

Taddeo: così l'intelligenza artificiale può inquinare il dibattito politico

ANALISI

La studiosa di etica digitale, docente a Oxford: la creazione e diffusione capillare e personalizzata di disinformazione adesso minaccia la nostra democrazia

ANDREA LAVAZZA

«La capacità di fare domande e la spinta naturale a comprendere il mondo che ci circonda» sono quello che caratterizza l'essere umano di fronte all'intelligenza artificiale (IA) che, oltre alle sue grandi potenzialità positive, rischia di diventare strumento per inquinare e falsare il dibattito pubblico. Risulta quindi sempre più necessaria un'etica digitale. Ed è esattamente Digital Ethics e Defence Technologies che insegna all'Internet Institute dell'Università di Oxford Mariarosaria Taddeo, studi a Bari e Padova e poi brillante carriera accademica in Inghilterra. L'esperta è stata Milano per la quarta edizione di "Il Verde e il Blu Festival", promosso da BIP e prodotto da Beulcke&Partners. **Professoressa Taddeo, perché l'intelligenza artificiale solo recentemente sembra essere balzata all'interesse generale? È l'effetto ChatGPT e quindi, passata la sorpresa e abituati a dialogare con strumenti di quel tipo, l'IA tornerà sotto traccia?**

Nell'ultimo anno c'è stata sicuramente un'attenzione maggiore dovuta alla diffusione dei large language models (LLMs) - ChatGPT appunto - che hanno richiamato l'attenzione per varie ragioni: gli LLMs svolgono molto bene compiti tipicamente attribuiti agli esseri umani, per esempio comporre un testo, creare una nuova immagine; le interazioni con questi modelli sono molto semplici, si danno comandi ai modelli attraverso il linguaggio naturale, quindi tutti quelli che sanno scrivere e leggere possono usarli direttamente; perciò vi è stata un'adozione larga e repentina. Questi modelli hanno un grande potenziale ma pongono anche rischi concreti (si pensi alla questione del copyright del materiale che producono o all'uso di questi modelli per diffondere disinformazione, per esempio). Servono quindi uno sforzo condiviso da diversi attori nelle nostre società (accademici, tech providers, legislatori) e un dibattito pubblico informato. Per questo credo che parleremo d'Intelligenza Artificiale (IA) e di etica dell'IA ancora per un po'.

Dalla sua prospettiva, ci sono reali motivi di preoccupazione per l'espansione dell'intelligenza artificiale, al di là della perdita di posti di lavoro in molti settori economici?

Ci sono rischi concreti da considerare. Chiariamolo subito, questi rischi non sono quelli relativi allo sviluppo di macchine realmente intelligenti, senzienti. Questa è fantascienza. I rischi sono diversi e in parte sono gli stessi che riguardano l'IA predittiva, per esempio, la possibilità per questi modelli di diffondere su larga scala errori e pregiudizi, la limitata predicibilità dei comportamenti dei modelli IA e la loro limitata trasparenza, due fattori che ne li-

mitano il controllo; i problemi nell'attribuire la responsabilità agli esseri umani per le azioni di queste macchine.

Ci sono tre aree che, a mio avviso, sono forse più rilevanti quando si pensa agli LLMs. La prima è l'integrazione di questi modelli nelle nostre attività quotidiane. Andiamo verso scenari in cui l'IA diventa un membro dei nostri team professionali, per esempio. Qui la questione è come integrare con l'IA nei processi professionali in modo da delegare i compiti senza erodere le capacità degli esseri umani di eseguire gli stessi compiti. Per esempio,

possiamo delegare all'IA la lettura di immagini diagnostiche, ma vogliamo che i medici continuino a saper leggere le stesse per capire quando l'IA sbaglia. La seconda area afferisce al dibattito pubblico e ai processi democratici. Abbiamo visto già con Cambridge Analytica, che la combinazione IA e social network può porre rischi molto seri per il dibattito pubblico e per lo svolgimento di processi democratici, per esempio elezioni e referendum, attraverso la creazione e diffusione capillare e personalizzata di disinformazione (fake news). Credo che

con gli LLMs questi rischi aumentino nella misura in cui questi modelli possono creare a basso costo testi personalizzati su certi tipi di utenti e indistinguibili da quelli creati da un essere umano. Come controlliamo le fake news, che responsabilità hanno gli sviluppatori di LLMs e che ruolo dovrebbero avere i gestori dei social networks?

La terza - conclude Taddeo - è la cybersecurity. L'IA è una tecnologia vulnerabile dal punto di vista della cybersecurity, gli LLMs non fanno eccezione. I dati ci dicono che è più facile attaccare un modello LLMs di altri tipi di IA. Questo perché è possibile manipolare questi modelli usando il linguaggio naturale. Stiamo integrando questi sistemi in maniera quasi capillare, si pensi a Google Bard che è stato reso disponibile per un miliardo di utenti. C'è da chiedersi se gli LLMs non siano un cavallo di Troia per le società digitali.

© RIPRODUZIONE RISERVATA



Superficie 26 %

ARTICOLO NON CEDIBILE AD ALTRI AD USO ESCLUSIVO DEL CLIENTE CHE LO RICEVE - L.1976 - T.1976