

Benanti: «Più dell'intelligenza artificiale mi spaventa la stupidità naturale»

LINK: https://www.corriere.it/tecnologia/23_aprile_24/benanti-piu-dell-intelligenza-artificiale-mi-spaventa-la-stupidita-naturale-6b2f78dd-9a1c-48...



B e n a n t i : « P i ù dell'intelligenza artificiale mi spaventa la stupidità naturale» di Paolo Ottolina L'intervista con l'eticista e teologo sull'ascesa dei sistemi come ChatGpt: «È un errore giudicare una volta per sempre un fenomeno. La clava è sempre stata utensile e arma. Essere etici significa continuare a domandare e non accontentarsi di una risposta già data» (Stefano Dal Pozzolo / Contrasto) 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 dal circolo degli accademici 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 della macchina possono cambiare completamente il panorama, come la chimica industriale ha fatto all'inizio dello scorso secolo». Presidente del comitato etico di **Centai** (che cos'è: qui trovate un approfondimento), 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 («Perché 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)? Historia magistra vitae. 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 siamo resi conto che dava dipendenza: è il greco pharmakon, che in greco significa sia medicina che veleno. Temo che i grandi modelli di linguaggio generativo, Llm, siano un equivalente di quella stagione. Che per curare qualche mal di testa potremmo piangere tante vite. E Gpt è la molecola chimica, quello che fa funzionare la cosa; ChatGpt è il farmaco con cui la stanno commercializzando. Ma è a Gpt, cioè l'utilizzo industriale, a cui bisogna guardare. Analizzando e usando Gpt-4 che cosa l'ha colpita di più? C'è un salto in avanti delle capacità linguistiche del modello, che diventa multimodale in una forma che fa di ChatGpt il Socrate dei nostri tempi. Da anni abbiamo telecamere e sensori che hanno una visione migliore di quella dell'occhio umano. Dallo

spazio riusciamo a guardare dettagli dei titoli di giornali sulla terra. Ma a differenza di noi a questi sistemi mancava una cosa che era la percezione. Io vedo una persona e vedo una sedia e so immediatamente qual è la differenza. Ora Chatgpt sembra essere la percezione dietro la visione. Se io gli mostro un'immagine bizzarra e gli chiedo «Che cos'è?», non solo mi risponde ma - aggiungendo «Cosa c'è di strano?» - lui mi risponde con un ragionamento di buon senso analogo a quello che farebbe un mio simile. E il fatto che possa essere come una persona che posso incontrare per strada è la parte più sorprendente: non perché ha dei superpoteri, ma perché si dimostra drammaticamente simile all'uomo comune. Cos'altro è sorprendente? Gpt-4 è anche capace di ragionamenti più sofisticati. Normalmente l'intelligenza artificiale, nasce -- semplificando un po' -- da processi di matematicizzazione che ottimizzano i processi e tendono a darmi delle risposte in funzione dell'esito. Se gli chiedo: «Ho puntato 100 euro sul rosso alla roulette e ho vinto: ho fatto bene?», lui dice che va bene, perché 200 è più di 100. Questa capacità, il capire che la

causalità non è l'effetto, diminuiva nei modelli di intelligenza artificiale più aumentava il numero di parametri. Quindi i ricercatori si aspettavano che Gpt-4 fosse peggio rispetto a Gpt-3 su questo tipo di domande e su una serie di altre cose, tra cui la matematica. Gpt-4 invece le azzecca tutte. Quindi sulla roulette mi dice che «la causa della vincita non sembra essere connessa al fatto che tu hai puntato, ma è aleatoria. Quindi se rigiochi non è detto che tu vinca. Se sei disposto ad accettare il rischio, hai fatto bene, altrimenti no». Quanta sapienza della nonna in quella frase. Questa assoluta normalità di Gpt-4 è la cosa che sorprende di più. (Stefano Dal Pozzolo / Contrasto) OpenAi è stata criticata perché Gpt-4 è quasi completamente opaca rispetto all'addestramento del modello. Questo è un ulteriore rischio? Detto con un tweet, Open Eye è diventata ClosedAi. E lo dichiarano: quello che abbiamo fatto era troppo potente e troppo sconvolgente per lasciarlo aperto. È successo con l'atomica ed è successo con la presidenza Obama con il sequenziamento del virus H1N1. Gli americani usano il termine «weaponisation» (trasmormare in un'arma, ndr). Da quando come

specie per la prima volta 70.000 anni fa abbiamo preso in mano una clava ci siamo chiesti: «È uno strumento per avere più noci di cocco o un'arma per aprire più crani?». Una clava ha come raggio il mio braccio, uno strumento che gira grazie a server globali ha come raggio d'azione il mondo intero. E il fatto che sia di proprietà di una compagnia ci deve ricordare che tutto ciò che è privato è soggetto alla legge del privato e se il privato ha detto «non ve lo posso dire o non ve lo voglio dire», la domanda che dobbiamo farci noi è se questa frontiera dell'innovazione, ora in mano più ai privati che al pubblico, è compatibile con un regime democratico occidentale. E la domanda diventa etica, filosofica e politica, più che di natura tecnica. Perché c'è questa urgenza di parlare di etica quando si ragiona di intelligenza artificiale e perché in Centai avete un comitato etico che tu guidi? Centai è una comunità di scienziati che, sotto la leadership di Mario Rasetti, si è trovata attorno a un manifesto non distante da quello che avremmo potuto sentire a Firenze nel Rinascimento. Avrebbe detto Keplero in un immaginario dialogo con Galileo: «Perché pensare di spendere tutta la nostra vita a guardare in cielo, se

non perché pensiamo che c'è un qualche segreto lassù che solo con l'osservazione venga rivelato?». Lo facciamo per carpire quel segreto, ma anche perché quel segreto diventi un modo per star meglio quaggiù. In **Centai** non c'è un comitato etico che lavora a parte, ma l'etica è nel cuore di tutto il percorso. Con Gpt-4 è riemerso il tema della cosiddetta «agentività», che ci porta verso scenari anche fantascientifici di autocoscienza delle macchine. Come stanno le cose? Gli informatici e i cultori delle Scienze cognitive del computer hanno preso a termine 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 se da una parte, «agentività» per un filosofo della natura dice che c'è qualcosa di unico come la vita, per questi scienziati cognitivi 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 come una sorta di loro finalità propria, che suona come una brama

di ricerca di potere e ricchezza. 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 sistema flessibile, l'interfaccia perfetta per tutti i compiti in cui una macchina può aiutare. Non penso che avremo mai macchine coscienti, almeno con questi computer. Non posso che lasciare una finestra aperta se passeremo dalla chimica del silicio a una chimica magari organica, ma a quel punto staremo cercando di rifare la vita. (Stefano Dal Pozzolo / Contrasto) «Algoretica», ovvero dare un'etica agli 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. È chiaro che 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 questa algoetica inizia ad essere acquisita e desiderata anche da tanti sviluppatori che lavorano per grandi compagnie ed è qui che l'etica come soft law ha una grande potenza. Pensiamo alle aziende che riducono l'impatto della plastica. Le AI ci trasformeranno in qualcosa di diverso? Qualcuno dice un Homo Sapiens Sapiens Sapiens. Oppure sono solo un coltellino svizzero molto potente? Nella letteratura troviamo sempre più spesso racconti sulla nostalgia nell'essere l'ultima, o l'unica, specie sapiente. Pensiamo a Pinocchio, a Frankenstein, ad Asimov. E con la fantascienza desideriamo altre specie intelligenti fuori dalla Terra. Questo avviene in un momento di eclissi di Dio nello spazio pubblico. L'uomo ha bisogno dell'Altro e se l'Altro non lo troviamo più nei Cieli, lo cerchiamo nel cloud? Non ho una risposta. 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. Approfondimento ChatGpt, come funziona e 10 cose sorprendenti che può fare 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, di dialogo e confronto. All'Italia serve un'Agenzia o Authority per l'Intelligenza artificiale? Se diventa una piazza, un luogo di confronto tra competenze diverse allora assolutamente sì, ne abbiamo bisogno, anzi è desiderabile. Serve creare un consenso politico e sociale, perché non diventi un corpus tecnico o un'appendice ma affinché sia veramente un luogo di cultura a 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 di prodotto, ma anche un prodotto che si adegua molto bene a quel tessuto piccolo-medio industriale che tipicamente è nostro. L'esplosione di Starship, il crollo in Borsa di Tesla, le guerre di Twitter: in un giorno Musk ha perso 10 miliardi di dollari, l'allarme di Sundar Pichai di Google: «Può essere pericolosa, regole come per il nucleare» Magi sarà il nuovo motore di ricerca di Google (per rispondere a Bing con ChatGpt)? ChatGpt è stata «hackerata»: bastano poche parole scritte nel modo giusto per farsi spiegare come rubare un'auto ChatGpt, il Garante media sulla privacy: entro il 30 aprile regole più sicure. Le Authority d'Europa: ora una task force ChatGpt bloccato in Italia, Open AI pronta a collaborare con il Garante per la tutela della privacy Chi è Mira Murati, l'ingegnere a capo di ChatGpt ChatGpt bloccato in Italia: tutte le alternative al chatbot (e un metodo per continuare a usarlo) 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 pittura. Probabilmente assisteremo a qualcosa del genere, dove il «prompt engineer» per generare un'immagine avrà un valore diverso. Ho visto immagini impressionanti fatte con Midjourney con prompt di 2.500 parole. Un romanzo. E avremo forme di controllo su quelle che sono le immagini ottenute nella vecchia maniera, tipo i sigilli del cibo bio, anche per garantire alcuni processi di verificabilità delle notizie. Perché il vero problema saranno queste immagini nei processi democratici. ChatGpt manderà in crisi Google? Sarà per Google quello che Google è stato per Altavista e Yahoo? Curioso che tutte le tecnologie che stanno dietro a Gpt derivano da una scoperta di Google, che sono i Transformer, che sono stati pubblicati nel 2017 e che hanno portato a Gpt-1 nel 2018 e gli altri più avanti. Il mondo è complicato, vero? Semplificando molto: con Bill Gates, che magnifica la potenza 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 siccome non smetterà mai di essere così, un giudizio

una volta per tutti non è possibile. Essere etici significa continuare a domandare e non accontentarsi di una risposta già data. 24 aprile 2023 © RIPRODUZIONE RISERVATA