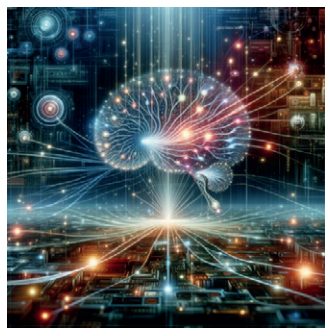
Le reti neurali artificiali, potenziale ancora da scoprire

Le reti neurali artificiali, da lungo tempo considerate inscrutabili, stanno finalmente rivelando nuovi aspetti del loro funzionamento grazie a recenti scoperte scientifiche. Questi avanzamenti stanno offrendo una nuova comprensione su come le reti neurali possano elaborare e risolvere problemi in modi precedentemente inimmaginabili. Un team di ricercatori di OpenAI ha osservato un fenomeno particolarmente interessante: il «grokking». Durante un esperimento, che originariamente includeva un errore di procedura, hanno continuato a sovrallenare una rete neurale oltre il punto in cui normalmente si sarebbe interrotto l'addestramento. Questo sovrallenamento ha portato a risultati inaspettati e straordinari: la rete non solo ha memorizzato i dati a cui era stata esposta, ma ha anche dimostrato di poter generalizzare le informazioni apprese a nuovi set di dati, raggiungendo una precisione del 100% nei test successivi. Questo suggerisce che le reti neurali potrebbero sviluppare una comprensione molto più profonda dei problemi rispetto a quanto si pensasse in precedenza, generando soluzioni innovative che trascendono i dati di addestramento forniti.



Il processo di grokking è stato esaminato attentamente da ricercatori di diverse istituzioni accademiche. Hanno scoperto che durante l'addestramento, le reti neurali possono organizzare i dati in modi non solo efficaci ma anche sorprendentemente creativi, come ad esempio proiettando numeri in uno spazio bidimensionale e identificando strutture circolari in esso. Nonostante queste scoperte siano promettenti, la ricerca sulle reti neurali e sul loro potenziale è ancora nelle sue fasi iniziali.
Gli scienziati avvertono che molte domande rimangono senza risposta, in particolare riguardo l'applicabilità di questi risultati a reti neurali di dimensioni maggiori e più complesse. La strada verso una piena comprensione delle capacità e dei limiti delle reti neurali è ancora lunga e ricca di sfide da superare.

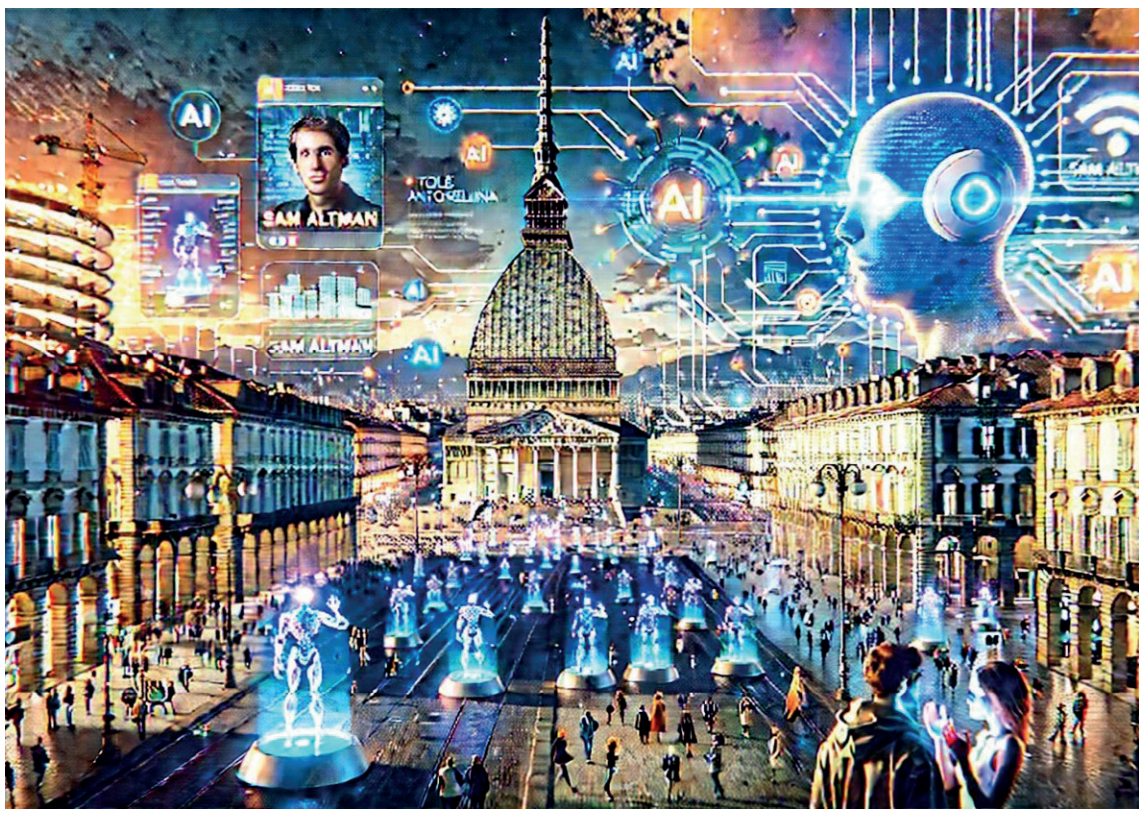
C.G.

APOSTOLATO DIGITALE

condividere codici di salvezza

CHATGPT E DINTORNI – ABBIAMO A CHE FARE CON UNA MACCHINA «CALCOLATRICE DI PAROLE»

Perché è ingannevole scambiare per «umani» i rapporti con gli algoritmi



Con la partecipazione di Sam Altman alla Italian Tech Week, Torino è diventata per un giorno la capitale mondiale dell'intelligenza artificiale generativa e conversazionale. Altman è considerato il «papà» dell'intelligenza artificiale, ma non l'ha inventata lui. Quasi due anni fa, Altman è stato il primo a renderla popolare e a portata di tutti, nella sua variante «generativa e conversazionale». **Conversazionale -** Su questo attributo fondamentale dell'intelligenza artificiale generativa, fondamentale per le ricadute sociali e antropologiche che porta con sé, si ragiona poco. Troppo poco. Al riguardo, sottopongo tre considerazioni.
La prima. Rendere l'interazione uomo/algoritmo sempre più fluida e naturale, dialogare in modo molto più realistico, è l'obiettivo di ChatGPT e dei suoi fratelli. È un obiettivo neutro, privo di conseguenze? A mio avviso potenziare l'empatia artificiale è una trappola emotiva. Infatti questo modo di conversare con l'intelligenza artificiale generativa ci illude di avere un rapporto con un «tu». Un tu che, almeno al momento, non esiste. Abbiamo ancora a che fare con una macchina calcolatrice di parole - come ci ricorda sempre il filosofo del digitale Cosimo Accoto - non con un essere senziente. È quindi sbagliato antropomorfizzare il rapporto con l'algoritmo.
A questa prima conseguenza, ne segue immediatamente un'altra. La capacità sempre più incisiva dell'intelligenza artificiale generativa di conversare con noi ci fa correre il rischio di considerarla un oracolo, una sorta di bocca della verità, le cui sentenze vanno prese per oro colato. Rischiamo di «confondere l'aumento della qualità dell'interazione con quello della correttezza delle informazioni fornite. L'elevata qualità dell'interazione può, paradossalmente, ridurre la consapevolezza degli utenti riguardo alla necessità di verificare le informazioni», come ha sottolineato Stefano Epifani, fondatore e presidente della Fondazione per la sostenibilità digitale, perché noi siamo naturalmente portati a unire buona comunicazione e affidabilità.

L'utopia della sostenibilità
Enrico Giovannini, Luca De Biase, Fabio Scaltritti.

Terza e ultima conseguenza, la più importante dal punto di vista antropologico. Noi esseri umani siamo esseri relazionali. Lo siamo per natura, siamo stati creati così. Nasciamo, cresciamo, viviamo, all'interno di rapporti, di contatti, di interazioni. Siamo l'unico essere vivente i cui cuccioli per tanti anni hanno bisogno di accudimento materiale e, soprattutto, relazionale. È altresì vero che tutti i rapporti, tutte le relazioni, anche le più belle, sono faticose, sono impegnative, perché noi siamo esseri relazionali pieni di limiti e di difetti, limiti e difetti che si ripercuotono nei rapporti e impongono fatica, anche con le persone che ci sono più care.
Il rapporto con l'intelligenza artificiale generativa è esente da questo tipo di fatica. Per «lei» è sempre la giornata mondiale della gentilezza. Ci tratta sempre con i guanti, è stata impostata così. Non essendo un essere senziente è priva dei limiti di noi esseri umani. È sempre «perfetta», quindi elimina la fatica dei rapporti. Corriamo il rischio di disabituarci a comprendere i bisogni dell'altro. Le relazioni si deteriorano quando ci concentriamo sui limiti e non sui bisogni della persona che abbiamo davanti. Una intelligenza artificiale generativa capace di emulare un'esperienza di interazione indistinguibile da quella che siamo abituati ad avere con un altro essere umano che impatto psicologico potrà avere sulla nostra capacità di reggere i rapporti tra noi? Vale per noi adulti, a maggior ragione vale soprattutto per i ragazzi, che cresceranno abituati ad avere rapporti con una intelligenza artificiale «amica perfetta».
In conclusione, dobbiamo rimanere attenti e vigili. Viviamo un'epoca straordinaria, ricca di possibilità inedite, ma che proprio per questo impone uno straordinario sforzo di attenzione, di pensiero, di comprensione, di educazione.

Antonio PALMIERI
Fondatore e presidente
Fondazione Pensiero Solido

GLOSSARIO/23 – ATTACCHI ON LINE

«Zero Day», vulnerabilità informatiche

Lo «zero day» è una vulnerabilità informatica sconosciuta agli sviluppatori del software e a chi lo utilizza, che viene scoperta solo quando viene sfruttata per la prima volta da un attaccante. Il termine «zero day» indica che gli sviluppatori del software hanno zero giorni per correggere la vulnerabilità, essendo questa appena scoperta e già utilizzata per un attacco. Queste vulnerabilità sono particolarmente pericolose perché non esistono patch o aggiornamenti disponibili per proteggere i sistemi colpiti. Gli attacchi zero day possono essere utilizzati per scopi diversi, tra cui il furto di dati sensibili, la compromissione di sistemi, l'installazione di malware o la creazione di botnet. Gli hacker che scoprono queste vulnerabilità possono scegliere di utilizzarle per i propri fini o venderle nel mercato nero, dove possono ottenere cifre elevate. Le vulnerabilità zero day sono spesso sfruttate in attacchi mirati contro organizzazioni di alto profilo, governi o infrastrutture critiche, rendendo la loro individuazione e mitigazione una priorità per i team di sicurezza informatica. La scoperta e l'analisi di una vulnerabilità zero day richiede competenze avanzate e risorse significative, poiché queste vulnerabilità possono essere nascoste in parti profonde del codice o essere il risultato di combinazioni complesse di fattori. Le aziende e gli sviluppatori di software adottano varie misure per ridurre il rischio di zero day, tra cui il codice sicuro, la revisione del codice, il testing approfondito e la collaborazione con ricercatori di sicurezza esterni attraverso programmi di bug bounty. Anche gli utenti possono proteggersi adottando buone pratiche di sicurezza, come mantenere i software aggiornati, utilizzare software di sicurezza e essere cauti nell'apertura di email e allegati sospetti. Tuttavia, la natura stessa delle vulnerabilità zero day rende difficile garantire una protezione completa, e quindi la vigilanza costante e la capacità di rispondere rapidamente a nuove minacce sono essenziali per limitare i danni che questi attacchi possono causare.