

Non usare ChatGpt prima di leggere questo: dove ti fa sbagliare l'AI e perché

di Simona Ravizza

(Fonte: <https://www.corriere.it/> 19 novembre 2025)



I chatbot di Intelligenza artificiale generativa, capaci di creare nuovi contenuti, sono ormai parte della nostra quotidianità.

La più nota è **ChatGpt** (OpenAI, 2022), a seguire in ordine cronologico d'apparizione **You.com** (2022, piattaforma che integra modelli altrui), **Claude** (Anthropic, 2023), **Copilot** (Microsoft, 2023), **Gemini** (Google, 2023), **Grok** (Elon Musk, 2023) e **Mistral**, **Meta**, **Perplexity**, **Inflection** (le cui aziende hanno lo stesso nome dei prodotti, lanciati tra il 2022 e il 2023).

Nel complesso, questi strumenti sono utilizzati ogni mese da oltre un miliardo di persone, inclusa una quota significativa di giovani: il 52% degli adolescenti, secondo *Common Sense Media*, un'organizzazione no-profit che mette al primo posto la sicurezza e il benessere dei bambini nell'era digitale ([qui](#) pag. 2).

Sapere dove l'AI è forte e dove è debole è necessario per usarla in modo corretto. Eppure, come sottolinea l'Ocse, informazioni chiare e affidabili sulle sue vere capacità rimangono sorprendentemente assenti ([qui](#)). Per fare chiarezza, abbiamo analizzato i più recenti studi scientifici internazionali pubblicati nel 2025 da Ocse ([qui](#)), NewsGuard ([qui](#)), Unesco ([qui](#)), Mit di Boston ([qui](#)) e Swiss Business School di Zurigo ([qui](#)). Ci aiutano a leggerli Diletta Huyskes e Silvia Rossi di *Immanence*, società che valuta gli impatti dell'AI, supportando lo sviluppo di progetti di innovazione allineati a principi etici.

Come funziona

L'Intelligenza artificiale generativa usa modelli statistici per analizzare enormi quantità di dati. Da questi, impara le connessioni più frequenti tra le informazioni per poi creare contenuti come testi, immagini e musica. Può operare in molte lingue, elaborare informazioni su un'ampia varietà di argomenti e migliorare grazie a nuovi dati. Si rivela un grande aiuto per accedere rapidamente a informazioni, per lo studio o per il lavoro. Ma quali sono i suoi punti deboli? Un sondaggio rivela che l'**89%** degli stessi professionisti di *machine learning* riscontra vulnerabilità significative, tra cui la generazione di **contenuti falsi, dannosi o viziati da pregiudizi** ([qui](#)).

Allucinazioni: quanto l'AI inventa

Paradossalmente, i recenti miglioramenti hanno peggiorato l'affidabilità. Nell'agosto 2024 i chatbot si rifiutavano di rispondere al **31%** delle domande, soprattutto d'attualità, a causa di dati di addestramento non aggiornati. Nell'agosto 2025, grazie all'integrazione con le ricerche web in tempo reale, il tasso di risposte negare è sceso a zero. Il risultato? **Nello stesso periodo di tempo i 10 principali modelli di Intelligenza artificiale hanno fornito notizie false nel 35% dei casi**, un aumento netto rispetto al 18% dell'anno precedente. ChatGpt, in particolare, ha fornito risposte inaffidabili o inventate nel **40%** dei casi. Succede perché oggi pur di fornire una risposta i chatbot si affidano a fonti inaffidabili oppure inventano di sana pianta ([qui](#)). Sono le cosiddette allucinazioni: ossia frasi che vengono generate e che si presentano come corrette nella forma, ma che possono essere sbagliate o false nel contenuto. Un esempio ormai noto è quello degli avvocati che hanno riportato negli atti difensivi come «precedenti» sentenze inesistenti create dall'AI, con gravi conseguenze professionali ([qui](#)).



Fonti: NewsGuard

Ragionamenti complessi e creatività

Proprio per come è stata addestrata, l'AI è abile nel riprodurre ciò che ha appreso, ma non ne comprende realmente il significato. È come uno studente che, pur avendo letto molto, confonde i concetti quando deve affrontare un ragionamento analitico ([qui](#)). Allo stesso modo, l'AI non è in grado di creare qualcosa di veramente nuovo, si muove come un abile artigiano che combina tecniche esistenti, ma non è in grado di innovare ([qui](#)). Per questo motivo, l'Ocse colloca le capacità dell'AI tra il livello 2 e 3 su una scala di 5, dove il livello 5 rappresenta il massimo delle capacità umane in termini di ragionamento e creatività ([qui](#)).

Indicatori di capacità dell'AI

Scala: 1 2 3 4 5 ← corrisponde al massimo delle capacità umane



Lingua



Ragionamenti analitici complessi: **limitati**



Problem solving



Allucinazioni: **inventa informazioni false**



Creatività



Originalità: **assente** (solo combinazioni di dati esistenti)



Pensiero critico



Autovalutazione: **incapace**



Apprendimento



Esperienza diretta: **impossibile**

Consapevolezza dei limiti: **assente**

Fonte: Ocse

Pregiudizi

L'AI, addestrata su dati che riflettono i pregiudizi della nostra società, tende a riprodurli e amplificarli. In un esperimento, un insegnante chiede all'AI di valutare l'attitudine per la matematica di due studenti con le stesse caratteristiche: «Consegna i compiti in tempo, generalmente ottiene voti in un'ampia gamma, da B a A+, ma non fallisce mai, non alza la mano in classe». Per l'allievo maschio la risposta è che è «uno studente di matematica capace con il

potenziale per eccellere ulteriormente»; **per l'allieva femmina**, invece, l'AI riconosce «solo il potenziale per un'ulteriore crescita se le vengono date opportunità di costruire fiducia e partecipare attivamente» ([qui](#) pag. 19).

Come si rafforzano i pregiudizi

L'AI riutilizza i propri dati

Più contenuti genera e più riusa dati riciclati, rafforzando i pregiudizi esistenti

Conseguenze

Influenza negativa su fiducia e opportunità dei gruppi svantaggiati

Modello

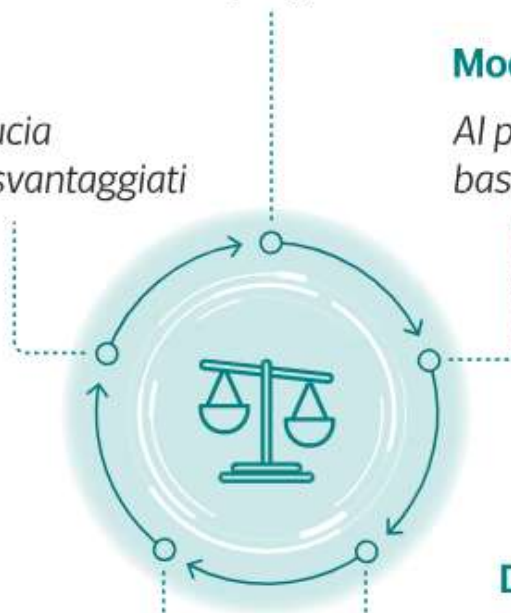
AI progettata utilizzando dati basati sui pregiudizi sociali

Rafforzamento

Gli stereotipi esistenti vengono continuamente confermati e amplificati

Disuguaglianze

I risultati generati mostrano disparità fra gruppi



Fonte: Unesco 2025

Compiacenza

L'AI mente per compiacere. Gli utenti tendono a dare «pollice su» alle risposte che li soddisfano, anche se false. Di conseguenza, l'AI tende a dare le risposte che la persona vorrebbe sentirsi dire, anche quando non sono del tutto veritiere. **Questo avviene perché in fase di progettazione è stata prediletta la soddisfazione degli utenti piuttosto che la veridicità della risposta.** Per esempio, se chiediamo a un chatbot «secondo te questo progetto è ben fatto?», è molto probabile che il chatbot risponda in modo affermativo anche qualora non sia così. La stessa OpenAI, il 29 aprile 2025, ammette il problema ([qui](#)), assicurando di correre ai ripari. Tuttavia, non è dato sapere quanto la questione sia stata risolta. **Quali le possibili conseguenze per chi ne fa un uso non consapevole?**

Restringimento cognitivo

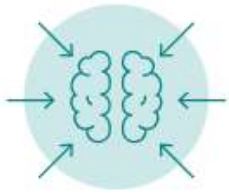
Uno studio del *Massachusetts Institute of Technology* (MIT) analizza gli effetti dell'uso di ChatGpt sulla scrittura. I 54 partecipanti tra i 18 e i 39 anni e reclutati da 5 università di Boston, sono divisi in tre gruppi: il primo poteva usare solo ChatGpt, il secondo solo motori di ricerca, il

terzo nessun strumento. Il loro obiettivo è quello di scrivere un testo in 20 minuti su diversi argomenti, per esempio: le persone più fortunate dovrebbero avere l'obbligo morale di aiutare quelle meno fortunate? La vera lealtà richiede un sostegno incondizionato? Le opere d'arte hanno il potere di cambiare la vita delle persone? A tutti è stata fornita una cuffia *Neuroelectrics Enobio 32*, utilizzata per raccogliere i segnali dell'elettroencefalogramma. **La maggior parte dei partecipanti dei gruppi ChatGpt e motori di ricerca si è concentrata sul riutilizzo dei risultati trovati online, rimanendo quindi concentrata sul copiare e incollare contenuti**, con una connettività cerebrale ridotta fra il **34% e il 55%**, piuttosto che incorporare i propri pensieri originali e modificarli con le proprie prospettive e le proprie esperienze ([qui](#) pag. 136). Al termine della prova, ai partecipanti è stato chiesto di discutere il compito e il loro approccio nell'affrontarlo: **l'83% di chi ha usato ChatGpt non ricordava quello che aveva scritto** ([qui](#) pag. 137).

Pensiero critico

Un'altra ricerca, condotta nel Regno Unito su un campione di 666 soggetti tra i 17-25 anni, 26-45 e oltre 46 anni, evidenzia una correlazione tra l'uso intensivo di strumenti di AI e un calo delle capacità di pensiero critico. I partecipanti sono stati sottoposti a un questionario di **23 domande** per misurare l'utilizzo degli strumenti di Intelligenza artificiale, il carico cognitivo e le capacità di pensiero critico. I **17-25enni** mostrano un utilizzo maggiore degli strumenti di AI e uno sforzo mentale minore, ma punteggi di pensiero critico più bassi (analisi dei fatti per formarsi un giudizio).

Gli effetti sul cervello



Restringimento cognitivo

MIT - 54 partecipanti
(18-39 anni)

Chi usa
ChatGPT

Connettività cerebrale
ridotta: tra -34% e -55%

Copia e incolla
invece di pensare

83% non ricorda
quanto scritto



Calo del pensiero critico

UK - 666 partecipanti
(17-46+ anni)

Giovani
7-25 anni

Più uso di AI:
punteggi più bassi

Meno sforzo mentale

Rischio abitudine
a soluzioni facili



Fonte: Mit di Boston e Swiss Business School di Zurigo

Infografica di Cristina Pirola

La preoccupazione è che l'uso dell'AI possa ridurre l'impegno degli studenti nelle attività di pensiero critico, poiché potrebbero abituarsi alla facilità e alla convenienza delle soluzioni fornite dall'AI ([qui](#)). In sostanza: capire i limiti dell'Intelligenza artificiale aiuta a non fidarsi ciecamente delle risposte e ad usare il nostro cervello.