



# SE L'IA ENTRA IN LABORATORIO

*L'avvento dell'IA nel campo della ricerca scientifica ha reso quest'ultima in qualche modo più accessibile e meno elitaria, consentendo tra l'altro un'accelerazione senza precedenti nel processo di raccolta, analisi e interpretazione dei dati, con conseguenze importanti, per esempio, nel campo della medicina e delle terapie personalizzate. Non mancano però i rischi, sia in termini di trasparenza, equità e autonomia della ricerca, sia rispetto al ruolo degli scienziati stessi.*

## SIMONE ROSSI

Nella sua *Poetica*, Aristotele l'avrebbe definito *καταστροφή* (catastrofe), ovvero un momento cruciale nella struttura narrativa della tragedia, che rappresenta il culmine della storia, quando si verifica un cambiamento improvviso, spesso drammatico, che – in quei contesti – di solito conduce a una conclusione fatale o al disastro. Certamente, l'avvento delle intelligenze artificiali (IA) una *καταστροφή* in senso storico, tecnologico, di epoca e di civiltà ce la sta facendo vivere dal vero, e non su un palcoscenico teatrale. Sta a noi, come genere umano, governare questo processo e scongiurare un finale da tragedia greca.

Il momento esatto di questa “catastrofe” è databile al novembre 2022, quando è stata resa disponibile a tutti la versione 3 di ChatGPT: è da allora che le nostre vite stanno cambiando e che siamo costretti a porci nuovi interrogativi, che solo qualche anno fa non avremmo nemmeno immaginato. Inizialmente definita «pappagallo stocastico»<sup>1</sup>, concetto successivamente ripreso, non senza un filo di disprezzo, anche da Noam Chomsky, l'IA si è subito vendicata, imponendo al genere umano la madre di tutte le domande: ma è proprio vero che l'essere umano è l'unico proprie-

<sup>1</sup> Emily Bender *et al.*, “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”, in *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, marzo 2021.

43



tario indiscusso del linguaggio e che, soprattutto, per parlare sia necessario pensare? I fatti sembrerebbero dire di no: nel senso che non possiamo più essere sicuri che sia così. L'IA parla bene, un po' stucchevole e perbenista in genere, ma lo fa correttamente e in tutte le lingue del mondo. E, soprattutto, non pensa (non può pensare), per parlare. Ma noi esseri umani, quindi, pensiamo? O meglio, è veramente necessario pensare per parlare, se un algoritmo, per lo più acefalo, lo fa come noi?<sup>2</sup>.

E, *mutatis mutandis*, d'ora in poi sarà necessario che gli scienziati continuino a pensare per fare una ricerca di ottima qualità o delegheranno i futuri sviluppi direttamente a IA, mettendo a riposo i propri network neurali dell'intelligenza naturale, della conoscenza, della creatività e dell'intuizione? Diminuirà l'unicità intellettuale, cioè il valore intrinseco del singolo ricercatore? Nessuno ha la sfera di cristallo per prevedere come andrà a finire questo processo, ma qualche considerazione possiamo farla.

#### ***Il valore dello scienziato: da creatore a curatore della ricerca?***

44

Di certo, lo sviluppo delle intelligenze artificiali sta trasformando profondamente il panorama della ricerca scientifica, ma già da più di un decennio il *machine learning* aveva iniziato a farlo, almeno in alcuni settori specifici. Questo cambiamento riguarda non solo l'accesso alle informazioni, la quantità di dati disponibili e la velocità con cui questi possono essere analizzati, ma anche il ruolo stesso dello scienziato e, in parte, la natura della ricerca, oltre alle implicazioni etiche legate all'uso di queste nuove tecnologie.

Finora, l'immagine comunemente condivisa dello scienziato è stata quella del principale creatore e interprete delle ipotesi sperimentali prima e dell'analisi dei dati successivamente. La ricerca classica ha spesso valorizzato le abilità individuali (o più raramente di gruppo) nel generare nuove ipotesi di lavoro o nuovi paradigmi di studio: come un vero e proprio "creatore di ricerca", lo scienziato – o almeno alcuni scienziati – hanno sovente dato origine a scuole di pensiero e di metodo, a volte suscitando anche vere e proprie rivalità scientifiche giunte persino alla ribalta dei media. Pensiamo alla tesa atmosfera fra Santiago Ramón y Cajal

<sup>2</sup> Per approfondimenti, si veda Riccardo Manzotti e Simone Rossi, *IO & IA. Mente, cervello e ChatGPT*, Rubbettino Editore, 2023.



e Camillo Golgi (qualcuno ha ipotizzato anche una vera e propria lite) durante la cerimonia per il conferimento del premio Nobel per la Fisiologia e la Medicina a Stoccolma nel 1906: Golgi era un sostenitore della “teoria reticolare”, secondo la quale il sistema nervoso è costituito da una rete continua di cellule, mentre Cajal era un fermo sostenitore di quella che a posteriori si sarebbe dimostrata più veritiera, la “teoria del neurone”, secondo la quale il sistema nervoso è composto da unità discrete, i neuroni, in grado di comunicare fra di loro. Oppure, nell’ambito della fisica, al dibattito fra Albert Einstein e Niels Bohr sulla meccanica quantistica: Einstein, nonostante ne fosse uno dei padri fondatori, era profondamente scettico riguardo alla sua interpretazione probabilistica. La sua famosa frase “Dio non gioca a dadi con l’universo” rifletteva la sua insoddisfazione per l’indeterminismo intrinseco della teoria. Bohr, invece, era un sostenitore della cosiddetta interpretazione di Copenaghen, che accettava l’indeterminismo e la probabilità come aspetti fondamentali della natura. Ma questi dibattiti, che forse non sarebbero nemmeno avvenuti se fosse stata già disponibile l’IA, hanno rappresentato delle vere e proprie pietre miliari per lo sviluppo della comprensione del sistema nervoso e della meccanica quantistica.

Ciò a cui stiamo assistendo è una progressiva perdita di questo ruolo guida prevalentemente individualistico della ricerca e della scienza, a favore di una transizione verso una figura di semplice “curatore” della ricerca: gli scienziati sono sempre più chiamati a progettare esperimenti e a interpretare i risultati prodotti dall’IA, piuttosto che a generare direttamente le ipotesi. Da creatore a curatore della ricerca, insomma, il passo è breve: un ruolo forse più comodo e anche più democratico, nel senso di più accessibile e meno elitario, ma non scevro di possibili pericoli e nuove responsabilità, come vedremo più avanti.

Storicamente, anche altri lavori hanno gradualmente perso il loro valore, in termini di capitale umano, a causa dell’evoluzione tecnologica: negli anni Sessanta, la figura del “contabile”, colui che teneva i conti, è stata soppiantata dalle calcolatrici; poi l’operaio della catena di montaggio è stato rimpiazzato da robot automatizzati; oggi i correttori di bozze stanno scomparendo dalle redazioni editoriali e così via. Nell’industria cinematografica, abbiamo recentemente assistito a una rivolta dei doppiatori di Hollywood, preoccupati per l’avvento degli algoritmi generativi, addestrati a riprodurre qualsiasi voce, in qualsiasi lingua del mondo. Vale la pena ricordare che l’acronimo GPT significa proprio *Generative*,

45



*Pre-trained, Transformer*: grazie all'enorme quantità di parametri immessi nell'algoritmo (P), che ora ha anche accesso diretto alla rete, e alla tecnologia Transformer (T)<sup>3</sup>, le IA possono generare (G) qualsiasi tipo di output venga loro richiesto, in pochi secondi. Questo cambiamento solleva ovviamente questioni fondamentali riguardo all'identità stessa dello scienziato. Cosa significa essere un ricercatore in un'epoca in cui le macchine che parlano possono formulare ipotesi, scrivere programmi e, in alcuni casi, scoprire nuove leggi della natura? La creatività umana, la capacità di estrapolare connessioni intuitive o divergenti e di contestualizzare i risultati in un quadro teorico più ampio, rimangono certamente competenze irrinunciabili, e rappresentano il valore intrinseco (questo sì, incommensurabile) del ricercatore in quanto persona. Però a esse si affiancano oggi strumenti che, se da un lato ampliano enormemente la portata dell'indagine scientifica, abbreviandone i tempi, dall'altro rischiano di omogenizzare – forse troppo – le modalità di ricerca. Concettualmente, niente di diverso da ciò che ha prodotto la globalizzazione nelle città del mondo, dove ormai per le strade del centro, da est a ovest e da sud a nord, si trovano negozi che sono riproduzioni all'infinito dello stesso stereotipo internazionale, localmente declinato solo in minime variazioni.

4  
6

### *Che ricerca con l'IA?*

Se, come abbiamo visto, la figura dello scienziato classico è in pericolo, non possiamo negare che l'avvento dell'IA nel campo della ricerca scientifica abbia reso quest'ultima in un certo senso più accessibile e meno elitaria, consentendo tra l'altro un'accelerazione senza precedenti nel processo di raccolta, analisi e interpretazione dei dati. In ambito neuroscientifico, ad esempio, nell'ultimo decennio lo sviluppo delle neuroimmagini è stato impressionante: è ora possibile migliorare la qualità delle immagini stesse, riducendone il rumore di fondo, ma anche integrare dati provenienti da più metodiche di indagine strutturale (risonanza magnetica nucleare) e funzionale (risonanza funzionale, magneto- o elettroencefalografia, tomografia a emissione di positroni), rendendo così i processi diagnostici e di ricerca più efficienti, accurati e personalizzati. Ed è oggi possibile anche “mettere insieme” analisi prove-

<sup>3</sup> Si veda l'articolo in cui viene definita e descritta la tecnologia Transformer: Ashish Vaswani *et al.*, “Attention Is All You Need”, arXiv, 2017.



nienti da migliaia di cervelli differenti, creando dei database a disposizione di tutti gli scienziati, utili sia per capire meglio il funzionamento del cervello sano, sia per comprendere ciò che cambia in condizioni patologiche. Con questo approccio, si stanno scoprendo veri e propri “indicatori” (*biomarkers*) di malattia, necessari per monitorare l’andamento della stessa o valutare l’efficacia di un trattamento, o individuare chi fra i pazienti affetti da una determinata patologia potrà beneficiare da un trattamento specifico. Il che, avendo la possibilità di saperlo in anticipo, è un bel vantaggio.

Quindi, algoritmi di *machine learning* e reti neurali avanzate sono oggi in grado di setacciare enormi quantità di dati, scovando pattern spesso così complessi che sarebbero stati difficilmente rilevabili dal solo occhio umano. Questo ha portato a scoperte più rapide e spesso inaspettate, in settori che vanno dalla medicina alla fisica delle particelle, dalla biologia alla climatologia. Due esempi su tutti: i *digital twins* e lo studio delle strutture proteiche.

I *digital twins* (o gemelli digitali) sono rappresentazioni virtuali esatte di oggetti, sistemi o processi fisici reali. Questi modelli fisici virtuali vengono creati utilizzando grandi quantità di dati raccolti da sensori e altre fonti per simulare il comportamento del gemello fisico quando è immerso nel mondo reale, tramite un’interfaccia utente. I gemelli digitali, ormai disponibili per numerosi campi di ricerca, dall’ingegneria industriale alla ricerca di nuovi materiali, dai dispositivi biomedicali alla gestione delle reti energetiche, possono essere utilizzati per monitorare, analizzare, ottimizzare e prevedere le prestazioni del loro corrispettivo fisico in tempo reale. I vantaggi dei *digital twins* sono quindi enormi in termini di riduzione di costi e tempi di sviluppo di prodotti o tecnologie, di ottimizzazione delle prestazioni del corrispettivo fisico o, nel campo medico, di personalizzazione di nuovi trattamenti “su misura” per i pazienti o per prevedere l’efficacia di una terapia specifica.

Una sfida particolarmente appassionante è rappresentata dallo sviluppo del *digital twin* del cervello umano. Una sfida di non poco conto, considerato che il cervello è il sistema più complesso dell’universo, composto da circa 60-80 miliardi di neuroni e un numero incalcolabile di sinapsi, che per di più varia in continuazione nel tempo in rapporto a ciò che l’individuo vive nell’ambiente, interno ed esterno, che lo circonda. Una sfida titanica, quindi, che possiamo però considerare intrapresa su scala mondiale grazie a progetti collaborativi come l’Human Brain Project o il Blue Brain Project, all’interno dei quali si cerca di ricostruire e

47



simulare dettagliatamente il funzionamento del cervello di piccoli roditori, fino all'obiettivo finale del cervello umano, creando modelli digitali in diverse scale, dalla singola sinapsi fra due neuroni fino all'intero cervello, per studiare la fisiologia delle funzioni percettive, motorie e cognitive, fino alla fisiopatologia delle malattie neurologiche (come Parkinson, Alzheimer o epilessia), sviluppando congiuntamente nuove tecnologie di ispirazione biologica. La sfida, dicevamo, è titanica, anche perché, a dire la verità, non è che come neuroscienziati conosciamo alla perfezione tutti i meccanismi che generano le cosiddette "proprietà emergenti" dei nostri network neurali, tipo l'attività onirica, la singolarità degli individui, i meccanismi del pensiero e della coscienza. Ma appare comunque chiaro che un *digital twin* completo e accurato (soprattutto se ben funzionante) del cervello umano potrebbe rivoluzionare le neuroscienze, permettendo ai medici (neurologi e psichiatri in primis) di testare trattamenti in un ambiente virtuale prima di applicarli ai pazienti reali, o di sviluppare terapie completamente nuove basate su simulazioni precise del cervello umano.

Per quanto riguarda lo studio delle strutture proteiche, i programmi AlphaFold, sviluppato da DeepMind, e il suo concorrente MetaAI, hanno reso disponibile gratuitamente alla comunità scientifica la struttura tridimensionale di oltre 200 milioni (secondo AlphaFold) e fino a 600 milioni (secondo MetaAI) di proteine. Questi dati comprendono praticamente tutte le strutture proteiche conosciute presenti negli organismi umani, nei batteri, negli animali, nelle piante, e altro ancora. Prima dell'introduzione dell'IA, la ricostruzione tridimensionale di una proteina, con le sue complesse sequenze di amminoacidi che si ripiegano per formare una struttura atomica tridimensionale, poteva richiedere mesi di lavoro. Qualsiasi ricercatore può ora prevedere rapidamente e gratuitamente le interazioni tra proteine o tra proteine e altre molecole, permettendo l'osservazione diretta dei processi biologici, anche quelli più complessi, nonché, ad esempio, l'effetto di nuovi farmaci o dei vaccini. Grazie a questi sistemi, sono stati recentemente chiariti alcuni meccanismi di resistenza dei batteri agli antibiotici ed è stato scoperto un enzima capace di degradare la plastica non riciclabile. Gli algoritmi complessi possono quindi simulare interazioni tra molecole e prevedere l'efficacia di nuovi composti, riducendo così la necessità di lunghe sperimentazioni iniziali su larga scala.



### *I rischi dell'IA nella scienza*

Se da un lato l'IA rappresenta un potente e oramai insostituibile alleato per la ricerca, dall'altro, come dicevamo, introduce anche rischi che non possono essere ignorati. La prima domanda che deve essere posta è: tutti i risultati prodotti dall'IA sono accurati, attendibili e riproducibili? Gli algoritmi di *machine learning*, per esempio, possono generare previsioni molto accurate, ma la loro "scatola nera" (provvista di reti neurali che generalmente auto-apprendono) può rendere difficile capire come queste conclusioni siano state raggiunte, sollevando problemi di trasparenza e riproducibilità, che sono i pilastri fondamentali del metodo scientifico. Non che prima delle IA non vi siano stati casi di mancata trasparenza o addirittura falsificazione dei risultati: è già accaduto in tutti i campi della scienza. Ma per il ricercatore contemporaneo, la fiducia acritica nelle risposte dell'IA ai suoi *prompt* deve essere necessariamente bandita. Anche perché, è bene ricordare, le modalità di funzionamento delle IA, che si basano essenzialmente su probabilità statistiche condizionate di tipo bayesiano che trasformano le parole in numeri e viceversa, sia in ingresso che in uscita, non consentono all'algoritmo operante nessun grado di comprensione di quello che viene prodotto, né tantomeno di porsi dubbi di correttezza o errore. In sostanza, a seconda di come viene posto il quesito e in che tipo di stringa numerica viene convertito, la risposta che può essere ottenuta (che deriva a sua volta dalla riconversione della stringa numerica in parole o immagini) può essere declinata dall'IA in un modo o nel suo esatto contrario, con estrema *nonchalance* e nessun senso di colpa.

Questa modalità operativa solleva, *ça va sans dire*, potenziali ma rilevanti interrogativi di tipo etico relativamente all'impatto sulla ricerca, così come sulla società in genere: le decisioni prese sulla base di modelli IA possono avere conseguenze significative, soprattutto se gli algoritmi utilizzati non sono progettati in modo equo o se riflettono *bias* preesistenti, fattori che possono portare a risultati distorti o a una sovra-semplificazione della realtà. Uno dei principali compiti dello scienziato contemporaneo, e che gli potrebbe restituire il valore dell'individualità, è quindi quello di assumersi la responsabilità di garantire che l'uso dell'IA sia guidato da principi etici robusti, che considerino non solo l'efficacia tecnica, ma anche l'equità, la trasparenza e, in ultima analisi, il pieno rispetto dei principi scientifici fondamentali e dei diritti umani.

49



L'altro compito che spetta al ricercatore è quello di cercare di preservare al meglio la funzionalità del suo cervello, in piena autonomia e consapevolezza. La mera, spesso eccessiva, dipendenza dagli algoritmi potrebbe infatti portare a un impoverimento del pensiero critico e alla riduzione della creatività nel fare ricerca: *use it or lose it*, dicono gli autori anglosassoni a proposito della funzionalità dei network cerebrali. La plasticità uso-dipendente delle nostre sinapsi, infatti, è bidirezionale, e comporta il rafforzamento dei circuiti costantemente utilizzati, ma anche l'indebolimento di quelli messi da parte per un po' di tempo. Ad esempio, le competenze linguistiche dei ragazzi di oggi, nativi digitali, sono molto ridotte rispetto a quelle delle generazioni precedenti, quando l'interazione umana avveniva di persona e non quasi esclusivamente attraverso una manciata di caratteri scritti su WhatsApp o sui vari social. Quasi più nessuno sa fare una divisione a tre cifre utilizzando carta e penna piuttosto che premere il tasto di una semplice calcolatrice. In compenso, le aree motorie che governano i movimenti dei pollici, utilizzati costantemente per chattare sulle tastierine dei cellulari, saranno probabilmente più ampie rispetto a quelle dei loro nonni. Se ne deduce che il ricercatore contemporaneo, magari già nativo digitale, non deve cadere nella tentazione di demandare tutto agli algoritmi; anzi, è auspicabile che, dal momento che la maggior parte del lavoro routinario può essere effettuato dall'IA (supervisionata), si ritagli più tempo a disposizione per accrescere le sue competenze che, ricordiamo, costituiscono il vero, unico valore della persona, specialmente quando sono inderogabili. Il tempo da dedicare a se stessi rappresenta esso stesso un valore, o un privilegio, incommensurabile. Infine, non è di poco conto il rischio che l'automazione del processo scientifico porti a una progressiva perdita di autonomia nella ricerca. Se le decisioni su cosa studiare e su come farlo saranno sempre più delegate a sistemi automatizzati, la libertà accademica potrebbe essere compromessa. Questo è particolarmente preoccupante in settori in cui la ricerca è strettamente legata a interessi commerciali o politici o, peggio ancora, militari, ambiti dove l'IA potrebbe essere utilizzata per indirizzare gli sforzi scientifici verso obiettivi specifici, anche scarsamente leciti o quanto meno scarsamente etici, piuttosto che verso una genuina e, permettetemi di dirlo, divertente esplorazione della conoscenza<sup>4</sup>.

<sup>4</sup> Nella versione online questo articolo è corredato di un'appendice in cui si dà la parola direttamente a ChatGPT, *n.d.r.*