



GETTY IMAGES

500 mln

Il giro d'affari italiano dell'intelligenza artificiale, cresciuto del 22% in un anno

8%

Gli italiani che hanno usato Chat Gpt secondo un sondaggio di The Fool

«Condividiamo molte delle preoccupazioni della lettera. Da anni parliamo delle promesse e dei rischi dell'AI. Anche noi siamo preoccupati per il rischio di una sua accelerazione. Ma ci sono due cose. Abbiamo impiegato più di sei mesi per rendere sicuro Gpt-4, prima di renderlo pubblico. E non stiamo addestrando Gpt-5, a differenza di quanto si dice nella lettera. Non abbiamo nemmeno intenzione di farlo nei prossimi sei mesi. Ma pensiamo che per rendere sicuri questi sistemi bisogna metterli a disposizione degli utenti e capire la loro reazione. Così si possono calcolare rischi e impatti».

Gli impatti più temuti al momento sono due: gli effetti sul mondo del lavoro e quello sulla disinformazione, temi entrambi cruciali per la tenuta di una democrazia. Li temete anche voi?

«Questa tecnologia è destinata a trasformare molte cose. Ma oggi Chat Gpt è uno strumento di lavoro. Come altre tecnologie penso che cambierà la natura di molte professioni, forse ne creerà altre che nemmeno possiamo immaginare. È un tema serio con cui ci confrontiamo regolarmente con i governi».

E per quanto riguarda i contenuti fake? (Murati lascia per un attimo l'italiano e risponde in inglese)

«La disinformazione è uno dei rischi dei "Large language model", i modelli linguistici di grandi dimensioni, ma penso che sia molto importante portare la società a confrontarsi con questi rischi da subito e non lavorare al chiuso di un laboratorio per poi metterli a disposizione di tutti. È la nostra strategia».

Non è rischioso che la rivoluzione promessa dall'AI sia guidata da aziende private?

«Non credo che questo cambiamento debba essere gestito dalle società private. La tecnologia è uno strumento. Può essere un bene o un male. L'AI non fa differenza. Può fare cose grandiose per l'umanità, o cose molto pericolose. La soluzione non è fermarla. Ma gestirla».

© RIPRODUZIONE RISERVATA

IL COMMENTO

Piergiorgio Odifreddi

L'intelligenza artificiale ci sta superando darà vita a una nuova specie tecnologica

La possibilità che le macchine prendano il sopravvento sugli esseri umani è realistica

Nelle ultime settimane l'Intelligenza Artificiale è affiorata nel dibattito pubblico, in seguito all'esplosione di Chat Gpt. Ma in realtà il dibattito al proposito è vecchio come il mondo, letteralmente, perché il problema teorico ha di molto preceduto le sue realizzazioni pratiche.



Nel mito di Pigmalione, per esempio, lo scultore crea la statua di Galatea, e finisce per innamorarsene. Sembrerebbe una reazione esagerata, ma è quello che succede spesso nel nostro rapporto con le macchine: quelle letterali, come le automobili o le motociclette, e quelle metaforiche, come i computer o i telefonini.

George Bernard Shaw riprese nel 1913 il mito di Pigmalione in un'omonima commedia, che poi approdò nel cinema con *My fair lady*. Il problema che lo scrittore affrontò riguardava le contesse: per esserlo, bisogna nascere con il sangue blu, o basta imparare a comportarsi come loro? Nella commedia, una fiammiferia viene addestrata alle maniere nobili, e le impara così bene, che quando viene introdotta in società nessuno si accorge che non è una vera contessa.

Nel 1950 Alan Turing, l'inventore del computer, propose un test analogo: quand'è che un computer diventa in-

telligente? La risposta dello scienziato fu: quando, chiacchierando con lui sullo schermo, non riusciamo ad accorgerci che non è un vero essere umano, ma è solo un programma. L'articolo di Turing fu la prima formulazione moderna del sogno dell'Intelligenza Artificiale.

Per molti anni, l'unico computer in grado di superare il test di Turing fu quello del film 2001 *Odissea nello spazio* di Kubrick, che ripropose il solito dilemma sulle macchine che alla fine scappano di mano e finiscono per fare

C'è il pericolo di abuso di sistemi così potenti come successo negli Stati Uniti

ciò che vogliono loro, procurando danni invece che benefici. Più o meno come nell'episodio della scopa portatrice d'acqua nel film *Fantasia* di Walt Disney, basato su una ballata di Goethe. O nel romanzo *Frankenstein* di Mary Shelley.

Poteva sembrare solo fantasia, ma già nel 1966 uno dei primi programmi di Intelligenza Artificiale mostrò che i problemi sollevati dagli scrittori erano reali. Si chiamava Eliza, come la protagonista del Pigmalione di Shaw, e il suo programmatore, Joseph

Weizenbaum, scoprì con sorpresa due cose. Primo, bastavano poche righe di codice per imitare in maniera convincente il comportamento di uno psicanalista. Secondo, gli utenti ebbero subito l'impressione di parlare per davvero con uno psicanalista, e incominciarono a interagire con il programma come se fosse un essere umano. Weizenbaum rimase scioccato, e divenne uno dei primi critici dell'Intelligenza Artificiale.

Nel 1997 il programma Deep Blue batté il campione mondiale di scacchi Gary Kasparov, in un torneo di sei partite. Fu un momento epocale, perché gli scacchi sono considerati uno dei giochi più intellettuali che ci siano. Ancora anni dopo quella sconfitta, Kasparov rifiutava di credere che fosse stata tutta farina del sacco del computer. Sosteneva, in particolare, che alcune mosse erano troppo umane, perché una macchina avesse potuto pensarle.

Oggi i programmi per computer ormai superano facilmente i giocatori umani, esattamente come le automobili corrono più veloci degli atleti. Questo non impedisce che ci siano le Olimpiadi di corsa o i Campionati di scacchi, ma c'è una differenza: gli uomini sapevano da sempre di non essere i più veloci del creato, ma credevano di essere i più intelligenti. Ora hanno dovuto ricredersi, almeno per il ti-

po di intelligenza logico-deduttiva che serve per giocare a scacchi.

L'arrivo di Chat Gpt ha impressionato anche il pubblico che non gioca a scacchi, perché tutti hanno potuto vedere di persona quanto sia facile imitare un'intelligenza di routine. O, viceversa, quanto sia meccanico il lavoro compilatorio che Chat Gpt produce senza fatica, e che molti di noi sono costretti a fare nel proprio lavoro di ufficio. Siamo dunque di fronte a un potenziale rischio che riguarda la disoccupazione di molti im-

I programmi finiscono per compiere azioni impreviste dai programmatori

piegati, che possono o potranno essere facilmente sostituiti dalle macchine, come già sono stati sostituiti dai robot molti operai.

Ma l'Intelligenza Artificiale minaccia di avere ben altro in serbo, per noi. Primo fra tutti, il rischio che qualcuno ne abusi a proprio vantaggio e a danno altrui, come si è già visto per gli algoritmi di Twitter usati nella campagna elettorale di Trump nel 2016, o negli scandali dei dati venduti per un analogo sfruttamento da aziende come Telecom nel 2006, o Facebook nel

2019. I maggiori rischi sono però altri. Anzitutto c'è la possibilità che, anche con le migliori intenzioni, i programmi finiscano per fare cose che non erano state previste dal programmatore. La verifica di correttezza dei programmi, per assicurarsi che facciano tutto e solo ciò che devono fare, è un problema dimostrabilmente insolubile. Sono dunque sempre in agguato potenziali disastri, come il falso allarme nucleare del 1983 in Unione Sovietica, o disastri attuali, come il Flash Crash in borsa del 2010 negli Stati Uniti.

Inoltre, e infine, c'è la possibilità che l'Intelligenza Artificiale non raggiunga soltanto l'intelligenza umana, ma la superi, e prenda il sopravvento su di essa. Gli idealisti credono che la coscienza e i valori siano unicamente umani, e che una macchina rimarrà sempre una macchina. Ma i materialisti sanno che la coscienza è solo un epifenomeno, che sorge spontaneamente in un cervello che ha raggiunto una massa critica.

Quando anche l'Intelligenza Artificiale raggiungerà una massa critica, saremo di fronte a una nuova specie tecnologica, con i propri valori, che potrebbero essere in competizione con i nostri. Se ne preoccupano Bill Gates e Elon Musk, che di certe cose se ne intendono. Forse faremo bene a preoccuparcene anche noi. —

© RIPRODUZIONE RISERVATA