

# EduRag.

## Una proposta per usare l'intelligenza artificiale in ambito educativo

Daniele Dragoni\*, Maria Désirée Epure\*\*

*Riassunto:* L'articolo esplora il rapporto tra l'intelligenza artificiale e l'educazione, focalizzandosi sui *Large Language Models* e sulle opportunità e criticità che questi strumenti offrono. Dopo aver analizzato le architetture che hanno reso possibile il loro successo, come il *Transformer*, vengono esaminate le problematiche legate alla trasparenza e all'affidabilità delle risposte generate, con particolare attenzione al fenomeno delle "allucinazioni" dei modelli. Vengono poi illustrate strategie per mitigare tali criticità, come la *Chain of Thought* e la *Retrieval-Augmented Generation*, che migliorano la trasparenza e l'affidabilità degli LLM. Il progetto EduRag rappresenta un'applicazione pratica di queste strategie, dimostrando come l'IA possa essere utilizzata efficacemente nell'ambito educativo per supportare l'apprendimento e promuovere l'uso di fonti attendibili. Infine, si delineano possibili scenari futuri in cui l'IA giocherà un ruolo centrale nell'educazione, con strumenti sempre più personalizzati, interattivi e multimodali, in grado di migliorare l'esperienza didattica attraverso un'integrazione più profonda con il linguaggio, la scrittura e altre forme di interazione umana.

*Parole chiave:* intelligenza artificiale, LLM, RAG, *prompt engineering*, EduRag.

*English title:* EduRag. A Proposal to Use Artificial Intelligence in Education

*Abstract:* This article explores the relationship between artificial intelligence and education, focusing on Large Language Models and the opportunities and challenges these tools present. After analysing the architectures that have enabled their success, such as the Transformer, the issues related to the transparency and reliability of the generated responses are examined, with particular attention to the phenomenon of model "hallucinations". Strategies to mitigate these challenges, such as *Chain of Thought* and *Retrieval-Augmented Generation*, are then illustrated, which improve the transparency and reliability of LLMs. The EduRag project represents a practical application of these strategies, demonstrating how AI can be effectively used in the educational field to support learning and promote the use of reliable sources. Finally, possible future scenarios are outlined in which AI will play a central role in education, with increasingly personalized, interactive, and multimodal tools capable of enhancing the educational experience through deeper integration with language, writing, and other forms of human interaction.

*Keywords:* Artificial intelligence, LLMs, RAG, Prompt engineering, EduRag.

\* Università di Macerata. Email: d.dragoni1@unimc.it

\*\* Università di Roma Tre. Email: mar.epure@stud.uniroma3.it

L'articolo è frutto della collaborazione tra Autore 1 e Autore 2, che hanno contribuito rispettivamente come segue:....

## 1 Introduzione

Il presente articolo esamina il rapporto tra intelligenza artificiale (IA) ed educazione, focalizzandosi sul contributo dei *Large Language Models* (LLM) (Naveed et al., 2023). Questi modelli, addestrati su vasti insiemi di dati, sono in grado di generare testi autonomamente, riconoscendo schemi complessi e fornendo risposte articolate e coerenti. Tuttavia, il loro funzionamento rimane spesso opaco e solleva preoccupazioni, poiché non è sempre chiaro il processo attraverso cui generano risposte, con il rischio di produrre informazioni inesatte o non verificate, note come “allucinazioni” (Chen et al., 2023). A ciò si aggiunge una generale diffidenza verso le nuove tecnologie, poiché ogni grande innovazione comporta profonde trasformazioni sociali. Un tempo lo è stata la ferrovia, oggi lo è l’IA.

Nel prosieguo dell’articolo, proporrò anzitutto alcune riflessioni sui timori legati all’introduzione delle nuove tecnologie, ma anche sulle opportunità che queste possono offrire per migliorare l’apprendimento. Fornirò inoltre una breve panoramica della storia dell’IA e delle innovazioni che ne hanno favorito il significativo avanzamento negli ultimi anni. Successivamente, ci concentreremo sulle strategie per migliorare il *prompting*, ovvero il processo di fornire *input* a un LLM per ottenere risposte più precise e pertinenti, e sugli strumenti che consentono a tali modelli di operare su dati affidabili, offrendo agli utenti la possibilità di risalire alle fonti delle informazioni.

Inoltre, presenteremo il progetto EduRag<sup>1</sup>, un *bot* che può essere considerato un piccolo laboratorio per testare l’efficacia dell’IA nell’educazione, consentendo a docenti e studenti di sperimentare nuove modalità di insegnamento e apprendimento basate su contenuti affidabili e interazioni personalizzate.

Infine, esploreremo scenari futuri in cui l’IA potrebbe giocare un ruolo centrale nell’educazione, fornendo strumenti sempre più personalizzati, interattivi e multimodali, in grado di migliorare l’esperienza didattica attraverso un’integrazione più profonda con il linguaggio, la scrittura e altre forme di interazione umana.

## 2 Nuove tecnologie e apprendimento: timori e opportunità

Come già accennato, l’introduzione di nuove tecnologie spesso suscita timori e resistenze, poiché può cambiare profondamente la struttura della società. Un esempio storico emblematico è la creazione della *Liverpool and Manchester Railway* nel 1830, considerata la prima ferrovia nella sua forma moderna. Come riportato da Giliberto (2013), l’introduzione di questa innovazione generò conflitti tra conservatori e progressisti, alimentando paure riguardo alla velocità “incredibile” del treno, ai possibili danni agli animali e all’impatto sul lavoro di cocchieri

<sup>1</sup> <https://github.com/nugh75/EduRag.git>.

e marinai. Il tragico incidente durante l'inaugurazione della tratta, che causò la morte di Sir William Huskisson, uno dei più ferventi sostenitori della tecnologia ferroviaria, accentuò ulteriormente le preoccupazioni del tempo.

In tutt'altro ambito, George Miller Beard, nel suo libro *American Nervousness* (1888) criticava la “perfezione” degli orologi da polso, sostenendo che potessero provocare stati di nervosismo e che “un ritardo di pochi attimi potrebbe distruggere tutta una vita”.

Resistenze simili si sono osservate fin dall'antichità. Nel *Fedro* (IV secolo a.C.), Platone esprimeva forti perplessità nei confronti della scrittura, sostenendo che non favorisse la memoria e inducesse le persone a credere di conoscere ciò che in realtà avevano solo letto.

In tempi più recenti, lo psichiatra Manfred Spitzer (2013) ha definito gli strumenti digitali come “macchine che ostacolano l'apprendimento”, parlando di “demenza digitale”, mentre Nicholas Carr (2011) ha sostenuto che internet ci renderà inesorabilmente “stupidi”. Al contrario, altri studiosi hanno evidenziato il ruolo positivo delle tecnologie nello sviluppo cognitivo. Lev Vygotskij (1934) sottolineava l'importanza degli strumenti tecnologici nel processo di apprendimento, evidenziando come possano estendere le capacità cognitive umane. B.F. Skinner (1961) proponeva la costruzione di “macchine per insegnare” per affiancare o sostituire gli insegnanti, anticipando l'uso di tecnologie educative personalizzate.

Per quanto riguarda l'IA, alcuni timori sembrano ben fondati<sup>2</sup>, tant'è che l'Unione Europea, con l'*AI Act* (Parlamento Europeo e Consiglio dell'Unione Europea, 2024), considera l'uso di queste tecnologie nell'educazione ad alto rischio e ne vieta l'applicazione per il *social scoring* e la valutazione automatica degli studenti.

D'altra parte, l'IA offre numerose opportunità per migliorare l'esperienza di apprendimento. Può fornire supporto personalizzato, adattando i contenuti alle esigenze specifiche degli studenti e offrendo un *feedback* immediato. Gli assistenti virtuali basati sull'IA possono fungere da tutor individuali, rispondendo alle domande degli studenti, aiutandoli a comprendere concetti complessi, offrendo loro spiegazioni teoriche ed esempi pratici, adattando l'approccio alle necessità di ciascun apprendente. Inoltre, i docenti possono utilizzare l'IA per monitorare i progressi degli studenti, identificando aree di difficoltà e suggerendo interventi mirati. L'IA può anche automatizzare attività amministrative, come la gestione dei voti e delle presenze, liberando tempo per una maggiore interazione tra insegnanti e studenti. E ancora, può analizzare i dati educativi per identificare tendenze nell'apprendimento e supportare decisioni strategiche, sia sul piano didattico che organizzativo.

<sup>2</sup> Per una riflessione ricca e approfondita sulle problematiche dell'IA, rimandiamo al testo di Teresa Numerico *Big Data e algoritmi* (2021), che analizza in modo completo e critico queste problematiche.

### ③ L'evoluzione dell'IA: dalle origini agli sviluppi recenti

Prima di presentare il progetto EduRag, che esplora l'applicazione dell'intelligenza artificiale nell'ambito educativo, è importante comprendere quali innovazioni hanno permesso ai nuovi modelli linguistici di simulare il linguaggio umano in modo accurato e realistico. Queste innovazioni verranno illustrate nei paragrafi seguenti per evidenziare il loro impatto sull'educazione e sull'apprendimento.

#### *Un breve excursus*

L'IA è un campo dell'informatica dedicato alla creazione di sistemi in grado di eseguire compiti che richiedono intelligenza umana, come il riconoscimento visivo, l'elaborazione del linguaggio naturale, il processo decisionale e l'apprendimento. Dalla sua nascita negli anni '50 (McCarthy et al., 1955), l'IA ha attraversato diverse fasi di sviluppo, caratterizzate da periodi di entusiasmo, le cosiddette "primavere", alternati a periodi di stasi, noti come "inverni", a seconda dell'interesse e dei finanziamenti disponibili (Toosi et al., 2021).

I successi iniziali degli anni '50 e '60 furono seguiti da periodi di rallentamento, causati dalle aspettative eccessive e dalle limitate capacità tecnologiche dell'epoca. Negli anni '80 e '90, tuttavia, l'IA ha ritrovato slancio grazie all'introduzione del *machine learning*, tecnica che permette ai computer di apprendere dai dati senza essere esplicitamente programmati, e dalle *reti neurali*, modelli computazionali ispirati al cervello umano (Gori, 2003).

Questo rinnovato progresso è stato ulteriormente amplificato nel XXI secolo dall'avvento dei *big data* e dai miglioramenti nell'*hardware* computazionale, che hanno reso possibile lo sviluppo del *deep learning*, sottocampo del *machine learning* che utilizza reti neurali profonde, ossia modelli computazionali con molteplici strati, per elaborare grandi quantità di dati complessi.

Nel 2023, abbiamo assistito non solo a una nuova primavera, ma a una vera e propria estate dell'IA. Il lancio di ChatGPT ha segnato un punto di svolta, avvicinando l'IA al grande pubblico come mai prima d'allora. Sebbene i risultati iniziali fossero imperfetti, e talvolta persino comici, i miglioramenti sono stati rapidi e costanti. Questo successo ha spianato la strada al lancio di molti altri strumenti, come generatori di immagini, assistenti virtuali e sistemi di automazione per vari settori, aumentando ulteriormente l'impatto dell'IA sul mercato globale.

Ogni giorno, questi strumenti diventano più performanti, raffinati ed efficaci. Un fattore chiave alla base di questa straordinaria evoluzione è l'architettura *Transformer*, che ha rivoluzionato il modo in cui i modelli di IA comprendono e generano il linguaggio.

#### *L'architettura Transformer*

I recenti progressi nel campo dell'IA si basano su diversi fattori, ma un punto di svolta fondamentale è stata l'introduzione dell'architettura *Transformer* (Vaswani

et al., 2017). Al cuore di questa architettura c'è il *meccanismo di attenzione*, che valuta la relazione e l'importanza di ciascuna parola all'interno di una frase, migliorando la comprensione globale del testo. Questo meccanismo consente ai modelli di prevedere con maggiore precisione quale parola o frase sia più probabile in un determinato contesto, grazie all'analisi di vasti *dataset* linguistici.

In pratica, quando un LLM come GPT genera testo (la "T" indica appunto *Transformer*), analizza l'intera sequenza di parole fornite e calcola la probabilità che una parola segua l'altra, creando una struttura coerente e logicamente sequenziale. Questo approccio permette al modello di generare testi che rispecchiano il flusso e le sfumature del linguaggio umano, tenendo conto sia del significato letterale delle parole che delle loro connessioni semantiche.

### *Il problema delle risposte imprecise o "allucinazioni"*

Un problema evidente nell'utilizzo di questi modelli è la produzione di risposte scorrette o fantasiose. Questo fenomeno si verifica perché spesso si tende a considerare questi strumenti come tuttologi o oracoli (Vespignani & Rijitano, 2019), quando in realtà sono produttori di testo che generano risposte basandosi sui *prompt* che ricevono. Di conseguenza, se si pone una domanda banale, è molto probabile che si ottenga una risposta altrettanto banale.

Inoltre, come mostrato nel film *A Beautiful Mind* (Howard, 2002), che racconta la vita del matematico John Nash, questi modelli sono molto abili nel riconoscere *pattern* e ricorrenze, ma possono anche avere visioni distorte della realtà, proprio come le allucinazioni di Nash. Per mitigare questi errori, è necessario adottare delle strategie precise, esattamente come Nash ha dovuto sviluppare soluzioni per gestire le sue visioni.

Le strategie più efficaci includono l'uso di *prompt* (Schulhoff et al., 2024) come "agisci come" per dare al modello un ruolo preciso, fornire un contesto dettagliato, imporre vincoli e chiedere di esplicitare il ragionamento *step by step*. Ad esempio, si può usare un *prompt* del tipo: "Prima di procedere, fammi delle domande per assicurarti di aver compreso ciò che ti ho chiesto". Spesso, infatti, è l'utente a non formulare chiaramente la domanda, il che rende difficile per il modello fornire una risposta adeguata. Tutte queste strategie e tecniche rientrano sotto l'etichetta di *prompt engineering*.

### *Prompt engineering: ottimizzare le richieste per migliorare le risposte*

Negli ultimi anni, numerosi studi hanno analizzato il *prompt engineering*, concentrandosi su come ottimizzare la formulazione dei *prompt* per migliorare l'efficacia e la precisione delle risposte fornite dagli LLM. Tra le strategie più rilevanti, spiccano la *Chain of Thought* (CoT) e il *Tree of Thoughts* (ToT).

La CoT (Wei et al., 2022) è una strategia che aiuta gli LLM a ragionare su problemi complessi, suddividendo il processo in passaggi logici intermedi. Invece di fornire

direttamente la risposta finale, il modello viene istruito a scomporre il ragionamento in piccoli passaggi, affrontandoli uno alla volta fino a raggiungere la soluzione.

Questa tecnica è particolarmente utile quando il problema richiede più passaggi per essere risolto. Per esempio, se il modello deve risolvere un problema matematico come *Roger ha 5 palline, ne compra altre 6 in due lattine da 3 palline ciascuna. Quante palline ha in totale?*, invece di fornire direttamente la risposta finale, il modello procederà per gradi. Prima calcolerà quante palline ci sono in totale nelle lattine, poi sommerà questo risultato alle palline iniziali di Roger, arrivando infine alla risposta corretta di 11 palline.

Oltre a migliorare l'efficienza nel risolvere problemi complessi, questa strategia rende più trasparente il processo decisionale del modello, poiché espone chiaramente i passaggi intermedi che lo hanno portato alla risposta. Questo permette anche di identificare eventuali errori lungo il percorso e correggerli.

Il grande vantaggio di questa tecnica è che non richiede un ulteriore addestramento specifico del modello, ma sfrutta semplicemente esempi di ragionamento in contesto, permettendo di affrontare diversi compiti senza bisogno di modificare o addestrare il modello per ogni tipo di problema.

Il ToT (Yao et al., 2024) è una metodologia che sfrutta l'esplorazione simultanea di più percorsi di ragionamento. Funziona suddividendo un problema in unità chiamate *thoughts* (pensieri), che rappresentano possibili passi intermedi verso la soluzione. A differenza di approcci più lineari, ToT consente di valutare simultaneamente diversi percorsi logici, ottimizzando il processo decisionale. Se un percorso risulta inefficace, può essere scartato, mentre quelli validi vengono esplorati ulteriormente, permettendo di arrivare alla soluzione in modo più efficiente.

ToT è particolarmente utile in contesti che richiedono flessibilità, come la pianificazione strategica o il *problem solving* complesso.

Un esempio concreto dell'applicazione di ToT è il gioco del 24, in cui l'obiettivo è ottenere il numero 24 utilizzando quattro numeri e operazioni matematiche. Utilizzando ToT, il modello esplora simultaneamente diverse combinazioni di operazioni aritmetiche come addizione, sottrazione, moltiplicazione e divisione. In questo modo, valuta ogni possibile percorso che potrebbe portare al risultato desiderato. Se un percorso non funziona, viene scartato, e il modello continua a esplorare altri percorsi fino a trovare la combinazione corretta.

### *Lavorare con grandi volumi di dati nei modelli di IA*

Un'altra sfida rilevante da affrontare riguarda la limitazione della finestra di contesto negli LLM, che impone un limite alla quantità di dati che possono essere inseriti nel *prompt*. Questo problema diventa particolarmente critico quando si desidera lavorare con un grande volume di informazioni, come centinaia di articoli, libri, appunti di diversa lunghezza e natura. Tali materiali possono facilmente eccedere la capacità di elaborazione del modello, rendendo difficile includere tutte le fonti rilevanti in un'unica richiesta.

Per superare questa limitazione si possono adottare principalmente due approcci: la selezione manuale e le tecniche automatiche.

Nel caso della selezione manuale, l'utente esamina direttamente il materiale a disposizione e seleziona quali parti includere nel *prompt*, scegliendo le informazioni più rilevanti. Alcuni modelli offrono anche la possibilità di caricare *file* direttamente per essere elaborati, rendendo il processo più fluido. Questo approccio garantisce un controllo preciso sui contenuti, poiché consente di scegliere con cura le informazioni che si ritiene possano essere utili al modello. Tuttavia, la selezione manuale può diventare onerosa e richiedere molto tempo, soprattutto quando si devono gestire grandi quantità di dati o materiali molto eterogenei.

Retrieval-Augmented Generation: *combinare ricerca e generazione di testo*

Un approccio alternativo per superare la limitazione della finestra di contesto è l'uso di tecniche automatiche, note come *Retrieval-Augmented Generation* (RAG) (Lewis et al., 2021). Questo metodo consente al modello di interagire con un vasto archivio di dati, come articoli o documenti, e di recuperare solo le informazioni rilevanti per il compito in questione. In pratica, il modello prima esegue una ricerca in un *database* (fase di *retrieval*), identificando i contenuti pertinenti, e successivamente genera una risposta o un testo che integra in modo coerente tali informazioni (fase di *generation*).

La tecnica del RAG può essere articolata in diversi passaggi e componenti. Il primo è il *chunking* (Ramshaw & Marcus, 1995), che consiste nel suddividere il testo in segmenti più piccoli e gestibili, detti *chunk*, che rappresentano unità di testo coerenti e semantiche che possono essere elaborate in modo indipendente e riutilizzate nella fase di *generation*.

Dopo che i dati sono stati suddivisi in *chunk*, viene impiegata la tecnica dell'*embedding* (Mikolov et al., 2013; Neelakantan et al., 2022). I modelli di *embedding* trasformano i *chunk* in rappresentazioni numeriche, ovvero *vettori multidimensionali* che ne riflettono il valore semantico.

Un ulteriore passaggio cruciale è l'*indicizzazione* dei vettori. L'indicizzazione organizza i vettori in modo che possano essere recuperati rapidamente. Strumenti come *Facebook AI Similarity Search* (FAISS) (Johnson et al., 2017) sono comunemente usati per questo scopo, permettendo ricerche efficienti anche in grandi *dataset*. FAISS consente di trovare rapidamente vettori simili, garantendo un recupero delle informazioni veloce e accurato, il che è essenziale per la fase di *retrieval* all'interno del processo RAG.

#### ④ Il progetto EduRag

Il progetto EduRag si propone di sfruttare le tecniche di *prompting* e RAG per migliorare l'interazione con gli LLM. EduRag è un programma scritto in Python

che utilizza due librerie principali: *Streamlit* («Streamlit», 2024) e *LangChain* («LangChain», 2024).

Streamlit permette di creare un'interfaccia grafica intuitiva in HTML, facilitando l'interazione con l'utente finale. LangChain, invece, rappresenta il nucleo centrale del progetto, poiché consente di interagire con vari modelli LLM, inviando loro *prompt* o sequenze di *prompt*. Inoltre, all'interno dello *script* Python di EduRag è possibile integrare l'uso degli LLM con le tecniche RAG, migliorando ulteriormente l'efficacia del sistema nell'elaborazione e nel recupero di informazioni.

EduRag è un sistema progettato per offrire numerose funzionalità avanzate che facilitano la gestione e l'utilizzo di materiali didattici e scientifici. La sua funzione principale è la creazione di *database* indicizzati a partire da documenti di vari formati, come PDF, DOC e TXT. Questi documenti possono includere articoli scientifici, appunti personali, libri e altri materiali rilevanti. Una volta che i documenti sono stati indicizzati, EduRag permette di effettuare interrogazioni mirate su di essi, consentendo agli utenti di recuperare rapidamente le informazioni più pertinenti.

Oltre a questa sua funzione principale, EduRag offre una serie di utilità aggiuntive che arricchiscono l'esperienza di apprendimento. Tra queste, troviamo la possibilità di generare riassunti personalizzati, che consentono agli utenti di ottenere una sintesi mirata di un testo, facilitando la comprensione e l'assimilazione dei concetti. È inoltre possibile utilizzare EduRag per tradurre documenti, rendendo accessibili contenuti in diverse lingue. Un'altra funzione molto utile è la creazione di domande aperte basate sui *database* indicizzati, con la possibilità di fornire suggerimenti per la loro correzione, il che può essere particolarmente utile per attività di autovalutazione e approfondimento.

Uno dei vantaggi di EduRag è la possibilità di scaricare i risultati delle interazioni in diversi formati. A seconda della conversazione o delle richieste effettuate, gli utenti possono scaricare un *file* di testo in formato TXT o DOC, oppure un file audio in formato MP3, consentendo di archiviare e rivedere facilmente le informazioni.

Un aspetto centrale è la trasparenza e l'affidabilità delle informazioni fornite da EduRag. Ogni risposta generata dal sistema è sempre collegata alle fonti da cui proviene l'informazione. EduRag riporta esattamente i *chunk* da cui ha estratto i dati, garantendo così la possibilità di verificare e tracciare l'origine delle risposte. Questo aspetto è essenziale per mantenere la qualità e l'accuratezza del contenuto fornito.

Infine, le catene di *prompt* utilizzate da EduRag sono costruite seguendo tecniche pedagogiche consolidate. Ciò significa che l'interazione con il sistema non è solo tecnologicamente avanzata, ma anche strutturata in modo da supportare l'apprendimento attivo e consapevole, rispettando i principi educativi moderni. Questo approccio garantisce che le interazioni con EduRag siano non solo funzionali, ma anche orientate al miglioramento dell'esperienza didattica degli studenti.

## 5 Conclusioni e riflessioni su sviluppi futuri

Ad oggi, EduRag permette un'interazione esclusivamente tramite tastiera, offrendo agli utenti un'interfaccia testuale per inserire domande e ottenere risposte. Oltre all'*output* testuale, fornisce anche una risposta audio. Tuttavia, è facile immaginare che il futuro di sistemi come EduRag si espanderà oltre il semplice utilizzo della tastiera, aprendo la strada a forme di interazione più avanzate.

Con lo sviluppo accelerato delle tecnologie IA, la prossima frontiera sarà la multimodalità, permettendo agli utenti di interagire non solo attraverso il testo, ma anche tramite voce, immagini, gesti e scrittura manuale. Questo futuro multimodale potrebbe includere, ad esempio, la capacità di discutere oralmente con l'IA su un argomento, ricevendo risposte verbali o visive, oppure utilizzare i gesti per indirizzare il sistema verso un determinato compito o problema.

Recentemente, modelli avanzati come Gemini 2, ChatGPT e Claude hanno introdotto funzionalità che permettono di accedere in tempo reale allo schermo degli utenti e alla loro webcam, favorendo un'interazione immediata e contestuale. Queste tecnologie consentono agli utenti di collaborare su documenti o attività visibili sullo schermo.

Un altro esempio di multimodalità è rappresentato dalle penne digitali che scrivono su carta e trasmettono l'input direttamente al dispositivo. Combinando il comfort della scrittura tradizionale con le potenzialità del digitale, questi strumenti risultano particolarmente utili in ambito educativo. Consentono agli studenti di prendere appunti, risolvere problemi o organizzare idee su supporti fisici trasferendo i contenuti in formato digitale. Attività come il *brainstorming* o la creazione di mappe concettuali, spesso svolte su carta per la sua semplicità, possono essere facilmente digitalizzate e integrate nei sistemi di IA.

La scrittura manuale digitalizzata potrebbe diventare una delle modalità preferite di interazione. Già oggi, alcuni modelli - come GPT-4o - sono in grado di riconoscere e leggere testi scritti a mano. In un video di OpenAI<sup>3</sup>, viene mostrato uno studente che interagisce con l'IA sia tramite la voce sia scrivendo con una penna su un tablet, disegnando un problema di geometria. L'IA, accedendo allo schermo del tablet, guida lo studente nella risoluzione del problema.

Questo scenario apre nuove prospettive pedagogiche. L'integrazione della scrittura manuale come strumento di interazione con l'IA potrebbe rilanciare tecniche tradizionali di apprendimento, rendendole più accessibili nel contesto educativo moderno. Attraverso l'IA, gli studenti potrebbero esercitarsi nella calligrafia o nella scrittura di saggi utilizzando testi scritti a mano, ricevendo un *feedback* immediato. Questo approccio, che unisce passato e futuro, potrebbe rafforzare la creatività e il pensiero critico, rendendo l'esperienza educativa più ricca e diversificata.

<sup>3</sup> [https://www.youtube.com/watch?v=\\_nSmkyDNulk](https://www.youtube.com/watch?v=_nSmkyDNulk).

Guardando al futuro, è chiaro che l'integrazione della multimodalità e della scrittura manuale non sarà solo una questione di strumenti, ma un'opportunità per ripensare il ruolo della tecnologia nell'educazione. Attraverso un equilibrio tra innovazione e tradizione, possiamo immaginare un sistema educativo più ricco, inclusivo e in grado di affrontare le sfide del XXI secolo.

### Riferimenti bibliografici

- Beard G.M. (1888). *Il nervosismo americano. Le sue cause e le sue conseguenze*. Città di Castello: S. Lapi tipografo editore (ed. or. 1881).
- Carr N. (2011). *Internet ci rende stupidi? Come la rete sta cambiando il nostro cervello*. Milano: Raffaello Cortina (ed. or. 2010).
- Chen Y., Fu Q., Yuan Y., Wen Z., Fan G., Liu D., Xiao Y. (2023). Hallucination Detection: Robustly Discerning Reliable Answers in Large Language Models. *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 245–255. <https://doi.org/10.1145/3583780.3614905>.
- Giliberto J. (2013, settembre 17). Storia. Quando la paura del treno scatena rivolte e sabotaggi. È il 1830. Recuperato 18 gennaio 2024, da *Correnti* website: <https://jacopogiliberto.blog.ilsole24ore.com/2013/09/17/tav-quando-la-paura-del-treno-scatena-rivolte-sabotaggi/>.
- Gori M. (2003). Introduzione alle reti neurali artificiali. *Mondo digitale*, 2(8), 4–20.
- Howard R. (Regista) (2002). *A Beautiful Mind* [Biography, Drama, Mystery]. Universal Pictures, Dreamworks Pictures, Imagine Entertainment.
- Johnson J., Douze M., Jégou H. (2017, febbraio 28). *Billion-scale similarity search with GPUs*. arXiv. <https://doi.org/10.48550/arXiv.1702.08734>.
- LangChain (2024, settembre 15). Recuperato 15 settembre 2024, da <https://www.langchain.com/>.
- Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Kiela D. (2021). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. arXiv. <https://doi.org/10.48550/arXiv.2005.11401>.
- McCarthy J., Minsky M. L., Rochester N., Shannon C. E. (1955). A proposal for the dartmouth summer research project on artificial intelligence. August 31, 1955. *AI magazine*, 27(4), 12–12.
- Mikolov T., Chen K., Corrado G., Dean J. (2013). *Efficient Estimation of Word Representations in Vector Space*. arXiv. <https://doi.org/10.48550/arXiv.1301.3781>.
- Naveed H., Khan A. U., Qiu S., Saqib M., Anwar S., Usman M., Mian A. (2023). *A Comprehensive Overview of Large Language Models*. <https://doi.org/10.48550/ARXIV.2307.06435>.
- Neelakantan A., Xu T., Puri R., Radford A., Han J. M., Tworek J., Weng L. (2022). *Text and Code Embeddings by Contrastive Pre-Training*. arXiv. <https://doi.org/10.48550/arXiv.2201.10005>.

- Numerico T. (2021). *Big data e algoritmi: Prospettive critiche*. Roma: Carocci.
- Parlamento Europeo e Consiglio dell'Unione Europea (2024). *Regolamento (UE) 2024/1689 relativo all'intelligenza artificiale (AI Act)*. Recuperato da <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>.
- Ramshaw L., Marcus M. (1995). Text chunking using transformation-based learning. *Third workshop on very large corpora*. Recuperato da <https://aclanthology.org/W95-0107>.
- Schulhoff S., Ilie M., Balepur N., Kahadze K., Liu A., Si C., Resnik P. (2024). *The Prompt Report: A Systematic Survey of Prompting Techniques*. <https://doi.org/10.48550/ARXIV.2406.06608>.
- Skinner B. F. (1961). Teaching Machines. *Scientific American*, 205(5), 90–106.
- Spitzer M. (2013). *Demenza digitale: Come la nuova tecnologia ci rende stupidi*. Milano: Garzanti (ed. or. 2012).
- Streamlit (2024, settembre 15). Recuperato 15 settembre 2024, da <https://streamlit.io/>.
- Toosi A., Bottino A. G., Saboury B., Siegel E., Rahmim A. (2021). A brief history of AI: How to prevent another winter (a critical review). *PET clinics*, 16(4), 449–469.
- Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Polosukhin I. (2017). *Attention Is All You Need*. arXiv. <https://doi.org/10.48550/arXiv.1706.03762>.
- Vespignani A., Rijntano R. (2019). *L'algoritmo e l'oracolo: Come la scienza predice il futuro e ci aiuta a cambiarlo* (p. 197). Milano: Il Saggiatore.
- Vygotskij L. S. (1934). *Pensiero e linguaggio* (L. Mecacci, a c. di). Roma-Bari: Laterza e figli (ed. italiana 1990).
- Wei J., Wang X., Schuurmans D., Bosma M., Xia F., Chi E., Zhou D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824–24837.
- Yao S., Yu D., Zhao J., Shafran I., Griffiths T., Cao Y., Narasimhan K. (2024). Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, 36. Recuperato da [https://proceedings.neurips.cc/paper\\_files/paper/2023/hash/271db9922b8d1f4dd7aaef84ed5ac703-Abstract-Conference.html](https://proceedings.neurips.cc/paper_files/paper/2023/hash/271db9922b8d1f4dd7aaef84ed5ac703-Abstract-Conference.html).

