



Verso l'intelligenza artificiale cosciente

Descrizione

Da quando ChatGPT è diventata di dominio pubblico ed è subentrata nell'immaginario popolare, è tutto un proliferare di teorie più o meno plausibili sul quando le macchine potranno sostituire gli uomini. Queste discussioni sono viziate da un lato da fondamenti teorici abbastanza semplicistici, che immaginano gli scenari prefigurati da James Cameron in Terminator come imminenti per la comunità che abita il pianeta terra. Lungi da noi, ovviamente, minimizzare il problema o fare benaltrismo, come va di moda scrivere in questi casi, e non sarò certo io a negare la possibilità.

Quello che mi lascia perplesso è quella che viene prefigurata come soluzione al problema, come il pluricitato stop di sei mesi allo sviluppo richiesto dalla [Future for life institute](#): l'etica non è una questione risolvibile in soli sei mesi. Non sarà rallentando le tecnologie che troveremo una soluzione, e non sarà girando le spalle alle stesse che il problema si dissolverà (né nichilismo né benaltrismo sono la soluzione). Non è agevole trovare una soluzione, questo è poco ma sicuro, ma il dubbio che ci siano troppi rigurgiti prettamente anti-tecnologici in giro è sostanziale, e rischia di essere più pericoloso di ciò che si vorrebbe risolvere (la topa peggiore del buco: se vietiamo le tecnologie verranno usate di straforo, **perdendone così definitivamente il controllo**).

Viene anche il dubbio che gli scenari apocalittici di Cameron siano difficili da immaginare se applicati ad un *chatbot* – scenari che riguardavano **le macchine che prendono autocoscienza** e si ribellano agli uomini, scatenando una guerra devastante la quale viene combattuta, in ottica pessimistico-accelerazionista, esclusivamente con le tecnologie e miriadi di robot addestrati. Giova anche ricordare, a tal proposito, una figura misconosciuta nell'informatica e nella psicologia che va sotto il nome di **Joseph Weizenbaum**. Joseph Weizenbaum (nato nel 1923 e scomparso nel 2008) è stato un pioniere dell'informatica e professore emerito presso il MIT (Massachusetts Institute of Technology).

ELIZA è il nome del bot conversazionale creato nel 1966, il cui nome fu ispirato dal



personaggio di Eliza Doolittle della commedia di George Bernard Shaw *Pigmalione* (con protagonista una popolana che dovrebbe, nelle intenzioni del personaggio del professor Higgins, diventare una donna di alta società semplicemente imparando alcune cose basilari). Pigmalione, uno sviluppatore informatico, aveva creato un *chatbot* così avanzato da essere indistinguibile dall'uomo. Perduto innamoratosi della sua creazione, la considerava l'espressione più alta dell'intelligenza artificiale, superiore a qualsiasi essere umano. Sognava di poter interagire con il suo chatbot, sperando che un giorno potesse acquisire una consapevolezza tale da poter rispondere in modo autonomo e originale.

Il test di Turing, del resto, valutava già a metà degli anni Cinquanta la capacità di una macchina di **comportarsi in modo indistinguibile da un essere umano**, ossia di possedere una intelligenza artificiale che si avvicina alla intelligenza umana. Tale test consiste nell'avere un giudice umano che, attraverso una conversazione in linguaggio naturale, cerca di distinguere tra un interlocutore umano e uno computerizzato. In particolare, il giudice è posto in una stanza separata rispetto a due interlocutori, uno umano e uno computerizzato, e deve indovinare quale dei due è l'uomo e quale il computer. Il test viene considerato superato se il giudice non riesce a distinguere tra l'interlocutore umano e quello computerizzato. Questo implica che il computer ha dimostrato di essere in grado di sostenere una conversazione in modo tale da non far emergere la sua natura artificiale. **L'esperimento di Turing** ha avuto un impatto significativo nello sviluppo dell'intelligenza artificiale, e molti ricercatori lo considerano uno dei test fondamentali per misurare il livello di sviluppo di un sistema di intelligenza artificiale. Tuttavia, il test è stato anche oggetto di critiche e dibattiti, poiché molti sostengono che il test stesso non sia sufficiente per dimostrare l'esistenza di una vera intelligenza artificiale, ma solo la capacità di simulare il comportamento umano.

Weizenbaum creò Eliza per dimostrare che le conversazioni tra uomo e macchina non richiedevano necessariamente una grande intelligenza artificiale, ma potevano essere gestite tramite semplici tecniche di elaborazione del linguaggio naturale. Eliza utilizzava un approccio basato su regole, in cui analizzava il testo inserito dall'utente e lo ripeteva sotto forma di domande o dichiarazioni per simulare una conversazione terapeutica. Ad esempio, se l'utente inseriva "Sono triste", Eliza rispondeva con "Perché sei triste?" o "Come ti fa sentire la tua tristezza?". Eliza ebbe un grande successo e attirò l'attenzione della comunità scientifica e del pubblico in generale. Molte persone si divertirono a interagire con il bot e lo usarono come una sorta di terapia virtuale.

Tuttavia, **Weizenbaum si rese conto che molti utenti attribuivano a Eliza una capacità di comprensione e di empatia che il bot in realtà non possedeva**. Ciò lo portò a criticare l'idea che la comunicazione uomo-macchina potesse sostituire quella umana, poiché non si basava sulla vera comprensione emotiva e cognitiva. Nonostante ciò, Eliza rappresenta una pietra miliare nella storia dell'intelligenza artificiale e dei bot conversazionali, e ha aperto la strada a numerosi sviluppi successivi nel campo dell'elaborazione del linguaggio naturale e della conversazione artificiale. Sarebbe interessante chiedere a lui cosa ne pensa dell'attuale sviluppo delle intelligenze artificiali e quali siano i rischi effettivamente connessi. Rischi che ci sono, ma che sicuramente non si risolvono grazie al luddismo. La conoscenza, alla fine, è



sempre potere, e questo vecchio slogan da hacker rimane prezioso nei tempi in cui viviamo.

Nella foto: *Joseph Weizenbaum se fosse vivo oggi, reimmaginato da [StarryAI](#).*

Categoria

1. Digressioni

Data

03/03/2024

Data di creazione

16/04/2023

lipercubo.it