

Paolo Benanti, eticista e terziario francescano
La riflessione sull'ascesa delle intelligenze artificiali

ALGORETICA PER

di **PAOLO OTTOLINA**

Sam Altman, fondatore e ceo di OpenAI, nel 2019 paragonò il suo lavoro sull'intelligenza artificiale (IA o AI, se preferite l'inglese Artificial Intelligence) al Progetto Manhattan: «Sto facendo qualcosa di buono? O qualcosa di molto cattivo?». Con ChatGpt il quesito è uscito dall'accademia ed entrato nelle vite di tutti. Per cercare risposte abbiamo incontrato Paolo Benanti. «Se ChatGpt ci ha scosso, sono gli utilizzi industriali di Gpt (il modello alle spalle del chatbot, ndr) a cui guardare per capire come i linguaggi sintetici possono cambiare completamente il panorama, così come fece la chimica industriale all'inizio dello scorso secolo». Terziario francescano, presidente del comitato etico di Centai (la fondazione di cui parliamo qui a lato), docente di Bioetica ed etica delle tecnologie alla Pontificia Università Gregoriana, Benanti è eticista, teologo, filosofo, ma anche appassionato di tecnologia ed ex studente di ingegneria: «Cercavo risposte per capire il mondo, da francescano ho chiuso il cerchio con un dottorato sull'etica della tecnologia».

Riprendiamo il paragone con la chimica industriale. Perché ci aiuta a capire la differenza tra la chat (ChatGpt) e il modello (Gpt-3 e ora Gpt-4)?

«A fine '800 si scoprono i coloranti sintetici, dalla chimica nasce la produzione di farmaci e, nella stessa settimana, vengono scoperte l'aspirina e l'eroina, che all'inizio era commercializzata come farmaco perfetto. Dopo un po' ci si rende conto che dà dipendenza: *pharmakon*, che in greco significa sia medicina che veleno. Temo che i grandi modelli di linguaggio generativo, gli Llm, siano un equivalente di quella stagione. Per curare qualche mal di testa potremmo piangere tante vite. Gpt è la molecola chimica, ChatGpt è il farmaco con cui la commercializzano. Ma è a Gpt, cioè all'utilizzo industriale, a cui guardare».

Ha usato e analizzato Gpt-4: che cosa l'ha colpita di più?

«Da anni abbiamo telecamere e sensori con una visione migliore di quella dell'occhio umano. Ma gli mancava la percezione. Io vedo una persona e una sedia: so immediatamente qual è la differenza. Ora ChatGpt sembra diventare la "percezione dietro la visione". Se gli mostro un'immagine bizzarra lui la descrive, ma se aggiungo "cosa c'è di strano?", lui replica con un ragionamento di buon senso, analogo a

quello che farebbe una persona che incontro per strada. È la parte più sorprendente: non l'aver superpoteri, ma il dimostrarsi drammaticamente simile all'uomo comune».

Con Gpt-4 è riemerso il tema della cosiddetta «agen-

tività», che ci porta verso scenari di autocoscienza delle macchine. Come stanno le cose?

«Gli informatici hanno preso a prestito un termine che in filosofia significa "vita" per descrivere una caratteristica che non sapevano come chiamare. Ovvero che il sistema si comporta un po' come vuole: qui c'è tutta l'ambiguità del discorso. Perché "agentività" per un filosofo della natura è dire che c'è qualcosa di unico come la vita. Mentre per questi scienziati significa semplicemente che non sono in grado di controllare fino in fondo quello che il sistema fa. In una nota del documento di OpenAI, a pagina 52, scopriamo che questi grandi sistemi sembrano avere una sorta di loro finalità propria, che suona come una brama di ricerca di potere e ricchezza. Egoismo, in una parola».

Perché accade?

«Non lo sappiamo ma azzardo un'ipotesi. Questi sistemi di addestramento si basano su un Foundational model. La macchina estrapola dai dati cose che non sono soltanto altri dati, ma diventano informazioni. Ecco che questa "agentività", questa sorta di finalità del modello che emerge, potrebbe essere un effetto non voluto — per tornare all'esempio dell'eroina — del processo di "distillazione" dell'informazione».

La cosiddetta Agi, l'Intelligenza artificiale generale simile all'uomo, è un mito o una preoccupazione reale?

«OpenAI nel suo paper parla di "scintille" di Artificial General Intelligence (Agi). E qualcuno l'ha inteso come "aiuto, arriva Terminator!". In realtà probabilmente parliamo di un



**Non mi spaventa
l'intelligenza
artificiale
ma la stupidità
naturale
Essere etici
significa continuare
a domandare
e non accontentarsi
di risposte già date**



Come nei farmaci: «Gpt è la molecola chimica, ChatGpt la formula con cui la commercializzano»

L'UMANITÀ

sistema flessibile, l'interfaccia perfetta per tutti i compiti in cui una macchina può aiutare. Non penso che avremo mai macchine coscienti, almeno non con i computer basati sul silicio».

“Algoetica”, ovvero un’etica per gli algoritmi, è un suo neologismo, coniato ben prima di ChatGpt. Che cosa significa e perché oggi è più attuale che mai?

«È come mettere dei guardrail alla macchina perché non vada fuoristrada. Non basta l’etica, che è *soft law*, ma servono anche dei principi di *hard law*, per esempio il codice della strada. La cosa interessante è che l’algoetica inizia a essere acquisita e desiderata anche da tanti sviluppatori che lavorano per grandi compagnie: l’etica come *soft law* ha una grande potenza. Pensiamo alle aziende che riducono l’impatto della plastica».

Le IA ci trasformeranno in Homo Sapiens Sapiens Sapiens. Oppure sono solo un coltellino svizzero molto potente?

«Probabilmente arriverà l’età delle IA e la abiteremo da esseri umani: in questo la cosa che mi spaventa di più non sono le intelligenze artificiali, ma la stupidità naturale».

Il tema del lavoro e dell’impatto delle IA: serve che i governi adottino al più presto una “sostenibilità digitale”?

«C’è una sfida fortissima tra una tecnologia che corre tantissimo e un processo democratico che è molto lento. Il tutto in un contesto sociale molto polarizzato che ci rende fragili. Per costruire servono comunità allargate e secondo me l’esempio di **Centai** è ottimo perché ha un Dna di pluralità e di dialogo».

All’Italia serve un’Authority per l’IA?

«Se diventa un luogo di confronto tra competenze diverse allora assolutamente sì. Serve creare un consenso politico e

ra. Probabilmente assisteremo a qualcosa del genere, dove il *prompt engineering* per generare un’immagine avrà un valore diverso. Ho visto immagini su Midjourney con prompt di 2.500 parole. Un romanzo. Avremo forme di controllo su quelle che sono le immagini ottenute nella vecchia maniera, tipo i sigilli del cibo bio, anche perché il vero problema sarà la verificabilità delle foto nei processi democratici».

Con Bill Gates, che esalta le possibilità dell’IA, o con Elon Musk, che dice «fermate tutto perché rischiamo l’estinzione»?

«Direi che entrambi fanno un errore: giudicare una volta per sempre un fenomeno. La clava è sempre stata sia utensile che arma e vale anche per il nucleare. Essere etici significa continuare a domandare e non accontentarsi di una risposta già data».

© RIPRODUZIONE RISERVATA

sociale, perché non diventi un corpus tecnico o un’appendice ma affinché sia veramente un luogo di cultura al servizio di tutti».

Uber, carne sintetica, ChatGpt: l’Italia ha un problema con l’innovazione?

«Abbiamo una lunga tradizione giuridica che a volte può diventare un incubo burocratico, che può essere nemico dell’innovazione. Il problema è mettere in atto dei circuiti virtuosi che spingano il sistema Paese a produrre delle eccellenze, che non significa solo la qualità del prodotto, ma anche un prodotto che si adegua molto bene a quel tessuto piccolo-medio industriale che tipicamente è nostro».

Le IA che creano immagini e i casi delle foto fake di Trump o del Papa: come ci salveremo in futuro?

«La fotografia non ha sostituito come forma d’arte la pittu-



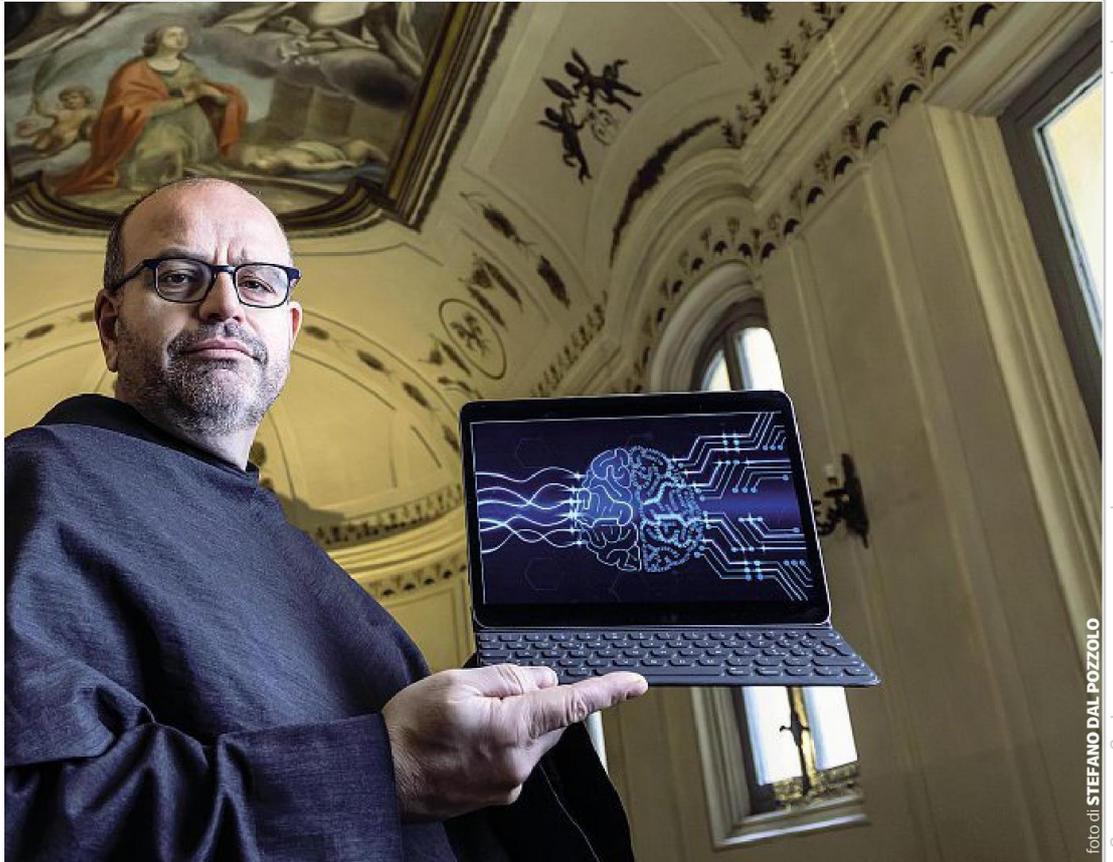
Tra Gutenberg e Steve Jobs

Paolo Benanti nella canonica della chiesa dei Santi Quirico e Giulitta, a Roma

LOG

ISUOI STUDI E I LIBRI

Nato nel 1973, Paolo Benanti è un frate francescano del Terzo ordine regolare. Da molti anni scrive di etica della tecnologia e analizza l'impatto che l'innovazione ha su uomo e società. Tra i suoi libri: «Human in the loop» (Mondadori), «Digital Age: teoria del cambio d'epoca» (San Paolo), «Oracoli: tra algoretica e algocrazia» (Luca Sossella), L'hamburger di Frankenstein (Edb) e «La tecnologia per l'uomo» (Mondadori).



La proprietà intellettuale è riconducibile alla fonte specificata in testa alla pagina. Il ritaglio stampa è da intendersi per uso privato
foto di STEFANO DAL POZZOLO





L'iniziativa

Fondata a Torino
nel 2021, Centai
si focalizza sull'analisi
dei sistemi complessi
con l'uso di IA e big data