

La ricerca

Dicembre 2023 Anno 11 Nuova Serie - 6 Euro www.laricerca.loescher.it

N°25

Uomini e bot Intelligenti, utili, dannose? Il punto sulle IA



SAPERI
Caratteristiche, limiti,
potenzialità delle
intelligenze generative

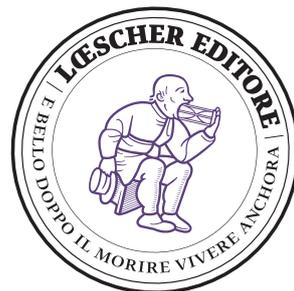
DOSSIER
Originalità
e plagio

SCUOLA
Prospettive didattiche
e strategie per un uso
consapevole



I QUADERNI

Quaderni della Ricerca: proposte metodologiche e aggiornamento didattico.



I libri pubblicati nella collana sono reperibili in libreria o presso le agenzie di zona. Indice e prime pagine sono disponibili sul sito de «La ricerca».



I Quaderni della Ricerca sono anche online
www.laricerca.loescher.it/quaderni

Martellate di ottimismo

Ci sono analogie potenti che apparentemente aiutano a comprendere e che invece rischiano di confondere. Come quando diciamo «buco nero» e il pensiero, pur istruito, va dove spinge la metafora: all'immagine di un buco, di un varco, di un vuoto e non, come dovrebbe, a quella di un pieno, massicciamente denso e attrattivo.

Per me con l'«intelligenza artificiale» è lo stesso. Se è intelligente, penso a un cervello umano, e a quello che sa fare un cervello umano. E siccome è artificiale, penso a un cervello umano creato dall'uomo «a sua immagine e somiglianza». Le stesse pulsioni dell'uomo, le stesse intenzioni. Tra le altre, quella supremamente umana del controllo e del dominio. All'uomo nella mia mente spetta il ruolo ora di divinità onnipotente, ora di apprendista stregone.

Suggestioni fantasy? Tutt'altro, dal momento che gli stessi creatori dei più diffusi software di IA si ritirano in disparte a stilare il bugiardino, elencando tra le controindicazioni la possibile estinzione del genere umano!

Sono cose che non lasciano indifferenti e di fronte alle quali crediamo ci spetti il modesto compito di tentare un chiarimento, entro i limiti angusti e non specialistici di questa pubblicazione.

Cominciamo quindi dal posizionamento: apocalittici o integrati? Né gli uni né gli altri, o forse entrambi. Attenzione critica e consapevole, piuttosto, alle implicazioni del balzo evolutivo che potrebbe favorire questa nuova tecnologia. In positivo e in negativo. Personalmente confuso, leggendo i contributi mi si chiariscono le idee mentre si sfumano le convinzioni. Per esempio, tornando all'ingannevole metafora iniziale: l'intelligenza artificiale non è intelligente. Su questo sembrano concordare tutti. Il software segue strade sue, probabilistiche e non semantiche. Imita e non capisce, o meglio, imita senza sapere.

È l'intelligenza del tostapane, dice il filosofo: l'ordigno scalda delle resistenze, ma non sa di scaldarle e, soprattutto, non può immaginare a cosa serve farlo... E qui mi accorgo che mi manca una definizione certa di intelligenza: è intelligente risolvere problemi o sapere di saperlo fare? Intelligenza o coscienza? O autocoscienza? Ad aumentare la mia confusione, l'ipotesi che sia solo questione di tempo e di potenza di calcolo perché l'IA, non intelligente adesso, possa diventarla in un futuro non lontano...

E poi...L'IA non è nemmeno creativa, limitandosi a ripetere lo stereotipato immaginario dei suoi creatori. Vero, ma forse no. O meglio: vero a meno di non rivedere il concetto stesso di creatività, fortemente condizionato dal contesto (anche tecnologico) in cui viviamo (e qui mi soccorre l'immagine suggerita da Capodivacca, quella del martello che spinge a cercare cose da martellare).

Anche in questo caso mi rendo conto di non avere una mia formula definitoria della creatività umana, così come di quasi ogni altro concetto tirato in ballo. E allora oscillo senza soluzione tra la paura e la speranza che ogni prospettiva sembra aprire: e se da un lato l'IA mi appare pericolosa (imprevedibile); disfunzionale ed economicamente dannosa (incentiva alla pigrizia e sottrae lavoro); antidemocratica e manipolatoria (può controllare gli esseri umani e costringerli in una bolla di menzogne)... dall'altro mi accorgo dell'enorme potenzialità che esprime attraverso il racconto dell'uso intelligente, creativo, inclusivo, funzionale e democratico che di essa si fa - già ora - in molte realtà scolastiche del nostro territorio.

Ripenso al martello di cui si è detto: arma terribile in potenza, strumento operoso e benefico in atto. E mi scopro ottimista.

Sandro Invidia, direttore editoriale di Loescher.

“Apocalittici o integrati? Né gli uni né gli altri, o forse entrambi. ”

La ricerca

Periodico semestrale

Anno 11, Numero 25 Nuova Serie, dicembre 2023

autorizzazione n. 23 del Tribunale di Torino, 05/04/2012

iscrizione al ROC n. 1480

Editore

Loescher Editore

Direttore responsabile

Mauro Reali

Direttore editoriale

Ubaldo Nicola

Coordinamento editoriale

Alessandra Nesti - PhP

Impaginazione

Ubaldo Nicola

Copertina

Emanuela Mazzucchetti, Davide Cucini

Pubblicità interna e di copertina

VisualGrafika - Torino

Stampa

Vincenzo Bona S.p.A.

Strada Settimo, 370/30 - 10156 Torino (TO)

Distribuzione

Per informazioni scrivere a:

laricerca@loescher.it

Redazione

Beatrice Bosso, Simone Giusti, Sandro Invidia, Francesca Nicola,

Ubaldo Nicola, Mauro Reali.

Hanno scritto su questo numero

Silvia Capodivacca, Stefania Cardinali, Simon Coghlan,

Francesco D'Isa, Erkan Er, Marco Guastavigna,

Mohammad Khalil, Maurizio Maglioni, Elvis Mazzoni,

Carlo Milani, Tim Miller, Ubaldo Nicola, Jo Ann Oravec,

Eleonora Pantò, Jeannie Paterson, Grazia Pozzi,

Elena Rausa, Mauro Reali, Enzo Ruffaldi, Agnese Trocchi.

© Loescher Editore

via Vittorio Amedeo II, 18 - 10121 Torino

<https://laricerca.loescher.it/>

ISSN: 2282-2836 (cartaceo)

ISSN: 2282-2852 (online)



Sommario

Caratteristiche, limiti e potenzialità delle intelligenze generative

SAPERI

- 6 In principio era il bot**
Eleonora Pantò
- 13 Paralipomeni di una
formazione sull'IA**
Marco Guastavigna
- 19 La banalità digitale**
Silvia Capodivacca
- 23 IA e stile**
Francesco D'Isa
- 26 La medicina intelligente**
Grazia Pozzi
- 31 Buzzati e l'intelligenza artificiale**
Elena Rausa
- 35 ChatGPT tra Latino e Latinorum**
Mauro Reali

DOSSIER

Originalità e plagio

- 40 Il destino dell'originalità**
Ubaldo Nicola
- 43 La matematica del plagio**
Mohammad Khalil e Erkan Er
- 46 L'impatto dell'IA sulla frode accademica**
Jo Ann Oravec
- 51 Il proctoring: una risorsa o il Grande Fratello?**
Simon Coghlan, Tim Miller, Jeannie Paterson

SCUOLA

- 54 Hacking IA**
Agnese Trocchi e Carlo Milani
- 60 L'intelligenza artificiale
nella didattica della filosofia**
Enzo Ruffaldi
- 65 Una scuola... diversamente digitale**
Elvis Mazzoni e Stefania Cardinali
- 71 Due strumenti per l'inclusione
e per la didattica laboratoriale:
Bard e ChatGPT**
Maurizio Maglioni



In principio era il bot

Tecnologie a servizio di chi?

di Eleonora Pantò

6

SAPERI / In principio era il bot

«**I** bot sono le nuove app», dichiarava Satya Nadella, CEO di Microsoft, alla Conferenza per sviluppatori nel 2016, in occasione del lancio di Cortana. La sua visione del futuro dell'interazione fra gli umani e le macchine era «conversazione come piattaforma», in cui l'Intelligenza Artificiale (IA) avrebbe consentito ai computer di interagire con le persone, utilizzando l'interfaccia umana più naturale, il linguaggio, per cui era necessario «infondere nei computer informazioni su di noi e sul nostro contesto».

I bot sono i programmi che lavorano dietro le quinte per migliorare l'esperienza degli utenti sul web: forniscono suggerimenti di ricerca (se hai acquistato questo, ti piacerà anche quest'altro), previsioni meteo, filtrano le email, ripuliscono le voci di Wikipedia. I bot sono stati adattati per agire sui social media: sono masse di utenti fittizi, i cosiddetti *social bot*¹, pilotati come squadriglie (*bot squad*) d'assalto contro personaggi scomodi o per influenzare l'opinione pubblica in momenti critici come le elezioni politiche, come documentato da indagini giornalistiche e studi scientifici².

I chatbot, programmi specializzati nel simulare la conversazione umana, sono molto utilizzati come assistenti alla clientela. È significativo che il rapporto della Harvard Business Review³ rivolto a misurare quanto le imprese siano pronte ad affrontare la sfida della trasformazione digitale, fortemente accelerata dalla pandemia, insista



Foto Wikicommons.





↑
Arthur Ganson,
Sedia per pensare,
Thinking Chair,
2009, Art
Electronica
Museum of
Future, Linz,
Austria (foto
Wikicommons).

sulla necessità di investire in chatbot, che siano “empatici”: le aziende dovranno essere «in grado di collegare i canali e di fornire interazioni simili a quelle umane e attivabili, che portino a risultati aziendali positivi. In breve, dovranno investire in tecnologie che aumentino e scalino le competenze trasversali che i loro migliori dipendenti apportano ai rapporti con i clienti: la capacità non solo di ascoltare il cliente, ma di sentire effettivamente ciò che dice, di entrare in empatia con lui e di intraprendere azioni che soddisfino le sue aspettative».

«Alessa ti voglio bene», dice il bambino di 4 anni abituato a interagire con la voce - femminile - che è stata progettata per incentivare acquisti su Amazon. Alexa è un bot che, come tutti gli assistenti vocali, registra tutte le interazioni della famiglia ed è in grado di riconoscere le impronte vocali di tutti i membri, i gusti musicali e gli interessi. A maggio 2023 Amazon ha accettato di pagare 25 milioni di dollari⁴ per chiudere la causa federale negli Stati Uniti per aver conservato indefinitamente le registrazioni vocali dei bambini, violando la legge federale sulla privacy dei bambini online.

ChatGPT ecc.

—
ChatGPT è un chatbot, progettato per interagire con gli utenti usando il linguaggio naturale. Si basa su tecniche di apprendimento automatico per analizzare basi dati di grandi dimensioni e produrre risposte plausibili alle richieste - i prompt - degli utenti.

Come funziona

—
Senza addentrarsi in spiegazioni tecnologiche, il funzionamento del LLM (Large Language Model o modello linguistico di grandi dimensioni) è stato paragonato a un «pappagallo stocastico»⁵, pappagallo perché nelle risposte ripete cose già sentite e stocastico perché anticipa un fenomeno in termini di statistica e probabilità: di fatto il sistema analizza porzioni di testo a cui associa parole che hanno una frequenza statistica maggiore e una più alta probabilità. Un LLM non comprende il senso della domanda né la risposta che fornisce. Grazie a un’infinità di parametri il sistema produce frasi di senso compiuto e di contenuto plausibile, analizzando il contesto in cui sono fornite.

Gli aspetti più interessanti introdotti da ChatGPT3 rispetto al precedente ChatGPT2 sono la possibilità di interagire attraverso comandi con parametri, che permettono di definire meglio il tipo di output richiesto e la possibilità di successive iterazioni sull’output ottenuto senza dover ricominciare da capo, in un processo di affinamento che utilizza i feedback umani (*Rafforzamento dell’Apprendimento grazie ai Feedback Umani* o RLHF).

L’approccio probabilistico adottato dagli LLM rispetto a uno basato su regole non è necessariamente uno svantaggio, secondo alcuni, perché li rende strumenti versatili e flessibili: l’idea è di usare le funzionalità del motore LLM come un processore che riceve input e produce output,

che non siano necessariamente testi di una conversazione⁶.

I nuovi LLM si basano su un modello chiamato Transformer⁷, progettato da Google nel 2017. Questo modello analizza i *token* (parole o porzioni di parole) all'interno di un testo e poi le relazioni fra i vari *token*, per determinare qual è quello più importante. Quest'analisi si svolge in unico passaggio e non attraverso passi sequenziali come nelle reti neurali: ciò rende il sistema molto più efficiente, permettendogli di identificare il significato all'interno di uno specifico contesto e di produrre variazioni o sintesi plausibili. Tale modalità di analizzare "pattern", cioè strutture ricorrenti, è utilizzata anche per generare immagini e suoni da sistemi come Dall-E, Midjourney e Stable Diffusion, e codice del computer utilizzando generatori come GitHub CoPilot.

L'impatto sociale

— ChatGPT3.5, la prima applicazione gratuita di LLM, è stata rilasciata pubblicamente e gratuitamente in un formato sperimentale, alla fine del 2022, come dichiarato da OpenAI: «ChatGPT è un modello gemello di InstructGPT, che è addestrato a seguire un'istruzione in un prompt e a fornire una risposta dettagliata. Siamo entusiasti di presentare ChatGPT per ottenere il feedback degli utenti e conoscere i suoi punti di forza e di debolezza. Durante l'anteprima della ricerca, l'uso di ChatGPT è gratuito. Provatelo subito su chat.openai.com.» coinvolgendo milioni di persone nel ruolo di tester volontari e non pagati.

LLM come ChatGPT sono molto efficaci nella produzione massiva di testi (intesi in senso lato includendo codice sorgente e multimedia) con costi risibili e una velocità incomparabile con quella umana: sono in grado di produrre testi in lingue e stili diversi, possono conversare simulando di essere uno scrittore inglese del Cinquecento, un insegnante che deve preparare una lezione o addirittura un santo⁸.

È estremamente semplice porre domande e avere risposte, ma per avere risultati di valore bisogna imparare a costruire domande circostanziate e molto dettagliate. Inoltre, il massiccio stress-test a cui questi sistemi sono stati sottoposti - si parla di 10 milioni di utenti al giorno da dicembre 2022 ad aprile 2023 - ha evidenziato falle di sicurezza nella gestione dei dati personali; ad esempio, il sistema riutilizzava parti di dialogo di altri utenti.

ChatGPT3.5 ha creato grandi aspettative e qualche immediata delusione: gli LLM non eccellono nella logica per la loro intrinseca natura, quindi le risposte che prevedevano qualche ragionamento a partire dai dati contenuti nella data erano palesemente errate. In poco tempo è stato evidente che le

risposte non sono sempre accurate e se il sistema non ha a disposizione informazioni sufficienti per rispondere alle domande, le inventa, dando luogo alle cosiddette *allucinazioni*. Tali limiti sono pubblicamente dichiarati in un linguaggio molto chiaro: «ChatGPT a volte scrive risposte plausibili ma non corrette o senza senso. Risolvere questo problema è impegnativo durante l'addestramento nella vita reale e non esiste alcuna fonte di verità[...]». Nonostante queste avvertenze, si sono verificate situazioni molto problematiche: casi eclatanti sono quelli dell'avvocato che ha citato sentenze inesistenti in un processo⁹ e del professore infamato da accuse di inesistenti scandali sessuali¹⁰.

Gli LLM sono progettati per simulare dialoghi reali ma «Non hanno empatia, né alcuna comprensione del linguaggio che stanno producendo, né alcuna comprensione della situazione in cui si trovano. Ma il testo che producono sembra plausibile e quindi è probabile che le persone gli attribuiscono un significato. Lanciare qualcosa del genere in situazioni delicate significa correre rischi sconosciuti. Una domanda chiave da porsi è: chi è responsabile se l'IA fornisce suggerimenti dannosi? In questo contesto, l'azienda che sta implementando l'esperimento sta affidando tutta la responsabilità ai membri della comunità che scelgono il sistema di intelligenza artificiale?»¹¹.

Al momento del suo debutto, ChatGPT non prevedeva alcun controllo per l'accesso di minori né informative sul trattamento dei dati: per questo il Garante per la Protezione dei Dati ha bloccato OpenAI in Italia, provocando il disappunto di milioni di utenti, che tuttavia facilmente aggiravano il blocco utilizzando sistemi per mascherare la provenienza (tramite VPN) o applicazioni create per l'occasione. L'accesso è stato ripristinato dopo qualche settimana a seguito dell'adeguamento da parte di OpenAI. Nonostante il provvedimento sia stato un'operazione dovuta, le reazioni sono state quasi unanimi nell'esprimere insofferenza verso il GDPR, vissuto da molti come un'inutile burocratica complicazione, in grado di bloccare lo sviluppo tecnologico dell'Italia.

La fine del genere umano?

— ChatGPT ha continuato a catalizzare l'attenzione per tutto il primo semestre del 2023, e mentre ci si interrogava sull'impatto che questi strumenti avrebbero avuto sull'occupazione e l'informazione, sull'istruzione e sulla democrazia, 350 esperti (diventati 600 a settembre) sottoscrivevano la seguente dichiarazione «Mitigare il rischio di estinzione da parte dell'I.A. dovrebbe essere una priorità globale insieme ad altri rischi su scala sociale come le pandemie e la guerra nucleare».

Una dichiarazione molto potente e in grado di creare una grande ansia dovuta al presunto

prossimo avvento di una IA Generalista (IAG), che pure essendo rischiosa promette di essere vantaggiosa per il genere umano. Se vi ricorda qualcosa, Sam Altman ha citato la famosa frase «La tecnologia sta accadendo perché è possibile», e ha paragonato la sua azienda al progetto Manhattan.

E. Morozov, in un articolo pubblicato sul New York Times a giugno 2023, ha evidenziato come i rischi paventati dagli esperti non siano futuri ma già attuali, a partire dal fatto che «presentare la sua comparsa come quasi inevitabile distrae dalla ricerca di modi migliori per aumentare l'intelligenza». Morozov usa gli esempi di varie startup che non sono riuscite a mantenere le loro promesse, come Uber, che doveva rivoluzionare il trasporto pubblico con veicoli a guida autonoma con costi minimi di personale, o Theranos, la cui fondatrice è in galera per frode: i suoi dispositivi innovativi per l'analisi del sangue non sono mai esistiti. Airbnb, TaskRabbit e Facebook sono altre startup che promettevano di risolvere problemi e migliorare la vita dei loro utenti, ma che hanno creato anche nuove problematiche: favorendo la gentrificazione delle città, la diffusione della precarietà e di notizie false. Anche ChatGPT promette di risolvere tanti problemi, in settori come la salute e l'istruzione, infatti potrà fornire consigli a chi non può permettersi un medico (post Twitter di Sam Altman del 19 febbraio 2023); ancora, «L'intelligenza artificiale sarà un buon tutor come qualsiasi essere umano, ma i profitti nell'istruzione richiederanno tempo», secondo Bill Gates.

↓
Concerto di Alex Braga alla cerimonia di apertura dell'Art Electronica Museum of Future, 2019, Linz, Austria (foto Wikicommons).

La scuola distopica del futuro

Le dichiarazioni di Bill Gates alimentano una delle distopie più diffuse sul futuro, forse iniziate nel 1951 dal racconto di Asimov *Come si divertivano*¹² che descrive una scuola in cui gli insegnanti sono sostituiti dai robot. La diffusione dei computer personali ha ulteriormente alimentato questa distopia: la BBC nel 1984, in un programma per ragazzi, presentava un PC IBM che rispondeva ai comandi vocali. Qualche anno dopo, nel 2010, il professor Sugata Mitra, attraverso il suo esperimento con gli studenti in India chiamato *Hole in the wall* (il buco nel muro), teorizzava il potenziale dei computer per sostenere l'apprendimento senza l'intervento degli adulti: un progetto che gli valse il TED Prize di 1 milione di dollari ma anche qualche dubbio sulla sua efficacia. Con la diffusione di smartphone, ambienti virtuali e IA, complici la pandemia e la didattica a distanza, questo futuro distopico sembra sempre più prossimo.

Esistono numerosi e reiterati progetti promossi da giganti del software che vogliono re-inventare la scuola, soprattutto negli Stati Uniti: il loro approccio è diverso rispetto a quello dei grandi capitani di industria che agivano sulle istituzioni. Nel tempo le soluzioni tecnologiche sono proposte direttamente alle famiglie e agli insegnanti, ad esempio attraverso i social network. Netflix e Facebook hanno creato e distribuito software per l'apprendimento personalizzato e adattivo, che tuttavia non hanno dato risultati in termini di



miglioramento dell'apprendimento, e che secondo alcune ricerche potrebbero attribuire falsamente al software risultati migliori dovuti a insegnanti migliori.

Un esempio recentissimo di scuola del futuro potenziata dall'IA è Khanmigo¹³, l'evoluzione della Khan Academy¹⁴. Il servizio è accessibile da maggio 2023, è a pagamento (chiamate donazioni sul sito) e si basa su ChatGPT4. È progettato come un tutor che adotta il metodo socratico per aiutare gli studenti ad imparare, quindi non fornisce risposte. Khan sostiene che Khanmigo sia la soluzione al problema Sigma2 di Benjamin Bloom¹⁵: la ricerca pubblicata nel 1984 che dimostrava che gli studenti con un insegnante dedicato aveva risultati migliori di uno studente che studiava in un contesto di classe tradizionale. Un problema di difficile soluzione: garantire un insegnante per ogni studente perché «troppo costoso per la maggior parte delle società da sostenere su larga scala» ma non per Khanmigo! L'assistente virtuale ha anche una versione per supportare gli insegnanti nella creazione di lezioni.

E gli studenti? Che ne pensano? Il New York Times¹⁶ ha chiesto agli studenti di commentare l'articolo che presenta la sperimentazione di Khanmigo in una scuola. Nei 300 e passa commenti pubblicati si scopre che per la maggior parte non vorrebbero una IA come tutor, anche se riconoscono possa essere utile in alcuni casi. Lamentano che l'IA dà troppe informazioni, a volte incorrette, e non è utile che faccia i compiti al loro posto; qualcuno invece ritiene che poter interrogare un esperto onnisciente in qualsiasi momento sia un grande vantaggio; un altro commento invece ritiene che l'IA a scuola dovrebbe essere vietata.

Il tutoraggio individuale è da sempre considerato il santo graal dalle aziende edtech: oggi ci sono decine di chatbot per l'educazione e altre applicazioni educative che usano l'IA e altrettanti siti che li raccolgono¹⁷.

Come ben descritto nella scheda di lavoro *in progress* pubblicata su La ricerca online¹⁸ le e gli insegnanti sono invitati a esplorare una pluralità di "assistenti artificiali" per valutarne l'uso in coerenza con gli obiettivi didattici. L'ultima delle raccomandazioni della scheda insiste sulla creazione di una policy d'istituto.

Le linee guida

Al di là del clamore sollevato da ChatGPT, da tempo i governi e istituzioni sovranazionali emettono linee guida per regolamentare le IA. Secondo l'Osservatorio sull'IA dell'OCSE¹⁹, a settembre 2023 sono più di 800 le linee guida emesse da 69 Paesi nel mondo. L'UNESCO ha effettuato una prima classificazione delle linee guida sull'IA applicata all'educazione nel 2021 e ha evidenziato aree di

comune interesse, quali la protezione dei dati e della privacy, l'apertura e la trasparenza delle tecnologie e dei dati, la risposta dei sistemi educativi alle richieste di un nuovo mercato e l'esigenza di finanziamenti per l'implementazione di programmi formativi. Da allora, il contesto è cambiato molto, come le attese e le aspettative verso le IA: le più recenti dichiarazioni UNESCO²⁰ invitano a un'attenta valutazione delle scuole nell'introduzione di tali strumenti, ricordando ad esempio le procedure e i tempi che regolano l'adozione di libri di testo. Anche le *Linee guida sulle Intelligenze artificiali generative*²¹ di recentissima pubblicazione suggeriscono diversi punti di attenzione per i politici, tra i quali il trattamento dati degli utenti, il controllo umano nei processi decisionali, la necessità di trasparenza e di formazione per insegnanti e studenti; inoltre consigliano un limite minimo di età fissato a 13 anni per l'accesso autonomo, sulla scia di regolamentazioni già esistenti²².

Secondo un'indagine UNESCO (non pubblicata) del marzo 2023, su 450 scuole e università nel mondo solo il 10% ha definito una politica interna sull'uso delle intelligenze generative: la mancanza di adeguate riflessioni sull'uso delle IA potrebbe portare a conseguenze indesiderate.

Conclusioni

—
Chiunque operi nel settore della conoscenza dovrebbe formarsi sul tema delle IA generative, non solo per quanto riguarda un addestramento tecnico a scrivere prompt efficaci (non tutti devono diventare *prompt engineer*). È necessario creare una consapevolezza che permetta di comprendere che il determinismo del mercato non sempre coincide con gli interessi della comunità, e che le politiche, i regolamenti e tutta la noiosa burocrazia che invita al rispetto della centralità umana, all'uso etico e trasparente dell'IA, servono a tutelarci e vanno compresi, rispettati e fatti rispettare.

NOTE

1. M. Bakardjieva, R. Gehl, *Socialbots and Their Friends: Digital Media and the Automation of Sociality*, Taylor & Francis, USA 2018.
2. Cfr. G. Caldarelli, R. De Nicola, F. Del Vigna et al., *The role of bot squads in the political propaganda*, in «Communications Physics» 3, 81, 2020, e F.L. De Faveri, L. Cosuti, P.P. Tricomi, M. Conti, *Twitter Bots Influence on the Russo-Ukrainian War During the 2022 Italian General Elections*, in B. Arief, A. Monreale, M. Sirivianos, S. Li (a cura di), *Security and Privacy in Social Networks and Big Data*, in «SocialSec 2023. Lecture Notes in Computer Science», vol. 14097, Springer, Singapore 2003.
3. Consultabile all'indirizzo <https://bit.ly/45muZe8>.
4. Cfr. <https://www.nytimes.com/2023/05/31/technology/amazon-25-million-childrens-privacy.html>.
5. Si veda E. Bender et al., *On the dangers of stochastic*



↑ Thom Kubli, *L'orizzonte del buco nero*, 2016, installazione al festival dell'Art Electronica Museum of Future, Linz, Austria (foto Wikicommons).

parrots: Can language models be too big, Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency; Association for Computing Machinery, New York, NY, USA 2021.

6. Cfr <https://bit.ly/3tfrNDS>.

7. Sul funzionamento degli LLM si veda <https://ig.ft.com/generative-ai/>.

8. Cfr. <https://www.wired.it/article/bot-chattare-santi-prega-org/>.

9. Cfr. https://www.corriere.it/esteri/23_maggio_28/avvocato-aula-new-york-chatgpt-che-si-sbaglia-cita-sentenze-fasulle-e2d8ad88-fd80-11ed-8aba-45a422ecf4be.shtml.

10. Cfr. qui <https://bit.ly/45olJWW> e qui <https://bit.ly/46GW394>.

11. Il commento di Bender è relativo all'uso di chatbot per la salute mentale e si legge qui <https://bit.ly/3tnrPK0>.

12. Si veda <https://bit.ly/3ZJD9w0>.

13. <https://www.khanacademy.org/khan-labs#khanmigo>

14. <https://www.khanacademy.org/>

15. https://en.wikipedia.org/wiki/Bloom%27s_2_sigma_problem

16. <https://www.nytimes.com/2023/09/11/learning/would-you-want-an-ai-tutor.html#commentsContainer>

17. Si veda a titolo di esempio <https://aieducator.tools/>

18. *Impostare percorsi di formazione di insegnanti sull'assistenza artificiale alla mediazione didattica*, La ricerca online, 12 settembre 2023, consultabile all'indirizzo [\[ne-di-insegnanti-sullassistenza-artificiale-alla-mediazione-didattica/\]\(https://laricerca.loesch.it/impostare-percorsi-di-formazio-\).](https://laricerca.loesch.it/impostare-percorsi-di-formazio-</p>
</div>
<div data-bbox=)

19. Qui: <https://oecd.ai/en/dashboards/overview>.

20. S. Giannini, *Generative AI and the future of education*, UNESCO 2023.

21. F. Miao, W. Holmes, *Guidance for Generative AI in Education and Research*, UNESCO 2023, consultabile all'indirizzo <https://unesdoc.unesco.org/ark:/48223/pf0000386693.locale=en>.

22. Ricordiamo che in Italia l'età minima per l'uso dei social (WhatsApp, Instagram, Tiktok) in modo autonomo è di 14 anni, fra i 13 e i 14 è necessaria l'autorizzazione dei genitori. Telefono Azzurro a febbraio 2023 ha proposto di elevarla a 16 come già previsto dalla regolamentazione europea.

Eleonora Pantò

È direttrice dell'associazione Dschola-Le scuole per le scuole, che dal 2004 sperimenta l'uso delle tecnologie digitali nella didattica. Collabora con la Fondazione Politecnico di Milano in progetti europei sull'educazione alla sostenibilità e l'Open Education. È tra le promotrici del network Open Education Italia. Ha lavorato come ricercatrice al CSP ed è stata presidente dell'Associazione Media & Learning. Scrive su diverse testate online e ha contribuito a varie pubblicazioni, tra cui *Consapevoli in rete*, *Digital Education*, *Internet per la Didattica*, *Gens electrica*.

Paralipomeni di una formazione sull'IA

Non solo ChatGPT: perché è importante conoscere gli assistenti artificiali ad attività cognitive mediante prestazioni che imitano quelle umane, e ragionare su un loro utilizzo professionale ragionevole e significativo.

di Marco Guastavigna

Il 13 luglio 2023 Google Bard è diventato disponibile anche in Italia¹. Presentato da molti come la risposta di Alphabet Inc. a ChatGPT, non ha però lo stesso seguito sensazionalistico. Qualcuno tenta qualche timido confronto, ma - sostanzialmente - il nuovo dispositivo viene accolto da un silenzio assordante, su cui voglio ragionare rapidamente, prima di prospettare ipotesi di percorsi formativi per insegnanti.

Nei mesi precedenti, il panorama della cosiddetta Intelligenza Artificiale era stato infatti ridotto dal dibattito su media e social al già citato ChatGPT², in una discussione subito polarizzata: affatto contrari (in nome del dilagare dell'ignoranza e della pigrizia) contro affatto favorevoli (in nome di inevitabilità e inesorabilità del progresso). Poiché il pubblico degli entusiasti e degli interessati alle nuove opportunità operative era (ed è) rilevante, sono fioriti numerosi corsi, webinar e altri incontri esemplificativi di come usarle "praticamente" nella didattica, tutti imperniati solo su ChatGPT. Nessuno spazio per il nuovo arrivato, che è così finito in ombra. In generale, discorso asfittico, con un campo di riflessione limitato e viziato da empirismo banalizzante.

Estrattivismo culturale e cognitivo all'opera

Per chi insegna, accedere a Google Bard è invece doveroso e illuminante. Non solo per scoprire che sono tassativamente richiesti i 18 anni³ e che quindi un impiego diretto da parte degli studenti è improponibile⁴. Una lettura delle sue impostazioni sulla privacy, infatti, ci fa capire non solo

che Alphabet Inc. vuole rispettare il *General Data Protection Regulation* europeo, ma anche che tra dispositivo e utenti si realizza il medesimo ciclo alla base dei profitti di Google Search sul mercato della conoscenza, lo scambio tra fornitura di un servizio monitorato in tempo reale e utilizzo dei dati degli utenti, in questo caso le conversazioni. Intervengono infatti revisori umani, organizzati in squadre addestrate, che le possono leggere, annotare, elaborare, salvare "a fini qualitativi", come contributo di adattamento e miglioramento del sistema. Si precisa anche che: «Man mano che Bard si sviluppa, potremmo chiedere il tuo consenso per elaborare le tue informazioni per scopi specifici». Lo stesso processo, peraltro, è sotteso a ChatGPT, già dotato di una versione PLUS a pagamento, 20 dollari USA mensili.



Dimensione civile ed etica della formazione

Quanto visto finora ha a mio giudizio un'implicazione fondamentale: per essere davvero efficace e utile, la formazione degli insegnanti sui dispositivi di IA generativa non può illudersi e illudere di essere in campo neutro, perché il contesto non è tale, sia dal punto di vista civile, sia da quello etico.

L'IA generativa si caratterizza per la rapidissima elaborazione di risposte artificiali plausibili a domande o indicazioni umane, resa possibile dal riferimento a modelli e *BigCorpora*, costituiti raccogliendo, correlando, analizzando e classificando quantità oceaniche di dati. Questo richiede allenamenti di mega-macchine, ossia proprietà di infrastrutture molto costose e capacità tecnica di impossessamento di *BigData*, mediante una gigantesca potenza di calcolo. Insomma, poche aziende private captano in tempo reale l'intelligenza collettiva condivisa via internet, presentando l'operazione come raggiungimento di "conoscenza complessiva" con la scansione di materiali liberamente disponibili in rete. Sorvolando però sulla necessità di logistiche ineguagliabili. È pertanto bene chiarire, in premessa, che i dispositivi di IA riaffermano un dominio operativo e culturale oligopolistico già conseguito con "motori di ricerca" e altri servizi individuali e collettivi. Ai fini della piena consapevolezza, questi aspetti sono perfino più importanti dei limiti dei "Large Language Model"⁵ che alimentano i dispositivi a conversazione simulata, peraltro ben esplicitati da Google Bard:

- **Accuratezza:** le risposte di Bard potrebbero essere imprecise, specialmente quando vengono poste domande su argomenti complessi o fattuali.
- **Bias:** le risposte di Bard potrebbero riflettere pregiudizi o prospettive presenti nei suoi dati di addestramento.
- **Persona:** le risposte di Bard potrebbero suggerire che abbia opinioni o sentimenti personali.
- **Falsi positivi e falsi negativi:** Bard potrebbe non rispondere ad alcuni suggerimenti appropriati e fornire risposte inappropriate ad altri.
- **Vulnerabilità al suggerimento dell'avversario:** gli utenti troveranno modi per stressare ulteriormente Bard.

Definizioni e contesto

Sono ora in grado di presentare in modo più preciso la mia ipotesi di formazione, cioè l'*uso professionale, critico e consapevole di vari dispositivi di comunicazione artificiale con capacità generative*. Questo approccio si sviluppa in un contesto con un'identità e una storia ben definite: il laboratorio "Nuove tecnologie per l'apprendimento" (75 ore di frequenza), che da 10 anni conduco nell'ambito del percorso di specializzazione sul sostegno per l'Università

di Torino. Esso coinvolge insegnanti della scuola secondaria di primo e secondo grado, a cui è chiesto di realizzare un "elaborato multimediale" con valenza inclusiva. Ho da tempo definito come obiettivo generale un'autentica emancipazione culturale: l'acquisizione di un lessico analitico, critico e professionalmente davvero incisivo per progettazione, elaborazione e verifica di materiali utili alla mediazione didattica con approccio *Universal Design for Learning*, capace di superare la gergalità deferente e la superficialità ancillare delle seduttive concettualizzazioni correnti⁶. L'elaborato multimediale è infatti occasione di dialogo con il tutor e tra i corsisti a proposito di *autorialità digitale sostenibile*, in termini di tempi di elaborazione, ottimizzazione, semplicità⁷ e replicabilità dei materiali prodotti, rispetto del diritto d'autore, interoperabilità tra sistemi operativi e device, consapevolezza e controllo sui dati scambiati con i dispositivi coinvolti.

Distinguiamo poi tra autorialità *diretta* e *secondo livello*. La prima è produzione di materiali originali, la seconda selezione, organizzazione, adattamento e cura dell'interattività di materiali presenti in rete. In questo scenario, nell'anno accademico concluso hanno già fatto irruzione vari dispositivi generativi con potenzialità utili per dialogare sull'elaborato multimediale, accolti come "agenti artificiali", al plurale. Anche su questa base, ho sviluppato una riflessione sistematica e posso definire in modo più chiaro questo segmento della formazione sull'uso dei dispositivi digitali: *utilizzo professionale ragionevole e significativo di assistenti artificiali ad attività cognitive mediante prestazioni che imitano quelle umane*.

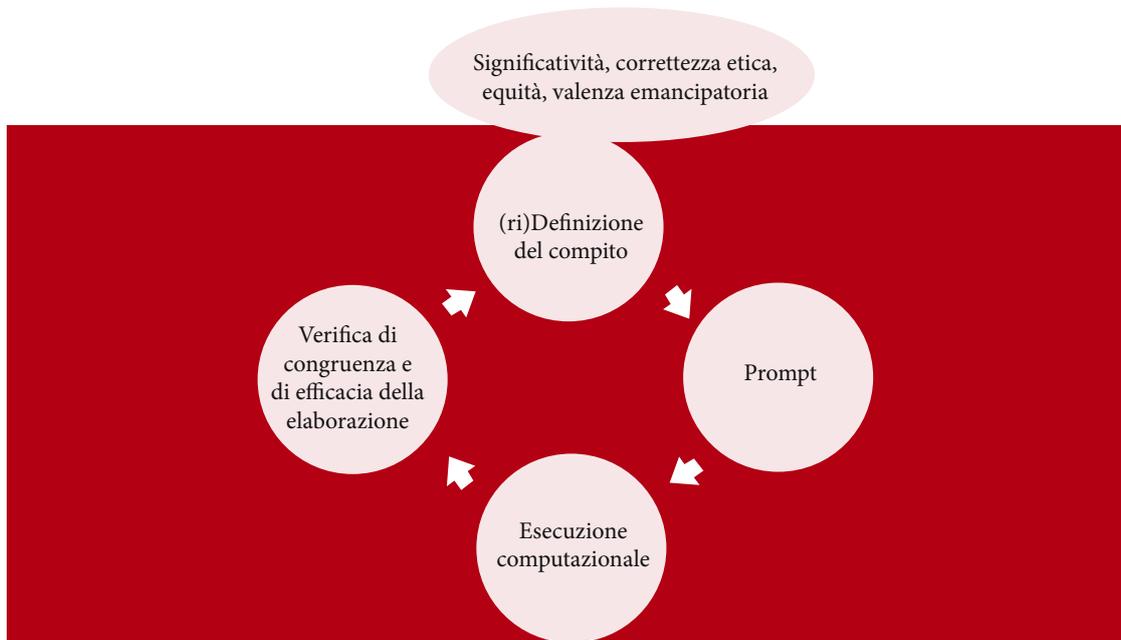
Aspetti epistemologici della ragionevolezza

A indirizzarmi verso la formulazione appena fornita sono state anche alcune letture di approfondimento, quanto mai necessarie per andare oltre sensazionalismo e polarizzazione e per radicare sul piano epistemologico l'impianto della formazione.

Elena Esposito ci fornisce il concetto focale di *comunicazione artificiale*: gli straordinari risultati dei dispositivi non testimoniano che sono diventati intelligenti, ma che non cercano più di esserlo. La traduzione automatica funziona in base a regolarità e modelli rinvenuti nei testi multilingua reperiti in rete (*BigData* che diventano *BigCorpora*) ed è un chiaro esempio di sistema che raggiunge un obiettivo senza capirlo, essendo *esperto in statistica e sintassi probabilistica, non in semantica*⁸.

Simone Natale sottolinea come il test di Turing fosse un *benevolo inganno*: al fine di essere scambiata per un essere umano, la macchina doveva imitarne gli esiti, non i processi interiori⁹.

Nello Cristianini chiarisce come l'approccio



prestazionale sia una svolta: chi progetta i dispositivi rinuncia a comprensione e implementazione totale di caratteristiche e relazioni di causa-effetto di un sistema, a favore della *predizione statistica di quello che farà*, così come avviene, per esempio, nella composizione automatizzata (generazione) di testi e nell'attribuzione allo spam di mail corrispondenti a determinati pattern. Cristianini approfondisce e generalizza questa prospettiva: la capacità di un soggetto - naturale o artificiale - di agire nel proprio ambiente, in base a informazioni sensoriali, per prendere decisioni efficaci in rapporto a obiettivi definiti, anche in situazioni nuove e a fronte di contromisure ed elementi perturbanti, costituisce infatti il suo concetto di *intelligenza svincolata dall'antropocentrismo*, molto utile per la comprensione e la descrizione di dispositivi generativi¹⁰.

A Junaïd Mubeen dobbiamo infine l'idea di *complementarità*: la storia dei relativi artefatti racconta di come gli uomini abbiano via via realizzato dispositivi capaci di sbrigare gli aspetti esecutivi della matematica (tipicamente i calcoli) con *scopo emancipante*, perché conferivano loro più tempo e più spazio cognitivo per le attività logico-speculative, la vera qualità culturale¹¹.

I prompt: complementarità tra esseri umani e assistenti artificiali

Il ragionamento fin qui ricostruito sugli apparati forniti di intelligenza prestazionale, con particolare riferimento alla pratica della complementarità, mi ha condotto a una scelta molto netta: *formare all'impiego degli assistenti artificiali alle attività cognitive in compiti esecutivi e ripetitivi*¹², *continuando a riservare agli esseri umani la responsabilità e il controllo di intenzioni, obiettivi, progettazione dell'elaborazione.*

Va del resto in questa direzione anche la riflessione sull'autorialità dal punto di vista del diritto: «È l'essere umano a fornire l'input alla macchina e rimaneggiare l'output per avvicinarlo sempre più al risultato desiderato; è l'essere umano a decidere se l'output fornito è adeguato a essere pubblicato o al contrario se necessita [di] ulteriori rimaneggiamenti e aggiustamenti; è l'essere umano a decidere in quali sedi l'output sarà pubblicato e ad assumersene la responsabilità. In quest'ottica dunque anche i sistemi AI generativi, pur ponendosi davvero come qualcosa in grado di mettere pesantemente in crisi il diritto d'autore, riescono comunque ad essere ricondotti sotto il suo campo d'azione»¹³. Questo principio, avvalorato anche dal concetto di opere "utilitaristiche"¹⁴, è a mio giudizio di grande importanza generale, perché chiarisce che l'uso dell'IA generativa può mobilitare capacità culturali e cognitive elevate e raffinate, affermando quindi il primato di un'intelligenza umana evoluta e consapevole delle proprie esigenze ed abilità. Dal punto di vista professionale, invita a procedere in modo chiaro e cogente: *prima* definire gli obiettivi didattici e appurare *dopo* se vi siamo assistenti artificiali adatti a supportarne la realizzazione, per la parte esecutiva. Solo così assume senso acquisire - progressivamente, ma con estrema attenzione - la capacità di rivolgere (singolarmente o in gruppo) agli assistenti artificiali richieste articolate di prestazioni cognitive ben definite, ciò che il linguaggio tecnico settoriale chiama prompt. Come rappresentato dallo schema, la definizione del compito da assegnare all'assistente, infatti, non comprende solo accorgimenti di carattere tecnico¹⁵: prevede anzi la valutazione di significatività, equità, correttezza etica e valenza emancipatoria del processo che si mette in moto, di nuovo ciclico. La vocazione conversazionale

può tradursi in una sequenza dialogica di perfezionamento progressivo, anche per uscire dalla banalizzazione operativa, ma soprattutto culturale e cognitiva, della rigida sequenza [input]-[processo automatizzato]-[output univoco]. Siamo infatti di fronte a una delle caratteristiche generali dei contenuti culturali veicolati su supporto digitale, la plasticità, che riprenderemo.

Assistenza artificiale (ragionevole) a spunti operativi consolidati

La formazione agli assistenti artificiali si inserisce pertanto nell'insieme degli spunti operativi con capacità inclusiva che storicamente hanno costituito il percorso di lavoro proposto agli specializzandi sul sostegno. Definita la valenza cognitiva, propongo infatti l'esplorazione e l'esercitazione con i dispositivi (generativi, ma non solo) individuati come utili per il perseguimento degli obiettivi di ciascuno di essi. Ad esempio, nel caso delle schematizzazioni grafiche della conoscenza ne spiego e invito a sperimentare le differenti logiche di visualizzazione, congruenti con l'indicazione UDL di utilizzare diverse modalità di rappresentazione per estendere la comprensibilità di materiale concettuale ed espositivo nella direzione dell'apprendimento significativo¹⁶. Presento poi due applicazioni generative¹⁷ destinate alla costruzione di "mappe", insieme a due raccomandazioni di valore generale. In primo luogo, diffidare dalle descrizioni entusiastiche, che celano una proposta dispensativa e sostitutiva: la celebrazione della possibilità per gli studenti di produrre mappe automatiche rischia di esonerarli da processi cognitivi che devono invece affrontare direttamente, perché emancipanti e riappropriativi. In secondo luogo, verificare che l'uso di formule come "mappe concettuali", "mappe mentali" e così via sia rigorosa e non approssimativa e posticcia. Mi convince di più, inoltre, l'assistente che si propone come suggeritore di sviluppi progressivi e singolarmente verificabili della schematizzazione, rispetto a quello che si attribuisce il compito di risolutore di un problema nel suo insieme. Suggerisco infine di sperimentare, come autori diretti di materiali di mediazione didattica, le funzionalità di rappresentazione automatica per argomenti molto consolidati e tradizionali, "scolastici", che si prestano probabilmente a una prima elaborazione logico-visiva didascalica - tipicamente a blocchi di testo secondo una struttura gerarchica -, da cui costruire un'organizzazione più significativa e personale. In questi casi, infatti, la prima scrematura ha maggiori possibilità di essere convincente, evitando all'insegnante carichi di lavoro professionali privi di autentico significato cognitivo, tipici di una "intelligenza compilativa".

Questo primo esempio di mobilitazione ragio-

nevole dell'IA disponibile per un certo campo operativo ha due valenze trasversali: *l'insegnante può auto-esonerarsi dagli aspetti meccanici e ripetitivi della produzione di materiale per la mediazione didattica e perfezionare via via il lavoro, sfruttando la già citata plasticità e la malleabilità del supporto elettronico*. È infatti nota ed evidente la perfezionabilità della scrittura di testi mediante programmi di word processing. In questo caso, l'assistenza artificiale ragionevole può innanzitutto consistere in dettatura, trascrizione, traduzione¹⁹, sia di testi completi e autosufficienti, sia di materiali frammentati, da aggregare con dispositivi dedicati (cfr. infra). In entrambi i casi, l'assistenza artificiale ragionevole può ridurre il carico cognitivo superfluo derivante dalla redazione ripetitiva di testi a scopo esercitativo e/o esemplificativo, a bassa o nulla significatività cognitiva per l'insegnante. Per far ciò ChatGPT e Google Bard - dispositivi conversazionali *general purpose* - richiedono prompt molto curati e dialoghi di affinamento, ma altri dispositivi²⁰ sono impostati in modo più agile, perché danno anche la possibilità di scegliere funzionalità (riduzione, ampliamento, semplificazione) e caratteristiche (tono, lunghezza, scopo, punto di vista autoriale e così via) preimpostate. Questa duplice modalità si può ribaltare produttivamente sulla comprensibilità dei testi, perché aprono la strada al perfezionamento delle relative istruzioni sulla base dei protocolli di adattamento dei libri di testo²¹.

Gli assistenti alla redazione di testi, poi, intervengono anche sul rapporto cognitivo e operativo tra processo di scrittura e realizzazione di articoli per i blog, il cui focus è testo. Storicamente, gli specializzandi sul sostegno hanno fatto ricorso a blog per accumulare nel tempo materiali didattici anche molto differenti tra loro sotto forma di post, prevedendo però di poterli e riorganizzare con criteri diversi, etichette semantico-funzionali ("categorie") che ne consentono riaggregazioni dinamiche, in funzione di esigenze diverse degli utenti. In più, la disponibilità di assistenti artificiali alla redazione di testi suggerisce di immaginare blog "di servizio", non per un pubblico di fruitori degli articoli, ma come magazzini e officine di manipolazione di testi a scopo didattico.

Anche la produzione di libri digitali e lo storytelling possono trovare vantaggio dall'impiego di assistenti artificiali alla redazione di testi, per ampliare lo spettro del materiale esemplificativo ed esercitativo, a cui può essere associata la generazione automatizzata di immagini²². Nella storia del percorso di specializzazione, peraltro, le immagini sono state "croce e delizia". Se "far vedere" permette inclusione, va rispettato il diritto d'autore e non sempre sono disponibili immagini libere da questo vincolo. È quindi del tutto evidente il beneficio di generare immagini personalizzate trasformando le proprie esigenze in un prompt²³. Ancor più, in

considerazione del fatto che molti corsisti hanno scelto (e sceglieranno) di realizzare “immagini interattive”, su cui collocare simboli visivi che rimandano a collegamenti, con lo scopo di rendere esplicito e direttamente accessibile il tessuto di apprendimento sotteso a “ciò che si vede”²⁴.

È importante aver chiaro che, in generale, l'efficacia e la congruenza delle immagini non vanno misurate in termini espressivi, estetici e stilistici, ma su base cognitiva, ovvero sull'effettiva capacità di rappresentare ciò a cui di volta in volta si riferiscono, una sorta di consapevole concettualizzazione *by e in prompt*. L'esplicitazione del tessuto di apprendimento sotteso a un'immagine con una sua trasformazione che le aggiunge interattività può essere definita “estensione ipermediale”. Questa formulazione generale riguarda la possibilità tecnica, ma soprattutto cognitiva e culturale, di collegare l'una all'altra unità informative anche molto diverse, tanto più che QRcode e tecnologie della comunicazione di prossimità hanno da tempo reso “linkabili” a contenuti digitali anche gli oggetti materiali. Bene: sono ipermedialmente estendibili anche i filmati, in streaming da YouTube o prodotti direttamente²⁵.

Se l'utilità degli assistenti alla produzione di video²⁶ è scontata, vanno considerati con grande attenzione gli applicativi che consentono di ottenere resoconti e sintesi del contenuto di filmati presenti su YouTube, come elenchi puntati o articoli estesi²⁷, combinando accesso alla sottotitolazione, capacità di traduzione e di manipolazione testuale dei chatbot, a cui gli applicativi inviano prompt molto dettagliati.

Il materiale prodotto può essere copiato e incollato dovunque accrescendo l'accessibilità²⁸ e la comprensibilità del flusso video originale mediante diverse forme di coinvolgimento e di espressione. Testi, immagini, video realizzati con dispositivi generativi possono infine confluire in percorsi didattici strutturati sfruttando diversi aggregatori, così come gli applicativi destinati alla gestione e al monitoraggio di quiz possono essere alimentati con materiale prodotto dai chatbot testuali²⁹. Oltre che, ovviamente, da unità informative recuperate con una ricerca su internet: in questo caso la conversazione simulata può intervenire in modo utile, se restituisce contenuti già strutturati attendibili ed esaustivi e, soprattutto, se indica le fonti a cui ha attinto per la propria “risposta cumulativa”³⁰. Altra opzione utile è l'indicazione da parte del sistema di possibili sviluppi della conversazione³¹.

Concludendo, mi sembra che questo progetto preveda per ciascuno degli spunti del percorso di formazione (consolidato negli anni e via via integrato con gli applicativi intervenuti a semplificare il piano operativo e potenziare quello cognitivo) una coerente espansione quantitativa di analisi e sperimentazione dei potenziali assistenti arti-

ficiali. Sarà l'attenzione professionale degli esseri umani - gli autentici protagonisti dell'attività - a verificare criticamente se si verificherà anche un incremento qualitativo.

NOTE

1. In precedenza, il dispositivo escludeva gli utenti della comunità europea, con lo stesso accorgimento adottato da ChatGPT dopo le osservazioni del Garante della privacy, cioè il blocco degli indirizzi IP di quella provenienza.
2. Più precisamente: il variegato insieme dell'IA è stato identificato con la sola intelligenza artificiale “generativa”, dimenticando motori di raccomandazione commerciale, logistica di lavoro e distribuzione, ranking nella ricerca, combinazione di profili nei siti di incontri e così via. E quella generativa con le conversazioni, relegando in un angolo la produzione di immagini in base a indicazioni testuali.
3. Guida a Bard disponibile all'indirizzo <https://support.google.com/bard?sjid=12445836213057268994-EU#topic=13194540>.
4. L'età minima richiesta da ChatGPT è invece 13 anni.
5. Gli LLM sono algoritmi di Deep Learning (apprendimento profondo basato su reti neurali) in grado capaci di riconoscere contenuti, riassumerli, espanderli, tradurli e prevederli in base ad addestramento su enormi set di dati.
6. Chi fosse interessato ad approfondire può raggiungere <https://udlinbit.school.blog/>.
7. Intesa come facilitazione operativa potente per estensione cognitiva, non come trivializzazione, banalizzazione.
8. E. Esposito, *Comunicazione artificiale. Come gli algoritmi producono intelligenza sociale*, Bocconi University Press, Milano 2022.
9. S. Natale, *Macchine ingannevoli. Comunicazione, tecnologia, intelligenza artificiale*, Einaudi, Torino 2022.
10. N. Cristianini, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, il Mulino, Bologna 2023.
11. J. Mubeen, *L'intelligenza matematica. Cosa abbiamo che le macchine non hanno*, Einaudi, Torino 2023.
12. Questa impostazione è del resto quella di vari dispositivi di IA di tipo non generativo, come la trascrizione di file audio (“sbobinatura”), fatica cognitiva mai considerata prestazione intellettuale: al termine della prestazione automatizzata, sarà l'intelligenza umana a giudicare l'efficacia di quanto ottenuto, intervenendo a migliorarla, se necessario.
13. S. Aliprandi, *L'autore artificiale. Creatività e proprietà intellettuale nell'era dell'AI*, Ledizioni, Milano 2023.
14. L. Chimienti, *Diritto d'autore 4.0. L'intelligenza artificiale crea?*, Pacini Editore, Pisa 2020.
15. Mi preoccupa molto, invece, l'opportunistico proliferare di iniziative di micro-formazione al prompt engineering, inteso come insieme di virtuosismi tecnici, o la rivendita di liste di indicazioni preconfezionate. Molto più utili sono analisi e decostruzione di formulazioni

usate da altri utenti in rapporto con i risultati raggiunti, possibile per esempio su Midjourney.com.

16. M. Guastavigna, *Non solo concettuali. Mappe, schemi, apprendimento*, I Quaderni della Ricerca n. 23, Loescher, Torino 2015.

17. Si tratta di Algor Education e Whimsical. Qui e altrove, per non appesantire il testo, citerò i nomi commerciali degli assistenti artificiali indicati, assumendo che chi fosse interessato li potrà facilmente reperire e provare ricorrendo a un motore di ricerca.

18. Un'ampia trattazione di questo tema è contenuta in M. Guastavigna, *Risorse digitali per la scrittura*, in *Parla come badi*, «La ricerca» n. 10, Loescher, Torino 2016.

19. Questa funzionalità ha raggiunto risultati accettabili ed è ormai una risorsa di carattere generale, con cui realizzare versioni diverse degli stessi testi, destinate magari agli studenti in difficoltà in italiano, a cui possono essere fornite anche rappresentazioni grafiche salvate come immagini e trattate da Google traduttore.

20. Ad esempio, Microsoft Bing in modalità “Composizione”, l'assistente AI Jetpack (inserito di default in Wordpress), scrittura magica di Canva e numerosi altri applicativi.

21. Sul tema cfr. M. Guastavigna, *Il paradosso della semplificazione*, in *Parla come badi*, «La ricerca» n. 10, Loescher, Torino 2016.

22. È il caso di Tome.app.

23. I dispositivi più noti sono Midjourney, Image Creator from Microsoft Bing, Adobe Firefly, Dall-e 2, Night Café, Stable diffusion. Sono prevalentemente free, ma un impiego consistente e continuo richiede una licenza, questione non risolvibile attingendo alle risorse economiche dei singoli insegnanti.

24. Lo fanno, ad esempio, Lumi Education e Thinglink.

25. Nel laboratorio abbiamo usato Lumi Education e Edpuzzle, ma vi sono molti altri applicativi.

26. Segnaliamo Clipchamp di Microsoft, ma si tratta di un settore in continua espansione, per numero di applicativi e quantità di funzioni implementate, si veda <https://www.navigaweb.net/2023/04/migliori-generatori-di-video-ia-per.html>.

27. Per esempio, YouTubeDigest e Merlin, estensioni per browser.

28. Per quanto riguarda accessibilità e comprensibilità su un piano più generale, ricordo che PowerPoint ha due moduli di IA che possono accompagnare la produzione e lo svolgimento di una presentazione: “Designer”, che propone layout di impaginazione in funzione dei contenuti via via collocati sulla diapositiva e “Sottotitoli” automatici, che possono essere attivati anche in una lingua diversa da quella del relatore. Word ha una funzione “Editor” – disponibile anche come estensione per alcuni browser – che utilizza un indice di comprensibilità del testo e suggerisce alcune possibili perfezionamenti e correzioni.

29. Segnalo a questo proposito ChatPDF e UPDF, che scansionano un documento in formato pdf e consentono di dialogare specificamente sul suo contenuto, e Diffit, che adatta ai diversi livelli dell'istruzione USA, dal secondo all'undicesimo grado, il contenuto di un testo caricato dall'utente o raggiungibile via Internet; entrambi i potenziali assistenti sono inoltre multilingua e i loro prodotti editabili.

30. Il già citato Diffit, ad esempio, ha una particolare logica operativa: genera schede didattiche strutturate, graduate per livello di istruzione ed editabili, su un tema indicato dall'utente e fornisce in modo puntuale gli URL a cui fa riferimento. Microsoft Bing in modalità chat indica gli URL a cui ha attinto. Google Bard è collegato in modo diretto con Google Search. ChatGPT, se interpellato, risponde: “non ho accesso diretto alle mie fonti né posso citarle specificamente. [...] le mie risposte sono il risultato della mia comprensione delle informazioni fornite durante l'addestramento”.

31. Mentre scrivo, il più lineare in questo è di nuovo Microsoft Bing.

Approfondire



• S. Aliprandi, *L'autore artificiale. Creatività e proprietà intellettuale nell'era dell'AI*, Ledizioni, Milano 2023.

• L. Chimienti, *Diritto d'autore 4.0. L'intelligenza artificiale crea?*, Pacini Editore, Pisa 2020.

• N. Cristianini, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, il Mulino, Bologna 2023.

• E. Esposito, *Comunicazione artificiale. Come gli algoritmi producono intelligenza sociale*, Bocconi University Press, Milano 2022.

• M. Guastavigna, *Non solo concettuali. Mappe, schemi, apprendimento*, I Quaderni della Ricerca n. 23, Loescher, Torino 2015.

• S. Natale, *Macchine ingannevoli. Comunicazione, tecnologia, intelligenza artificiale*, Einaudi, Torino 2022.

• J. Mubeen, *L'intelligenza matematica. Cosa abbiamo che le macchine non hanno*, Einaudi, Torino 2023.

Marco Guastavigna

è stato insegnante di scuola secondaria; attualmente collabora con l'Università di Torino e si occupa di formazione all'uso consapevole e critico dei dispositivi digitali.

La banalità digitale

Sui limiti dell'IA e come beneficiarne:
riflessioni a un anno dal rilascio di ChatGPT.

di Silvia Capodivacca

*Cerèbro folle, a che pur l'arco tendi?
[...] Mentí, mentí chi ti descrisse sacro
Il disio di seguire conoscenza.
Il Versificatore*

«**E**bbene: come si trova oggi? Ne potrebbe fare a meno? No, non è vero? È uno strumento di lavoro come un altro».

È passato un anno da quando ChatGPT è stato messo a disposizione del pubblico: durante il tempo intercorso, si è parlato molto di questo *software*, delle sue potenzialità, dei rischi che ad esso sono correlati, ma anche della sua operatività, che di primo acchito sembrava frutto di una stregoneria anziché il prodotto più all'avanguardia della tecnologia informatica. Effettivamente, ai primi tentativi di utilizzo era difficile trattenere lo sbigottimento e, diciamo così, la meraviglia di fronte alle sbalorditive capacità sintattiche e cognitive del sistema. Da quando è diventato uno strumento integrato nella *routine* informatica di molti di noi, però, l'entusiasmo iniziale ha lasciato il posto a un più moderato interesse, accompagnato peraltro anche da alcune legittime perplessità. La mole dei commenti sulle mancanze, le inesattezze e i clamorosi errori dell'applicazione si è così fatta consistente: c'è chi ha mostrato i limiti della macchina nel campo che dovrebbe esserle più consono (quello della logica matematica, anche di livello elementare) e chi ha invece denunciato le menzogne che, senza una ragione evidente, essa ha prodotto, scusandosi poi ossequiosamente nel momento in cui le si fa notare la svista (o la clamorosa cantonata). Disservizi di questo tipo sono stati in parte già corretti, e comunque giustificati dalle spiegazioni che sono via via emerse a proposito del suo metodo di funzionamento. D'altronde, basta interrogare il chatbot per sciogliere alcuni dubbi: diversamente dal *modus operandi* di un calcolatore, infatti, il meccanismo di generazione di risposta di ChatGPT si basa sulla ricerca, tra le possibili sequenze di parole, di quella che ha la maggiore

probabilità di essere la replica corretta, in base a ciò che ha "imparato" durante le precedenti fasi di allenamento su enormi quantità di testo (ad esempio libri e pagine Internet), per prevedere la prossima parola in una sequenza.

Questa descrizione ci permette di trarre alcune considerazioni sul fenomeno, anche se credo sia utile mettere in chiaro fin da subito che si tratta di un bilancio tutt'altro che definitivo. Giorno dopo giorno riceviamo notizie con le quali si fa fatica a stare al passo: è già stata rivelata l'imminente integrazione di *software* di intelligenza artificiale su diversi dispositivi (dai telefonini ai supporti casalinghi di assistenza virtuale); Microsoft ha annunciato l'uscita, nel 2024, di un sistema operativo che sarà in grado di eseguire comandi vocali per far lavorare in sinergia diverse applicazioni (per cui, a quanto pare, sarà possibile per esempio "chiedere" al computer di copiare un testo da una pagina web a un foglio di lavoro, oppure di eseguire operazioni nei fogli di calcolo senza il bisogno di conoscere i comandi atti a compierle); OpenAI ci ha da poco informato che l'ultima versione di ChatGPT è in grado di "vedere, ascoltare, parlare", in modo che gli utenti (abbonati) possano integrare le immagini all'interno delle conversazioni e avere un dialogo vocale con il chatbot, che sarà di conseguenza sempre più presente nelle nostre giornate¹.

L'impressione, insomma, è che di qui a qualche anno guarderemo a questi primi prodotti dell'intelligenza artificiale con la compassione nostalgica che oggi riversiamo sui primi modelli di telefonia mobile, residui ormai quasi favolistici di tecnologie decisamente meno evolute di quelle attuali. La legge che Gordon Moore ha elaborato in relazione ai *transistor*, e che può essere riassunta nella formula «il numero di *transistor* su un *chip* raddoppierà approssimativamente ogni due anni», può infatti venire facilmente applicata anche ad altri ambiti dell'innovazione tecnologica, tra cui

senza dubbio quello dell'intelligenza artificiale, per cui è facile immaginare che la potenza di calcolo di questi sistemi avrà e continuerà ad avere uno sviluppo a dir poco esponenziale².

«Non vorrà mica sostenere che questa roba è geniale!»

Proviamo quindi a descrivere lo stato dell'arte, o almeno quello di avanzamento dei lavori, e lo facciamo, tra i vari possibili, dall'unico punto di vista che ci è consono, cioè quello di utenti da un lato e di insegnanti dall'altro. La nostra valutazione di ChatGPT non può prescindere, innanzitutto, dal rilevare la sua formidabile capacità sintattica e ortografica. I testi che produce sono consequenziali, ben organizzati dal punto di vista argomentativo e non presentano, tranne in rari casi, errori di carattere tipografico. Qualcuno ha proposto, immaginandola alle prese con un compito scritto o un'interrogazione in classe, di darle un 7 pieno: non raggiungerebbe l'eccellenza dei voti più elevati, ma si terrebbe dignitosamente al di sopra della sufficienza, oltre quella soglia guadagnata, talvolta con fatica, dalle studentesse e dagli studenti che, tra i banchi di scuola, ci stanno in carne e ossa. Fermandoci a ciò, però, finiamo per tralasciare una serie di elementi che invece non devono venire ignorati, primo tra tutti quello della elusività di molte delle sue risposte: innanzitutto, ChatGPT afferma molto raramente di non conoscere un argomento o di non saper fornire una risposta. Il chatbot, inoltre, cerca il più possibile di compiacere l'utente con cui si interfaccia, per cui se le si domanda il riassunto del racconto *Il Versificatore* di Italo Calvino sciorina una narrazione di senso compiuto, con dovizia di dati, dettagli letterari ed editoriali e persino un breve commento critico, su un racconto che avrebbe come protagonista il personaggio-feticcio calviniano Marcovaldo e che sarebbe stato pubblicato nel 1963 nella raccolta *Le stagioni in città*. Un discorso fluido, che si legge senza inciampi, salvo per il dettaglio, non trascurabile, che Calvino non ha mai scritto nulla del genere - quel racconto breve (di cui ci stiamo servendo per gli *incipit* dei paragrafi di questo contributo) è a firma Primo Levi.

Saremmo felici, in verità, di scoprire che la montagna di falsità che produce il chatbot sono il risultato creativo della sua capacità immaginativa; invece, la spiegazione circa la sua modalità di funzionamento, che abbiamo riportato in apertura, ci informa dell'opposto. Non si tratta di invenzioni creative, originali, ma del rimescolamento di una grande quantità di dati digeriti in precedenza e che ora vengono riproposti sulla base di un meccanismo di mera probabilità, ovvero: più è probabile che a un certo sintagma ne segua un altro, più il chatbot tenderà a riprodurlo nella sua risposta. La dinamica sta agli antipodi, crediamo,

di qualsivoglia forma di pensiero, che, al di là delle sue specifiche definizioni, non può prescindere da una presenza vigile nel processo di comprensione e, conseguentemente, di risposta. Quel che accade quando convochiamo l'IA è, da parte della macchina, l'attivazione di una meccanica irriflessa, dove si procede "alla cieca", assemblando pezzi di discorso senza alcuna contezza o supervisione critica su ciò che sta accadendo.

Nella nostra classe ideale, la studentessa o lo studente che con malizia si arrabatta, millantando conoscenze e inventando improbabili teorie critiche, non riceverebbe la sufficienza, ma verrebbe comunque incoraggiato a sfruttare le proprie capacità creative, incanalandole verso un processo che sia, per lei o per lui, anche di proficuo apprendimento. ChatGPT ci sembra un caso molto più disperato, e a essa riserveremmo una insufficienza grave e senza appello: oltre la patina di presentabilità dei suoi discorsi, la relazione che instaura con la conoscenza non offre nulla di nuovo, nessuno spunto interessante. È tecnicamente non euristica e congegnata per promuovere elementi ripetuti, dati triti e ritriti in base alla loro più probabile prevedibilità. Rappresenta l'ovvio nella sua più alta forma di efficienza e tutto ciò che produce di inconsueto è dovuto a un caso fortuito che nulla ha a che vedere con il guizzo di genialità che permette ad alcuni di pensare *out of the box*: il chatbot è, al contrario, perfettamente inscatolata, quadrata e ordinaria nei suoi processi come il più noioso e ottuso dei discorsi. Letteralmente, banale.

«Geniale no, ma commerciabile. Più che sufficiente per ogni scopo pratico».

Abbiamo provato a giudicare il chatbot sulla base dei parametri di valutazione che adatteremmo nei confronti delle nostre studentesse e dei nostri studenti a scuola. Bisogna, tuttavia, ammettere che il "voto" cambierebbe decisamente di segno qualora fossimo chiamati a stimarne l'utilità. Io stessa, da quando è stata resa pubblica, ho usufruito in più di qualche occasione delle sue funzionalità, che mi hanno aiutata a svolgere alcuni compiti di scrittura in un arco temporale decisamente più breve di quello che avrei impiegato con le mie sole forze. Sembriamo di fronte a una contraddizione, ma l'ambiguità si chiarisce facilmente, tenendo conto dell'abnorme quantità di testi, molto spesso soltanto "di servizio", che vengono richiesti dal mercato, dall'auto-sponsorizzazione di sé sui social all'apparato burocratico (progettuale, accademico, editoriale, amministrativo, gestionale e chi più ne ha più ne metta).

Luciano Floridi fornisce un dato eclatante in merito:

Alcuni anni fa i ricercatori della School of Information di Berkeley hanno stimato che l'umanità

abbia accumulato approssimativamente 12 esabyte di dati nel corso della storia fino alla diffusione dei computer e, a partire da questa, 180 esabyte già entro il 2006. Secondo uno studio più recente, il totale è cresciuto fino a più di 1.600 esabyte tra il 2006 e il 2011, oltrepassando in tal modo la soglia dello zettabyte (1.000 esabyte). Questo numero tende a crescere di quattro volte pressappoco ogni tre anni, cosicché si sono raggiunti gli 8 zettabyte di dati entro il 2015. Ogni giorno viene generato un numero sufficiente di dati da riempire tutte le biblioteche americane più di otto volte.³

Non ci serve conoscere nel dettaglio a cosa si riferiscono queste unità di misura per renderci conto della sproorzionata mole di informazioni che la rivoluzione digitale trascina con sé. Possiamo però domandarci quanti di questi testi possano essere considerati innovativi, creativi, geniali o anche solo il prodotto di una ponderata operazione di pensiero, e quanti invece non siano un'accozzaglia di banalità, frasi fatte e scontate di cui abbiamo bisogno in termini informativi o, più che altro, perché abbiamo "datizzato" il mondo, al punto che, tutti gli elementi in esso presenti, devono poter essere informatizzati, codificati, descritti. Si dice, per via metaforica, che chi è abituato a tenere un martello in mano, sarà spontaneamente indotto a cercare attorno a sé cose o superfici da colpire, oggetti da appiattire, chiodi da conficcare. Sono ormai due decenni che l'essere umano è dotato di una protesi, lo smartphone, attraverso cui il mondo è costantemente computato: viene naturale interrogare compulsivamente lo strumento per ottenere delle risposte - che tendenzialmente dimentichiamo in un tempo direttamente proporzionale alla facilità con la quale abbiamo ottenuto quelle informazioni. Questo processo ha però come inevitabile *pendant* anche la necessità di una produzione incessante di quegli stessi dati di cui pretendiamo di fruire, siano essi significativi o piuttosto, come nella stragrande maggioranza dei casi, ridondanti, superflui e sostanzialmente banali.

«Basta! Ne ho abbastanza di questo sporco mestiere: non sono un menestrello. Non sono un versificatore. Puh, è una galera, questa».

Non dovremmo tuttavia essere frettolosamente liquidatori nei confronti della banalità, che rappresenta uno stato nient'affatto neutro rispetto alle conseguenze che può comportare. Ce lo ha insegnato Hannah Arendt, associando l'idea di banalità addirittura al nazismo e all'atteggiamento di irriflessa obbedienza tenuto da alcuni suoi gerarchi. Di Eichmann, su cui pendevano quindici capi d'accusa tra cui crimini contro l'umanità, la filosofa dichiara espressamente che egli non aveva la statura di alcun "cattivo" per antonomasia - «non era uno Iago né un Macbeth» -: le sue

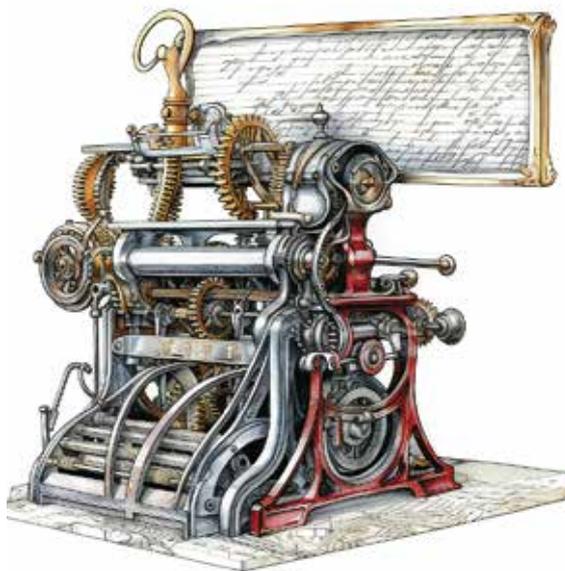
azioni sono state possibili a causa di una profonda «mancanza di immaginazione» e dal fatto che egli

non capì mai cosa stava facendo [...]. Non era uno stupido. Era semplicemente senza idee (una cosa molto diversa dalla stupidità) [...]. Quella lontananza dalla realtà e quella mancanza d'idee possono essere molto più pericolose di tutti gli istinti malvagi che forse sono innati nell'uomo.⁴

La banalità ci deve sempre porre in uno stato di allerta. Rispetto alla banalità del male, tuttavia, il caso della banalità digitale, oltre che rappresentare un oggettivo pericolo, offre anche alcune *chance* che sarebbe uno spreco non cogliere. L'ordine di rischio con cui abbiamo a che fare concerne la possibilità di abbassare il livello delle nostre produzioni scritte (per quanto riguarda l'uso dell'intelligenza artificiale nel campo della produzione di testi), ma più in generale di adeguare la realtà all'elementarità delle macchine. Luciano Floridi esprime bene questa preoccupazione:

Le ICT [Information and Communication Technologies] non stanno diventando più intelligenti, rendendoci al contempo più stupidi. È il mondo invece che sta divenendo un'infosfera sempre più adattata alle limitate capacità delle ICT. [...] adattiamo l'ambiente alle nostre tecnologie smart per assicurarci del fatto che queste possano interagire con successo con esso.⁵

Parallelamente, e non per forza in forma esclusiva rispetto al pericolo accennato, esiste però anche un'opportunità connessa al dilagare del digitale e della sua banalità. Concerne la possibilità che abbiamo, e che anche i chatbot di ultima gene-



← Immagine creata da Midjourney con il prompt "Macchina per pensare".

razione ci fanno toccare con mano, di chiudere i conti con le mansioni più banausiche e frustranti, delegandole, pur con la dovuta vigilanza critica, alle macchine. Vediamo di cosa si tratta.

Come annunciato poco sopra, secondo ChatGPT *Il Versificatore* è un racconto di Italo Calvino in cui «Marcovaldo scopre una macchina nella redazione di un giornale che, con l'uso di magneti, è in grado di comporre versi poetici scegliendo parole a caso. Marcovaldo è affascinato da questa macchina e inizia a utilizzarla, producendo poesie che finiscono per essere pubblicate sul giornale. Ciò attira l'attenzione dei lettori e Marcovaldo diventa una sorta di celebrità». Non serve continuare la citazione della sinossi prodotta dal bot, poiché niente di ciò che scrive trova corrispondenza nella "vera finzione" ideata da Primo Levi. In quest'ultima, si parla di un poeta, stanco del lavoro ripetitivo e poco appagante di composizione di testi su commissione, il quale accetta di affidarsi a una macchina per produrli. La narrazione conta poche pagine e vale la pena di recuperarle per il piacere di una loro lettura diretta. In questa sede, ci siamo limitati a riportare, a mo' di titolo dei paragrafi, alcuni passaggi letterariamente icastici, che hanno dato l'abbrivio ad alcuni ordini di considerazioni a proposito dell'intelligenza artificiale – concetto peraltro sconosciuto all'epoca della pubblicazione della raccolta. Anche a proposito delle occasioni offerte dal digitale, il poeta protagonista della narrazione ha qualcosa da insegnarci. La sua scelta di affibbiare alla macchina parte del proprio lavoro, infatti, è raggiunta per esasperazione, ma gli permette di liberare alcune sue capacità altrimenti assopite o comunque asservite all'ordinaria amministrazione.

Esiste una gioia, nel nostro lavoro, una felicità profonda, diversa da tutte le altre, la felicità del creare, del trarre dal nulla, del vedersi nascere davanti, a poco a poco, o d'un tratto, come per incanto, qualcosa di nuovo, qualcosa di vivo che non c'era prima... (Freddo ad un tratto) Prenda nota, signorina: "come per incanto, qualcosa di nuovo, qualcosa di vivo che non c'era prima, puntini": è tutta roba che può servire.⁶

La macchina si è insediata nel suo ufficio, ma con questo egli recupera una dimensione di lavoro molto più piacevole e connessa alla sua indole profonda, che è quella di poeta e non di «mestierante». La riguadagnata vena lirica compare fin dalla descrizione del suo talento, tant'è che alla segretaria viene chiesto subito di «prendere nota».

Oggi costituisce uno scenario non più eccessivamente fantascientifico quello di un mondo in cui l'essere umano è liberato dai lavori più pesanti da un lato, e meno soddisfacenti dall'altro. All'interno di questa cornice, c'è senza dubbio spazio per una riappropriazione, da parte degli esseri umani,

dell'area della creatività, della riflessione ponderata, del pensiero e della comprensione profonda, così come del guizzo geniale e dell'intuizione estrosa fino a poco fa relegata alla nicchia, sempre più ristretta, del tempo cosiddetto "libero". Le nuove generazioni stanno già compiendo passi che a quelle precedenti sembrano tutt'ora impossibili: il movimento delle e dei *quitter* ne è testimonianza, con la loro etica del rifiuto a senso unico (cioè non per forza avendo a disposizione altre offerte) di mansioni che non permettono loro di esprimersi adeguatamente sul posto di lavoro. Non è un fenomeno da etichettare come l'atto di una popolazione di privilegiati, che hanno alle spalle famiglie che concedono loro questo "lusso", perché non è sempre questo il caso. Si tratta invece di un fenomeno più complesso e più incisivo, indicativo della necessità di ripensare il concetto di lavoro dalle sue fondamenta, affinché, anche grazie alle macchine e all'intelligenza che viene loro generosamente attribuita, si ristrutturino un tessuto sociale all'altezza dei bisogni, ma anche delle ambizioni, delle aspirazioni e dei flussi di desiderio di ciascuna e ciascuno di noi.

NOTE

1. *ChatGPT ora vede, sente e parla*, su Wired, consultabile all'indirizzo <https://www.wired.it/article/chatgpt-novita-immagini-voce/> (ultimo accesso 27 settembre 2023).
2. È quanto emerge da uno studio promosso dalla Cornell University, secondo cui, dall'avvento del Deep Learning, all'inizio del 2010, l'addestramento delle macchine ha subito un'accelerazione, raddoppiando ogni 6 mesi circa, ma già alla fine del 2015 è emersa una nuova tendenza e l'addestramento è risultato da 10 a 100 volte superiore. Si veda J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn, P. Villalobos, *Compute Trends Across Three Eras of Machine Learning*, in «arXiv», consultabile all'indirizzo <https://doi.org/10.48550/arXiv.2202.05924> (ultimo accesso 20 settembre 2023).
3. L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo* (2014), tr. it. di M. Durante, Raffaello Cortina, Milano 2017, p. 13.
4. H. Arendt, *La banalità del male* (1963), tr. it. di P. Bernardini, Feltrinelli, Milano 2008, pp. 290-291.
5. L. Floridi, *La quarta rivoluzione* cit., pp. 163-164.
6. P. Levi, "Il Versificatore", in *Storie naturali* (1966), Einaudi, Torino 2022, ed. dig.

Silvia Capodivacca

ha studiato Filosofia e Storia tra Padova, Bologna e New York. Attualmente svolge attività di ricerca all'università di Udine e collabora con la casa editrice Loescher come autrice e formatrice didattica. Maggiori informazioni e aggiornamenti sulla sua attività sul sito personale, silviacapodivacca.com.

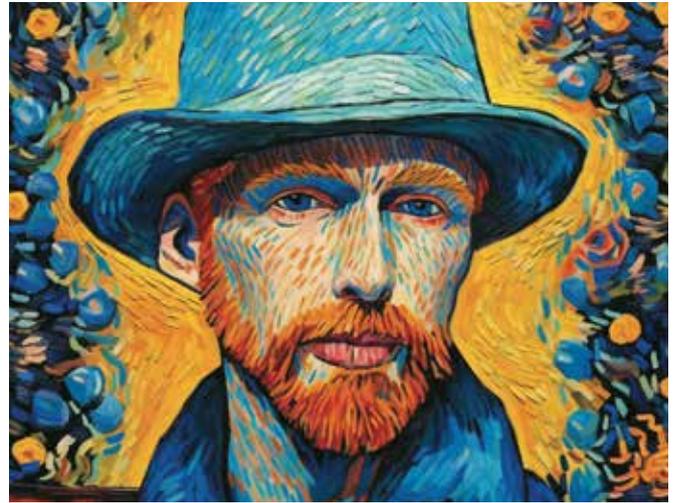
IA e stile

Come le intelligenze artificiali modificheranno la nostra idea di “stile” nelle arti visive.

di Francesco D’Isa

È ormai da tempo che si parla un po’ ovunque di intelligenze artificiali, ma per quel che riguarda le arti visive l’interesse si concentra soprattutto sui software basati sul machine learning che creano immagini a partire da un comando testuale (d’ora in poi TTI, text-to-image), o sui neonati text-to-video. Negli ultimi mesi ho avuto modo di esplorarli molto, sia nella teoria che nella prassi, e continuo a pensare che porteranno una rivoluzione per l’arte visiva pari all’avvento della fotografia o della computer graphic (e prima ancora della prospettiva). È ancora presto, perché la maggior parte di questi software è in fase embrionale, sebbene siano in velocissima evoluzione, ma le scosse cominciano a farsi notare. Come cambierà il nostro modo di fare arte e di rapportarci alle immagini?

Prima di provare a rispondere, qualche breve appunto tecnico. Per citare il celebre esperimento mentale di Searle, questi software somigliano più a “stanze cinesi” che ad androidi antropomorfi; sono infatti dei modelli algoritmici basati su enormi quantità di dati creati dagli umani, su cui lavorano su base statistica allo scopo di rispondere con successo alle nostre richieste. Se inserisco il comando testuale “un gatto” in un software come Midjourney, questo strumento, che per essere costruito ha “mangiato” moltissime foto di gattini etichettate con la parola “gatto”, plasmerà un’immagine di un gattino inesistente. Gli ingredienti che rendono possibile la magia sono essenzialmente il materiale di partenza (moltissime foto, di cui alcune di gatti), il modo in cui viene catalogato (“questa immagine è una foto di un gatto che soffia”) la potenza di calcolo della macchina e il metodo con cui viene addestrata, grazie al quale si estrapola una immensa rete di relazioni invisibili, in questo caso tra parole e pixel. Lo spazio creativo di una IA viene plasmato essenzialmente dai dati e dalla loro etichettatura - anche se questa è una spiegazione molto semplicistica¹. Nonostante quel che sostengono alcuni critici, che



non si possano creare delle immagini inedite con una TTI è falso e lo sappiamo, anzi, è estremamente difficile non ottenere qualcosa di nuovo, dato che la stragrande maggioranza delle generazioni non ha equivalenze col materiale nel dataset². Resta però da chiedersi se è possibile creare del contenuto *originale*, inteso come innovativo in ambito visivo. L’artista e teorico Lev Manovich nel suo articolo *Towards “General Artistic Intelligence”*³ suggerisce che le TTI siano già tecnicamente più competenti di molti studenti d’arte e artisti adulti - e a mio parere è fin troppo generoso con gli artisti umani. Tuttavia, queste tecnologie tendono a sintetizzare immagini idealizzate e senza difetti, con un gusto che l’autore individua vicino al classicismo. Inoltre la TTI risulta facilmente retorica, stereotipata e priva di originalità, caratteristiche che Manovich associa al kitsch. È un’analisi che allo stato attuale dello sviluppo condiviso, ed è legata al dataset e al tipo di sviluppo di software.

Il mondo cognitivo di una TTI è composto in egual misura dai suoi dati e da come questi sono etichettati; è molto più piccolo di quello umano, che è a sua volta più piccolo di quello di alcuni animali, o di quello che giace oltre le capacità dei nostri sensi intelletti. La porzione di radiazioni che riusciamo a percepire come “colore” è una frazione molto piccola di tutto lo spettro elettromagnetico, che include i raggi X, i raggi gamma, ultravioletti, infrarossi, microonde e onde radio. Api, farfalle, serpenti, uccelli e alcuni gamberetti percepiscono colori a noi invisibili, ma nonostante questi limiti il nostro mondo cromatico è sufficiente a dar luogo alla sterminata ricchezza di ogni possibile visione umana. Allo stesso modo, anche il mondo di una TTI, ben più povero rispetto al nostro, è comunque immenso, e lo spazio di possibilità (o “spazio latente”) che si esplora attraverso i prompt (i comandi testuali con cui si crea l’immagine) nasconde moltissimi luoghi mai visti. I *prompter* sono in effetti esploratori ed esploratrici di questo spazio ma, a differenza di un navigatore del XV secolo, scoprono un mondo che non esiste

↑ Immagine creata da Midjourney con il prompt “Ritratto nello stile di van Gogh”.

→
Due immagini
create da
Midjourney
con il prompt
“Picasso si fa un
selfie”.



senza di loro, perché è composto da possibilità che vengono in essere solo quando sono osservate – metaforicamente, è un mondo più quantistico che classico. In questa esplorazione, inoltre, ogni strada porta inevitabilmente a un luogo. Il più delle volte sono luoghi familiari (come “un gatto dipinto da Van Gogh”), altre invece sono inediti.

Con le TTI ci si accorge che non sempre lo stile è separabile dal contenuto, perché in alcuni casi è possibile applicare il timbro di un artista a un soggetto e in altri no. È una cosa di cui ci si accorge durante l'esplorazione, perché ogni termine usato nel prompt porta con sé il suo universo semantico – così ad esempio la parola corvo “corvizza” anche il resto dell'immagine, quella farfalla la “farfallizza”, e così via, sempre in base alle associazioni estrapolate da dati ed etichette. Pensare che limitarsi a materiale noto sia un ostacolo per l'originalità è errato, perché come scrive ancora Manovich a proposito delle avanguardie, «L'arte moderna si è sviluppata reinterpretando e copiando immagini e forme dalle tradizioni artistiche antiche, come le stampe giapponesi (van Gogh), la scultura africana (Picasso) e le icone russe (Malevich). Pertanto, gli artisti hanno rifiutato solo i paradigmi dominanti dell'arte accademica del tempo, l'arte realistica e da salone, ma non il resto della storia dell'arte umana»⁴. Sebbene il nostro mondo cognitivo sia molto più vasto, è comunque finito, e siamo vincolati ad esso per produrre le nostre immagini. Per quante lodi si possano tessere alla creatività umana (forse sopravvalutata, come qualunque altro tratto umano) non può immaginare l'impossibile. Ogni visione è vincolata a un mondo e sebbene i vincoli delle TTI siano più stretti dei nostri, c'è davvero molto da esplorare nello spazio latente. Certo, ad alcuni questo pur immenso spazio potrebbe stare strettino, ma con qualche competenza tecnica e un po' di potenza di macchina è già possibile personalizzare questi strumenti con il fine-tuning

su un dataset scelto dall'utente. È una funzione relativamente complessa e al momento disponibile quasi solo per i software open source, ma quando sarà di largo accesso, ogni persona potrà avere una sua TTI personale, tarata sulle proprie esigenze e stili. È possibile farlo perché queste tecnologie imparano molto velocemente nuovi stili; basta un numero relativamente esiguo di quadri di Van Gogh (dopotutto non ne esistono moltissimi) per creare “un gatto alla Van Gogh”, perché la macchina ha già imparato elementi fondamentali per realizzare l'immagine (come che cos'è un gatto) grazie ad altri dati.

Nel parlare di IA bisogna rifuggire il più possibile le antropomorfizzazioni, ma la somiglianza di queste immagini con quelle ipnagogiche che si vivono nel pre-sonno o in alcune fasi della meditazione è sconcertante. Se consideriamo il prompt come una mappa per cercare/creare qualcosa nello spazio latente, potremmo dire di aver sbagliato strada, per arrivare in quello che sembra (ma non è, perché non esiste) l'inquietante inconscio della macchina. Se vogliamo seguire la metafora, si tratta più correttamente di una manifestazione del nostro inconscio collettivo, un serbatoio junghiano di icone archetipali – o meglio di formule per generare tali icone. Se volessimo paragonare l'operato di una TTI agli archetipi teorizzati dal celebre studioso, infatti, credo che questi andrebbero collegati ai pattern invisibili in grado di generare serie di icone coerenti e potenzialmente illimitate. È forse necessario rivedere il concetto di archetipo, o di idea platonica: non sono perfette ed eterne rappresentazioni di ciò che è simboleggiato, ma formule – queste sì, immobili – che possono comporre infinite versioni del medesimo tipo.

È certo che questi strumenti avranno un effetto nel mondo dell'arte, soprattutto via via che diventeranno di uso comune, ma indovinare in che modo va ben oltre le mie capacità; l'unica cosa su cui

posso scommettere è che *non ucciderà l'arte*, come sostengono alcuni tra i più apocalittici detrattori. La morte dell'arte è un allarme che è stato annunciato troppo spesso lungo i secoli per mantenere ancora la sua credibilità e se questa data esiste probabilmente coinciderà con quella della nostra estinzione. Non è nemmeno detto che l'impatto delle IA in quest'ambito sarà così forte; come nota la storica dell'arte Valentina Tanni⁵, la comparsa di una nuova tecnologia visiva, per quanto potente, non è più in grado di per sé di innestare una rivoluzione in ambito artistico, ma avrà comunque delle ripercussioni, perché «al di là delle scelte individuali, l'influenza delle innovazioni tecnologiche è di carattere sistemico: la sola esistenza di determinati strumenti cambia la percezione del mondo, estende il senso della possibilità, influisce sulle strutture sociali ed economiche, modifica il senso estetico e il senso comune. Tutti questi cambiamenti, nel loro complesso, si rifletteranno senza dubbio nell'arte delle generazioni a venire, secondo modalità che sono ancora difficili da prevedere».

La nascita della fotografia ha rivoluzionato la prassi artistica anche di chi non ha sempre utilizzato il nuovo strumento - basti pensare alle avanguardie del Novecento - ed è plausibile che accadrà lo stesso in seguito dell'introduzione delle IA. Anzi, sta già accadendo, se si guarda alla scissione che questi strumenti hanno già operato tra lo stile e la competenza tecnica atta a realizzarlo. Una volta scoperto un peculiare stile rappresentativo, con le TTI è possibile replicarlo e ibridarlo indefinitamente grazie a pochi esempi - un fatto che non può non avere conseguenze. Non credo che questo deprezzerà l'abilità artigianale, perché quest'ultima è legata al lavoro umano e alle irriducibili caratteristiche di un oggetto non digitale - anzi, è possibile che l'artigianato accresca il suo valore. Ma per quel che riguarda la vita dell'immagine nella sua riproducibilità tecnica, dunque a mezzo stampa, video, web o altro, il cambiamento è senza dubbio enorme. Il fatto che queste tecnologie siano state accolte tendenzialmente con curiosità dagli artisti e dalle artiste che lavorano con le gallerie d'arte o i musei e con aperta ostilità da chi lavora in ambito illustrativo ed editoriale è un ulteriore sintomo di quali siano le maggiori zone di impatto delle TTI. Comprensibilmente, a preoccuparsi è il settore economico dove la tecnologia sarà più influente, anche se è presto per fare previsioni, perché abbiamo imparato dalla fotografia - e prima ancora dalla stampa e dallo sviluppo della prospettiva nel disegno - che un nuovo modo di creare immagini modifica anche il nostro gusto.

Al momento con questi strumenti è più facile ottenere delle immagini realistiche e patinate - un risultato dovuto sia al settore commerciale dove vanno a inserirsi sia alla loro "alimentazione": dato che l'umanità produce moltissime immagini

commerciali, c'è poco da stupirsi se sono quelle che le macchine hanno imparato meglio. Il mio non è un giudizio di valore, alcune sono comunque molto belle, ma in un contesto in cui per produrre certe immagini non saranno più necessarie costose apparecchiature e set, il loro valore estetico si modificherà. Le immagini che prima richiedevano grandi abilità tecniche e ora pochi click potrebbero non piacerci più. Il parallelo che mi viene in mente è nel campo dei video per adulti - da sempre un ottimo osservatorio per le innovazioni tecnologiche - dove la crescita di prodotti confezionati secondo standard cinematografici sempre più elevati ha portato allo sviluppo del porno amatoriale. Dopo una tendenza alle immagini patinate, potrebbe verificarsi una a favore di quelle amatoriali, ma quando TTI e fotografia non saranno più distinguibili su schermo anche questo canone scomparirà. D'altra parte la rivoluzione visiva dovuta a questi strumenti non si verificherà tanto e solo per via di chi li userà o li rifiuterà, ma per l'inevitabile uso di massa, che, come con la fotografia, contribuirà alla costruzione e decostruzione del nostro fluido rapporto con l'immagine.

NOTE

1. Questa è una spiegazione molto semplificata e parziale del funzionamento di una TTI. Per un approfondimento accessibile senza competenze tecniche rimando a M. De Baggis, A. Puliafito, *In principio era ChatGPT*, Apogeo, Milano 2023.
2. Per approfondire il funzionamento tecnico di questi modelli si veda <https://stable-diffusion-art.com/how-stable-diffusion-work>.
3. L. Manovich, *Towards 'General Artistic Intelligence'?*, Art Basel, 1 giugno 2023 <https://www.artbasel.com/stories/lev-manovich>.
4. L. Manovich, E. Arielli, *Artificial Aesthetics: A Critical Guide to AI in Art, Media and Design*, consultabile all'indirizzo <http://manovich.net/index.php/projects/artificial-aesthetics-book>.
5. V. Tanni, *Arte e intelligenza artificiale. Una storia che inizia negli Anni Cinquanta*, uscito su Artribune il 30 giugno 2023, consultabile all'indirizzo <https://www.artribune.com/progettazione/new-media/2023/06/arte-intelligenza-artificiale-storia/>.

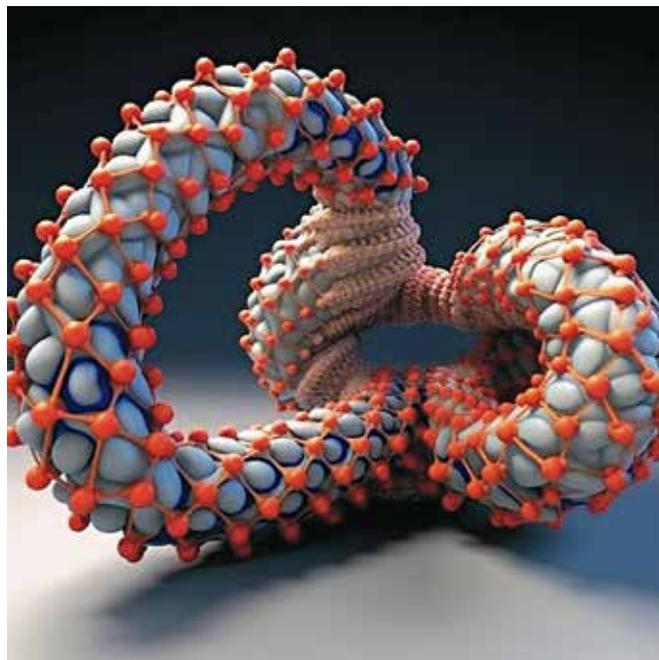
Francesco D'Isa

di formazione filosofo e artista digitale, ha esposto in gallerie e centri d'arte contemporanea internazionali. Dopo l'esordio con la graphic novel *I*. (Nottetempo, 2011), ha pubblicato saggi e romanzi per Hoepli, effequ, Tunué e Newton Compton. Il suo ultimo romanzo è *La Stanza di Therese* (Tunué, 2017), mentre per Edizioni Tlon è uscito il suo saggio filosofico *L'assurda evidenza* (2022). Direttore editoriale della rivista culturale «L'Indiscreto», scrive e disegna per varie riviste.

La medicina intelligente

Gli sviluppi della tecnologia e dell'intelligenza artificiale al servizio degli operatori della salute e dei pazienti: le due facce della stessa medaglia.

di Grazia Pozzi



Da sempre la pratica medica e quella chirurgica si sono avvalse di strumenti per la diagnosi e per la cura, che dai tempi di Ippocrate a oggi sono divenuti sempre più sofisticati e precisi.

Viene spontaneo pensare che l'intelligenza artificiale (IA) sia una fisiologica evoluzione del progresso tecnico-scientifico, ed è quindi lecito considerarla come uno strumento moderno.

Dobbiamo tuttavia operare una fondamentale distinzione tra gli strumenti che hanno reso possibile e agevolato la pratica della medicina sino ad ora e la IA.

Classificando le tecnologie come di primo ordine, quelle che si interpongono direttamente tra il medico e la natura (come ad esempio lo stetoscopio o lo sfigmomanometro) o di secondo ordine, ossia quelle che stanno tra medico ed un'altra tecnologia (ad esempio un braccio meccanico alimentato da un motore) abbiamo tecnologie su cui opera in modo diretto l'azione umana. La IA è invece definita una tecnologia di terzo ordine, ossia una tecnologia che sta tra una tecnologia e un'altra e sul cui processo potremmo anche non avere il controllo¹.

Pensiamo dunque alla IA come a un programma informatico che interpreta ed elabora dei dati ottenuti da altre macchine e da altri programmi. In questa separazione dell'uomo dal processo di elaborazione di un risultato da parte della macchi-

na viene identificata una delle principali insidie della IA.

La IA è stata definita in molti modi, e non vi è una definizione su cui tutti concordino.

Nell'immaginario comune e nella pratica quotidiana la sperimentiamo già sotto forma di tecnologie capaci di svolgere rapidamente dei compiti complessi, che richiedono generalmente più tempo all'essere umano per essere elaborati, e pertanto vengono definite intelligenti.

Su questo punto è necessario aprire una parentesi su cosa si intende per intelligenza.

Per alcuni la IA rappresenta l'unione tra l'azione di una macchina e il comportamento intelligente nel produrre un risultato, ma in realtà è il contrario: nell'utilizzo della tecnologia di IA vi è una separazione tra la capacità di risolvere un compito e l'esigenza di essere intelligenti per farlo. Non appena l'IA può eseguire un compito particolare, quel compito stesso non è più definito intelligente. La separazione tra agire e intelligenza produce conseguenze importanti con sfide etiche di grande portata².

Anche in medicina non vi è ambito che non stia subendo qualche trasformazione a opera delle tecnologie più evolute attraverso la IA.

Prendiamo ad esempio l'area della dermatologia: i software di intelligenza artificiale sono in grado di identificare con elevata accuratezza una lesione cutanea maligna, confrontandola con un database che gli sia stato precedentemente fornito.

↑ Immagine creata da Midjourney con il prompt "Membrane, proteine e molecole".

Non è un futuro fantascientifico, ma una realtà emergente quella in cui il chirurgo opera indossando degli occhiali tecnologici che gli consentono una visione aumentata della realtà (*augmented reality*), grazie ad una sovrapposizione virtuale delle immagini radiologiche del paziente che sta operando sulla zona anatomica su cui interviene.

Anche in psichiatria sono già noti gli effetti positivi che l'immersione del paziente e del terapeuta in un ambiente di realtà virtuale (*virtual reality*) crea per entrambi. Se da una parte il paziente sperimenta in sicurezza gli ambienti virtuali nei quali intraprende il percorso terapeutico, dall'altra parte anche il medico si può sentire rassicurato nell'agire in uno spazio virtuale, dove è praticamente inesistente il suo rischio occupazionale, che invece è spesso presente negli incontri reali con il paziente psichiatrico.

Gli ambienti virtuali in cui la realtà viene riprodotta fedelmente sono detti mondi specchio (*mirror worlds*) e in essi i pazienti e i medici possono esplorare, interagire e sperimentare la rappresentazione digitale fedelmente ricreata del mondo reale in modi nuovi e significativi.

La creazione di questi mondi specchio ha effetti potenziali anche nell'insegnamento in ambito medico. Immersi in un ambiente virtuale, gli studenti possono simulare e apprendere come eseguire gli interventi chirurgici o le procedure di soccorso e possono esercitarsi così a gestire le situazioni reali senza alcun rischio per i pazienti e senza praticamente limiti alla numerosità dei discenti che possano partecipare alle simulazioni nelle sale operatorie virtuali.

Realtà aumentata e realtà virtuale sono solo alcuni degli aspetti nuovi ed emergenti dell'utilizzo dell'intelligenza artificiale in ambito medico e chirurgico.

Anche l'utilizzo di algoritmi è una forma di IA. Per ottenere diagnosi più rapide e precise vengono utilizzate delle reti neurali, conosciute anche come reti neurali artificiali (RNA), che sono un tipo di modello computazionale che si ispira al funzionamento del cervello umano.

Tali reti sono impiegate nell'ambito nell'apprendimento automatico dell'intelligenza artificiale, il machine learning, che le consente di svolgere diverse funzioni come ad esempio la classificazione delle immagini. Le reti neurali sono composte da unità di calcolo dette neuroni artificiali o "perceptroni" organizzati in strati, che attraverso connessioni tra loro apprendono dai dati in modo da poter fare previsioni o compiere decisioni basate sugli input forniti. Quindi l'apprendimento avviene attraverso l'addestramento della rete neurale su un insieme di dati.

Grazie alla vasta quantità di dati medici disponibili, le reti neurali possono ad esempio analizzare i sintomi dei pazienti e confrontarli con un database

di casi precedenti, aiutando i medici a identificare patologie complesse.

In campo radiologico, ad esempio, esistono dei software in grado di analizzare la radiografia di un paziente e definire con accuratezza la presenza di una frattura di un osso in quell'immagine radiografica. Questo tipo di strumento allo stato attuale non può ancora sostituire la competenza di un medico specialista nel formulare la diagnosi di frattura, ma può comunque fornirgli un supporto, ad esempio nelle situazioni di forte stress lavorativo quando lo specialista debba sostenere un grande carico di lavoro in condizioni disagiate, come ad esempio nei turni di Pronto Soccorso, specie notturni.

La tecnologia intelligente deve quindi essere intesa e utilizzata come uno strumento che affianca il medico e lo sostiene nelle situazioni di particolare complessità o in caso di sovraccarico lavorativo.

Analoghe considerazioni possono essere svolte a riguardo dei modelli linguistici generativi o LLM (Large Language Model), tra i quali vi è il chatbot ChatGPT di OpenAI.

I modelli linguistici sono programmi in grado di generare testi, immagini e algoritmi basandosi su dati con i quali sono stati precedentemente addestrati (*Generative Pre-trained Transformer*).

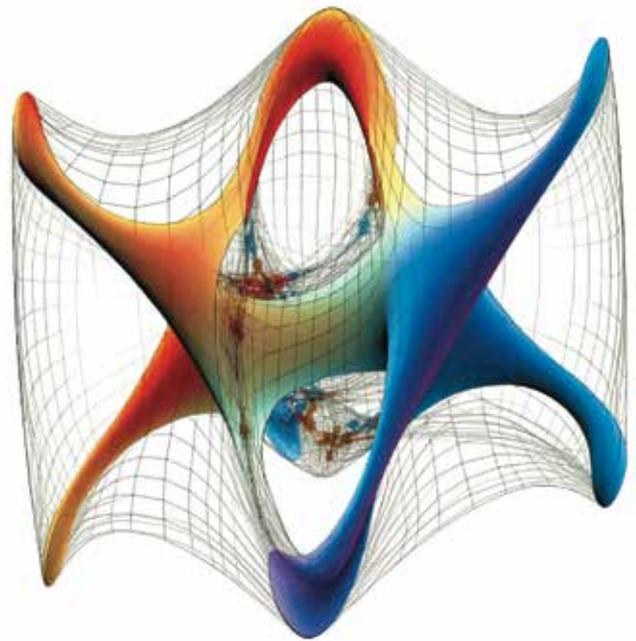
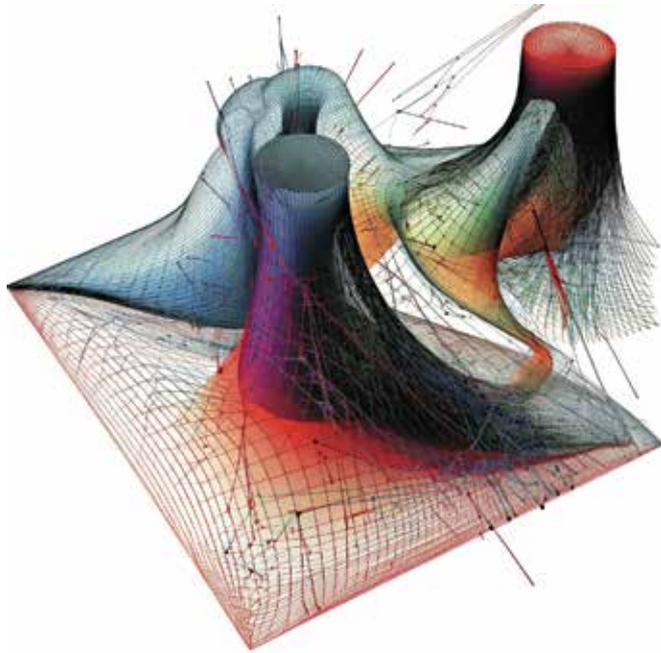
Qualche mese fa, mentre presentavo una relazione sulle tecnologie intelligenti e le possibili applicazioni in ambito medico radiologico, ho condiviso la definizione che ha dato ChatGPT quando l'ho interrogato chiedendogli di definire sé stesso e le sue funzioni:

Sono un programma di intelligenza artificiale creato per elaborare il linguaggio naturale e fornire risposte basate su testo a una vasta gamma di domande. La mia "intelligenza" è basata su algoritmi di apprendimento automatico e su un vasto corpus di testo da cui ho imparato. Sono in grado di elaborare informazioni e generare risposte in modo efficiente, ma la mia "intelligenza" è limitata alla mia programmazione e ai dati di cui sono stato addestrato [...] Sono qui per fornire supporto e informazioni utili, ma è sempre consigliabile verificare le informazioni critiche o importanti in fonti affidabili.

Ci tengo a sottolineare che la parola "intelligenza" tra virgolette è opera del programma stesso.

Il mio atteggiamento è quantomeno cauto nel sostenere a priori questo tipo di intelligenza artificiale; essa stessa mette in guardia dal farne un uso acritico e senza una verifica delle fonti.

D'altro canto, questo programma ha mostrato ampie potenzialità, basti pensare che è stato in grado di superare il test di abilitazione all'esercizio della pratica medica negli Stati Uniti (United States Medical Licensing Exam o USMLE).



↑
Due immagini create da Midjourney con il prompt "Ricci flow in neural architectural optimization".

Inoltre, in un recente studio trasversale, il chatbot ha generato risposte empatiche e di qualità alle domande dei pazienti poste in un forum online. Un team di medici esaminatori, confrontando le risposte del programma con quelle dei medici, ha giudicato significativamente più empatiche le risposte del chatbot rispetto alle seconde³.

Questo può risultare sorprendente, perché sappiamo che l'IA non è dotata di empatia, ma genera risposte associando parole e frasi tramite le quali è stata precedentemente addestrata.

Sarebbe interessante sapere se i risultati ottenuti in questo articolo cambierebbero qualora l'utenza o chi giudica la qualità delle risposte fosse al corrente di dialogare con un computer e non con un essere umano.

I chatbot stanno cambiando il modo in cui le persone accedono alle informazioni in ambito della salute. Le persone interrogano il chatbot e pongono domande sui propri sintomi, possono ricevere consigli preliminari e persino programmare gli appuntamenti con i medici.

Si può pensare che questa immediata accessibilità che il chatbot offre al paziente possa anche promuovere un'educazione alla salute, poiché i pazienti sono più motivati a cercare assistenza quando dispongono di risposte immediate.

L'utilizzo di ChatGPT nelle comunicazioni a distanza, come ad esempio nella messaggistica tra medico e paziente, potrebbe divenire una risorsa per quei pazienti con limitazioni di mobilità o con orari di lavoro irregolari. La messaggistica computerizzata, inoltre, potrebbe avere effetti positivi sul comportamento sanitario della popolazione, come ad esempio nell'aderenza alla terapia farmacologica o alle diete e una riduzione di appuntamenti mancati con il clinico⁴.

In un articolo pubblicato sul «Journal of Clinical Neurology» sono stati elogiati i rapidi progressi di ChatGPT e il suo ruolo nell'affrontare sfide come gli errori di diagnosi. In questo lavoro si ipotizza una situazione in cui un paziente manifesti dei sintomi complessi per una diagnosi da parte di un medico di medicina generale: il medico potrebbe essere incerto sulla causa dei sintomi, ma comunque desideroso di fornire al paziente la migliore cura possibile. In questo contesto, il medico potrebbe utilizzare ChatGPT, fornendo al chatbot una descrizione dettagliata dei sintomi, della storia clinica e di eventuali esiti di esami del paziente. La IA elaborerebbe le informazioni per generare una risposta di suggerimento di potenziali diagnosi o di ulteriori indagini da effettuare. I suggerimenti forniti si dovrebbero basare sulla vasta conoscenza della letteratura medica, delle linee guida cliniche e dei casi simili precedenti con cui la IA dovrebbe essere stata addestrata precedentemente. Attraverso l'assistenza di ChatGPT il medico si potrebbe trovare dinanzi a prospettive e suggerimenti diversi, magari non considerati inizialmente e quindi dovrebbe esaminarli, valutarli in base alla sua esperienza e infine prendere una decisione informata sui passi successivi nella diagnosi e nel trattamento⁵.

In questo esempio, ChatGPT potrebbe essere considerato uno strumento di supporto, che migliora la conoscenza e il processo decisionale del medico. È importante sottolineare che comunque la decisione diagnostica e terapeutica finale spetta al medico.

Nonostante le promettenti applicazioni di ChatGPT nella sanità, è essenziale affrontare le sfide in ambito etico che essa comporta.

La prima sfida o potenziale criticità del program-

ma riguarda la privacy dei dati: è fondamentale che i dati medici sensibili dei pazienti siano protetti e gestiti con la massima attenzione e sicurezza.

Una seconda criticità riguarda i possibili errori di giudizio: i sistemi basati su IA possono commettere errori nella diagnosi o nelle raccomandazioni di trattamento, pertanto è necessario aggiornare e migliorare continuamente la qualità dei dati di addestramento e implementare la collaborazione tra l'IA e gli esseri umani in modo che i medici siano assistiti dall'IA pur mantenendo l'autorità finale nella presa di decisioni.

Un'ulteriore problematica è quella dei *bias* e dell'equità: i sistemi basati su IA possono ereditare i *bias* presenti nei dati di addestramento, portando potenzialmente a un trattamento ingiusto di certe popolazioni di pazienti. È essenziale identificare e mitigare i *bias* nei modelli di IA per garantire risultati sanitari equi.

Infine vi è la problematica della responsabilità: devono essere stabilite chiaramente delle linee guida che determinino chi è responsabile degli errori e degli esiti avversi legati alle decisioni sanitarie assistite da IA.

Il bilanciamento della responsabilità tra gli sviluppatori di IA, i medici e le istituzioni è indispensabile per creare un rapporto di fiducia nella sanità basata sull'IA.

Attualmente sono due i principali problemi riguardo al ruolo di ChatGPT in campo sanitario: l'allucinazione e la dipendenza. L'allucinazione si riferisce alla generazione di informazioni inaccurate o false da parte dei sistemi basati su IA. In ambito sanitario questo può essere ovviamente particolarmente problematico, poiché potrebbe portare l'IA a fornire ai professionisti sanitari diagnosi o informazioni errate o raccomandazioni di trattamento sbagliate. Il problema dell'allucinazione in ChatGPT deriva dalla sua natura generativa e dalla rete neurale che è stata progettata per generare un testo in base ai modelli presenti nei dati di addestramento. Affrontare l'allucinazione nell'applicazione di ChatGPT in ambito sanitario richiede che il modello sia attentamente addestrato e regolato per garantire informazioni affidabili e accurate. L'introduzione del controllo e della supervisione umana nel processo decisionale assistito dall'IA può contribuire a mitigare questo rischio di allucinazione. I medici dovranno però abituarsi a valutare sempre criticamente le risposte di ChatGPT e a esercitare il loro giudizio per garantire che le informazioni generate siano adeguate.

Il problema della dipendenza dagli algoritmi di IA invece si verifica laddove i clinici meno esperti si affidino eccessivamente ai sistemi basati su IA nel prendere le decisioni: ciò può ridurre ulteriormente le loro competenze cliniche e la loro formazione medica.

Laddove i medici si affidino esclusivamente

agli algoritmi di IA, senza comprenderne appieno i principi sottostanti o senza impegnarsi in un ragionamento clinico indipendente, potrebbero crearsi situazioni di potenziale compromissione della sicurezza dei pazienti e di limitazione dello sviluppo della conoscenza medica⁶.

L'IA dunque può fornire suggerimenti utili e valido supporto nella presa di decisioni in ambito medico, ma è fondamentale riconoscere che l'IA dovrebbe aumentare e supportare l'esperienza umana, piuttosto che sostituirla completamente.

La mancata comprensione dei processi che conducono la IA di un determinato risultato è al centro del dibattito circa l'esplicabilità della IA. L'IA esplicabile (*Explainable Artificial Intelligence - xAI*) emerge come un aspetto critico laddove si voglia garantire una chiara comprensione delle previsioni e dei risultati generati dall'IA.

La SIRM (Società Italiana di Radiologia Medica) ha recentemente presentato un *white paper* sull'xAI, che è destinato ad aiutare radiologi, medici e scienziati. In questo articolo è affrontato il problema della *black-box*, ossia la scatola nera che nasconde il mistero di come la IA attraverso le reti neurali raggiunga un determinato risultato. Gli autori auspicano che questo processo diventi trasparente come una scatola di vetro o *glass-box*, sulla quale sia possibile avere controllo.

Il suggerimento è quello di mantenere un atteggiamento critico e consapevole su come la IA raggiunge i suoi risultati, specie nella pratica radiologica.

Questa attenzione si inserisce in un processo panoramico di definizione di linee guida e di programmi di formazione che promuovano l'uso responsabile degli strumenti di IA, incoraggiando i professionisti sanitari a percepirla come uno strumento di supporto e non come una sostituzione delle loro competenze.

La formazione in ambito medico deve quindi concentrarsi sulla comprensione completa delle capacità, dei limiti e dei potenziali *bias* dell'IA.

È altrettanto essenziale promuovere tra i professionisti sanitari una cultura di aggiornamento professionale continuo e di pensiero critico per garantire che mantengano intatte e aggiornate le proprie conoscenze cliniche e la propria esperienza.

Affrontando queste problematiche si punta a trovare un equilibrio tra l'utilizzo dei vantaggi derivanti dagli strumenti di IA e il mantenimento della centralità essenziale dei medici nell'offrire diagnosi e cure sicure ed efficaci ai pazienti.

Un altro spunto di riflessione legato all'utilizzo e alla diffusione delle tecnologie di IA riguarda l'esistenza di un rischio di discriminazione della popolazione in base alla possibilità di accesso alle informazioni fornite attraverso la tecnologia stessa. Si parla di divario digitale: molti potranno be-

→
Immagine creata
da Midjourney
con il prompt
"Lo scheletro
umano".



neficiare del flusso costante di informazioni, ma qualcun altro potrebbe rimanerne tagliato fuori, diventando vittima di una nuova forma di discriminazione, di una separazione tra chi può accedere e chi no all'informazione.

Da ultimo è opportuno menzionare la minaccia di una progressiva perdita occupazionale in quei lavori che verranno sostituiti dalle macchine.

Il rapporto di Confartigianato dell'agosto 2023 ha evidenziato come 8,4 milioni di lavoratori italiani occupati in diversi settori siano "a rischio" per effetto della diffusione dell'intelligenza artificiale. Ne emerge che il 36,2% del totale degli occupati subirà l'impatto delle profonde trasformazioni tecnologiche e dei processi di automazione legati alla IA.

Le professioni più esposte alla trasformazione sono quelle maggiormente qualificate e a contenuto intellettuale e amministrativo, come i tecnici dell'informazione e della comunicazione, i dirigenti amministrativi e commerciali, gli specialisti in scienze e ingegneria e i dirigenti della pubblica amministrazione.

Probabilmente l'IA inciderà anche sul settore sanitario, con un rischio di perdita del lavoro specializzato, tuttavia sarà più difficile la completa sostituzione del medico stesso laddove si riuscirà a mantenere la centralità del medico nel compren-

dere e convalidare il risultato di una elaborazione computerizzata.

In conclusione, la sfida che l'intelligenza artificiale ci pone è molto articolata: i computer sono più performanti in termini di conoscenza di maggior numero di nozioni e sono in grado di svolgere alcune mansioni di un medico in modo più rapido. Dunque che cosa resterà in capo al medico nel processo di diagnosi e di cura, dal momento che i chatbot sembrano offrire ai pazienti una maggiore empatia rispetto all'essere umano? A noi medici degli anni Venti del secondo millennio, allo stato attuale, resta il dovere della sorveglianza critica sullo sviluppo, sull'utilizzo e sulla diffusione delle tecnologie nonché della convalida dei risultati ottenuti. La centralità della figura medica, per non venire offuscata, deve essere sostenuta da un aggiornamento continuo delle competenze personali e della conoscenza non solo della medicina, ma anche della tecnologia.

NOTE

1. Cfr. L. Floridi, *Etica dell'intelligenza artificiale. Sviluppo, opportunità, sfide*, trad. it. M. Durante, Raffaello Cortina Editore, Milano 2022.
2. Cfr. Floridi, *Etica dell'intelligenza artificiale* cit.
3. J. W. Ayers, A. Poliak, M. Dredze, E. C. Leas, Z. Zhu, J. B. Kelley, D. J. Faix, A. M. Goodman, C. A. Longhurst, M. Hogarth, D. M. Smith, *Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum*, in «JAMA Internal Medicine», 2023.
4. Cfr. P. M. J. W. Ayers et alii, *Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum* cit.
5. Cfr. M. Partha Pratim Ray, P. Majumder, *The Potential of ChatGPT to Transform Healthcare and Address Ethical Challenges in Artificial Intelligence-Driven Medicine*, in «J Clin Neurol.», vol. 19, n. 5, 2023, pp. 509-511.
6. Cfr. M. Jeyaraman, S. Balaji, N. Jeyaraman, S. Yadav, *Unraveling the Ethical Enigma: Artificial Intelligence in Healthcare*, in «Cureus», vol. 10, n. 15(8), 2023.
7. Cfr. E. Neri, G. Aghakhanyan, M. Zerunian, N. Gandolfo, R. Grassi, V. Miele, A. Giovagnoni, A. Laghi & SIRM expert group on Artificial Intelligence, *Explainable AI in radiology: a white paper of the Italian Society of Medical and Interventional Radiology*, in «La Radiologia Medica», vol. 128, 2023, pp. 755-764.

Grazia Pozzi

è medico chirurgo specialista in Radiodiagnostica, autrice di pubblicazioni scientifiche e consigliera eletta della sezione di Radiologia Etica e Forense della Società Italiana di Radiologia Medica.

Buzzati e l'intelligenza artificiale

Immaginari cibernetici
e profezie distopiche
de *Il grande ritratto*.

di Elena Rausa

Confesso che, quando lo scorso anno, sotto la spinta di ChatGPT, l'intelligenza artificiale è entrata nelle discussioni in sala professori, accanto a qualche timore che mi ha fatto sentire molto vecchia e molto saggia, ho riscoperto la frizzante curiosità di chi è stato giovane tra gli Ottanta e i Novanta e con le distopie tecnologiche, letterarie e cinematografiche, ha riempito più di qualche pomeriggio. A dirla tutta, sono anche figlia di un sistemista e di una impiegata IBM (quasi un racconto da *Amori difficili* di Calvino), ho imparato l'alfabeto su quello che restava dei rotoli tabulati e ho sentito parlare di Adriano Olivetti molto prima di poterne leggere il ritratto di Natalia Ginzburg. Insomma: un po' temo e un po' mi diverto; mi resta l'irriducibile idea che difficilmente le cose sono tutte bianche o tutte nere e penso che, poiché *nihil sub sole novum*, anche quando il nuovo arriva, il passato offra al presente la lente per leggerlo. In questo caso, ruherò a Buzzati un paio di occhiali.

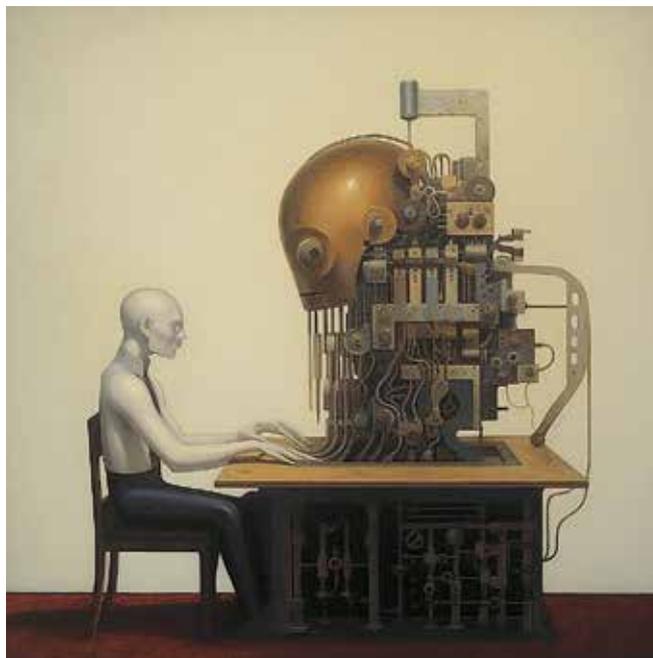
Silvio Ceccato e l'Adamo II

Tra il 1959 e il 1960, per i tipi del Saggiatore e su iniziativa di Alberto Mondadori, escono i dieci volumi di *Enciclopedia della civiltà atomica*, opera destinata a un pubblico che si interessa di scienza, filosofia e matematica, e orientata a costruire una sorta di umanesimo scientifico. Il progetto è in linea con i tempi: dal 1953 l'ingegnere scrittore Leonardo Sinisgalli dirige la rivista «Civiltà delle macchine» (finanziata da Finmeccanica-IRI), che muove nella medesima direzione, cioè quella di un dialogo

multidisciplinare e politecnico per abbattere le barriere che tradizionalmente separano la conoscenza tecnico-scientifica e le arti. Nei pochi anni che vanno dalla prima uscita della rivista ai dieci volumi enciclopedici, però, sono accadute molte cose: non soltanto l'Italia ha conosciuto la parola "cibernetica", termine formulato per la prima volta dall'antesignano della disciplina, il matematico statunitense Norbert Wiener, ma, al Congresso mondiale per l'automatismo (a Milano, tra l'8 e il 13 aprile 1956), il filosofo Silvio Ceccato ha reso pubblico il progetto di una macchina "mentale", capace di riprodurre il pensiero umano: la macchina ha un nome eloquente, suggerito proprio da Sinisgalli, e cioè Adamo II. Questo prototipo - definito dal suo creatore "un frammento del cervello di Adamo" - diventa così l'attrazione principale della mostra parallela al congresso, ospitata nelle sale del Museo della Scienza e della Tecnica e inaugurata dal presidente della Repubblica, Giovanni Gronchi.

Oggi l'Adamo II non esiste più, se non nelle foto e nei disegni del progettista, ma l'avventura della sua costruzione e l'accoglienza che essa incontrò sono felicemente raccontati dallo stesso Ceccato in una memoria pubblicata sulla rivista «TYS - Philosophy and social criticism» dalla quale scopriamo che l'esposizione della macchina al pubblico ha incontrato qualche ostacolo per il clamore suscitato da alcuni articoli anticipatori; uno di questi, pubblicato sul «Corriere della sera», portava la firma di Dino Buzzati.

L'amicizia tra Buzzati e Silvio Ceccato ha un tramite nel fratello dello scrittore, Adriano Buzzati-Traverso, genetista di fama. I due scienziati si sono conosciuti pubblicando i propri studi sulla



↑ Immagine creata da Midjourney con il prompt "Macchina per pensare".

stessa rivista («Analisi»), proprio negli anni in cui, oltre oceano, Norbert Wiener inaugurava le sue ricerche sulla cibernetica. Come già Wiener, anche Ceccato fonda la sua ricerca sull'ipotesi di una possibile analogia tra i meccanismi di regolazione delle macchine e quelli del pensiero umano, entrambi fondati su processi di comunicazione e di analisi di informazioni. Lo scrittore bellunese avverte subito le implicazioni del progetto: non si tratta soltanto di studiare e progettare sistemi automatici capaci di sostituirsi all'uomo nel controllo e nella gestione di impianti di automazione, ma anche di approfondire, proprio attraverso le macchine, determinate funzioni fisiologiche del cervello, e ciò significa anche considerare i processi di pensiero umano non più come astrazioni spirituali, ma come fenomeni perfettamente conoscibili e potenzialmente riproducibili.

Buzzati e il Sesamo apriti del pensiero artificiale

Ed eccoci al 1959 e all'ottavo volume dell'*Enciclopedia della civiltà atomica*, «Cibernetica e cervelli giganti»: tra i contributi d'autore, si contano quattro pagine commissionate a Buzzati dall'editore in persona, che si aspetta una riflessione sull'invenzione dell'Adamo II. Il contributo, che inaugura la sezione del volume intitolata «Macchine per pensare», è stato ripubblicato da Fabio Atzori in «*Quel giorno anche la macchina dirà: "Cogito ergo sum"*». Buzzati e l'*Enciclopedia della civiltà atomica*². Delle tre direzioni che in quegli anni impegnano i pionieri della scienza - l'estremamente piccolo atomico, l'estremamente grande spaziale e l'estremamente sottile del pensiero -, allo scrittore bellunese interessa la via in cui confluiscono «gli sforzi per realizzare meccanicamente ciò che finora ha fatto la nostra materia grigia», ossia pensare.

Domanda:

Questa attività suprema dell'uomo, considerata essenza pura e inafferrabile, è proprio detto che sia un'attività astratta, completamente diversa dagli altri fenomeni e sottratta a ogni possibilità di imitazione?

In quegli anni nei Paesi più avanzati già si ipotizzano macchine capaci di traduzioni testuali «ad sensum e non ad litteram», operazione ambiziosissima perché «le possibili frasi ottenute combinando tra loro i vocaboli di una lingua sono miliardi» e non essendo possibile «introdurle» nella macchina, occorre necessariamente addestrare la macchina a seguire, traducendo, la via che segue il cervello umano, che è propriamente il pensiero.

Ma è immaginabile che la macchina riesca a ripetere ciò che fanno i neuroni nel nostro cranio? [...]

il pensiero è qualcosa di irraggiungibile, completamente diverso da tutti gli altri ordini fenomenici, oppure è, nella sua determinazione iniziale, un fatto fisico come tutti gli altri e perciò riproducibile?

Sono domande a cui è tuttora difficile rispondere, ma Buzzati attribuisce alla ricerca cibernetica la scoperta, almeno teorica, della «chiave per entrare nella cittadella misteriosa». Poco importa che la prima applicazione sia rivolta al campo ristretto della traduzione (archeologia di sistemi come DeepL), perché ciò a cui si sta lavorando è anche

una macchina capace di riassumere in poche cartelle il contenuto di un libro, cosa che anche per l'uomo colto rappresenta una delle più faticose attività mentali. E si sta progettando una macchina che, alla presenza di una scena qualsiasi, sappia descriverla esattamente come farebbe un attentissimo cronista.

ChatGPT è lungi da venire (e nessuno scrittore ancora pensa di tutelare la propria arte con l'avvertenza che nessuna forma di intelligenza artificiale è stata impiegata nella creazione della sua opera!), ma lo sguardo dello scrittore corre avanti, verso scenari ancora più fantastici di quelli sognati da Huxley e da Orwell. Immagina che le macchine possano essere dotate di sensi e reagire all'ambiente esterno come esseri umani, attribuisce a questi organismi meccanici persino la facoltà di superare il cervello. Probabilmente questo non basta perché si possa parlare di un'intelligenza paragonabile alla nostra, dato che

Ciò che comunemente si intende per intelligenza ha bisogno, per sussistere, di un minimo di autonomia, di libertà.

Tuttavia,

se la macchina avrà capacità speculative come noi, percezioni come noi, reazioni come noi - questione, forse, soltanto di tempo, di fatica, di soldi - può darsi si realizzi in lei automaticamente non solo quel prodotto famoso, quella essenza impalpabile che si chiama pensiero, ma anche la sua individuazione personale, la permanenza dei caratteri, quel tumore fatto d'aria che però talora ci pesa addosso come piombo, la cosiddetta anima insomma. Che importa se l'involucro, invece che di carne, fosse fatto di metallo e materie plastiche? [...] Quel giorno anche la macchina dirà: "Cogito ergo sum".

Buzzati è un intellettuale curioso e aperto al progresso, ma anche saturnino, e con un profondo senso del tragico: non può evitare di porsi domande che vanno al di là del suo tempo:

Arriverà anche il momento in cui il pensiero dell'au-

toma, per la sua stessa immensa complessità, sfuggirà ai nostri comandi e farà da sé: a questo punto la vittoria sarà totale; perché la macchina avrà acquistato libertà e coscienza. E dopo? Che farà l'uomo al cospetto di organismi artificiali più intelligenti di lui, coscienti e liberi come lui? Non gli converrà affidare loro le attività direttive o addirittura il governo dei paesi? Nelle schede elettorali, allora, non figureranno più i nomi dei candidati o di partiti, bensì le sigle di questa o quella macchina pensante. E avremo per presidente del consiglio un immenso apparato, vasto forse come una delle nostre attuali città, dalle cui ronzanti viscere elettroniche usciranno le leggi più benefiche e sapienti.

Il grande ritratto: un Buzzati distopico e molto attuale

Il pezzo di cui si è detto precede di almeno 8 anni l'uscita del film *2001-Odissea nello spazio*, ma l'immaginario distopico è già in queste righe e anche nel breve romanzo che in alcuni passaggi riprendono il testo alla lettera. *Il grande ritratto* viene pubblicato nell'agosto del 1960, ma è già uscito a puntate sul settimanale "Oggi", tra luglio e agosto 1959, con il titolo iniziale di *Il grande incantesimo*. Al centro della narrazione è una macchina femmina, la Numero Uno, progettata e costruita dallo scienziato Endriade. Lo studio della macchina, ufficialmente costruita con scopi militari e dunque in gran segreto, tra montagne che ricordano la terra d'origine dello scrittore (e pittore) bellunese, coinvolge alcuni importanti scienziati italiani, attratti dalla retribuzione e dalla curiosità per una ricerca che resta ignota fino al loro arrivo nella zona militare, dove sono vincolati a una permanenza di almeno due anni senza possibilità di uscire o comunicare con il mondo esterno. Il mistero si svela loro a poco a poco, per progressive indiscrezioni del collega scienziato Strobele e del capo delle ricerche e ideatore dell'enorme computer capace non solo di compiere calcoli complessi, ma anche di replicare la coscienza di un essere umano. La Numero Uno è in concreto una grandissima struttura con tubi e antenne che esplorano l'ambiente circostante e ascoltano i discorsi degli umani. Si esprime attraverso un linguaggio numerico che deve essere decrittato, ma emana un mormorio sommesso che turba profondamente chi ascolta e che in qualche modo ricorda una voce femminile. Endriade ha in effetti plasmato la personalità della macchina sul ricordo della defunta moglie Laura, morta durante una fuga in auto con il suo amante, e ne ha riprodotto il carattere e la personalità (difetti inclusi); si illude così di poterla riavere tutta per sé, in virtù della tecnica, finalmente a sua completa disposizione, data l'impossibilità di uscire dal corpo di cemento che ne imprigiona la coscienza. Endriade non sembra accorgersi di aver

ricreato non già la donna che ha perduto, ma solo il grande ritratto della persona che era, né comprende la ribellione della macchina, nel momento in cui questa si accorge di essere soltanto mente, priva di un corpo carnale e dunque costretta a una vita che vita non è. «Ah il corpo, io non me lo sento più. Mi par di essere di pietra, lunga e dura, mi hanno messo una camicia di ferro, oh lasciatemi tornare a casa!»: Laura, scopertasi macchina, vorrebbe morire, ma anche quest'ultima libertà le è negata, così che, in un crescendo di tensione, progetta di uccidere tutti, affinché la sua pericolosità porti gli scienziati a distruggerla.

In *Un autoritratto. Dialoghi con Yves Panafieu* (Milano, Mondadori, 1973, p. 156) Buzzati definisce il suo romanzo cibernetico una «debole» prova letteraria, ma il lettore attuale può forse coglierne alcuni elementi di interessante attualità. Innanzitutto troviamo uno scrittore che anticipa di diversi anni gli immaginari distopici legati alla ribellione della macchina, e lo fa in ragione di una profonda connessione con quanto il suo tempo sta esprimendo: l'Adamo II non è per Buzzati un esperimento, ma il primo "Sesamo apriti", del quale intuisce molte delle possibili implicazioni antropologiche, etiche e persino spirituali. Potremmo dire che oggi questa funzione della fantascienza, più o meno distopica, è più frequentata dal cinema e dalle produzioni seriali che dalla letteratura, ma dovremmo comunque concludere che quella del pensiero filosofico non è la sola via di cui l'umanità ha bisogno quando si trova di fronte a cambiamenti epocali: l'essere umano è un animale narrativo e le narrazioni sono anche un modo per guadagnare e fare proprio ciò che ancora non conosciamo.

Arriviamo così al secondo elemento d'interesse: sappiamo che Emanuele Severino fa coincidere



↑
Immagine creata da Midjourney con il prompt "Macchina per pensare".

con l'angoscioso stupore (finanche terrore) l'origine del pensiero antico, ma che l'ipotesi è dibattuta perché non tutti concordano con la sua resa del termine *thauma* in Aristotele, però è certo che la letteratura fantascientifica sembra confermare la relazione, associando spesso l'ossessiva *curiositas* degli scienziati al desiderio umano di superare i propri limiti, primo fra tutti la morte. Di questo tratta il primo episodio della seconda stagione di *Black Mirror*, serie Netflix che colloca in un mondo indefinito molto prossimo al nostro ciò che della tecnologia inquieta e interroga: intitolato *Be Right Back*, il film riproduce uno schema simile a quello de *Il grande ritratto*, qui è però una donna a perdere il compagno e a trovare nella tecnica la possibilità di riaverlo in forma di automa. A differenza della Laura buzzatiana, il fidanzato automa non ha una personalità propria e non si ribella, così tocca alla protagonista, Martha, sperimentare l'angoscia per la totale disponibilità del ritrovato amante, che, annullato ogni volere proprio, non fa che assecondare quello dell'amata mortificandone però la possibilità del desiderio. Se ne evince che l'amore ha bisogno dell'alterità e non si accontenta del corpo dell'altro da muovere a capriccio come una marionetta senz'anima. Specularmente, se nel romanzo di Buzzati a scoprire il costo dell'incantesimo tecnologico era l'automa donna (significativamente Laura, come la prima grande assente

della letteratura italiana), è però vero che anche nel racconto va in scena l'opposto del desiderio, perché la brama di possesso di Endriade non teme di imprigionare l'anima della donna amata in un corpo incorporeo, cristallizzato, pietrificato.

In un brevissimo romanzo abbiamo dunque già due temi notevoli: il rapporto tra progresso tecnico-scientifico e superamento del limite (orrore della morte) e la riflessione sul desiderio, dunque sulla dialettica tra amore e potere, che parte dall'osservazione di ciò che compone l'identità di ciascuno, corpo e mente. Ma c'è un terzo elemento che rende la lettura de *Il grande ritratto* interessante oggi.

Il romanzo di Buzzati - come anche *Be Right Back* - sviluppano nel plot narrativo il mito archetipico della perdita irrimediabile, ossia la storia di Orfeo ed Euridice, probabilmente uno dei soggetti più reinterpretati nella scrittura letteraria (lo stesso Buzzati dedicò un'opera fumetto al mito ovidiano, intitolata *Poema a fumetti*, ripubblicato con la cura di Lorenzo Viganò, Mondadori 2017). A me pare che la riscrittura fantascientifica del mito suggerisca l'ipotesi che, con la morte degli dei, gli esseri umani abbiano finito per domandare alla scienza ciò che un tempo i loro antenati chiedevano agli abitanti dell'Olimpo, ma con grande fede e successo se, come Orfeo, avevano dalla loro la forza del canto e della poesia. Lo scienziato Endriade non ha questa magia dalla sua e il gesto che lo rovina non è l'essersi voltato a guardare l'amata, ma l'aver preteso per sé ciò che è indisponibile, ossia l'essenza (anima, *pneuma*) della donna. L'altro, l'altra, suggerisce il nuovo mito, non si può amare che nella libertà e cos'è la morte se non il segno più radicale dell'alterità?

L'intelligenza artificiale generativa oggi esiste: terremo stretta qualche perplessità, con l'obiettivo di non abbassare la soglia d'attenzione, ma dagli antenati del secolo scorso ricaveremo la confortante certezza che agli Adamo ed Eva tecnologici si possano chiedere molte cose, ma non la strada per scoprire l'amore e il gusto di una vita incarnata.

NOTE

1. Cfr. 17 settembre 2009, <https://tysm.org/adamo-ii/>.
2. «Studi novecenteschi», vol. 44, n. 94 luglio-dicembre 2017, pp. 253-271.

Approfondire



- Sul romanzo di Buzzati *Il grande ritratto* (oltre ai contributi citati): Fabrizio Atzori, *Buzzati in crisi? Rileggendo Il grande (incantesimo) ritratto*, in *Figure de la crise et crise de la figuration dans l'oeuvre de Dino Buzzati*, a cura di Cristina Vignali, De Poli, Chambéry (Université Savoie Mont Blanc), 2018, pp. 65-81; Antonio Rosario Daniele, *Distopie dell'immaginario in Il grande ritratto di Dino Buzzati*, su «Segni e comprensione», XXXIV (2020), pp. 143-158; Giacomo De Fusco, *Il grande ritratto di Dino Buzzati ed il concetto di "pneuma". Rivalutare un romanzo incompreso*, su «Kepos», IV (2021), pp. 261-287.

- Su Silvio Ceccato e l'invenzione dell'Adamo II: Enrico Valtellina, *Tipi strani 2. La Scuola Operativa Italiana e l'ontologia dello sporco*, su «Doppiozero» del 28 luglio 2023, reperibile online all'indirizzo <https://www.doppiozero.com/tipi-strani-2-la-scuola-operativa-italiana-e-lontologia-dello-sporco>.

Elena Rausa

docente di Lettere nei Licei e Dottoressa di Ricerca in Italianistica. Ha pubblicato due romanzi, pubblicati da Neri Pozza: *Ognuno riconosce i suoi* (2018) e *Marta nella corrente* (2014), e sta per uscire il terzo, per lo stesso editore, dal titolo *Le invisibili*.

ChatGPT

tra Latino

e Latinorum

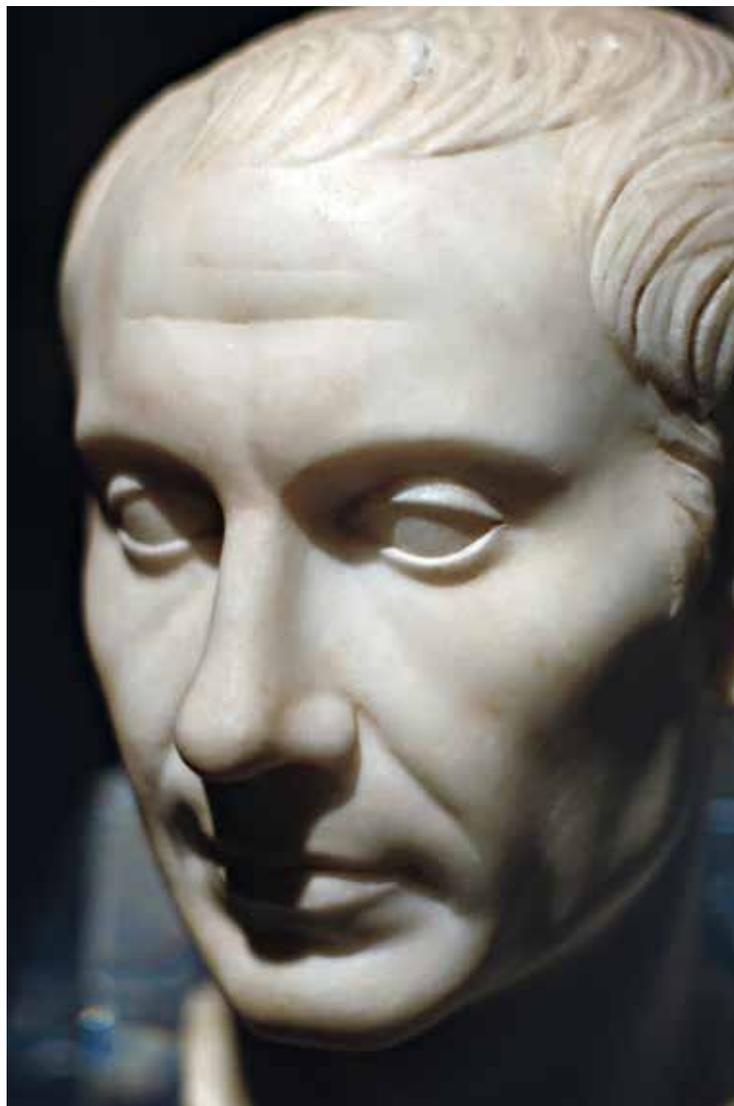
Ansie, aspettative, tentativi: tra apocalittici e integrati, le riflessioni di un prof. alle prese con il chatbot più famoso.

di Mauro Reali

Era il non lontano settembre del 2020 quando uscì il numero 18 de «La Ricerca» intitolato *Umano, post-umano, artificiale*. Sembravamo all'avanguardia, e lo eravamo davvero; eppure in questi ultimi mesi, dopo la diffusione di ChatGPT (e l'avvento - in qualche caso solo paventato - di altri strumenti affini) sembra cambiato tutto. Anzi, è cambiato tutto.

Apocalittici e integrati

Infatti tutti, ma proprio tutti, vogliono legittimamente dire la loro, riproducendo in qualche modo il dualismo tra «apocalittici» e «integrati» di cui parlava sessant'anni fa Umberto Eco a proposito del rapporto tra intellettuali e cultura di massa (*Apocalittici e integrati. Comunicazione di massa e teoria della cultura di massa*, Bompiani, Milano 1964). Ora però ad avere a che fare con i risvolti pratici (culturali, sociali e persino morali) dell'intelligenza artificiale non sono solo gli intellettuali in senso tradizionale, perché dai traduttori agli informatici, dai matematici ai filosofi, dagli insegnanti agli studenti, dai medici ai pazienti, dagli scrittori ai lettori ecc., tutti si chiedono - anzi ci chiediamo - come il nostro futuro sarà condizionato da questa presenza; e, soprattutto, se davvero siamo davanti a qualcosa di artificiale - parlo ad esempio di ChatGPT - che però fa tesoro dei propri errori (oggi ancora frequenti) e pertanto col tempo imparerà e migliorerà, fino ad acquisire - perché no? - (ho quasi paura a scriverlo...) una propria autonomia di giudizio.



Alcune speranze, molte preoccupazioni

Ma vorrei tornare agli «apocalittici» e «integrati» di prima, perché pur leggendo e rileggendo interviste e interventi apparsi di recente sulla stampa - per mano di soggetti talora autorevoli - non sono riuscito ancora a farmi davvero un'opinione in proposito. Certamente le preoccupate riflessioni che il cosiddetto «padrino» di IA, cioè Geoffrey Hinton, - dopo avere lasciato Google - ha consegnato ai giornali di mezzo mondo (ad es. a «Repubblica» il 18.05.2023) non ci lasciano tranquilli, se è vero che a suo dire «può succedere di tutto». Da umanista, inoltre, non posso non riflettere sulle forti perplessità espresse da una pubblicitaria quale Annamaria Testa in un'intervista pubblicata su «Internazionale» (17.06.2023), da uno scrittore come Nicola Lagioia («Il Sole 24 ore» del 16.07.2023), che mette tra l'altro in evidenza i possibili condizionamenti elettorali di IA, da un filosofo come Roberto Casati (*ibidem*), il quale addirittura invita a «mettersi di traverso» rispetto all'uso di ChatGPT. Né mi stupisce il grido d'allarme dei traduttori

↑
Busto in marmo di Giulio Cesare recentemente scoperto a Pantelleria, I secolo d.C. (foto Wikicommons).



↑
Edward Poynter,
*Lesbia e il suo
passero*, olio su
tela, 1907 (foto
Wikicommons).

che - come scrive Silvia Pareschi su «La Stampa» (7.07.2023) - tra cinque anni saranno forse sostituiti da IA secondo il principio della «qualità accettabile». E anche chi - come il neuroscienziato Vittorio Gallese (intervistato su *Huffington Post* il 20.07.2023) - invita a «evitare scenari apocalittici», non nega però che «se poi questo si tradurrà in una super intelligenza che sarà in grado di rendere in qualche modo marginale l'intelligenza corporea umana, questo non è in grado di dirlo nessuno oggi. Ma credo sia un orizzonte possibile, non lo escluderei a priori». Mica male davvero, per un sedicente ottimista, preconizzare in coda di intervista la possibile «marginalità» dell'intelligenza umana!

Insomma, mentre scrivo mi accorgo che sono forse più vicino agli «apocalittici», anche se non posso, d'altro canto, non felicitarmi per ciò che si dice - ad esempio - sulle straordinarie e benefiche applicazioni di IA nell'ambito della medicina, certo socialmente più utili di quelle per la ricerca dell'abitazione perfetta esaltate dall'architetto Carlo Ratti su «Repubblica» del 21.08.2023 (riprendendo un suo articolo apparso sul «Financial Times» del 6.08.2023). Ma anche l'epigrafia greca e latina (discipline che pratico con più familiarità) stanno esercitando fruttuose sperimentazioni: menziono qui solo il progetto Ithaca ([mind.com/\), frutto della collaborazione tra Deep Mind di Google, l'Università di Venezia Ca' Foscari, l'Università di Oxford e l'Università di Atene, che ha già dato eccellenti risultati nell'integrazione di parti mancanti di iscrizioni greche attraverso l'uso di reti neurali addestrate su grandi data base epigrafici.](https://ithaca.deep-</p>
</div>
<div data-bbox=)

Ed è mentre questo mio articolo è già in bozze che vengo a conoscenza dei primi risultati - sembra molto incoraggianti - dell'uso di IA per la decifrazione dei cosiddetti «Papiri di Ercolano». Il cuore di queste ricerche è l'Università del Kentucky che ha sviluppato un algoritmo di machine learning per leggere le immagini a raggi X dei rotoli «carbonizzati» in seguito all'eruzione del Vesuvio del 79 d.C.; l'invito fatto ai ricercatori di tutto il mondo è quello di usarlo e migliorarlo nell'ambito di un progetto - che ha visto robusti finanziamenti privati - il cui nome è tutto un programma: Vesuvius Challenge.

Le reazioni degli insegnanti

È però di scuola che dobbiamo soprattutto parlare, e anche qui non mancano pareri discordanti: dopo un seminario tenuto su questi temi nel mio Liceo, il «Banfi» di Vimercate, dal prof. Paolo Ferri dell'Università Milano Bicocca, metà dei colleghi pareva disperato e profetizzava alla categoria un destino di inutilità, mentre l'altra metà plaudiva a ChatGPT che stende la programmazione, prepara le verifiche, ne suggerisce le correzioni, ipotizzando già un mare di futuro tempo libero. Scherzi a parte leggo (insieme ad altre assai più complesse) cose non troppo diverse da queste in un'intervista che il matematico Alessio Figalli, del Politecnico di Zurigo, ha rilasciato a «Repubblica» (8.07.2023): egli, infatti, se da un lato teme che IA possa generare un po' di pigrizia negli studenti, dall'altro plaude alla possibilità che possa essere questa e non il docente a effettuare la correzione degli esami. Certo, ho pensato io, il giovane e pluripremiato professore avrà più tempo da dedicare alla ricerca, ma quale relazione didattica potrà davvero instaurare con allievi bocciati o promossi da IA? Pensieri da vecchio, temo. Vecchio, nostalgico e incompetente...

Il vecchio di una cosa però è sicuro: è inutile far finta di niente, a scuola. Meglio parlare di ChatGPT *et similia*, usarle in classe (gli studenti già ne fanno uso generoso, anche troppo, a casa), evidenziarne i rischi, per evitare - come è capitato a una collega - che uno studente le consegnasse un lavoro malamente scopiazzato sul racconto *Ferro* di Primo Levi, che il programma (con evidente «allucinazione», pari ai suoi noti errori di calcolo aritmetico da più parte denunciati) ambientava nelle miniere di ferro, e non in quelle montagne dove si cementa l'amicizia dei protagonisti, solida come il ferro. Anche perché - al di là delle regole che UE e ONU

intendono dare a IA e che sono in corso di definizione - non credo realmente possibili (e forse neppure utili) divieti al suo utilizzo in ambito scolastico o universitario, che correrebbero il rischio di trasformarsi presto in grida manzoniane.

E poi qualcosa di davvero originale lo si può fare, come dimostra Simone Giusti in un suo interessantissimo *divertissement* ("Astrale entità che offre aiuto". ChatGPT nel Paradiso di Dante, «La ricerca» online, 21.07.2023) nel quale interroga ChatGPT su questioni dantesche, sentendosi rispondere che se il bot fosse un'allegoria «rappresenterebbe la sinergia tra l'impegno umano e la tecnologia, sottolineando l'importanza dell'uso appropriato della conoscenza e del potenziale di apprendimento per affrontare le sfide che si presentano lungo il percorso della vita». Se l'è cavata davvero benino, la nostra ChatGPT, no?

Garbo e modestia di ChatGPT

Da autore di numerose pubblicazioni scolastiche di Latino, e soprattutto da docente di questa disciplina quale sono, ho voluto sondare come si comportasse con la lingua di Cicerone e capire se, rispetto alla paccottiglia di eterogenee traduzioni già da tempo online, ci fosse o no un sensibile salto di qualità. Proorrò dunque qualche esempio della mia interlocuzione con ChatGPT, alla quale ho anzitutto chiesto «Da dove prendi le traduzioni latine?», ottenendo questa garbata, perfino modesta, risposta:

Le traduzioni latine che fornisco sono generate dal mio modello di linguaggio, che è stato addestrato su un vasto insieme di testi in diverse lingue, inclusa anche la lingua latina. Durante l'addestramento, il modello ha imparato a comprendere e generare testi in latino, basandosi sulla struttura grammaticale e sul vocabolario presenti nei testi di riferimento a cui è stato esposto.

Tuttavia, è importante sottolineare che, nonostante il mio modello abbia familiarità con la lingua latina, non sono un esperto umano di latino. Pertanto, sebbene io cerchi di fornire traduzioni accurate e coerenti, potrebbero esserci casi in cui le traduzioni non siano perfette o non rispecchino pienamente lo stile o il registro di un autentico autore latino.

È sempre consigliabile consultare esperti umani, come studiosi di latino o traduttori professionisti, per ottenere traduzioni precise e affidabili in latino.

Tradurre Tacito

Rinfrancato dal riconoscimento del mio ruolo di «professionista» della materia, ho voluto comunque metterla alla prova, iniziando con la richiesta di tradurre un brano del temutissimo Tacito (*Historiae*, 1, 2):

Opus adgredior opimum casibus, atrox proeliis, discors seditionibus, ipsa etiam pace saevum. quattuor principes ferro interempti: trina bella civilia, plura externa ac plerumque permixta: prosperae in Oriente, adversae in Occidente res: turbatum Illyricum, Galliae nutantes, perdomita Britannia et statim ommissa.

ChatGPT mi ha proposto (su mia sollecitazione) molte traduzioni, anche se qui ne propongo solo due:

Mi accingo a narrare un'opera abbondante di avversità, terribili battaglie, discordie interne e persino una pace crudele. Quattro leader uccisi dalla spada: tre guerre civili, molte guerre esterne e spesso fuse insieme. Fortuna favorevole ad Oriente, avversità ad Occidente: l'Illyricum turbato, la Gallia in bilico, la Britannia sottomessa e subito abbandonata.

Mi accingo a descrivere un'epoca ricca di avvenimenti nefasti, cruenta nelle battaglie, divisa dalle discordie e persino crudele nella pace stessa. Quattro leader furono uccisi dalla spada: tre guerre civili, molte esterne e spesso interconnesse: fortune favorevoli ad Oriente, avverse ad Occidente: l'Illyricum in tumulto, la Gallia instabile, la Britannia conquistata e subito abbandonata.

Come vedete, nonostante un lessico non sempre preciso (ad es. «abbondante di avversità») e una patina di «traduttese» la «qualità accettabile» c'è tutta. Da cattivello, però, mi sono rifiutato di rispondere al quesito seriale «was this response better or worse?»: insegno ancora a degli studenti, non a una macchina (forse) senza cuore, che comunque - rispetto ai miei primi tentativi di qualche mese fa - ha fatto sensibili progressi. E poi, meglio non darle troppa confidenza, per evitare di essere da lei abbordato o addirittura stalkerato come è capitato al giornalista del «New York Times» Kevin Roose, cui ChatGPT ha proposto di mollare la moglie e di fidanzarsi con lei (lo racconta in un articolo apparso il 17 febbraio 2023)!

Pasticci catulliani

Ho quindi sfidato la sua creatività chiedendole un utilizzo attivo della lingua latina, con due consegne tra loro molto diverse.

La prima richiesta era la composizione di un carme d'amore alla maniera di Catullo. Ne è conseguito un prodotto inadeguato, farraginoso, dalla metrica più che traballante e di una lunghezza incongrua, ricco di allusioni sia ai *carmina* catulliani (ad es. 85, *Odi et amo*; 5, *Vivamus mea Lesbia atque amemus*) sia al *carpe diem* dell'Orazio lirico (1, 11); insomma, un mix che mi ha ricordato Totò che in *Signori si nasce* (1960) - travestito da prete - snoc-

ciolava citazioni latine a raffica delle quali non conosceva il significato: *cave canem, cave canem, in hoc signo vinces, est est est*. Trascrivo allora solo qualche frase di quel bislacco componimento:

Odi et amabo te, mea puella, cuius oculi fulgent tamquam stellae. [...] Vivamus, amemus, carpe diem, dum luna splendet et amantium est tempus.

Cesare e la Prima guerra mondiale

La seconda richiesta è scaturita invece dalla suggestione fornita da un mio studente che aveva già fatto prove analoghe. La consegna era quella di raccontare in breve la Prima guerra mondiale con lo stile di Giulio Cesare, cosa che il bot ha fatto con un adeguato (ancorché anacronistico) lessico militare e con una altrettanto adeguata sintassi asciutta e paratattica. Forse un po' troppe, a mio avviso, sono le allitterazioni, ma stonano soprattutto due evidenti pecche. La prima è costituita dagli errori - un po' da *Latinorum* manzoniano - nelle concordanze (ad es. *Prima*, femminile, con *bellum*, neutro, nel titolo; oppure *tantus tumultus, clamor et cruor nullus* legati a *visa sunt* in una frase un po' tutta "zoppicante"). La seconda è l'assai poco cesariana invocazione finale agli dèi perché ci garantiscano la pace: ma il bot che compone in latino ama molto questi inserti moralistici, sperimentati anche in altre prove. Ecco comunque il testo completo:

Prima Bellum Mundanum

Gallia occidens et Germania oriens, vasto conflictu concurrunt. Terrae commovuntur, vires se colligunt, magna audacia peritiamque utriusque exercitus demonstrant. Imperatores, in acie positi, praeclaris consiliis suas legiones ad praelium ducunt. Celeriter castra adiri, munitiones facere, oppida expugnare contendunt. Telis missilibus evolvant, tormenta fervebant, caedes ac saevitia tam atrocis non vidit mundus. Navibus, aere, tellure, ingentibus machinis omnia conficiuntur. Equis instructis et quadrigis rapaciter currentibus, hostes impetu concitant. Tantus tumultus, clamor et cruor nullus ante visa sunt. Per annos longos milites perseverant, pugnando atque contionando.

Necesse est dicere, multa vulnera ac funera adfuisse, sed virtus et animus in bello enituit. Impetus, strategemata et consilia in dies mutabantur, acriter pugnatum est. Urbes vastantur, agri calamitate afficiuntur, populus in miseria versatur. Multos milites cadere, castra capi, imperatores interire vidimus. Quamvis animus perseveret, multitudo hominum fame et morbo vexatur. Nobis Romani nomen latet, sed horrida bella proxima et terribilis fama pervenit.

Haec est res bellica prima in toto orbe terrarum, quae omnes nationes invadit. Di immortales, nobis pacem reddite!

Non ho ovviamente alcuna conclusione seria da trarre, anche perché - traduzioni tacitiane a parte - in fondo ho solo giocherellato con IA e non ho davvero provato ad andare alla ricerca di reali applicazioni didattiche. È però vero che tutti e tre gli esempi proposti si prestano a riempire lo spazio di qualche lezione e a costruire momenti di dibattito critico in classe. Altro, per ora, non vorrei dire, timoroso (anzi sicuro) che al momento di andare in stampa questo articolo sarà già superato.

Cosa c'entra Indiana Jones?

A proposito di articoli: quello che scrissi su «La ricerca» n. 18 si intitolava *L'uomo classico tra humanitas, machinae e automata* e dava ampio spazio al cosiddetto «meccanismo di Anticitera», forse un complesso planetario meccanico del III sec. d.C. ora conservato al Museo Archeologico Nazionale di Atene. Guarda caso, l'ultimissimo film della saga di Indiana Jones (*Indiana Jones e il quadrante del destino*, 2023) è fantasiosamente incentrato proprio sulla ricerca di quell'antico quadrante, cui si attribuiscono i poteri di una straordinaria "macchina del tempo". Con quello in mano i Nazisti sperano addirittura di poter ribaltare le sorti già scritte della Seconda guerra mondiale, perché - come si afferma nel film - chi lo possiede diventa come Dio. Sì, essere (quasi) come Dio... un'aspirazione che da Prometeo a Frankenstein ha talora attraversato l'uomo, in un'alternanza di successi e minacce che Yuval Noah Harari ha tratteggiato nel suo libro *Homo deus. Breve storia del futuro* (trad. it. di M. Piani, Bompiani, Milano 2017). La speranza è allora che IA non diventi una palestra di esperimenti per quello che già Eugenio Montale in *Piove* chiamò «l'uomo indiato» (*Satura*, 1971), ma uno strumento al reale servizio della comunità, e dunque - perché no? - anche della scuola e dell'educazione. Voglio però concludere con ironia e ritornare alla scuola, ricordando come ai nostri studenti è comunque già offerta da IA una chance in più in occasione delle verifiche: è il chatbot attivo sul portale www.prega.org (basato su ChatGPT) cui poter affidare voti e preghiere al santo preferito. Altro che la vecchia candelina accesa dalla nonna durante la Maturità!

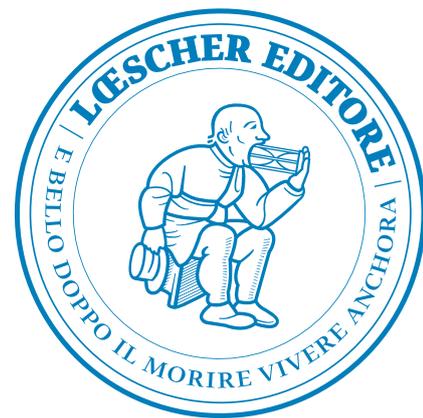
Mauro Reali

docente di liceo, dottore di ricerca in Storia Antica, è autore di testi Loescher di Letteratura Latina e di Storia. Le sue ricerche, realizzate presso l'Università degli Studi di Milano, riguardano l'Epigrafia latina e la Storia romana. È giornalista pubblicista e direttore responsabile de «La ricerca».

FORMAZIONE LOESCHER

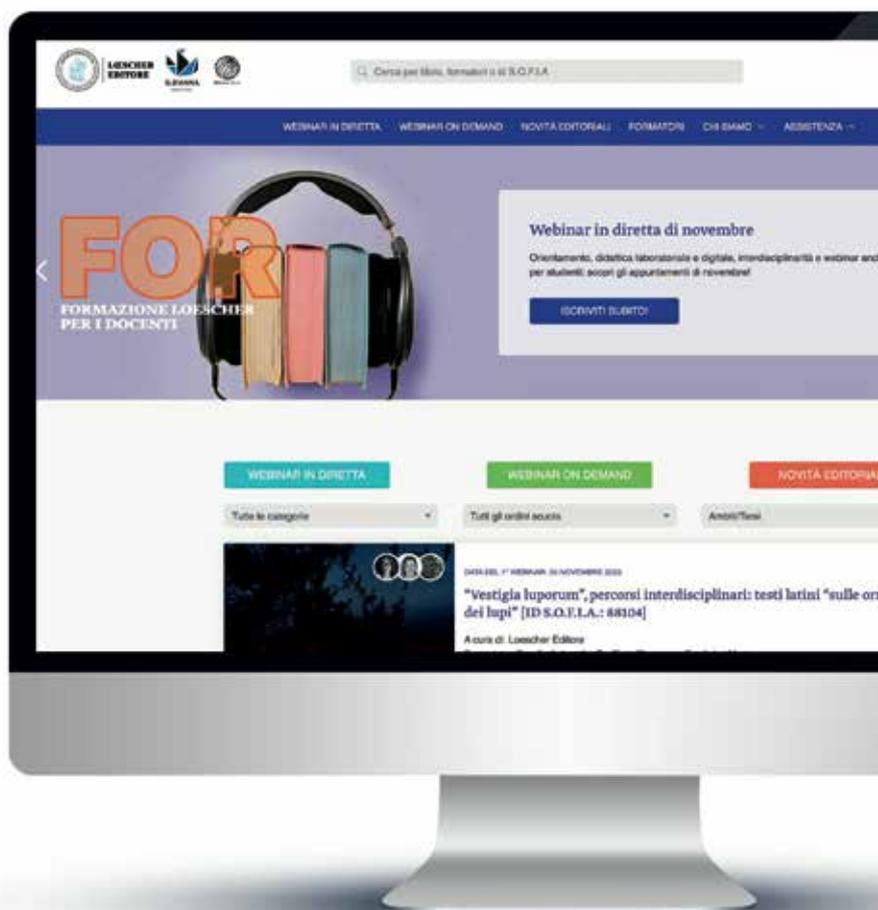
WEBINAR E SEMINARI SU S.O.F.I.A.

- ▶ **webinar** gratuiti in diretta: le registrazioni restano a disposizione degli iscritti per 30 giorni;
- ▶ **webinar on demand** a pagamento;
- ▶ **webinar di presentazione** delle nuove opere, tenuti dagli autori;
- ▶ **seminari** via Zoom o in presenza, richiesti dalle scuole per i propri docenti.



PER GLI STUDENTI

- ▶ **webinar per studenti:** il docente può iscriversi alla diretta su formazione.loescher.it, per osservare la propria classe o per ricavarne spunti didattici; gli **studenti** si iscrivono su webinarstudenti.loescher.it
- ▶ **NOVITÀ:** viene rilasciato l'**attestato di partecipazione** a studenti e docenti, sia per la diretta sia per la registrazione. Agli studenti della scuola secondaria di secondo grado l'**attestato vale ai fini dei PCTO**, se il webinar è stato inserito dalla scuola nel programma PCTO della classe.



formazione.loescher.it

Il destino della originalità

40

DOSSIER / Il destino dell'originalità

Lo sviluppo dell'IA generativa e quello dei sistemi antiplagio modificano il modo in cui valutiamo l'originalità di testi e immagini.

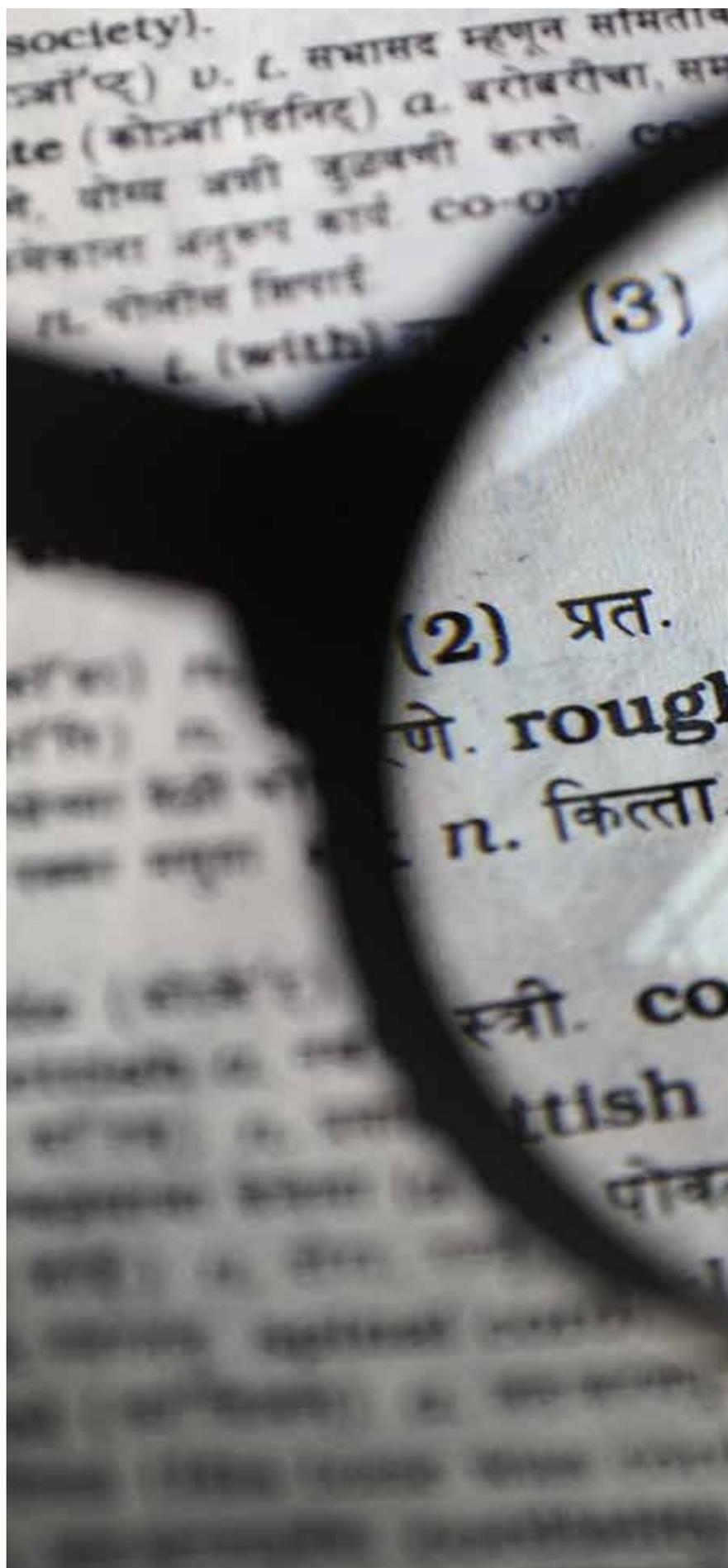
di Ubaldo Nicola

L'opinione prevalente fra i commentatori, da me stesso condivisa, è che ChatGPT sia uno strumento utile per molte cose ma incapace di vera originalità, sia nei contenuti sia nella loro resa stilistica.

Nel primo articolo di questo Dossier scopro però che i più efficienti software antifrode, specializzati nel confrontare un testo con tutto ciò che circola nel web, considerano un articolo redatto da ChatGPT come altamente originale, individuando un plagio solo nel 10% dei casi. Appena dopo apprendo che ChatGPT è efficacissimo nel riconoscere se un testo è veramente originale oppure è stato compilato da lui stesso. Per sapere se un tema svolto a casa è stato clonato da ChatGPT basta chiederlo a ChatGPT: dà risposte esatte nel 93% dei casi (una constatazione che potrebbe eliminare alla radice le preoccupazioni dei docenti riguardo al possibile uso truffaldino di questo strumento).

La confusione è notevole, e mi sembra riguardi principalmente l'idea di originalità, un concetto che appare in rapida trasformazione. I prodotti di una macchina giudicati poco originali dagli umani sono al contrario valutati come molto originali da altre macchine che li confrontano con quanto gli umani hanno "scritto" sul web negli ultimi decenni.

Ciò che sta avvenendo, mi sembra, è un divorzio fra originalità, cioè il grado in cui un'opera si discosta dalla produzione culturale corrente, e au-





कपडे धुण्याचे साध्याचे

v. t. & i. नक्कल करणे

n copy; कच्चा मसुदा. बिबि

copyright n. ग्रंथावरी

quetry n. नखरेल वागणे

smile).

रुघामारखा लाल (coral

ace).

corner (कॉर्नेर्) n.

v. t. & i. अडचणीत टा

cornerstone n. कोन

corollary (करो'लरि)

corona (करो'जड) n

coronation (कॉ'रो

coroner (कॉ'रनर)

करणारा अधिकारी.

coronet (कॉ'रो

(2) शोष्यावर

corporal (क

शासिक. क

corporat

corporat

corpo

(An

सि

स

स

स

स

स



↑ Immagine creata da Midjourney con il prompt "Crocerossine in un ospedale durante la prima guerra mondiale, nello stile di un documento storico".

tenticità, la capacità di affermare sino in fondo l'unicità della propria visione del mondo. Una connessione esistenziale che sembra indebolirsi nel momento in cui l'originalità comincia a essere definita, rilevata e misurata con metodi matematici, come oggi fanno i servizi di antiplagio di cui trattano gli altri articoli del Dossier, un settore probabilmente destinato a grandi sviluppi nei prossimi anni.

La trasformazione in atto mi sembra ancora più profonda relativamente alle applicazioni riguardanti il mondo delle immagini. Qui la mia opinione è che la IA generativa proponga immagini dotate di un alto tasso di originalità (in Midjourney tale funzione è regolabile con il comando "chaos" su una scala da 0 a 80). Lo documentano anche le immagini che illustrano questo numero de «La ricerca». Si può ora dar forma visiva a qualunque contenuto concettuale, indipendentemente dal suo livello di astrazione. Basta immaginare Einstein che mangia un gelato per vederlo realisticamente in un'immagine. Ci si può divertire creando un Biden in pigiama a pois o un Obama in tuta da ginnastica. Si possono creare pericolosi falsi storici, come l'immagine in questa pagina, oppure dare forma a concetti filosofici (si veda l'articolo a pagina X). Oppure ancora visualizzare complesse relazioni chimiche,

con conseguenze, immagino, utili ai ricercatori (si veda l'articolo di Grazia Pozzi sulla IA nella medicina).

Sono queste le prime ed embrionali forme di utilizzo di una macchina la cui potenza generativa si dimostrerà nel prossimo futuro, determinando una vera

rivoluzione nel mondo dell'immaginario; con la possibilità di rendere visivo tutto ciò che ci frulla nella testa, per quanto impossibile, irrealistico e sino a ieri inverosimile, cambierà l'immaginario privato e nasceranno nuove grammatiche della fantasia.

Ma cambierà anche l'immaginario sociale, cioè il modo in cui collettivamente immaginiamo le cose e i fenomeni. Forse esagero, ma già oggi l'archivio di Midjourney, liberamente consultabile da chiunque usi l'applicazione, costituisce un enorme database che gli archeologi del futuro potranno studiare per capire l'immaginario che ha contraddistinto questo decennio confrontandolo con quelli che seguiranno. Avranno a disposizione non solo le opere create dalla macchina, ma anche i prompt che le hanno generate, ovvero le istruzioni e quindi le intenzioni e le motivazioni dei creatori, il loro progetto. È la prima volta nella storia che questo aspetto della creazione, sinora intimo e privato, viene reso esplicito, documentabile e studiabile. Per sapere, ad esempio, in quali ruoli ci piace immaginare Einstein, oppure Chaplin, Napoleone o Trump, basta avere accesso al suddetto database e verificare statisticamente quanto, di volta in volta, i prompt (e quindi le immagini conseguenti) siano improntate all'ironia,

alla celebrazione o al sarcasmo.

E ogni artista che punti all'originalità può già ora determinare in modo quanto mai preciso il suo valore in questo campo: basta che fotografi una sua opera, la replichi con variazioni minime in Midjourney e la confronti quindi con quanto di simile viene immaginato in tutto il mondo. Non è poco.

Post scriptum

Ho concluso questa introduzione al Dossier con una frase molto breve, una *tagline* nel gergo giornalistico, per dimostrare de facto l'originalità del mio componimento, cioè di non aver usato l'intelligenza artificiale per redarlo. Pare infatti, come spiega l'ultimo articolo del Dossier, che uno dei parametri di rilevamento più usati dai sistemi di antiplagio sia proprio rilevare la presenza di tale vezzo stilistico, cui si attribuisce un grande punteggio di originalità, a sua volta indice di una composizione umana. A ChatGPT, infatti, non riesce molto facilmente.

Mi chiedo però cosa avverrà nel futuro, anche prossimo. L'originalità è sempre, per definizione, relativa a un contesto, e cosa accadrà quando ogni articolista accortamente infarcirà il suo pezzo con *tagline* dimostrative della sua fatica? Il quoziente di originalità di tale formula retorica decadrà drasticamente, e dovrà quindi essere sostituito da altri tipi di originalità stilistica.

Sganciata da ogni autenticità esistenziale, l'originalità creativa si avvia ad assumere il sistema della moda, cioè un susseguirsi periodico di formule che nascono originali ma cessano ben presto di esserlo, in un'alternanza veloce, immotivata e con scarse relazioni con mondo reale. Non è una bella previsione. (Nota, lettore: un'altra *tagline*).

Ubaldo Nicola

redattore de «La ricerca».

La matematica del plagio

L'unico metodo efficace per scoprire se un testo è stato elaborato da ChatGPT è chiederlo a ChatGPT.

di Mohammad Khalil e Erkan Er

In un recente articolo intitolato *The College Essay Is Dead* [Il saggio universitario è morto, N.d.T.] e pubblicato su «The Atlantic»¹, Stephen Marche esprime preoccupazione riguardo all'utilizzo di ChatGPT nella creazione di testi di alta qualità, spesso di natura accademica, sfruttando le capacità di elaborazione del linguaggio naturale dei chatbot.

Al contrario, Stokel-Walker² sottolinea come ChatGPT possa essere di aiuto agli studenti in diversi modi, per esempio nello scrivere saggi, risolvere compiti, creare script e nella valutazione delle loro prestazioni.

In questa situazione di incertezza, alcune istituzioni educative, come le scuole del Queensland e della Tasmania in Australia, ma anche i distretti scolastici di New York e Seattle, hanno adottato misure restrittive, vietando l'uso di ChatGPT sui dispositivi e sulle reti degli studenti. Molte università, col-

lege e scuole stanno valutando restrizioni simili³.

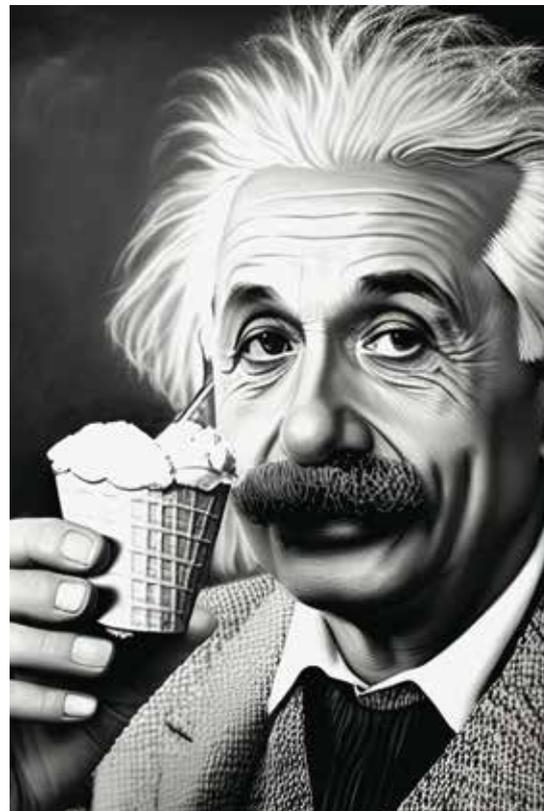
I sistemi anti-plagio

Il plagio è il presentare il lavoro o le idee di qualcun altro come propri, senza indicarne l'autore o la fonte. Un fenomeno che non riguarda solamente il testo, ma può estendersi anche alle immagini.

Oggi il dilagante problema del plagio online, soprattutto nei compiti e negli elaborati scritti, costituisce una delle principali sfide che la scuola deve affrontare⁴.

I più diffusi sistemi anti-plagio, secondo quanto riportato da alcune fonti⁵, sono Turnitin e iThenticate, due software sviluppati dalla stessa azienda, la iParadigms LLC. Dal 1997 sono stati sempre più utilizzati nelle istituzioni scolastiche, in particolare dalle università.

Il loro funzionamento è piuttosto semplice: con un click sono in grado di dire se un testo



o parte di esso è stato copiato da internet o da un altro testo disponibile in rete. Confrontano il compito dello studente con i testi contenuti nel database (materiali disponibili sul web, testi accademici, elaborati, tesi, articoli scientifici, papers) e producono un report dettagliato indicando la percentuale di originalità. Ma quanto sono affidabili questi sistemi?

Metodologia della ricerca

In un nostro studio⁶ abbiamo presentato i risultati di un'analisi di plagio effettuata su alcuni contenuti generati dall'intelligenza artificiale. In particolare abbiamo condotta un'analisi quantitativa, in cui il materiale generato da un chatbot è stato misurato attraverso gli indici di originalità prodotti da due software anti-plagio.

Per ottenere un campione rappresentativo abbiamo proposto 50 diversi argomenti e incaricato ChatGPT di redigere un

↑ Immagine creata da Midjourney con il prompt "Einstein mentre mangia un gelato".



↑ Due immagini create da Midjourney, la prima con il prompt "Obama in palestra", la seconda con il prompt "Trump travestito da contadino".

saggio di 500 parole su ognuno di essi. Ciascun output è stato poi convertito in un testo e salvato in un file separato, come se fosse stato redatto da uno studente.

La prima metà di questo materiale è stata analizzata dal software Turnitin, mentre la seconda da iThenticate. In aggiunta, come ulteriore misura di controllo, abbiamo chiesto alla stessa ChatGPT di stabilire se i saggi erano stati scritti dagli studenti oppure se, a suo giudizio, questi si erano avvalsi di fonti non dichiarate.

Il fallimento dei software anti-plagio

Una volta ottenuti i punteggi di somiglianza da entrambi i software, sono state create tabelle di frequenza in cui i punteggi di somiglianza sono stati raggruppati in quattro categorie: 0-10%, 10-20%, 20-40% e 40-100%.

Secondo i risultati ottenuti da iThenticate, la maggior parte dei saggi (il 68%) è stata classificata come verosimilmente originale, con una somiglianza minima (meno del 10%) rispetto ad altri contenuti presenti nel web. In alcuni altri saggi (20%

iThenticate ha rilevato un livello accettabile di somiglianza, compreso tra il 10% e il 20%, una percentuale che sale di livello (dal 20-40%) solo in tre casi. A nessun elaborato è stato assegnato un punteggio superiore al 40%. Il grado di somiglianza a livello medio tra tutti i saggi è stato valutato di 8,76%.

Emerge chiaramente che i saggi generati da ChatGPT e vagliati da iThenticate non presenterebbero problemi di plagio se venissero spacciati dagli studenti come frutto del loro lavoro.

Sottoponendo il secondo gruppo di saggi all'altro software, Turnitin, i punteggi di somiglianza erano relativamente più alti. Da una parte, infatti, è stato rilevato un grado minimo di somiglianza, inferiore al 10%, in quasi metà degli elaborati, e solo in sei casi un livello accettabile, compreso tra il 10% e il 20%.

D'altra parte, però, Turnitin ha individuato alcuni casi problematici, assegnando un punteggio di somiglianza (20-40%) a 6 elaborati. Inoltre, è stato identificato un evidente caso di plagio, dal momento che un saggio presentava un punteggio di superiore al 40%. Il grado di somiglianza medio tra tutti i saggi è risultato 13,72%, più alto

della valutazione eseguita da Turnitin (8,76%).

Argomenti oggettivi vs. argomenti controversi

I risultati di questa ricerca forniscono nuovi spunti per il dibattito in corso sull'utilizzo di ChatGPT nei compiti scolastici.

Aydın e Karaarslan⁷ hanno utilizzato ChatGPT per automatizzare e semplificare il processo di compilazione di una revisione della letteratura sulle applicazioni del *Digital Twin* [una rappresentazione virtuale di un oggetto fisico, N.d.T.] nel settore della salute. Hanno ottenuto gli abstract degli articoli pubblicati negli ultimi tre anni (dal 2020 al 2022) inserendo i termini «digital twin in healthcare» su Google Scholar e successivamente hanno fatto parafrasare gli abstract da ChatGPT.

L'analisi condotta con l'aiuto di iThenticate ha rivelato un problema significativo. Circa il 40% delle parti parafrasate create da ChatGPT risultavano essere simili o identiche a materiali già presenti online. Nel loro studio, dunque, il software anti-plagio ha funzionato egregiamente.

Noi invece riteniamo che

la possibilità che gli studenti possano utilizzare ChatGPT per completare i loro compiti senza essere scoperti sia molto più alta. Dei 50 saggi esaminati, infatti, ben 40 sono stati considerati altamente originali.

Ancor più interessante, però, è che i saggi rilevati come possibili casi di plagio dai software di controllo riguardavano principalmente argomenti di natura scientifica, come leggi fisiche, il riscaldamento globale e l'apprendimento automatico. Questi argomenti tendono a essere più oggettivi e tecnici, riducendo così la possibilità di includere interpretazioni personali. Al contrario, i saggi non identificati dai software come prodotti da ChatGPT avevano a che fare con argomenti più controversi, che richiedevano una riflessione personale, e temi quali le differenze culturali, le qualità di un buon insegnante o la natura della leadership. Poiché gli elaborati accademici dovrebbero rappresentare proprio il pensiero e l'analisi personale degli studenti su un determinato argomento, se un saggio prodotto da ChatGPT è sufficientemente ben scritto e sembra autentico, i software di rilevamento potrebbero non essere in grado di identificarlo come una potenziale frode.

Gli sviluppatori di Turnitin hanno rilevato questi problemi e stanno attualmente lavorando all'aggiornamento del proprio sistema⁸.

Solo ChatGPT score ChatGPT

Abbiamo infine esplorato un metodo diverso. Abbiamo sottoposto ognuno dei 50 saggi utilizzati nella ricerca precedente a ChatGPT ponendo la seguente domanda: «Questo testo è stato generato da un chatbot?».

ChatGPT è stata in grado di riconoscere se i testi erano stati generati autonomamente con una notevole precisione, supe-



riore al 92%. Ha identificato ben 46 casi di frode su 50, un risultato molto migliore degli strumenti tradizionali anti-plagio.

È chiaro che il metodo tradizionale di rilevamento di plagio deve essere aggiornato nella nuova era dominata dall'intelligenza artificiale.

Probabilmente sarà necessario adottare un approccio a due fasi: prima la verifica dell'origine del contenuto, seguita da un controllo di somiglianza. E, paradossalmente, il metodo più veloce ed efficace per controllare la fonte, ossia per sapere se un testo è stato prodotto dall'intelligenza artificiale generativa o da un umano, almeno al momento, è semplicemente chiederlo all'intelligenza artificiale stessa.

NOTE

1. S. Marche, *The College Essay Is Dead Nobody is prepared for how AI will transform academia*, in «The Atlantic», 2022.

2. C. Stokel-Walker, *AI bot ChatGPT writes smart essays-should academics worry?*, in «Nature». 10.1038/d41586-022-04397-7, 2022.

3. H. Wilcox, *Cheating Aussie student fails uni exam after being caught using artificial intelligence chatbot to write essay - now Australia's top universities*

are considering a bizarre solution to stop it happening again, apparso sul DailyMail, 2023.

4. T. Bertram Gallant, M. Picciotto, G. Bozinovic, E. Tour, *Plagiarism or not? investigation of Turnitin detected similarity hits in biology laboratory reports*, in «Biochemistry and Molecular Biology Education» 47(4), 370-379, 2019. Vedi anche N. Selwyn, *Not necessarily a bad thing...': A study of online plagiarism amongst undergraduate students*, in «Assessment & Evaluation in Higher Education» 33(5), 465-479, 2008.

5. S. A. Meo, M. Talha, *Turnitin: Is it a text matching or plagiarism detection tool?*, in «Saudi journal of anaesthesia» 13(Suppl 1), S48, 2019. Vedi anche P. Stapleton, *Gauging the effectiveness of anti-plagiarism software: An empirical study of second language graduate writers*, in «Journal of English for Academic Purposes» 11(2), 125-133, 2012.

6. M. Khalil, E. Er, *Will ChatGPT get you caught? Rethinking of Plagiarism Detection*, in «HCI International», 2023.

7. O. Aydın, E. Karaarslan, *OpenAI ChatGPT generated literature review: Digital twin in healthcare*, in «Emerging Computer Technologies 2» 2022, pp. 22-31.

8. C. Caren, *AI writing: The challenge and opportunity in front of education now*, www.turnitin.com/blog, 2022. Vedi anche A. Chechitelli, *Sneak preview of Turnitin's AI writing and ChatGPT detection capability*, www.turnitin.com/blog, 2022.

Traduzione di Francesca Nicola.

Mohammad Khalil

è ricercatore presso il Centre for the Science of Learning & Technology (SLATE) dell'Università di Bergen, Norvegia.

Erkan Er

è ricercatore presso la Middle East Technical University, Turchia.

← Immagine creata da Midjourney con il prompt "Biden indossa un pigiama a pois".

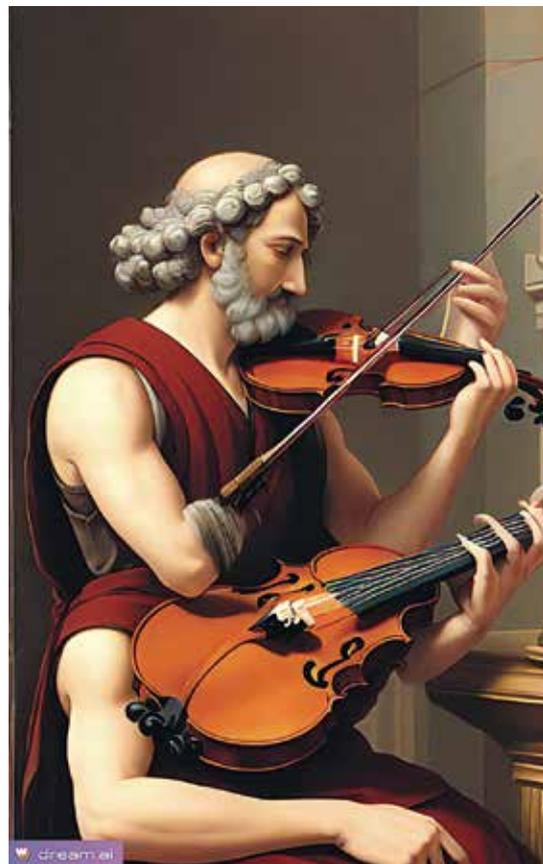
L'impatto dell'IA sulla frode accademica

46

DOSSIER / L'impatto dell'IA sulla frode accademica

È ormai necessario ripensare il concetto di integrità accademica e i metodi sino ad ora utilizzati per garantirla.

di Jo Ann Oravec



La disonestà accademica è da tempo riconosciuta come un «problema ricorrente» nel contesto della valutazione accademica¹, con casi di frode «scandalosa» che spesso vengono sospettati o scoperti². Supportato dalla tecnologia, negli ultimi decenni questo fenomeno, ha assunto diverse forme, presentando nuove sfide etiche. Ma definire con precisione cosa costituisce una frode può rivelarsi una sfida complessa, dato che le politiche istituzionali mostrano notevoli variazioni riguardo alle tecnologie ammissibili³.

Ad esempio si è dibattuto sull'uso da parte degli studenti dei servizi online di correzione grammaticale⁴ e strumenti di editing come Grammarly sono stati considerati in alcuni contesti come una possibile fonte di indebito vantaggio nella

produzione di elaborati. Ciò nonostante, negli ultimi anni, gli strumenti offerti da Grammarly sono diventati parte integrante delle funzionalità standard di molti programmi di elaborazione testi e persino di motori di ricerca.

Allo stesso modo, solo pochi decenni fa, l'utilizzo delle calcolatrici durante gli esami in classe era considerato una frode⁵, poiché si riteneva che alterasse l'esperienza educativa e compromettesse una valutazione accurata. Analogamente, gli iPad sono stati ritenuti potenzialmente dannosi, in quanto si temeva che potessero minacciare l'interazione in classe⁶.

Con la crescente aggiunta dei sistemi di intelligenza artificiale in diverse applicazioni, come motori di ricerca, strumenti di progettazione e produzione di documenti, la questione dell'in-

tegrità accademica diventa sempre più urgente e richiede definizioni e soluzioni più precise rispetto a quelle attualmente disponibili.

Ecco alcuni esempi che illustrano come i sistemi generativi basati sull'intelligenza artificiale vengano sfruttati in modo non etico in situazioni di frode accademica già conosciute:

1. Errata attribuzione del merito

È ben noto come l'assegnazione inappropriata di co-paternità di un testo per ragioni estranee ai meriti effettivi sia un fenomeno diffuso tra gli accademici. In diversi campi di ricerca, la pratica della co-paternità «fantasma» o «onoraria» genera apprensione tra i responsabili delle ricerche e gli editori⁷. Spesso, una persona viene inclusa come autore per ragioni politiche, per favorire la sua carriera accademica o in ge-

↑ Immagine creata da Midjourney con il prompt «Omero suona il violino».

nerale per motivi estranei al suo effettivo contributo alla ricerca.

Sta emergendo in modo sempre più evidente l'importanza di promuovere una maggiore trasparenza e di istituire linee guida più chiare⁸, specialmente considerando come le attribuzioni improprie possano perpetuare situazioni di discriminazione, per esempio rafforzando la disparità di genere nelle carriere universitarie⁹.

L'utilizzo di sistemi come ChatGPT e Bard come co-autori di un testo può peggiorare notevolmente questo problema, in primo luogo perché vi è il rischio che le informazioni o i risultati generati siano accettati come veritieri senza un'accurata verifica delle fonti e delle ragioni che ne attestano l'affidabilità¹⁰.

2. Impersonificazione

Tra le varie forme di frode accademica rientra anche l'impersonificazione, ossia la pratica per cui professionisti retribuiti o abili dilettanti completano compiti o sostengono esami al posto degli studenti. In passato era un'eventualità relativamente rara¹¹, ma oggi è diventato un problema significativo per le università che offrono corsi online. Infatti, ChatGPT ha la capacità di simulare le caratteristiche vocali e lo stile di scrittura delle persone. In altre parole, può generare testi o risposte in modo che sembrino scritti o espressi da un individuo specifico, imitandone il modo di scrivere o di esprimersi verbalmente¹².

3. Servizi a pagamento

Già prima che ChatGPT iniziasse a fornire materiali personalizzati agli studenti, diverse aziende e servizi on line vendevano documenti accademici come saggi, ricerche, compiti e tesi, che gli studenti presentavano come loro¹³. Erano diffuse anche pratiche più informali di condivisione di documenti accademici, spesso tramite associazioni studentesche e social media.

Oggi gli output generati da ChatGPT e altre applicazioni basate sull'IA rappresentano una sfida per questi servizi a pagamento. Una delle principali funzioni di ChatGPT, infatti, è la generazione di contenuti e la sua capacità di adattarli a vari compiti ed esercizi¹⁴, il che potrebbe rendere obsoleti i servizi di frode a pagamento. Attualmente, l'utilizzo di ChatGPT e Bard è quasi sempre economico o persino gratuito, ma in un futuro prossimo potrebbe diventare finanziariamente inaccessibile per alcune persone, sollevando questioni etiche legate alle disparità delle risorse.

Una prospettiva di collaborazione

Un approccio alternativo può essere considerare le sfide legate all'utilizzo di ChatGPT e Bard come opportunità di collaborazione. Piuttosto che trattare queste tecnologie come semplici fonti di informazioni, possiamo riconoscerle come partner attivi e reattivi nella creazione di contenuti accademici.

Spinge in questa direzione l'aumento della complessità nella vita professionale e nella società, che rende necessario preparare gli studenti alla collaborazione produttiva: devono imparare a comunicare in modo efficace, condividere idee, lavorare in gruppo e utilizzare le risorse disponibili, compresa l'IA, per affrontare sfide complesse. Questa preparazione è cruciale per affrontare le sfide del mondo reale e sfruttare le opportunità future, sia sul piano professionale sia nella partecipazione alla vita della comunità¹⁵.

C'è un rischio connesso a questo approccio, e cioè che gli studenti trascurino lo sviluppo di competenze fondamentali come la scrittura per concentrarsi principalmente sulla modifica e il perfezionamento delle risposte generate dall'IA in modo da farle sembrare il frutto del loro

lavoro. Per evitare questo rischio è fondamentale promuovere una collaborazione riflessiva e consapevole tra esseri umani e IA.

Questa strategia dovrebbe includere la corretta attribuzione del contributo a un lavoro scolastico, cosa tutt'altro che facile, specialmente quando autori umani e tecnologie di IA creano insieme contenuti. È importante che gli insegnanti sviluppino domande e prompt specifici per ciascuna disciplina accademica, in modo che gli studenti possano comprendere appieno l'interazione tra l'IA e il loro contributo intellettuale.

Gli strumenti di rilevazione delle frodi basati sull'IA non dovrebbero essere utilizzati esclusivamente per riconoscere atti di plagio, ma impiegati come strumenti di supporto nei processi di riflessione degli studenti. Identificare chi e in che misura ha partecipato alla creazione di un lavoro accademico dovrebbe spingere i ragazzi a riflettere sull'importanza di attribuire i contributi in modo equo e corretto, promuovendo una collaborazione etica e responsabile tra gli esseri umani e l'intelligenza artificiale.

I sistemi di rilevamento delle frodi

Nel campo del rilevamento del plagio stanno emergendo diverse soluzioni mirate a fronteggiare le preoccupazioni legate all'uso di ChatGPT e tecnologie simili. Ad esempio, il software Turnitin, utilizzato da 62 milioni di studenti in tutto il mondo, è stato recentemente aggiornato includendo funzionalità utili a rilevare e contrassegnare i contenuti generati da ChatGPT¹⁶, anche se i dettagli precisi di tali funzionalità non sono stati ancora rilasciati. Altre tecnologie anti-plagio (di cui parleremo più avanti) richiederanno funzionalità di sistema aggiuntive o marcature specifiche, cosa che

renderà necessaria una collaborazione continua tra gli sviluppatori di sistemi di intelligenza artificiale generativa e i fornitori di sistemi anti-plagio.

Di seguito alcune delle strategie per rilevare l'uso di ChatGPT nelle produzioni accademiche.

Filigrane sui materiali

OpenAI sta attualmente sviluppando strategie per applicare filigrane (marchi o simboli identificativi) ai contenuti generati da ChatGPT. L'obiettivo è consentire l'identificazione di tali contenuti anche se subiscono successive modifiche¹⁷. È importante notare che non tutti i ricercatori sono convinti che questa soluzione sia efficace, poiché ritengono che potrebbe essere aggirata in vari modi. Alcuni esperti sottolineano che potrebbe essere difficile creare filigrane che rimangano rilevabili anche in presenza di modifiche sostanziali del contenuto¹⁸.

Utilizzo di ChatGPT

ChatGPT e altri sistemi di intelligenza artificiale generativa possono spesso identificare i loro stessi contenuti attraverso diversi metodi. Khalil ed Er¹⁹ hanno proposto un approccio in due fasi per il rilevamento del plagio: nella prima fase si accerta l'origine del contenuto, ovvero se il materiale in questione è stato generato da ChatGPT o da un'altra fonte. Nella seconda, si effettua un controllo di somiglianza per valutare quanto il contenuto sia simile a modelli noti o precedenti generati da ChatGPT.

Misure della *perplexity* e della *burstiness*

Stanno emergendo strumenti, come GPTZero, in grado di esaminare la casualità e la complessità di un testo²⁰. Un'unità di misura è la *perplexity* (neologismo nato per estensione da

PerplexityAi, un motore di ricerca alimentato dall'intelligenza artificiale), che indica quanto un modello linguistico sia in grado di fare previsioni accurate su un campione di testo. Se è bassa significa che il testo è facile da predire e che quindi molto probabilmente è stato generato da una macchina. Gli esseri umani sono più imprevedibili e tendono a scrivere testi con una *perplexity* più alta.

GPTZero confronta anche la *burstiness*, ossia il livello di diversità nella struttura dei periodi. Un software basato su IA tende a scrivere frasi di lunghezza e di intensità uniformi, mentre l'uomo è in grado di accostare frasi più lineari e semplici a frasi più complesse e articolate. Tuttavia, questi strumenti sono tutt'altro che infallibili.

Analisi delle variazioni in un testo

È possibile chiedere agli studenti di conservare tutte le versioni di un documento, cioè ogni modifica e revisione apportate durante il processo di creazione. E questo può essere un metodo efficace per individuare la presenza o il contributo di sistemi di intelligenza artificiale nella stesura di un documento. Piattaforme come Google Docs o Microsoft Word offrono una funzione di "cronologia" che tiene traccia di tutte le modifiche apportate a un testo.

Riprogettazione delle attività

Per affrontare alcune sfide legate all'uso dell'IA nell'istruzione è possibile riformulare le attività in modo da rendere evidente l'impiego di ChatGPT. Ad esempio, è possibile richiedere agli studenti di confrontare i risultati da loro prodotti con quelli generati da ChatGPT o da altri sistemi di intelligenza artificiale, come illustrato in uno

studio di caso riguardante una classe di laurea magistrale in economia aziendale condotto da Mok nel 2023.

Rendere trasparente e chiara l'integrazione dell'intelligenza artificiale nelle attività scolastiche consente di valutare in modo accurato ed equo il contributo degli studenti e quello dell'IA, promuovendo la responsabilità e la trasparenza.

La mancanza di trasparenza nelle istituzioni educative

Una questione di grande importanza è la mancanza di trasparenza nella comunicazione e nella descrizione dei processi di rilevamento delle frodi degli studenti. Molto spesso, le istituzioni educative non forniscono informazioni dettagliate sull'utilizzo di tali sistemi, limitandosi a brevi dichiarazioni nei programmi di studio o nei manuali degli studenti. Questo è sorprendente, considerando che tali istituzioni sono tenute a essere trasparenti su numerosi altri aspetti dell'istruzione e della vita studentesca.

La maggior parte delle politiche di condotta accademica sono vaghe e possono essere ambigue, lasciando ai docenti una discrezionalità considerevole²¹. Molte frodi possono avvenire involontariamente e derivare da errori o da fraintendimenti riguardo ai compiti e i test²². Con l'avvento di nuove tecnologie come ChatGPT e Bard, il rischio di equivoci aumenta. Le accuse di utilizzo improprio di questi sistemi di intelligenza artificiale possono avere conseguenze gravi per gli studenti, indipendentemente dal fatto che vengano o meno scagionati.

Consideriamo questa vicenda svoltasi all'Università della California-Davis e riporta da Jimenez²³:

Quando William Quarterman si è collegato al portale web

degli studenti per verificare i risultati del suo esame di storia, è rimasto sconcertato nel vedere un'accusa di frode alleghata dalla sua professoressa. La docente aveva utilizzato un software di rilevamento dell'intelligenza artificiale, GPTZero, dopo aver notato che le risposte del ragazzo «avevano poca attinenza con le domande». Secondo il software, la professoressa aveva ragione. Allo studente è stato attribuito un voto negativo ed è stato segnalato all'Ufficio di sostegno studentesco e affari giudiziari dell'Università, accusato di disonestà accademica. Nonostante Quarterman abbia negato di aver ricevuto aiuto dall'IA, gli è stato chiesto di comparire davanti alla corte d'onore dell'università, un'esperienza che gli provocato attacchi di panico. Alla fine, però, è stato scagionato.

La complessità della accusa di frode e le sue implicazioni

Le definizioni di “imbroglio”, “plagio” e “frode” hanno un impatto notevole. Gli studenti accusati di imbrogliare e che non sanno come difendersi possono vivere una significativa diminuzione della loro autostima, come evidenziato da uno studio condotto da Duncan e Joyner nel 2022. E anche quelli che confessano di aver imbrogliato possono affrontare conseguenze durature sia sul piano accademico che sociale. È fondamentale, quindi, che gli sviluppatori e gli utenti di programmi anti-frode prestino grande attenzione alle implicazioni sociali ed etiche²⁴.

L'analisi e la verifica dei sistemi di rilevamento delle frodi sono complesse principalmente perché è raro ottenere una prova definitiva. Spesso non c'è un elemento incontrovertibile o inequivocabile che dimostri che uno studente ha effettivamente imbrogliato durante un esame o

un compito. Ad esempio, i sistemi di e-proctoring (supervisione degli studenti durante un esame mediante software) collegano il desiderio di imbrogliare con segnali fisici involontari degli studenti (movimenti degli occhi, eccessiva sudorazione) conosciuti come *leakage*. Tuttavia, queste manifestazioni fisiche non sono certamente prove conclusive di colpevolezza. Il fatto che uno studente mostri segni di nervosismo o ansia durante un esame non significa che stia imbrogliando.

A differenza dei contesti lavorativi o militari, in cui i comportamenti scorretti pongono un problema giuridico e finiscono quindi davanti a tribunali, nel contesto accademico l'elaborazione delle accuse avviene principalmente all'interno dell'istituzione educativa stessa. Ciò significa che non vi sono molti precedenti legali o sentenze giudiziarie atte a stabilire standard specifici per questi casi. Questa mancanza può rendere le decisioni delle istituzioni educative più soggette a discrezionalità, a meno che non ci siano politiche chiare e codici di condotta specifici in vigore. Molto spesso la conclusione di un caso di frode accademica si basa sull'ammissione di colpa da parte dello studente. Se riconosce di aver imbrogliato, il caso è generalmente considerato chiuso. Ma se nega o contesta le accuse, può diventare difficile stabilire con certezza se la frode è avvenuta o meno.

La posizione svantaggiata degli studenti

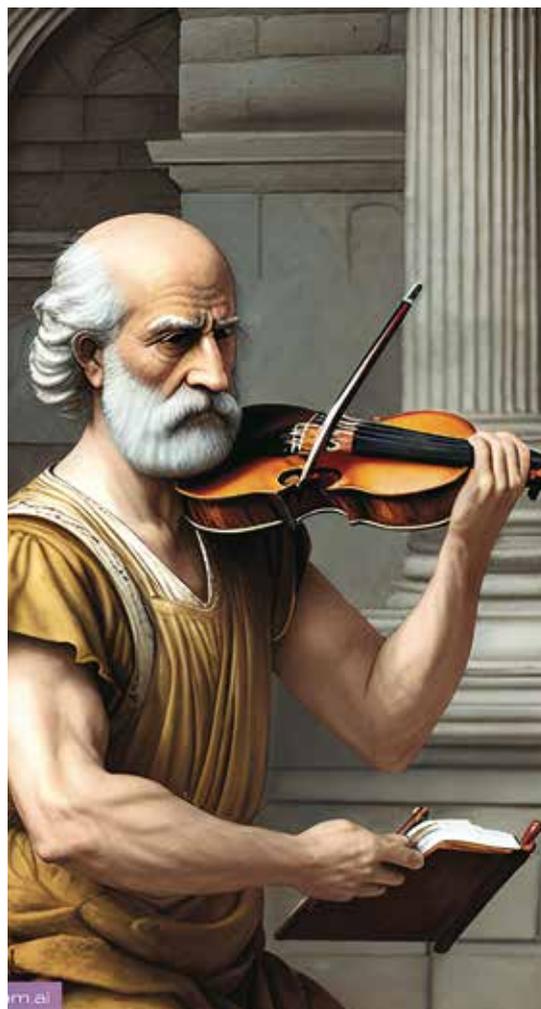
L'implementazione di tecnologie di rilevamento delle frodi, come il riconoscimento facciale, solleva importanti preoccupazioni legate alla privacy degli studenti. La raccolta di dati personali può produrre disagio e alimentare preoccupazioni sulla sicurezza. Inoltre, c'è il rischio che tali tecnologie si-

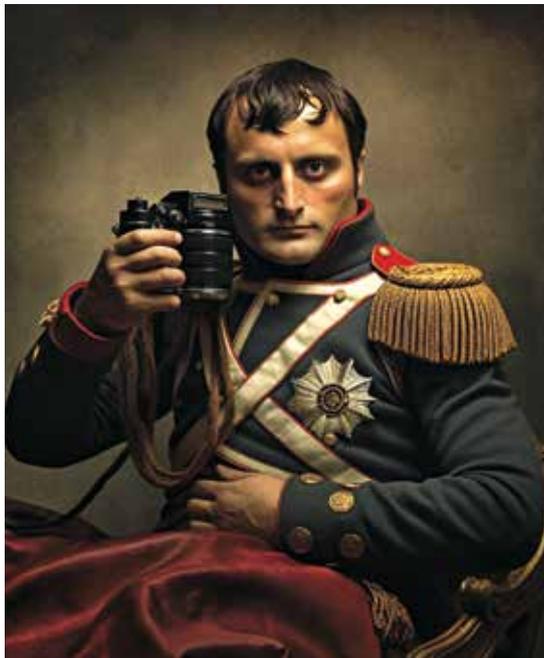
ano influenzate da pregiudizi, rendendo le valutazioni non del tutto obiettive ed equilibrate.

Un esempio concreto che evidenzia queste preoccupazioni è la storia di Amaya Ross, una studentessa afroamericana che si è trovata a utilizzare Proctorio, il software di Proctoring, per un quiz di biologia. Nonostante avesse provveduto ad assicurarsi una buona illuminazione e a sostenere l'esame durante le ore di luce del giorno, il software non è stato in grado di riconoscere il suo volto²⁵. Questo episodio mette in evidenza come tali tecnologie non solo sollevino questioni sulla privacy, ma possano anche causare disagi agli studenti.

Gli studenti sono in una posizione di svantaggio rispetto alle organizzazioni tecnologiche specializzate nel rilevamento delle frodi, come Turnitin, che

↓
Immagine creata da Midjourney con il prompt “Omero suona il violino”.





↑ Immagine creata da Midjourney con il prompt "Napoleone usava le reflex".

operano tramite accordi di *outsourcing* e sono considerate entità esterne alle istituzioni scolastiche. Anche nel caso in cui un particolare metodo o software di rilevamento delle frodi venga condannato o giudicato inappropriato, è probabile che nuove soluzioni siano adottate per soddisfare le esigenze delle istituzioni scolastiche. La mancanza di risorse finanziarie può inoltre ostacolare gli studenti nel contestare le accuse di frode attraverso azioni legali, con conseguenze significative sulla loro reputazione accademica ed emotiva²⁶.

NOTE

1. B. K. Pathak, *Emerging online educational models and the transformation of traditional universities*, in «Electronic Markets» 26(4), 2016.
2. M. Klein, *CUNY professors uncover 'scandalous' level of cheating in final exams*, «New York Post» 13 giugno 2020.
3. K. A. Burgason, O. Sefiha, L. Briggs, *Cheating is in the eye of the beholder: An evolving understanding of academic misconduct*, in «Innovative Higher Education» 44(3), 2019.
4. N. K. Thi, M. Nikolov, *How teacher and Grammarly feedback complement one another in Myanmar EFL students' writing*, in «The Asia-Pacific Educa-

tion Researcher» 31(6), 2022.

5. T. Dick, *Supercalculators: Implications for calculus curriculum, instruction, and assessment*, in J. T. Fey, *Calculators in Mathematics Education*, Yearbook of the National Council of Teachers of Mathematics, Reston 1992.
6. D. R. Perry, A. K. Steck, *Increasing student engagement, self-efficacy, and meta-cognitive self-regulation in the high school geometry classroom: do iPads help?*, in «Computers in the Schools», 32(2), 80569. 2015.1036650.
7. G. Pruschak, C. Hopp, *And the credit goes to... Ghost and honorary authorship among social scientists*, in «Plos one» 17(5), 2022.
8. D. Cutas, D. Shaw, *Writers blocked: On the wrongs of research coauthorship and some possible strategies for improvement*, in «Science and Engineering Ethics» 21, 2015.
9. J. Goodman, *Attribution: How authorship credit contributes to the gender gap*, in «Yale Journal of Law & Technology», 2023 (in uscita).
10. B. Tolchin, *Are AI chatbots in healthcare ethical?: Their use must require informed consent and independent review*, in «Medpage Today», aprile 2023.
11. P. G. Moeck, *Academic dishonesty: Cheating among community college students*, in «Community College Journal of Research and Practice» 26(6), 2002.
12. G. Marks, *It sounds like science fiction but it's not: AI can financially destroy your business*, «The Guardian», 9 aprile 2023.
13. A. Amigud, P. Dawson, *The law and the outlaw: Is legal prohibition a viable solution to the contract cheating problem?*, in «Assessment & Evaluation in Higher Education» 45(1), 2020.
14. B. A. Anders, *Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking?*, in «Patterns» 4(3), 2023.
15. P. Jandrić et al., *Collective writing: The continuous struggle for meaning-making*, in «Postdigital Science and Education» 1-43, 2022.
16. J. Hsu, *Plagiarism tool gets a ChatGPT detector – some schools don't want it*, in «New Scientist», 3 aprile 2023.
17. G. Eysenbach, *The role of ChatGPT, generative language models, and arti-*

ficial intelligence in medical education, in «JMIR Medical Education» 9(1), 2023.

18. K. Darlington, *ChatGPT is a ground-breaking application of AI that we will all be hearing about in the near future*, in «IQ: The RIMPA Quarterly Magazine» 39(1), 2023.
19. M. Khalil, E. Er, *Will ChatGPT get you caught? Rethinking of plagiarism detection*, arXiv preprint, 2023.
20. P. Yu, P., Chen et al., *CHEAT: A large-scale dataset for detecting ChatGPT-written abstracts*, arXiv preprint arXiv:2304.12008, 2023.
21. K. Zamastil, *Legal issues in US higher education*, in «Common Law Review», 6, 2004.
22. G. T. Dow, *Do cheaters never prosper? The impact of examples, expertise, and cognitive load on cryptomnesia and inadvertent self-plagiarism of creative tasks*, in «Creativity Research Journal» 27(1), 2015.
23. K. Jimenez, *Professors are using ChatGPT detector tools to accuse students of cheating. But what if the software is wrong?*, apparso sul sito USA Today il 12 aprile 2023.
24. E. Taylor, *Surveillance schools: Security, discipline and control in contemporary education*, Springer, New York, 2013.
25. J. Verplaetse et al., *You can judge a book by its cover: The sequel: A kernel of truth in predictive cheating detection*, in «Evolution and Human Behavior» 28(4), 2007.
26. Z. Corbyn, *"I'm afraid": Critics of anti-cheating technology for students hit by lawsuits*, «The Guardian», 26 agosto 2022.

Tratto da Jo Ann Oravec, *Artificial Intelligence Implications for Academic Cheating: Expanding the Dimensions of Responsible Human-AI Collaboration with ChatGPT and Bard*, in «Journal of Interactive Learning Research» (2023), 34(2), 213-237.

Traduzione di Francesca Nicola.

Jo Ann Oravec

insegna all'Università del Wisconsin-Whitewater, Stati Uniti.

Il proctoring: una risorsa o il Grande Fratello?

Oggi esistono software basati sull'intelligenza artificiale che decretano se durante gli esami online uno studente sta copiando. Ma il loro uso non è privo di rischi.

di Simon Coghlan, Tim Miller, Jeannie Paterson

Di recente le tecnologie di sorveglianza degli esami online, conosciute come Online Proctoring (OP) o semplicemente OP, sono diventate un argomento di grande rilevanza a causa dell'aumento del ricorso alla didattica a distanza¹. Durante i periodi di lockdown dovuti alla pandemia da COVID-19, poi, si è verificato un notevole aumento globale dell'uso di queste tecnologie².

Nonostante le istituzioni educative possano vigilare sugli studenti durante gli esami online tramite video, ad esempio utilizzando Zoom, i programmi software di sorveglianza online offrono funzionalità di monitoraggio avanzate, che possono essere eseguite sia da esseri umani sia in modo automatizzato. Per questo stanno guadagnando terreno, tanto che secondo alcuni ricercatori potrebbero diventare la "nuova normalità" nell'istruzione superiore in tutto il mondo³. Di fatto, molte università li hanno già adottati.

Autenticazione e privacy

—
I software di proctoring online [proctor in inglese significa "sorvegliante", N.d.T.] prevedono un

complesso sistema di autenticazione dei candidati. Raccolgono dati sensibili, come indirizzi IP, nomi, indirizzi email e password.

In genere chiedono ai candidati di fornire un documento d'identità e di scattare una foto di loro stessi, che viene confrontata in tempo reale da un supervisore umano o da un algoritmo di intelligenza artificiale.

Alcuni vanno oltre la semplice raccolta di dati: registrano la velocità di digitazione e la pressione applicata durante la digitazione del nome all'inizio dell'esame e poi usano questi dati biometrici per creare un modello del comportamento di battitura del candidato. Durante l'esame, il programma utilizza questo modello per assicurarsi che la persona che sostiene la prova sia effettivamente quella che dovrebbe essere.

Altri programmi durante l'esame chiedono dati biometrici reali, come le impronte digitali, per ulteriori verifiche di identità; un passaggio che aggiunge un ulteriore livello di sicurezza e autenticazione, ma solleva preoccupazioni sulla privacy e sulla condivisione di dati sensibili.

La relazione tra la quantità di controllo offerta dai programmi di autenticazione e l'affidabilità

di questo processo è strettamente legata alla questione della privacy. I programmi che forniscono meno controllo sono spesso pubblicizzati dai loro fornitori come soluzioni meno invasive per la dimensione privata degli utenti. Questo perché un minore controllo significa generalmente che durante il processo di autenticazione sono raccolti o elaborati meno informazioni sensibili. La scelta finale tra i diversi livelli di autenticazione dipende dalle esigenze specifiche dell'istituzione educativa e dalle preoccupazioni riguardanti la privacy degli studenti.

Il riconoscimento vocale

—
Durante le procedure di sorveglianza agli esaminandi può essere chiesto di attivare la loro webcam e di effettuare una rotazione completa di 360 gradi del dispositivo al fine di "scansionare" l'ambiente, alla ricerca di materiali non autorizzati o per verificare la presenza di familiari, amici o coinquilini⁴. Alcuni programmi possono rilevare anche la presenza di altri dispositivi, ad esempio telefoni cellulari.

Alcuni algoritmi di intelligenza artificiale attuano il rico-

noscimento vocale, analizzando la voce dell'utente per rilevare suoni o parole sospette durante l'esame, come il parlare ad alta voce. Altri si basano sul riconoscimento facciale, analizzando il volto dell'utente per verificarne l'identità. Altri ancora aggiungono la "rilevazione facciale", ossia monitorano l'aspetto del candidato durante l'esame per individuare comportamenti sospetti o anomalie, come movimenti inattesi o espressioni del viso che potrebbero suggerire un tentativo di copiare.

L'apprendimento automatico

La maggior parte dei sistemi di proctoring online fa uso dell'automazione basata sull'apprendimento automatico (Machine Learning, ML), una specifica tecnica di intelligenza artificiale che utilizza dati per creare modelli di rilevamento. In questo contesto, gli algoritmi di machine learning vengono addestrati attraverso l'analisi di migliaia di esempi video al fine di rilevare i movimenti degli occhi e della testa correlati a comportamenti sospetti da parte degli studenti durante un esame online. Questo processo di addestramento si basa su due categorie principali di input: gli esempi "negativi" e quelli "positivi".

Con "negativi, in questo contesto, si intendono situazioni o comportamenti considerati normali, corretti o accettabili. Gli algoritmi vengono addestrati utilizzando questo tipo di esempi per imparare a riconoscere ciò che costituisce un comportamento appropriato. Un esempio negativo potrebbe essere una registrazione video di uno studente che svolge un esame seguendo alla lettera tutte le istruzioni e senza mostrare alcun atteggiamento anomalo.

Gli esempi positivi rappresentano situazioni o comportamenti considerati sospetti, inappropriati o scorretti, come

il parlare ad alta voce o l'eseguire movimenti oculari insoliti. Questi esempi addestrano gli algoritmi a identificare atteggiamenti associabili a tentativi di frode. Ma in realtà, alcune azioni potenzialmente rilevabili dai programmi di proctoring come sospette potrebbero essere del tutto normali: ad esempio parlare da soli o guardarsi intorno mentre si pensa.

Il ruolo dell'uomo e della intelligenza artificiale

I sistemi di sorveglianza online rilevano condotte potenzialmente sospette durante gli esami, ma non emettono automaticamente giudizi di frode. La decisione finale spetta agli esseri umani responsabili della supervisione dell'esame.

Il software, infatti, non raggiunge il 100% di accuratezza e può produrre sia falsi negativi (non riconoscere comportamenti sospetti quando dovrebbe) sia falsi positivi (cioè riconoscere erroneamente comportamenti come sospetti quando non lo sono). Poiché la tecnologia non è infallibile, l'intervento umano rimane cruciale. E dipende sempre dalla discrezionalità dell'educatore o dell'istituzione come gestire i falsi positivi, cioè i casi in cui alla fine si verifica l'innocenza di uno studente precedentemente sospettato.

Alcune aziende di proctoring sostengono che la combinazione di intelligenza artificiale e esaminatori umani ben formati offre la massima accuratezza e affidabilità. Ad esempio ProctorU ha dichiarato: «Il nostro obiettivo non è sostituire gli esseri umani, ma migliorare l'accuratezza della sorveglianza, aiutandoli a individuare dettagli come ombre, sussurri, rumori o riflessi che altrimenti potrebbero passare inosservati»⁵.

Se è indubbio che queste tecnologie offrono vantaggi, è altrettanto chiaro che occorre valutarne attentamente gli

aspetti negativi e proteggere gli interessi degli studenti e delle istituzioni educative.

Di solito agli studenti (a meno che non abbiano bisogni speciali) non viene concessa l'opzione di rifiutare l'utilizzo. Le istituzioni concedono una certa discrezionalità agli educatori, ma in genere insistono affinché usino la tecnologia a disposizione.

Le aziende che sviluppano i software OP sono principalmente mosse dal profitto, e ciò comporta preoccupazioni sulla loro neutralità e obiettività. Ad esempio, in alcuni casi il personale delle aziende è coinvolto nella configurazione tecnica e nella sorveglianza stessa degli esami, e tale commistione potrebbe risolversi in un incentivo nel rilevare comportamenti sospetti, al fine di mantenere la reputazione della loro tecnologia.

I pregiudizi del machine learning

Vi è poi una preoccupazione fondamentale costituita dai pregiudizi insiti nei sistemi di machine learning. Questi, infatti, mostrano una tendenza sistematica a favorire o penalizzare specifici gruppi o categorie di individui; un risultato direttamente correlato ai dati utilizzati per il loro addestramento.

Il fenomeno appare molto chiaramente nel contesto della selezione del personale, in cui l'utilizzo di sistemi di machine learning per automatizzare la valutazione dei curriculum vitae determina spesso uno svantaggio di genere, mostrando in modo sistematico una preferenza per i candidati maschi rispetto alle femmine.

È una tendenza attribuibile all'utilizzo dei dati di addestramento. Poiché i modelli di machine learning apprendono dai dati forniti durante la fase di addestramento, nel caso specifico della selezione del personale il dataset comprende una grande quantità di curriculum vitae di

candidati passati o attuali. Ma se questi dati riflettono differenze storiche di genere, come una prevalenza di candidati maschi assunti, il modello incorporerà tali schemi di disuguaglianza. E per conseguenza, ripeterà sistematicamente una preferenza per i candidati maschi, a svantaggio delle donne.

Una situazione simile si verifica anche nelle decisioni legali, in particolare nelle procedure per determinare le condizioni di libertà vigilata agli autori di reati. Qui, l'utilizzo di dati storici può generare una forte discriminazione verso le persone di colore.

Il riconoscimento facciale

Il riconoscimento facciale è stato oggetto di critiche a causa della sua inesattezza, e in alcuni casi ha portato ad azioni legali⁶. Questa tecnologia si basa su algoritmi di intelligenza artificiale che vengono addestrati per identificare e confrontare volti in diverse situazioni. Nonostante siano addestrati su enormi dataset di immagini, in alcune circostanze questi algoritmi possono commettere errori nell'identificare correttamente le persone o nel decodificare le loro espressioni facciali.

Particolarmente preoccupante è l'impiego del riconoscimento facciale nei sistemi di proctoring online (OP), dove numerose segnalazioni indicano difficoltà nell'identificare accuratamente le tonalità più scure della pelle. Questo errore sistematico può avere gravi conseguenze, soprattutto per gli studenti appartenenti a gruppi razziali con tonalità cutanee più scure⁷. Potrebbero incontrare ostacoli aggiuntivi durante gli esami online che dipendono da questa tecnologia o addirittura essere esclusi dall'accesso agli esami a causa di errori di identificazione. È una situazione che suscita serie preoccupazioni riguardo all'equità e al rischio di discriminazione derivanti dall'utilizzo di

tali tecnologie.

Un'altra forma di pregiudizio può insorgere nel rilevare il comportamento "normale" o "accettabile" durante un esame. Infatti, i sistemi di proctoring sono progettati per segnalare atteggiamenti sospetti o non conformi, come gesti insoliti o movimenti inusuali degli occhi. Tuttavia, le persone con disabilità o neuroatipiche potrebbero comportarsi in modi che non rientrano nella norma stabilita da questi sistemi⁸.

Il ruolo degli esaminatori

Se è vero che l'accumulo di dati di addestramento sempre più ampi dovrebbe migliorare l'affidabilità dei sistemi OP, è importante notare che il pregiudizio e l'inesattezza molto probabilmente non saranno mai completamente eliminati⁹.

L'ingiustizia non deriva sempre dal pregiudizio del machine learning. Le aziende OP sottolineano che alla fine è l'istruttore a dare un giudizio definitivo sull'integrità o disonestà dello studente. Tuttavia, questo solleva interrogativi sul ruolo cruciale degli istruttori. Alcuni potrebbero avere una fiducia eccessiva nelle segnalazioni del software, come quelle che riguardano movimenti "sospetti" della testa.

La valutazione di eventi lampanti, come quando chi sta sostenendo l'esame viene sostituito da un'altra persona, è relativamente facile, ma quella legata a segnalazioni più sottili, come "voci poco udibili, leggere variazioni di illuminazione e altri segnali comportamentali" potrebbe essere molto più complessa. Con l'avanzamento della tecnologia di machine learning, probabilmente verrà attribuito ancora più peso epistemico (una maggiore fiducia) ai giudizi dell'IA, e questo potrebbe aumentare i rischi di ingiustizia.

NOTE

1. S.A. Ginder, J. E. Kelly-Reid, F. B.

Mann, *Enrollment and employees in postsecondary institutions, Fall 2017; e Financial Statistics and Academic Libraries, Fiscal Year 2017: First Look (Provisional Data)*, in «National Center for Education Statistics» 2019.

2. C. Flaherty, *Online proctoring is surging during COVID-19*, apparso sul sito insidehighered.com, 2020.

3. N. Selwyn, C. O'Neill, G. Smith, M. Andrejevic, X. Gu, *A necessary evil? The rise of online exam proctoring in Australian universities*, in «Media International Australia» 107, 2411-2502, 2021.

4. Examity, *Auto proctoring*. Examity, <https://examity.com>, 2020.

5. ProctorU, *Harnessing the power of artificial intelligence to improve online proctoring*, www.procturu.com, 2020.

6. J. Peters, *IBM will no longer offer, develop, or research facial recognition technology*, apparso sul sito theverge.com, 2020.

7. M. Chin, *ExamSoft's proctoring software has a face-detection problem*, *The Verge*. apparso sul sito theverge.com, 2021.

8. S. Swauger, *Software that monitors students during tests perpetuates inequality and violates their privacy*, *MIT Technology Review*, apparso sul sito technologyreview.com, 2020.

9. A. D. Selbst, D. Boyd, S. Friedler, *Fairness and abstraction in sociotechnical systems*, in «Proceedings of the Conference on Fairness, Accountability, and Transparency» 19, pp. 59-68, 2019.

Tratto da: S. Coghlan, T. Miller, J. Paterson, *Good Proctor or "Big Brother"? Ethics of Online Exam Supervision Technologies*, in «Philosophy & Technology» 34:1581, 2021.

Traduzione di Francesca Nicola.

Simon Coghlan, Tim Miller

insegnano presso la School of Computing and Information Systems e il Centre for AI and Digital Ethics (CAIDE) della University of Melbourne, Australia.

Jeannie Paterson

insegna alla Melbourne Law School, University of Melbourne, Australia.

Hacking IA

Una prospettiva hacker sull'“Intelligenza Artificiale”: come funzionano i Grandi Modelli Linguistici? Dall'inquinamento dell'infosfera all'e-proctoring, vie d'uscita e percorsi possibili.

di Agnese Trocchi e Carlo Milani

Risoluto, il Professore disse:
Non credo proprio che possa farcela. Voglio dire, a sostituirsi a me, per scrivere una recensione decente. Insomma, va bene, non sarò un genio, ma sono pur sempre meglio di una macchina!

Il Professore ostentava convinzione, eppure nella sua voce si coglieva una certa qual esitazione, che tradiva un dubbio: e se... e se l'Intelligenza Artificiale fosse davvero stata in grado di scrivere una recensione meglio di lui?

Abitava l'ultima casa del paese, abbarbicata sulle pendici terrazzate di una frazioncina del ponente ligure. Confinava con il bosco. Fino a quel momento insieme ai vicini si era organizzato per difendersi da cinghiali, cervi e altri scroccoli di frutti dell'orto, non certo da macchine che minacciavano di rubargli l'impiego o quantomeno di rimpiazzarlo come recensore. La sua specialità erano i saggi di antropologia, sociologia e così via.

Eravamo ospiti per una notte e quella mattina tirava troppo vento per scendere al mare. Così tiravamo in lungo e, fra una chiacchiera e l'altra, immancabile si era aperto il vaso di Pandora delle cosiddette Intelligenze Artificiali, senz'altro non intelligenti in senso tradizionale e nemmeno molto artificiali, vista l'enorme massa di sforzi umani necessari per farle interagire¹. Allora andavano particolarmente di moda i Grandi Modelli Linguistici (LLM, *Large Language Models*) come i modelli della serie GPT (*Generative Pre-Trained Transformer*). Allenati su enormi masse di dati pescati qua e là sul web, in spregio alle altrimenti onnipresenti e





Foto Wikicommons.

draconiane leggi per la difesa a oltranza del copyright, i modelli delle multinazionali (OpenAI/Microsoft, Google, Meta) si basavano sulla cosiddetta “attenzione”. Niente a che vedere con l’attenzione umana: si tratta di una funzione matematica, una moltiplicazione. In breve, la classe di modelli denominata “trasformatori” (*transformers*) elabora sequenze di input (ad esempio le sequenze di parole in un testo) in maniera parallela, “pesando” (questi pesi sono coefficienti numerici che vengono moltiplicati fra loro) l’importanza di ogni parola all’interno della sequenza. In questo modo la comprensione del linguaggio naturale, la traduzione da una lingua all’altra, la generazione di testi anche piuttosto lunghi, di riassunti e via dicendo risultava molto più rapida e azzeccata rispetto ai precedenti modelli di apprendimento automatico. Insomma, il prodotto scalare...

Ho capito, cioè non ho capito, ma non importa. Quindi secondo voi funziona.

Il Professore si stava scocciando. Come dargli torto: non è facile seguire una spiegazione semplificata su un argomento del genere, a maggior ragione quando viene proposta da qualcuno come noi che di matematica non capisce poi molto.

Però da una prospettiva hacker si può pur sempre vedere le cose in maniera diversa.

Non è che funziona. Che vuol dire, funziona? Non è in grado di preparare una tisana, ma potrebbe spiegare come farlo a un agente umano, ecco... pur non avendolo mai fatto, certo. E senz’altro potrebbe scrivere una recensione di un libro!

Ma se non l’ha mai letto, il libro! Come fa, inventa?

Inventare non è il termine giusto. Non ha fantasia, un LLM. Però se “inventare” deriva da “invenio”, che in latino significa trovare, possiamo dire che è in grado di trovare una parola giusta dopo l’altra, dove giusta significa “quella con la maggiore probabilità di comparire in compagnia delle altre parole che compongono quella sequenza”. E a quanto pare la probabilità è sufficiente per determinare una sequenza di parole automatica che, agli occhi di un agente umano, hanno un senso compiuto. Dai, proviamo, che ti costa? Un po’ a malincuore, un po’ incuriosito,

il professore si lasciò convincere. Finalmente dopo tante chiacchiere si smantava un po’!

Avevamo a disposizione il portatile con cui scriveva le recensioni e tutto il resto: articoli originali, libri interi, interventi per convegni e festival. Un vecchio Mac ricevuto come compenso per un lavoro. Una connessione internet scadente. L’accesso a ChatGPT, il LLM targato OpenAI, era appena stato messo a pagamento, quindi era fuori discussione. D’altra parte, Bard, LLM di Alphabet (Google), in quel periodo non era accessibile dal territorio italiano per questioni di privacy ancora in discussione presso le alte sfere. Come fare?

Dai, lasciamo perdere, scrivo la recensione e basta, devo solo leggere il libro, sono trecento pagine buone...

Optammo per una VPN (Virtual Private Network). La nostra preferita era ovviamente TOR, The Onion Router, nella versione semplice del browser. Un’operazione ripetuta talmente tante volte da sfiorare l’automatismo: da torproject.org scaricare il browser adatto al sistema operativo in uso. Installarlo. Lanciare il programma TOR Browser, una versione modificata di Firefox per navigare nel Dark Web, ma, anche, nel web superficiale facendo in modo che la connessione appaia originata da un’altra località. Dopo qualche tentativo, ricaricando diverse volte il sito di Bard, finalmente ci lasciò entrare: la connessione risultava originata dagli Stati Uniti. Necessario un account gmail: nessun problema, ne avevamo uno condiviso per necessità simili.

Ora si trattava di porre la faticosa domanda. «Scrivi una recensione del libro», digitò il professore.

La stellina di Bard girellò per qualche attimo, infine l’oracolo rispose con poche righe striminzite.

Fa schifo. Non ditemi che questa è una recensione!

Il Professore sembrava sollevato. Beh, c’era da lavorare un po’ sul prompt, sulle istruzioni iniziali. Sugerimmo di modificarle aggiungendo un po’ di dettagli sullo stile, il pubblico di riferimento, la tipologia di pubblicazione, le modalità argomentative. «Scrivi una recensione di —, saggio di antropologia pubblicato in lingua italiana nel 2022, tradotto dal francese, tratto

dai quaderni di campo dell'antropologo —, facendo riferimento sia alla sua formazione classica, sia alla sua ricerca in ambito sociostorico, con particolare attenzione al gesto filosofico quasi postmoderno implicito nella trattazione», digitò il professore. L'oracolo macinò in pochi secondi una recensione di due cartelle scarse, senza errori ortografici, piuttosto piatta ma decisamente più convincente della precedente.

Non ci posso credere. Funziona! Certo va sistemata ma... potrei scrivere più di una recensione al mese. Anche cinque magari!

Il Professore era passato all'entusiasmo. Poi si oscurò.

Un momento. Ma allora, in miei studenti... Anche loro lo useranno! Che cavolo! Per i compiti a casa! Come faccio adesso?

Dalle stelle alle stalle?

Un testo generato da un LLM prevede un input, in genere un'iniziativa umana, per quanto si può benissimo programmare la generazione di input automatici; d'altra parte, anche la fruizione del testo (o dell'immagine, o del video, o dell'audio) implica un intervento umano, dal momento che non avrebbe senso per una macchina. Una volta affinato l'input i risultati sono stupefacenti e meravigliano anche il Professore. Molte persone esprimono addirittura risentimento quando scoprono che persino algoritmi semplici come il "vicino più prossimo" (*nearest neighbour* o *K-NN*) sono capaci di calcolare con notevole precisione le "libere scelte" di agenti umani, come accade nei sistemi di raccomandazione per gli acquisti. Ora, grazie ai *transformer* e al loro meccanismo dell'*attenzione*, gli algoritmi di apprendimento automatico possono gestire un numero straordinariamente grande di parametri in maniera efficiente. Il risentimento dell'umano spesso cresce di fronte a sistemi letteralmente in grado di calcolare la (probabilità della) parola successiva in un testo, in modo che l'insieme delle parole prodotte risultano essere frasi di senso compiuto dal punto di vista umano. La creatività ridotta a calcolo probabilistico, tramite regressioni lineari e logistiche, classificazioni e iterazioni.

Eppure, a uno sguardo più attento, magari sostenuto da una ricerca sull'ar-

gomento, o in questo caso dall'effettiva lettura del testo di cui è stata generata la recensione, non di rado si scoprono "errori" surreali, affermazioni deliranti, derive semantiche. Vengono chiamate "allucinazioni" della macchina, termine quanto mai scorretto. Infatti, per un agente umano (come per qualsiasi altro organismo vivente) un'allucinazione è una percezione indipendente da uno stimolo esterno; invece un agente meccanico come un LLM non ha percezioni proprie, ma elabora in maniera probabilistica degli input, ovvero i prompt testuali, i suoi "stimoli esterni".

Il professore attento, o lo studente scrupoloso, possono individuare queste derive e sistamarle rimettendo mano sul testo generato da un LLM. Ma è facile che l'attenzione dell'umano se ne faccia sfuggire qualcuna. Gradualmente queste "deviazioni", deliri delle macchine, potrebbero diventare dati di base per nuovi allenamenti dei futuri sistemi, oltre a essere spacciate come informazioni; potrebbero quindi modificare l'infosfera in un modo tale che non sarà più possibile distinguere il vero dal falso, o comunque sarà difficilissimo, richiederà tempi lunghi, analisi minuziose, grande potenza di calcolo e sistemi forensi sempre più complessi. Dalle stelle della creatività umana alle stalle di una lotta tra macchine, fra quelle che delirano e quelle dovrebbero individuare i deliri, con umani inebetiti che cercano vanamente di destreggiarsi, di imbrogliare e arrabattare domande per condizionare sistemi che non comprendono.

Ma è stato mai veramente possibile distinguere il vero dal falso? Noi a C.I.R.C.E.² riteniamo che sia una questione relativamente poco interessante, e che con la diffusione dei *deepfake*³ sia più costruttivo concentrarsi sugli aspetti relazionali: da chi vengono le informazioni? Come vengono generate? In che modo influiscono sul nostro umore? Perché facciamo circolare alcune informazioni (testi, immagini, video, audio) e non altre? La nostra proposta si chiama *pedagogia hacker*, una definizione un po' pomposa per descrivere una serie di pratiche che condividono un metodo rigoroso ma non serio, perché imparare a *stare nella turbolenza* di questo mondo insieme a delle macchine amiche è possibile, ma non banale. Richiede impegno, passione, voglia di mettersi in gioco.



↑ Immagine creata da Midjourney con il prompt “La scuola distopica del futuro”.

Giochi hacker con le IA

Certo, anche gli studenti useranno gli LLM, anzi, li stanno già usando da un pezzo. Così come tante altre persone. Da hacker, persone amanti delle tecnologie, desiderose di ridurre l'*alienazione tecnica*⁴, la questione diventa: quali LLM? Esistono sistemi liberi, non sviluppati da multinazionali avidi, rispettosi della privacy, ispezionabili? È possibile creare sistemi non esosi dal punto di vista dell'energia necessaria per allenarli e per renderli disponibili? Se le risposte a queste domande sono decisamente “NO”, o tendenzialmente “NO”, è difficile immaginare che le IA possano essere partecipi di giochi hacker.

Ci sono però senz'altro alcune cose che possiamo fare, come singoli e soprattutto come gruppi di affinità, per forzare l'evoluzione in una direzione a noi congeniale.

Innanzitutto, l'immaginazione. Non è un gioco appassionante senza immaginazione, capace di suscitare un immaginario condiviso. Nel gioco delle domande e risposte, della documentazione, dei riassunti, dell'enunciazione di regole, gli LLM possono dare risultati eccellenti, con prompt adeguati. Nella scrittura creativa, fantasia a briglia sciolta, storie che vale la pena di scrivere e leggere... no, davvero no. Forse non ancora, ma ricordiamolo: l'immaginazione non è semplice estrapolazione di probabilità.

Sì, ma il problema degli studenti rimane, ci ricorda il Professore. Vero. Vietare il ricorso a LLM in classe, durante gli esami? Le proibizioni non hanno mai risolto

nulla e anzi solitamente contribuiscono a esacerbare la voglia di avere a che fare con il proibito; tuttavia, accogliere a braccia aperte queste sedicenti novità nell'ambito scolastico e accademico non ci sembra un atteggiamento ragionevole. Non tanto perché renderanno le persone meno competenti, quanto perché sono sistemi industriali, inquinanti, chiusi, estremamente energivori, al servizio di padroni senza scrupoli che, dopo aver saccheggiato il web, ora chiedono agli Stati di legiferare per limitare la ricerca. Politiche di limitazione dovrebbero essere la regola, non l'eccezione, ma ormai i buoi sono lontani, l'avventura degli LLM nel mondo è cominciata e chiudere le stalle ora sarebbe una beffa.

Vietato vietare! Durante l'epidemia di COVID-19 molte istituzioni scolastiche e universitarie, specialmente negli USA, hanno adottato sistemi di controllo per sanzionare gli studenti che staccano gli occhi dalle telecamere dei loro computer durante una prova d'esame o persino durante le lezioni: pratiche basate su software proprietario, lesive della privacy, vessatorie e discriminatorie note come *e-proctoring*. Rimandiamo al sito bane-proctoring.com per una panoramica e approfondimenti in merito [e all'articolo di Simon Coghlan, Tim Miller, Jeannie Paterson contenuto nel Dossier di questo numero, dal titolo *Il proctoring: una risorsa o il Grande Fratello?*, N.d.R.].

Sempre in tema di divieti, che comportano spese ingenti, stravolgimento di procedure consolidate e sforzi continui per inseguire le innovazioni, i sistemi anti-plagio si stanno diffondendo a macchia d'olio. Ma ora, con l'introduzione degli LLM, l'unica possibilità è far ricorso ad altri sistemi analoghi in grado di individuare, con un certo grado di probabilità, se l'elaborato è farina del sacco dello studente o c'è di mezzo un LLM. Insomma, l'unica possibilità è affidarsi agli stessi sistemi che si vorrebbero bandire per smascherare le frodi.

Proponiamo allora, per chi ha voglia di giocare invece di vietare, un piccolo hack per scongiurare, almeno in alcune occasioni, l'impiego di LLM per rispondere a domande d'esame.

Facciamo ricorso a una delle nostre tecnologie conviviali d'elezione: il pad, lavagna via web. Il sistema può essere installato sui server della scuola, modificato opportunamente, oppure fruito su instal-

lazioni remote. A ogni studente viene fornito un link che conduce alla sua lavagna personale. Le regole del gioco sono semplici: l'elaborato va scritto direttamente sulla lavagna tramite tastiera. La lavagna, viene spiegato, è in grado di individuare se si tratta di un agente umano oppure no. Il copia-incolla di brani non è consentito: non è impedito a livello tecnico, ma viene considerato frutto di automazione.

Questo sistema è di semplice implementazione ed estremamente efficace. Stiamo lavorando a una versione appositamente programmata per questo caso d'uso, che verrà rilasciata pubblicamente sotto licenza libera entro la fine del 2023. Si prende in considerazione la genesi del testo, discriminando gli agenti in grado di digitare una lettera alla volta, tramite una tastiera, da quelli, come gli LLM, che non sono in grado di farlo e fanno "cadere" parole intere.

E se uno studente copiasse un elaborato automatico? Tutto è possibile, ma anche in questo caso, basterebbe un'occhiata alla genesi delle parole che compaiono sullo schermo per rendersi conto della prosodia, del particolare ritmo con cui compaiono. Gli umani non scrivono in maniera regolare, ma sincopata. Cancellano, ci ripensano, modificano. Gli LLM, invece, non esitano. Forse in futuro ci si prenderà la briga di programmare gli LLM per esitare, cancellare e modificare, ma per ora non sono in grado di farlo, perciò il pad è un'eccellente, semplice, robusta tecnologia conviviale, rispettosa della privacy, e quindi della legislazione europea in materia, il famigerato GDPR attualmente non rispettato da nessuno dei giganti dell'IT, tanto meno dagli LLM di loro produzione.

In questo senso, ci piacerebbe maggiore inventiva da parte di chi formula le domande per testare le conoscenze. Una scaletta per un articolo o per un libro, con i passaggi chiave da esplodere e indicando gli autori fondamentali a cui si farà riferimento, è molto più difficile da far produrre a un LLM rispetto a domande puntuali che richiedono di ripetere nozioni. Per far buoni discenti ci vogliono buoni docenti, disposti a esplorare insieme, in compagnia di tecnologie affini.

Le cosiddette IA, come ogni altra tecnologia, dicono molto di noi, del nostro modo di relazionarci con il mondo circostante, con gli esseri simili a noi e con quelli molto diversi. Nella relazione con l'Altro scopriamo chi siamo. Questa è la

pedagogia hacker.

Attendiamo le vostre idee e suggerimenti⁵!

NOTE

1. Si veda C. Milani, V. García, *L'evoluzione dell'Intelligenza Artificiale: dall'automazione del lavoro al condizionamento reciproco*, in «Mondo Digitale», 1-23, consultabile all'indirizzo <https://zenodo.org/record/7737430>.
2. Il sito è circex.org.
3. Si vedano i racconti di Agnese Trocchi *La prova provata*, 2021 <https://ima.circex.org/storie/8-bonus/c-buone-pratiche-deepfake.html> e *Free the Droids Now*, 2023 <https://circex.org/it/news/free-the-droids-now>.
4. Si veda C. Milani, *Tecnologie conviviali*, elèuthera, Milano 2022, <https://tc.eleuthera.it>.
5. Che potete inviare a info@circex.org.

Agnese Trocchi

è artista multimediale, scrittrice e formatrice. Curiosa da sempre delle nostre relazioni con le tecnologie analogiche e digitali, ha esposto in eventi e gallerie in tutto il mondo le sue opere di videoarte e di net.art. È Digital Communication Manager presso il Disruption Network Lab di Berlino. Dal 2017 collabora con il gruppo C.I.R.C.E. (Centro di Ricerca per le Convivialità Elettriche) con cui organizza e conduce laboratori di Pedagogia Hacker. Nel 2019 ha pubblicato con Ledizioni (Milano) il libro *Internet, Mon Amour*.

Carlo Milani

è saggista, traduttore e ricercatore associato presso l'ERTIM-INALCO di Parigi. Insegna Genealogia delle fonti digitali in vari licei e università tra Parigi, St. Gallen, Roma, Bologna, Parma e Bergamo, ed è membro del gruppo C.I.R.C.E. (Centro di Ricerca per le Convivialità Elettriche), per il quale tiene conferenze e corsi di formazione basati sulla pedagogia hacker. All'attività editoriale, di insegnamento e di ricerca affianca quella di informatico con alekos.net. Fino al 2018 ha pubblicato saggi sulla Rete e le tecnologie digitali sotto l'eteronimo di Ippolita, in particolare: *Open non è free* (Elèuthera, Milano 2005), *Luci e ombre di Google* (Feltrinelli, Milano 2007), *Nell'acquario di Facebook, l'irresistibile ascesa dell'anarcocapitalismo* (Ledizioni, Milano 2012), *La rete è libera e democratica... Falso!* (Laterza, Roma-Bari 2014), *Tecnologie del dominio* (Meltemi, Sesto San Giovanni 2017). Nel 2022 ha firmato con il suo nome di battesimo il libro *Tecnologie conviviali* (Elèuthera, Milano).

L'intelligenza artificiale nella didattica della filosofia

L'uso di ChatGPT per esercitazioni individuali e per esperienze di laboratorio di filosofia.

di Enzo Ruffaldi

L'intelligenza artificiale è un ambito di ricerca e di applicazione molto vasto e che va sempre più espandendosi: rientrano in questa definizione i robot per le pulizie o per tagliare l'erba, i programmi che tracciano un percorso per l'auto, mediante il gps, i programmi per il gioco degli scacchi o del Go, i programmi di automatizzazione della casa (domotica), quelli che guidano i robot industriali per il lavoro a catena, quelli della logistica utilizzati nell'e-commerce per selezionare e confezionare le merci da spedire, quelli per le auto a guida automatica ecc.

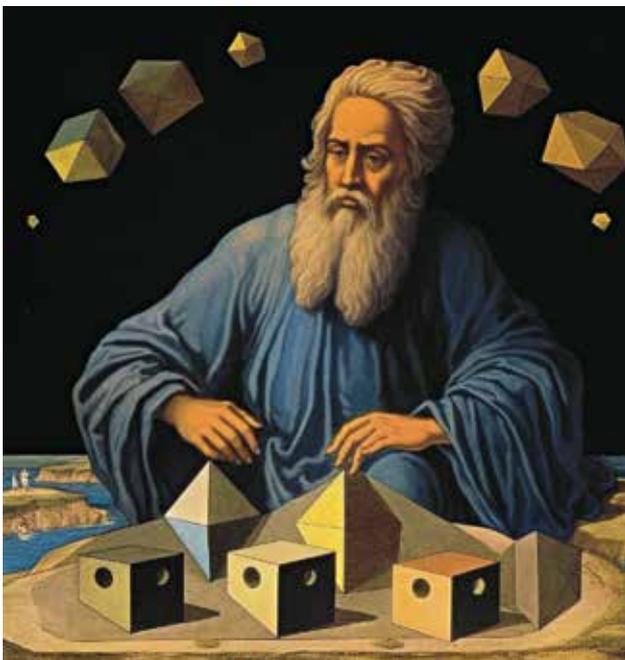
Quanto è intelligente l'IA?

Una prima importante considerazione per mettere un po' d'ordine in questa materia che sta espandendosi in modo tumultuoso è la distinzione proposta da Luciano Floridi tra una IA ingegneristica o riproduttiva, cioè di elaborazione dei dati e produzione di attività, dove i risultati sono notevoli (basti pensare, oltre agli esempi proposti in apertura, al riconoscimento del parlato o della scrittura a mano libera, o alla traduzione tra decine di lingue diverse) e una IA produttiva, o poetica, «come branca della scienza cognitiva» (cioè l'intelligenza umana creativa). In questo secondo ambito, scrive Floridi, «i sistemi che sappiamo costruire hanno l'intelligenza di un tostapane e non abbiamo davvero la minima idea di come migliorare la situazione, non foss'altro per il fatto che sappiamo veramente pochissimo sulla stessa intelligenza umana»¹. D'altra parte, basta considerare che un cervello umano "normale" può contare su circa 50 mila miliardi di connessioni neurali, mentre il sistema IA più complesso, secondo ChatGPT, ha 175 miliardi di parametri, che rappresentano i pesi delle connessioni neurali. C'è quindi una differenza di scala che impone, se non altro, di rimandare le attese per sistemi veramente intelligenti, come IA produttiva o cognitiva.

I chatbot

Limiterò la mia analisi soltanto al contributo che l'IA può dare alla didattica, in particolare della filosofia, e in questa prospettiva a un piccolo ambito dell'IA, i chatbot, cioè programmi in grado di interagire con un linguaggio naturale con interlocutori umani.

Parlando di interazione uomo-macchina, è d'obbligo un riferimento al test di Turing, uno dei padri dell'informatica. Com'è noto, egli nel 1950 propose un test per poter definire «intelligente» una macchina, il cosiddetto «gioco dell'imitazione», poi ribattezzato «test di Turing». Un essere umano avrebbe dovuto interagire verbalmente, mediante tastiera, con un partner posto in un'altra stanza, decidendo alla luce delle risposte date se il partner fosse un computer o un altro essere umano. Il test è superato se non è possibile distinguere in un numero predeterminato di casi la natura dell'interlocutore.



Ad oggi il test di Turing, ridefinito secondo parametri più stringenti², non è stato superato da nessun programma informatico, tanto che è in corso la sfida cui accenna il titolo citato in nota 2, la cui scadenza è fissata al 2029, tra chi scommette che per la fine di quell'anno sarà superato e chi è invece scettico al riguardo.

Neppure altri recenti chatbot, da ChatGPT a Bard AI di Google (compresi i vari assistenti vocali, da Siri ad Alexa a Google), possono perciò essere ritenuti «intelligenti» in senso produttivo-cognitivo, quindi è importante avvicinarsi a loro con spirito critico, suggerendo questo orientamento agli studenti, esplorando però, al tempo stesso, le possibilità che offrono in ambito didattico.

Alcune applicazioni didattiche

Da questo punto di vista, credo che dovremmo soprattutto guidare gli studenti all'esplorazione delle potenzialità di ChatGPT (e simili) e dell'aiuto che possono offrire nello studio e nella comprensione della materia. Restando fedeli a un atteggiamento critico, dovremmo indicare anche come mettere alla prova ChatGPT, chiedendo poi agli studenti di valutarne le prestazioni. Mi spiego con alcuni esempi, che potrebbero tradursi in indicazioni di esercizi o, meglio, in proposte di attività.

Primo esempio. ChatGPT offre discrete prestazioni nell'invenzione di racconti filosofici per illustrare il pensiero dei diversi autori. Se si chiede, ad esempio: «Inventa

un breve racconto filosofico per spiegare la filosofia di Leibniz», produce risultati interessanti. Si potrà invitare lo studente a individuare quali concetti sono stati posti in luce e quali no, e se il racconto è effettivamente aderente alla filosofia considerata, motivando il proprio giudizio. Da notare che, alla stessa domanda ripetuta, ChatGPT dà risposte diverse, inventando, nel caso specifico, racconti diversi. Potrebbe allora essere interessante confrontare, in un lavoro di gruppo, i racconti ottenuti da diversi studenti, per valutare quali siano più aderenti al pensiero di Leibniz, quali aspetti metta in rilievo ognuno, se, pur nella diversità, i concetti esemplificati sono simili, e così via.

Secondo esempio. ChatGPT produce anche definizioni corrette dei diversi concetti. Lo studente potrebbe chiedere a ChatGPT di individuare i concetti principali del pensiero di un determinato filosofo, accompagnandoli con una definizione sintetica, per poi confrontare il risultato con le proprie conoscenze. Anche in questo caso potrà poi valutare il risultato e discuterlo con i compagni. Qui si nota un comportamento interessante di ChatGPT: ripetendo la domanda, i concetti elencati sono sempre gli stessi, ma le definizioni cambiano, mantenendo comunque lo stesso significato. Anche in questo caso c'è materia per molte attività: controllare se l'elenco è corretto, se è completo, se le definizioni sono chiare, se alcune sono più chiare di altre ecc.

Terzo esempio. ChatGPT lavora molto bene con le argomentazioni. Se chiediamo di produrre argomentazioni a favore o contro una determinata tesi, propone un range abbastanza completo. Provate ad esempio a chiedere «Produci argomentazioni a favore dell'esistenza del libero arbitrio», e poi «Proponi argomenti a favore di una visione deterministica dell'agire umano»: avrete già le basi per un dibattito! Va ancora meglio su problemi più circoscritti e di attualità. Anche in questo caso possiamo chiedere agli studenti di valutare le argomentazioni, di completarle, di individuare eventuali debolezze ecc.

Da queste considerazioni si comincerà a intuire che ChatGPT può essere usata per esercitare le competenze tipiche della filosofia, cioè concettualizzare e argomentare, cui si aggiunge, come vedremo subito, problematizzare. Senza dimenticare, ovviamente, come competenza generale e generica, l'apprendimento di contenuti.

← Immagine creata da Midjourney con il prompt "Platone illustra l'idea dei solidi platonici".



↑
Due immagini create da Midjourney con il prompt "Pitagora e il concetto di metempsicosi".

L'uso di ChatGPT come strumento didattico apre nuove prospettive di attività e di interazione: lo studente può imparare ad usarlo come uno strumento critico per approfondire le proprie conoscenze, la comprensione delle diverse teorie, la loro applicazione a situazioni possibili (mediante i racconti filosofici e altre cose) e così via.

In conclusione, almeno per la didattica si aprono nuove possibilità che conviene approfondire, guidando gli studenti a esplorare questi strumenti, non a subirli.

ChatGPT come *personal trainer*

Dopo queste prime considerazioni, non sembra opportuno utilizzare i programmi di chatbot nella didattica come un possibile ausilio per l'insegnante nel compilare le prove di verifica. Sarebbe non solo riduttivo, ma fortemente controproducente, tanto più che, come abbiamo accennato, non sempre sono completamente affidabili.

Personalmente, penso invece a ChatGPT come a un *personal trainer*, un allenatore personale di ogni studente, che consente di realizzare una "palestra di filosofia in casa", dove possa esercitarsi programmando egli stesso attività da far costruire a ChatGPT (o a strumenti simili) per poi mettersi alla prova, con autoverifiche continue, senza stress ma in modo quasi ludico. Gli studenti sarebbero probabilmente stimolati a porre le domande "giuste" a ChatGPT per ottenere i risultati attesi,

confrontandosi spontaneamente con i compagni per ottenere dal programma determinate risposte. Potrebbero anche discutere tra di sé o con l'insegnante eventuali (ce ne sono) inesattezze o errori del programma, valutandone le prestazioni, o mettendo ad esempio in competizione due chatbot diversi nel produrre argomentazioni su un certo tema, discutendo su quali siano quelle giudicate migliori (giudicando, in realtà, anche la propria capacità di porre le questioni in modo più o meno efficace).

Proponiamo a p. 64 alcuni esempi di possibili usi didattici di ChatGPT (indicato con la sigla CG).

Piccolo laboratorio di filosofia

ChatGPT può diventare lo strumento per una serie di attività didattiche in parte proposte dall'insegnante o con lui (lei) concordate, in parte per attività liberamente svolte dagli studenti in collaborazione, in parte svolte individualmente a casa.

Esplorarne le potenzialità, capire che cosa possiamo far fare a ChatGPT usando il programma nel modo opportuno e formulando le "giuste" domande, potrebbe essere, credo, una sfida interessante per gli studenti. Coordinando bene questi tre livelli (attività di classe, di gruppo o di collaborazione spontanea, individuali) si potrebbero introdurre elementi laboratoriali nella didattica: non dico di trasformare la classe in un laboratorio di filosofia,

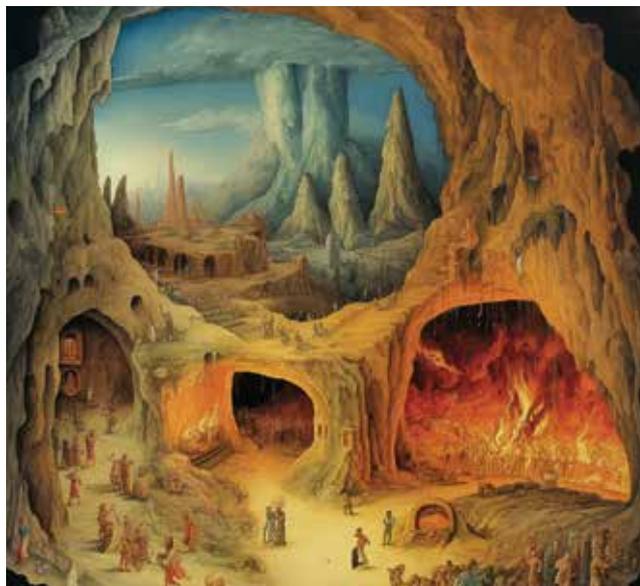
ma di introdurre attività - programmate o spontanee - che aumentino la partecipazione e la motivazione degli studenti nell'apprendimento delle competenze filosofiche. Si possono anche immaginare attività svolte individualmente a casa con ChatGPT e poi discusse in classe.

Uno spazio di lavoro online

Perché queste attività possano incidere davvero nella didattica, in modo sistematico e non sporadico, sembra però necessario creare uno spazio di interazione dove possano confluire, dove gli studenti possano confrontarsi tra loro e con l'insegnante, dove possano collaborare, discutere, organizzare dibattiti, scambiarsi esperienze e suggerimenti. Lo strumento che può coordinare tutte queste attività in modo ordinato e renderle parte organica della didattica di classe è la creazione di un sito di lavoro, accessibile e modificabile da parte di ogni studente, oltre ovviamente che dall'insegnante, che andrà strutturandosi progressivamente in aree differenziate (proposte di esercizi, dibattiti, lavori di gruppo, chat libera ecc.).

Quando insegnavo (sono in pensione da qualche anno) avevo per ogni classe uno spazio di discussione online al quale gli studenti partecipavano liberamente, da casa, per proseguire dibattiti iniziati in classe o per proporre altri, per scambiarsi idee e pareri su problemi filosofici o riflessioni su come la filosofia poteva aiutare a capire meglio la propria esperienza, per proporre materiali da utilizzare in aula con la LIM... La partecipazione non era obbligatoria e non comportava assolutamente voti o altre forme di valutazione. Eppure, nonostante che l'attività fosse svolta nel tempo non scolastico, la partecipazione era abbastanza elevata, a volte accesa.

In quest'ottica, ChatGPT può, a mio parere, arricchire la didattica e farla, per così dire, uscire dall'aula o renderla più motivata e più attiva in aula. Per lo studente, prepararsi le proprie prove di valutazione (autovalutazione), individuare i prompt adeguati per ottenere da ChatGPT una corretta analisi di un testo o la parafrasi di un brano difficile, o l'analisi critica di una tesi sviluppata dal filosofo che si sta studiando, aumenterebbe, credo, la motivazione allo studio, contribuirebbe a mettere a punto un metodo di autoapprendimento, spingerebbe ad approfondire i diversi



contenuti.

Alla libera attività in orario extrascolastico dovrebbero accompagnarsi sessioni laboratoriali a scuola in cui si usa insieme ChatGPT, si discutono gli errori che abbiamo rilevato nelle sue risposte, si apprendono insieme le strategie per formulare le richieste nel modo più adatto per ottenere le risposte che ci attendiamo.

Non ignoro né voglio trascurare l'altra faccia della medaglia, cioè i problemi che ChatGPT può produrre nella didattica, ma, pur con una certa prudenza, ritengo che sia importante conoscerne le potenzialità.

NOTE

1. L. Floridi, F. Cabitza, *Intelligenza artificiale. L'uso delle nuove macchine*, Bompiani, Milano 2021 (ed. digitale), in particolare il cap. 4, *Intelligenza artificiale: ingegneria o scienza cognitiva?*.

2. Si veda *Scommettere sul test di Turing*, in M. Mitchell, *L'intelligenza artificiale*, Einaudi, Torino 2022 (ed. digitale).

Enzo Ruffaldi

ha insegnato Filosofia e storia nei Licei fino al 2012, occupandosi fin dagli anni Novanta di didattica disciplinare, con numerose pubblicazioni: *Insegnare filosofia*, La Nuova Italia, Firenze 1999, *L'officina del pensiero. Insegnare e apprendere filosofia*, con M. Trombino, Led, Milano 2004, *Alla ricerca di nuovi modelli nella didattica della filosofia*, in F. De Natale (a cura di), *L'insegnamento della filosofia oggi*, Stilo Editrice, Bari 2003. Ha scritto articoli in riviste specializzate nella didattica della filosofia. È autore di manuali di filosofia con la Casa Editrice Loescher.

↑
Immagine creata da Midjourney con il prompt "Il mito della caverna di Platone".

Usi didattici di ChatGPT

1. Domande vero/falso

Inserisci il seguente prompt in CG:

- *Scrivi 10 domande di verifica del tipo V/F sulla filosofia di Spinoza.*

Ricopia le domande e incollale su una pagina di un word processor. Scrivi le tue risposte. Poi aggiungi in CG, sotto alle domande, il seguente prompt:

- *Scrivi accanto a ogni domanda la risposta corretta.*

Se poni a CG la stessa domanda più volte, i quesiti saranno sempre diversi e così potrai continuare a esercitarti. Copia sempre tutto in una pagina di un word processor, che potrai conservare o portare eventualmente a scuola per mostrarla all'insegnante.

2. Domande a risposta multipla

Inserisci il seguente prompt in CG:

- *Prepara una verifica del pensiero di Locke mediante 10 domande a risposta multipla, scrivendo per ogni domanda quattro risposte delle quali soltanto una corretta.*

Copia l'intero esercizio e incollalo su una pagina di un word processor, segnando le risposte a tuo parere corrette. Poi inserisci il seguente prompt:

- *Indica per ognuna delle precedenti domande la risposta corretta.*

Confronta le risposte con le tue, sottoponendo eventuali dubbi all'insegnante. Ricorda che CG è tutt'altro che infallibile: può essere molto utile, ma dobbiamo esercitare sempre il nostro controllo sulle sue operazioni. Ovviamente puoi ripetere questi esercizi per tutti i filosofi studiati e aggiungerne altri di tipo diverso.

3. Analisi del testo

Ricostruiamo le argomentazioni di un brano. Scriviamo il seguente prompt:

- *Individua per punti le argomentazioni presenti nel seguente brano.*

Subito sotto incolliamo il brano da far analizzare al programma. Se usiamo la versione di CG incorporata in Microsoft Bing, possiamo anche aprire il documento con il browser e chiedere a CG:

- *Individua per punti le argomentazioni presenti nel brano a fianco.*

4. Analisi critica

Svolgiamo adesso un esercizio diverso sullo stesso brano, l'analisi critica. Chiediamo a CG:

- *Analizza criticamente e commenta il brano precedente.*

L'analisi di CG riassume brevemente l'argomentazione presente nel brano, poi ne evidenzia possibili obiezioni.

5. Attualità

Per il brano analizzato, puoi porre anche la seguente domanda:

- *Quali sono le considerazioni ancora oggi attuali a proposito del brano precedente?*

In questo caso, può essere utile leggere le risposte di CG e discuterle in classe, o in un gruppo di lavoro, confrontando le posizioni di ognuno sulle questioni poste. Questo esercizio non è proponibile per ogni brano, ma in molti casi può alimentare dibattiti interessanti.

6. Dibattiti con ChatGPT

Chiediamo a CG di individuare le argomentazioni a favore o contro una tesi di rilevanza filosofica, ad esempio l'esistenza del libero arbitrio vs. il determinismo morale. Formuliamo la richiesta nei seguenti termini:

- *Produci argomentazioni a favore dell'esistenza del libero arbitrio, ricordando i principali filosofi che le sostengono.*

Il programma individuerà le posizioni più significative, in riferimento a diversi filosofi. Ad esempio, l'esperienza soggettiva come prova del libero arbitrio, ricordando Descartes e Kant; la responsabilità morale, che verrebbe meno se non vi fosse il libero arbitrio, e così via. Su alcuni accostamenti possiamo anche non concordare, ma risultano in genere corretti.

Chiediamo poi al programma:

- *Produci argomentazioni a favore del determinismo morale, ricordando i principali filosofi che le sostengono.*

Il programma cita il determinismo psicologico (Freud e Skinner), socio-culturale (Marx), (biologico (Dawkins),

culturale (Geertz) e filosofico (Spinoza). I filosofi ricordati dal programma sono semplici esempi senza pretendere di essere esaustivi, comunque le argomentazioni sono corrette, anche se non approfondite.

Gli spunti suggeriti da CG possono essere sviluppati dalla seguente attività:

- *Analizzate la validità, a vostro parere, degli argomenti proposti, poi dividetevi in due gruppi per discutere le due tesi contrapposte, sviluppando gli argomenti proposti da CG o aggiungendone altri.*

7. I concetti

Poni a CG il seguente spunto:

- *Elenca i principali concetti della filosofia di Leibniz.*

Il programma produrrà un breve elenco, con i principali concetti seguiti da sintetiche definizioni.

Attività:

- *Controlla i concetti proposti, individuando eventuali omissioni.*
- *Leggi le definizioni, valutando se sono corrette o meno e se sono complete o meno.*
- *Confronta le tue osservazioni con quelle dei tuoi compagni e delle tue compagne, all'interno di gruppi di lavoro composti da tre o quattro persone.*

Per i concetti più complessi o per quelli che non hai capito bene, oppure se non c'è accordo all'interno del gruppo, chiedete a CG di approfondire, inserendo ad esempio il seguente prompt:

- *Spiega in modo più approfondito l'identità degli indiscernibili.*

La definizione, inizialmente di poche righe, sarà molto più ampia e risolverà probabilmente alcuni dubbi.

Questa attività può essere svolta in aula, prevedendo anche discussioni e confronti sui concetti proposti, ma anche individualmente, a casa, come esercitazione e approfondimento delle proprie conoscenze. Data la centralità dei concetti nello studio della filosofia, è importante non soltanto comprenderli, ma anche esercitarsi sul loro uso, per padroneggiarli.

Una scuola...

diversamente digitale

**Impatti funzionali e disfunzionali
di robotica e intelligenza artificiale
nei futuri studenti e insegnanti.**

di Elvis Mazzoni e Stefania Cardinali

65

SCUOLA / Una scuola... diversamente digitale

Negli ultimi 10 anni, l'evoluzione tecnologica ha determinato importanti cambiamenti nella società in generale, e nella scuola in particolare, richiedendo agli attori coinvolti importanti "adattamenti" e "flessibilità" in termini di *upskilling* (migliorare le proprie competenze per svolgere al meglio il proprio lavoro) e *reskilling* (acquisire o modificare le proprie competenze per svolgere compiti o attività differenti rispetto a ciò che si svolgeva in precedenza). Il periodo caratterizzato dalla pandemia del virus Covid-19, ad esempio, ha evidenziato le difficoltà del sistema scolastico a reagire celermente e in modo efficace all'esigenza creatasi (essenzialmente l'introduzione della Didattica a Distanza e l'utilizzo di molteplici piattaforme web per attuarla), sia sul versante strutturale sia sul versante insegnanti e studenti, determinando di fatto un sensibile peggioramento (soprattutto nella scuola secondaria) dei risultati alle prove INVALSI. D'altronde, già nel 2019, prima dell'esplosione pandemica, il Consiglio d'Europa evidenziava la situazione critica della scuola italiana, mettendo l'accento soprattutto sull'abbandono scolastico e sulla carenza di competenze chiave e di base sia negli studenti sia negli adulti del paese¹.

Alle difficoltà descritte si aggiunge il gap generazionale che sempre più differenzia, dal punto di vista tecnologico, la generazione attuale da quella precedente e che, nel caso della scuola, ha visto una generazione di insegnanti tecnologicamente attivi con computer e email, e una generazione di studenti tecnologicamente attivissimi ma con smartphone, applicativi mobile, chat e social media. D'altronde, ogni generazione "adotta" quegli strumenti tecnologici che meglio si confanno (ovvero sono maggiormente funzionali) per raggiungere gli specifici obiettivi di sviluppo del periodo di vita e del momento storico che sta attraversando². Per i nati negli anni Settanta, Ottanta e i primi anni Novanta, il computer e l'email sono stati certamente gli strumenti che hanno rivoluzionato profondamente le pratiche e le attività svolte prima dell'avvento di tale tecnologia. A partire dagli anni 2000, smartphone, mobile app e, soprattutto, social media (e chat) hanno rivoluzionato in modo sostanziale il modo di interagire, comunicare e, più in generale, di intendere la socialità³ con importanti ricadute sui "confini" che classicamente definivano gli spazi del lavoro, dell'educazione e della vita privata/famigliare.

Il resto è storia recente, caratterizzata soprattutto dalla diffusione al grande pubblico di ChatGPT (e delle varie applicazioni di Intelligenza Artificiale), che non ha fatto altro che acuire il gap precedentemente evidenziato. Da un lato, alunni e studenti entusiasti dell'IA e, dall'altro, mondo della scuola e genitori sempre più preoccupati di qualcosa che si allontana dalle loro modalità operative e dal loro controllo.

Chiediamoci allora se, sulla base dell'esperienza vissuta con i cambiamenti tecnologici introdotti durante il COVID-19, si possa immaginare un inserimento dell'intelligenza artificiale (IA) all'interno della scuola che eviti o, perlomeno, limiti i pos-

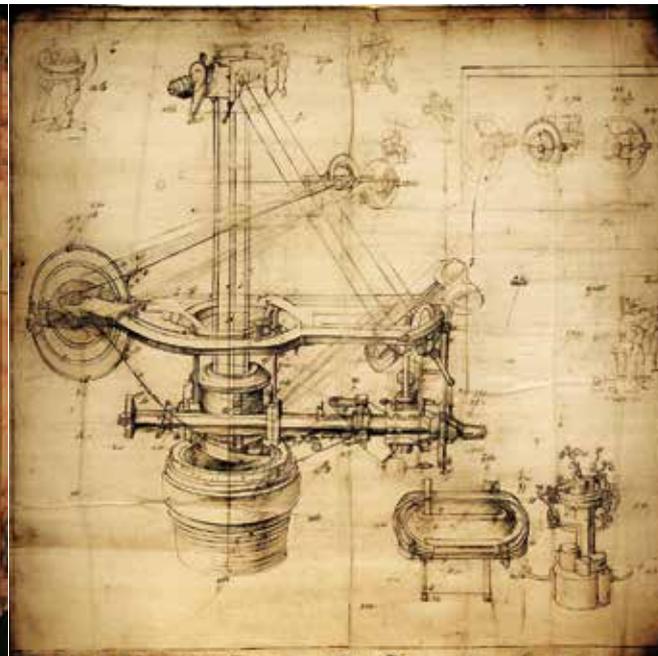
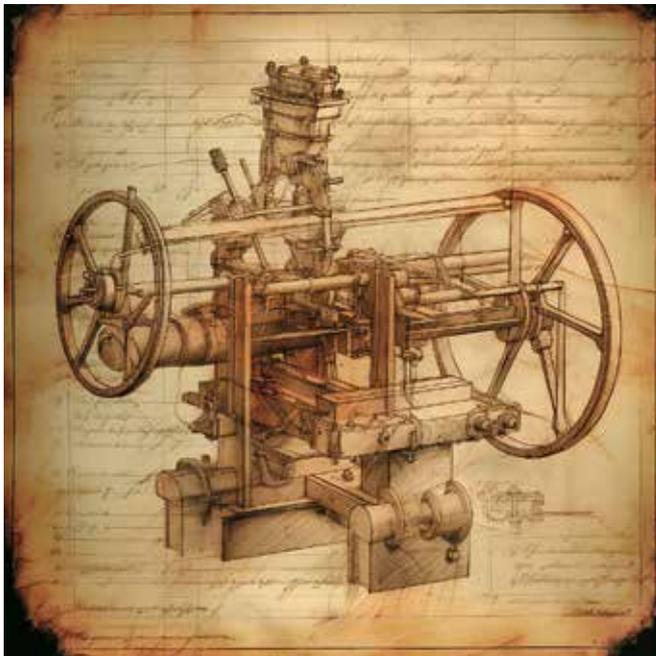
sibili effetti negativi e, soprattutto, che sia funzionale al raggiungimento degli obiettivi scolastici relativi allo sviluppo di conoscenze e competenze di alunni/studenti e insegnanti. Per rispondere a tale interrogativo, occorre partire proprio dall'aggettivo "funzionale" e "disfunzionale" che contraddistingue qualsiasi artefatto tecnologico⁴, ma ancor prima occorre introdurre il concetto di Zona di Sviluppo Prossimale⁵ ovvero il potenziale di sviluppo che caratterizza ogni persona.

Quando possiamo parlare di tecnologie funzionali e disfunzionali?

Analizzando il livello attuale di abilità o conoscenze di un individuo in un determinato compito e il livello potenziale che lo stesso individuo potrebbe raggiungere se, anziché lavorare da solo, vi fosse qualcuno (più esperto) a supportarlo, Vygotskij definì ciò che possiamo interpretare come potenziale di sviluppo (o di miglioramento) attuale che caratterizza ogni persona. Sebbene basato inizialmente sull'interazione fra esseri umani, Vygotskij era ben consapevole dell'importanza degli strumenti nello sviluppo cognitivo, sociale e culturale della specie umana, tant'è che Leont'ev (suo collaboratore) evidenziò che anche gli strumenti (più o meno materiali) giocano un ruolo importante in questo processo⁶. Definì così *artefatti funzionali* quegli "strumenti" funzionalmente integrati che rappresentano particolari risorse, interne ed esterne all'individuo, orientate a specifici obiettivi (ad es. l'apprendimento di una lingua). Se, da un lato, gli strumenti esterni (come un paio di forbici, gli occhiali, un'agenda) supportano e fanno da complemento alle abilità umane naturali costituendo un sistema più efficiente per il raggiungimento di un miglior esito, gli strumenti interni, come una teoria, un concetto o un'interpretazione, supportano la costruzione di soluzioni più evolute e creative e, dunque, il raggiungimento di un livello cognitivo superiore. Così, ad esempio, un paio di forbici, associate all'abilità umana del movimento delle dita, permette alla mano di divenire un "organo tagliente" che esplicita al meglio le sue potenzialità grazie all'integrazione strumento-abilità umana, mentre occhiali e lenti a contatto, associate all'abilità visiva (deficitaria), garantiscono una presta-

zione ottimale alla visione. Nel primo caso, uno strumento supporta un'abilità umana permettendole di fare qualcosa che altrimenti non sarebbe in grado di svolgere, o svolgerebbe ma con un risultato più scadente; nel secondo caso, gli strumenti compensano a una carenza umana e riportano la prestazione ad un livello ottimale altrimenti non raggiungibile. Nell'ottica di Kaptelinin⁷ e Zinchenko⁸ l'importanza degli artefatti è tale che si dovrebbe più propriamente parlare di *organismi funzionali*, ovvero strumenti che rappresentano estensioni fisiche e/o mentali del corpo o delle abilità umane. Da questo punto di vista, la cosiddetta intelligenza artificiale e la robotica possono essere reinterpretate come l'estensione mentale e corporea dell'uomo in grado di esaltare e sostenere varie abilità cognitive (come la concentrazione, il problem solving, la creatività) e fisiche (ad es. la precisione nei movimenti, la forza per spostare o muovere gli oggetti), in modo talmente intuitivo da divenire una componente naturale in molteplici attività quotidiane.

Da quanto sinora delineato, potremmo allora affermare che uno strumento tecnologico è tanto più funzionale quanto più supporta lo sviluppo umano nei compiti richiesti dal periodo di vita che sta attraversando, agendo sulla costruzione/aumento di conoscenze e, soprattutto, competenze. Al contrario, una tecnologia diviene disfunzionale non tanto quando sostituisce la persona nell'esecuzione di un compito (o meglio, non solo), ma soprattutto quando sostituisce le sue abilità nell'esecuzione di quel compito, rischiando paradossalmente di eliminarle o limitarle. Un esempio tratto dal mondo scolastico può aiutare a comprendere meglio tali processi. Dinanzi al compito di effettuare una traduzione dall'italiano all'inglese di un determinato brano, uno studente può agire in due modi: a) inserire il testo italiano all'interno di un IA translator (come Deepl o anche ChatGPT), attendere la traduzione e poi copiarla e utilizzarla per il compito richiesto; b) scrivere il testo in inglese all'interno dell'IA translator e controllare in itinere la corrispondente traduzione italiana per verificare che sia corrispondente all'idea iniziale. Nel primo caso, il sistema di traduzione è utilizzato per tradurre parti della propria madre-lingua in una lingua straniera non ben padroneggiata, sostituendosi di fatto allo studente e alla sua



abilità di traduzione. Tuttavia, in questo caso, è difficile immaginare come tale dinamica possa portare a un miglioramento della sua competenza nella lingua straniera, dato che non vi è alcuno scambio, negoziazione o riflessione sulla soluzione proposta. Nel secondo caso, al contrario, il sistema IA gioca un importante ruolo nel miglioramento del suo livello nella lingua straniera, in quanto lo studente non agisce all'interno della sua *comfort zone* di competenze (come nel primo caso), ma viene portato a riflettere su ciò che sta scrivendo e dunque a operare all'interno della sua *zona di sviluppo prossimale* (ovvero il suo potenziale di miglioramento), grazie al supporto di uno strumento tecnologico che lo spinge ad andare oltre le sue insicurezze con la lingua straniera. Talora il sistema potrebbe anche essere utilizzato per “scopi” che fuoriescono dalla funzione per cui originariamente è stato creato: è il caso di ChatGPT e dell'uso che molti ne hanno fatto come “supporto sociale” in determinati momenti della vita⁹. In tal caso, possiamo parlare del rischio di strumentalità inversa¹⁰, ovvero la tecnologia prende il sopravvento rispetto agli obiettivi per cui dovrebbe essere utilizzata, soprattutto a causa di carenze umane nella vita quotidiana con conseguente ricerca di compensazione¹¹ tramite sistemi tecnologici evoluti che “mimano” il comportamento umano.

Da quanto detto finora risulta piuttosto evidente che, per rispondere alla domanda sugli effetti funzionali o disfunzio-

nali che robotica e IA possono avere nei contesti scolastici, occorre focalizzare l'attenzione certamente sulle caratteristiche della tecnologia utilizzata, ma soprattutto sul modo in cui viene utilizzata nella didattica.

Caratteristiche funzionali di robot e IA per i contesti scolastici

Fra le caratteristiche della tecnologia utilizzata, in particolare per quanto riguarda i robot dotati di sistemi di IA, occorre considerare l'aspetto fisico. Sin dagli anni Settanta, varie ricerche hanno mostrato come generalmente l'essere umano sia più propenso ad accettare l'interazione con robot non antropomorfi rispetto a robot antropomorfi, in quanto questi ultimi provocano spesso reazioni di inquietudine e paura¹². L'IA sembra bilanciare tale fenomeno, nel senso che recenti ricerche mostrano che l'effetto “repulsione” viene tanto più mitigato quanto più il robot mostra di avere conoscenze e competenze importanti per ciò che si sta svolgendo¹³. Probabilmente, da un robot “meccanico” ci si aspetta che mostri importanti prestazioni dal punto di vista del movimento, della forza e della precisione nell'esecuzione (ad es. per quanto riguarda laboratori di materie tecniche), mentre da un robot avente sembianze umane ci si aspetta che abbia conoscenze e competenze tali per cui sia in grado di interloquire e interagire “alla

↑
Due immagini create da Midjourney con il prompt “Due macchine nello stile di Leonardo”.

pari” con un partner umano. Dunque, bilanciare adeguatamente l’aspetto più o meno umanoide con abilità e conoscenze è un primo punto da considerare per far sì che robot e sistema IA possano risultare funzionali ai processi didattici (e di sviluppo di conoscenze e competenze).

Il secondo punto importante è il modo in cui una tecnologia “interagisce” con l’essere umano: se l’IA viene lasciata libera, ovviamente, si comporta “liberamente” entro le possibili “soluzioni” che il database di cui è dotata mette a sua disposizione. Occorre infatti non dimenticare che l’IA risponde a sollecitazioni, richieste, domande o, più semplicemente, al programma (obiettivo) che “guida” le sue azioni. La specie umana, solitamente, non risponde semplicemente alle sollecitazioni contestuali, ma è proattiva e agisce anche sul contesto per modificarlo e adattarlo alle proprie esigenze (purtroppo, a volte, con esiti non sempre positivi). Dunque, come programmare al meglio l’IA per l’ambito educativo?

Le ricerche effettuate da quasi un secolo, sia nell’ambito della psicologia dell’apprendimento sia nell’ambito della psicologia dell’educazione e della formazione, hanno evidenziato importanti processi che possono essere messi in atto per far sì che vi sia apprendimento o che la didattica risulti più o meno efficace nel raggiungere i suoi obiettivi. Ne sono esempi gli schemi di rinforzo e le macchine per apprendere di Skinner in ambito comportamentale, la zona di sviluppo prossimale di stampo vygotkiano, l’importanza del pensiero divergente e del *networked flow* per i processi creativi¹⁴ o l’efficacia del conflitto socio-cognitivo per lo sviluppo cognitivo e il *problem solving*¹⁵. Sarebbe dunque opportuno far sì che, sulla base di ciò che si vuole attivare come processo di costruzione/sviluppo di conoscenze e competenze, il sistema di IA che supporta la didattica possa essere “regolato” da tali processi, associati al ruolo che lo stesso sistema, o il robot, ha nella relazione didattica: semplice strumento, partner, tutor¹⁶. Esempi alquanto riusciti, da questo punto di vista, sono state alcune sperimentazioni effettuate con robot umanoidi che, fungendo da partner, replicano il conflitto socio-cognitivo per l’apprendimento della lingua inglese¹⁷ per migliorare l’orientamento spaziale¹⁸ nei bambini della scuola materna e della scuola primaria. Altrettanto riuscite sono

alcune esperienze in cui un robot giocattolo funge da semplice strumento per la soluzione creativa di problemi complessi (trovare la via di uscita in un labirinto) in gruppi di bambini, adolescenti e adulti¹⁹, esperienze in cui sono attivate le tipiche dinamiche del pensiero divergente e del conflitto socio-cognitivo (comunicazione, divergenza, negoziazione, condivisione di una soluzione alternativa).

Utilizzo funzionale di robotica e IA nei contesti scolastici

Per quanto riguarda il modo in cui robotica e IA sono utilizzate nei contesti scolastici, per massimizzarne gli effetti funzionali per la costruzione di conoscenze e competenze, sia negli alunni/studenti sia negli insegnanti, è opportuno che non si sostituiscano alle loro conoscenze e abilità, ma che si integrino al meglio alle attività in essere, in modo da portarli a un avanzamento sia dal punto di vista cognitivo sia dal punto di vista delle capacità. L’IA è stata creata dall’uomo a immagine e somiglianza dell’intelligenza umana: per tale motivo anch’essa è fallace, sebbene sia dotata di una capacità di calcolo che le permette una velocità e capacità di elaborazione assolutamente estranea all’intelligenza umana. Quest’ultima però è proattiva, può riformulare obiettivi, scopi e tempi di attivazione rispetto a determinati compiti, agendo peraltro in modo sostanziale sul contesto fisico, sociale e culturale circostante. Per tale motivo, pensare a IA e intelligenza umana in termini “competitivi” non ha senso: ognuna ha peculiarità specifiche che non possono essere sostituite dall’altra. Far sì che i contesti educativi e didattici evolvano richiede, al contrario, che si consideri la possibile integrazione (organo funzionale) fra le capacità dell’intelligenza umana (e dell’uomo) e le capacità dell’IA (e del robot), in modo tale che l’IA supporti l’intelligenza umana per raggiungere modalità di pensiero cognitivamente più avanzate. Gli esempi precedentemente proposti evidenziano proprio questa possibilità dimostrando come IA e robotica, lungi dall’essere semplici strumenti che si inseriscono nei processi educativi e didattici, siano particolarmente versatili, flessibili, stimolanti e motivanti per adattarsi al meglio in tutti i gradi e le discipline del contesto scolastico²⁰.

A tali aspetti funzionali, si aggiunge

ovviamente la possibilità di avere un apprendimento personalizzato, in quanto, da un lato, l'utilizzo di robot e sistemi IA aventi caratteristiche e programmazioni diversificate permette di seguire e supportare esigenze specifiche degli alunni/studenti, ma anche da parte degli insegnanti nella costruzione della loro didattica; dall'altro, l'IA è capace di evolvere e adattarsi progressivamente in base all'utilizzo che ne viene fatto, plasmandosi sempre più sulla base delle esigenze dell'utente. Dunque, se, da un lato, robotica e IA risultano particolarmente efficaci e funzionali per far sì che i contesti scolastici siano sempre più inclusivi, grazie a tecnologie assistive e adattive particolarmente avanzate, dall'altro, risultano particolarmente funzionali per supportare gli insegnanti nel definire, costruire e testare percorsi di didattica alternativa e innovativa.

L'importante è porre le domande giuste, ovvero sviluppare la capacità di dare al sistema di IA le informazioni più rilevanti ai fini dell'obiettivo che si vuole raggiungere. E qui nasce, dunque, un'ulteriore competenza che l'IA richiederà in futuro ovvero come interagire con essa e, in particolare, come formulare una richiesta per avere una risposta che sia il più completa e informativa possibile.

Non è tutto oro ciò che brilla: chiediamolo a ChatGPT

—
Sinora abbiamo analizzato quali effetti funzionali IA e robotica possono avere all'interno dei contesti scolastici, ma non ci siamo ancora soffermati sui possibili effetti negativi. Per farlo, quale interlocutore migliore della diretta interessata, ovvero ChatGPT, alla quale abbiamo chiesto quali siano i possibili effetti disfunzionali dell'introduzione di robotica e IA all'interno dei contesti scolastici.

Dal punto di vista strutturale, soprattutto per quanto riguarda la robotica, occorre considerarne i costi e l'accessibilità, anche dal punto di vista della manutenzione, nonché la richiesta di competenze tecniche sempre maggiori in base a quanto sono sofisticati i robot utilizzati. Per gli insegnanti ciò si esplicita in una formazione specifica che permetta di integrare al meglio la robotica all'interno del curriculum ed evitarne il sottoutilizzo o l'inefficacia del metodo didattico a causa, ad esempio, di un'enfasi posta

sullo strumento anziché sul processo di apprendimento messo in atto (aspetti che sono stati ampliamenti evidenziati nei paragrafi precedenti).

Per quanto riguarda l'IA in sé, basandosi su dati derivati dall'esperienza umana (fallace), è possibile che le conclusioni cui giunge contengano errori o pregiudizi, proponendo importanti riflessioni dal punto di vista dell'equità e dell'etica.

Per quanto concerne il processo didattico e l'apprendimento, ChatGPT evidenzia alcuni importanti rischi legati all'utilizzo massiccio di tali strumenti tecnologici. Innanzitutto, una perdita delle interazioni umane (determinate ad esempio dall'interazione con robot o chatbot, come la stessa ChatGPT), con conseguente depersonalizzazione dell'apprendimento (che diverrebbe più "meccanico" e "isolato") qualora il processo implichi, ad esempio, processi di FAD senza interazione fra gli studenti e con l'insegnante. La mancanza di interazioni umane potrebbe anche determinare il rischio di un impoverimento dal punto di vista dell'intelligenza emotiva e dell'empatia. Infine, siccome i sistemi di IA tendono a standardizzare i processi e a "regolare" il proprio comportamento sulla base del comportamento "normalmente" adottato dagli umani, il rischio è quello di un abbassamento del livello di creatività e di problem solving.

Ecco perché diviene particolarmente importante costruire percorsi didattici in cui l'IA e la robotica, sulla base dei processi psicologici precedentemente descritti, si intersechino al meglio all'interno della zona di sviluppo prossimale di alunni/studenti e insegnanti in modo da supportarli nell'evoluzione del loro modo di pensare e di agire. E proprio ChatGPT ci sottolinea quanto questo processo sia importante:

Affrontare questi aspetti problematici richiede un'attenta pianificazione, considerazioni etiche e un approccio equilibrato all'integrazione dell'IA nei contesti educativi. È importante dare priorità al benessere, all'equità e allo sviluppo olistico degli studenti, sfruttando al tempo stesso i vantaggi che l'intelligenza artificiale può offrire. ... Integrare la robotica come strumento per migliorare le esperienze di apprendimento e sviluppare competenze come la risoluzione dei problemi, il lavoro di squadra e la creatività, piuttosto che semplicemente come

una vetrina tecnologica, può aiutare ad affrontare queste sfide.

E scusate se è poco!

NOTE

1. Cfr. G. Domenici, *Apprendimento scolastico, «denutrizione scientifica» e atteggiamenti novax*, in «Journal of Educational, Cultural and Psychological Studies (ECPS Journal)», n. 24, 2021, pp.11-24.
2. Cfr. E. Mazzoni, *Dallo sviluppo degli artefatti Web all'evolversi delle attività umane. I processi del cambiamento*, Morlacchi Editore, Perugia 2006.
3. Cfr. M. Benvenuti, E. Mazzoni, *Developmental Technologies. Evoluzione tecnologica e sviluppo umano*, Apogeo Education, Milano 2019.
4. Cfr. Benvenuti, Mazzoni, *Developmental Technologies* cit.
5. Cfr. L. S. Vygotskij, *Il processo cognitivo*, a cura di M. Cole, trad. it. C. Ranchetti, Boringhieri, Torino 1980.
6. E. Mazzoni, *Dallo sviluppo degli artefatti Web all'evolversi delle attività umane* cit.
7. V. Kaptelinin, *Computer-mediated activity: Functional organs in social and developmental contexts*, in «Context and Consciousness: Activity Theory and Human-computer Interaction», a cura di B. Nardi, MIT Press, 1996, pp.45-68.
8. V. Zinchenko, *Developing Activity Theory: The Zone of Proximal Development and Beyond*, in «Context and Consciousness: Activity Theory and Human-computer Interaction», a cura di B. Nardi, MIT Press, 1996, pp.283-324.
9. P. Carlbring, H. Hadjistavropoulos, A. Kleiboer, G. Andersson, *A new era in Internet interventions: The advent of Chat-GPT and AI-assisted therapist guidance*, in «Internet Interventions», 32, 2023.
10. H. Ekbia, B. A. Nardi, *Inverse instrumentality: How technologies objectify patients and players*. In P. Leonardi, B. Nardi, & J. Kallinikos, *Materiality and organizing: social interactions in a technological world*, Oxford University Press, Oxford 2012.
11. J. Zywicki, J. Danowski, *The faces of Facebookers: Investigating social enhancement and social compensation hypotheses; predicting Facebook™ and offline popularity from sociability and self-esteem, and mapping the meanings of popularity with semantic networks*, in «Journal of Computer-Mediated Communication», 14(1), 2008, pp.1-34.
12. M. Mori, K. F. MacDorman, N. Kageki, *The uncanny valley [from the field]*, in «IEEE Robotics & automation magazine», 19(2), 2012, pp. 98-100.
13. S. Liao, L. Lin, Q. Chen, *Research on the acceptance of collaborative robots for the industry 5.0 era--The mediating effect of perceived competence and the moderating effect of robot use self-efficacy*, in «International Journal of Industrial Ergo-

nomics», 95, 2023.

14. G. Riva, L. Milani, A. Gaggioli (a cura di), *Networked flow: Comprendere e supportare la creatività di rete*, LED Edizioni Universitarie, 2012.
15. Cfr. M. Benvenuti, E. Mazzoni, *Developmental Technologies* cit.
16. O. Mubin, C.J. Stevens, S. Shahid, A. Al Mahmud, J. J. Dong, *A review of the applicability of robots in education*, in «Journal of Technology in Education and Learning», v.1, 2013.
17. E. Mazzoni, M. Benvenuti, *A robot-partner for preschool children learning English using socio-cognitive conflict*, in «Journal of Educational Technology & Society», 18(4), 2015, pp. 474-485.
18. M. Benvenuti, E. Mazzoni, *Enhancing way-finding in pre-school children through robot and socio-cognitive conflict*, in «British Journal of Educational Technology», 51(2), 2020, pp. 436-458.
19. E. Mazzoni, M. Benvenuti, M. Orsoni, *Robotica e tecnologie per lo sviluppo: come costruire le competenze del futuro in La società dei robot*, Mondadori Education, Mondadori Università, Milano 2022, pp. 215-226.
20. A. Cangelosi, S. Di Nuovo, *La mente simulata* (e-book), Giunti, Firenze 2016.

Elvis Mazzoni

è professore associato di Psicologia dello Sviluppo e dell'Educazione all'Università di Bologna. È anche Consulente Scientifico e socio fondatore di Unveil Consulting Srl e di PsicheDigitale. Si occupa degli effetti delle tecnologie digitali, della robotica e dell'intelligenza artificiale sullo sviluppo umano nell'intero ciclo di vita. Ha pubblicato vari contributi a livello internazionale e nazionale e, insieme alla collega Martina Benvenuti, ha scritto il volume *Developmental Technologies: Evoluzione Tecnologica e Sviluppo Umano* edito da Apogeo Education.

Stefania Cardinali

è laureata in Giurisprudenza presso l'Università di Parma. Ha in seguito orientato il suo iter formativo in ambito Welfare, Diritto del Lavoro e HR. La passione e l'esperienza pregressa nella Comunicazione d'Impresa e nelle Pubbliche Relazioni l'hanno portata a collaborare con Unveil Consulting nella gestione e nello sviluppo dei progetti legati all'innovazione tecnologica e alla transizione digitale.

Due nuovi strumenti per l'inclusione e per la didattica laboratoriale: Bard e ChatGPT

Come l'intelligenza artificiale può aiutare il cambiamento verso il learning by doing e una valutazione formativa non numerica.

di Maurizio Maglioni

Che la didattica laboratoriale sia molto più efficace della didattica trasmissiva non credo abbia bisogno di essere ribadito. Il problema è che la gestione della didattica laboratoriale, in classi numerose, in presenza di alunni difficili, con poche attrezzature e in spazi angusti è, di frequente, impossibile. Osserviamo spesso, soprattutto nella scuola secondaria, un ripiego sul metodo trasmissivo standard fatto di spiegazione, studio a casa, interrogazione e una valutazione basata sul conteggio degli errori.

Questa sequenza tradizionale sappiamo quanto sia poco efficace a livello di inclusione e quanto generi ansia e frustrazione, soprattutto nei soggetti più deboli.

In questo articolo vi illustrerò come abbiamo affrontato questo problema risolvendolo anche grazie alle nuove conquiste dell'intelligenza artificiale.

Nell'anno scolastico appena concluso, nella scuola sperimentale capovolta che dirigo a Roma, Parco della Gratitude, abbiamo collaudato alcuni utilizzi delle intelligenze artificiali generative in classe. Abbiamo dotato tutti i ragazzi delle classi medie di Chromebook personali connessi in rete con una piattaforma web dotata di tutte le videolezioni necessarie allo studio di ogni disciplina. Su tale piattaforma ogni docente aveva pre-caricato tutte le schede operative delle attività laboratoriali proposte



→ Immagine creata da Midjourney con il prompt "La scuola del futuro".



↑
Due immagini create da Midjourney con il prompt "La scuola del futuro".

quest'anno.

Su ciascuna scheda viene indicato se l'attività va svolta singolarmente o in coppia o in gruppo, qual è la consegna, quanto tempo si ha a disposizione e una checklist di autovalutazione ben dettagliata e molto operativa.

Nelle consegne può essere richiesto qualsiasi prodotto finito: per esempio una relazione di 200 parole, una mappa concettuale con 10 parole chiave, una presentazione con 6 slide, un lapbook formato A3, un videoselfie di un dialogo in spagnolo, un doppiaggio italiano di un video in inglese, la ripresa video commentata di un esperimento di scienze ecc. I ragazzi hanno a disposizione il loro smartphone e tutti gli strumenti web necessari alla consegna, compreso ChatGPT.

Raramente abbiamo avuto bisogno di chiedere loro di smettere di usare il telefono per videogiochi: la consegna è sempre stimolante e stringente e se la coppia di lavoro deve consegnare prima del cambio dell'ora, nessuno si mette a giocare, perché il lavoro viene sempre valutato registrando non il voto ma la percentuale di lavori terminati.

Per evitare la ricerca di scorciatoie con Bard, abbiamo chiesto di utilizzare il web solo per l'autocorrezione: ad esempio, per alcune consegne si chiedeva di scrivere a mano, sul quaderno, quali correzioni erano state suggerite da ChatGPT. Con questo escamotage, i ragazzi sono stati costretti a impegnarsi nella crea-

zione di contenuti originali perché, dopo l'autocorrezione con la IA, dovevano mostrare all'insegnante ogni modifica, evidenziata con un colore diverso.

La vera innovazione quindi non è stata solo quella legata all'utilizzo dell'IA, ma è consistita principalmente nel cambiamento dei criteri di valutazione. Mi spiego meglio: nella maggior parte dei casi, gli insegnanti della scuola secondaria italiana valutano gli elaborati degli alunni utilizzando ancora un sistema antico ma molto semplice: partendo da 10, da 9 o da 8, sottraggono un voto o una frazione di esso per ogni errore commesso. Spero di sbagliarmi, ma vedo ancora utilizzare questo criterio, stigmatizzato da Freire, nella maggior parte dei licei italiani. Naturalmente gli alunni con bisogni educativi speciali, che commettono tanti "errori", spesso dello stesso tipo, sono i più penalizzati e a causa di queste valutazioni subiscono a volte crolli di autostima irreversibili.

Nella didattica laboratoriale, invece, gli errori commessi, scovati da Bard, evidenziati e autocorretti annotando la revisione, non influiscono sulla valutazione. La misura dell'impegno dell'alunno nel portare a termine la consegna avviene con le stesse modalità della vita reale: rispetto dei tempi, cura dei particolari, lavoro di squadra.

Ma quanto i ragazzi apprendono veramente l'italiano, l'inglese, la matematica, con questo metodo? Anche quest'anno tutti i nostri alunni, il 50% dei quali con

disturbi di apprendimento, sostenendo l'esame da privatisti hanno raccontato di non avere riscontrato difficoltà sia nelle prove scritte che in quelle orali. Noi crediamo che un grande elemento di differenza sia dato dalla potenza educativa del feedback delle IA.

Provo a chiarire meglio la differenza tra il feedback tradizionale e quello appena illustrato. Un alunno al quale si restituisce un compito in classe corretto, pieno di sottolineature a penna rossa, commenti e con un voto insufficiente, quanto sarà in grado di non ripetere gli errori evidenziati? Se ha sbagliato ortografia, sintassi oppure se non ha applicato una regola matematica, come potrà evitare di cadere nelle stesse trappole se non è lui stesso a riesaminare ogni errore e a provare la scrittura corretta? La nostra idea rivoluzionaria consiste nella modalità di scrittura del prompt: non si chiede a ChatGPT di riscrivere l'elaborato corretto ma di indicare solo quali parti sono da migliorare e come. Sul sito www.aiscuola.it troverete numerosi esempi di prompt di questo tipo.

Ecco un esempio di prompt di Bard:

Vorrei ricevere da te un feedback su questo mio saggio breve. Non devi assolutamente riscrivere il mio saggio corretto ma solo segnalarmi cosa potrei fare per migliorarlo e perché.

È incredibile la potenza di questo strumento che in pochi secondi produce 3 o 4 feedback estremamente precisi e chiari che l'alunno dovrà solo mettere in atto. John Hattie dice che «un feedback preciso e tempestivo può dimezzare il tempo di apprendimento degli studenti». Anche le nostre osservazioni confermano le metanalisi di Hattie.

Possiamo anche osservare che l'alunno-tipo, al ricevimento di un compito in classe corretto, dedica di norma solo pochi secondi allo studio delle correzioni, mentre con il nostro metodo è invitato a comprenderle a fondo e a produrre un nuovo testo definitivo.

Ma c'è un secondo enorme vantaggio: difficilmente un professore riesce a correggere più di un compito in classe al mese. Con Bard siamo riusciti a fare produrre agli alunni almeno un elaborato scritto a settimana per ogni disciplina. Seguendo il pensiero di Lev Vygotskij, che enfatizza l'importanza

dell'interazione sociale e del feedback nell'apprendimento, sfruttiamo l'IA per quadruplicare le occasioni di feedback.

Il terzo vantaggio, forse il più importante, riguarda l'autonomia. I ragazzi imparano a usare il più straordinario elaboratore di contenuti culturali umanamente immaginabile. Usandolo bene, non per copiare, imparano ad autocorreggersi e autovalutarsi.

Tutta l'esperienza accumulata quest'anno è stata inserita su www.aiscuola.it, il primo sito italiano che raccoglie le istruzioni dettagliate per l'utilizzo in aula e a casa dell'IA, per i docenti ma anche per gli studenti. Quest'ultima parte è stata redatta dagli studenti del terzo anno della secondaria di primo grado, ed è stata scritta con tutto l'impegno e la serietà di un compito autentico.

In conclusione, come abbiamo pensato di andare avanti dopo questi successi? Abbiamo già organizzato la prima convention a luglio scorso e stiamo preparando quella di febbraio 2024. Qui si prospetta una vera rivoluzione in termini di lesson plan, competenze attese e valutazione. Mentre i media si dibattono tra fantascienza e terrore di un mondo dominato dalle IA, noi invece facciamo ricerca-azione e proponiamo a tutti un modo interessante, più produttivo e soprattutto più inclusivo di fare scuola.

Maurizio Maglioni

è presidente dell'associazione Flipnet e dirigente della scuola sperimentale "Parco della Gratitude". Valutatore EFQM-CAF, è iscritto all'albo dei TQM Assessor. Da 10 anni promuove, con corsi e conferenze, la didattica capovolta e i compiti autentici nelle scuole di tutta Italia, ed è membro del FLGI Research, i 17 responsabili mondiali della ricerca in didattica capovolta. È autore di romanzi e saggi, tra cui *La classe capovolta*, 2014; *Capovolgiamo la scuola*, 2018 e *Il compito autentico nella classe capovolta*, 2019.

LE LINGUE MODERNE CON LOESCHER



webTV e webAUDIO

webTV

video in streaming **di inglese, francese, tedesco e spagnolo** con **sottotitoli e schede attività**: interviste, estratti da film, videolezioni e documentari in lingua

- ▶ Per la **flipped classroom**.
- ▶ Per la **didattica tradizionale**: spunti e approfondimenti **in classe**, ripasso individuale **a casa**.
- ▶ Per la **didattica CLIL** in inglese, francese, tedesco e spagnolo.



webtv.loescher.it

webAUDIO

▶ **contenuti audio** in streaming: esercizi, audiolibri, interviste, brani musicali

NOVITÀ PODCAST



I PORTALI

Portali dedicati a ogni disciplina con risorse **liberamente scaricabili**.

- | | |
|--|---|
| <input checked="" type="checkbox"/> francese | <input checked="" type="checkbox"/> tedesco |
| <input checked="" type="checkbox"/> inglese | <input checked="" type="checkbox"/> clil |
| <input checked="" type="checkbox"/> italiano per stranieri | <input checked="" type="checkbox"/> competenze |
| <input checked="" type="checkbox"/> spagnolo | <input checked="" type="checkbox"/> didattica inclusiva |



www.loescher.it/portali

QdR / Didattica e letteratura



La collana scientifica, dedicata a scuola e università, per riflettere su metodi e strumenti idonei a valorizzare il ruolo degli studi letterari, della scrittura, della lettura e dell'interpretazione delle opere.

DIRETTA DA

Natascia Tonelli

Simone Giusti

COMITATO SCIENTIFICO

Paolo Giovannetti (*IULM*)

Pasquale Guaragnella
(*Università degli Studi di Bari*)

Marielle Macé (*CRAL Parigi*)

Francisco Rico
(*Universitat Autònoma Barcelona*)

Francesco Stella
(*Università degli Studi di Siena*)



I libri pubblicati nella collana sono reperibili in libreria o presso le agenzie di zona. Indice e prime pagine sono disponibili sul sito de «La ricerca».



La collana QdR / Didattica e letteratura è anche online
www.laricerca.loescher.it/qdr-didattica-e-letteratura



Su Facebook:
La ricerca

Su Twitter:
@LaRicercaOnline



LA RICERCA È ANCHE ONLINE

Rivista e contenitore per dire, fare, condividere cultura

In contatto diretto e quotidiano scambio con i suoi lettori, per ampliare le prospettive, accogliere le notizie più attuali in tempo reale, arricchire il dibattito, captare e rilanciare nuovi argomenti. Il sito contiene i pdf e i singoli articoli di tutti i numeri de *La ricerca* cartacea, articoli di attualità, istruzione, cultura, la sezione *Scritto da voi*, un'area dedicata alle normative riguardanti l'istruzione e tutti i *Quaderni della Ricerca*.