



14 Maggio 2023

## **Intelligenza Artificiale, «tutti sulla Terra moriranno»**

L'umanità non è preparata a sopravvivere a un incontro con un'Intelligenza Artificiale molto più intelligente. Eliezer Yudkowsky, esperto riconosciuto delle prospettive catastrofiche dell'AI, lo oramai ripete da decenni.

Arrestare lo sviluppo di sistemi avanzati di intelligenza artificiale in tutto il mondo e punire severamente coloro che violano la moratoria è l'unico modo per salvare l'umanità dall'estinzione, avverte Yudkowsky, co-fondatore del Machine Intelligence Research Institute (MIRI).

Yudkowsky ha scritto un articolo di opinione per la rivista *TIME* il mese scorso, spiegando perché non ha firmato una petizione che invitava «tutti i laboratori di Intelligenza Artificiale a sospendere immediatamente per almeno sei mesi l'addestramento di sistemi di intelligenza artificiale più potenti di GPT-4», un modello di linguaggio multimodale rilasciato da OpenAI recentemente e che ha già mostrato scenari inquietanti, e anche molto.

Yudkowsky ha sostenuto che la moratoria per fermare la ricerca IA firmata da a Elon Musk e Steve Wozniak e altri, «chiede troppo poco per risolvere» il problema posto dallo sviluppo rapido e incontrollato dell'IA. (Alla moratoria, ricordiamo, si è opposto Bill Gates)

«Il risultato più probabile della costruzione di un'Intelligenza Artificiale sovrumana, in qualcosa di lontanamente simile alle circostanze attuali, è che letteralmente tutti sulla Terra moriranno», scrive lo Yudkowsky.

Sopravvivere a un incontro con un sistema informatico che «non si prende cura di noi né della vita senziente in generale» richiederebbe «precisione e preparazione e nuove intuizioni scientifiche» che l'umanità manca al momento ed è improbabile che ottenga nel prossimo futuro, ha affermato.

«Un'Intelligenza Artificiale sufficientemente intelligente non rimarrà confinata a lungo nei computer», ha avvertito Yudkowsky, spiegando che il fatto che sia già possibile inviare stringhe di DNA ai laboratori per la

produzione di proteine consentirà probabilmente all'intelligenza artificiale di «costruire forme di vita artificiale o avviarsi direttamente alla produzione molecolare postbiologica» e uscire nel mondo.

Secondo il ricercatore, deve essere introdotta immediatamente una moratoria a tempo indeterminato e mondiale sulle nuove importanti corse di addestramento dell'IA. «Non ci possono essere eccezioni, anche per governi o militari», ha sottolineato.

Dovrebbero essere firmati accordi internazionali per porre un limite alla quantità di potenza di calcolo che chiunque può utilizzare per addestrare tali sistemi, insiste lo Yudkowsky.

«Se l'Intelligence dice che un Paese al di fuori dell'accordo sta costruendo un cluster GPU (unità di elaborazione grafica), bisogna avere meno paura di un conflitto a fuoco tra le Nazioni che della violazione della moratoria; bisogna essere disposti a distruggere un datacenter canaglia con un attacco aereo», ha scritto.

Come riportato da *Renovatio 21*, Yudkowsky propone di bombardare materialmente i computer per fermare l'ascesa di un'IA in grado di mettere fine all'esistenza degli uomini.

La minaccia dell'Intelligenza Artificiale è così grande che dovrebbe essere reso «esplicito nella diplomazia internazionale che la prevenzione degli scenari di estinzione dell'IA è considerata una priorità rispetto alla prevenzione di uno scambio nucleare completo».

La prospettiva di Yudkowsky è meno radicale di quanto sembri.

Elon Musk ha dichiarato durante la sua recente intervista con Tucker Carlson che il fine di Google – nemmeno così occulto – è quello di creare un dio-Intelligenza Artificiale. Quando Musk ha ribattuto ad un entusiasta fondatore di Google Larry Page specificando i rischi dell'AI, il Page gli ha dato dello «specista», termine del gergo animalista che indica una persona che mette la specie umana sopra le altre.

L'idea che l'AI diventerà una sorta di dio circola da anni e investe direttamente personale di OpenAI come il capo scientifico e cofondatore Ilya Sutskever, il quale ha chiesto ai suoi followers se le super-IA avanzate dovrebbero essere rese «profondamente obbedienti» ai loro creatori umani, o se questi algoritmi dovrebbero «amare veramente profondamente l'umanità». L'anno scorso, Sutskever affermò che «può darsi che le grandi reti neurali di oggi siano leggermente consapevoli».

Da Google si è recentemente licenziato l'esperto britannico Geoffrey Hinton, divenendo così libero di dare voce ai suoi dubbi per i rischi essenziali che corre l'umanità con la creazione delle macchine superintelligenti. Pochi mesi fa un ingegnere di Google, Blake Lemoine, ha dichiarato pubblicamente che il chatbot AI dell'azienda, LaMDA, è «senziente», ossia dotato di una coscienza comparabile a quella degli esseri umani.

Come riportato da *Renovatio 21*, Musk, oramai quasi una decina di anni fa aveva parlato del fatto che creare l'AI era «come evocare un demone».

L'ultimo germe di incubo pochi giorni fa, quando si è visto che qualcuno ha equipaggiato un robocane Boston Dynamics con ChatGPT.

La minaccia è reale, e *Renovatio 21* ve lo ripete da anni.