




Stefano Moriggi
Mario Pireddu

L'Intelligenza
Artificiale
e i suoi fantasmi

Vivere e pensare
con le reti generative

IM

Il Margine



L'evoluzione delle Intelligenze Artificiali mostra come le reti generative siano in grado di imparare e di gestire enormi quantità di dati finendo per diventare «scatole nere» che forniscono risposte senza spiegare come. *L'Intelligenza Artificiale e i suoi fantasmi* si propone di aprire queste scatole per indagarne il funzionamento e decostruire pregiudizi e ideologie che vi aleggiano intorno come ectoplasmi, fino a insinuare che lo stesso concetto di intelligenza possa essere il più insidioso tra tutti gli spettri.

Nei quattro capitoli che compongono questo breve saggio, Moriggi e Pireddu spaziano da Giordano Bruno a Pac-Man ripercorrendo la storia che conduce alle attuali reti generative — fatta di idee, scoperte, illusioni, primavere e inverni della ricerca — per raffreddare i facili entusiasmi di chi crede ciecamente nelle magnifiche sorti e progressive, ma anche per spegnere sul nascere i roghi ai server delle aziende tech invocati dai tecnoscettici. Gli autori illustrano in modo accessibile e rigoroso le logiche e gli utilizzi concreti dei sistemi generativi — alla base degli strumenti che sempre più si utilizzano e si utilizzeranno per produrre contenuti di ogni tipo (testuali, iconografici, multimediali, ecc.) — offrendo un inquadramento teorico della creatività aumentata dalle tecnologie, al fine di comprendere meglio l'IA e le sue ricadute reali nella quotidianità di tutti.

Stefano Moriggi

1972

Professore associato di Cittadinanza digitale e di Società e Contesti Educativi Digitali presso l'Università degli Studi di Modena e Reggio Emilia. È membro della McLuhan Foundation (Toronto).

Mario Pireddu

1977

Professore associato di Tecnologie per la Formazione e di Apprendimento in rete e Gestione della conoscenza presso l'Università degli Studi della Tuscia di Viterbo, dove è Delegato per l'Innovazione nella Didattica.

Pluralità. Eterogeneità. Apertura.

Il Margine, progetto editoriale di Erickson, dà voce a punti di vista dissonanti e originali per metterli al centro del pensiero e del dibattito.

IN COPERTINA Eugène Leroux, litografia di un dipinto di Alexandre Gabriel Decamps (XIX secolo), *Scimmia allo specchio*, Smithsonian American Art Museum, gift of Mrs. Nigel Cholmeley-Jones, 1967.

PROGETTO GRAFICO Bunker

€ 10,00



Stefano Moriggi
Mario Pireddu

L'Intelligenza
Artificiale
e i suoi fantasmi

Vivere e pensare
con le reti generative

IA

L'Intelligenza
Artificiale
e i suoi fantasmi

Introduzione

La più umanistica delle domande

Il fantasma nella macchina ha un accento americano. Così all'inizio degli anni venti del XXI secolo un gruppo di ricercatrici e ricercatori intitolava un articolo dedicato al rapporto tra linguaggio e sistemi di calcolo (Johnson et al., 2022). Studiando GPT-3, modello di terza generazione della serie GPT creato da OpenAI e alla base di ChatGPT, Johnson e colleghi hanno analizzato la relazione tra valori culturali prevalenti contenuti nel dataset del modello linguistico con particolare attenzione alle differenze tra i prompt di input e gli output generati. Attraverso stress test e prove mirate, si è evidenziata una tendenza del modello ad allinearsi con quelli che hanno definito «valori dominanti» negli USA — in accordo con il World Values Survey, progetto di ricerca globale che monitora novanta paesi ed esplora i valori e la fiducia delle persone nel tempo e il relativo impatto politico e sociale.

Se i concetti e i termini che abbiamo usato in queste prime righe vi mandano almeno un poco

in confusione, questo che vi accingete a leggere è il libro che fa per voi. Cercheremo infatti di aiutarvi a capire di cosa parliamo quando parliamo di Intelligenza Artificiale (IA), e lo faremo evocando a più riprese *fantasmi*, spettri di ogni tipo che da sempre abitano le macchine e i discorsi sulle macchine.

Il modello GPT citato è un modello linguistico di grandi dimensioni capace di produrre testo simile al linguaggio umano (*natural language processing*), frutto di decenni di ricerca nell'ambito della linguistica computazionale e tra i più associati — nel dibattito pubblico degli ultimi anni — all'idea di Intelligenza Artificiale. I modelli basati su architettura trasformativa come GPT sono divenuti noti al grande pubblico dopo la pubblicazione, da parte dei ricercatori di Google Brain, di un paper dal titolo *Attention is all you need* (Vaswani et al., 2017), e in seguito soprattutto per la disponibilità di sistemi come ChatGPT di OpenAI.

Nei capitoli che vi accingete a leggere parleremo dell'importanza dei meccanismi di «attenzione» delle macchine e soprattutto delle *reti neurali* alla base dei sistemi che sempre più utilizziamo per generare contenuti di ogni tipo (testuali, iconografici, multimediali, ecc.). La storia che conduce alle attuali reti genera-

tive è fatta di idee, scoperte, illusioni, primavere e inverni della ricerca, fantasmi concettuali, entusiasmi e risultati che sino a pochi anni fa venivano visti come esclusivo appannaggio di racconti e film di fantascienza.

Nel 1950 il matematico inglese Alan M. Turing si chiese se le macchine possono pensare,¹ in quello che è diventato uno degli articoli più noti sull'argomento, dal titolo *Computing machinery and intelligence*. Non potendo offrire una definizione univoca di intelligenza e pensiero, propose il celebre *imitation game* — passato alla storia come test di Turing — che spostava la questione sul versante della percezione umana e sulla possibilità per la macchina di ingannare l'essere umano fingendosi una persona. Le prime ricerche sulle reti neurali artificiali risalgono però agli anni quaranta del xx secolo, con biologi e psicologi al lavoro su teorie che vedevano intelligenza e apprendimento connessi in qualche modo alla trasmissione di segnali tra i neuroni nel cervello umano.

¹ «La più umanistica delle domande», scrive Nello Cristiani nel suo *Machina sapiens* (2024). Sulla stessa questione insisteva già nel 1933 l'articolo dello scienziato statunitense Thomas Ross dal titolo *Machines that think* pubblicato su «Scientific American».

In realtà, alcuni fanno addirittura risalire l'origine di questi studi al 1890, quando il filosofo e psicologo William James sembrò prefigurare l'idea che l'attività di un neurone fosse determinata dall'accumulo degli stimoli ricevuti da altri neuroni, e che la forza di queste connessioni fosse modellata dalle esperienze passate. La premessa centrale alla base degli studi degli anni quaranta sosteneva effettivamente che i legami tra determinati neuroni si intensificassero con l'aumentare della frequenza delle loro interazioni. In base a questo principio si può spiegare il perché, quando proviamo a fare qualcosa di nuovo per la prima volta possiamo sentirci inadeguati, per poi acquisire sicurezza con la pratica e l'esercizio: gli atti compiuti con successo tendono a consolidare la connessione tra i neuroni coinvolti nell'esecuzione di quella specifica attività.

Nel 1943 due ricercatori, Warren McCulloch e Walter Pitts, pubblicarono i risultati di un lavoro pionieristico sulla natura e il funzionamento dei neuroni, che fu seguito dalla realizzazione pratica di una rete neurale elementare attraverso l'uso di circuiti elettrici. Il loro modello innovativo gettò le basi per lo sviluppo della ricerca sulle reti neurali, sia per quel che riguarda lo studio dei processi biologici che avvengono nel

cervello, sia per ciò che riguarda l'evoluzione degli studi sull'Intelligenza Artificiale.

McCulloch era uno psichiatra e neurofisiologo, mentre Pitts era un matematico: la collaborazione dei due studiosi condusse alla descrizione del calcolo logico della rete neurale che unisce la neurofisiologia alla logica matematica. La rete neurale artificiale è dunque una macchina progettata per simulare almeno in parte il funzionamento del cervello umano, realizzata materialmente tramite componenti elettronici e operante attraverso l'uso di software specifici.

All'epoca, e per lungo tempo, la maggior parte degli scienziati si diceva convinta che un'autentica Intelligenza Artificiale potesse emergere soltanto attraverso un approccio *top-down*, ovvero essendo dotata *a priori* di tutte le regole indispensabili per eseguire i propri compiti. Questa visione operava in maniera diametralmente opposta rispetto all'approccio evolutivo di chi pensava i cervelli elettronici come sistemi flessibili in grado di apprendere in modo indipendente e di identificare autonomamente modelli nei dati forniti.

Nel 1958, a due anni di distanza da un incontro fondativo per gli studi sull'Intelligenza Artificiale di cui parleremo in seguito, il «New York Times» pubblica un articolo dal titolo *Electronic*

'brain' teaches itself, in cui si descrive un cervello elettronico in grado di insegnare a se stesso, di imparare autonomamente, e di diventare in breve tempo un sistema in grado di percepire, riconoscere e identificare ciò che lo circonda, senza bisogno di controllo o addestramento da parte dell'uomo.

L'enfasi data dal quotidiano alla notizia è comprensibile: inizialmente, l'ambizione che guidava lo sviluppo delle reti neurali era quella di concepire un sistema computazionale capace di emulare la capacità del cervello umano nell'affrontare e risolvere problemi. Nel tempo, però, l'interesse dei ricercatori si è orientato verso l'impiego delle reti neurali per eseguire specifiche funzioni, allontanandosi da un'impostazione puramente biologica (Signorelli, 2017). Le reti neurali sono state così applicate a una vasta gamma di campi, inclusi la visione artificiale, il riconoscimento vocale, la traduzione automatica, il supporto nelle diagnosi mediche, la moderazione dei contenuti sui social network, i giochi da tavolo e i videogiochi.

La ricerca sull'IA subì una rapida accelerazione a partire dalla metà degli anni settanta del xx secolo, quando l'informatico e ricercatore giapponese Kuniyuki Fukushima pose le basi per lo sviluppo della prima vera rete neurale multistr-

to alla base di quel che viene oggi definito *deep learning*, definendola capace di auto-organizzazione (*self-organizing*) (Fukushima, 1975).

Il libro che avete tra le mani è pensato per essere una guida nella comprensione delle reti generative e delle cosiddette Intelligenze Artificiali: in particolar modo vuole aiutare a capire perché di frequente ciò di cui parliamo non è (spesso) intelligenza e non è (forse) così artificiale come pensiamo.

Da più parti si invoca un «nuovo umanesimo» come soluzione al dilagare di tecnologie sempre più performanti, pervasive e artificiali. «Rimettere l'uomo al centro» pare essere l'urgenza di chi è pronto a scommettere su valori e principi che le macchine (e soprattutto quelle «pensanti») sembrano minacciare. Le macchine che «pensano», sosteneva Turing, anche se tenute in posizione subordinata, per molte persone evocherebbero una sorta di diminuzione della nostra specie. Ma se da un lato molti ricorrono a una screditata caricatura dell'umanesimo storico per anestetizzare l'angoscia che l'innovazione tecnologica innesca nelle anime belle e nelle coscienze inquiete, dall'altro ancora latita una proposta laica in grado di far davvero i conti con i fantasmi del passato e con le ombre

che da un futuro ignoto e incerto si proiettano sui nostri tempi.

Imparare a (con)vivere con le reti generative è il dovere di chi — da concreto umanista — ha compreso la natura tecnologica del bipede implume, e cerca di forgiare categorie e tattiche che lo aiutino a studiare (e in parte a scrivere) la storia di quel futuro ignoto e incerto che ci attende. Naturalmente per gradi: cominciando intanto a fare i conti con i fantasmi che da sempre accompagnano la storia delle macchine, e in particolare con quelli più terrificanti che affollano l'incubo ricorrente di una intelligenza non umana.

Cominciare per gradi può significare, ad esempio, monitorare criticamente la progressiva — e spesso inavvertita — incorporazione di sistemi di Intelligenza Artificiale generativa nei contesti professionali, negli ambienti educativi e anche nelle nostre vite private. E sforzarsi di indagarne le nuove dinamiche di traduzione ed esperienza del mondo, nonché l'inedita interazione tra il fare creativo umano e la capacità generativa di tali dispositivi. Il tutto, possibilmente, dotandosi di un quadro epistemologico adeguato.

Le reti generative, da questo punto di vista, rappresentano una svolta radicale: creano con-

tenuti originali in vari formati (immagini, testi, codici, suoni, video), e sfidano la tradizionale distinzione tra ideazione umana e produzione tecnologica. Il software diventa un co-creatore di conoscenza, di contesti e di esperienze. Questo solleva interrogativi sul confine — ammesso che sia tracciabile — che delimiterebbe (e tutelerebbe?) l'opera di una intelligenza umana da quella di una intelligenza «artificiale». Sfidare — e dunque prendere sul serio — i fantasmi che aleggiano attorno all'IA significa anzitutto confrontarsi con noi stessi e con i nostri pregiudizi sulla tecnica, ma ancor prima sulla stessa specie umana e sui valori entro i cui confini, insensatamente, continuiamo a custodirla come una reliquia.

Infatti, fare i conti con i fantasmi significa, ad esempio, scoprire e prendere atto che generatività e creatività non sono prerogative esclusivamente umane: possono essere individuate tra gli animali, come sappiamo, ma occorre anche riconoscerle ai software, alle macchine e alle reti neurali, che dimostrano abilità generative in ambiti come l'arte figurativa, la musica e la scrittura.

Fare i conti con i fantasmi significa inoltre (e forse soprattutto) imparare a vedere nelle tecnologie — e nelle paure che generano — il riflesso

più o meno distorto di un'umanità vacillante, fragile, timorosa di navigare in un mare aperto e sconosciuto. Un'umanità, questa sì, più simile a quella descritta e dipinta dai grandi umanisti italiani. Un'umanità al contempo cosciente — come avrebbe detto Giordano Bruno — del fatto che a una nuova visione del mondo deve per forza corrispondere una nuova visione dell'uomo.²

Tutte le parti e le tesi di questo libro sono state discusse e condivise da entrambi gli autori. Più nel dettaglio: l'Atto I (Prendere sul serio i fantasmi) è a cura di Stefano Moriggi; l'Atto II (Confabulazioni) a cura di Mario Pireddu.

² Il riferimento è qui al film *Giordano Bruno* diretto da Giuliano Montaldo nel 1973 (si veda in particolare: <https://www.youtube.com/watch?v=DOasVVHYcp4>) [tutti i link e siti Internet citati nel volume sono stati consultati l'ultima volta il 9 aprile 2024, ndr].