

Alla ricerca della nuova relazione tra emozioni, umanità e intelligenza artificiale.

di: [PIETRO MONARI](#) e [LUCA BARRETTA](#)

Sommario

Questo articolo esamina l'intersezione tra emozioni umane e intelligenza artificiale (AI), analizzando come l'interazione uomo-macchina influenzi la comprensione delle emozioni e del pensiero. Attraverso una prospettiva interdisciplinare, si esplorano le sfide etiche e scientifiche dell'integrazione delle emozioni nell'AI, criticando modelli come quello di Ekman. Successivamente, si analizza l'evoluzione dei chatbot, da ELIZA a ChatGPT, per approfondire la ridefinizione della relazione uomo-macchina. Infine, si sottolinea l'importanza di un'educazione emotiva che preservi la complessità dell'esperienza umana nell'era dell'AI, promuovendo un approccio critico e consapevole all'uso della tecnologia.

Abstract

This article examines the evolution of the study of human emotions and the development of artificial intelligence (AI), highlighting how the interaction between humans and machines is influencing our understanding of emotions and thought processes. Through an interdisciplinary perspective, it explores the ethical and scientific challenges of integrating emotions into AI systems, critically assessing theoretical models like that of Paul Ekman. The discussion reflects on the potentials and limitations of current systems and emphasizes the importance of emotional education to preserve the complexity of human experience in the age of AI.

Parole chiave

emotions, Artificial Intelligence, Human-Machine Interaction, chatbot, technological awareness

1. Introduzione

Lo studio delle emozioni umane è un campo di ricerca vasto e multidisciplinare, la cui evoluzione riflette la complessità del suo oggetto di studio. Le radici di questa indagine affondano nella filosofia greca antica, dove pensatori come Platone e Aristotele consideravano le emozioni elementi essenziali dell'esperienza umana. Questa riflessione ha gettato le basi per secoli di dibattito sul ruolo delle emozioni nella vita e nella morale, influenzando profondamente le successive indagini scientifiche.

Con l'avvento della psicologia come disciplina scientifica nel XIX secolo, l'interesse per le emozioni ha subito una significativa trasformazione. Pionieri come William James e Carl Lange hanno spostato l'attenzione dalla speculazione filosofica all'indagine empirica, aprendo la strada a uno studio scientifico delle emozioni. Questo cambio di paradigma ha portato, nel XX secolo, a una proliferazione di approcci diversificati, ognuno dei quali ha contribuito a illuminare differenti aspetti dell'esperienza emotiva.

La psicoanalisi di Freud, ad esempio, ha evidenziato l'importanza delle emozioni inconsce nel plasmare il comportamento umano, mentre il behaviorismo si è focalizzato sulle risposte emotive osservabili e misurabili. Questi approcci, seppur contrastanti, hanno contribuito a una comprensione più sfaccettata delle emozioni, mettendo in luce sia i loro aspetti interiori che le manifestazioni esterne. L'avvento della psicologia cognitiva ha segnato una svolta significativa, introducendo l'analisi dell'interazione tra processi cognitivi ed emozioni. Questa prospettiva ha superato la tradizionale dicotomia tra ragione e sentimento, offrendo una visione integrata del funzionamento mentale. Esplorando come pensieri e interpretazioni influenzino le risposte emotive, ha gettato nuova luce sulla complessità della mente umana.

Parallelamente, le neuroscienze hanno fornito profonde intuizioni sui meccanismi cerebrali sottostanti l'esperienza emotiva. Questa prospettiva biologica ha arricchito la nostra comprensione, offrendo una base fisiologica alle teorie psicologiche e filosofiche sulle emozioni. L'integrazione di questi filoni di ricerca ha

portato a una visione olistica delle emozioni, riconoscendole come fenomeni complessi che coinvolgono mente, corpo e ambiente.

In questo contesto, la sociologia e l'antropologia hanno aggiunto un'ulteriore dimensione allo studio delle emozioni, esaminando come il contesto culturale e sociale le modelli e influenzi. La sociologia delle emozioni, emersa negli anni '70, ha introdotto prospettive innovative sulla natura sociale delle emozioni, sfidando concezioni puramente individualistiche. Arlie Russell Hochschild, nel suo influente libro *The Managed Heart* (1983), ha esplorato i concetti di "regole del sentimento" e "lavoro emotivo", evidenziando come le emozioni siano profondamente radicate nelle strutture sociali e nelle interazioni quotidiane.

Tra le varie prospettive teoriche, l'interazionismo simbolico spicca per il suo ruolo anticipatore nello studio delle nuove forme di interazione sociale. L'interazionismo simbolico si focalizza sul modo in cui le persone creano significati attraverso le interazioni sociali. Questo approccio, sviluppato da studiosi come George Herbert Mead, Herbert Blumer ed Erving Goffman, sostiene che la realtà sociale è costruita attraverso il linguaggio e i simboli, e che gli individui interpretano e rinegoziano costantemente tali significati nelle loro relazioni quotidiane. La prospettiva drammaturgica di Goffman, in particolare, ha arricchito questa visione, considerando le emozioni come parte della *facciata* che gli individui presentano nelle loro interazioni quotidiane. Le emozioni, in questa visione, non sono semplicemente esperienze soggettive, ma vengono costruite e interpretate in base ai contesti sociali e culturali, influenzando e regolando i comportamenti degli individui. Anche le metodologie utilizzate hanno apportato elementi preziosi: tecniche come l'osservazione partecipante e le interviste in profondità hanno permesso di esplorare le modalità con cui la sfera emotiva viene negoziata nella quotidianità.

Negli ultimi decenni, le dinamiche analizzate attraverso la lente dell'interazionismo simbolico si sono ulteriormente trasformate con il diffondersi di strumenti simbolici come i social media e le piattaforme digitali, che hanno ampliato le modalità di costruzione e condivisione di significati emotivi collettivi in nuovi contesti di interazione sociale.

La panoramica fin qui accennata, seppur parziale e troppo sintetica per essere esaustiva, vuole offrire solo alcuni spunti sulle diverse prospettive nello studio delle emozioni umane, evidenziando come l'integrazione di approcci filosofici, psicologici, neuroscientifici e sociologici possa fornire una visione olistica delle emozioni, considerandole fenomeni che coinvolgono mente, corpo e società. Seguendo questa prospettiva interdisciplinare, la metodologia adottata si è basata su un'analisi approfondita della letteratura esistente, integrando ulteriori prospettive dall'informatica e dall'intelligenza artificiale. Si è deciso di condurre una revisione sistematica di pubblicazioni accademiche recenti, rapporti tecnici e casi di studio, con particolare attenzione agli sviluppi degli ultimi cinque anni nel campo dell'intelligenza artificiale generativa. L'analisi di report industriali, come l'Artificial Intelligence Index Report 2024, ha inoltre arricchito la comprensione dello stato dell'arte e degli investimenti nel settore, permettendo di tracciare un quadro multidimensionale dell'intersezione tra l'intelligenza artificiale e le emozioni umane.

2. Intelligenza artificiale ed emozioni

L'intelligenza artificiale, di seguito AI, è un ambito di ricerca transdisciplinare nato dal desiderio umano di creare macchine capaci di pensare, cercando di replicare artificialmente l'intelligenza che distingue l'essere umano dalle altre specie viventi.

L'inizio dello sviluppo tecnico dell'AI può essere ricondotto alla metà del XX secolo, con l'invenzione del neurone artificiale di McCulloch e Pitts (McCulloch, W. S., Pitts, W., 1943), uno dei primi modelli teorici alla base delle moderne reti neurali artificiali. Warren McCulloch, neurofisiologo, e Walter Pitts, logico matematico, collaborarono per creare un modello formale di neurone artificiale, ispirato al comportamento dei neuroni biologici. La loro sfida consisteva nel rappresentare matematicamente la logica del

funzionamento cerebrale, semplificando la complessità biologica e costruendo un sistema che simulasse l'attivazione o inibizione di un neurone.

L'approccio computazionale adottato da McCulloch e Pitts ha radici comuni con le teorie esplorate da Alan Turing nella sua teoria della computabilità del 1936, illustrata dal concetto di macchina di Turing, un dispositivo teorico capace di eseguire operazioni logiche sequenziali (Turing, A. M., 1936). Entrambi i modelli condividono l'idea che processi mentali come il calcolo e il ragionamento possano essere rappresentati matematicamente; tuttavia, riconoscono anche l'esistenza di limiti intrinseci nella computabilità, come evidenziato dal teorema di incompletezza di Gödel e dai risultati di Turing.

McCulloch e Pitts propongono inoltre una giustificazione psicologica per l'idea che i processi mentali possano essere rappresentati come operazioni logiche, suggerendo che la psicologia possa fungere da ponte tra la dimensione quantitativa delle operazioni logiche e la complessità qualitativa delle reti neurali biologiche. Pur non eliminando completamente la distanza tra macchina e umano, questa giustificazione psicologica permette di integrare il calcolo formale con la comprensione dei processi mentali complessi.

Nel 1950, Alan Turing contribuì significativamente al dibattito sulla relazione tra macchine e intelligenza umana, ponendo la domanda: le macchine possono pensare? (Turing, A. M., 1950). Per evitare ambiguità, Turing propose il "gioco dell'imitazione", oggi noto come "test di Turing", come criterio per valutare se una macchina possa imitare il comportamento umano al punto che un interrogatore non riesca a distinguerla da un essere umano. Questo approccio ha spostato l'attenzione dalla macchina all'utente umano, mettendo in primo piano la percezione e la soggettività interpretativa (Natale S., 2021). Tale cambio di paradigma, che pone l'umano al centro dell'ecosistema tecnologico, ha avuto successo in ambito pubblico e privato, ma risulta in contrasto con lo sviluppo tecnologico che rimane fortemente tecno-centrico.

L'anno di nascita attribuito all'AI è il 1956, quando, presso il Dartmouth College, si svolse il primo progetto congiunto tra università e aziende di ricerca sull'intelligenza artificiale. Il termine "Intelligenza Artificiale" venne utilizzato per la prima volta nella proposta del progetto di ricerca estivo del 1956, nota come "Proposta di Dartmouth" (McCarthy, J. et al., 1955). Sin dalla sua nascita, l'AI è stata un terreno di ricerca e sviluppo congiunto tra università e aziende, come dimostrano i firmatari della proposta – J. McCarthy, M. L. Minsky, N. Rochester e C.E. Shannon – che sono rispettivamente affiliati ai seguenti enti: Dartmouth College, Università di Harvard, I.B.M. Corporation e Bell Telephone Laboratories. Tra i sette aspetti critici su cui focalizzare il progetto di ricerca, si legge la volontà di approfondire il lavoro di modellizzazione di McCulloch e Pitts sulle reti neurali, e affrontare la complessità del pensiero umano, con particolare attenzione all'astrazione e alla casualità nella creatività. Su quest'ultimo problema si cita la traduzione del testo originale perché mette in luce la difficoltà di categorizzare il pensiero e ridurlo a un modello matematico: «Un'ipotesi piuttosto interessante, ma chiaramente incompleta, è che la differenza tra il pensiero creativo e il pensiero competente, ma non immaginativo, risieda nell'introduzione di una certa dose di casualità. La casualità deve essere guidata dall'intuizione per essere efficiente. In altre parole, l'ipotesi plausibile o l'intuizione includono una casualità controllata all'interno di un pensiero altrimenti ordinato».

Il risultato concreto della conferenza di Dartmouth fu un programma chiamato Logic Theorist (Newell, A., Simon, H. A., 1956), in grado di costruire dimostrazioni matematiche da zero e considerato il primo programma di AI. Tuttavia, utilizzava metodi di "forza bruta" per risolvere problemi, rimanendo distante dal sogno di una macchina pensante. Negli anni successivi, lo sviluppo dell'AI ha vissuto alti e bassi sia in termini di risultati scientifici che di finanziamenti, dipendenti da promesse non mantenute e dall'evoluzione geopolitica mondiale; mentre, a partire dagli anni '90, grazie al miglioramento delle performance di calcolo, allo sviluppo di nuovi algoritmi e alla disponibilità di dati è stato possibile iniziare sperimentazioni concrete sulle teorie dell'apprendimento automatico e delle reti neurali.

La possibilità di apprendere autonomamente da dati d'esempio ha accelerato lo sviluppo dell'AI, semplificando il processo umano di interpretazione e modellizzazione della realtà. Entrando più nello

specifico, il machine learning impara analizzando ed elaborando un insieme di dati, chiamato in gergo dataset, con algoritmi matematici che ne estraggono relazioni e pattern. Il processo di elaborazione è statistico-probabilistico e segue la logica induttiva: il sistema elabora una regola generale a partire dall'analisi degli esempi analizzati. L'intervento umano è fondamentale nella selezione degli esempi, nella scelta dell'algoritmo di apprendimento più adatto e nella valutazione critica dell'apprendimento del sistema, passo necessario per ottenere un modello che replichi più fedelmente possibile la realtà indicata dai dati. Tuttavia, la complessità dei sistemi attuali rende spesso questi modelli delle *scatole nere* (black box), motivo per cui si eviteranno ulteriori approfondimenti tecnici per concentrare l'attenzione sulla connessione tra la dimensione umana e quella delle macchine.

2.1 La modellizzazione delle emozioni

Nel 1978, Paul Ekman e Wallace V. Friesen svilupparono il Facial Action Coding System (FACS), un metodo per valutare e misurare l'intensità delle espressioni facciali e, quindi, delle emozioni. Questo *atlante* del volto umano scompone le espressioni in singole unità di movimento, note come *Action Units* (AU), classificandole con un'intensità variabile secondo la seguente scala crescente:

- A: tracce deboli dell'azione (Trace)
- B: evidenza leggera dell'azione (Slight)
- C: segni marcati o pronunciati del movimento (Marked o Pronounced)
- D: segni intensi o estremi dell'azione (Severe o Extreme)
- E: intensità massima del movimento.

L'evoluzione tecnologica dell'AI ha intensificato il dibattito sulla natura delle emozioni nelle macchine e sul loro impatto sulla sfera emotiva umana. Questo dibattito ha messo in luce problematiche fondamentali nella teoria dell'espressione delle emozioni di Ekman, che ha dominato la psicologia delle emozioni per decenni ed è stata integrata nei sistemi di AI per il riconoscimento facciale. Kate Crawford, nel suo libro *Né intelligente né artificiale* (Crawford, K., 2021), critica duramente questa teoria, evidenziandone i limiti scientifici e gli effetti discriminatori. Secondo Crawford, l'idea che sia possibile riconoscere con certezza le emozioni dall'espressione facciale e discernere il vero dal simulato è una pretesa più *poliziesca* che scientifica. Questo approccio trascura la complessità del rapporto tra emozione interna ed espressione esterna, ignorando sfumature culturali, contestuali e individuali (Barrett, L. F. et al., 2019).

Ekman ha basato i suoi studi sulla fisiologia muscolare del volto, proseguendo il lavoro di Duchenne de Boulogne e spinto dal desiderio di confutare le considerazioni di Charles Darwin, ritenute più aneddotiche che scientifiche. La metodologia di Duchenne, che induceva espressioni facciali attraverso stimoli elettrici per fotografarle, è stata criticata per la mancanza di spontaneità e veridicità e Ekman, dichiaratosi seguace di Duchenne, ha accettato le sue tesi senza sufficienti dubbi metodologici. Critici come Margaret Mead e Gregory Bateson hanno evidenziato come questi studi non abbiano considerato adeguatamente la complessa relazione tra il mondo interiore delle emozioni e la loro manifestazione esterna.

Il ruolo della fotografia nella teoria di Ekman è un altro aspetto problematico. Il nesso tra fisiognomica e fotografia, stabilito da Duchenne e ripreso da Ekman, ha creato un'illusione di oggettività scientifica, ignorando che la fotografia può essere un mezzo di interpretazione soggettiva piuttosto che di pura documentazione.

La diffusione del sistema di Ekman in ambiti delicati, come il riconoscimento di potenziali terroristi negli aeroporti o l'utilizzo nella terapia per bambini con diagnosi di spettro autistico, solleva serie questioni etiche. La mancanza di una critica approfondita alla povertà epistemologica di queste teorie ha contribuito alla loro accettazione acritica. È necessario un approccio più ampio e critico allo studio delle emozioni,

evitando di sezionare un aspetto così pervasivo della psiche umana. La riflessione sull'AI offre un'opportunità per ripensare radicalmente ciò che è umano, aprendo la porta a sperimentazioni più audaci e creative nel campo delle emozioni e della loro espressione.

Mentre l'AI evolve nel tentativo di comprendere e replicare le emozioni umane, il dibattito sulle emozioni nelle macchine rimane più rilevante che mai. Il rischio è di sbilanciarsi eccessivamente verso la computazione di un aspetto intimo e mutevole dell'essere umano, tentando di estrarre valore da ogni sfaccettatura dell'esperienza umana, con effetti concreti sull'esistenza di ciascuno. La sfida per il futuro sarà sviluppare approcci che riconoscano la complessità e la ricchezza delle emozioni umane, evitando semplificazioni riduttive e rispettando l'unicità e la soggettività dell'esperienza emotiva.

3. Intelligenza artificiale generativa e la nuova relazione

L'intelligenza artificiale generativa (GenAI) è una branca dell'AI caratterizzata dalla capacità di creare nuovi contenuti, come testo, immagini, suoni, video e codici di programmazione. Contrariamente all'AI tradizionale, focalizzata sull'analisi e classificazione di dati preesistenti, la GenAI utilizza modelli statistici avanzati e reti neurali per apprendere dai dati e generare nuovi contenuti apparentemente unici. Il tema dell'unicità percepita del risultato generato è molto dibattuto. Dal punto di vista tecnico, l'AI generativa viene allenata con grandi quantità di dati, dai quali impara estraendo pattern e relazioni utilizzate per generare gli output.

Una metafora ornitologica che ha avuto successo mediatico è quella del "pappagallo stocastico" (Bender, E. M. et al., 2021), riferita ai sistemi generativi di linguaggio naturale, perché evidenzia come la generazione delle risposte sia una concatenazione di parole imparate per imitazione, seguendo logiche probabilistiche senza comprendere la semantica umana. Il riconoscimento dell'unicità del risultato generato da una macchina amplia la complessità della questione tecnica e legale e vari studi hanno mostrato la necessità di ridefinire le regole del copyright in questa nuova era di collaborazione tra umani e macchine (Watiktinnakorn, C. et al. 2023; Fenwick, M., Jurcys, P., 2023; Napoli, G., Naccarato, M., 2023; Catarinicchia S., Scalzini A., 2023; Mazzi, F., 2024).

Secondo l'*Artificial Intelligence Index Report 2024* (Maslej, N. et al., 2024), gli investimenti in AI generativa sono saliti alle stelle. Nonostante un calo degli investimenti privati complessivi in AI lo scorso anno, i finanziamenti per l'AI generativa sono aumentati, quasi ottuplicandosi dal 2022 fino a raggiungere i 25,2 miliardi di dollari. I principali attori nel settore, tra cui OpenAI, Anthropic, Hugging Face e Inflection, hanno segnalato sostanziali round di raccolta fondi. Attualmente, il settore tecnologico dell'AI è fortemente focalizzato sullo sviluppo di modelli di base, foundation models, necessari per realizzare applicazioni di GenAI; infatti, nel 2023 sono stati rilasciati 149 modelli di base, che rappresentano più del doppio dei rilasci dell'anno precedente.

Il punto di partenza della rivoluzione generativa dell'AI, a livello popolare e globale, è il 30 novembre 2022, giorno in cui viene rilasciato online ChatGPT (OpenAI, 2022). Il chatbot sviluppato da OpenAI ha ottenuto un record di diffusione senza precedenti: in soli cinque giorni dal suo rilascio ha raccolto un milione di utenti (Forbes Italia, 2022), superando i social network più popolari in termini di adozione e velocità. La gratuità, la semplicità d'uso, l'aumento delle capacità di processamento del testo e la possibilità di conversare in varie lingue sono elementi chiave per giustificare l'ampia adozione e l'impatto mediatico.

A quasi due anni dal rilascio, il termine ChatGPT viene spesso usato impropriamente al posto di GenAI; infatti, l'AI generativa è un ambito che include applicazioni specifiche come ChatGPT, ma non solo. ChatGPT, Gemini sviluppato da Google, Claude sviluppato da Anthropic, Pi sviluppato da Inflection AI e altri sono chatbot basati su modelli di linguaggio di grandi dimensioni (LLM), che generano la migliore risposta alla richiesta (prompt) dell'utente. Altre applicazioni di IA generativa includono sistemi T.T.I. (Text To Image) che

generano immagini da descrizioni testuali, o generatori di video e audio a partire da specifiche caratteristiche o descrizioni.

Per approfondire la relazione umano-macchina in questa nuova fase di sviluppo dell'AI, è importante evitare l'errore di ridurre il campo di ricerca dell'AI generativa al solo strumento ChatGPT o simili. L'AI generativa odierna si sviluppa a partire dal machine learning, frutto di un intenso lavoro nel campo del deep learning. Ricordando che il machine learning è un sottoinsieme dell'AI che utilizza algoritmi in grado di apprendere autonomamente da dati (dataset), senza necessità di programmazione esplicita, è possibile definire il deep learning come quel ramo del machine learning che apprende tramite reti neurali artificiali, organizzate in architetture complesse. È importante notare che le reti neurali artificiali sono ispirate al funzionamento del cervello umano, ma non ne replicano la complessità e la plasticità (Dehaene, S., 2019). Nonostante le somiglianze concettuali, i principi operativi fondamentali sono profondamente differenti.

Essere consapevoli che il pensiero delle macchine segue un processo diverso rispetto a quello umano è fondamentale quando si considera la dimensione emotiva della relazione. Attualmente, gruppi di ricerca neuroscientifica e ingegneristica stanno realizzando reti neurali artificiali sempre più simili a quelle biologiche, chiamate reti neuromorfiche, che rappresentano una frontiera promettente dell'AI (Ielmini, D., 2018; Davies, M., et al., 2021; Intel, 2024)

3.1 Da ELIZA a ChatGPT

Il termine "chatbot" è particolarmente rilevante nell'approfondimento della relazione tra umani e macchine. Un chatbot è un'applicazione di AI generativa specializzata nella gestione del linguaggio naturale e nel sostenere conversazioni scritte. Il termine deriva dalla fusione di *chat*, chiacchierare, e *bot*, abbreviazione di *robot*. Interessante notare che *robot* deriva dal ceco *robota*, nel senso di lavoro servile, utilizzato per la prima volta nel 1920 dallo scrittore Karel Čapek per denominare gli automi nel suo dramma fantascientifico *R.U.R.* (Treccani).

Il primo esempio di chatbot nella storia della GenAI moderna è ELIZA, realizzato negli anni '60 al Massachusetts Institute of Technology (MIT) da Joseph Weizenbaum. ELIZA utilizzava tecniche primordiali di processamento del linguaggio naturale per elaborare i messaggi dell'utente e generare risposte; il sistema non faceva inferenze da una base di conoscenza predefinita e non aveva una reale comprensione dei messaggi scambiati (Weizenbaum, J. 1966). Weizenbaum rimase colpito da come ELIZA venisse antropomorfizzata, creando un collegamento empatico tra l'utente e la macchina (Weizenbaum, J. 1976), basato sull'illusione percettiva dell'utente umano (Natale, S., 2021).

Dagli anni '60 ad oggi, l'approccio alla gestione del linguaggio naturale è evoluto notevolmente. L'ultima frontiera è rappresentata dai modelli di rete neurale denominati Transformer, introdotti nel 2017 con la pubblicazione dell'articolo "Attention is All You Need" da parte di ricercatori di Google (Vaswani, A., et al., 2017). Questo approccio ha rivoluzionato il deep learning grazie alla capacità di processare sequenze di parole in modo più efficiente e preciso. Il modello Transformer, per le sue caratteristiche innovative e versatilità, è considerato un modello di base (foundation model), in grado di svolgere compiti generici e adattarsi a differenti applicazioni e, l'ultima frontiera di sviluppo raggiunta, permette al modello di processare input differenti tra loro, come testo, immagini, suoni, video. L'ampliamento delle modalità di input del modello si indica con il termine *multimodale* e si può considerare come un passo molto importante nell'evoluzione di una maggiore capacità sensoriale e percettiva della macchina (MIT Technology Review, 2024).

3.2 La nuova relazione

Prima dell'avvento di ChatGPT, sono stati realizzati molti chatbot che hanno suscitato scalpore mediatico. Il 23 marzo 2016, Microsoft rilasciò Tay, un chatbot basato su AI per testare le sue capacità di interazione sui

social media. Tay imitava una ragazza di 19 anni su Twitter, rispondendo ai messaggi degli utenti; tuttavia, l'esperimento si concluse dopo sole sedici ore perché il chatbot rispose con commenti razzisti e politicamente scorretti (Bennato, D., 2018). Questo evento sollevò un dibattito sull'influenza dei dati di addestramento e delle interazioni degli utenti nel comportamento del modello.

Un altro esempio rilevante è Replika, un'app chatbot rilasciata nel 2017 che realizza la profezia raccontata nella puntata "Be Right Back" della serie *Black Mirror* (2013). Come raccontato dalla fondatrice del progetto, il chatbot nasce dall'ispirazione di ricreare digitalmente la personalità del migliore amico perduto in un incidente stradale, utilizzando i dati delle conversazioni sui social (Replika). Le prime sperimentazioni mostrarono la facilità con cui gli utenti si aprivano emotivamente al chatbot, ispirando gli sviluppatori a collegare lo sviluppo della personalità del chatbot all'interazione con l'utente.

La nuova relazione con le macchine è integrata e multimodale: attraverso dispositivi mobili e indossabili, vengono forniti ai sistemi di AI parametri biometrici e ambientali, utilizzati per prevedere il nostro stato di salute e le nostre necessità. L'interazione avviene attraverso più canali sensoriali, abbassando le difficoltà comunicative legate al linguaggio scritto. I sistemi di AI generativa hanno quindi più dati relativi all'utente e al contesto, e possono essere plasmati per favorire la comunicazione, in un gioco di illusione empatica e dimensione non giudicante.

La nuova relazione diventa così un delicato equilibrio tra la dimensione reale/analogica e quella artificiale/digitale, dove i linguaggi umani vengono elaborati da sistemi tecnologici che non sono neutri (Floridi, L. 2023), non vivono una dimensione emotiva e non hanno una comprensione semantica del contesto (Baraldi, A., et al., 2023), ma interpretano la realtà attraverso sequenze numeriche elaborate seguendo regole matematiche (Floyd, C., 2023).

4. Conclusione

L'intersezione tra emozioni umane e intelligenza artificiale rappresenta un campo di ricerca in rapida evoluzione, sollevando questioni profonde sulla natura dell'esperienza emotiva e sul futuro delle interazioni uomo-macchina. Lo sviluppo di reti neurali neuromorfiche, che mirano a replicare il funzionamento cerebrale nell'elaborazione delle emozioni, apre nuove frontiere nella simulazione dei processi emotivi. Tuttavia, è cruciale interrogarsi sulle finalità e sulle implicazioni etiche di questa evoluzione tecnologica, poiché l'interazione e integrazione crescente con sistemi di AI pone profondi interrogativi sull'impatto psicologico a medio e lungo termine.

Il fenomeno dell'antropomorfizzazione, già osservato da Weizenbaum, assume nuove sfumature con chatbot e assistenti virtuali sempre più sofisticati. Il paradigma tradizionale, in cui l'uomo si relazionava alla macchina come a un'entità subordinata, sta cambiando: i sistemi di AI non solo eccellono in compiti ripetitivi, ma mostrano capacità di ragionamento e creatività che, in certi casi, superano quelle umane. Questa evoluzione tecnologica comporta rischi maggiori rispetto alle rivoluzioni del passato e va approfondita nelle varie sfumature di percezione e interpretazione del ruolo che si è soliti attribuire alle macchine e alle scelte fatte dai sistemi di intelligenza artificiale (Hidalgo, C. A., 2021).

Come evidenziato dallo psicoterapeuta Miguel Benasayag (Benasayag, M., 2016), è fondamentale preservare la singolarità del vivente di fronte all'avanzata dell'AI, in quanto la tecnologia potrebbe semplificare eccessivamente i processi cognitivi e neurali, atrofizzando varie abilità umane. Questo processo è stato osservato neuroscientificamente, ad esempio, nei tassisti che utilizzano sistemi di navigazione GPS rispetto agli stessi che si orientano senza aiuti digitali, o confrontando l'attivazione neuronale semplificata di chi legge testi su uno schermo rispetto a chi legge su carta.

Considerando che l'ibridazione digitale del vivente è già in atto e sta modificando i processi di pensiero, l'emotività rimane un ambito prettamente umano e l'educazione emotiva assume un ruolo ancora più

importante nella protezione della singolarità del vivente. La continua esposizione a contenuti virali su piattaforme social come TikTok, in cui l'utente può essere considerato un "otre vuoto" da riempire e iperstimolare, incide sulla sfera emotiva e attentiva, senza dare la possibilità di conoscere la razionalità degli algoritmi che scelgono i contenuti da profilare attraverso le nostre interazioni.

È essenziale promuovere strategie educative che sviluppino una consapevolezza critica delle interazioni dell'uomo con la tecnologia, contrastando la tendenza verso un pensiero automatizzato e preservando la complessità dell'esperienza emotiva umana. Ogni elemento che entra in contatto con i sensi ha il potenziale di evocare emozioni, influenzando la percezione e coscienza. L'interfaccia uomo-macchina e l'esperienza utente non dovrebbero solo ottimizzare la performance, ma considerare l'intero spettro dell'esperienza umana, inclusa la dimensione emotiva.

Ritornando ai chatbot e alla sfida contemporanea, la dimensione di manipolazione dell'utente da parte del sistema di AI risulta molto più sottile e subdola: l'illusione di poter plasmare la personalità del chatbot, l'apparente neutralità e astensione dal giudizio, la semplicità dell'interazione che passa attraverso la semplificazione della comunicazione portano a sottostimare l'effetto e l'impatto sulle abilità e competenze umane. Le rigidità nell'esecuzione di alcuni compiti portano l'utente a modificare il proprio modo di formulare il problema, semplificando le richieste, applicando strategie di scomposizione del problema in piccoli passi da ricombinare, abituandosi a comunicare in modo semplificato e disimparando a gestire la complessità. *Pensieri robotizzati* è la metafora che più si avvicina a questo processo in cui la macchina influenza sempre di più l'umano fino quasi a scambiare i ruoli e rapporti di intelligenza e controllo (Dominici, P., 2022).

La nuova relazione tra l'umano e la macchina è un futuro già presente, in cui l'intelligenza artificiale ha un ruolo centrale come mediatore dell'esperienza umana della realtà. È essenziale mantenere un approccio critico e riflessivo, riconoscendo quella singolarità del vivente che lo porta all'esistenza e non a un mero funzionamento (Benasayag, M., 2021). Non si può escludere che in futuro sarà possibile ricondurre il funzionamento della mente umana, emozioni incluse, a un processo descrivibile matematicamente e replicabile seguendo le logiche binarie o quantistiche dell'elettronica sviluppate fino ad oggi. Tuttavia, non si nega che l'utilizzo dell'intelligenza artificiale in un'ottica abilitante e inclusiva per il pensiero umano (Monari P., Baraldi L., 2022) sia fondamentale per una crescita globale e umanocentrica, in contrasto con l'avanzare indifferente della tecnica.

Aumentando la consapevolezza tecnologica, imparando a usare la tecnologia senza essere usati dalla tecnologia e riconoscendo che le emozioni appartengono a una complessità non ancora riconducibile a dati o equazioni, è possibile sottolineare che i *pensieri robotizzati* rimarranno solo un ricordo distopico presente nei racconti di fantascienza.

Riferimenti bibliografici

- Bennato D. (2018). Intelligenze artificiali razziste e psicopatiche: studiarle per non esserne vittima. *Agenda Digitale* <https://www.agendadigitale.eu/cultura-digitale/intelligenze-artificiali-razziste-e-psicopatiche-studiarle-per-non-esserne-vittima/> (ultimo accesso settembre 2024)
- Baraldi, A., Sapia, L. D., Tiede, D., Sudmanns, M., Augustin, H. L., & Lang, S. (2023). Innovative analysis ready data (ARD) product and process requirements, software system design, algorithms and implementation at the midstream as necessary-but-not-sufficient precondition of the downstream in a new notion of Space Economy 4.0 – Part 1: Problem background in Artificial General Intelligence (AGI). *Big Earth Data*, 7(3), 455-693. <https://doi.org/10.1080/20964471.2021.2017549>

- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68.
- Benasayag, M. (2016). *Il cervello aumentato, l'uomo diminuito*. Erickson.
- Benasayag, M. (2021). *Funzionare o esistere?*. Vita e Pensiero.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? . Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623.
- Catarinicchia, S., & Scalzini, A. (2023). Intelligenza Artificiale e Diritto d'Autore: Scenario Normativo in Italia. *Mondaq*. <https://www.mondaq.com/italy/copyright/1491912/intelligenza-artificiale-e-diritto-dautore-scenario-normativo-in-italia> (ultimo accesso settembre 2024)
- Crawford, K. (2021). *Né intelligente né artificiale. Il lato oscuro dell'IA*. Il Mulino.
- Davies, M., Srinivasa, N., Lin, T. H., China, G., Cao, Y., Joshi, P., Tse, J., Chang, L., Narayanan, P., & Tuan, T. (2021). Advancing neuromorphic computing with Loihi: A survey of results and outlook. *Proceedings of the IEEE*, 109(5), 911-934. <https://doi.org/10.1109/JPROC.2021.3067593>
- Dehaene, S. (2019). *Imparare. Il talento del cervello, la sfida delle macchine* (L. Ducceschi, Trad.). Raffaello Cortina Editore.
- Dominici, P. (2022). The Digital Mockingbird: Anthropological Transformation and the “New” Nature. *World Futures*, 78(6), 343–371. <https://doi.org/10.1080/02604027.2022.2028539>
- Fenwick, M., & Jurcys, P. (2023), Originality and the Future of Copyright in an Age of Generative AI. Available at SSRN: <https://ssrn.com/abstract=4354449>
- Floridi, L. (2023). On good and evil, the mistaken idea that technology is ever neutral, and the importance of the double-charge thesis. *Philosophy & Technology*, September 2023. Available at SSRN: <https://ssrn.com/abstract=4551487> or <http://dx.doi.org/10.2139/ssrn.4551487>
- Floyd, C. (2023). From Joseph Weizenbaum to ChatGPT: Critical Encounters with Dazzling AI Technology. *Weizenbaum Journal of the Digital Society*, 3(3). <https://doi.org/10.34669/WI.WJDS/3.3.3>
- Forbes Italia. (2022). ChatGPT raggiunge un milione di utenti in cinque giorni. <https://www.forbes.com/advisor/it/business/trend-ai-statistiche/> (ultimo accesso settembre 2024)
- Hidalgo, C. A., Orghian, D., Albo Canals, J., de Almeida, F., & Martin, N. (2021). *How humans judge machines*. MIT Press.
- Ielmini, D. (2018). Intelligenza artificiale: l'approccio neuromorfico. *Mondo Digitale*, 79, 1–8. https://mondodigitale.aicanet.net/2018-6/articoli/01_MD79_ielmini_Intelligenza_artificiale.pdf
- Intel (2024) Intel Builds World's Largest Neuromorphic System to Enable More Sustainable AI <https://www.intel.com/content/www/us/en/newsroom/news/intel-builds-worlds-largest-neuromorphic-system.html> (ultimo accesso settembre 2024)
- Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). The AI Index 2024 annual report. *AI Index Steering Committee*, Institute for Human-Centered AI, Stanford University.

- Mazzi, F. (2024). Authorship in artificial intelligence-generated works: Exploring originality in text prompts and artificial intelligence outputs through philosophical foundations of copyright and collage protection. *The Journal of World Intellectual Property*, 1–18. <https://doi.org/10.1111/jwip.12310>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf> (ultimo accesso settembre 2024)
- McCulloch, W. S., & Pitts, W. H. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4), 115–133. doi:10.1007/BF02478259
- MIT Technology Review (2024) Multimodal: AI's new frontier <https://www.technologyreview.com/2024/05/08/1092009/multimodal-ais-new-frontier/> (ultimo accesso settembre 2024)
- Monari, P., & Baraldi, L. (2022). AI: A super-tool for fostering human thought. *Atti del Convegno Nazionale DIDAMATICA 2022* (pp. 437-444)
- Napoli, G., & Naccarato, M. (2023). Copyright protection for AI-generated works: Exploring originality and ownership in a digital landscape. *Asian Journal of International Law*, 12(1), 1–30.
- Natale, S. (2021). *Macchine ingannevoli*. Giulio Einaudi Editore.
- Newell, A., & Simon, H. A. (1956). The logic theory machine: A complex information processing system. *IRE Transactions on Information Theory*, 2(3), 61–79.
- OpenAI. (2022). ChatGPT. <https://openai.com/index/chatgpt/> (ultimo accesso settembre 2024)
- Replika. Our Story. <https://replika.com/about/story> (ultimo accesso settembre 2024)
- Shen, Y., Jin, J., Zhang, B., & Xu, Q. (2022). How Does ChatGPT Understand and Generate Language? *International Journal of Computational Linguistics*, 13(2), 123–135.
- Stanford University. (2024). Artificial Intelligence Index Report 2024. <https://aiindex.stanford.edu/report/> (ultimo accesso settembre 2024)
- Treccani. Robot. <https://www.treccani.it/vocabolario/robot/> (ultimo accesso settembre 2024)
- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2), 230–265.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://academic.oup.com/mind/article/LIX/236/433/986238>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000–6010. NIPS.
- Watiktinnakorn, C., Seesai, J. & Kerdvibulvech, C. Blurring the lines: how AI is redefining artistic ownership and copyright. *Discov Artif Intell* 3, 37 (2023). <https://doi.org/10.1007/s44163-023-00088-y>
- Weizenbaum, J. (1966). ELIZA – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.

- Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgment to Calculation*. W. H. Freeman.

[PIETRO MONARI](#)

Pietro Monari – Project Manager Education nel team Human Sciences di Ammagamma Part of Accenture Fisico e musicista con un'esperienza pluriennale come insegnante di Matematica e Fisica alle Scuole Superiori, negli ultimi anni ha condotto varie sperimentazioni in ambito di educazione multidisciplinare STEAM. È ideatore e promotore di progetti quali "Educare a Pensare: umanità dell'IA" e LUCY, prima scuola sperimentale di IA in Italia realizzata in collaborazione con l'I.C. 3 Mattarella di Modena. In Ammagamma è project manager per il coordinamento, lo sviluppo e l'implementazione delle attività educative e direttore del centro di ricerca "The Energy of Data". (p.monari@ammagamma.com).

[LUCA BARRETTA](#)

Luca Barretta – Project Manager nel team Human Sciences di Ammagamma Part of Accenture Sociologo esperto di culture digitali e della comunicazione, con una forte passione per le tecnologie moderne e i paradigmi di interazione uomo-macchina. Ha maturato una solida esperienza nel terziario avanzato e in contesti multinazionali in ambito software e servizi professionali, dove si è focalizzato sulle dinamiche relazionali nel settore delle vendite e di consulenza ai clienti. Attualmente è Project Manager nel team Human Sciences di Ammagamma, dove integra le sue competenze sociologiche con le tecnologie emergenti.
(l.barretta@ammagamma.com)