal prof. Giulio Natta, presso il Politecnico di Milano, che hanno portato all'invenzione del polipropilene e che hanno consentito alla Montecatini, che finanziava il Centro di ricerca, di diventare leader mondiale nella produzione del nuovo prodotto chimico e a Natta di essere insignito del premio Nobel⁶.

La successiva trasformazione a partire dagli anni '70 dell'Accademia in università "di massa", unita a una maggiore diffidenza da parte dell'industria nella ricerca universitaria, ha comportato un drastico ridimensionamento delle attività di ricerca pubblica con il conseguente allontanamento degli studi mi-

ratati a sostenere un indirizzo industriale dalle attività di ricerca.

Solo in anni relativamente recenti, è iniziata una nuova considerazione del ruolo dell'università nello sviluppo economico.

3. UNIVERSITÀ E TRASFERIMENTO TECNOLOGICO

Uno dei motivi principali di quella trasformazione, di cui si è parlato nel paragrafo precedente, è rappresentato dalla necessità di risorse per contrastare la riduzione dei fondi di ricerca pubblici. La tendenza internazionale verso la diminuzione dei finanziamenti pubblici nella ricerca è testimoniata dal grafico (Figura 2), nel quale emerge la riduzione che ha interessato il sistema statunitense e francese: rispettivamente del 40% e del 20%. Nel decennio considerato, l'Italia non ha subito tale riduzione e l'incidenza degli investimenti pubblici si è mantenuta su valori appena superiori al 50%.

Tale riduzione delle risorse sul fronte pubblico non è stata, tuttavia, compensata da investimenti da parte delle imprese, le quali hanno pre-

⁶ Per una completa descrizione del ruolo del Politecnico di Milano nel sostegno allo sviluppo dell'industria lombarda si rimanda al testo di Campo-dall'Orto [2].

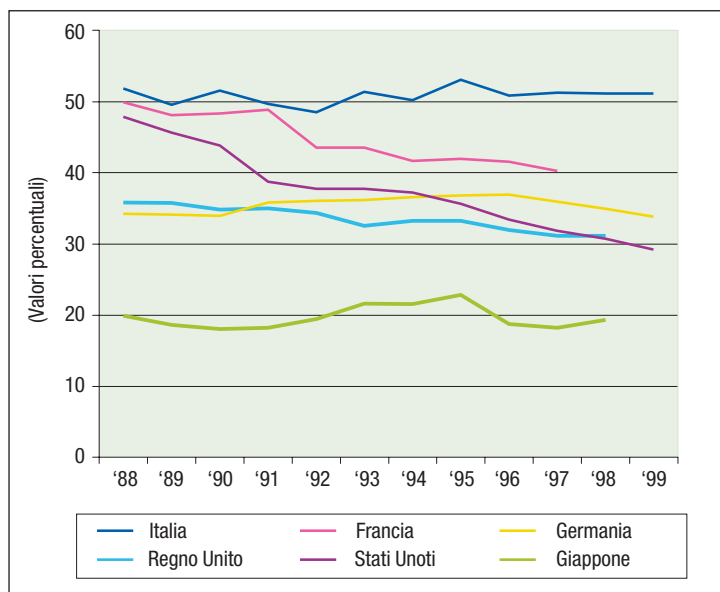


FIGURA 2 Spesa per R&S (Research and Study), globale, nei principali Paesi industrializzati (incidenza percentuale e finanziamento pubblico sul totale nazionale).
Fonte: AIRI

ferito privilegiare gli investimenti a breve termine, piuttosto che quelli basati sulla ricerca di base o pre-competitiva e ciò ha determinato un ridimensionamento delle quote di investimento complessivo nella ricerca scientifica.

Questi fenomeni, più generali, sono stati accompagnati, nei diversi paesi, da interventi normativi più specifici sul sistema universitario e della ricerca. Per quanto riguarda l'Italia, il processo di cambiamento ha inizio con l'emanazione della legge 168/89 che, istituendo il MURST (Ministero dell'Università, della Ricerca Scientifica e Tecnologica), oggi denominato MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca), agli articoli 6 e 7 fissa, in attuazione dell'art. 33 della Costituzione, i capisaldi dell'autonomia didattica, scientifica, organizzativa, finanziaria e contabile delle università, prevedendo che quest'ultime si diano ordinamenti autonomi con propri statuti e regolamenti. Questa norma ha, di fatto, determinato un cambiamento nella filosofia d'azione: le università non si limitano più a fare ciò che è consentito, bensì si attivano per fare ciò che non viene proibito. In applicazione di questa legge, esse si danno statuti propri che, accanto alle attività tradizionali della ricerca e della didattica, iniziano a prefigurare diverse iniziative e nuovi modi d'azione per valorizzare le proprie risorse. Il processo subisce un'accelerazione con la legge 537/93 che, ridefinendo le modalità di finanziamento delle università, introduce sostanziali cambiamenti relativamente all'erogazione dei finanziamenti

centrali (quanto a entità, aggregazione e politiche di riequilibrio tra le sedi), all'eliminazione di vincoli di spesa, alla libertà di utilizzo dei fondi, alla responsabilizzazione nei confronti delle uscite, al coinvolgimento per il reperimento delle fonti di finanziamento, alla libertà nel fissare le tasse di iscrizione previste per gli studenti. Oltre che a una maggiore attenzione ai processi decisionali relativi all'impiego delle risorse, le università sono così orientate, concretamente, ad agire sul mercato per conseguire il proprio equilibrio economico-finanziario.

Per far fronte a questo cambiamento, le università hanno iniziato a utilizzare, in forma più articolata, strumenti *ad hoc*, quali uffici dedicati al collegamento con il mondo industriale organizzati secondo i modelli statunitensi degli *industrial liason office*, i *patent e licence office* per la valorizzazione delle conoscenze sviluppate nei centri di ricerca universitari, gli incubatori per l'offerta di servizi alle nuove imprese nate da spin-off della ricerca accademica.

Università e industria possono, dunque, procedere congiuntamente sulla strada della collaborazione, ma la velocità, con cui i due sistemi possono cooperare, può essere influenzata dalle scelte dell'Amministrazione Pubblica centrale o locale. Basti pensare al ruolo che rivestono: l'ammontare degli investimenti dedicati alle R&S o ai programmi di innovazione tecnologica; l'imposizione fiscale e, in particolare, il regime di detassazione delle imposte; gli scambi tra ricercatori pubblici e industria sovvenzionati dallo Stato; la presenza o meno di aree attrezzate quali i parchi scientifici o incubatori tecnologici.

L'inserimento del terzo attore, il governo, che affianca i due precedenti (università e industria), ha comportato un insieme di reazioni conosciute, in campo accademico, come *Triple Helix model* [3]. La prima forma di tale rapporto si basa sulla preminenza del sistema di governo (Figura 3, fase 1) ed è caratteristico dei sistemi monarchici ed ex-socialisti. Il secondo (Figura 3, fase 2), prevede un intervento del Governo verso gli altri attori lasciando al mercato la possibilità di intervenire tra di essi. Infine, il terzo stadio evolutivo, quello più avanzato, prevede una forte interconnessione nell'operato dei tre soggetti con un conseguente incremento della complessità dell'intero sistema in cui ogni azione di un soggetto si ripercuote immediatamente sugli altri.

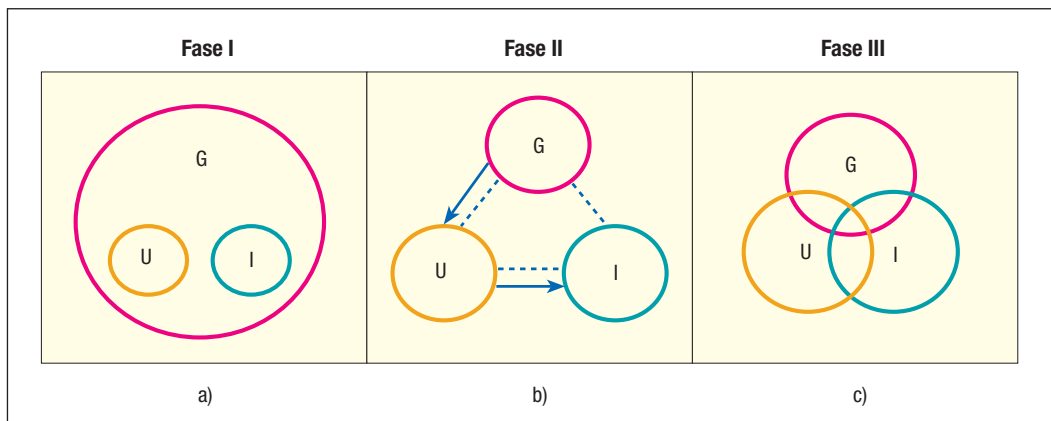


FIGURA 3
Triplex Helic model
Fonte: H. Etzkowitz

Ma, soprattutto, l'ingresso del Governo, oltre a condizionare i sentieri della ricerca, ha indotto l'università a ricoprire, accanto alle tradizionali funzioni, posizioni-chiave nello sviluppo economico territoriale, facendole assumere sempre più il ruolo di "università imprenditoriale". Questo ruolo imprenditoriale dell'università si è manifestato chiaramente con la rapida crescita di spin-off derivati dalla ricerca pubblica [1]. Incubatori tecnologici e nuove imprese sono due delle caratteristiche della moderna università e rappresentano un vero motore di sviluppo economico basato su una nuova modalità di trasferimento tecnologico. Anche in Italia i maggiori atenei: il Politecnico di Torino, l'Università di Bologna, la Sant'Anna di Pisa e il Politecnico di Milano, hanno avviato esperienze, proprio nel campo dello sviluppo imprenditoriale.

4. UN CONSORZIO PER IL TRASFERIMENTO TECNOLOGICO

Politecnico Innovazione⁷ nasce nel marzo del 2000 come immagine stessa del processo d'innovazione interno al Politecnico di Milano. Lo scopo è di integrare le esperienze precedenti portate avanti direttamente dai Dipartimenti dell'Ateneo e da altre strutture preposte al supporto dei rapporti dell'Ateneo con le imprese. L'interazione dell'università

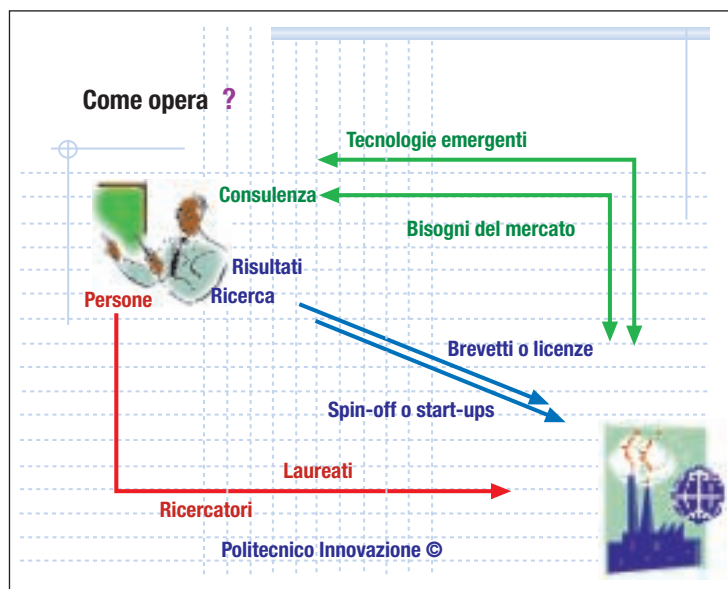


FIGURA 4
Schema di funzionamento del trasferimento tecnologico universitario

con il sistema di sviluppo sociale (Figura 4) si concretizza attraverso:

- la trasmissione di sapere alle persone, prima di tutto tramite i laureati o i ricercatori che fanno ingresso nel sistema industriale;
- il trasferimento di *know-how* tramite i servizi a terzi operati dai centri di conoscenza degli atenei;
- la cessione o licenza di brevetti;
- la convergenza di persone e sapere con la creazione di nuove imprese.

Durante questo processo, lo scambio avviene anche dall'impresa verso le università le

⁷ Politecnico Innovazione è un Consorzio controllato dal Politecnico di Milano e sostenuto da: ApiComo, ApiLecco, ApiMilano, Associazione Ind.le di Piacenza, Assolombarda, CCIAA di Como, CCIAA di Lecco, CCIAA di Milano, CNAMilano, Istituto Scientifico Breda SpA, Finlombarda, Polo Scientifico e Tecnologico Lombardo, Unione Ind.li di Como, Unione Ind.li di Lecco.

quali, dovendo affrontare in maniera più diretta il mercato, recepiscono inevitabilmente i bisogni di quest'ultimo in termini di formazione e di priorità negli indirizzi da dare alla ricerca applicata.

Dopo due anni di attività, Politecnico Innovazione mantiene questo aspetto trasversale incorporando quattro aree nelle quali si ritrovano i progetti di trasferimento tecnologico, del servizio brevetti, del supporto ai finanziamenti, della creazione d'impresa e di formule innovative di tirocinio.

Le collaborazioni dell'Area Servizi alle Imprese sono prevalentemente attivate con PMI e consistono in consulenze su specifiche esigenze tecnologiche segnalate dalle imprese stesse e, inoltre, prevedono prove, misure e studi di anticipazione tecnologica. In particolare, è piuttosto frequente che tali esigenze tecnologiche siano relative a settori per i quali non esista una vera e propria competenza in materia o dove l'integrazione sistemica richieda di integrare differenti competenze. Questo accade tipicamente nella progettazione delle macchine industriali, per le quali sarebbe auspicabile, per ottenere un prodotto veramente competitivo, poter unire know-how a:

- funzioni del processo;
- caratteristiche delle materie prime o semilavorati da processare;
- specifiche strutturali ed efficienza cinematica della macchina;
- opzioni alternative in funzione di criteri come la leggerezza;
- azionamenti elettro/oleo/pneumatici;
- controllo avanzato e sulle connessioni della macchina nei sistemi informativi aziendali;
- criteri ambientali relativi al ciclo di vita della macchina;
- movimentazione, logistica e sicurezza;

■ compatibilità elettromagnetica e spesso anche sull'idraulica.

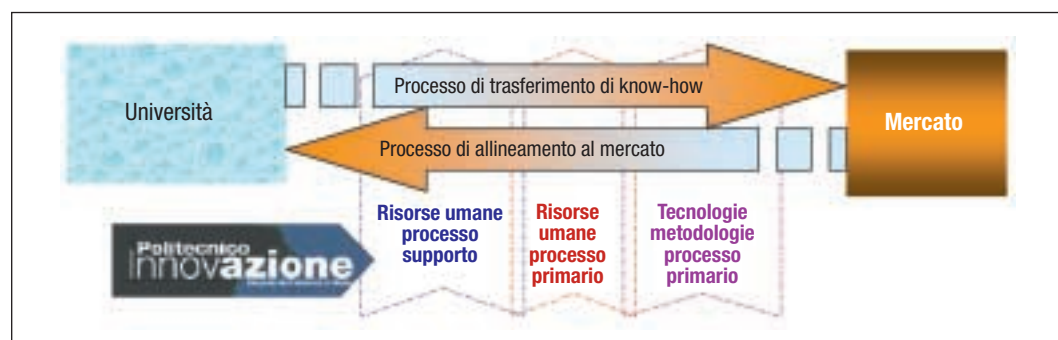
5. LA GESTIONE DEL PROCESSO DI TRASFERIMENTO TECNOLOGICO

L'innovazione, come introduzione di una nuova metodologia o tecnologia, comporta un beneficio solo se l'evoluzione presenta sufficienti caratteristiche per poter attrarre ed essere sostenibile nel contesto in cui viene introdotta. Obiettivo di Politecnico Innovazione è attivare e coordinare i processi di supporto al processo primario di cessione della conoscenza (Figura 5).

In parte, le attività delle divisioni di *Consorzio Politecnico Innovazione* (CPI) si concentrano sul *competence mapping*, sia per distribuire consulenze e progetti dedicati alle esigenze delle imprese, con l'eventuale supporto di agevolazioni finanziarie, sia per verificare la possibilità di presentare domande di brevetto e la successiva fase di *licensing*. L'intervento di Politecnico Innovazione genera un'organizzazione temporanea "a matrice" tra i gruppi di ricerca e gli operatori interni del Consorzio; quest'ultimo assume, inoltre, le responsabilità di *project management*, per la configurazione e il governo del progetto, per anticipare i vincoli, costruire il *team* e gestire le relazioni, attivare un sistema di controllo e per fornire i servizi amministrativi (Figura 6). Ciò consente ai ricercatori e ai docenti in università di operare su progetti con aziende esterne, rimanendo all'interno di un'organizzazione allargata dell'Ateneo: questa opportunità di operare diminuisce inoltre il rischio di *brain drain* [11] dalle università e dal Paese.

Il supporto del Consorzio al processo permette

FIGURA 5
Schematizzazione
del processo
di trasferimento
tecnologico



di attivare anche un sistema di controllo *in itinere* sulle dimensioni, sugli aspetti temporali e sulla qualità del progetto. I parametri di controllo che vengono utilizzati si riferiscono ai principali *drive* dei processi collaborativi: efficacia, adeguatezza, efficienza, sforzo e funzionalità [10]. Tanto più nel progetto innovativo sono presenti processi complessi e interconnessi, tanto più le dimensioni aziendali influiscono sulla convenienza per l'azienda di esternalizzare parte di questi progetti (Figura 7).

Le esigenze che le imprese esprimono vengono spesso ricalibrate, in funzione dell'effettivo intervento dei gruppi di ricerca, ma evidenziano questa necessità di supporto nei progetti complessi.

Anche se il *target* è l'azienda, la comunicazione è rivolta ai decisori tecnologici e, nella tipicità del sistema industriale dell'ambito geografico in cui opera il Politecnico di Milano, si sposta dai vertici imprenditoriali verso i responsabili dell'area R&D delle PMI. Una caratteristica che rende efficace l'atteggiamento di collaborazione con l'università e che si riscontra nelle imprese, è il riconoscimento dell'im-

parzialità che l'istituzione universitaria ha se confrontata con i consueti fornitori di servizi tecnologici presenti sul mercato. Questo atteggiamento crea opportunità di collaborazione negli studi di anticipazione e confronto tra diverse tecnologie e per la credibilità, e quindi portabilità, delle analisi e misure effettuate in collaborazione con centri universitari.

6. CREARE IMPRESE: UN NUOVO MODO PER TRASFERIRE COMPETENZE

I sistemi economici dei principali paesi occidentali hanno compreso che lo sviluppo industriale di una specifica area geografica si esplicita anche attraverso l'incentivazione della nascita di nuove imprese *high-tech* derivate da ricerche sviluppate all'interno dei centri e dipartimenti universitari.

Seguendo questa linea di condotta, Politecnico Innovazione ha deciso di avviare un Acceleratore di Impresa nel proprio *campus* di Bovis. Nel primo biennio di attività l'Acceleratore ha contribuito allo sviluppo di 10 *start up*, tut-

FIGURA 6

CPI nel ruolo di project manager

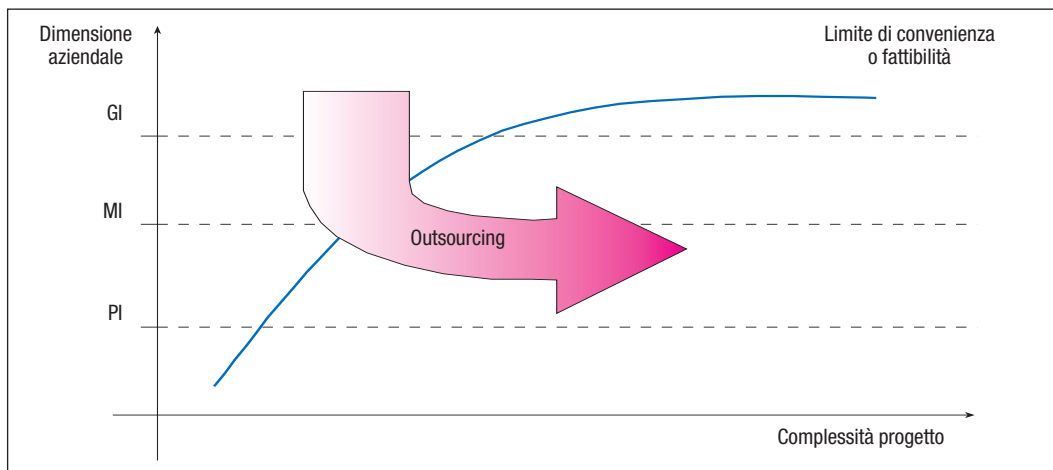
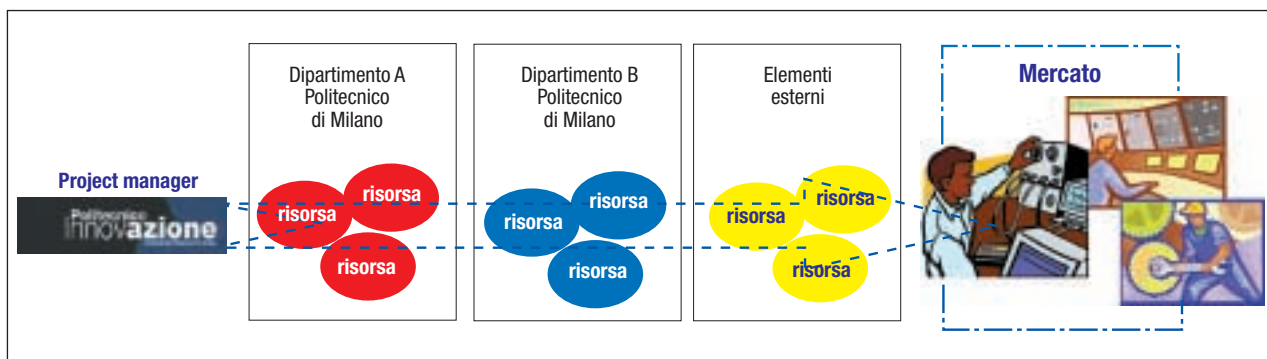


FIGURA 7

Outsourcing in funzione di dimensione aziendale e complessità progetto

te focalizzate nelle attività *internet based*.

Alle imprese incubate, che accedono tramite una selezione basata sulla presentazione di un *business plan*, Politecnico Innovazione offre l'utilizzo degli spazi fisici arredati, l'infrastruttura informatica (server, posta elettronica) e di telecomunicazione (telefoni, accesso Internet). Inoltre, vengono offerti servizi comuni, come segreteria, *reception*, utilizzo di sala riunioni, nonché servizi esclusivi quali *tutoring* a supporto della pianificazione e organizzazione del progetto di sviluppo imprenditoriale; marketing; consulenza finanziaria.

Nel corso del primo anno di attività dell'Acceleratore, è stata fornita assistenza informativa a 142 neo imprenditori, di questi 87 hanno richiesto un intervento di consulenza nella realizzazione del business plan. Al termine di questa fase, 34 neo imprenditori hanno fatto domanda di ammissione ai servizi dell'Acceleratore e, da questi, sono state individuate otto aziende da incubare.

Per alcune delle imprese, che attualmente usufruiscono dei servizi dell'Acceleratore, si aprono interessanti prospettive di sviluppo nel mercato.

Attualmente, sono incubate presso l'Acceleratore numerose aziende.

□ **Neptuny.** Prima impresa italiana specializzata nella fornitura di servizi di monitoraggio e analisi delle prestazioni dei siti Internet dal punto di vista degli utenti finali, mettendo a disposizione dei propri clienti, aziende e istituzioni per le quali la qualità del servizio offerto rappresenta uno dei fattori critici di successo, un'infrastruttura di misurazione ritagliata sulla realtà italiana.

□ **Alfaproject.net.** ASP (*Application Service Provider*) per le aziende che intendono sviluppare il proprio Business attraverso Internet per quanto concerne lo studio e la realizzazione del progetto, sviluppando e implementando soluzioni di comunicazione e commercio elettronico.

□ **DesignVillage.it.** Primo portale internazionale interamente dedicato al Design in tutte le sue dimensioni culturali, produttive ed economiche.

□ **TRE-Telerilevamento Europa.** Elaborazione dati ottenuti da sistemi radar satellitari di tipo SAR (*Synthetic Aperture Radar*), tramite la Tecnica PS, un sofisticato algoritmo sviluppato e brevettato presso il Politecnico di Milano e in uso esclusivo. Le informazioni acquisite permettono di individuare movi-

menti crostali del terreno e ricostruire modelli altimetrici su vasta scala.

□ **Wise-Lab.** Servizi di analisi, progettazione e sviluppo di applicazioni Internet complesse, mediante l'utilizzo di una soluzione architetture innovativa (*WiseModel*) di propria ideazione.

□ **Gaia.** Servizi di sicurezza e controllo remoto via web.

Per rispondere alla maggiore richiesta di nuove imprese *high tech*, provenienti dal sistema università, il Consorzio Politecnico Innovazione ha individuato in uno stabile sito in via Garofalo, angolo via Paisiello, uno spazio adatto a creare una nuova ulteriore sede dell'Acceleratore. Il nuovo Acceleratore è situato in prossimità della sede storica del Politecnico di piazza Leonardo da Vinci, ciò consentirà di mantenere quella contiguità fisica che facilita le interrelazioni e gli scambi fra università e neo-imprese.

7. CONCLUSIONE

L'esperienza di Politecnico Innovazione dimostra che la collaborazione tra università e industria è possibile, pur con tutte le difficoltà derivate dal diverso ambito culturale in cui le due realtà interagiscono.

Se in altri Paesi, come per esempio gli Stati Uniti o nei Paesi Scandinavi, tali opportunità di collaborazione sono già ampiamente sviluppate, e ne è una dimostrazione le aree imprenditoriali cresciute a lato delle principali università tecnologiche, in Italia non si è finora verificato un reale *link* di attività.

La risposta alla domanda sintetizzata nel titolo del presente articolo non può, quindi, che essere positiva. Rimane da definire le modalità e soprattutto il periodo temporale necessario affinché l'Università si attrezzi per sviluppare una reale attività di trasferimento tecnologico. Non sono sufficienti, infatti, i contratti di ricerca affidati dalle imprese all'Università per modificare rapidamente lo *status quo*, sono invece necessari interventi di struttura sia interni che esterni agli Atenei. Per quanto concerne l'aspetto interno, è necessario che la valutazione sui singoli ricercatori non sia solo attraverso il numero di pubblicazioni, ma occorre anche porre considerazione sulla forma di collaborazione che il ricercatore ha con l'industria e le attività che egli svolge per sviluppare e proporre idee applicabili al processo produttivo.

Per quanto concerne l'aspetto esterno, sono indispensabili sostegni pubblici mirati a supportare la collaborazione università-industria, come i finanziamenti destinati alle metodologie di *transfer by head* o per lo sviluppo di nuove imprese *high-tech*.

Bibliografia

- [1] Bergami M, Sombrero M: *Il nuovo ruolo del sistema universitario*. Articolo apparso sul portale Impresando (www.impresando.it), 14 febbraio 2001.
- [2] Campodall'Orto S: *Innovazione e Sviluppo a Milano*. AIM, 1996, a cura di.
- [3] Etzkowitz H, Leydesdorff L: *The Triple Helix as a Model for Innovation Studies*. The 2nd International Triple Helix Conference Report, New York City, gennaio 1998, pubblicato in *Science & Public Policy*, Vol. 25, 1998.
- [4] European Research Area: *Thematic priorities*. European Commission, 2001.
- [5] Furman JL, Porter ME, Stern S: The determinants of national innovative capacity. *Research Policy*, Vol. 31, 2002.
- [6] Grant RM: *L'analisi strategica per le decisioni aziendali*. il Mulino, Bologna, 1999.
- [7] Pero L: *Corso di organizzazione aziendale*. MIP Politecnico di Milano 2000.
- [8] Piccaluga A, Varaldo R: Un ponte tra industria e università per rilanciare la ricerca. *L'Impresa*, n. 8, 1994.

- [9] Raymond E: *The Cathedral and the Bazaar*. Paperback Edition, 2001.
- [10] Roveda C, Vercesi P, et al.: *The mobility of academic researchers*. JRC/IPTS-ESTO Study, 2001, p. 25.
- [11] Trott P, Cordey-Hayes M, Seaton R: Inward technology transfer as an interactive process. *Tech-novation*, Vol. 15, 1995.

SERGIO CAMPODALL'ORTO si è laureato in Economia e Organizzazione Aziendale presso l'Università degli Studi di Milano. Dal 1997 insegna Economia e Organizzazione Aziendale presso il Politecnico di Milano. Dal 2000 è Consigliere Delegato del Consorzio Politecnico Innovazione. La sua attività scientifica è rivolta al trasferimento di tecnologia innovativa, creazione di nuova imprenditoria, start-up, sviluppo e cooperazione università-industria per progetti di ricerca congiunti. Fa parte della Commissione Scientifica CeS&T della Fondazione Rosselli. È autore di numerose pubblicazioni e libri apparsi sulle più autorevoli riviste nazionali e internazionali.
e-mail: campodallorto@cpi.polimi.it

PAOLO VERCESI laureatosi in Ingegneria al Politecnico; attualmente è docente di Fisica Tecnica e seminari-sta nel corso di Missilistica, ed è responsabile dell'area servizi alle imprese di Politecnico innovazione. Ha svolto e porta avanti un intenso programma di consulenze nei settori dell'innovazione tecnologica, dell'e-business e della Qualità con aziende private e istituzioni. Collabora con la Fondazione Rosselli in progetti nazionali ed europei sul foresight strategico e sulla valutazione dell'impatto delle tecnologie emergenti sui sistemi industriali.
e-mail: paolo.vercesi@polimi.it

Un esempio di collaborazione, tra centri di ricerca (il Dipartimento di Elettronica e Informazione del Politecnico e il CNR), con un centro di sviluppo applicativo nell'azienda Italtel riguarda uno studio di fattibilità per innovare radicalmente la filosofia di sistema nel mondo della telefonia fissa. Gli obiettivi del progetto sono stati, da una parte, verificare una possibile realizzazione di un sistema *open source*, che mantenesse però le caratteristiche *real time* e *fault tolerant* del sistema in uso e, dall'altra, valutare le prestazioni e la garanzia di affidabilità del sistema. Vittorio Rampa, docente della sezione di Telecomunicazioni (Dipartimento di Elettronica e Informazione) del Politecnico di Milano spiega come: "L'approccio innovativo, proprio di ogni progetto *open source*, sia chiaramente delineato nel documento *La Cattedrale ed il Bazar* di E.S. Raymond" [9]. Secondo questa filosofia, anziché puntare, nello sviluppo di un sistema, a un modello centralizzato (la Cattedrale), attentamente curato da pochi geni o artisti, è preferibile un grande e confuso Bazar pullulante di nuove idee, concetti e approcci differenti apportati da un gran numero d'entusiasti utenti/sviluppatori. In altre parole, si può affermare che "dato un numero sufficiente di occhi, tutti i *bug* vengono a galla". Quindi, un sistema *open source* può svilupparsi molto più velocemente di uno proprietario. Nel caso di Italtel, è stato possibile realizzare velocemente un prototipo riutilizzando il codice disponibile per un'altra applicazione, verificare in modo altrettanto rapido le tecniche di software *fault tolerant* emerse durante lo studio, constatarne le prestazioni *real time* e scoprirne i limiti applicativi. Con l'approccio *open source* è importante saper riconoscere le buone idee progettuali altrui, per poterle integrare con le proprie. Questo non implica tanto il fatto che, per lo sviluppo avanzato del software *open source*, non saranno più essenziali le buone idee o la capacità personale, quanto che tutte queste qualità possono essere ulteriormente amplificate tramite la costruzione di una comunità volontaria d'interesse. Liberare le sinergie tra mondo accademico e industriale è uno dei compiti primari del Consorzio Politecnico Innovazione: la filosofia *open source* si sposa con questo compito, evidenziando le peculiarità di ciascun attore ed esaltandone i ruoli e le capacità. Per l'azienda significa, inoltre, avere un'apertura, anche internazionale, e ottenere una presenza nei *forum* specifici. Graziano Damele di Italtel ricorda che alle caratteristiche dell'attuale sistema si uniranno quelle derivanti dall'adozione di un *kernel Linux* che incrementerà le possibilità di impiego nei prodotti Italtel. L'adozione di prodotti software derivanti dal mondo *open source*, avrà particolare riguardo per il sistema operativo Linux, applicato a tutti i livelli di elaborazione, elaboratori periferici, comandi di modulo ed elaboratori di servizio. Grazie all'alta portabilità di tale software su diverse piattaforme commerciali si dovrebbero ottenere drastici abbassamenti dei costi di investimento e manutenzione, usufruendo dei prodotti acquisiti in accordo con la licenza GPL (Gnu Public Licence).

INVESTIRE NELL'INNOVAZIONE TECNOLOGICA

La sopravvivenza e il successo delle imprese dipendono dalla loro capacità di rinnovarsi e di ricercare nuove strade per la creazione di valore. La maggiore ragionevolezza e ponderazione delle scelte in tecnologie innovative, hanno portato a un utilizzo più strategico delle stesse, con l'obiettivo principale di ottenere un ritorno sugli investimenti. L'intervento è focalizzato sulle componenti che hanno reale valore aggiunto e sulla gestione della conoscenza che diventa un punto centrale in azienda.

1. LE DINAMICHE NELLE DECISIONI DI INVESTIMENTO IN IT DELLE AZIENDE

Le nuove tecnologie dell'informazione e della comunicazione sono da considerarsi la *conditio sine qua non* per poter concretamente agevolare la trasformazione dei modelli organizzativi e di *business* dell'impresa, ma qualcosa è cambiato nella domanda del settore IT (*Information Technology*).

Se è vero che le recenti dinamiche del mercato IT hanno portato a un pesante rallentamento negli investimenti in tecnologie, è altrettanto vero che si è assistito anche a un profondo cambiamento nella spesa IT delle aziende.

Queste ultime si muovono, infatti, con più prudenza e maggiore ragionevolezza nella scelta degli investimenti, puntando a un utilizzo più strategico delle tecnologie con l'obiettivo principale di ottenere un ritorno sugli investimenti effettuati.

Le tecnologie non vengono più considerate come la soluzione di tutto, ma come comple-

tamento delle leve competitive tradizionali e come supporto dei vantaggi competitivi già esistenti.

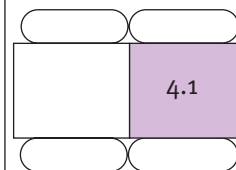
Si tratta di un atteggiamento conservatore e cauto, ma giustificato dalla delusione delle aspettative del precedente periodo di grande euforia della *New Economy*.

Le aspettative di crescita, oltre misura, del mercato IT e gli investimenti in tecnologia sono stati superiori a quanto fisiologicamente richiesto dalle imprese. L'utilizzo così intensivo dell'IT, ha provocato, infatti, una sorta di "indigestione" del mercato, arrivando a un eccesso di disponibilità rispetto all'usabilità nei principali mercati (USA, Giappone, Germania), con inevitabili ripercussioni sulle altre nazioni.

Trascinati dall'entusiasmo generale, in molti hanno creduto che Internet potesse trasformare tutto rendendo completamente obsolete le tradizionali regole competitive, trascurando spesso di supportare gli investimenti in tecnologie con adeguati piani di riorganizzazione, ridisegno dei processi e gestione del cambiamento, con effetti a vol-



Paolo Bordin
Alessandra Selva



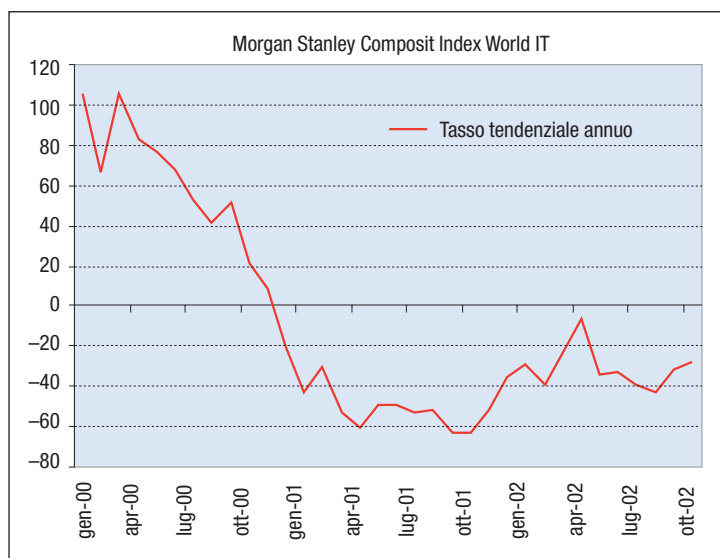


FIGURA 1
Il settore
dell'Information
Technology:
andamento
dell'indice MSCI
(Fonte: Datastream)

te addirittura controproducenti all'organizzazione stessa.

Se si analizza il settore dell'Information Technology, rappresentato convenzionalmente dall'indice di Morgan Stanley, MSCI (Morgan Stanley Capital International Indices) World Index (Figura 1), si evince come, dopo la lunga fase di contrazione del tasso tendenziale annuo (iniziata a marzo 2000 e che per tutto il 2001 ha assunto valori negativi), il settore ha cominciato a mostrare un forte rialzo agli inizi del 2002, guidato dalle aspettative di ripresa dell'economia mondiale. Ma le prospettive, moderatamente rosee, di una ripresa del settore tecnologico sono state nuovamente deluse e ridimensionate da una crescita economica inferiore alle aspettative portando l'indice ancora al ribasso.

Si riconferma, quindi, un contesto caratterizzato da un'economia sempre più interconnessa, incerta e discontinua in cui le aziende si muovono, con sempre più prudenza e concretezza, guardando agli investimenti con un'ottica di lungo periodo e allungando i cicli di decisione di investimento.

2. LA TENDENZA DELL'ATTUALE SPESA IT

Il rallentamento non si traduce in un taglio indiscriminato della spesa, ma in una inversione di rotta nel processo di investimento.

Non più, quindi, rivoluzioni radicali delle

aziende con modelli di business nati e pensati per le reti, ma realtà imprenditoriali consolidate con prodotti e mercati ben definiti che chiedono a Internet un supporto di valore per migliorare i sistemi informativi tradizionali nell'ottica di ottimizzare la gestione, la distribuzione e la commercializzazione del loro business.

Gli investimenti in Information Technology, siano essi costituiti da tecnologie, architetture o applicazioni, hanno lo scopo di integrare le nuove soluzioni Web, adattandole funzionalmente alle specificità dei processi aziendali interni e delle relazioni verso fornitori e partner.

In particolare, la soluzione IT deve essere in grado di influire sui processi *core* dell'azienda. L'investimento IT deve, inoltre, creare i presupposti perché l'azienda possa crescere nel tempo e favorire l'innovazione tecnologica (Figura 2).

Un particolare punto di attenzione tra le soluzioni IT è da rivolgere agli strumenti per la gestione della conoscenza aziendale, in quanto possono costituire una vera leva strategica per la conquista e il mantenimento del vantaggio competitivo.

Il valore degli *intangible asset* prevarrà sempre di più rispetto a quello dei *tangible*, passando da un'economia basata sull'industria e sui servizi a una fondata sulla conoscenza; ed è per questa ragione che gli investimenti in innovazione tecnologica convergeranno, soprattutto, su strumenti per la gestione della conoscenza.

3. IL KNOWLEDGE MANAGEMENT

Il processo di abbattimento delle barriere spazio-temporali che limitano le comunicazioni a livello aziendale, istituzionale e sociale, insieme con lo sviluppo delle più recenti tecnologie digitali e telematiche hanno portato a un'intensificazione e a un potenziamento dei flussi di comunicazione.

Si tratta di un fenomeno di massa nel quale tutte le aziende, grandi e piccole, sono coinvolte. Per sopravvivere, non potranno far altro che sostenere processi di cambiamento, e di adeguamento, molto rilevanti e in tempi sempre più ridotti per poter mantenere la

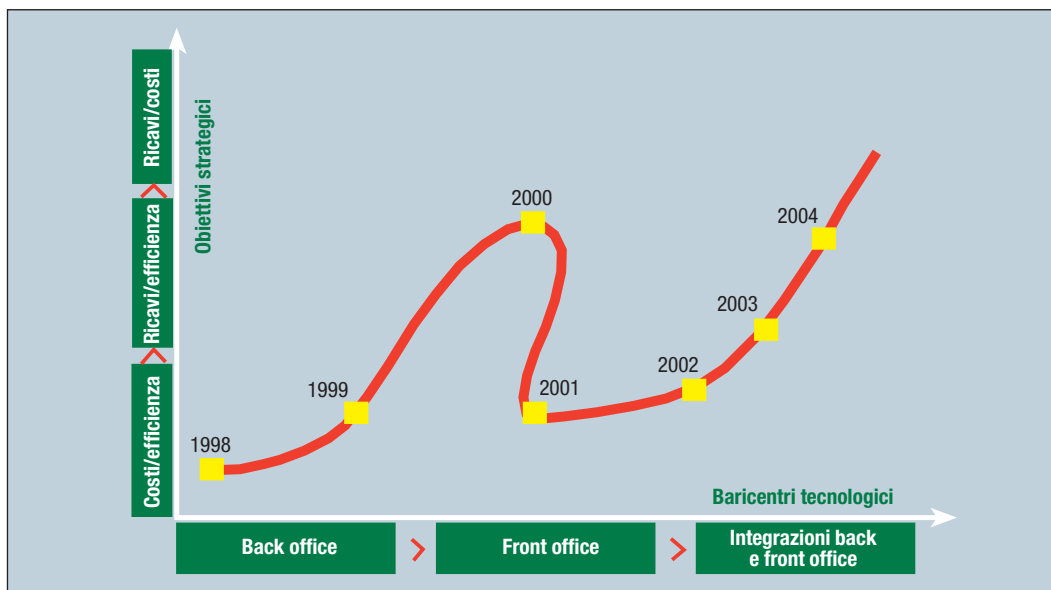


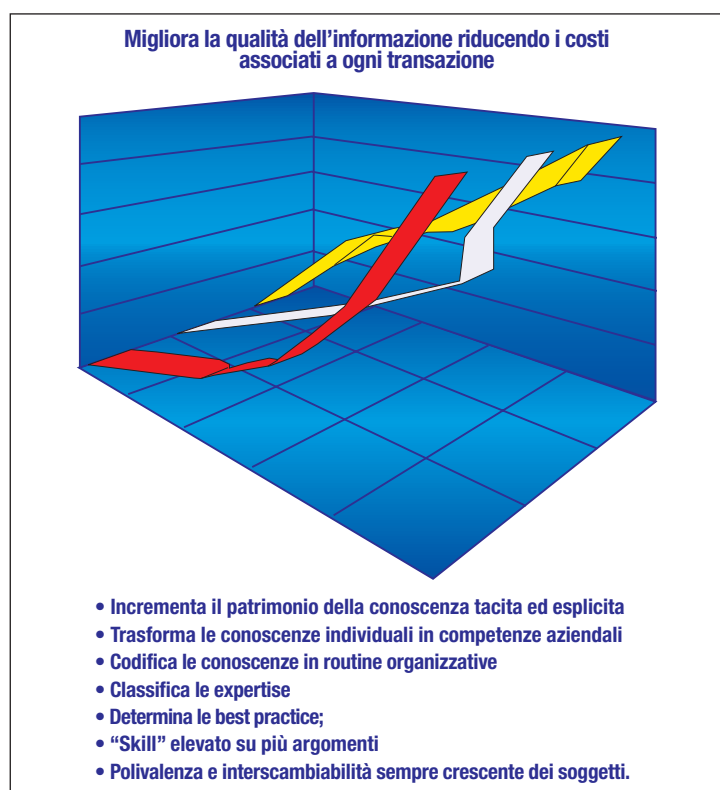
FIGURA 2
Evoluzione
del baricentro
strategico
e tecnologico
nelle aziende
italiane, 1998-2004

propria competitività. Il differenziale quindi per poter cogliere e sfruttare le opportunità, contrastando le minacce del mercato, consiste proprio nella capacità di gestire e integrare, in modo efficace, e efficiente una mole crescente di informazioni e di conoscenze. Molti autori [5, 7, 8, 18] sono concordi nel ritenere che la risorsa chiave per competere sia la conoscenza, unica fonte di vantaggio per le organizzazioni.

Le imprese che acquisiscono, gestiscono e accrescono la conoscenza, sviluppano un vantaggio competitivo che determina benefici per l'organizzazione nel suo complesso e per i singoli individui che la compongono. La generazione di *knowledge*, in ambito aziendale, può essere, infatti, interpretata come un processo che amplifica la conoscenza creata dagli individui e la cristallizza come parte della rete di patrimonio informativo aziendale, attraverso un processo di elaborazione intelligente delle risorse disponibili e degli *input* esterni.

È possibile definire il *Knowledge Management* (KM) come un processo di recupero, organizzazione e sistematizzazione di tutto ciò che, indicato genericamente come conoscenza, si traduce in azienda in quel vasto insieme di competenze, esperienze, informazioni che costituiscono l'anima del vantaggio competitivo.

Questo patrimonio informativo include i *database*, i documenti, le procedure, ma anche



le competenze, l'esperienza tacita e l'intuito dei singoli dipendenti.

Attraverso le tecnologie di knowledge management è possibile gestire, a costi sempre più contenuti, la raccolta da una pluralità di fonti eterogenee, l'elaborazione, la traslazione nel tempo e nello spazio della risorsa-conoscenza complessiva (sia tacita

FIGURA 3
I vantaggi
del Knowledge
Management

che esplicita), a beneficio di tutta l'organizzazione aziendale, migliorando la qualità e la velocità dei processi decisionali interni ed esterni (Figura 3).

In questo modo l'azienda e quindi tutte le persone che la compongono, possono affrontare e risolvere qualsiasi tipo di problematica, dalla banale routine giornaliera a una più complessa, utilizzando nel miglior modo possibile il capitale di esperienze e le capacità intellettuali aziendali.

Se è vero che il knowledge management permette una serie di economie, il suo valore reale viene conseguito soprattutto nelle organizzazioni più adattabili e orientate al futuro. Infatti, l'ostacolo più frequente che si frappone al successo di tali programmi, riguarda il fattore umano che spesso pone ancora delle resistenze al processo di "trasmissione e acquisizione di conoscenza" del KM, per mancanza di abitudine, incentivi ecc. Inoltre, è fondamentale il presidio dedicato a progetti di questo tipo durante tutte le sue fasi, dallo sviluppo al consolidamento e all'aggiornamento continuo.

4. UN CASO DI SUCCESSO: BANCA INTESABCI

L'attività di knowledge management in Banca Intesa è nata nel 1998 con l'obiettivo di implementare un servizio *Help Desk* (HD), di assistenza *on-line*, per agevolare l'attività degli operatori delle filiali.

In particolare, attraverso il supporto di un **motore inferenziale** (CBR – *Case Base Reasoning*) vengono costruite

Il **motore inferenziale** è un programma costituito da un interprete che decide quale regola applicare per poter aumentare la base di conoscenza e da uno schedatore che organizza le regole da sviluppare e il loro ordine di esecuzione. Il compito del motore è estrarre le regole utili alla soluzione del problema secondo un meccanismo di riconoscimento e attivazione delle stesse, eseguendo un esame delle regole nella base di conoscenza. Successivamente il motore seleziona la regola più appropriata, la esegue e registra nella memoria di lavoro l'azione ripetendo tutte queste operazioni finché non trova la possibile soluzione.

basi di casi, ovvero strutture di domande e risposte preimpostate attraverso le quali l'utente viene guidato sino alla soluzione finale. Le problematiche inserite sono quelle che, secondo l'esperto di help desk, si presentano con maggiore frequenza e presentano le maggiori difficoltà.

La struttura di una base di conoscenza viene rappresentata tramite un albero; anche se in realtà, i singoli

rami possono avere legami tra loro che permettono di trovare una soluzione anche saltando da un ramo all'altro.

A monte di tale lavoro l'esperto KM prende visione della materia trattata e individua i problemi.

L'attività che si richiede all'esperto HD esula dalla normale attività svolta: gli si domanda, infatti, di ripercorrere l'*iter* mentale che (inconsapevolmente) lo porta a individuare il problema del collega e di spiegarlo verbalmente. A fronte di una richiesta di assistenza, infatti, non sempre i dati forniti dall'operatore di filiale sono completi, sufficienti o corretti al punto tale da dare, all'operatore HD, un quadro chiaro della difficoltà incontrata dal collega e dell'effettiva problematica presentatasi. L'abilità dell'operatore di HD, consiste proprio nel saper formulare domande mirate a ottenere tutti i dati necessari alla comprensione e alla successiva risoluzione del problema.

Le domande formulate, dunque, vengono riprodotte dall'esperto KM nella **struttura ad albero** che permette la ramificazione degli argomenti e il loro

successivo sviluppo sino alla individuazione della soluzione finale. È lo stesso esperto HD (massimo esperto HD di quella materia) che inserisce i contenuti nei documenti: tale intervento rende la soluzione qualificata e certificata (dall'esperto stesso). Dopo aver costruito le basi dei dati, seguono la fase di

test e la valutazione dell'effettiva efficacia delle stesse, per poi rilasciarle e consentire a tutti gli operatori di help desk di accedervi. Da questo momento è importante l'attività di aggiornamento del prodotto da parte degli operatori e dall'esperto di KM.

Nel corso dei quattro anni trascorsi, gli esperti KM di Callnet (poi Amì, ora Bizmatica) hanno costruito circa 30 basi di casi uti-

Si dice albero una struttura costituita da uno o più nodi e da segmenti che li uniscono. I segmenti possono unire solo nodi posti a livelli diversi; il nodo più in alto si dice padre, quello più in basso figlio. La **struttura ad albero** è utilissima per rappresentare, in determinati problemi, i legami logici tra i dati dei problemi stessi. In particolare, gli alberi si prestano in maniera ottimale per descrivere tutti quei problemi di natura ricorsiva; esiste, infatti, un legame naturale tra la struttura ad albero e la ricorsività.



lizzate nei due help desk: Cariplo e Banco Ambroveneto.

Tali strumenti hanno, in sintesi, consentito:

- diminuzione dei tempi di risposta;
- aumento del livello di servizio;
- ottimizzazione delle risorse;
- riduzione dei costi.

Nel mese di ottobre 2001, si è attuato un progetto "pilota" su trenta filiali test (15 ex Cariplo e 15 ex BAV). Il prodotto di KM è stato messo a loro disposizione per la consultazione su 14 basi di casi. Anche in questo caso, sono stati raggiunti gli obiettivi già acquisiti in HD ed è stato riconosciuto il valore aggiunto di un prodotto di KM rispetto ai manuali in linea, presenti nelle filiali, alle circolari, consultabili on-line da parte degli operatori di HD o di Filiale o ai manuali cartacei.

In particolare, la soluzione adottata ha consentito di raggiungere i seguenti risultati:

- =>riduzione del 30% dei contatti pervenuti presso l'HD da parte delle filiali;
- =>risoluzione delle problematiche in modalità *self service*;
- =>maggiore disponibilità, in termini di tempo, degli operatori dell'HD per risolvere le problematiche più complesse;
- =>soddisfazione altissima e unanime espressa dalle filiali interessate.

A fronte di tali risultati, la banca ha deciso di mettere, quindi, il prodotto a disposizione degli operatori di filiale, i quali possono ora trovare individualmente (in modalità *self service*) la soluzione al loro problema evitando, così, di contattare l'HD.

Tale pratica, se effettuata con continuità, consente di ridurre le chiamate in entrata all'HD e permette agli operatori di filiale di contattarli solo per quelle problematiche più complesse che richiedono effettivamente il loro intervento.

5. CONCLUSIONI

Come visto, il Knowledge Management attraverso un processo di raccolta-strutturazione-elaborazione-utilizzo della risorsa conoscenza che va ad incidere sulla struttura organizzativa e tecnologica dell'azienda, permette di ottenere vantaggi significativi, tra i quali: migliore gestione del patrimonio informativo

aziendale, rapidità di risoluzione dei problemi, maggiore efficacia e rapidità dei processi decisionali, maggiore visibilità sulle nuove opportunità di business.

Il processo di innovazione generato dal KM tende, per sua natura, ad autoalimentarsi all'interno della rete aziendale e richiede continuamente uno sforzo organizzativo per completarsi. Solo in questo modo si ottengono i principali benefici, in termini di condivisione dei patrimoni conoscitivi e di attivazione delle competenze disponibili.

L'investimento in innovazione tecnologica costituisce, quindi, uno strumento fondamentale per l'acquisizione di competenze distintive e l'individuazione di sempre nuovi vantaggi competitivi con cui affrontare il proprio business e, come tutti gli investimenti, va valutato, gestito, monitorato, soprattutto in una fase come quella attuale.

Si sta, infatti, attraversando una fase di scoperta e applicazione delle nuove tecnologie quali la telefonia mobile ed Internet, in un momento di profonda incertezza e fino a quando non si saranno chiariti meglio i fini delle nuove tecnologie continueremo a soffrire di incertezze confusioni nelle scelte strategiche, con la conseguente maggiore volatilità dei mercati azionari del mondo.

È solo a questo punto che diventeranno chiare quali potranno essere, in senso più ampio, le conseguenze e le implicazioni delle nuove tecnologie nella vita economica, sociale e politica e solo allora si avrà un nuovo paradigma che darà maggiore stabilità.

Bibliografia

- [1] Auckland M: Competing through knowledge. *Knowledge Management Review*, February, 1999, p. 14-19.
- [2] Austin R: Putting change into context. *Knowledge Management Review*, February, 1999, p. 10-13.
- [3] Caldwell N, Clarkson J: Web-Based Knowledge Management for Distributed Design. *IEEE Intelligent Systems*, May - June 2000.
- [4] Cothrel J, Williams R: On-line communities. *Knowledge Management Review*, February, 1999.
- [5] Davenport TH, Prusak L: *Working knowledge: how organizations manage what they know*. Harvard College, USA, 1998.

- [6] Dietinger T, Gütl G, Maurer H, Schmaranz K: *Intelligent Knowledge Gathering and Management as New Ways of an Improved Learning Process*. Proceedings of AACE WebNet'98 Conference, Orlando, Florida USA. 7-12 November, 1998.
- [7] Drucker PF: Knowledge-worker productivity: The biggest challenge. *California Management Review*, Winter, 1999.
- [8] Grant JH, Gnyawali DR: Strategic process improvement through organizational learning. *Strategy and Leadership*, May-June, 1996.
- [9] Hansen MT, Nohria N, Tierney T: What's Your Strategy for Managing Knowledge?. *Harvard Business Review*, March-April, 1999.
- [10] Kaplan RS, Norton D: Strategic learning & the balanced scorecard. *Strategy and Leadership*, September-October, 1996, p. 18-24.
- [11] Klein M: *A Knowledge-Based Approach to Handling Exceptions in Workflow Systems*. Center for Coordination Science, Mit, Boston 1998.
- [12] Mayo A: Making human capital meaningful. *Knowledge Management Review*, February, 1999, p. 26-29.
- [13] Massa S, Merlino M, Puliafito PP: *Knowledge Management e vantaggio competitivo*. Sviluppo & Organizzazione, n. 173, Maggio-Giugno 1999.
- [14] Mesenzani M, Ottaiano C: *Il Knowledge Management in Andersen Consulting*. AICA Notizie, Anno XIII, n. 2, Giugno 2000.
- [15] Normann R: *Reframing Business: When the Map Changes the Landscape*. Chichester: John Wiley, 2001; edizione italiana, *Ridisegnare l'impresa: Quando la mappa cambia il paesaggio*. Milano: Etas, 2002.
- [16] Porter M: *The Competitive Advantage of Nations*. London: Macmillan, 1990.
- Porter M: *Strategia e competizione. Come creare, sostenere e difendere il vantaggio competitivo di imprese e nazioni*. Il Sole 24ore libri, 2001.
- [17] Von Krogh G: *Enabling Knowledge Creation*. Oxford University Press, 2000.

PAOLO BORDIN Laureato in Economia presso l'Università Cattolica del Sacro Cuore di Milano. Ha maturato 7 anni di esperienza nel campo dell'Information Technology. Prima di Bizmatica ha collaborato con Data Management e con Oracle. In entrambe le aziende ha svolto attività di prevendita di applicazioni e progetti ERP; tra i più importanti: il Gruppo Fiat, Gruppo Unicredit, SAI, Gruppo ENI, B-Ticino, Cartiere Burgo.
e-mail: Paolo.Bordin@Bizmatica.com

ALESSANDRA SELVA Laureata in Economia presso l'Università Statale di Milano-Bicocca e dopo un'esperienza di dottorato in economia e strategia aziendale, ha maturato 4 anni di esperienza nel campo dell'Information Technology. Attualmente consulente aziendale e di organizzazione in Bizmatica. Ha partecipato a diversi progetti, tra i più importanti: Uvet spa; Transmec Servizi, Elitel e Banca Intesa.
e-mail: alessandra.selva@bizmatica.com

EXTREME PROGRAMMING: PROGRAMMAZIONE ESTREMA O REVISIONISMO ESTREMISTA?

L'XP, nel processo di sviluppo del software, ha l'obiettivo di evitare la produzione di semilavorati diversi da quelli necessari alla realizzazione delle applicazioni. Presentata come un approccio del tutto nuovo e originale, ha in realtà le sue radici in metodi già noti e sperimentati. Secondo chi scrive, inoltre, pur introducendo utili concetti e offrendo un approccio valido in alcuni specifici contesti, l'XP è troppo spesso proposta come una soluzione universale ai problemi della moderna produzione industriale di software¹.

1. INTRODUZIONE

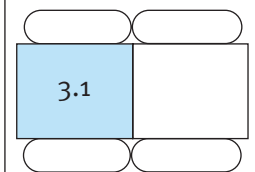
L'ingegnere del software è, periodicamente, bombardato dagli annunci di approcci rivoluzionari che cambieranno definitivamente il suo modo di lavorare, dandogli, finalmente, la possibilità di conseguire quegli obiettivi di produttività e qualità che ha sempre sognato. Le cose stanno, in realtà, in maniera differente, ma non è sempre facile distinguere quanto effettivamente ci sia di innovativo e originale perché la rivoluzione è, innanzitutto, terminologica e spesso si usano parole nuove per ribadire concetti già noti. In un famoso articolo del 1987 [7], F. Brooks analizzò, a fondo, quali fossero le difficoltà intrinseche e ineliminabili del software (*essential difficulties*), argomentando, in modo molto efficace e convincente, che il progresso, fino ad allora, aveva solo aggredito gli aspetti più facili (*accident*). Le conclu-

sioni che venivano tratte erano che non esisteva una panacea capace di rendere facile lo sviluppo del software. Secondo l'efficace metafora di Brooks, non esiste il rassicurante *silver bullet* capace di eliminare "il lupo mannaro delle favole dei bambini". Purtroppo, il mito del silver bullet ha continuato ad affacciarsi sulla scena della tecnologia software, manifestandosi attraverso la sopravvalutazione o la generalizzazione di metodi, tecniche e strumenti che si dimostravano efficaci in certi contesti, ma che venivano, ingiustificatamente, proposti come soluzioni generali definitive. L'entusiasmo dei ricercatori e, ancor più spesso, la propaganda commerciale, hanno in molti casi generato false aspettative di soluzioni miracolistiche che, alla prova dei fatti, hanno subito mostrato i propri limiti. Uno dei settori in cui più frequentemente si è assistito a questo fenomeno è quello dei metodi e strumenti di supporto al processo di produzione del software. In tale ambito, per l'appunto, si colloca l'*eXtreme Programming* (XP). Il motivo dell'interesse per questo tema è evidente, data la sua rilevanza econo-

¹ Questo lavoro è stato svolto in parte nell'ambito del progetto *Sahara* cofinanziato dal MIUR (Ministero dell'Istruzione dell'Università e della Ricerca).



Carlo Ghezzi
Mattia Monga



mica. L'obiettivo, infatti, è quello di definire la struttura del processo di sviluppo in modo da renderlo, il più possibile, efficace e controllabile. Nei metodi tradizionali, la controllabilità è ottenuta dettagliando fasi e attività attraverso cui il processo deve evolvere, definendo le regole di transizione da una fase alla successiva e documentando ognuna di esse secondo procedure standardizzate. Le persone coinvolte nello sviluppo, perciò, sono responsabili della produzione non solo del codice sorgente dell'applicazione, ma anche di molti altri semilavorati, che non sono necessariamente rivolti alla pura generazione del codice. Ogni applicazione, però, pone problemi diversi, ed è difficile e irrealistico (se non addirittura impossibile) definire un processo universalmente valido, senza cadere nei generici richiami al buon senso. Quando si cerca di adottare una qualche forma di processo senza progettarne l'adattamento alle specifiche esigenze del caso, il lavoro aggiuntivo finisce per essere percepito come burocrazia inutile, che influisce in maniera negativa sulla flessibilità e l'efficienza. L'XP promette di mantenere la controllabilità del processo pur riducendo il lavoro di supporto e convogliando il massimo dello sforzo sulla mera produzione dell'applicazione.

Un altro problema ricorrente nella produzione del software è che spesso i risultati ottenuti da anni di lavoro finiscono per soddisfare solo le esigenze degli sviluppatori, anziché quelle dei committenti.

L'XP promette di fornire i meccanismi perché gli sviluppatori possano acquisire, *durante lo sviluppo*, la consapevolezza che ciò che stanno costruendo soddisferà pienamente, al momento della sua posa in opera, le esigenze di chi l'ha commissionato.

Quanto detto finora costituisce una premessa teorica al presente articolo che sarà suddiviso in tre parti. Nella prima (Paragrafo 2) si fornirà una breve presentazione delle principali caratteristiche dell'XP. Nella seconda (Paragrafo 3) si cercherà di collocare il metodo nel contesto dei metodi di definizione dei processi software. Con ciò si intende mostrare come l'XP, che spesso viene presentata come un approccio del tutto nuovo e originale, in realtà abbia le sue radici in metodi noti da e sperimentati da mol-

ti anni. Infine, nella terza (Paragrafo 4) si fornirà una valutazione di XP, cercando di spiegare perché, secondo chi scrive, essa pur introducendo utili concetti e ponendo reali problemi, possa essere complessivamente valutata come uno dei tanti silver bullet falliti. Alcune lezioni che essa insegna devono senz'altro essere tenute in seria considerazione dagli ingegneri del software ma, altrettanto chiaro, deve risultare il fatto che non si tratta della soluzione definitiva ai problemi dello sviluppo delle moderne applicazioni software di qualità, come tanta letteratura sull'argomento vorrebbe far credere. In altri termini, se è vero che alcune delle raccomandazioni dell'XP si inquadrano correttamente nel solco delle proposte di processi di produzione di software agili e flessibili e che in quanto tali, forniscono un bagaglio concettuale utile al progettista di oggi, nella visione radicale che realizza la sua formulazione data dai propri paladini costituisce solo un revisionismo estremistico dei metodi tradizionali, che rischia di produrre un danno anche maggiore di quello che vorrebbe eliminare.

2. XP: UN TUTORIAL

Con eXtreme Programming si intende una tecnica, proposta da K. Beck, per organizzare il processo di sviluppo del software con l'esplicito obiettivo di evitare la produzione di semilavorati diversi da quelli strettamente necessari alla realizzazione dell'applicazione. Dettagliate specifiche formali, approfondite analisi e puntigliosa documentazione sono considerate attività troppo costose rispetto ai benefici apportati, in quanto limitano la flessibilità del processo che deve poter modificare i propri obiettivi in ogni momento. La produzione di un'applicazione, secondo Beck, non è un'attività che possa essere analizzata e precisamente pianificata a priori. Invece, esattamente come quando si guida l'automobile, la condotta complessiva è il risultato di un gran numero di minimi cambiamenti di rotta che il pilota decide in base alla sua istantanea percezione di curve ed ostacoli. Il lavoro dei programmatori procede organizzando quattro attività fondamentali che si ripetono per tutto il corso del progetto:



- | scrittura del codice dell'applicazione (*coding*);
- | verifica delle funzionalità (*testing*);
- | osservazione dell'ambiente, inteso come desideri del committente, opportunità tecnologiche, sviluppi di mercato (*listening*);
- | progetto dell'applicazione (*design*).

I programmatori sono responsabili non solo della codifica dell'applicazione, ma anche della sua verifica. Anzi, particolare enfasi è posta proprio su questa attività: nessuna istruzione dovrebbe essere considerata veramente parte dell'applicazione finché non ne sia stato verificato l'effetto secondo le attese. La costruzione dell'applicazione procede in maniera iterativa, cogliendo le reazioni dei committenti e riprogettando, in maniera adeguata, le sue funzionalità e i meccanismi adottati per ottenerle.

Perché queste attività possano essere svolte efficacemente, vengono identificate alcune prassi fondamentali che aiutino i programmatori a rendere il loro lavoro il più efficiente possibile.

Pianificazione delle attività (*Planning the game*). Lo sviluppo dell'applicazione è accompagnato dalla stesura di un piano di lavoro. Il piano è definito e, continuamente aggiornato, a intervalli brevi e regolari dai responsabili del progetto, secondo le priorità aziendali e le stime dei programmatori, che partecipano, in modo attivo, alla pianificazione. Il meccanismo per attuare, efficacemente, la pianificazione è articolato come un gioco in cui sono coinvolti utenti responsabili del progetto e sviluppatori per stabilire un equilibrio dinamico fra le esigenze di tutti gli attori coinvolti. Gli utenti finali dell'applicazione presentano gli obiettivi da raggiungere descrivendo una serie di scenari (storie) che il sistema deve soddisfare. Gli sviluppatori stimano il tempo necessario per la realizzazione di ogni storia: qualora ciò non sia possibile, la storia viene suddivisa in storie più semplici. Le storie vengono ordinate da utenti e responsabili secondo la loro priorità di realizzazione, dopo che gli sviluppatori ne hanno stimata la rispettiva difficoltà. Dalla sintesi di queste valutazioni i responsabili del progetto generano la pianificazione delle attività, intesa come l'insieme di storie che dovranno essere realiz-

zate per il prossimo rilascio e le date previste: sarà, inoltre, loro principale responsabilità misurare e controllare l'andamento delle attività rispetto alla pianificazione stessa. Questo processo viene ripetuto dopo ogni rilascio per pianificare il successivo. Una pianificazione più dettagliata viene poi decisa dagli stessi sviluppatori, i quali definiscono i "compiti" elementari necessari all'implementazione delle singole storie. Per ogni compito uno sviluppatore deve stimare i giorni necessari al suo completamento e assumersi la responsabilità della sua realizzazione. Il carico di ciascuno sviluppatore non deve superare quello concesso dalla pianificazione del rilascio, dopo essere stato pesato secondo un fattore di efficienza di gruppo che tiene conto delle necessità di cooperazione tramite riunioni, incontri con il committente ecc..

Rilasci frequenti (*Short releases*). La vita e lo sviluppo dell'applicazione sono scanditi dai rilasci di versioni del prodotto funzionanti, nel senso che realizzano qualcuna delle storie che ne descrivono gli obiettivi. Ogni rilascio rappresenta il punto conclusivo di un'iterazione di sviluppo e l'inizio di una nuova pianificazione. Per poter tener conto di cambi di prospettiva, errori di valutazione, nuovi requisiti, restrizioni di bilancio, ogni iterazione dovrebbe durare non più di qualche settimana (in genere, da due a quattro).

Metafora condivisa (*Metaphor*). Ogni progetto è guidato da una metafora condivisa da responsabili e sviluppatori. La metafora non è altro che una descrizione semplificata, ma efficace, del sistema nel suo complesso. Serve a fornire un vocabolario comune a tutte le persone coinvolte, senza scendere nei dettagli implementativi.

Progetti semplici (*Simple design*). La struttura dell'applicazione deve essere la più semplice possibile. L'architettura del sistema deve essere comprensibile da tutte le persone coinvolte nel progetto. Non devono esserci parti superflue o duplicazioni. Le parti che compongono il sistema devono essere, soltanto, quelle strettamente necessarie alle esigenze correnti. Solo quando nuove circostanze lo richiederanno, verranno progettati nuovi componenti, eventualmente riprogettando anche quelli già esistenti.

Ristrutturazione del codice (*Refactoring*).

Come si è detto precedentemente, l'applicazione necessita di continue riprogettazioni per eliminare parti divenute superflue e per adattare il sistema alle nuove esigenze. Questa attività è detta *refactoring*, ristrutturazione, a intendere che, ogni volta che si intravede la possibilità di eliminare parti superflue o di semplificarne l'organizzazione, l'intera struttura del codice va adattata ai nuovi principi progettuali.

Verifica di ogni funzionalità (*Testing*). Ogni funzionalità va sottoposta a verifica, in modo che si possa acquisire una ragionevole certezza sulla sua correttezza. Ciò sia a livello di sistema (test di sistema) sia a livello del singolo metodo (test di unità). I test di sistema sono costruiti sulla base delle storie concordate con il committente che dice l'ultima parola sulla convalida del sistema. I test di unità devono poter essere rieseguiti automaticamente, con tempi dell'ordine dei minuti: allo scopo è utile avvalersi di strumenti opportuni [10]. Ogni ristrutturazione o modifica del codice deve mantenere inalterato il risultato dei test già considerati. I test vengono, generalmente, scritti prima della codifica della funzionalità. Il metodo prescrive, addirittura, che il codice da sviluppare debba soddisfare tutti i test e niente di più. Secondo Beck, la verifica è sostanzialmente una falsificazione di stampo "popperiano"²: una teoria (il codice dell'applicazione) viene messa alla prova grazie ad una serie di osservazioni (casi di test) che essa deve spiegare (ovvero, non deve disattendere le aspettative). Ogni nuova teoria deve rimanere coerente con le osservazioni descritte precedentemente.

Programmazione a coppie (*Pair programming*). La scrittura vera e propria del codice è fatta da coppie di programmatori che lavorano al medesimo terminale. Le coppie non sono fisse, ma si compongono associando le

migliori competenze per la risoluzione di uno specifico problema. Il lavoro in coppia permette, scambiandosi periodicamente i ruoli, di mantenere mediamente più alto il livello d'attenzione. I locali dove si svolge il lavoro devono permettere senza difficoltà di lavorare a coppie.

Collettivizzazione del codice (*Collective ownership*). Il codice dell'applicazione può essere liberamente manipolato da qualsiasi sviluppatore. Ciò è possibile grazie al fatto che l'organizzazione è la più semplice possibile e che il codice è scritto rispettando regole condivise da tutti. La collettivizzazione è anche un meccanismo per stimolare la semplificazione delle parti più oscure del codice, che, essendo incomprensibili a tutti, fuorché agli autori, hanno un'alta probabilità di essere eliminate. Naturalmente ogni modifica non deve pregiudicare la correttezza dei test.

Integrazione continua (*Continuous integration*). Non sono previste sessioni particolari di integrazione. In effetti, dato che il codice dell'intera applicazione è sotto il controllo di tutto il gruppo di sviluppo, l'integrazione del lavoro dei singoli è continua. Deve essere costantemente possibile ottenere una versione funzionante dell'applicazione sulla quale operare le verifiche. Una piattaforma pronta per l'integrazione è sempre disponibile per i programmatori.

Rinuncia al lavoro straordinario (*40-hour week*). Lo sviluppo di applicazioni con i metodi dell'eXtreme Programming è un'attività che richiede grande concentrazione, entusiasmo e creatività. Il lavoro in condizioni di stress, con ripetute sessioni straordinarie, provoca necessariamente un deterioramento della qualità dell'impegno e deve pertanto essere evitato. Per questo motivo, la pianificazione coinvolge anche gli sviluppatori ed è aggiornata di frequente per tener conto di errori e imprevisti.

Partecipazione del committente (*Onsite customer*). Il committente deve essere coinvolto nello sviluppo perché è l'unica fondamentale fonte di convalida del sistema. Partecipa perciò alla stesura dei test di sistema e verifica, periodicamente, che il sistema realizzato corrisponda effettivamente alle proprie esigenze. Inoltre, è la principale fonte di informazione per la conoscenza sul dominio di applicazione. Secondo chi propone il meto-

² Il filosofo di origine austriaca Karl Popper (1902-1994) sostenne che un'ipotesi scientifica non può mai essere verificata da risultati sperimentali, che possono tutt'al più semplicemente corroborare la nostra fiducia in essa. Viceversa, un risultato sperimentale contrario alle aspettative (falsificazione) permette di sviluppare nuova conoscenza, obbligando al rifiuto dell'ipotesi di partenza.



do, la comunicazione orale tra progettisti e committenti, unita alla definizione a priori dei test, allevia la necessità di produrre una formulazione scritta e precisa dei requisiti.

Uso di standard per la codifica (Coding standards). Il codice deve esplicitare le astrazioni dei programmatori ed è il loro principale strumento di comunicazione. Deve, pertanto, essere scritto in maniera uniforme e omogenea. Tutti gli sviluppatori devono essere in grado di capire e modificare ogni linea di codice scritta da altri.

3. IL CONTESTO DELLE METODOLOGIE DI PROCESSO

Dalla fine degli anni '60 in poi, con la nascita dell'ingegneria del software come disciplina, ci si è posti il problema di definire modelli del ciclo di vita del software, dalla concezione iniziale dell'idea del prodotto fino al suo sviluppo, al suo rilascio e, infine, alla sua dismissione. Tutti i prodotti industriali di cui si vogliono assicurare adeguati livelli di qualità vengono, infatti, sviluppati secondo processi sistematici ben definiti. L'esistenza di un processo di produzione ben definito viene, addirittura, considerato come un prerequisito necessario perché il prodotto industriale possa aspirare a raggiungere un livello di qualità certificabile. Il primo e più noto modello di processo è noto come *ciclo di vita a cascata*. Il ciclo di vita a cascata, nato come esperienza all'interno di sistemi militari sviluppati fin dagli anni '50, è caratterizzato da una progressione lineare attraverso fasi (per l'appunto, *in cascata*), quasi si trattasse di una catena di montaggio attraverso cui far procedere le attività di sviluppo del software.

Nel seguito, vengono presentate le caratteristiche fondamentali di un'organizzazione a cascata del ciclo di vita.

■ **La sequenzialità del processo.** Secondo i proponenti, ciò dovrebbe garantire una migliore controllabilità e prevedibilità del processo e, quindi, la capacità di tenere i costi e i tempi di sviluppo sotto controllo. I ritorni all'indietro, infatti, vengono considerati dannosi ai fini di tenere il processo sotto controllo.

■ **La definizione di fasi e attività standard.** Per ciascuna fase, viene definito, con esattezza,

il criterio di successo che permette di certificare il completamento e, quindi, l'autorizzazione a procedere alla fase successiva. Per di più, spesso vengono prescritti metodi specifici che i progettisti sono tenuti a seguire per lo svolgimento delle attività della fase.

■ **Gli standard documentali.** Si afferma, spesso, che un ciclo di vita a cascata è *document driven*. Con ciò si intende che i semilavorati di ciascuna fase, che costituiscono *input* per la fase successiva, sono documenti che devono seguire certi standard che l'organizzazione produttiva prescrive. Il completamento di ciascuna fase, pertanto, viene a identificarsi con l'approvazione, da parte dei responsabili del progetto, dei documenti che rispettano gli standard prescritti. Il ciclo di vita a cascata, posto ancor oggi, in molti casi, come modello generale di riferimento per lo sviluppo di software, è in realtà un modello irrealistico e, spesso, del tutto inadeguato. L'organizzazione lineare del processo non può quasi mai avverarsi in pratica ma, anzi, il risultato dello svolgimento di certe fasi comporta, spesso, la ripetizione di quanto è stato fatto in fasi precedenti. Ciò è causato, principalmente, dalle difficoltà connesse con l'acquisizione e specifica dei requisiti. È, in generale, illusorio pretendere di congelare i requisiti in una specifica esaustiva che viene prodotta all'inizio del processo, in base alla quale si procede poi al progetto dell'architettura e all'implementazione dell'applicazione. I requisiti sono spesso, inizialmente, confusi e, dunque, non sono noti in maniera precisa e, qualora lo siano, è facile che vengano fraintesi dall'ingegnere del software, che potrebbe non avere alcuna competenza circa il dominio applicativo nel quale l'applicazione dovrà operare.

Con un modello a cascata, errori o fraintendimenti nella specifica dei requisiti verrebbero scoperti molto tardi, quando l'applicazione viene rilasciata al committente o immessa sul mercato. Il risultato è dunque che, immediatamente, inizia una fase di "manutenzione", che in realtà tende solo a rimediare gli errori introdotti all'inizio e, tornando indietro, a rielaborare l'analisi e la specifica dei requisiti. È anche stato osservato che molto spesso i requisiti si chiariscono solo dopo che una qualche forma dell'applicazione viene posta nelle mani del committente e che,

quindi, risulta illusorio ipotizzare che il processo di sviluppo possa procedere in modo lineare, basato su conoscenze che possono solo in parte essere note *a priori*.

In conclusione, il ciclo di vita a cascata nasce con l'obiettivo primario di ridurre il rischio che i costi e i tempi di sviluppo non siano controllabili. Nel cercare di limitare questo rischio, vengono ignorati i rischi ancora più gravi che sono dovuti alla scarsa conoscenza e all'instabilità dei requisiti, che fanno sì che la riduzione dei costi di sviluppo sia solo illusoria, e che si riversi, da un lato, in elevati costi di manutenzione e, dall'altro, in insoddisfazione dei committenti. Un modo più astratto di caratterizzare un ciclo di vita a cascata definisce il processo come una scatola nera, che riceve, in ingresso, la specifica dei requisiti dell'applicazione e che, in uscita, fornisce l'applicazione finale, ottenuta attraverso la sequenza lineare di fasi. Poiché il processo di acquisizione e specifica dei requisiti soffre dei problemi intrinseci che sono stati riassunti in precedenza, qualunque sia la scomposizione in fasi che viene scelta per articolare le attività della scatola nera, si arriverà sempre, e troppo tardi, a scoprire che occorre procedere a modifiche e ricalibrature, ripercorrendo in tutto o in parte il processo di sviluppo.

È anche possibile definire modelli flessibili, agili e incrementali in cui si abbandona l'idea dello sviluppo dell'applicazione intesa come un prodotto monolitico, con un processo teso a un'unica data finale di consegna. Al contrario, il processo di sviluppo viene visto come una successione di rilasci incrementali, che adatta lo sviluppo in maniera flessibile ai requisiti man mano che questi vengono esplicitati.

Viene così abbandonata l'idea che debba esistere un unico modello di riferimento da adottare come standard immutabile e universale, ma viene invece accettato che il modello di processo debba essere di volta in volta definito in base alle specifiche caratteristiche dell'applicazione.

Pur senza entrare in dettagli che esulano dallo scopo di questo lavoro (e per una trattazione dei quali si rimanda a un testo di Ghezzi et al. [8]), si ricorda che i modelli alternativi proposti di volta in volta hanno preso il nome di modelli basati su *rapid prototyping*, di mo-

delli *user driven*, di modelli "a spirale" (per contrasto con la linearità del ciclo a cascata [1]). Esempi ben noti, e diversi tra loro, di organizzazione innovativa del ciclo di vita del software sono offerti dal cosiddetto *Unified Development Process*, sviluppato nel contesto di UML (*Unified Modeling Language*) [12], dal metodo agile e flessibile adottato da Microsoft, descritto da Cusumano e Selby [6] ed efficacemente definito come *synchronize & stabilize* e, infine, dai processi di tipo "bazaar" seguiti talvolta per lo sviluppo di software *open-source* [17]. È stato anche osservato [7] che nell'era di Internet sempre più spesso chi sviluppa software innovativo può seguire la strategia di sviluppare versioni iniziali rendendole, in qualche modo, disponibili, anche in forma gratuita, attraverso Internet. Ciò può incoraggiare molti potenziali utenti, da un lato, a sperimentare l'uso dell'applicazione e, dall'altro, a fornire utili suggerimenti per sue modifiche che potrebbero essere, successivamente, incorporate nel prodotto.

In questo modo, si generano comunità virtuali di *early adopter* che possono essere fidelizzati al prodotto e che lo adotteranno quando questo verrà poi messo sul mercato, o che commissioneranno servizi aggiuntivi, stabilendo così un forte legame anche di natura economica con il produttore di software. La proposta dell'XP, si colloca, dunque, in un contesto generale che, da tempo, ha riconosciuto l'improponibilità sia di metodi generali e universalmente adottabili per lo sviluppo delle applicazioni, sia l'inadeguatezza, se non in casi molto specifici, di metodi rigidamente sequenziali, quali il ciclo di vita a cascata. In particolare, si colloca all'interno dei metodi agili e flessibili, che vogliono rispondere in maniera efficace sia all'instabilità dei requisiti che all'esigenza di rapide risposte al mercato.

4. UN'ANALISI CRITICA DELL'XP

In che senso la produzione di software, secondo la modalità proposta da Beck, deve essere considerata estrema? Probabilmente, l'intenzione dei proponenti era quella di suggerire l'idea che la costruzione di applicazioni è un'attività che deve essere svolta in condizioni particolarmente rischiose, in

cui i progettisti e gli implementatori, a differenza di quanto accade in altre discipline ingegneristiche, devono essere in grado di reagire, prontamente, a ogni genere di eventualità impreviste: cambiamenti dei requisiti, stravolgimenti tecnologici e dell'ambiente in cui il sistema verrà utilizzato, *turn-over* dei lavoratori ecc..

Ecco, allora, che in un contesto simile il "meglio" diventa nemico del "bene" e l'unica cosa che conta è generare, prima possibile, un'applicazione utilizzabile. A parere di chi scrive, invece, il termine "estremo" caratterizza bene un approccio che non si limita a proporre una collezione di buone tecniche che l'ingegnere del software può scegliere di volta in volta, ma che fornisce una soluzione radicale in cui le singole tecniche sono integrate in un approccio estremista.

L'ipotesi, implicitamente accettata, sembra essere che l'analisi sia, in sé, un appesantimento del progetto. Il credo dei programmatori estremi è che "ad ogni giorno deve bastare la sua pena", ovvero non serve prevedere i possibili cambiamenti, perché la previsione è considerata troppo incerta per valere il suo costo. L'ingegneria del software tradizionale insegna che cambiare idea ha un costo che cresce esponenzialmente nel corso del progetto. La curva di figura 1, che si trova su molti testi classici di ingegneria del software [16], descrive, in maniera qualitativa³, quello che è normalmente considerato l'andamento dei costi delle varianti in corso d'opera. In sostanza, motiva la ragionevolezza del vecchio adagio ingegneristico (e non solo) secondo cui "prevenire è meglio che curare". Secondo Beck, invece, al giorno d'oggi, i progressi della tecnica nella produzione del software (soprattutto la diffusione di linguaggi orientati agli oggetti e la disponibilità di strumenti automatici di verifica e di refactoring) fanno sì che l'andamento dei costi sia meglio descritto da una curva come quella di figura 2, per cui è possibile rischiare una variante tardiva per risparmiare tempo prezioso nelle prime fasi dello sviluppo. Questa curva, peraltro, non appare giustificata da reali dati di natura

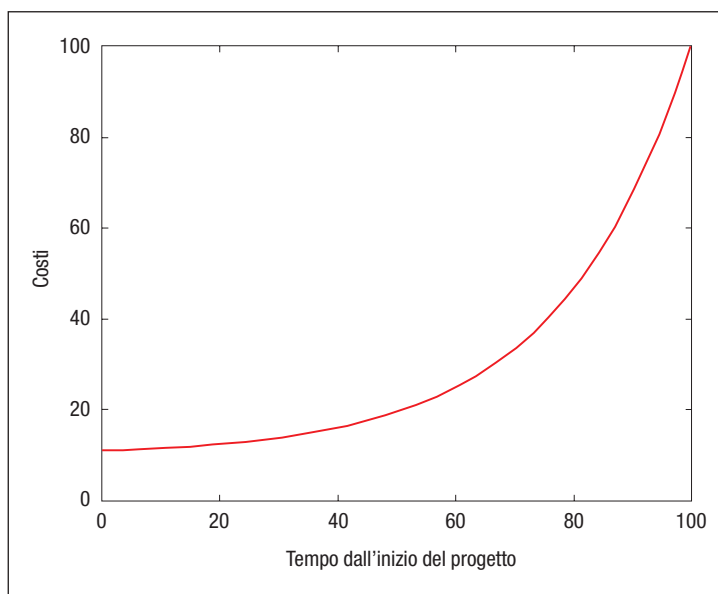


FIGURA 1

Relazione tradizionale fra costi delle modifiche e momento in cui vengono attuate

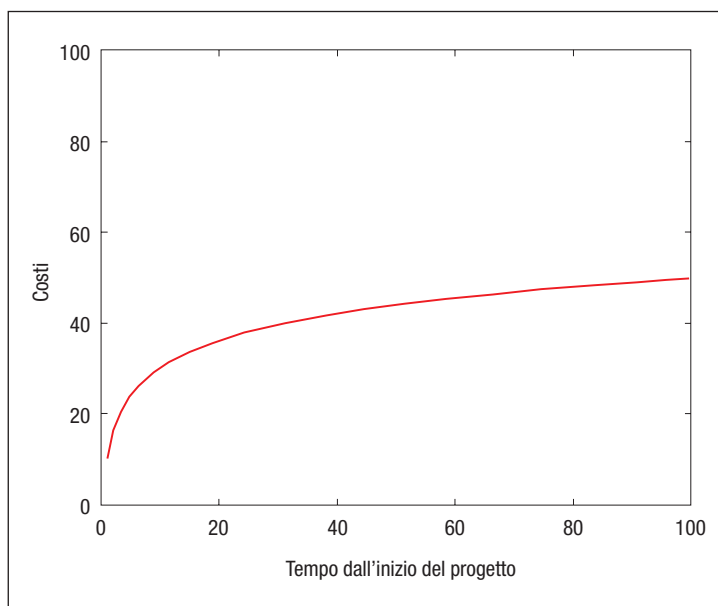


FIGURA 2

Relazione fra costi delle modifiche e momento in cui vengono attuate secondo XP

empirica, ma viene presentata come la conseguenza indiscutibile dei progressi tecnici e metodologici sopra ricordati, che per atto di fede assumono così il ruolo di nuovi silver bullet. Si può ammettere l'esistenza di classi di applicazioni per le quali l'approccio dell'XP può risultare vantaggioso: sistemi non critici di piccola-media dimensione, di tipo esplorativo (per esempio, nella ricerca e sviluppo) o, comunque, fortemente caratterizzati da requisiti poco definiti e instabili. Ma per tutti i sistemi di una certa complessità che vengono

³ Per una curva rappresentante dati reali si veda l'articolo di Bohem [2].

costruiti prevedendone una certa sopravvivenza nel tempo, l'approccio dell'XP deve essere stemperato in forme meno estreme, in particolare, attraverso un maggiore investimento nelle attività di acquisizione, analisi e specifica dei requisiti. Sarcasticamente, Beck vede in tali attività una pura perdita di tempo: una produzione di graziosi diagrammi di dubbia utilità che nessuno utilizza durante lo sviluppo dell'applicazione. Non c'è dubbio che, in molti casi, questa fase cruciale venga interpretata come il burocratico aderire a standard aziendali che prevedono la produzione di moduli cartacei e di diagrammi talvolta superflui. Forse perché la metodologia proposta risulta essere particolarmente difficile da applicare e facile ai fallimenti? Non sembra opportuno credere che Beck abbia scelto questo nome per insinuare alcunché di simile, né per sottolineare il suo approccio integralista nel rifiuto delle pratiche correnti dell'ingegneria del software. È anche vero che, in assenza di una focalizzazione mirata, esiste il rischio di concentrarsi su aspetti dell'applicazione che non hanno alcun interesse o scarsa priorità per il committente. Ma esiste un rischio anche più serio che lo sviluppo dei prodotti software degeneri in un processo senza fine del tipo *code & fix* [8]. La specifica dei test che devono essere superati da ciascuna nuova versione del prodotto è, di fatto, solo un criterio illusorio di accettazione. Per loro natura, i test definiscono solo un campione finito di possibili comportamenti del prodotto: non possono, dunque, esaurirne né la specifica né l'accettabilità. In sistemi dalle caratteristiche critiche e, a maggior ragione, in sistemi *safety critical*, inoltre, è del tutto improponibile che non si ponga l'accento sulla necessità di un'analisi a priori. La letteratura è ricchissima di esempi di malfunzionamenti che sono da ascrivere ad un'errata comprensione iniziale o a un'inefficace formulazione dei requisiti [19, 20]. Un'analisi fondata su metodi rigorosi o formali potrebbe, invece, ridurre i rischi o del tutto eliminare le cause dei malfunzionamenti. Ciò è coerente con quanto accade in settori ingegneristici dalla tradizione maggiormente consolidata, che hanno spesso reso obbligatorie alcune forme di analisi preliminari: si pensi ai calcoli del cemento armato o ad altre forme di verifica che devono

essere svolte, a priori, per fare convalide progettuali che devono necessariamente precedere la realizzazione. È, quindi, opportuna l'enfasi che l'XP pone nel ricordare che costi, tempi, qualità dei risultati e generalità del prodotto non sono variabili indipendenti l'una dall'altra, ma si influenzano a vicenda: progettisti e manager tendono a volte a dimenticarlo provocando l'esplosione dei costi. Il suggerimento di Beck è, in sintesi, quello di sacrificare la generalità, rinunciando ad anticipare aleatorie e imprevedibili evoluzioni future e, quindi, rinunciando a produrre soluzioni progettuali che facilitino tale evoluzione. Così facendo, però, Beck cade in contraddizione con ciò che Parnas ha insegnato attraverso il principio di *design for change* [17] e che generazioni di ingegneri del software hanno (con maggiore o minor successo) messo in pratica nell'ultimo ventennio. Lo sforzo di individuazione delle possibili evoluzioni future del sistema non deve certo diventare un esercizio sterile e del tutto teorico che rischia di prolungare i tempi di analisi senza reali riscontri pratici. Tuttavia, una costante attenzione alle probabili future evoluzioni costituisce uno dei principi basilari su cui si fonda l'ingegneria del software. Dunque non si può che convenire sull'opportunità che i progettisti approfondano molte energie e tutta la loro esperienza e sensibilità nell'identificazione di quelle parti del sistema suscettibili di maggiori modifiche, strutturando le applicazioni in modo tale che i cambiamenti più probabili siano anche i meno costosi, cosicché i loro sforzi siano ripagati nel tempo e gli investimenti ammortizzati.

Malgrado la proposta dell'XP, intesa come metodo generale unitario sia, da chi scrive, ritenuta poco convincente, si deve constatare che essa ha recentemente riscosso una certa popolarità.

Le ragioni sono duplici: da un lato, ragioni effimere che derivano dalla moda e dalla novità dell'approccio, dall'altra ragioni più profonde che derivano dall'aver riproposto in una forma nuova principi e metodi consolidati. Le ragioni del primo tipo sembrano dovute principalmente a due fattori collaterali, più che al loro intrinseco valore:

■ l'uso della metafora e dello slogan teorizzato come mezzo per comunicare e condividere gli

obiettivi; come insegna la realtà quotidiana, però, metafore e slogan, pur essendo formidabili metodi di aggregazione di massa, sono spesso delle grossolane semplificazioni che lasciano spazio a pericolose ambiguità, e trascurano aspetti fondamentali dei problemi;

l'uso di semplici ricette empiriche per gestire la complessità dello sviluppo; se è vero che ogni sforzo deve essere fatto per dominare l'intrinseca complessità, Brooks ha insegnato che non esistono facili scorciatoie. Lo sviluppo di software è una complessa attività di progettazione che si basa *in primis* sulle capacità individuali delle persone e sulla loro attitudine a cooperare nel lavoro di gruppo. Non esistono ricette generali prefissate che si possono dare a supporto di questa attività, ma solo una serie di principi, metodi, tecniche e strumenti che gli ingegneri del software, di volta in volta, devono essere in grado di aggregare in un processo atto a sviluppare lo specifico progetto sul quale sono impegnati. Purtroppo, invece, periodicamente, il fascino illusorio della semplificazione estrema dato dai ricettari standard riappare come un silver bullet nello scenario dell'ingegneria del software.

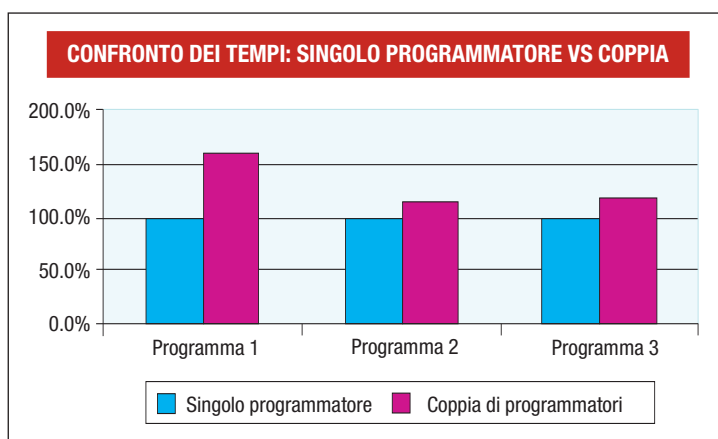
Si esaminano ora, invece, le ragioni positive della diffusione dell'XP. Innanzitutto, si osserva che, malgrado si sia cercato in passato di introdurre metodi sistematici di sviluppo, la produzione artigianale di software (code & fix) è un approccio ancora, estremamente, diffuso. Le metafore e le ricette (magari sostenute dall'uso di strumenti semplici e opportuni) proposte dall'XP possono essere un modo per insinuare alcune idee dell'ingegneria del software in ambienti altrimenti restii. Un po' come è successo recentemente con javadoc, lo strumento che Sun distribuisce con Java e che permette di ottenere documentazione direttamente dai commenti del codice: la *literate programming* non è, di per sé, un'idea rivoluzionaria, né una tecnica nata con Java, tuttavia, si deve riconoscere l'indubbia efficacia che ha avuto la diffusione di un simile strumento sulla qualità della documentazione del software prodotto in ambito accademico o di ricerca e sviluppo. Così, per esempio, l'organizzazione dello sviluppo basata sulle "storie" riflette la necessità ben nota di coinvolgere il committente nella vali-

dazione del sistema, fin dalle prime fasi della sua costruzione, e l'opportunità dell'uso di elementi tangibili per la pianificazione e il tracciamento dei progressi del lavoro. La collaborazione del committente allo sviluppo, quando è praticabile, è senz'altro auspicabile, anche perché riduce la conflittualità contrattuale. Molto spesso, però, quello che si vorrebbe davvero è il coinvolgimento degli utenti finali che potrebbero avere obiettivi diversi da quelli identificati dal committente e, soprattutto, diversi fra loro. Non convince, invece, l'idea che la mera partecipazione del committente all'interno del progetto riduca la necessità di un'attenta documentazione di specifica. Se si può convenire che tale specifica possa non essere necessaria sul momento, la sua necessità si giustifica quando occorre revisionare e far evolvere l'implementazione. In che modo è possibile risalire dal codice alle sue motivazioni? Come si può risalire dall'implementazione ai principi ispiratori delle scelte di progetto?

La tecnica del pair programming appare utile in molti casi. La programmazione è, infatti, tradizionalmente considerata un'attività solitaria praticata da persone introversive e scostanti anche se geniali. Tant'è che viene da chiedersi quanto questa mitologia del *real programmer* [18] abbia influito sulla cronica scarsità di presenza femminile fra gli sviluppatori. L'XP propone la programmazione a coppie che può, in effetti, essere un modo per migliorare la comunicazione di tecniche e obiettivi all'interno del gruppo di lavoro rendendolo più omogeneo e per introdurre ripetute ispezioni del software volte a migliorarne la qualità. Le figure 3, 4 e 5 mostrano i ri-

FIGURA 3

Confronto fra il tempo di realizzazione con e senza pair programming [4]



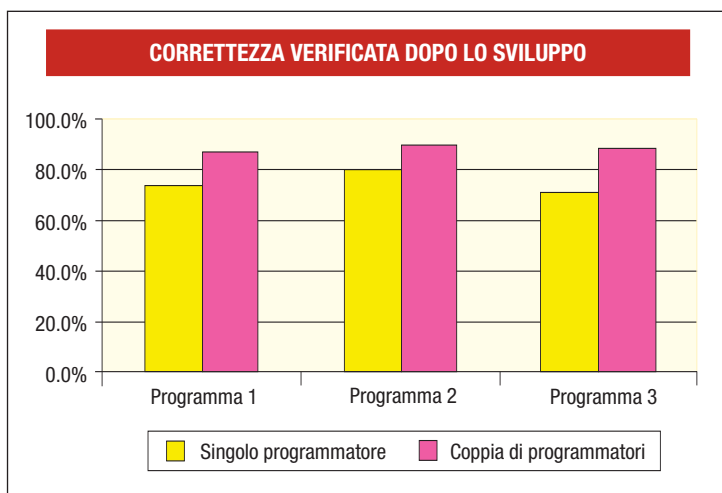


FIGURA 4

Confronto fra la correttezza con e senza pair programming [4]

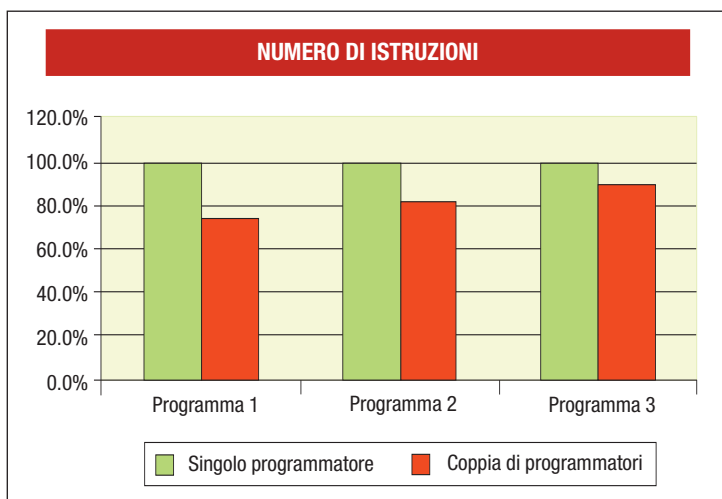


FIGURA 5

Confronto fra la compattezza del programma con e senza pair programming [4]

sultati di un esperimento [4] condotto per valutare gli effetti dell'introduzione della programmazione a coppie. La correttezza (Figura 4) e la compattezza (Figura 5) del codice aumentano, significativamente, a prezzo di una modesta diminuzione della produttività (Figura 3). Secondo J. Nosek [14], il tempo risparmiato assegnando a due programmatori il lavoro che potrebbe fare uno solo è in media il 29%. Altri studi [11] mostrano che la programmazione a coppie permette di far crescere velocemente ed efficacemente le capacità dei novizi e migliora la qualità del lavoro perché le persone coinvolte beneficiano dei rapporti interpersonali.

L'uso delle tecniche di pair programming comporta, dunque, un aumento del costo del personale, ma permette di essere più

prontamente sul mercato e nel caso di progetti in cui il *time to market* è un fattore critico, può effettivamente portare vantaggi economici [13]. Inoltre, accoppiando programmatori di diversa esperienza, essa risulta essere un utile strumento didattico, anche se abbastanza costoso per tutte le parti in causa.

Un'altra tecnica che ha raggiunto una certa diffusione è la scrittura di test di unità prima della codifica dell'unità stessa. La popolarità di tale tecnica la si deve soprattutto alla disponibilità di numerosi e pratici strumenti per la loro esecuzione automatica (il più noto dei quali è Junit [10]). Nella sua essenza essa può essere considerata una forma di programmazione per contratto, in cui i contratti non sono descritti in maniera dichiarativa, ma procedurale. Pur vantaggioso, l'approccio procedurale, comporta principalmente due problemi:

- una specifica dichiarativa (sostanzialmente *intensiva*) è normalmente più sintetica e descrittiva di una procedurale (*estensiva*); per di più, una specifica procedurale, data attraverso casi di test in numero finito, risulta necessariamente parziale;

- alcune proprietà sono difficili da descrivere in maniera procedurale; si pensi ad un programma Java che faccia uso di *thread*: per scrivere dei test che mettano in evidenza delle corse critiche bisognerebbe poter manipolare la macchina virtuale Java.

La tecnica di anticipare la scrittura dei test è, talvolta, ritenuta efficace per facilitare la stessa scrittura del codice dell'applicazione. Uno studio recente [12] mostra che, in generale, ciò non accade, ma coerentemente con l'interpretazione della programmazione per contratto, la presenza dei test facilita il riuso del codice in contesti nuovi o diversi. L'integrazione continua garantisce che esista, in ogni momento, un prototipo o una versione dell'applicazione funzionante. Questa è, in generale, una condizione assai desiderabile, come del resto da tempo mettono in evidenza molti autori⁴, perché per-

⁴ Il processo *synchronize & stabilize* adottato da Microsoft [6] si basa fondamentalmente su questo principio.



mette di convalidare continuamente ciò che è stato costruito. Non va dimenticato però, che affidando la convalida dei requisiti sempre e solo all'esecuzione dell'applicazione, la possibilità di esplorare soluzioni alternative è fortemente ridotta. L'integrazione continua e la proprietà condivisa di tutto il codice, poi, sembrano tecniche del tutto inapplicabili quando il numero di programmatori sale e il numero dei conflitti diventa ingestibile. Considerazioni simili valgono per le ristrutturazioni del codice. Al momento gli strumenti per operare queste ristrutturazioni in maniera semi-automatica sono ancora piuttosto rozzi. Pertanto, attualmente, sembra possibile ristrutturare in maniera consistente solo porzioni di codice di dimensioni relativamente modeste.

5. CONCLUSIONI

L'eXtreme Programming è nata, negli ultimi anni, come approccio radicalmente nuovo al processo di sviluppo del software. In questo articolo, dopo averne illustrato i tratti salienti, si è cercato di collocare XP all'interno del quadro complessivo dei metodi di supporto al processo. È stato inoltre evidenziato come XP abbia le proprie radici all'interno della classe di metodi agili, flessibili ed incrementali che da molti anni sono stati proposti, e largamente applicati, in alternativa ai processi a cascata che si dimostravano invece inadeguati nelle situazioni, per altro molto frequenti, caratterizzate da incertezza e variabilità nei requisiti o dalla necessità di risposta in tempi rapidi alle esigenze di un mercato in rapida e continua evoluzione.

Il giudizio che viene espresso è che, se le motivazioni alla base dell'approccio risultano largamente condivisibili, assai criticabile appare il tentativo di congelare la risposta in un metodo di processo che, a sua volta, si presenta come predefinito e universale. Mentre alcuni singoli suggerimenti e tecniche dell'XP pare che possano costituire utili strumenti nel bagaglio degli attrezzi di cui può disporre l'ingegnere del software, si è, invece, fondamentalmente critici riguardo alla loro aggregazione in un metodo unitario che venga presentato come la soluzione di tutti i problemi del processo software.

Questa idea, come si è detto, è fallita in passato ed è destinata a fallire in quanto non tiene conto delle specificità che ogni singolo progetto e ogni singola organizzazione caratterizzano [5].

Da ultimo, si vuole, invece, evidenziare quello che, secondo il parere degli autori del presente articolo, è il contributo più importante dell'XP al dibattito scientifico all'interno dell'ingegneria del software. In sintesi, XP ricorda con grande determinazione che il fine ultimo dell'ingegneria del software è produrre programmi: il codice, spesso considerato semplicemente come il risultato finale di un lungo ed elaborato processo, è il vero obiettivo e su di esso (che lo si voglia o no) finiscono con il concentrarsi gli sviluppatori.

Si deve riconoscere che, troppo spesso, l'ingegneria del software ha prodotto metodi inutilmente elaborati e onerosi e che, troppo spesso, gli standard aziendali si sono dimostrati eccessivamente meticolosi e burocratici. In molti casi, gli sviluppatori non sono riusciti a vedere vantaggi sensibili nella loro adozione, ma solo un intralcio nel procedere verso l'implementazione. Gli strumenti di *Computer Aided Software Engineering* (CASE) - ovvero, un altro silver bullet del passato - sono in larga misura falliti proprio perché nel codice prodotto mancava l'evidenza della sua relazione con le fasi di progetto e, quindi, alla fine, non portavano a un codice migliore, più facile da far evolvere, ma semplicemente sviluppato più velocemente.

La lezione da trarre per il futuro è, dunque, che i metodi e gli strumenti di supporto al processo dovranno consentire una rapida transizione al codice e dovranno fornire visibili benefici in termini di verificabilità e modificabilità del codice. La documentazione di analisi e di progetto non solo dovrà automaticamente essere legata al codice ed evolvere con esso, ma dovrà giocare un ruolo attivo, sia nel generare automaticamente parte dell'applicazione che nel generare automaticamente strumenti per la sua convalida.

Sembra opportuno, a questo punto, concludere sottolineando come le affermazioni di efficacia fatte dai paladini dell'XP, siano

quasi sempre argomentate su basi puramente ideologiche e perciò prive di supporto sperimentale.

Gli studi recenti apparsi sull'argomento [4, 12, 13] corroborano solo in parte le argomentazioni dell'XP. Molto lavoro di ricerca è, dunque, ancora necessario per valutare gli effetti delle singole tecniche proposte, in modo che l'ingegnere del software possa veramente considerarle nuove frecce a disposizione del suo arco.

Bibliografia

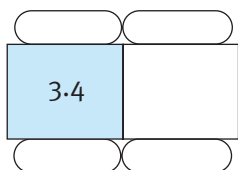
- [1] Boehm BW: A spiral model of software development and enhancement. *IEEE Computer*, Vol. 21, n. 5, 1988.
- [2] Bohem BW: Software engineering. *Transactions on Computers*, Dec.1976.
- [3] Brooks F: No silver bullet: Essence and accidents of software engineering. *IEEE Computer*, Vol. 20, n. 4, Apr. 1987.
- [4] Cockburn A, Williams L: *The costs and benefits of pair programming*. In *Extreme Programming Examined*. Addison Wesley, 2001.
- [5] Cugola G, Ghezzi C: *Software processes: a retrospective and a path to the future*. Software Process 14 Improvement and Practice 4, 3, Sept. 1998.
- [6] Cusumano MA, Selby RW: *Microsoft Secrets*. The Free Press, New York, NY, 1995.
- [7] Cusumano MA, Yoffie DB: Software development on Internet time. *IEEE Computer*, Vol. 32, n. 10, Oct. 1999.
- [8] Ghezzi C, Jazayeri M, Mandrioli D: *Fundamentals of Software Engineering, second ed.* Prentice Hall PTR, Apr. 2002.
- [9] Jacobson I, Rumbaugh J, Booch G: *The Unified Software Development Process*. Object Technology Series. Addison-Wesley, Reading/MA, 1999.
- [10] Junit: <http://www.junit.org>.
- [11] Lippert M, Roock S, Wolf H, Zullighoven H: JWAM and XP: *Using XP for framework development*. In *Extreme Programming Examined*. Addison Wesley, 2001.
- [12] Muller M M, Hagner O: *Experiment about test-first programming*. In *Conference on Empirical Assessment In Software Engineering EASE '02* (Keele, Apr. 2002).
- [13] Muller MM, Padberg F: *Extreme programming from an engineering economics viewpoint*. In *Fourth International Workshop on Economics-Driven Software Engineering Research (EDSER)*, Orlando, Florida, May 2002.
- [14] Nosek J: The case for collaborative programming. *Communications of the ACM*, Vol. 41, n. 3, Mar. 1998, p. 105-108.
- [15] Parnas DL: *Designing software for ease of extension and contraction*. In *Proceedings of the Third International Conference on Software Engineering*, 10-12 May 1978,
- [16] Pressman RS: *Principi di Ingegneria del software*. Second ed. Mc Graw Hill, 1997.
- [17] Raymond, ES: The Cathedral and the Bazaar. [Online.] Available: <http://www.ccil.org/~esr/writings/cathedral-paper.html>, Nov. 1997.
- [18] Real programmers:<http://www.cirr.com/~barkley/jokes/realprog.html>.
- [19] Neumann PG: *Computer Related Risks*. ACM Press 1995.
- [20] Jackson M: *Software Requirements and Specification: a lexicon of practice, principles, and prejudices*. Addison-Wesley, 1995.

CARLO GHEZZI è professore ordinario di Ingegneria del Software presso il Politecnico di Milano e responsabile scientifico dell'area di ricerca sull'ingegneria del software presso il CEFRIEL. Autore di numerosi articoli scientifici e libri, è editor in chief della rivista "ACM Transaction on Software Engineering and Methodology".
e-mail: ghezzi@elet.polimi.it

MATTIA MONGA ha conseguito il dottorato di ricerca in Ingegneria Informatica e Automatica presso il Politecnico di Milano nel 2001. I suoi interessi di ricerca riguardano l'ingegneria del software in ambito Internet e i linguaggi di programmazione orientati agli aspetti.
e-mail: monga@elet.polimi.it



Paolo Ceccherini
Gianni Grazia



LA RIVOLUZIONE NELLA RETE: ATM ED IP SONO DAVVERO NEMICI?

Le reti del futuro saranno basate su IP oppure su ATM e IP? Ovvero: in un mondo di servizi ed applicazioni sempre più orientate verso la multimedia e monopolizzato da IP, c'è ancora spazio per ATM? La rivoluzione in atto nelle telecomunicazioni, suggerisce di analizzare le origini delle due soluzioni nell'intento di prevederne l'evoluzione.

1. INTRODUZIONE

Appare oggi naturale parlare di linguaggi e protocolli come se da sempre esistessero in natura.

Nel secolo appena concluso si sono fatti grandi passi sia nella teoria dell'informazione [17] sia nella teoria della comunicazione [18]. Entrambe trovano ora nel *bit* e nelle sue manipolazioni la base per rappresentare ogni realtà e per trasmettere fedelmente a distanza tali rappresentazioni.

Il flusso di dati che ne consegue deve essere accompagnato da informazioni di controllo che ne garantiscano la fruibilità nel punto di destinazione: questo è il compito svolto dal **protocollo**.

Il mondo della comunicazione telefonica ha sem-

pre privilegiato la rapidità rispetto all'accuratezza. Questo perché il destinatario umano è in grado di comprendere anche una voce molto degradata, mentre il computer tollera più facilmente i ritardi che non gli errori e quindi nel mondo dati, si ricorre di preferenza alla ripetizione del messaggio anziché alla sua cancellazione quando è ricevuto con troppi errori.

Le due diverse esigenze: rapidità e accuratezza sono soddisfatte con compromessi diversi nei due protocolli oggi a confronto: ATM (*Asynchronous Transfer Model*), che viene dalla telefonia, privilegia la rapidità mentre IP (*Internet Protocol*) è invece teso a garantire l'accurato recapito di ciascun pacchetto anche a costo di ritardi indotti da ripetizioni e percorsi multipli.

Il **protocollo** nell'ambito delle telecomunicazioni, rappresenta l'insieme delle regole che devono essere applicate o realizzate affinché una trasmissione abbia luogo. I protocolli sono spesso associati ai livelli del modello OSI anche se, non necessariamente un livello deve essere rappresentato da un protocollo o viceversa. L'insieme dei protocolli che risultano coinvolti nel flusso di elaborazione delle informazioni dai livelli applicativi a quelli di rete e viceversa viene chiamata "Pila Protocolli" (Protocol Stack).

2. ATM ED IP: DUE PROTOCOLLI A CONFRONTO

ATM e IP rappresentano due tecnologie (o "filosofie", come qualcuno ama definirle) che, per anni, sono coesistite in modo indipendente l'una dall'altra. Sviluppatesi su seg-

menti di mercato e applicazioni completamente differenti, oggi stanno convergendo. Da più parti si avverte, tuttavia, il disagio provocato dall'incertezza sulla prossima evoluzione: "le reti del futuro saranno basate su IP oppure su ATM e IP"? In altre parole, in un mondo di servizi e applicazioni sempre più orientate verso la multimedialità e monopolizzato da IP, c'è ancora spazio per ATM?

Si cercherà, qui, di fare chiarezza sulle caratteristiche e peculiarità di queste tecnologie allo stato attuale e nella loro imminente evoluzione.

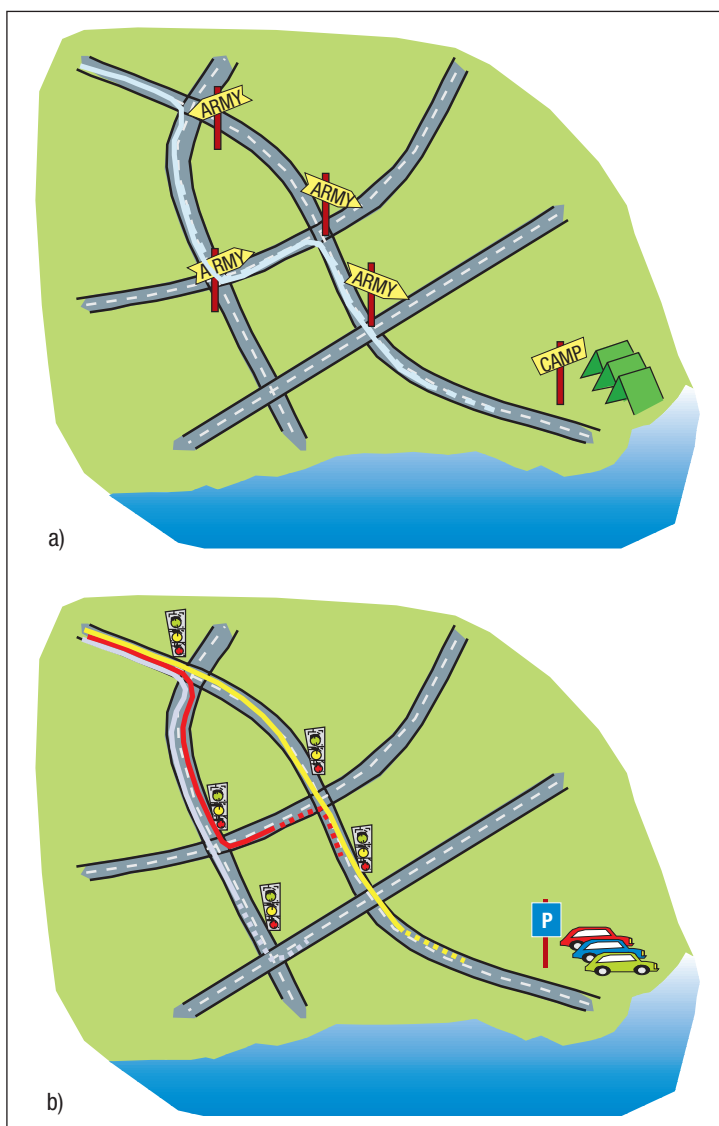
Oggi, tutti i sistemi di telecomunicazione operano sul principio della commutazione di pacchetto e possono essere classificati in sistemi *Connection-Oriented* oppure *Connection-less*.

Operare in modalità *connection-oriented* implica la necessità che i sistemi, prima di iniziare la trasmissione, realizzino una connessione logica fra trasmettitore e ricevitore. Durante questa fase, che viene definita di *setup*, sulla base del tipo di servizio e del tipo di informazioni che devono essere scambiate, nell'ambito dell'intera rete viene richiesto un determinato livello di qualità del servizio in funzione del quale viene definito "l'instradamento" per il percorso dei pacchetti. Al termine della trasmissione, la connessione viene "abbattuta" (*released*), in modo da liberare le risorse precedentemente allocate.

In contrapposizione con la precedente modalità operativa, esiste la cosiddetta *Connection-less*. Questa, in sostanza, non prevede alcuna fase di *setup*: le informazioni sono inviate, immediatamente, dal trasmettitore verso la rete la quale, non essendo "preparata" a riceverle potrebbe non accettarle, oppure, inoltrarle attraverso percorsi differenti all'interno della rete stessa dovendo, comunque, spendere del tempo per processare ogni singolo pacchetto in ogni nodo.

Per comprendere la sostanziale differenza tra le due modalità operative è di aiuto pensare a come un gruppo di amici e una Compagnia dell'esercito, organizzerebbero un viaggio verso una destinazione assegnata.

La Compagnia invierebbe una jeep di esplo-



ratori per tracciare la strada con un'apposita segnaletica a ogni incrocio (gli switch che attuano il *setup*) e solo dopo aver concluso questa fase, il convoglio partirebbe sul percorso prefissato preoccupandosi solo di seguire le indicazioni; in questo modo tutti i mezzi effettuano lo stesso tragitto e arrivano nello stesso ordine nel quale sono partiti (Figura 1 a).

Nel caso del gruppo di amici, stabilita la destinazione, ci si dà appuntamento all'arrivo: ognuno può procedere in modo autonomo scegliendosi il percorso che preferisce. Il risultato è che il gruppo di amici arriverà a destinazione in ordine sparso e bisognerà fare l'appello per ricomporre il gruppo (Figura 1 b).

FIGURA 1

Differenza tra un "percorso" *Connection-Oriented* a) e uno *Connection-less* b)

3. IL PROTOCOLLO ATM

Sviluppato in seno ad ITU-T (*International Telecommunication Union*), nell'ambito del progetto Broadband-ISDN (B-ISDN) [1], era stato, originariamente, concepito come una tecnologia per la trasmissione ad alta velocità su reti pubbliche di dati a supporto di applicazioni eterogenee. In seguito, ATM Forum (ATMF) ha esteso questa visione promuovendone l'impiego anche su reti private.

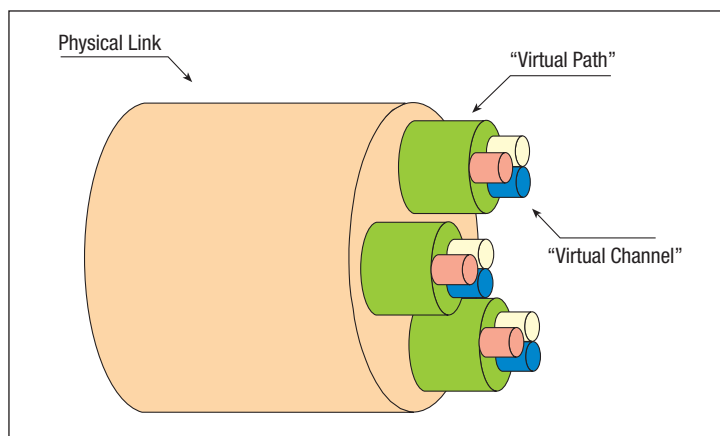
L'ATM, è un protocollo di comunicazione a commutazione di pacchetto di tipo Connection-Oriented.

I pacchetti o, più propriamente, le celle ATM hanno dimensione fissa e pari a 53 byte e sono composte da un *header* e da un *payload*. L'header è standard, mentre la parte di payload può contenere qualsiasi tipo di informazione, voce, video, testo, dati, grafica ed altro ancora. ATM offre un sistema omogeneo per il trasporto di tutte queste informazioni o servizi, in modo assolutamente trasparente alla rete. Come verrà esplicato in seguito, in funzione della natura delle informazioni che devono essere trasportate i dati vengono inseriti nel payload delle celle nella maniera più appropriata associandovi la necessaria qualità del servizio.

Una rete ATM è composta da *switch* e da *end-point* (EP): i primi provvedono al transito delle celle attraverso la rete, mentre i secondi smistano il traffico sulle terminazioni. In pratica, lo switch riceve le celle da un EP, ne legge e aggiorna l'header e, immediatamente, reindirizza la cella verso l'opportuna interfaccia di uscita.

FIGURA 2

Rappresentazione di un "Virtual Path" e di un "Virtual Channel" ATM



Su un collegamento ATM, è trasmesso un flusso di celle e, in qualsiasi momento, un altro utente può decidere di trasmettere qualsiasi tipo di informazione; questa filosofia spiega il sostantivo "asincrono" del nome. D'altro canto, questo significa che il ricevitore non è in grado, come avviene per i sistemi basati su TDM (*Time Division Multiplexing*), di distinguere le informazioni provenienti da un determinato utente semplicemente estraendole da una trama temporale (*timeslot*). Per questo nell'header delle celle sono memorizzati due campi denominati VPI (*Virtual Path Identifier*) e VCI (*Virtual Channel Identifier*).

Le "connessioni" possono essere gestite in modo statico o dinamico. Nel primo caso, vengono chiamate PVC (*Permanent Virtual Circuit*); nel secondo SVC (*Switched Virtual Circuit*). In entrambi, i casi ogni connessione (VP) contiene un certo numero di Virtual Channel (VC), ed è univocamente individuata dai valori VPI e VCI contenuti nell'header della cella (Figura 2).

Come detto precedentemente, l'ATM opera in modalità Connection-Oriented, il che implica che la trasmissione sia preceduta da una fase di setup per "tracciare" il percorso tra i punti da connettere. Durante il setup avviene la verifica sulla disponibilità delle risorse di rete e viene negoziata la qualità di servizio da garantire (per esempio, il CLR, *Cell Loss Ratio*).

Un'altra importante conseguenza di questa modalità operativa è il fatto che la sequenza temporale delle celle viene mantenuta; si limita così uno dei principali problemi della trasmissione a pacchetto di informazioni in tempo reale: il **jitter**. In sistemi *connection-less*, infatti, potendo i pacchetti percorrere differenti percorsi non è possibile garantire un ritardo costante durante la ricezione né, al limite, il relativo ordine.

Il **jitter** nei sistemi di trasmissione dati a commutazione di pacchetto rappresenta la varianza del tempo impiegato dai singoli pacchetti per giungere a destinazione. Se per applicazioni non *real time* (per esempio e-mail) l'effetto del jitter è trascurabile, questo risulta essere un parametro fondamentale nel caso di applicazioni *real time* (per esempio *Voice over IP*), dato che, al limite, può portare alla perdita (scarto) del pacchetto stesso.

Come detto la qualità del servizio viene specificata al momento della creazione del VC. Attualmente sono state definite le seguenti classi di servizio:

■ **Constant Bit Rate (CBR)**: classe di servizio adatta al trasporto di flussi di traffico isocroni a bit rate costante. Tale classe è caratterizzata da valori di CLR, CDV (*Cell Delay Variation*) e CTD (*Cell Transfer Delay*) molto stringenti e viene tipicamente utilizzata ai fini della emulazione di CDN (*Collegamenti Digitali Numerici*, esempio PCM $N \times 64$ kbit/s).

■ **Real Time - Variable Bit Rate (rt-VBR)**: classe di servizio adatta al trasporto di flussi di traffico isocroni a bit rate variabile. Tale classe è caratterizzata da valori di CDV e CTD molto stringenti, mentre il CLR non risulta particolarmente severo; viene di solito utilizzata per il trasporto del segnale video digitale compresso (MPEG, H.261 ecc.).

■ **Non-Real Time - Variable Bit Rate (nrt-VBR)**: classe di servizio adatta a flussi di traffico non isocroni a bit rate variabile che richiedano banda garantita; essa risulta particolarmente adatta all'interoperabilità tra reti ATM e reti Frame Relay in quanto il traffico offerto da queste ultime è un traffico dati (bit rate variabile non isocrono) ma necessita di banda garantita.

■ **Unspecified Bit Rate (UBR)**: classe di servizio adatta a flussi di traffico non isocroni a bit rate variabile che non richiedono alcuna garanzia su banda e latenza. Tale classe consente all'utente di trasmettere qualunque volume di traffico desideri, entro il limite consentito dalla PCR (*Peak Cell Rate*),

senza tuttavia assicurare alcunché circa CLR, CTD e CDV.

■ **Available Bit Rate (ABR)**: classe di servizio adatta a flussi di traffico non isocroni a bit rate variabile che non richiedono banda garantita. Tale classe è analoga alla UBR con la differenza che essa prevede un meccanismo di controllo di flusso per migliorare il *throughput* minimizzando la perdita di celle.

Le prime tre classi, CBR, rt-VBR e nrt-VBR, sono anche chiamate *Specified QoS Class* in quanto specificano caratteristiche peculiari del traffico; per motivi diametralmente opposti le ultime due sono anche denominate *Unspecified QoS Class*. La tabella 1 riassume quanto riportato precedentemente. Da notare che, nelle intenzioni degli autori, essa ha solo una funzione guida e non vuole assolutamente essere limitativa.

La natura connection-oriented, la relativa semplicità del protocollo (non è previsto né controllo d'errore, né di flusso) e la lunghezza fissa delle celle fanno sì che sia possibile commutarle in modo estremamente rapido (basti pensare, per esempio, che la capacità aggregata di elaborazione di un moderno Switch ATM è nell'ordine delle centinaia di Gbit/s). Occorre, tuttavia, tener presente che la procedura di *setup* è alquanto complessa e richiede una certa quantità di tempo ogni qual volta venga effettuata.

Senza voler scendere troppo nel dettaglio è utile descrivere il modello di riferimento dell'ATM [2]. Esso risulta composto da 3 strati: il primo, denominato *physical layer*, corrispon-

Aree di applicazione	CBR	rt-VBR	nrt-VBR	ABR	UBR
Dati critici	✓✓	✓	✓✓✓	✓	NA
Emulazione/Interconnessione LAN	✓	✓	✓✓	✓✓✓	✓✓
Trasporto dati (IP, FR)	✓	✓	✓✓	✓✓✓	✓✓
Emulazione di circuito	✓✓✓	✓✓	NA	NA	NA
Videoconferenza POTS/ISDN	✓✓✓		NA	NA	
Audio compresso	✓	✓✓✓	✓✓	✓✓	✓
Distribuzione video	✓✓✓	✓✓	✓	NA	NA
Multimedia interattivo	✓✓✓	✓✓✓	✓✓	✓✓	✓
Indicatore di "Vantaggio" per utente: Ottimo: ✓✓✓, Buono: ✓✓, Discreto: ✓, Non Applicabile: NA					

TABELLA 1

Aree di applicazione per le categorie di servizio ATM (Fonte: Libro bianco Forum ATM su "ATM Service Categories: The Benefits to the User")

Il **modello OSI** non è uno standard tecnologico ma un modello di riferimento comune per le tecnologie e i servizi della ICT. Viene rappresentato tramite una pila *stack* di 7 livelli, chiamati: *application, presentation, session, transport, network, data link, physical*. Il livello più basso descrive le caratteristiche fisiche che il sistema deve avere per interagire correttamente con il mezzo fisico (media); il livello più alto rappresenta l'applicazione utilizzata dall'utente (per esempio un *browser* per navigare in Internet o un *client* di posta elettronica).

de al livello 1 del **modello OSI** (*Open System Interconnection*) e gestisce tutti gli aspetti correlati al mezzo di trasmissione (rame, fibra, ...). Il secondo e il terzo, detti rispettivamente ATM Layer e AAL (*ATM Adaptation Layer*) [3], corrispondono all'incirca al livello 2 della pila OSI. L'ATM Layer è responsabile della condivisione contemporanea dei VC rispetto allo strato sottostante (*cell multiplexing*) e della trasmissione delle celle attraverso la rete (*cell relay*). L'AAL svolge il ruolo di isolare i protocolli di livello superiore dai dettagli dei processi ATM adattando i dati per poter essere segmentati e trasformati in *payload* di 48 byte. Nella figura 3 sono rappresentati il modello OSI e i relativi livelli di riferimento ATM.

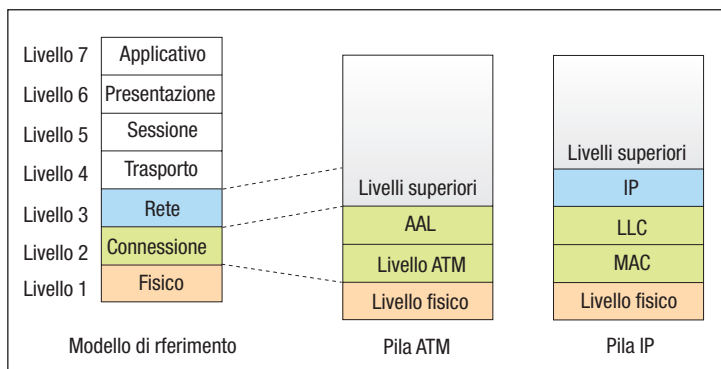
Riassumendo, dunque, ATM è un protocollo di rete di livello 2 (rispetto allo standard OSI), a commutazione di pacchetto, multiprotocollo (nel senso che attraverso AAL è in grado di supportare payload adatti al trasporto di voce, dati o video), e in grado di garantire una determinata qualità del servizio (QoS, *Quality of Service*).

FIGURA 3

Il Modello OSI
e le relative
trasposizioni su
ATM e IP

4. IL PROTOCOLLO IP

IP è, vista la molteplicità delle applicazioni basate su Internet, il protocollo più diffuso al mondo.



Intorno agli anni '70, il D-ARPA (*Defense Advanced Research Projects Agency*) avviò lo sviluppo di una rete a commutazione di pacchetto per connettere i diversi centri di ricerca presenti negli USA, realizzando la rete ARPANET. Verso il 1975, D-ARPA iniziò una cooperazione con l'università di Stanford che portò sul finire degli anni '70 alla nascita della rete Internet [4].

L'IP è un protocollo a commutazione di pacchetto di tipo connection-less. Facendo riferimento alla pila OSI, viene definito come un *Network-Layer protocol*, ovvero, un protocollo di livello 3 che, trasportando informazioni relative all'indirizzo e al controllo dei pacchetti, sovrintende all'instradamento di ogni singolo pacchetto all'interno di una rete. Una rete IP opera secondo il principio del *best effort*, il che significa che non esiste alcuna garanzia sul fatto che, una volta trasmesso, un pacchetto arrivi a destinazione.

Le dimensioni dei pacchetti IP possono variare: per esempio, l'header di un pacchetto è composto da 20 byte (40 byte, nel caso di pacchetti conformi alla più recente versione dello standard, la 6), mentre il payload può essere compreso tra 0 e 65535 byte (dimensioni tipiche sono comprese tra 576 o 1500 byte).

La natura connection-less della rete implica la necessità di prevedere risorse di elaborazione più complesse nei vari nodi per poter instradare i singoli pacchetti, da un punto all'altro, senza però dover effettuare alcuna procedura di setup.

L'instradamento dei pacchetti è realizzato sulla base delle informazioni contenute nel campo *Destination IP Address* dell'header del pacchetto: si tratta di un campo di 32 bit che consente, quindi, di indirizzare al massimo circa 4 miliardi di indirizzi. In generale, gli indirizzi IP sono unici, nel senso che ogni *host* connesso ad una rete IP ha un proprio indirizzo che univocamente lo identifica. In realtà, lo scenario è leggermente più complicato dalla presenza di altri meccanismi [5, 6] utilizzati per aumentare la sicurezza delle reti, che però esulano dallo scopo di questo articolo.

Una rete IP è composta, oltre che dai terminali (*host*), dai cosiddetti *router*, ai quali spetta il compito di instradare i pacchetti

dal mittente al destinatario. I router IP sono organizzati in maniera gerarchica e sono classificati in *Interior* o *Exterior*, a seconda che consentano lo scambio di informazioni tra sistemi chiusi o aperti. Questi router comunicano tra loro attraverso particolari protocolli tra i quali i più noti sono il RIP (*Routing Information Protocol*), e il BGP (*Border Gateway Protocol*).

I protocolli di instradamento IP operano in modo dinamico. Periodicamente, i router si scambiano quelle che comunemente sono chiamate *Routing Table* che consistono in coppie di valori *Destination Address/Next Hop*, dove Next Hop rappresenta il successivo tratto di rete che il pacchetto IP dovrà attraversare [7].

Come detto, il protocollo IP è un protocollo di livello 3 (Figura 3). Si vedano ora quali sono le funzioni svolte dai livelli inferiori; in generale, il compito del livello 2 della pila OSI (data link) è quello di trasmettere al media (mezzo fisico) i dati strutturati per poter identificare il destinatario. Si può dire che l'indirizzamento avviene quindi già a livello 2 (indirizzamento fisico) in modo indipendente dal protocollo di comunicazione.

La differenza fra l'indirizzamento fisico (layer 2) e l'indirizzamento logico (layer 3) sta proprio nell'utilizzo del tipo di indirizzo. A titolo di esempio, nel caso del trasporto di IP su Ethernet, il livello 2 utilizza l'indirizzo MAC (*Media Access Control*) che è del tutto indipendente dal protocollo di comunicazione adottato a livello 3.

Per meglio comprendere questa importante differenza, è necessario mettere in evidenza che il livello 2 del protocollo IP è composto da due sottolivelli noti come LLC (*Logical Link Control*), e MAC, appunto. Il sottolivello LLC (definito nella specifica IEEE 802.2), supporta sia i servizi connection-less che quelli connection-oriented. Esso definisce una serie di campi nella trama di *Data Link* che permettono ai protocolli di livello superiore di condividere lo stesso link fisico. LLC aggiunge ulteriori controlli ai dati ricevuti nel pacchetto IP dal livello 3. Questi ultimi riguardano informazioni sull'indirizzamento 802.2 e in particolare: il *Destination Service Access Point* (DSAP) e il *Source Service Access Point* (SSAP).

Questo re-impacchettamento del pacchetto IP verrà consegnato al sottostante livello MAC che si occupa dell'accesso concorrente al media e della scelta delle priorità di accesso al mezzo fisico. Tale attività può essere svolta solo in presenza di una conoscenza degli host fisicamente attaccati al *media* fisico. La conoscenza avviene attraverso un identificativo/indirizzo fisico chiamato appunto *MAC address*.

Sebbene il protocollo IP sia stato sviluppato circa 30 anni fa, il successo delle applicazioni basate su Internet ha fatto sì che sia stato soggetto ad una continua e rapida evoluzione. In particolare, le aree di sviluppo più fertili sono:

■ **IPSec**: lo sviluppo di un sistema in grado di aumentare la sicurezza e la riservatezza delle informazioni scambiate attraverso Internet;

■ **IP Multicast**: un protocollo che consente di inviare, evitando duplicazioni e dunque mantenendo una elevata efficienza, le stesse informazioni da un server ad un certo numero di client, consentendo di poter utilizzare Internet per applicazioni tipo la TV broadcast;

■ **IPv6**: la versione attuale del protocollo IP, la 4, non è in grado di garantire un sufficiente numero di indirizzi visto che il numero di utilizzatori di Internet è in continua crescita, per questo IPv6 prevede l'estensione del campo indirizzi [8].

■ **Quality of Service**: molti sforzi vengono profusi per cercare di offrire una qualche QoS su una rete IP in modo tale da permettere il trasporto di servizi o applicazioni real-time. Attualmente, sono stati definiti due modelli: il primo denominato *Integrated Services* (IntServ) [9] basato sul protocollo RSVP (*Reservation Protocol*) [10] che consente di allocare banda verso una determinata destinazione. Le difficoltà di implementare su larga scala il protocollo RSVP ha portato al secondo modello denominato *Differentiated Services* (DiffServ), basato sul principio che, in funzione del tipo di informazione trasportata, i pacchetti possono essere marcati *tagget* con un'indicazione relativa alla loro priorità.

Riassumendo, IP è un protocollo di livello 3 di tipo connection-less che non offre QoS. D'altra parte però è da sottolineare che l'efficienza del sistema di indirizzamento e dei meccanismi di instradamento e la diffusione pervasiva che continua ad avere, rappresentano

aspetti di assoluto rilievo tanto che oggi praticamente tutti gli apparati connessi in rete sono dotati di almeno un indirizzo IP piuttosto che di uno di tipo E.164 (la raccomandazione ITU-T che definisce il sistema di indirizzamento ATM).

5. IP SU TUTTO, ANCHE SU ATM

Oggi un pacchetto IP può essere trasmesso praticamente su qualsiasi media.

Per quanto illustrato finora, è facile verificare che tra IP e ATM esista un punto di contatto [11]. Facendo riferimento ai rispettivi modelli OSI, è evidente come a livello 2 lo strato ATM AAL coincida col corrispondente livello IP: LLC (Figura 3).

I metodi più comuni per effettuare questo "imbustamento" sono stati definiti in seno a IETF (*Internet Engineering Task Force*) e sono: LLC *Encapsulation e Virtual Connection Multiplexing*.

La prima è una modalità di imbustamento dei protocolli TCP/IP applicata, diffusamente, su molte tecnologie tradizionali; dal punto di vista architetturale, essa richiede di inserire il livello LLC tra lo strato di protocollo AAL5 e quello IP. I pacchetti IP sono imbustati in pacchetti LLC a loro volta inseriti nelle trame AAL5 [12]; la busta LLC consente di condividere lo stesso canale logico distinguendo, grazie ai campi *Destination Service Access Point* (DSAP) e *Source Service Access Point* (SSAP), il protocollo di livello superiore trasportato. Poiché il protocollo IP non è tra quelli standardizzati da OSI, l'adattamento tra LLC e AAL5 viene effettuata utilizzando l'estensione *Sub-Network Access Protocol* (SNAP).

Nel secondo caso, i pacchetti di livello rete sono trasportati all'interno delle trame AAL5 senza imbustamenti aggiuntivi; per ogni protocollo di livello rete è creata una diversa istanza di AAL5. Per distinguere protocolli differenti si utilizzano connessioni virtuali differenti, anche tra le stesse macchine, spostando la moltiplicazione fuori banda. Gli estremi delle connessioni virtuali sono i protocolli di livello rete, non i nodi della rete.

L'architettura descritta è conosciuta con il nome di *Overlay Model*. Tale modello potreb-

be essere rappresentato da un certo numero di router connessi per mezzo di una rete magliata tramite PVC ATM che offrono elevata banda e bassa latenza.

Tale modello, per quanto elegante e per molti aspetti efficace, non risulta essere né sufficientemente scalabile né facilmente amministrabile. Infatti, lo strato di rete IP è governato dai propri protocolli di instradamento, così come lo strato di rete ATM ha i suoi meccanismi di gestione, segnalazione, instradamento [13].

In particolare, per quanto riguarda la scalabilità dell'architettura è sufficiente notare che nella suddetta ipotesi di rete completamente magliata, se N è il numero di nodi, per connettere tutti i router sono necessari $N(N-1)$ PVC. Tanto per dare un ordine di grandezza, se fossero presenti 200 router occorrerebbero all'incirca 40.000 connessioni!

6. VERSO L'INTEGRAZIONE: LA COMMUTAZIONE DI ETICHETTA

Il principale collo di bottiglia dei meccanismi di routing del protocollo IP è rappresentato dal loro funzionamento *hop-per-hop*, secondo il quale prima di poter decidere il cosiddetto *Next-Hop* è necessario che il router estragga tutte le informazioni contenute nel pacchetto ricevuto fino al livello 3.

Un meccanismo per aggirare questo problema è di instradare i pacchetti direttamente a livello 2 eliminando la costosa elaborazione a livello 3 che impegna risorse e capacità di elaborazione. Nel contesto IP su ATM, questo si traduce nella possibilità di trasferire pacchetti IP senza che vengano processati da alcun router.

Sono state sviluppate diverse architetture per realizzare quanto sopra, come per esempio NHRP (*Next Hop Resolution Protocol*), e MPLS (*Multi Protocol Label Switching*) [14, 15], sviluppati da IETF o, ancora, PAR (*Public-NNI Augmented Routing*) e I-PNNI (*Integrated Public-NNI*), sviluppati da ATM Forum. La figura 4, riassume tali protocolli mettendoli a confronto.

Sebbene nei loro intenti sia IETF che ATM Forum perseguissero l'obiettivo di realizzare un sistema ad alte prestazioni di instradamento del traffico IP su ATM [16], è bene sottolineare

che le loro motivazioni di fondo erano ben diverse. IETF mirava ad aumentare la banda del traffico best effort IP, mentre ATMF mirava a potenziare il più complesso, ma più efficace protocollo PNNI adattandolo a quello che si prospettava potesse diventare il principale cliente, l'IP.

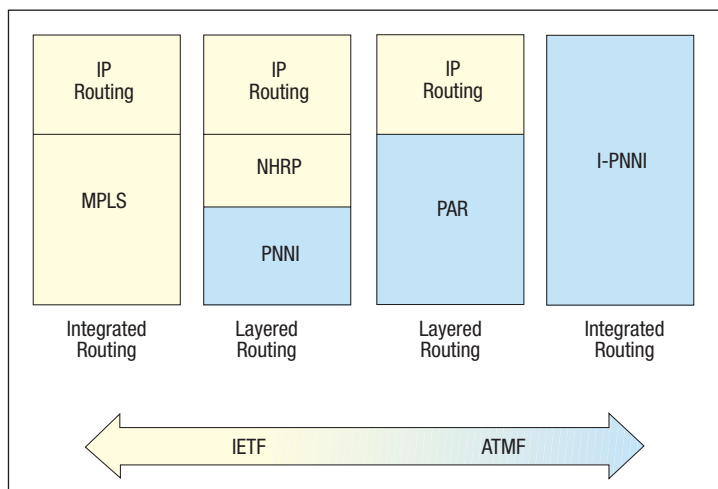
Sebbene ognuna delle suddette soluzioni abbia dei pro e dei contro, ad oggi, sembra che il mercato delle telecomunicazioni stia guardando sempre più insistentemente verso MPLS.

Per completezza, è bene sottolineare che l'obiettivo delle suddette tecnologie non è tanto quello incrementare le prestazioni dei sistemi, cosa peraltro ottenibile a livello *hardware* attraverso meccanismi di commutazione di livello 3 (Layer 3 Switching), i quali utilizzando opportuni ASIC (*Application-Specific Integrated Circuit*) instradano i pacchetti IP semplicemente "leggendone" l'indirizzo senza dover processare l'intera pila protocolare, quanto quello di aggiungere alle esistenti infrastrutture di rete l'intelligenza necessaria per garantire la richiesta qualità di servizio attraverso meccanismi che consentono il cosiddetto *Traffic Engineering* (TE).

Senza voler assolutamente affrontare i dettagli di questa tecnologia, né aver la pretesa di esaurirne in poche righe tutti gli aspetti, può, comunque, essere utile offrirne una brevissima descrizione.

Ai pacchetti che entrano in una rete MPLS vengono assegnate delle etichette (*label*) dai router periferici, chiamate *Label Edge Router* o LER. Le stesse LER si occupano di eliminare le etichette in uscita dalla rete MPLS. Le etichette vengono assegnate non solo in base a caratteristiche come l'ampiezza di banda o l'indirizzo di destinazione, ma anche guardando i campi dell'header del pacchetto IP come l'indirizzo sorgente e il *Type of Service*. All'interno della rete MPLS, il percorso del pacchetto - chiamato *Label Switching Path* o LSP - viene definito dalla nuova etichetta, non dall'indirizzo IP di destinazione originale.

Lungo un LSP, i router intermedi o *Label Switch Router* (LSR) si occupano di scambiare l'etichetta entrante con una uscente (in modo assolutamente identico a come opera uno switch ATM). Mediante lo smistamento su opportuni LSP, è possibile gestire



più facilmente diversi flussi di dati in base alla priorità e al tipo di servizio che si vogliono ottenere.

FIGURA 4
Architetture di integrazione tra IP e ATM

7. CONCLUSIONI

Per quanto detto è evidente che, in realtà, ATM ed IP possano essere considerati complementari tra loro. Il primo opera più vicino al livello fisico ed è diretto a fornire prestazioni di trasmissione ad alta velocità e basso ritardo su percorsi pre-tracciati. Esso è orientato ad ospitare i livelli superiori dei protocolli orientati alle applicazioni che chiedano per questo precisi requisiti al canale di trasmissione. ATM attuerà i meccanismi migliori per offrire quanto richiesto in funzione della tecnologia usata. Garantire il tasso d'errore in un collegamento radio od ottico è cosa totalmente diversa e ha soluzioni anch'esse diverse. D'altra parte, il protocollo IP si basa sull'assunto che tutte le funzioni di instradamento e correzione debbano coinvolgere il livello 2, in tal modo IP attua dinamicamente percorsi multipli in modo tale che in qualunque momento possa venir meno anche il *link* fisico. Ciò dà all'IP una grande "capacità di sopravvivenza" anche in condizioni ostili, ma obbliga ogni router ad un superlavoro che ritarda lo smistamento e sovraccarica il nodo.

Sul piano dei servizi offerti, questi si stanno muovendo sempre più verso la multimedia: tant'è che oggi si sente sempre più frequentemente parlare del cosiddetto *triple play*, ovvero, la possibilità per un operatore di convogliare telefonia, dati e video sulla

stessa rete. In particolare, i servizi telefonici e televisivi “consegnano” un prodotto finale che risente della qualità del canale su cui hanno viaggiato, in termini di tollerabilità nell’ergonomia di un’interazione umana.

Una possibile visione dell’evoluzione dunque, è che le funzioni di livello 2 si avvicinino progressivamente alle funzioni di trasmissione offrendo alla parte più prossima all’applicazione il compito di identificare le caratteristiche del contesto in cui i propri pacchetti vengono trasportati. Questo porterà ad una ridistribuzione delle funzionalità nei livelli OSI assegnando ai livelli inferiori il compito di garantire il contesto di trasmissione, e quelli superiori il compito di sceglierne i parametri di trasporto in funzione del contenuto che affidano alla rete.

D’altra parte, non è forse vero che si applica il francobollo “posta prioritaria” o le etichette: “fragile”, “non capovolgere” alle merci e lettere? Ci si fida poi del fatto che il vettore le tratti secondo tali etichette senza entrare nel merito di quanto vi è all’interno.

Come si è visto sopra, i processi evolutivi nella storia delle telecomunicazioni hanno chiesto sempre qualche decennio. Questo passo ormai avviato con MPLS, non potrà che totalizzare almeno un paio di lustri, soprattutto, a causa del gigantismo della rete che coinvolge, più che per la velocità di sviluppo della tecnologia disponibile.

Bibliografia

- [1] AA.VV.: *B-ISDN asynchronous transfer mode functional characteristics*. Racc. I.150. ITU-T. Febbraio 1999.
- [2] AA.VV.: *B-ISDN ATM layer specification*. Racc. I.361. ITU-T. Febbraio 1999
- [3] AA.VV.: *B-ISDN ATM adaptation layer (AAL) functional description*. Racc. I.362. ITU-T. Marzo 1993.
- [4] Postel J: *Internet Protocol - Darpa Internet Program Protocol Specification*. RFC791. IETF. Settembre 1981.
- [5] Mogul J: *Internet Subnets*. RFC917. IETF. Ottobre 1984.
- [6] Postel J: Mogul J, *Internet Standard Subnetting Procedure*. RFC950. IETF. Agosto 1985.
- [7] Katz D, Piscitello D: *NBMA next Hop Resolution Protocol (NHRP)*. Work in progress: draft-ietf-rolc-nhrp-03.txt. IETF. Novembre 1994.

- [8] Bradner S, Mankin A: *The Recommendation for the IP Next Generation Protocol*. RFC 1752. IETF. Gennaio 1995.
- [9] Braden R, Clark D, Shenker S: *Integrated Service in the Internet Architecture: an Overview*. RFC 1633. IETF. Luglio 1994.
- [10] Zhang L, Estrin D, Herzog S, Jamin S: *Resource ReSerVation Protocol (RSVP)*. Ed. Work in progress: draft-ietf-rsvp-spec-05.ps. IETF. Marzo 1995.
- [11] Laubach M: *Classical IP and ARP over ATM*. RFC 1577. IETF. Gennaio 1994.
- [12] Atkinson R: *Default IP MTU for use over ATM AAL5*. RFC 1626. IETF. Maggio 1994.
- [13] Heinanen J: *Multiprotocol encapsulation over ATM adaptation layer 5*. RFC1483, IETF. Dicembre 1994.
- [14] Rosen E, Viswanathan A, Callon R: *Multiprotocol Label Switching Architecture*. RFC3031. IETF. Gennaio 2001
- [15] Davie B, Lawrence J, McCloghrie K, Rosen E, Swallow G, Rekhter Y, Doolan P: *MPLS using LDP and ATM VC Switching*. RFC3035. IETF. Gennaio 2001
- [16] Perez M, Liaw F, Grossman D, Mankin A, Hoffman A: *ATM signalling support for IP over ATM*. RFC 1755. IETF. Gennaio 1995.
- [17] Nycholas P: *Negroponte Being Digital*. 1995 Sperling & Kupfer Editori SpA.
- [18] Bell labs: <http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html>

GIANNI GRAZIA laureato in Ingegneria Elettronica all’Università di Bologna nel 1966.

Ha frequentato corsi di commutazione telefonica e sviluppo software presso Michigan State University (nel 1970) e University of California Los Angeles (nel 1974). Negli anni ’70 ha progettato il controllo dell’Autocommutatore Fonia e Dati (AFD) del quale ha poi curato lo sviluppo software, pubblicando vari articoli sull’argomento.

È attualmente Direttore Tecnico di ALCATEL ITALIA e responsabile di Business Analysis e delle attività di Integrazione Servizi nell’ambito del Marketing di Alcatel Italia.

e-mail: Gianni.Grazia@alcatel.it

PAOLO CECCHERINI laureato in Ingegneria Elettronica all’Università “La Sapienza” di Roma nel 1995, si occupa di Architetture e Soluzioni di rete nell’ambito del Marketing di Alcatel Italia.

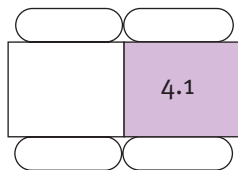
Precedentemente, ha svolto attività di ricerca e progettazione passando dallo sviluppo di ASIC, alla progettazione di apparati per trasmissione dati fino all’integrazione di sistemi complessi.

e-mail: Paolo.Ceccherini@alcatel.it



T.R.E.: UNA SPIN-OFF UNIVERSITARIA PER L'ELABORAZIONE DI DATI RADAR SATELLITARI

Fabio Rocca
Carlo Colesanti
Alessandro Ferretti
Raffaella Ratti



L'articolo ha l'obiettivo di documentare l'esperienza imprenditoriale di Tele-Rilevamento Europa, spin-off commerciale nata presso il Politecnico di Milano per offrire al pubblico prodotti e servizi sviluppati a partire da venti anni di ricerca all'avanguardia nel settore del telerilevamento radar. Dopo una sezione introduttiva sui sistemi radar e sull'interferometria, si illustrano i fondamenti scientifici della tecnica dei diffusori permanenti (Permanent Scatterers).

1. TELE-RILEVAMENTO EUROPA: DALLA RICERCA ALL'IMPRESA

L'esperienza imprenditoriale di Tele-Rilevamento Europa – T.R.E. Srl, *spin-off* commerciale del Politecnico di Milano, nasce dall'idea di offrire al pubblico i prodotti di una tecnologia d'avanguardia, frutto di una lunga tradizione di ricerca ad alto livello nel settore del telerilevamento radar satellitare.

Il concetto di spin-off universitario, ovvero una società commerciale che nasce in seno a una realtà accademica per trasformare in *business* quanto conseguito con un serio lavoro di ricerca scientifica, è, purtroppo, estraneo alla tradizione italiana. Si tratta, invece, di una prassi consolidata nei paesi occidentali in generale e negli Stati Uniti in modo particolare. Anche senza varcare l'Oceano Atlantico, vale la pena di osservare che nella vicina Svizzera sono numerose le realtà commerciali sorte presso le maggiori istituzioni universitarie per valorizzare i risultati più promettenti della ricerca scientifica.

In questo senso Tele-Rilevamento Europa costituisce realmente una novità nel panorama italiano: si tratta del primo spin-off commerciale del Politecnico di Milano, e probabilmente, è una *première* a livello nazionale. Ad ogni modo, la realtà italiana sembra essere finalmente matura per questo genere di tentativi imprenditoriali, tanto che al Politecnico di Milano, come del resto presso diverse altre prestigiose sedi universitarie, stanno già nascendo diverse nuove spin-off.

L'aspetto più interessante e appassionante di una tale esperienza è l'occasione di seguire la nascita di un prodotto e/o di un servizio attraverso tutte le fasi della sua gestazione. Da un'idea innovativa, all'intensa attività di ricerca volta a dimostrarne l'applicabilità e a descriverne formalmente e rigorosamente potenzialità e limiti, a nuove idee che dischiudono (o, talvolta, chiudono) orizzonti più vasti. Si tratta del lavoro affascinante di chi ha l'opportunità di dedicarsi alla ricerca scientifica, ma è solo la prima tappa, anche se la più congeniale, per chi è solito lavorare in un contesto universitario.



Si rende, quindi, necessario identificare il mercato potenziale per il nuovo risultato scientifico che, proprio in quanto nuovo, è praticamente sconosciuto al pubblico. Si tratta, quasi simultaneamente di *marketing* e di definizione e implementazione di prodotti e servizi in base alla risposta dei gruppi di clienti potenziali che via via si individuano e che prendono forma nelle rispettive peculiari esigenze, attese e necessità.

È un lavoro impegnativo e appassionante, nel corso del quale, come in sede di ricerca scientifica, non mancano imprevisti e le risposte più importanti o significative non corrispondono necessariamente alle aspettative iniziali.

Ci si propone di illustrare nel seguito le tappe più significative di nascita e crescita di Tele-Rilevamento Europa. Si ritiene però opportuno, innanzitutto, mettere a fuoco il contesto scientifico in cui è sorta la società.

2. INTRODUZIONE AI SISTEMI SAR

Si forniscono ora alcuni brevi cenni introduttivi ai sistemi radar e in particolare, ai sistemi radar ad apertura sintetica (*Synthetic Aperture Radar*, SAR).

2.1. Cenni storici

Radar è acronimo di **radio detecting and ranging** [1]. Esattamente un anno fa è stato festeggiato il centenario dell'ultimo di una straordinaria sequenza di esperimenti di Marconi: la prima trasmissione radio tra le sponde opposte dell'Atlantico, avvenne infatti il 12 dicembre 1901 [11 - 13]. Furono proprio i risultati di Marconi a suggerire di utilizzare le onde radio per localizzare oggetti.

In genere, l'invenzione del radar è ascrivita all'inglese Watson-Watt, 1935. In realtà, già nel 1904 il tedesco Hülsmeyer aveva messo a punto il "telemobiloscopio", apparecchio con cui era in grado di localizzare navi nella nebbia fino ad una distanza di un paio di chilometri [11]. Come spesso accade a chi precorre troppo i tempi, nessuno intuì le potenzialità dell'oggetto, che venne ben presto dimenticato.

Negli anni 1936-1939, mentre gli Stukas (velivoli da bombardamento in picchiata) tede-

schì martellavano le città spagnole durante la guerra civile, l'importanza di un sistema in grado di localizzare aerei nemici in avvicinamento divenne assolutamente evidente.

Nel corso di pochissimo tempo fu messa a punto "Chain Home", la prima rete radar di avvistamento destinata a sorvegliare la costa meridionale della Gran Bretagna. È superfluo ricordare il ruolo chiave giocato da queste stazioni radar durante la Battaglia d'Inghilterra nel 1940 e nel corso dell'intero conflitto [9].

Alla fine degli anni '50, negli Stati Uniti maturò l'idea di sistemi ad apertura sintetica che prevedevano l'impiego di un radar coerente montato su piattaforma mobile (all'inizio aerei, ora anche satelliti), in grado di illuminare l'area oggetto di interesse dalle diverse posizioni successivamente occupate dalla piattaforma lungo una ben precisa traiettoria.

Come sarà illustrato più avanti, ciò si traduce nella facoltà di generare immagini a elevata risoluzione spaziale pur utilizzando antenne di dimensioni contenute.

I primi sistemi SAR montati su aerei furono inizialmente utilizzati per scopi di ricognizione militare.

Ben presto si cominciarono a generare immagini radar del territorio [7].

Nel 1978 fu lanciato Seasat, il primo satellite per uso civile equipaggiato con un radar SAR, destinato allo studio degli oceani. Nonostante la breve durata della missione, compromessa dopo soli cento giorni da un cortocircuito, i dati ottenuti rivelarono potenzialità straordinarie.

Negli anni 1981, 1984 e 1994 si sono susseguite missioni SAR basate sull'impiego dello Space Shuttle (*Shuttle Imaging Radar*, SIR-A, SIR-B, SIR-C).

Nel luglio 1991 l'Agenzia Spaziale Europea (*European Space Agency*, ESA) lancia il satellite ERS-1 (*European Remote Sensing Satellite*) equipaggiato con sensore SAR. Nel 1995 gli viene affiancato ERS-2.

Nel contempo anche giapponesi (JERS, 1992) e canadesi (Radarsat, 1995) avviano missioni SAR.

Nel 2000, una nuova missione Shuttle (*Shuttle Radar Topography Mission*) viene dedicata all'acquisizione di dati SAR per la

ricostruzione della topografia delle terre comprese tra i paralleli 60°N e 56°S [8]. Diverse nuove missioni sono già state pianificate: tra le altre, si desidera menzionare ENVISAT (Agenzia Spaziale Europea, ESA) e Radarsat-2 (Agenzia Spaziale Canadese).

2.2. Sistemi radar ad apertura sintetica

Il principio di funzionamento di un sistema radar è abbastanza semplice: un apparecchio trasmittente illumina lo spazio circostante con un'onda elettromagnetica che incide su eventuali oggetti subendo un fenomeno di riflessione disordinata (diffusione, *scattering*). Una parte del campo diffuso torna verso la stazione trasmittente, equipaggiata anche per la ricezione.

Il ritardo temporale tra l'istante di trasmissione e quello di ricezione consente di valutare la distanza (*ranging*) a cui si trovano i singoli bersagli radar, cioè di localizzare gli oggetti lungo la direzione detta di *range*.

La **direttività dell'antenna** utilizzata per trasmettere e ricevere il segnale radar, e cioè la selettività nell'illuminazione dello spazio circostante consente di localizzare l'oggetto anche lungo l'altra dimensione (detta di *azimuth*).

Quanto più grande è l'antenna, tanto più stretta è la sua impronta e, di conseguenza, tanto meglio viene localizzato il bersaglio. Chiaramente ciò avviene a scapito dell'estensione dell'area illuminata.

Per ovviare questo inconveniente, le antenne molto direttive usate per i radar militari e per applicazioni di aviazione civile ruotano in modo da "spazzare" tutta l'area circostante alla loro posizione.

L'idea alla base del SAR consente di aggirare la stessa limitazione (compromesso riso-

luzione-estensione dell'area osservata). Combinando coerentemente (cioè tenendo conto di modulo e fase del segnale SAR, concetti il cui significato fisico si introdurrà in seguito) i dati acquisiti dal sensore nelle posizioni successivamente occupate, si sintetizza un'antenna fittizia di grandi dimensioni detta, appunto, apertura sintetica. È proprio questo procedimento a garantire un'elevata risoluzione anche nella direzione di azimuth [10].

Si farà ora riferimento ai sensori SAR montati sui satelliti ERS-1 ed ERS-2 dell'Agenzia Spaziale Europea (Figura 1). ERS-1 ha acquisito dati dalla fine del 1991 a marzo 2000. ERS-2 è operativo dall'inizio del 1995. I satelliti ERS seguono orbite lievemente inclinate rispetto ai meridiani, illuminando una striscia di terreno larga approssimativamente 100 km (*swath*). La direzione parallela all'orbita, detta azimuth, coincide all'incirca con la direzione nord-sud. La risoluzione in azimuth è di circa 5 m. La direzione della congiungente satellite-bersaglio radar, perpendicolare all'orbita, è detta *range* oppure *line of sight (LOS)*. La risoluzione in range vale circa 9 m. Le immagini radar si sviluppano lungo le direzioni di range e azimuth, dette usualmente coordinate SAR.

I satelliti ripercorrono la stessa orbita nominale ogni 35 giorni (*revisiting time*) illuminando attivamente la medesima area a terra con un segnale alla frequenza di 5.3 GHz (lunghezza d'onda $\lambda = 5.66$ cm). A differenza dei sistemi ottici, i sensori ERS-1/2 sono in grado di acquisire dati con qualsiasi condizione meteorologica e di notte.

Le immagini radar ottenute con **sistemi coerenti** si compongono di modulo (o ampiezza) e fase. Il modulo individua la quota di campo elettromagnetico incidente che ogni singolo bersaglio illuminato riflette verso il sensore. La fase racchiude, invece, l'informazione relativa alla distanza tra sensore e bersaglio. L'analisi interferometrica consiste nello studio dell'evoluzione della fase tra due distinte acquisizioni. L'interferogramma riassume in sé i diversi aspetti responsabili di variazioni di fase. I più importanti sono la topografia, eventuali fenomeni di movimento della superficie terrestre e la variazione delle condizioni atmosferiche.

Si usa l'espressione **direttività dell'antenna** per descrivere il funzionamento delle antenne. Si usano diversi parametri, il più importante è detto *funzione di direttività* ed individua quale porzione di spazio e con quale intensità di campo l'antenna è in grado di illuminare.

Sistemi coerenti. Un sistema si dice coerente se è in grado di acquisire dati di ampiezza e fase. La coerenza è ipotesi chiave per poter sfruttare il principio dell'apertura sintetica.

Tutti i sistemi radar sono pensati per stimare la distanza sensore - bersaglio con un'accuratezza, in genere, dell'ordine del metro.

L'impiego dell'informazione di fase (e quindi la coerenza) è indispensabile per stimare variazioni molto contenute (millimetriche) della distanza sensore - bersaglio da un'immagine radar alla successiva.



I principali obiettivi dell'interferometria SAR sono le ricostruzioni di modelli altimetrici numerici e di mappe di deformazione della superficie terrestre.

Per conseguire ogni singolo obiettivo si rende necessaria l'analisi del corrispondente contributo di fase. Gli altri termini, in particolare modo la distorsione di fase introdotta dalle condizioni atmosferiche, agiscono come fattori di disturbo.

La lettura degli interferogrammi è notevolmente complicata da fenomeni di decorrelazione che sono legati a variazioni della riflettività della superficie illuminata, che si traducono in un ulteriore contributo di fase casua-

le che rende inaccessibili i termini che individuano topografia e movimento. La riflettività locale (e in particolare, la sua fase) può infatti variare nel tempo, ad esempio con il susseguirsi delle stagioni in un'area vegetata, introducendo decorrelazione temporale. In genere, solo in aree ad elevata urbanizzazioni si riesce ad estrarre informazione da interferogrammi che coinvolgono immagini acquisite a distanza di anni.

La riflettività varia in modo ancor più significativo in funzione della geometria di acquisizione (un effetto del cambio di prospettiva). Le orbite reali del satellite si discostano dall'orbita nominale introducendo decorrelazio-

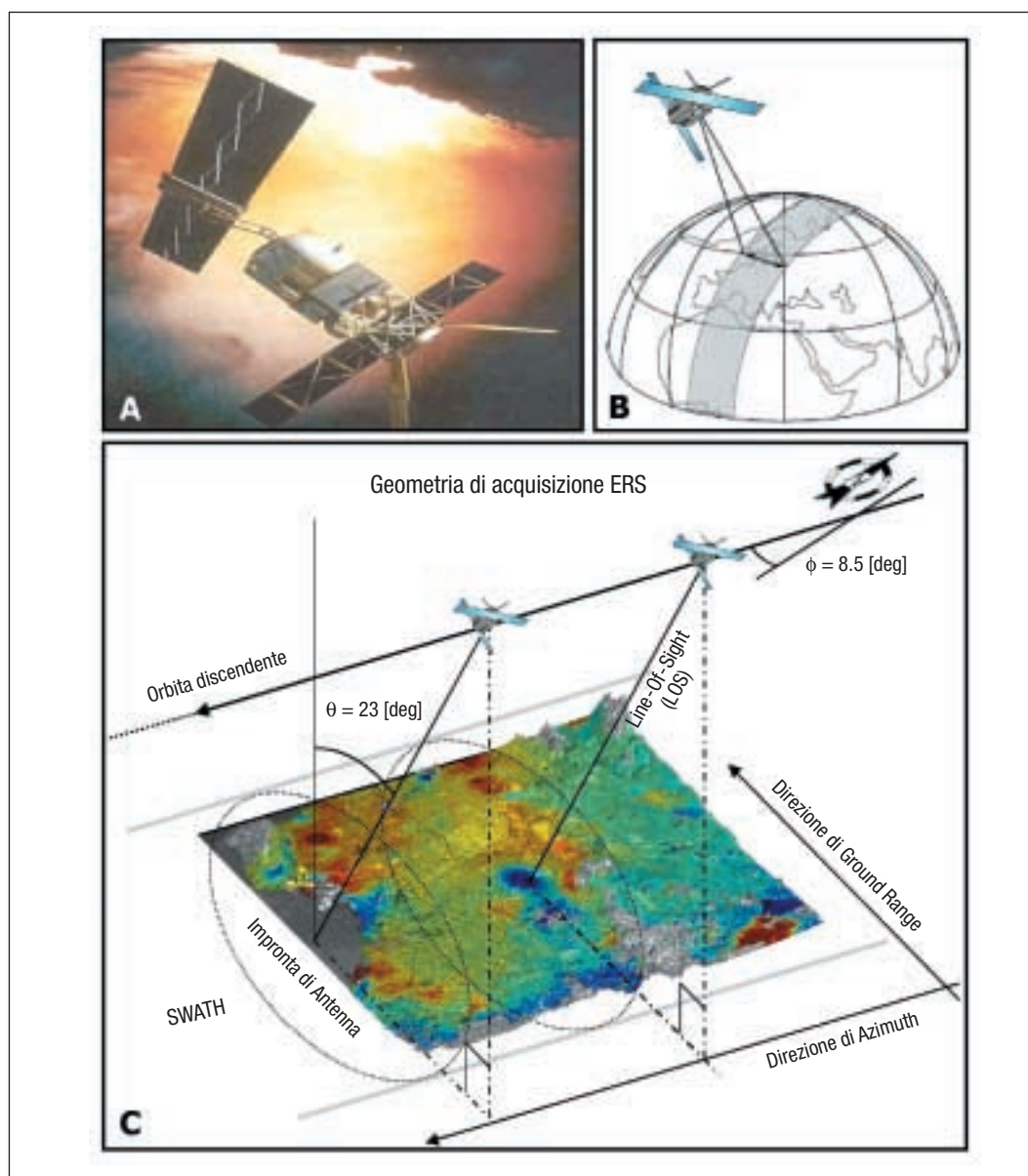


FIGURA 1

A Il satellite ERS.
B e C Geometria di acquisizione ERS

ne geometrica. È sufficiente una distanza di poche centinaia di metri tra le orbite relative a una coppia interferometrica a renderne completamente illeggibile la fase.

I limiti dell'interferometria SAR classica sono proprio legati ai fenomeni di decorrelazione e al contributo atmosferico di fase che sovente genera effetti difficili da distinguere da fenomeni di movimento del terreno e/o da profili altimetrici.

2.3. La tecnica dei diffusori permanenti

Entrambe le difficoltà sono state risolte tramite la tecnica dei diffusori permanenti (*Permanent Scatterers, PS*), messa a punto al Politecnico di Milano [3, 4].

L'approccio PS è basato sull'osservazione che un piccolo sottoinsieme di bersagli radar, costituito appunto dai diffusori permanenti, è praticamente immune agli effetti di decorrelazione. I PS preservano l'informazione di fase nel tempo e al variare della geometria di acquisizione. Possono, inoltre, essere utilizzati per ricostruire e compensare efficacemente il disturbo atmosferico sull'intera immagine radar, sfruttando il fatto che le condizioni atmosferiche variano lentamente nello spazio. Chiaramente è necessario che la densità spaziale di PS sia sufficientemente elevata (maggiore di 5-10 PS/km²), vincolo sempre verificato in aree urbane. In corrispondenza dei PS è, inoltre, possibile separare con estrema accuratezza il termine di fase dovuto alla topografia da quello dovuto al movimento del suolo. Per far ciò è necessario avvalersi di *data set* consistenti in almeno 25-30 immagini ERS. In aree ad elevata urbanizzazione, la densità spaziale di PS raggiunge valori molto alti: 100 – 400 PS/km².

In corrispondenza di ogni singolo PS si ricava il trend medio di deformazione con accuratezza compresa tra 0.1 e 1 mm/anno. L'accuratezza è in funzione del numero di immagini e della "qualità" del PS stesso, cioè di quanto l'informazione di fase disponibile presso il PS è immune a fenomeni di disturbo.

È, inoltre, possibile ricostruire l'intera serie temporale di deformazione del PS; l'accuratezza arriva (nei punti migliori) a 1 mm su ogni singola misura.

Un ulteriore aspetto vantaggioso è costituito dalla ricchezza di dati disponibili nell'archivio ERS dell'ESA. È possibile avviare un'analisi PS oggi, avendo a disposizione dati acquisiti a partire del 1992 e potendo, quindi, ricostruire la storia passata dell'area di interesse.

I limiti consistono nella facoltà di apprezzare solo la deformazione lungo la direzione di LOS, cioè approssimativamente lungo la verticale, e nel fatto che per portare a termine con successo l'analisi PS è necessario che l'area oggetto di studio presenti una densità sufficiente di diffusori permanenti (quanto meno lieve urbanizzazione oppure presenza di rocce esposte). Da ultimo, l'analisi di fenomeni di deformazione con evoluzione particolarmente rapida (maggiore di 80 mm/anno) è possibile solo disponendo di informazioni a priori sui fenomeni in atto.

Risultati significativi ottenuti con la tecnica PS sono stati validati facendo uso di dati di livellazione ottica, di misure effettuate presso stazioni GPS (*Global Positioning System*) e attraverso lo studio di fenomeni di dilatazione termica di strutture.

In definitiva, i PS costituiscono una sorta di "rete geodetica naturale" che consente l'analisi di fenomeni di deformazione superficiale. Si tratta di uno strumento di monitoraggio senza eguali per accuratezza, densità spaziale di punti di misura e competitività economica, in particolare nello studio di fenomeni di subsidenza urbana e di deformazione lungo faglie sismiche in aree di estensione di centinaia o migliaia di chilometri quadrati.

3. DALLA RICERCA AL MERCATO

Risulta interessante a questo punto ripercorrere i passi fondamentali di nascita e crescita di Tele-Rilevamento Europa, mettendo a fuoco, in particolare, il mercato cui sono rivolti prodotti e servizi PS.

3.1. Risultati scientifici e nascita di T.R.E.

Come già accennato, T.R.E. nasce dal lavoro di ricerca nell'ambito del telerilevamento radar condotto presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano dove, da circa venti anni, è attivo un gruppo di ricer-



ca guidato dai professori F. Rocca e C. Prati. Il *team* si occupa di elaborazione numerica dei dati acquisiti da sistemi SAR e può vantare un ampio riconoscimento a livello internazionale, documentato da numerosi articoli su riviste scientifiche, da brevetti nazionali e internazionali, e dalla partecipazione a numerosi progetti promossi, in particolare, dall'Agenzia Spaziale Europea (ESA), dall'Agenzia Spaziale Italiana (ASI) e dall'Unione Europea. I primissimi tentativi nella direzione di un'analisi PS risalgono agli anni 1996-1997, quando ormai l'interferometria SAR convenzionale cominciava a essere considerata una disciplina matura e diversi scienziati erano quasi convinti che non molto fosse ancora da sviluppare.

Negli anni successivi, si sono precisate le linee fondamentali dell'approccio PS e tra il 1998 e il 1999 sono stati conseguiti i primi notevoli risultati nell'area della Grande Frana di Ancona e sulle pendici dell'Etna (studio dei fenomeni di compattazione lavica nella Valle del Bove).

Nel 1999, la tecnica è stata brevettata a livello nazionale dal Politecnico di Milano ed è cominciata la pratica per il brevetto internazionale.

L'analisi delle potenzialità della tecnica, parallelamente all'accumularsi di una crescente messe di risultati, hanno portato, nel marzo 2000, alla nascita di Tele-Rilevamento Europa- T.R.E. Srl che è divenuta effettivamente operativa verso la fine dell'anno e a cui è stato affidato in esclusiva lo sfruttamento del brevetto.

T.R.E. è stata accolta per circa un anno e mezzo presso l'Acceleratore di Impresa del Politecnico di Milano [2], che ha messo a disposizione spazi e risorse per agevolare la crescita della società.

3.2. Il mercato e i prodotti di T.R.E.¹

Già da tempo esiste - e si va sempre più consolidando - un mercato legato al monitoraggio della deformazione della superficie terrestre (subsidenza, movimento lungo faglie

sismiche, frane e instabilità di versanti ecc.). Si pensi, per esempio, ai fenomeni di subsidenza (sprofondamento del suolo) che possono essere indotti dal prelievo di acqua della falda freatica (80% dei casi [6]), dall'estrazione di petrolio o gas naturale, oppure da fenomeni naturali di compattamento di sedimenti, in particolare lungo il corso di fiumi.

Nei soli Stati Uniti, l'area affetta da subsidenza supera i 45.000 km² [6] e produce costi che è difficile quantificare, ma che stime prudenti suggeriscono superare i 500 milioni di dollari l'anno [5].

In Europa, la situazione è ancora più drammatica a causa di un maggiore livello di urbanizzazione.

Un discorso analogo vale per i fenomeni franosi e di instabilità dei versanti.

Si tratta di tematiche strettamente connesse alla sicurezza del cittadino (Protezione Civile) a cui l'opinione pubblica è sempre più sensibile e verso cui viene convogliata una mole crescente di risorse pubbliche e private.

È ormai ampiamente riconosciuto il ruolo chiave del monitoraggio ambientale nella definizione di scenari di rischio realistici, indispensabili per la pianificazione di efficaci interventi a tutela della sicurezza.

Un ruolo analogo comincia, finalmente, ad essere ascrivito al monitoraggio preventivo anche per quanto concerne lo studio preliminare e la pianificazione di interventi infrastrutturali di grande rilievo (tratte ferroviarie e autostradali, tunnel, viadotti, linee metropolitane ecc.).

In particolare, il monitoraggio dei fenomeni di deformazione superficiali è generalmente di competenza della Pubblica Amministrazione, sia a livello locale, sia a livello nazionale (Legge 225 del 24 febbraio 1992, in relazione a fenomeni franosi). È quindi nella pubblica amministrazione che si è riconosciuto, almeno in una prima analisi, l'utente per eccellenza di prodotti e servizi PS.

Anche le compagnie petrolifere sono tenute, in diversa misura, al monitoraggio del suolo nelle zone di estrazione e costituiscono, di conseguenza, potenziali clienti. Prodotti e servizi PS possono, peraltro, rivelarsi preziosi anche per progettare nuove perforazioni

¹ L'analisi di mercato per prodotti e servizi PS è stata svolta con la fattiva collaborazione del Consorzio Politecnico Innovazione.

(dall'eventuale *pattern* di subsidenza indotto si può inserire dati su forma, profondità ed estensione del giacimento sottostante).

Più in generale, è nella grande industria che si individua un secondo sbocco per l'offerta PS. Si pensi, a titolo di esempio, al settore energia con dighe, oleodotti e gasdotti, eventualmente in aree a rischio, oppure ad impianti chimici la cui sicurezza è indispensabile garantire sotto ogni punto di vista.

Un terzo significativo segmento potenziale di mercato è legato alle Compagnie di Assicurazione, sia per quanto concerne il risarcimento di danni a beni immobili assicurati (per esempio, danni provocati dal cantiere relativo a un'opera pubblica come lo scavo per una linea di metropolitana), sia in sede di stipulazione di contratti e determinazione dei premi (richiede cautela assicurare, per esempio, una struttura che insiste su un terreno soggetto a deformazione che si traduce in una serie di rischi e costi aggiuntivi).

Peraltro, l'attività della legislazione italiana sembra possa dare un notevole impulso al mercato del monitoraggio di fenomeni di deformazione del suolo legati a problematiche assicurative (Legge 109, 11 febbraio 1994, articolo 30, in materia di obbligo di assicurazione di cantieri per opere pubbliche e Disegno di Legge 3326, articolo 8, in materia di introduzione nei contratti di assicurazione di beni immobili contro incendio, di clausole per la copertura del rischio sismico o derivante da altre calamità naturali).

Un quarto segmento di mercato è, infine, costituito da privati interessati a dati di deformazione relativi a singoli edifici o piccole aree (si citano, per esempio, le amministrazioni condominiali e le richieste di perizie di parte da usarsi in sede di vertenze giuridiche).

Vale la pena di osservare che il dibattito sulla possibile introduzione del Fascicolo del Fabbricato (una sorta di "cartella clinica" che ha lo scopo di certificare stabilità e "stato di salute" di ogni singolo edificio) è tuttora aperto e potrebbe tradursi in un letterale *boom* di questo ultimo segmento di mercato.

I dati di deformazione non sono, né si propongono, come sostitutivi di una perizia strutturale *in situ*, ma forniscono, senza dubbio, informazioni estremamente preziose per

chi si trova a dover garantire o meno la stabilità strutturale di un edificio.

L'analisi del mercato e, in particolare, una serie di incontri con potenziali clienti hanno permesso di mettere in luce che a partire da pochi essenziali *output* della catena di *processing* PS è possibile e agevole una vera e propria formattazione dei prodotti finali in ragione delle esigenze del cliente stesso ("customizzazione"). C'è, in definitiva, una forte domanda di dati di deformazione, coniugata con il desiderio che tali dati siano integrati nel sistema cartografico già in uso presso chi commissiona l'analisi.

Il prodotto PS è, quindi, generato a partire da una messe di dati puntuali di deformazione relativi al passato (trend medio di deformazione e serie temporale completa) che viene rappresentata in base alle esigenze del cliente, ad esempio, su cartografia tecnica regionale (CTR), su immagini ottiche ad elevata risoluzione (aerofotogrammetriche o satellitari), oppure, ed è questa la soluzione più efficace, in ambiente *GIS* (*Geographic Information System*), cioè nel contesto più ampio di un *database* che permette la consultazione interattiva di dati geografici, eventualmente anche di natura molto eterogenea.

La disponibilità di un archivio storico che raccoglie tutti i dati ERS a partire dal 1992 si è rivelata un elemento vincente che sovente garantisce una posizione di *leadership*: là dove non sono state condotte campagne di misura con altre tecniche, l'analisi PS (e l'interferometria più in generale) è l'unico strumento disponibile per misurare fenomeni di deformazione in atto, ottenendo dati estremamente preziosi al fine di quantificare il rischio attuale.

Vale la pena di osservare che non esistono veri e propri prodotti concorrenti. Le tecniche tradizionali di monitoraggio di deformazione (*in primis*, livellazione ottica e GPS) presentano caratteristiche diverse e, almeno in parte, complementari all'approccio PS. Peraltro, appare difficile trovare tecniche che, come i PS, permettono di studiare il movimento del terreno anche nel decennio passato!

Sono, quindi, auspicabili e prevedibili politiche di sinergia che possono realmente di-



schiodere nuovi orizzonti allo studio di movimenti della superficie terrestre.

La tecnica PS fornisce risultati unici, per accuratezza e densità spaziale, dei punti di misurazione ed è estremamente competitiva dal punto di vista economico, purché sia applicata su aree sufficientemente estese (da qualche decina di km² fino a scala regionale o addirittura nazionale).

È, infatti, necessario abbattere il costo dei dati grezzi che, in genere, ammonta a qualche decina di migliaia di Euro. I dati ERS vengono commercializzati sotto forma di "scene" da 100 × 100 km². Per applicazioni commerciali il costo di una scena è nell'ordine di qualche centinaio di Euro, nel caso di acquisto del numero elevato di scene tipicamente necessario a intraprendere un'analisi PS (in genere, da 40-50 fino a 150-200 qualora si desiderino combinare i risultati relativi a più geometrie di acquisizione).

Vale la pena di osservare che tali costi si riferiscono all'acquisizione dell'intero archivio storico (per esempio dal 1992) e vanno ripartiti sulle decine o centinaia di migliaia di punti di misura radar.

Si ottengono importi estremamente bassi rispetto a quelli che sarebbero necessari per rilevazioni GPS o di livellazione "a tappeto" (in effetti, condotti di rado), per i quali il costo di una singola misura presso un unico caposaldo è nell'ordine delle centinaia di Euro. Il servizio basato sulla tecnologia PS consiste nell'aggiornamento periodico dei dati di deformazione con i nuovi passaggi di ERS-2. A seconda delle esigenze del cliente, l'aggiornamento può essere effettuato ogni volta che si rende disponibile un nuovo dato (ogni 35 giorni) o con cadenza semestrale o annuale.

Problemi di assetto di ERS-2 durante l'anno 2001 hanno differito l'effettiva offerta del servizio di monitoraggio continuo. Si ha motivo di ritenere che tali difficoltà siano comunque in via di soluzione.

Vale, inoltre, la pena di osservare che ENVISAT (il cui lancio è previsto per il 1 marzo 2002) sarà presto in grado di garantire la continuità delle misure ERS. Sarà, inoltre, possibile ridurre da 35 a 3-4 quattro giorni il tempo di rivisitazione di uno stesso sito, aprendo, letteralmente, nuovi orizzonti all'approccio PS.

4. UN CASO APPLICATIVO: ANALISI DI DEFORMAZIONE NELL'AREA URBANA DI MILANO

Si ritiene interessante illustrare molto brevemente qualche risultato ottenuto nel corso di un'analisi di deformazione dell'area urbana di Milano commissionata dall'Agenzia Spaziale Europea.

L'analisi è stata svolta avvalendosi di 64 immagini radar acquisite dai sensori SAR ERS-1/2 tra maggio 1992 e dicembre 1999.

I principali obiettivi che ci si è prefissi all'inizio dello studio sono i seguenti:

- acquisizione di dati di deformazione su larga scala per la valutazione dell'eventuale impatto dell'innalzamento della falda freatica nell'area urbana milanese; fenomeno particolarmente accentuato verso sud-est (San Donato e San Giuliano) [12];

- acquisizione di dati puntuali di deformazione per identificare fenomeni che interessano singoli edifici o isolati;

- integrazione delle misure PS in ambiente GIS, per consentire all'utente di consultare e rappresentare i dati in base alle proprie esigenze. L'analisi ha consentito di individuare centinaia di migliaia di punti di misura affidabili (PS); la densità spaziale media è superiore 200 PS/km². In corrispondenza di ogni PS si ricava il valore di altimetria con un'accuratezza inferiore al metro. Ciò consente di georeferenziare la posizione dei PS (passaggio da coordinate SAR a coordinate geografiche) con un'accuratezza sufficiente a proiettare correttamente i PS su corrispondenti strutture e edifici.

Per ogni PS si è ricavato il trend medio di deformazione nell'intervallo temporale 1992-1999, attribuendo, inoltre al PS un parametro di qualità (con valore compreso tra 0 e 1) che riassume l'accuratezza (tra 0.1 e 1 mm/anno) del trend determinato.

Per i punti di misura ritenuti più interessanti si è ricostruita l'intera serie temporale di deformazione con accuratezza compresa tra 1 e 3 mm su ogni singolo dato.

I parametri relativi ai PS (coordinate SAR, coordinate geografiche, altimetria, trend medio di deformazione, parametro di qualità ed, eventualmente, serie temporale completa) sono stati raccolti in un *database* (.dbf) accessibile in ambiente ArcView GIS. I dati pos-

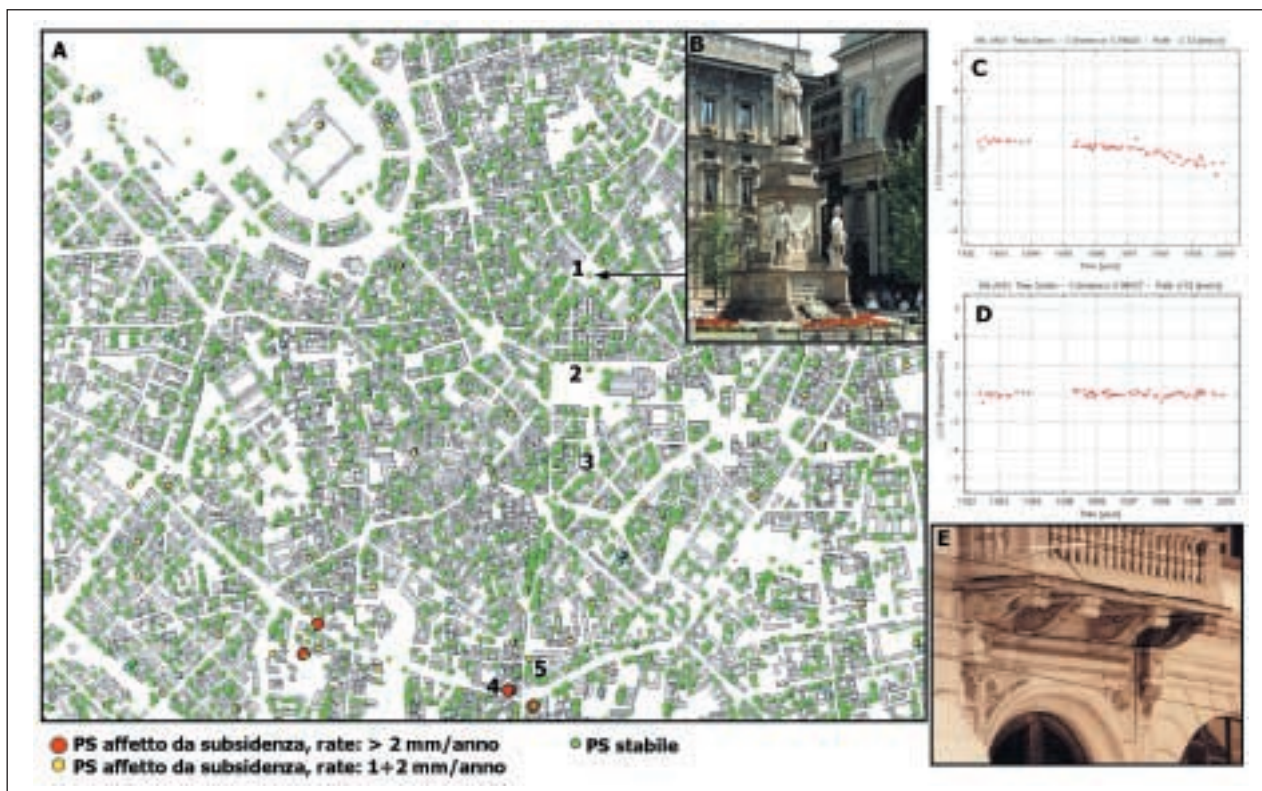


FIGURA 2

A Un tipico prodotto PS: il trend medio di deformazione in ambiente GIS. B Un Permanent Scatterer d'eccezione: il monumento a Leonardo da Vinci a Milano. C e D Serie temporali di deformazione. E Danni strutturali a un edificio a Milano

sono essere agevolmente sovrapposti ad altri layer, in particolare al tessuto urbano e al grafo stradale, in modo da rendere estremamente pratica la consultazione dei risultati dell'indagine PS (Figura 2 A).

L'analisi PS consente di affermare che non sono in atto significativi fenomeni di deformazione del terreno su larga scala. Il forte innalzarsi della falda freatica nel corso degli ultimi anni non ha (per il momento) innescato fenomeni di deformazione superficiale, quantomeno su aree vaste.

Ciò è perfettamente in linea con quanto ci si aspetta: si tratta della ricarica della falda freatica a valle di un intenso prelievo idrico protratto per anni che ha causato subsidenza irreversibile (compattamento inelastico [6]).

In figura 2 A è raffigurato il centro storico della città con i PS la cui deformazione ben si lascia riassumere in un trend medio. Si è fatto uso di dati GIS relativi al territorio del Comune di Milano [14], i PS sono stati sovrapposti alla cartografia urbana con le ri-

spettive velocità rappresentate tramite un codice colori.

Una quota significativa di edifici presenta caratteristiche di PS. Anche altre strutture, eventualmente molto particolari, possono costituire punti di misura radar affidabili: per esempio, il monumento a Leonardo da Vinci (1 in Figura 2 A) nel centro di Piazza alla Scala, oppure uno degli ingressi della metropolitana in piazza Duomo (2 in Figura 2 A), oppure, ancora, il monumento ai Carabinieri di Piazza Diaz (3 in Figura 2 A).

Nell'identificazione dei PS si è richiesta la coerenza di fase nell'intero intervallo 1992-1999. Per questo motivo non sono stati individuati PS in corrispondenza di edifici sottoposti a significativi lavori su facciata e tetto come il Teatro alla Scala. Analogamente, il Duomo mostra parecchi PS concentrati verso la facciata. L'abside e le navate laterali sono state, infatti, oggetto di protratti interventi di restauro nell'ultimo decennio. Gli interventi alla facciata, invece, sono stati

eseguiti negli anni '80, prima che cominciasse le acquisizioni ERS.

Chiaramente è possibile effettuare analisi locali per individuare PS tenendo conto di eventuali interventi di manutenzione che influenzano il comportamento dei bersagli radar.

Si osserva subito che buona parte dei punti affetti da movimento nel centro storico è ubicata lungo la cerchia dei Navigli, cioè lungo il corso del canale Naviglio, interrato ormai da diversi decenni.

Si riporta in figura 2 C la serie temporale di deformazione dell'edificio sull'angolo tra Corso Italia e Via Molino delle Armi (4 in Figura 2 A). Il trend medio di deformazione ammonta a -2.3 mm/anno. Vale però la pena di osservare che a partire dall'inizio del 1997 si assiste a un fenomeno di subsidenza ad evoluzione più rapida: si passa da -1.5 mm/anno a oltre -4 mm/anno.

L'edificio mostra, inoltre, evidenti danni strutturali, lasciando presumere che probabilmente è stato soggetto a deformazione anche precedentemente l'inizio delle acquisizioni ERS (Figura 2 E).

Il palazzo immediatamente di fronte (5 in Figura 2 A) è invece perfettamente fermo, come si riconosce dalla corrispondente serie temporale rappresentata in figura 2 D.

L'accuratezza delle misure è tale da rendere apprezzabili i fenomeni di dilatazione termica stagionale di singole strutture. Si rappresenta, a titolo di esempio, la serie temporale di deformazione dell'edificio in cui ha sede la Scuola Elementare di Via Massaua, nei pressi di piazza Tripoli (Figura 3 A – 3 D). Si riporta anche il grafico delle temperature medie nei giorni in cui sono state effettuate le acquisizioni ERS (Figura 3 B). Si osserva immediatamente la forte correlazione tra i due dati.

Una semplice analisi probabilistica consente di escludere che si tratti di un effetto casuale, anche alla luce del fatto che sono decine gli edifici che presentano analoghi fenomeni di dilatazione termica. Nel caso riportato è stato addirittura possibile inferire dalle misure satellitari e dai dati di temperatura il materiale con cui è stata realizzata la parte di struttura cui corrisponde il PS in esame.

Dalle misure di deformazione si ricava l'entità del fenomeno di dilatazione; dai valori

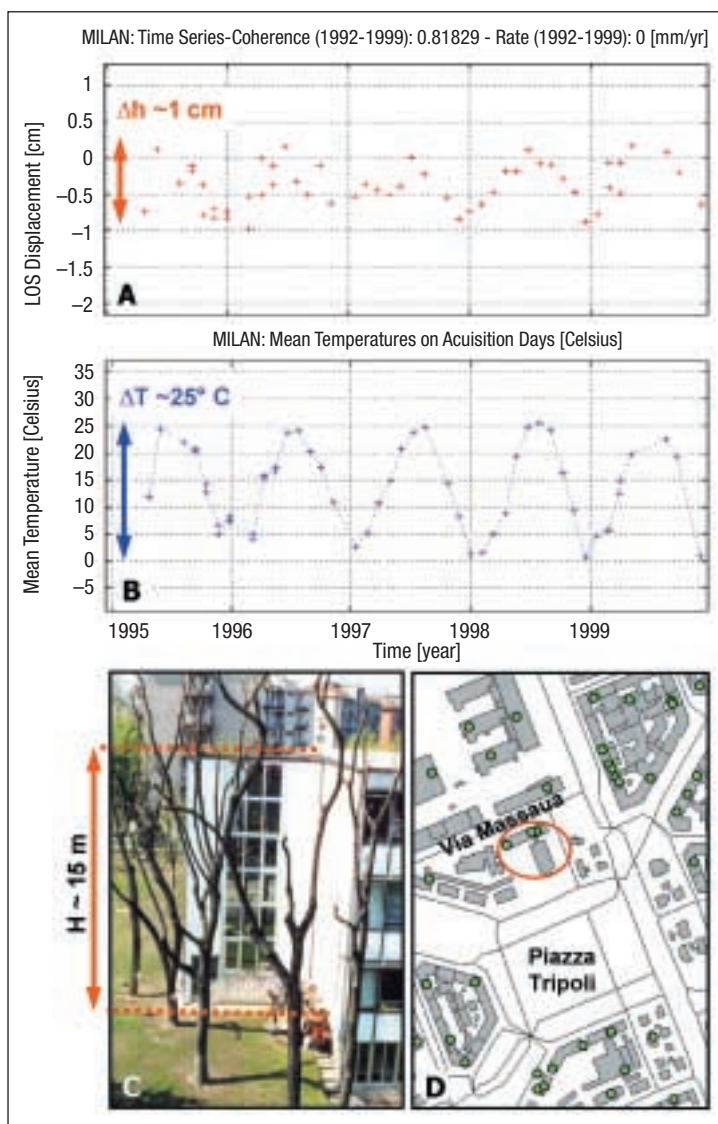


FIGURA 3

A Serie temporale di deformazione: evidente il "respiro" della struttura dovuto alla dilatazione termica stagionale. B Temperatura media nei giorni di acquisizione ERS. C Fotografia della struttura in alluminio. D Rappresentazione in ambiente GIS dei PS in prossimità della Scuola Elementare di Via Massaua.

di quota accurati (disponibili presso i PS nell'intera area circostante) si stima l'altezza dell'edificio. Note le temperature, si è in grado di stimare il coefficiente di dilatazione termica del materiale, che, nella fattispecie è compatibile solo con l'ipotesi di una struttura in alluminio.

Un sopralluogo, effettuato a posteriori (Figura 3 C), ha confermato la correttezza delle deduzioni. Chiaramente si tratta solo di un risultato dimostrativo; è sufficiente che parti

della struttura siano verniciate di nero per far sì che ci sia una notevole differenza tra la temperatura dell'aria e quella della struttura stessa, in particolare nelle giornate estive.

Ringraziamenti

Gli autori sono profondamente grati all'Agenzia Spaziale Europea (ESA) che ha fornito i dati ERS e ha finanziato il progetto di analisi PS su Milano. Si ringrazia inoltre il Prof. C. Prati e tutto lo staff T.R.E., in particolare gli Ingegneri R. Locatelli, F. Novali, A. Menegaz, M. Basilico, S. Cespa e D. Colombo. Infine, un grazie anche all'Arch. E. Fiorina per le fotografie.

Bibliografia

- [1] AAVV: *Dizionario Enciclopedico Italiano*. Istituto dell'Enciclopedia Italiana, Roma, 1955.
- [2] Acceleratore di Impresa del Politecnico di Milano – Sito Web: http://www.cpi.polimi.it/it/acceleratore/pre_acceleratore.htm
- [3] Ferretti A, Prati C, Rocca F: Permanent Scatterers in SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, n. 1, 2001, p. 8-20.
- [4] Ferretti A, Prati C, Rocca F: Non-linear Subsidence Rate Estimation Using Permanent Scatterers in Differential SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, n. 5, 2000, p. 2202-2212.
- [5] Freeze RA: *Social Decision Making and Land Subsidence*. In: Proceedings of the 6th International Symposium on Land Subsidence, 2000, Vol. I, p. 353-384.
- [6] Galloway D, Jones DR, Ingebritsen SE (Ed.): *Land Subsidence in the United States*. U.S. Geological Survey, Circular 1182, 1999.
- [7] Hanssen RF: *Radar Interferometry. Data Interpretation and Error Analysis*. Kluwer Academic Publishers, Dordrecht, 2001.
- [8] Jet Propulsion Laboratory (JPL), Shuttle Radar Topography Mission (SRTM), - Sito Web: <http://www.jpl.nasa.gov/srtm/statistics.html>
- [9] Liddel Hart BH: *Storia Militare della Seconda Guerra Mondiale*. Mondadori Verona, 1970.
- [10] Mensa DL: *High Resolution Radar Cross-Section Imaging*. Artech House, Norwood, 1991.
- [11] Nebeker F: The Electric Century.. *IEEE Spectrum*, n. 6, 2000, p. 68-74.
- [12] Raffaelli L, Raimondi P, Rosti G: *Finanziamento di interventi prevedibili per affrontare i problemi connessi al sollevamento della falda a Milano*. 1997. http://www.provincia.milano.it/ambiente/progettispeciali/pub/cap_97_f.pdf

[13] Taylor LS: *A Thumbnail History of Electronics*. Electrical Engineering Department, University of Maryland. <http://www.ee.umd.edu/~taylor/Electrons.htm>

[14] Urban Explorer, realizzato da Assimpredil in collaborazione con il Comune di Milano, - Siti Web: <http://www.assimpredil.ance.it>
<http://www.comune.milano.it>

FABIO ROCCA professore ordinario di Elaborazione Numerica dei Segnali presso il Politecnico di Milano. Membro del Comitato ASAR (Advanced Synthetic Aperture Radar) dell'ESA (Agenzia Spaziale Europea). Coordina per l'ESA un Expert Support Laboratory nel quadro delle attività scientifiche di sostegno dei satelliti di telerilevamento ERS-1/2 ed ENVISAT. Associate Editor delle riviste *Signal Processing* e *Journal of Seismic Exploration*. Premio Italgas per le Telecomunicazioni del 1995. Premio 1999 (condiviso con L. Chiariglione) per le Tecnologie della Rhein Foundation, RFT. Laurea Honoris Causa (2001) in Geofisica dell'Institut Polytechnique de la Lorraine, Nancy. Socio fondatore di Tele-Rilevamento Europa.
e-mail: rocca@elet.polimi.it

CARLO COLESANTI consegue la laurea in Ingegneria delle Telecomunicazioni nell'aprile 1999. Attualmente è dottorando presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano, dove svolge ricerca in materia di tecnica dei diffusori permanenti e di sistemi SAR passivi. Collabora con Tele-Rilevamento Europa fin dalla nascita della società.
e-mail: colesant@elet.polimi.it

ALESSANDRO FERRETTI si laurea in Ingegneria Elettronica nel 1993. Nello stesso anno consegue il Master in Information Technology presso CEFRIEL. Nel 1997 conclude il Dottorato di Ricerca presso il Politecnico di Milano. Svolge lavoro di ricerca nel settore dell'elaborazione di dati SAR. Si è occupato della ricostruzione di modelli digitali di elevazione con tecniche interferometriche. Insieme ai Prof. C. Prati e F. Rocca ha ideato e sviluppato la tecnica dei diffusori permanenti. È socio fondatore e amministratore delegato di Tele-Rilevamento Europa.
e-mail: aferrett@elet.polimi.it oppure
e-mail: alessandro.ferretti@treuropa.com

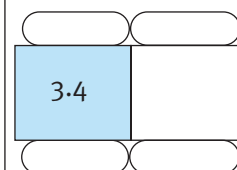
RAFFELLA RATTI si laurea in Ingegneria delle Telecomunicazioni nel 2000. Svolge attività di ricerca presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano e presso Tele-Rilevamento Europa. In particolare, si occupa di phase unwrapping e ricostruzione di modelli altimetrici digitali a partire da dati SAR.
e-mail: raffaella.ratti@treuropa.com

IL PASSAGGIO DALLA BANDA AI SERVIZI



La disponibilità di connettività ad elevate prestazioni e a costi decrescenti, nonché la standardizzazione delle architetture intorno al protocollo IP e al modello Web-Browser, ha creato una nuova idea di business chiamata ASP. Il modello ha generato entusiasmi e investimenti colossali, a cui, nella logica della “new economy”, sono seguite le immancabili disillusioni. Tutto sbagliato? No, semplicemente l’ASP è il punto di arrivo, e non il punto di partenza, nella catena del valore dei servizi offerti da un Internet Service Provider.

Giacomo Zanotti
Giulia Niccolai



1. INTRODUZIONE

C'era una volta il tempo in cui la banda era considerata un bene preziosissimo. L'economia della rete, più conosciuta come *net economy*, doveva ancora assumere dei contorni definiti e la sua diffusione sembrava fortemente legata alle capacità trasmissive dei vari operatori sul mercato. Era la seconda metà degli anni Novanta, periodo in cui si affermava Internet, e con esso si celebrava un'importantissima evoluzione concettuale nel panorama delle telecomunicazioni: il passaggio dal puro trasporto, in forma dedicata o commutata, a un'architettura complessa in grado di costituire una rete planetaria dalle potenzialità incommensurabili. In questo scenario proliferavano gli *Internet Service Provider* (ISP), fornitori di banda e connettività che, attraverso circuiti e indirizzi IP (*Internet Protocol*), portarono Internet, in poco tempo, non solo negli uffici ma anche nelle prime abitazioni.

A interrompere la fortuna di tali operatori, oltre la liberalizzazione delle telecomunicazioni che rese il settore molto affollato e alta-

mente competitivo, furono in un certo senso i ritmi dell'innovazione tecnologica. L'evoluzione dei processori e dei sistemi operativi, l'introduzione di applicativi di tipo *client-server*, la nascita di nuovi dispositivi *palm* e in particolare la legge di Gilder, che postulava il triplicarsi annuo della banda per i successivi 25 anni, furono tutti elementi che richiamarono l'attenzione sulla necessità di riconsiderare la banda come una *commodity* e di spostarsi su nuovi servizi.

2. LA SOLUZIONE ASP

Alla luce dei repentini mutamenti tecnologici e delle spinte di mercato, molti Internet Service Provider si riquificarono alla fine degli anni '90 come *Application Service Provider* (ASP), ovvero, erogatori di “software a consumo”. Nulla fu mai più confuso, nella storia dell'*Information Communication Technology* (ICT), della definizione di ASP. Per questo motivo si riportano le specifiche di dettaglio, così come erano presenti nella formulazione originale:

il software diventa un puro servizio: il cliente non conosce l'applicazione ma solo il servizio da essa erogato;

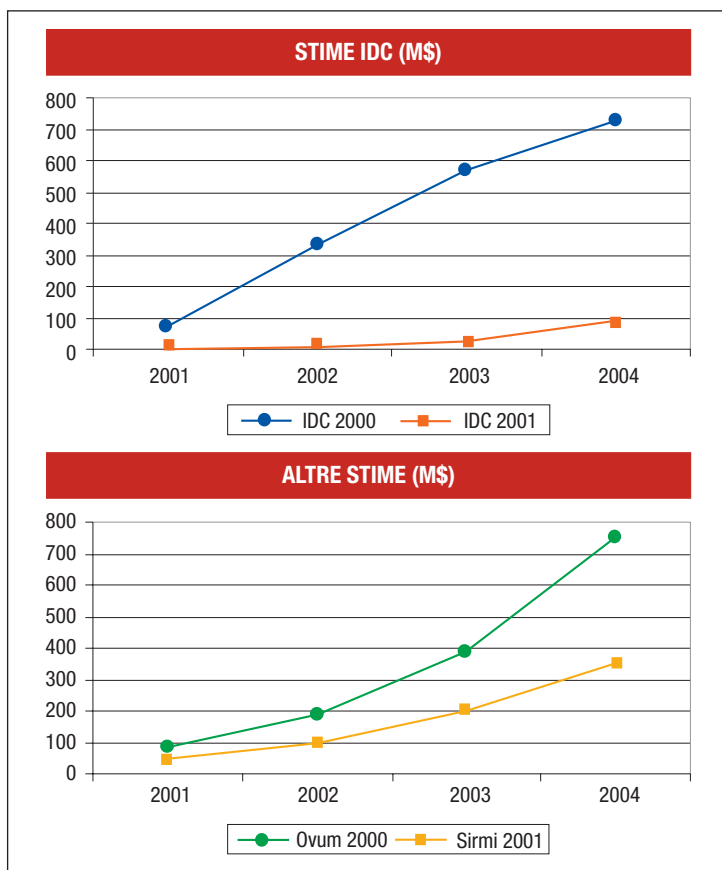
il cliente non possiede l'applicazione, ma la paga sulla base dell'effettivo consumo: gestione, aggiornamento e manutenzione dell'applicazione sono, quindi, a carico dell'ASP;

l'accesso all'applicazione deve essere possibile da qualunque postazione di lavoro dotata di un browser e di un accesso a Internet;

il servizio viene distribuito in logica *one-to-many*, per realizzare economie di scala.

Le ragioni che spinsero gli operatori ad orientarsi in questa direzione furono, prevalentemente, legate a uno studio sulle nuove tendenze aziendali: da un lato, le grandi aziende sembravano ben disposte a dare in *outsourcing*, la gestione di parti dei propri sistemi informativi aziendali, dall'altro, anche le PMI (*Piccole Medie Imprese*) sembravano dare i primi segnali di recepimento dei mutamenti tecnologici in atto, ampliando così, significativamente, il mercato della nuova idea di business.

FIGURA 1
Il mercato ASP,
stime a confronto
(fonte: IDC)



A supportare il concetto del *software as a service* contribuivano, inoltre, gli innumerevoli vantaggi legati sia alla **tecnologia thin client**, che rende fruibile ogni applicativo per mezzo di un semplice browser, sia alla connettività 24×7 , che garantisce assistenza continua *on-line*. La modalità *one-to-many*, cioè l'erogazione di un software condiviso da molti, ad alto livello di standardizzazione, avrebbe, inoltre, consentito al crescere dei clienti acquisiti l'instaurarsi di economie di scala vantaggiose sia per gli ASP stessi in termini di costi, sia per i clienti in termini di tariffa mensile. Sull'onda di questo entusiasmo nacquero molte forme di ASP, corrispondenti a diversi modelli di business: gli ASP orizzontali, specializzati nell'erogazione di servizi *cross-industry* (posta, messaggistica, storage,...), gli ASP verticali (legati ai processi di business di un particolare segmento di mercato), i fornitori di Infrastruttura (AIP, *Application Infrastructure Provider*, evoluzione naturale dei tradizionali Internet Service Provider e legati ai servizi di *hosting* e *housing*).

3. DALL'ENTUSIASMO ALLA DISILLUSIONE

In un clima di generale euforia legata agli andamenti della cosiddetta *new economy*, l'Application Service Provisioning si diffuse rapidamente e assunse svariate forme, in alcuni casi, come ampliamento di servizi dei precedenti ISP, in altri, come *core business* di coraggiose *start-up*. Nel frattempo, stime di mercato preannunciavano che il mercato ASP avrebbe raggiunto nel 2004, solo in Europa, i 700 milioni di dollari con tassi di crescita annuale vicini al 100% (fonte: IDC) (Figura 1).

L'inizio del nuovo millennio non deluse le aspettative, ma nel 2001 dilagò la crisi. Molti fra i principali ASP statunitensi dichiararono fallimento, avvennero numerose acquisizioni da parte di *Telco* e *Hosting Service Providers* (HSP, ovvero operatori specializzati nella gestione di Data Centres con servizi di *housing* ed *hosting*), i modelli di business, precedentemente focalizzati, sul *pure application provisioning* si ampliarono e differenziarono. Anche gli analisti di merca-

to corressero il tiro, mettendo chiaramente in evidenza come il passaggio dall'entusiasmo alla disillusione fu quanto mai traumatico. Alcune stime a confronto dimostrano quanto affermato: a distanza di un semestre esse divergono di quasi un ordine di grandezza!

4. LE RAGIONI DEL FALLIMENTO

Se tante erano le aspettative e le speranze legate al modello ASP quali sono state, dunque, le ragioni del fallimento? Senza dubbio un elemento chiave, che ha peraltro influenzato negativamente l'andamento di tutti i mercati della new economy, è stata l'esplosione della bolla speculativa creata attorno a Internet. Oltre a questo, comunque, da un'attività di monitoraggio del mercato condotta a livello mondiale (dicembre 2001-aprile 2002), emergono altri fattori coinvolti.

❑ In molti casi, il problema sostanziale è stato nella scelta degli applicativi. I primi ASP infatti hanno tentato di commercializzare soluzioni legate ai processi core aziendali quali sistemi ERP (*Enterprise Resource Planning*), CRM (*Customer Relationship Management*), SCM (*Supply Chain Management*) per la grande impresa, non tenendo conto, da un lato, del timore di perdere il controllo sugli applicativi, dall'altro dei rischi legati a eventuali incidenti.

❑ In altri casi, è stata la rete di vendita a non funzionare. Le aziende, che hanno assunto come *target* di riferimento le PMI, molto spesso, non hanno saputo trovare un adeguato contatto con il cliente: la vendita via web è risultata più che mai utopistica, mentre gli accordi con i rivenditori non hanno portato i risultati attesi.

❑ Altra questione concerne gli investimenti. Molti ASP hanno ritenuto opportuno dotarsi di infrastrutture per accaparrarsi anche la gestione dell'hosting e della connettività, sottostimando l'incidenza di tali investimenti sulla profittabilità e innescando una controproducente competizione con altri operatori.

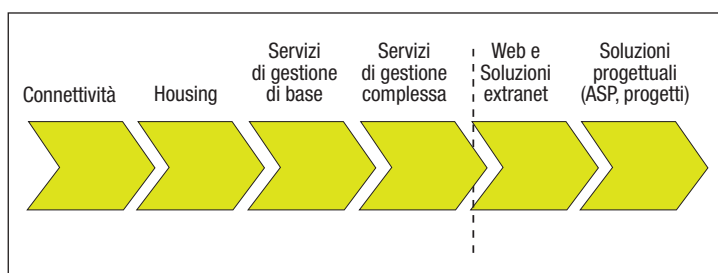
❑ Infine, anche un evidente ostruzionismo da parte dei *software vendor*, timorosi di perdere il contatto con il cliente finale, ha contribuito alla caduta del modello ASP.

Se questi sono stati i fattori contingenti che hanno, di volta in volta, determinato il fallimento di diversi progetti imprenditoriali, l'elemento più rilevante e, a parere di molti, maggiormente sottovalutato, è senza dubbio di tipo culturale. Chi ha ipotizzato un naturale passaggio dall'internalizzazione alla completa esternalizzazione del sistema informativo aziendale non ha fatti i conti, prima di tutto, con resistenze di tipo psicologico e umano. Chi ha creduto che la sola garanzia di affidabilità contenuta entro SLA (*Service Level Agreement*) a *five nines* (99,999%), peraltro, tutt'oggi difficile da realizzare, riuscisse a innescare un rapporto fiduciario fra committente e operatore tanto forte da consentire la totale delega gestionale dell'IT (*Information Technology*) aziendale, non ha considerato l'importanza della gradualità nelle scelte. In sintesi, si potrebbe affermare, in modo estremamente prosaico, ma efficace, che l'errore fondamentale è stato quello di partire... dal punto di arrivo! Un punto di arrivo che potrebbe anche non venire mai raggiunto. È, infatti, molto probabile (come si vedrà nei paragrafi successivi), che si affermi un nuovo modello di business, basato su un equilibrio tra servizi Web e processi eseguiti localmente sui server aziendali (in altri termini, un'estensione dell'attuale *extranet*).

Ed è proprio in risposta a una tale esigenza che si affermano gli xSP (*Extended Service Provider*). La catena del valore, rappresentata in figura 2, mostra il graduale processo di esternalizzazione richiesto prima di giungere all'erogazione del software a consumo.

L'xSP è una forma estesa di erogazione di servizi attraverso la rete. Il punto di *break* (rappresentato dalla linea verticale) è variabile e rappresenta dunque il livello di "asp-izzazione" del servizio offerto.

FIGURA 2:
XSP, la catena del valore



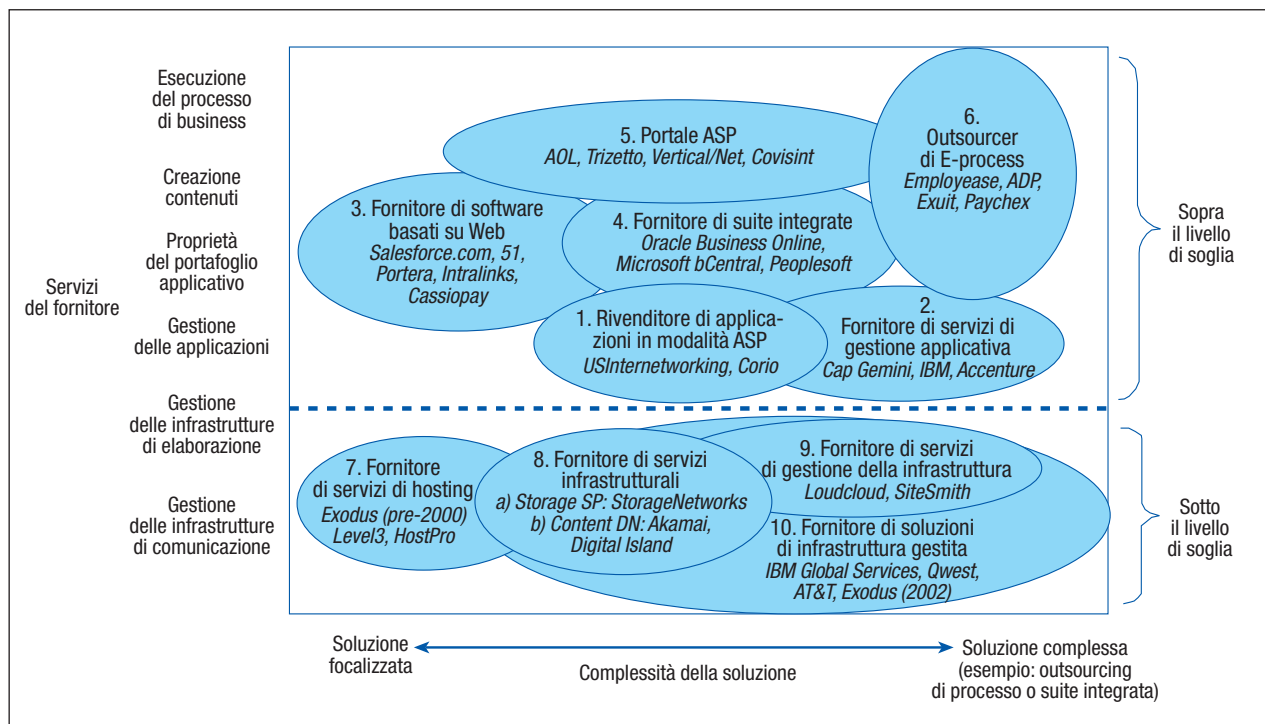


FIGURA 3 5. GLI ATTORI SUL MERCATO

*Il mercato XSP:
la visione IBM
(fonte:IBM)*

La fornitura di servizi gestiti implica non solo l'esigenza di un'infrastruttura di rete, ma anche il possesso o l'accesso a un Internet Datacenter e, eventualmente, di software da erogare. Da ciò, si deducono facilmente gli attori coinvolti nel settore, che sono i *Network Service Provider* (NSP), gli *Hosting Service Provider* e i *Software Vendor*. La diversa combinazione di ruoli e competenze dei tre soggetti dà vita a una miriade di soggetti economici differenziati, di caso in caso, per tipo di servizi offerti, per strutturazione dei canali di vendita, e per ripartizione dei ricavi. L'attuale condizione del mercato internazionale è ben rappresentata da IBM e IDC (Figure 3 e 4).

Di particolare interesse, in questo contesto instabile e quanto mai complesso, è il processo evolutivo che alcune tipologie di soggetti hanno compiuto per superare la crisi. Gli *ASP pure player*, nati come interfaccia cliente per l'erogazione di soluzioni integrate di software, connettività e soluzioni di hosting gestito, hanno avuto non pochi problemi a rimanere in vita. Gli elementi di differenziazione che hanno risollevato le sorti delle principali aziende a livello internazionale sono stati:

■ le partnership, che hanno consentito di portare sul mercato soluzioni integrate di connettività, hosting e applicazioni, senza richiedere investimenti insostenibili (per esempio, Corio);

■ l'integrazione di servizi ASP con servizi a valore aggiunto (ad esempio, Netstore con soluzioni di *back up* e Document Management, DM);

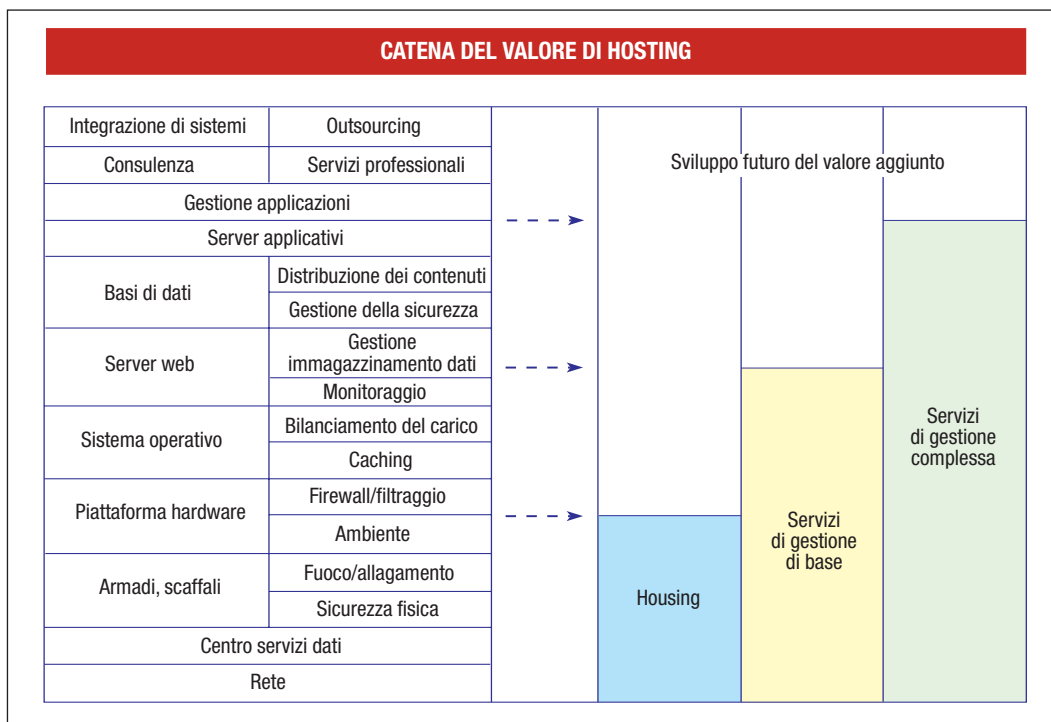
■ la concentrazione su settori verticali (per esempio, portali con software gestionali per studi professionali);

■ la specializzazione in determinate aree critiche di gestione aziendale (per esempio, Salesforce.com per applicazioni di Sales Force Automation e CRM);

■ la concentrazione su soluzioni di *e-business* (dai *marketplace* ai sistemi di *intelligence*).

Riguardo a IT player e Telco, l'errore più ricorrente è stato quello di estendere, autonomamente, il proprio business alla fornitura di applicativi. Il fallimento del colosso USInternetworking è stato esemplare a proposito.

Molto più oculata è stata, invece, la scelta di quegli operatori che hanno concentrato il proprio business sull'efficienza dei puri servizi, in particolare, su quelli di supporto al-

**FIGURA 4**

*Il mercato XSP:
la visione IDC
(fonte: IDC)*

l'erogazione del software a consumo quali:

■ connettività a banda garantita (per esempio, un ERP "leggero" richiede circa 12 kbit/s per user);

■ SLA superiori al 98,5% (99,95% per applicativi *mission critical*);

■ sicurezza, non solo di ambienti, ma anche costante monitoraggio anti-intrusione;

■ scalabilità dei sistemi.

Digex, per esempio, che ha creato un piano di *partnership* specifico per gli ASP (*app-link program*) la cui convenienza aumenta al crescere degli utenti connessi) rappresenta, senza dubbio, un caso di successo.

Anche le software house e i software vendor hanno subito un processo interessante. Superato il momento in cui esse vedevano con estremo scetticismo il rilascio delle licenze agli ASP puri, attualmente, le tendenze di mercato sembrano essere diverse. In alcuni casi, l'operatore ASP è visto come un nuovo reseller, in grado di penetrare nuove fette di mercato, in altri le *software house* stesse entrano nel mercato ASP proponendo la propria soluzione come alternativa per i loro clienti. Ancora oggi, non risulta chiaro se e in quali ambiti il contatto diretto o l'intermediazione risulterà conveniente. Certo è che, se si riuscirà a trovare un equilibrato model-

lo di ripartizione dei ricavi fra le parti, ci saranno buone prospettive di profitti per tutti gli attori in gioco.

6. IL MERCATO IN ITALIA

Così come nel contesto internazionale, anche in Italia il mercato ASP sembra, ancora, piuttosto lontano dalla fase di consolidamento. Gli operatori coinvolti sono molteplici e con modelli di business molto diversi fra loro. L'impressione ricorrente, maturata a fronte di una serie d'indagini sul campo e non, è che ogni singolo attore si sia predisposto a soddisfare un quadro molto ampio di esigenze, non solo per gli applicativi, ma anche per servizi gestiti, in attesa di ricevere dal mercato *feedback* significativi sulla profittabilità delle varie iniziative.

Tranne qualche sporadica iniziativa di verticalizzazione (per esempio, Axioma per il gestionale sul turismo o Netesi per l'automazione di pratiche comunali) o di specializzazione (per esempio, Zucchetti per la gestione dei cedolini paga), la cui validità è ancora tutta da dimostrare, l'offerta ASP ruota ancora molto attorno ai gestionali ERP, ai sistemi di CRM e ai programmi di Office Automation. Sul versante web, l'offerta

di partecipazione a marketplace attraverso pagine web preconfigurate e predisposte ad attività di *e-commerce* è assolutamente dominante.

Generalmente, quasi tutti gli operatori si propongono come interfaccia unica per la fornitura non solo di applicativi a consumo, ma anche di tutti i servizi a monte sulla catena del valore. A questo proposito, l'atteggiamento predominante sembra quello di costituire partnership strategiche allo scopo di concentrare il core business in quegli ambiti dove il *know-how* è maggiormente consolidato.

Fra gli ASP *enabler* (fornitura d'infrastruttura) e gli erogatori del servizio, ciò che rimane in molti casi da stabilire è chi curerà il contatto con il cliente, tenendo conto della complessa questione di come non ledere il business tradizionale delle software house con il passaggio alla logica ASP.

7. PROSPETTIVE FUTURE

Superata la fase d'entusiasmo che ha indotto a sovrastimare il valore di un settore, fin troppo innovativo e compreso che esistono, comunque, degli elementi per la ripresa, ci si chiede allora quale sia il giusto approccio nell'affrontare oggi il tema dell'ASP.

Tornando all'affermazione fatta inizialmente, l'errore del modello ASP è stato quello di concepire un modello di *business on-off*, pensando di passare da una situazione in cui il cliente possedeva tutte le risorse (*hardware*, software di base, software applicativo e personalizzazione progettuale), a una forma a consumo totalmente spersonalizzata. E allora? Semplice, si deve pensare a ciò che serve veramente all'impresa e non ai modelli teorici e ai relativi *business plan*, ancora più teorici.

Il primo passo verso l'ASP è stato l'esternalizzazione di siti Web (in forma più o meno estesa, ossia dal puro housing all'outsourcing totale). È il primo fondamentale processo volto a sfruttare le potenzialità della rete e dei servizi di *storage* di un Network Service Provider: il Web è la "vetrina" dell'azienda, creata per aumentare il mercato indirizzabile. Non fa parte del core business aziendale; deve essere aperta e ac-

cessibile da qualunque postazione connessa alla rete planetaria 24 h al giorno; deve essere - per ragioni di sicurezza - il più possibile disgiunta dagli altri processi aziendali (a parte semplici funzioni di import di cataloghi dei prodotti/servizi): l'esternalizzazione in logica ASP è, dunque, la forma più conveniente.

Nel secondo passo, così come è stato concepito dalla quasi totalità degli attori, si annida la causa principale degli insuccessi. Infatti si è pensato di associare *ex-abrupto*, alla pura vetrina, l'erogazione dei vari processi aziendali: contabilità, logistica, CRM, reportistica, strumenti di produttività individuale, *workflow* di procedure e, infine, commercio elettronico. A costo di apparire disacranti, si affermerà che, proprio il passaggio al commercio elettronico e al mercato digitale, è stata la causa più clamorosa degli insuccessi. Nelle applicazioni di commercio elettronico confluiscono, infatti, oltre alle tipiche difficoltà presenti nella logica ASP, anche forme completamente nuove di rapporto tra fornitori e venditori.

Il vero secondo passo da compiere è quello insito nelle origini di Internet. Internet non è solo una biblioteca universale ma anche, e soprattutto, un luogo per incontrarsi e condividere delle informazioni. Il passaggio successivo alla pura vetrina consiste nella creazione di un ambiente collaborativo in cui il Web diventa l'archivio condiviso dei documenti che vengono prodotti dai vari processi aziendali. Ogni processo aziendale può essere visto come un insieme di attività elaborative, che continuano a essere locali all'azienda, e un insieme di *output* documentali, che ne costituiscono il risultato e che debbono essere pubblici e condivisi. Questi ultimi possono essere utilmente archiviati, classificati e distribuiti attraverso il Web, creando un'interazione più forte tra processi aziendali, Web, rete di comunicazione e attori interessati al processo specifico. Il Web diventa quindi *backup*, archivio strutturato, punto di accesso, *workflow* documentale, *gateway* tra forme diverse di comunicazione (fax, posta elettronica, SMS, fisso, mobile). Non sembra opportuno fare esempi perché limitativi: ogni processo aziendale, in ogni caso, creando documenti "pubblici" (ossia, consultabili e condivisibili

tra più individui) può, dunque, aderire a questo modello. Si sta risalendo, in questo modo, la catena del valore, ma in maniera lenta e graduale, senza alterare gli equilibri innati nei vari processi di business aziendali, ma soprattutto sfruttando il vero valore di Internet e dei relativi *Internet Data Center* (IDC).

E dopo? A questo punto, il processo di estensione è naturale. Si possono cominciare a considerare i documenti non come entità statiche, ma attive, ossia partecipi di un processo di business. Gli esempi vanno dai servizi di Web-call center (dalla vetrina all'interazione con i consulenti aziendali), all'autenticazione del documento per mezzo di firma digitale, al riconoscimento del suo contenuto per una validità nell'ambito di una transazione commerciale. Se il documento si chiamerà "ordine" o "fattura" si è arrivati al commercio elettronico, se il documento è l'*output* di un processo produttivo e si dà all'utente la possibilità di interagire mediante un sistema di *trouble ticketing*, allora, si è ASP-izzato una parte di ERP.

Quando e se mai si arriverà a queste forme è un'incognita. Già da ora, usare Internet come punto di incontro del business e chiamarlo ASP è una realtà.

Bibliografia

- [1] Allaboutasp: <http://www.allaboutasp.com>. *Il portale del consorzio industriale ASP*.
- [2] Application Service Providers: Market strategies for Telco and ISP's. *Report Ovum* 2000.
- [3] Aspisland: <http://www.aspisland.com>
- [4] Aspnews: <http://www.aspnews.com>
- [5] Aspstreet: <http://www.aspstreet.com>
- [6] Deconstruct and Rebuild: Winning ASP Strategy for Europe. *Report IDC* 2001.
- [7] European ASP Services, Review and Forecast, 2000-2005. *Report IDC* 2001.
- [8] European Web Hosting Services, 2000-2005. *Report IDC* 2001.
- [9] Idc: <http://www.idc.com/enewsletters>.
- [10] *Il mercato Italiano dell'ICT*. Sirmi, Convegno Assintel 2001.
- [11] Internet: <http://www.internet.com>
Una bibbia per tutto ciò che è internet-oriented.
- [12] Market forecast and dynamics-XSP. *Report Butler* 2001.
- [13] Web Hosting: Usage and Markets. *Report Ovum* 2001.
- [14] Xspstreet: <http://www.xspstreet.com>
Portali con informazioni sul mondo ASP, analisi di mercato, forum di discussione, newsletter periodiche gratuite.

GIULIA NICCOLAI laureata in Scienze della Comunicazione con una tesi sul ruolo dell'innovazione tecnologica per lo sviluppo territoriale. Ha frequentato da marzo 2001 a marzo 2002 il Master in Convergenza e Società dell'informazione del Politecnico di Milano. Durante questo periodo ha seguito un project work dal titolo 'Lo sviluppo di un'offerta convergente nel settore degli Application Service Provider presso Wind-Infostrada. Attualmente segue un progetto internazionale di Datawarehouse in Accenture per un'azienda di largo consumo e-mail: g.niccolai@libero.it

GIACOMO ZANOTTI laureato in Ingegneria elettronica al Politecnico di Milano, ha seguito dal nascere il tema dell'ICT. Ha ricoperto incarichi di direttore marketing e di responsabile di Business Unit di system integration in primarie aziende di informatica e telecomunicazioni. Attualmente collabora in Itnet (Gruppo Wind) alla definizione delle strategie di sviluppo business. Ha interesse all'ambiente accademico e tecnico-scientifico, che sviluppa come membro del consiglio AICA, come membro della commissione Europea IST-prize e con attività di docenza in Master organizzati dal Politecnico di Milano.
e-mail : g.zanotti@it.net