

LA LETTERA

Il Gotha della tecnologia avverte: «Sui chatbot ci serve una pausa»

PIETRO SACC

Appello con oltre mille firme, compresi Musk, Wozniak e Bengio: stop all'addestramento dei sistemi di intelligenza artificiale, per renderli più sicuri e scrivere le regole «Rischiando di perdere il controllo della nostra civiltà» Milano. Alcuni dei più grandi esperti al mondo nel campo dell'intelligenza artificiale hanno firmato un appello per mettere in pausa lo sviluppo di ChatGPT e dei suoi simili. «Chiediamo a tutti i laboratori di intelligenza artificiale di fermare immediatamente per almeno sei mesi l'addestramento di sistemi di intelligenza artificiale più potenti di GPT-4. Questa pausa dovrebbe essere pubblica e verificabile, e includere tutti gli attori chiave. Se questa pausa non può essere attuata rapidamente, i governi dovrebbero intervenire e istituire una moratoria» recita la lettera aperta pubblicata dal Future of Life Institute, organizzazione non profit americana che ha l'obiettivo di «indirizzare le tecnologie trasformatrici lontano da rischi estremi e su larga scala, e verso il beneficio della vita». Tra le oltre mille firme di questo appello ci sono i nomi di centinaia di imprenditori, studiosi ed esperti di intelligenza artificiale. Tra i

firmatari spiccano Elon Musk (ceo di Tesla, SpaceX e Twitter), Steve Wozniak (co-fondatore di Apple), luminari dell'intelligenza artificiale come Yoshua Bengio (vincitore del premio Turing, una sorta di Nobel per le scienze informatiche) e Stuart Russell (autore del testo sull'intelligenza artificiale più studiato nelle università) e lo storico israeliano Yuval Noah Harari, famoso per i suoi bestseller sulla storia dell'uomo e della civiltà. Musk, attraverso la sua fondazione, è anche il principale finanziatore del Future of Life Institute, a cui ha donato 3,5 milioni di dollari nel 2021. Un altro mezzo milione di dollari sono arrivati da Founders Pledge, l'organizzazione filantropica preferita dagli imprenditori della Silicon Valley. Musk è allo stesso tempo uno dei fondatori di OpenAI, la società che ha sviluppato ChatGPT, e continua a collaborare con l'organizzazione, anche se ha lasciato il consiglio di amministrazione nel 2018. L'appello pone domande chiave e solleva molti dei problemi posti dal rapido successo di sistemi come ChatGPT, che secondo le rilevazioni di Similarweb è stato usato un miliardo di volte solo a febbraio, confermandosi come la nuova app tecnologica a più forte crescita della storia. Gli esperti si chiedono ad esempio se dobbiamo lasciare che le macchine inondino di propaganda e bugie i nostri canali informativi, o se dobbiamo lasciare che l'automazione spazzi via tutti i lavori, anche quelli gratificanti. «Dovremmo rischiare di perdere il controllo della nostra civiltà?». Da qui la richiesta di una pausa. Almeno sei mesi per sviluppare e applicare protocolli di sicurezza sul futuro dell'intelligenza artificiale, focalizzando la ricerca e l'attenzione sul rendere i sistemi già disponibili più sicuri, trasparenti e affidabili. La politica potrebbe usare questa pausa per definire



Avvenire

il contesto legale dell'intelligenza artificiale, ad esempio istituendo autorità per vigilare sul settore, garantendo al pubblico di potere conoscere la provenienza umana o robotica di un contenuto, assicurando che si possa sapere chi è il responsabile di eventuali danni provocati dall'intelligenza artificiale. « Dopo essere riusciti a creare potenti sistemi di intelligenza artificiale, ora possiamo goderci una "estate dell'intelligenza artificiale" in cui raccogliamo i frutti, progettiamo questi sistemi per il chiaro vantaggio di tutti e diamo alla società la possibilità di adattarsi – è la conclusione dell'appello –. La società ha messo in pausa altre tecnologie con effetti potenzialmente catastrofici sulla società (qui si fa riferimento ad esempio alla clonazione o all'alterazione genetica dell'essere umano, ndr). Possiamo farlo qui. Godiamoci una lunga estate AI, non precipitiamoci impreparati nell'autunno» RIPRODUZIONE RISERVATA Una schermata di ingresso di ChatGPT /Reuters.