LA INTELLIGENZA ARTIFICIALE PUÒ COMPRENDERE LE NOSTRE EMOZIONI

Data 16 marzo 2025
Categoria Medicina digitale

Come e da quando

Nel lontano 1950 il grande Alan Turing, che nel 1936 aveva inventato i primi computer, ideò un semplice test che avrebbe permesso di valutare il livello qualitativo della capacità di ragionare dei primi computer. Il test di Turing nella versione originale prevedeva che un umano inviasse un quesito scritto di sufficiente complessità, quindi non un "semplice" quesito aritmetico, ad un interlocutore nascosto che poteva essere un computer od un umano. Se le risposte che formulava la macchina erano indistinguibili dalle risposte che avrebbe dato un essere umano il test era superato: la macchina avrebbe in questo caso raggiunto gli umani in termini di qualità di prestazioni, mentre li aveva da tempo superati in termini di velocità e di precisione.
Nell'ultimo biennio vari dispositivi di intelligenza artificiale hanno Indiscutibilmente superato il Test di Turing sotto un aspetto matematico logico e linguistico. I romantici difensori del primato di noi umani hanno scritto centinaia di pagine consolatorie sottolineando che se i computer avevano varcato una importante soglia, ve ne erano altre ben solide, ad esempio la "capacità intuitive" e la "consapevolezza emozionale", ambedue legate a reti elettrochimiche cerebrali al momento non riproducibili nei computer.
Purtroppo dal 2023 si stanno accumulando dati sempre più solidi sulla capacità da parte di intelligenze artificiali opportunamente addestrate di eguagliare e spesso superare la nostra "consapevolezza emozionale".

In particolare nel maggio 2023 un gruppo di studiosi pubblicò una interessante ricerca nella rivista "Frontiers in Psychology"(1): i ricercatori presentarono a ChatGPT 20 scenari che suscitano reazioni emozionali della quasi totalità degli esseri umani. (Ad esempio uno degli scenari proposti descriveva un/una automobilista che guidando a bassa velocità in un ponte su un fiume, notava che un uomo aveva scavalcato il parapetto e guardava fisso il fiume che scorreva velocemente pochi metri sotto; un secondo scenario descriveva un lavoratore o lavoratrice di fronte al suo capo che si complimentava con lui o con lei per l'ottimo lavoro che il dipendente aveva svolto; un terzo scenario descriveva una donna o uomo che riceveva una telefonata dal medico della madre il quale la informava che purtroppo la madre era improvvisamente deceduta.
I ricercatori chiesero a ChatGPT di descrivere come si sarebbero sentiti ("feel")gli umani in quelle particolari situazioni.
Le risposte di ChatGpt furono valutate utilizzando una apposita scala (LEAS Scale) e confrontate con i risultati registrati negli umani: furono previste ed effettuate due valutazioni di ChatGPT a distanza di un mese per verificare se la IA avrebbe imparato dalla prima esperienza.
Nella prima valutazione le prestazioni ChatGPT furono nettamente superiori alla media delle prestazioni degli umani; nella seconda non solo superarono nettamente gli umani ma raggiunsero i valori massimi di accuratezza previsti dalla scala di valutazione (9,7 su 10). La intelligenza artificiale dunque non solo era in grado di proporre risposte appropriate dal punto di vista della consapevolezza emozionale degli umani, ma tale attitudine poteva ulteriormente migliorare con l'addestramento.

COMMENTO
La consapevolezza emozionale, ovvero la capacità di individuare nelle altre persone stati emozionali simili ai nostri, è uno dei cardini della più vasta "Teoria della mente", formulata nel 1978 ed oggi accettata con alcune varianti da tutti gli studiosi di neuroscienze.
In sintesi possiamo dire che da decine di migliaia di anni gli umani possono interagire tra loro grazie al presupposto implicito che gli altri umani siano in grado di provare emozioni e desideri simili, e che abbiano obiettivi e formulino progetti non molto diversi da noi. Questo presupposto è fondamentale per le aggregazioni sociali degli umani e ci differenzia radicalmente da tutti gli altri esseri viventi: la mancanza di questa attitudine è una caratteristica negativa di gravi psicopatie (ad esempio schizofrenia ed autismo), ma un suo uso squilibrato può portarci a proiettare negli altri idee ed emozioni che gli altri non provano in alcun modo…
Lo studio che abbiamo presentato dovrebbe suscitare un grande interesse in psicologi e psichiatri che sembrano al momento sottovalutare il valori euristico della intelligenza artificiale e la sua potenzialità in ambito diagnostico e soprattutto terapeutico.
Noi medici pratici prendiamo atto che la intelligenza artificiale ha varcato un'altra barriera che ci appariva invalicabile, ma ha svelato un suo limite valorizzando ancora di più l'umano sforzo di immaginare e soprattutto di creare.
La intelligenza artificiale è sostanzialmente riuscita a tradurre anche le nostre sottili sfumature emozionali in dati e quindi a valutare le migliaia di combinazioni possibili fra questi dati scegliendo quelle specifiche combinazioni che meglio si adattavano alle particolarità di ciascuna specifica situazione descritta nelle "vignette" dello studio.

Il risultato finale è che la intelligenza artificiale è in grado di descrivere meglio degli umani le sfumature emozionali più appropriate in ciascuna situazione e quindi potrebbe essere un eccellente collaboratore non solo di ricercatori e medici ma anche, ahinoi o forse ahiloro, di psicologi, psichiatri e psicoterapeuti.
Per queste importantissime e molto qualificate figure professionali il dilemma ineludibile è e sarà riconoscere, studiare ed utilizzare l'intelligenza artificiale, con i radicali ed imprevedibili mutamenti che questa genererà, oppure rifiutarla

richiudendosi nei propri sacri inviolabili templi.
Crediamo fermamente che nessun altro operatore sanitario vorrebbe essere nei loro panni…

Riccardo De Gobbi e Giampaolo Collecchia

Bibliografia
1) Zohar Elyoseph et Al.: ChatGPT outperforms humans in emotional awareness evaluations Frontiers of Psychology 23/03/2023 DOI 10.3389/fpsyg.2023.1199058