

Interazione e relazione utente-chatbot. Dove inizia l'esperienza umana e quando la finzione?

di Giorgia Altobelli*

L'articolo esplora l'uso dell'IA generativa e dell'*Affective Computing* nella rivoluzione educativa, evidenziando l'impatto su apprendimento, accessibilità e personalizzazione. Analizza l'interazione uomo-macchina, con focus su manipolazione emotiva e Effetto Eliza, che porta gli utenti ad attribuire qualità umane alle macchine. In un contesto tecnologico in rapida evoluzione, si sottolinea l'importanza delle competenze socio-emotive e di un approccio etico nella progettazione di esperienze utente sicure, efficaci e responsabili.

Parole chiave: intelligenza artificiale generativa; affective computing; ricerca educativa; effetto Eliza; competenze socio-emotive; tecnologia responsabile.

User-chatbot interaction and relationship. Where does the human experience begin and when does fiction start?

The article explores the use of generative AI and *Affective Computing* in the educational revolution, highlighting their impact on learning, accessibility, and personalization. It analyzes human-machine interaction, focusing on emotional manipulation and the Eliza Effect, which leads users to attribute human qualities to machines. In a rapidly evolving technological context, it emphasizes the importance of socio-emotional skills and an ethical approach in designing safe, effective, and responsible user experiences.

Keywords: generative artificial intelligence; affective computing; educational research; Eliza effect; socio-emotional skills; responsible technology.

1. Le macchine possono amare?

La riflessione di Alan Turing (1950) sulla capacità delle macchine di pensare ha stabilito i fondamenti teorici dell'intelligenza artificiale moderna. La questione rimane aperta e richiede una riflessione approfondita da parte della comunità scientifica e accademica sull'affidabilità dei dispositivi tecnologici personali, che occupano un ruolo sempre più centrale nella nostra vita quotidiana e affettiva. Il “Test di Turing” suggerisce che la domanda

DOI: 10.5281/zenodo.17523984

* Università degli Studi di Macerata. altobelligiorgia@gmail.com.

Giorgia Altobelli

principale non è se le macchine siano capaci di pensiero autonomo, ma piuttosto se gli esseri umani siano disposti ad attribuire valore di pensiero alle loro azioni (Turing, 1937). Nel 1956, al Dartmouth College nel New Hampshire, si tenne un workshop di due mesi che propose una nuova prospettiva, come evidenziato da McCarthy *et al.*:

Lo studio si baserà sulla congettura che ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza possa, in linea di principio, essere descritto con tale precisione da consentire a una macchina di simularlo. Si cercherà di capire come far sì che le macchine usino il linguaggio, formino astrazioni e concetti, risolvano tipi di problemi oggi riservati agli esseri umani e si migliorino (1955: 2).

L'intelligenza, definita da McCarthy, non consiste solo nell'imitare l'intelligenza umana, ma nel dimostrare l'intelligenza come una capacità computazionale in grado di raggiungere risultati nella realtà. All'epoca, era impensabile prevedere che l'avvento del Web, l'utilizzo dei big data e dell'intelligenza artificiale avrebbero portato allo sviluppo di algoritmi di apprendimento avanzati, capaci di emulare la creatività umana e generare contenuti tramite machine learning, dando vita alla cosiddetta GenAI¹. L'intelligenza artificiale generativa² a differenza dalle altre forme di IA riesce a produrre risposte sempre più efficienti e in forme destrutturate. Gli assistenti vocali incarnano questa evoluzione, poiché le macchine non solo eseguono comandi, ma anticipano esigenze e forniscono risposte sempre più personalizzate (Cohen *et al.*, 2004). Rosalind Picard (1977), nel suo libro '*Affective Computing*', evidenziò la necessità per le macchine di riconoscere e comprendere le emozioni umane, superando i limiti di un'interazione rigida e frustrante, dando vita all'affective computing. Termine traducibile come "calcolo affettivo", permette alle macchine di simulare emozioni umane e riprodurre interazioni sempre più simili a quelle dell'essere umano. Ciò include risposte empatiche e comprensione, riproducendo comportamenti emotivi complessi. Come sottolinea Borgna (1999), ogni manifestazione umana, anche quelle più discrete o silenziose, veicola forme di comunicazione e interazione. Brazelton e Greenspan (2001) affermano che la gamma emozionale umana è condizionata dalle esperienze pregresse; pertanto,

¹ Generative Artificial Intelligence.

² Il primo programma di Generative AI, ChatGPT, è stato lanciato sul mercato nel novembre 2022 dall'impresa statunitense OpenAI. Dopo solo una settimana, la piattaforma registrava più di un milione di utenti al giorno. ChatGPT, il primo tool di AI Generativa, sfruttando gli algoritmi di intelligenza artificiale di apprendimento automatico, è in grado di svolgere molissime funzioni.

risulta difficile esperire emozioni mai conosciute in precedenza. Analogamente, la capacità di instaurare relazioni profonde e durature dipende dall'aver vissuto esperienze significative di attaccamento e intimità nel corso della propria vita. La rielaborazione, la discussione e lo scambio rappresentano elementi fondamentali per l'integrazione dei saperi in contesti di apprendimento virtuale e mediato, consentendo una più efficace acquisizione di conoscenze (Lavanga, Mancaniello, 2022). Tuttavia, l'evoluzione verso un'interazione più fluida e naturale solleva anche questioni sulla dipendenza tecnologica e sul potenziale rischio di delega delle capacità cognitive umane a sistemi automatizzati (Floridi, 2022). Il dialogo multidimensionale tra il sé, l'altro e l'ambiente ha giocato un ruolo chiave nell'evoluzione cognitiva e sociale umana. Attraverso l'analisi di processi di domanda e risposta, è possibile comprendere come tale capacità abbia promosso la differenziazione dell'umano dal mondo animale e favorito lo sviluppo di competenze cognitive e sociali avanzate.

2. Quale relazione empatica con la GenAI?

Nel mondo educativo l'intelligenza artificiale sta apportando rapidi cambiamenti, rivoluzionando il mondo dell'educazione in quella che viene conosciuta come la “quarta rivoluzione educativa”³, definendo nuovi ambienti educativi accessibili e inclusivi, migliorando l'esperienza e la qualità dei percorsi di apprendimento, rendendoli personalizzati e modellati in relazione a specifici bisogni. Don Milani (1990) sottolineava quanto fosse necessaria, in un rapporto educativo, la relazione empatica d'amore tra educatore ed educando. Da un punto di vista didattico, diventa fondamentale cercare un equilibrio tra la dimensione spontanea e la necessità di educare, tendendo verso un percorso educativo fatto di benessere e felicità (Locatelli *et al.*, 2013). Brazelton e Greenspan (2001) sostengono che le emozioni sono gli artefici, le guide o gli organizzatori interni delle nostre menti. Infatti, la relazione è al centro dell'educazione, tanto quanto lo sono le emozioni, che guidano il *training* formativo. La dinamica relazionale e l'esperienza emotiva sono elementi cruciali nel contesto educativo e appare sempre più chiaro come i dispositivi tecnologici raccolgano sistematicamente informazioni personali, spesso senza piena consapevolezza degli utenti, consentendo una conoscenza dettagliata degli stessi che può superare quella dei familiari più prossimi (Bucchi, 2020). La ricerca educativa è chiamata ad intervenire sulla sfida

³ Coniata nel 2018 da Seldon e Abidoye.

Giorgia Altobelli

alla trasformazione dell'esperienza digitale umanizzante, in cui sono messe in discussione: la dimensione relazionale, lo scambio tra pari, l'ascolto attivo, il riconoscimento reciproco, la cura e la negoziazione del conflitto. Mariani (2006) sottolinea che l'azione educativa efficace richiede specifiche condizioni, quali cura, fiducia e amorevolezza, in quanto l'engagement e il riconoscimento sono elementi determinanti per il successo dell'azione formativa. Tali qualità sono intrinsecamente umane e si sviluppano all'interno di relazioni reciproche che si consolidano nel tempo. L'apertura alla realtà e in particolare all'alterità rappresenta il fondamento essenziale dell'attività relazionale, che a sua volta è sostenuta e alimentata dalla libertà. Le ricerche nel campo delle neuroscienze confermano l'intuizione di Vygotskij (1987) sull'influenza significativa delle esperienze sullo sviluppo cerebrale e l'importanza dell'interazione tra individuo e ambiente. Dunque, la comunità educante non ha solo il compito di promuovere un ambiente positivo, basato sulla condivisione di valori e sull'utilizzo di linguaggi emotivi ed affettivi, ma anche di favorire relazioni significative e diffondere il senso di appartenenza dei suoi membri come elementi unici e irripetibili. Bosi (2002) sottolinea che la professionalità dell'educatore richiede una profonda comprensione delle proprie dinamiche emotive e della soggettività personale, riconoscendo come queste influenzino gli atteggiamenti e le azioni educative. La relazione educativa è contraddistinta dalla reciprocità tra le parti coinvolte, con processi psicologici che si attivano inevitabilmente. Di conseguenza, l'educatore deve essere sensibile alle dinamiche emotive proprie delle interazioni educative, considerando sia il proprio benessere emotivo sia quello degli altri. La capacità di riconoscere e gestire le proprie emozioni è fondamentale per educare al riconoscimento e al rispetto delle emozioni altrui, promuovendo un ambiente di apprendimento emotivamente sicuro e supportivo. Tale consapevolezza è il risultato di un processo di crescita personale e professionale che richiede impegno costante, auto-riflessione critica e disponibilità a confrontarsi con le proprie vulnerabilità, al fine di sviluppare una maggiore sensibilità emotiva e una più efficace pratica educativa.

Giorgia Altobelli

3. Dall'illusione di manipolare all'essere manipolato

Nel 1966, Joseph Weizenbaum sviluppò ELIZA⁴, uno dei primi chatbot, basato su un algoritmo di ripetizione delle affermazioni dell'utente sotto forma di domanda, ispirato al modello di terapia centrata sulla persona di Carl Rogers. Questo lavoro pionieristico ha contribuito significativamente allo sviluppo di chatbot e sistemi di intelligenza artificiale conversazionali in grado di elaborare il linguaggio naturale basato su testo, fornendo risposte prestabilite in relazione a ciò che gli veniva chiesto, simulando uno psicoterapeuta. Wang definisce i chatbot:

computer programs designed to simulate human communication through text or voice interaction. They are a type of artificial intelligence technology that uses natural language processing (NLP) and machine learning algorithms to understand and respond to user queries. Chatbots can be used in a variety of applications, such as customer service, healthcare, and e-commerce, to provide instant responses and personalized experiences (2024: 57).

La storia dell'IA è caratterizzata da una forte componente antropomorfa, evidente nei riferimenti a elementi umani come 'cervelli' e 'reti neurali', che testimonia la sua intrinseca propensione all'antropomorfismo (Natale, 2021). Negli ultimi cinque decenni, i chatbot hanno subito una trasformazione significativa, passando da sistemi rudimentali a chatbot avanzati con abilità simili a quelle umane, come evidenziato da Rudolph, Tan e Tan (2023). L'aumento della capacità computazionale, la disponibilità di dati digitali e gli avanzamenti nel data mining, nel machine learning e nell'elaborazione del linguaggio naturale hanno contribuito a rinnovare l'interesse per Eliza⁵ (Natale, Ballatore, 2020; Ballatore, Natale, 2023). Nonostante la superficialità dell'interazione, numerosi utenti che utilizzarono inizialmente Eliza, quando il programma fu creato, furono convinti di aver interagito con uno psicologo umano piuttosto che con una macchina. Questo fenomeno, noto come Effetto Eliza, evidenzia la tendenza degli utenti a trattare i programmi di intelligenza artificiale come interlocutori empatici e comprensivi. Le macchine, pur non possedendo una vera consapevolezza dell'inganno, possono

⁴ Il nome Eliza trae ispirazione dalla protagonista del 'Pygmalion' di George Bernard Shaw, dove Eliza Doolittle si trasforma linguisticamente e culturalmente per inserirsi nell'alta società, simboleggiando rinascita e adattamento.

⁵ Il 21 dicembre 2024, Eliza è stata riportata in funzione con successo, dimostrando non solo la piena funzionalità del software, ma anche la conservazione della sua efficacia originaria, un risultato significativo a conferma della validità del progetto.

Giorgia Altobelli

simularlo efficacemente, sfruttando la fiducia dell’utente e compromettendo la relazione tra l’utente e il sistema. Weizenbaum (1976), il creatore di Eliza, ha espresso una forte critica verso l’eccessiva fiducia riposta negli strumenti informatici. Secondo lui, una conoscenza approfondita dei meccanismi interni di un programma ne riduce la complessità percepita e ne demistifica l’aura di misteriosità. L’Effetto Eliza si sta manifestando nuovamente nell’era dei chatbot terapeutici e non, suscitando preoccupazioni riguardo alla percezione di empatia simulata. I chatbot, basati su modelli predefiniti, possono generare dipendenza negli utenti, i quali potrebbero non riconoscere i limiti di tali strumenti, come l’incapacità di comprendere pienamente il contesto emotivo e la complessità della vita. Ciò potrebbe comportare risposte inadeguate e insufficienti, con ripercussioni negative. Inoltre, l’uso crescente di chatbot come compagni emotivi solleva interrogativi sulla loro capacità di contrastare la solitudine e sulla potenziale sostituzione delle relazioni umane. Un recente caso di cronaca⁶ ha evidenziato i possibili rischi di un attaccamento emotivo troppo forte alle tecnologie di compagnia, come le app progettate per simulare relazioni affettive e intime. Sebbene queste applicazioni stiano guadagnando popolarità, è fondamentale indagare se tali strumenti possano effettivamente alleviare la solitudine o se invece contribuiscono a peggiorarla. La personalizzazione delle risposte delle chatbot potrebbe limitare l’opportunità di sperimentazione e responsabilizzazione, aspetti cruciali per lo sviluppo di relazioni significative e per la crescita personale degli utenti.

4. Le competenze socio-emotive e la coscienza dell’IA

Le SES⁷ interessano diverse organizzazioni internazionali, tra cui l’OCSE⁸, che ha distinto le competenze trasversali in competenze sociali ed emotive, inaugurando nel 2017 un progetto di ricerca internazionale⁹ con lo scopo di fornire alle città e ai paesi partecipanti informazioni solide e affidabili sulle competenze sociali ed emotive degli studenti (OECD, 2017). La teoria di Heckman, premio Nobel per l’Economia, offre un framework per calcolare i ritorni economici degli investimenti educativi, evidenziando effetti positivi indotti, tra cui vantaggi economici e sociali, aumento della partecipazione civica e democrazie più attive. Secondo lo scienziato sociale ed

⁶ <https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html>

⁷ Social and Emotional Skills.

⁸ Organizzazione per la Cooperazione e lo Sviluppo Economico.

⁹ Study on Social and Emotional Skills.

Giorgia Altobelli

economista: «le competenze socio-emotive, la salute fisica e mentale, la perseveranza, l'attenzione, la motivazione e la fiducia in sé stessi sono importanti fattori determinanti del successo socio-economico» (Heckman, 2008: 3-4).

La velocità dell'evoluzione tecnologica richiede la capacità di adattarsi ai cambiamenti, affrontare sfide innovative e rispondere efficacemente a nuovi ambienti e tecnologie (Dweck, 2023). Basandosi sul modello teorico di Goleman (1995), risulta essenziale analizzare come le competenze socio-emotive incidano sulle scelte di vita e sulla capacità di pensiero critico, in relazione al contesto socioculturale in cui l'individuo si sviluppa. Il sistema degli algoritmi influenza la vita quotidiana, ma la loro progettazione non neutrale può avere effetti discriminatori e negativi sulla società, soprattutto se utilizzati per gestire questioni sociali (Eubanks, 2017; O'Neil, 2017). Alcuni studiosi invitano a non enfatizzare eccessivamente “il dramma algoritmico”, ovvero l’idea che le piattaforme digitali siano entità onnipotenti e incomprensibili, al fine di mantenere una prospettiva più equilibrata e realistica sulla loro influenza e sui loro limiti (Ziewitz, 2016). Gli algoritmi operano secondo programmi predefiniti, privi di coscienza, ignorando implicazioni etiche e morali, e risultano vulnerabili a errori o malfunzionamenti. La responsabilità permane in capo all’essere umano, che progetta, sviluppa e gestisce l’algoritmo, ed è essenziale che mantenga la propria autonomia decisionale e libertà. L’essere umano esperisce il mondo attraverso una complessa interazione di sentimenti ed emozioni, che si originano dall’interiorità e si manifestano esternamente. Secondo Faggin (2024) la natura intrinsecamente soggettiva dell’esperienza cosciente pone ostacoli significativi alla sua quantificazione e spiegazione scientifica. La consapevolezza che le narrazioni sulle tecnologie plasmano il nostro rapporto con esse suggerisce che il modo in cui le rappresentiamo abbia un impatto significativo sul loro utilizzo (Berger, Luckmann, 1966). Nell’ambito della rapida evoluzione tecnologica e dell’intelligenza artificiale, la regolamentazione risulta spesso inadeguata, evidenziando vulnerabilità e preoccupazioni legate alla manipolazione emotiva e cognitiva. Gli algoritmi, creati da esseri umani con limitazioni e bias cognitivi, possono influenzare la percezione della realtà e la libertà decisionale degli utenti. Si pone quindi la questione sulla consapevolezza degli utenti riguardo alle implicazioni dei prodotti utilizzati e sulla possibilità di garantire esperienze protette da errori e abusi. Per affrontare queste sfide, è cruciale chiedersi: come possiamo assicurarci che i sistemi di IA siano progettati per tutelare la libertà decisionale degli utenti? E, di conseguenza, in che modo è possibile sviluppare un approccio umano-centrico che bilanci innovazione tecnologica e protezione degli utenti?

Giorgia Altobelli

La distinzione fondamentale tra esseri umani e macchine non è oggetto di dibattito; tuttavia, l'idea che le macchine possano essere equiparate agli esseri umani solleva interrogativi sulla potenziale perdita di identità e auto-coscienza umana.

Riferimenti bibliografici

- Ballatore A., Natale S. (2023). Technological failures, controversies and the myth of AI. In Lindgren S. (ed.), *Handbook of Critical Studies of Artificial Intelligence* (pp. 237-244). Northampton (MA): Edward Elgar Publishing.
- Berger P.L., Luckmann T. (1966). *The social construction of reality: a treatise in the sociology of knowledge*. New York: Anchor Books.
- Borgna E. (1999). *Noi siamo un colloquio*. Milano: Feltrinelli.
- Bosi R. (2002). *Pedagogia al nido. Sentimenti e relazioni*. Roma: Carocci.
- Brazelton T., Greenspan S. (2001). *I bisogni irrinunciabili dei bambini*. Milano: Raffaello Cortina.
- Bucchi M. (2020). *Io & Tech, piccoli esercizi di tecnologia*. Milano: Bompiani.
- Cohen M.H., Giangola J.P., Balogh J. (2004). *Voice user interface design*. Boston: Addison-Wesley Professional.
- Dweck C.S. (2023). *Mindset. Cambiare forma mentis per raggiungere il successo*. Milano: FrancoAngeli.
- Eubanks V. (2017). *Automating inequality: how high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.
- Faggini F. (2024). *Oltre l'invisibile. Dove scienza e spiritualità si uniscono*. Milano: Mondadori.
- Floridi L. (2022). *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*. Milano: Raffaello Cortina.
- Goleman D. (1995). *Emotional intelligence: why it can matter more than IQ*. New York: Bantam Books.
- Heckman J.J. (2008). Schools, skills, and synapses. *SSRN Scholarly Paper* ID 1139905. Rochester (NY): Social Science Research Network.
- Lavanga F., Mancaniello M.R. (2022). *Formazione dell'adolescente nella realtà estesa. La pedagogia dell'adolescenza nel tempo della realtà virtuale, dell'intelligenza artificiale e del metaverso*. Genova: Libreriauniversitaria.it.
- Locatelli L., Pavone S., Belvedere G.C., Aldi G., Coccagna A. (2013). *Un'altra scuola è possibile*. Roma: Edizioni Enea.
- Mariani L. (2006). *La motivazione a scuola. Prospettive teoriche e interventi strategici*. Roma: Carocci.
- McCarthy J., Minsky M.L., Rochester N., Shannon C.E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence.
- Milani L. (1990). *Lettera a una professoressa*. Firenze: Libreria Editrice Fiorentina.
- Natale S. (2021). *Deceitful media: artificial intelligence and social life after the Turing Test*. Oxford: Oxford University Press.
- Natale S., Ballatore A. (2020). Imagining the thinking machine: technological myths and the rise of artificial intelligence. *Convergence*, 26(1): 3-18.

Giorgia Altobelli

- O’Neil C. (2017). *Weapons of math destruction: how big data increases inequality and threatens democracy*. London: Penguin Books.
- OECD (2017). *Social and emotional skills: well-being, connectedness and success*. Paris: OECD.
- Picard R.W. (1997). *Affective computing*. Cambridge (MA): MIT Press.
- Rogers C. (1951). *Client-centered therapy*. Boston: Houghton Mifflin.
- Rudolph J., Tan S., Tan S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1): 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Seldon A., Abidoye O. (2018). *The fourth education revolution. Will artificial intelligence liberate or infantilise humanity*. London: University of Buckingham Press.
- Turing A.M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1): 230-265.
- Turing A.M. (1950). Computing machinery and intelligence. *Mind*, 59(236): 433-460.
- van Deursen A.J.A.M., Helsper E.J. (2015). The third-level digital divide: who benefits most from being online? *Communication and Information Technologies Annual*, 10: 29-52.
- Vygotskij L.S. (1987). *Il processo cognitivo*. Torino: Bollati Boringhieri.
- Wang K. (2024). From ELIZA to ChatGPT: a brief history of chatbots and their evolution. *Applied and Computational Engineering*, 39: 57-62. <https://doi.org/10.54254/2755-2721/39/20230579>
- 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. <https://doi.org/10.1145/365153.365168>
- Weizenbaum J. (1976). *Computer power and human reason: from judgment to calculation*. San Francisco: W.H. Freeman.
- Ziewitz M. (2016). Governing algorithms: myth, mess, and methods. *Science, Technology, & Human Values*, 41(1): 3-16.