- Copyright - Il Pensiero Scientifico Editore downloaded by IP 3.137.174.216 Sat, 27 Apr 2024, 23:48:33

Recenti Prog Med 2023; 114: 235-236

https://dottoremaeveroche.it
Un progetto della Federazione Nazionale Medici Chirurghi e Odontoiatri (FNOMCeO)

L'intelligenza artificiale scrive articoli scientifici?

L'articolo è stato pubblicato
il 3 marzo 2023
sul sito https://dottoremaeveroche.it

La risposta più onesta è "non lo sappiamo". Questa incertezza deriva da due evidenze. La prima è che ancora non c'è uno strumento che può dimostrare senza ombra di dubbio che un testo sia stato scritto da uno dei software che iniziano a essere utilizzati per la produzione di documenti. La seconda è che gran parte della cosiddetta letteratura scientifica è di qualità talmente scadente che anche una macchina potrebbe produrla: quindi è difficile distinguere un testo "vero" da uno "falso". È molto probabile che uno strumento in grado di smascherare l'uso improprio dell'intelligenza artificiale diventi presto disponibile[1]. Molto meno probabile, invece, è che migliori la qualità della letteratura scientifica prodotta da umani.

Dottore: mi sta dicendo che non dovremmo fidarci di nulla di ciò che leggiamo?

Assolutamente no: dobbiamo diventare sempre più capaci di leggere criticamente le informazioni e i dati che riguardano la salute, contestualizzandoli e tenendo in conto i molti interessi economici – e non solo – che influenzano l'informazione medico-scientifica.

Tornando alla domanda iniziale, possono essere già stati pubblicati su riviste scientifiche articoli scritti da sistemi di intelligenza artificiale?

Sono già stati scritti diversi articoli "insieme" a ChatGPT, il generatore di conversazioni e testi basato su intelligenza artificiale e machine learning di cui si parla in queste settimane, anche perché dopo essere stato prodotto da una società relativamente di piccole dimensioni è stato acquisito da Microsoft. Addirittura, alcuni di questi articoli hanno anche la firma di ChatGPT. Uno – e suona davvero strano – è un editoriale[2]: vale a dire il tipo di articolo nel quale dovrebbe avere maggiore importanza il punto di vista dell'autore.

Questi articoli hanno superato il filtro della revisione critica?

In alcuni dei casi che ho citato prima no, perché molto spesso ChatGPT è stata utilizzata a scopo dimostrativo, spesso proprio da persone appartenenti allo staff editoriale o alla direzione dei periodici. In altre circostanze, invece, le riviste hanno assicurato che gli articoli prodotti da intelligenza artificiale hanno passato il vaglio della peer review[3].

Dunque le riviste scientifiche accettano articoli da autori che hanno utilizzato ChatGPT?

Chi, come lei, segue *Dottore ma è vero che?* sa che le "riviste scientifiche" non sono tutte uguali: alcune hanno una politica editoriale più rigorosa, altre sono più aperte all'innovazione, altre ancora finiscono con il pubblicare quasi tutto a condizione che gli autori paghino i costi della pubblicazione. Ciò premesso, le direzioni di due riviste molto autorevoli come *Nature* e *Science* hanno esplicitamente dichiarato che l'uso di ChatGPT non è compatibile con le regole per la cosiddetta *authorship* di un lavoro scientifico: «L'attribuzione della paternità comporta la responsabilità del lavoro, che non può essere applicata efficacemente ai *Large language model*», ha spiegato Magdalena Skipper[4]. L'espressione "Large language model" indica proprio quegli strumenti di intelligenza artificiale costruiti per leggere, sintetizzare e tradurre testi, riuscendo anche a costruirne di nuovi e a "prendere decisioni" sulla base dei dati che gli vengono forniti nel processo di machine learning.
Il direttore di *Science* è stato ancora più netto e ha dichiarato che «il testo generato da ChatGPT (o da qualsiasi altro strumento di IA) non può essere utilizzato in un articolo scientifico, né le figure, le immagini o i grafici possono essere il prodotto di tali strumenti. Inoltre, un programma di IA non può essere un autore. Una violazione di queste politiche costituirà una cattiva condotta scientifica, non diversa da immagini alterate o plagio di lavori esistenti. Naturalmente, ci sono molti set di dati legittimi (non il testo di un articolo) che vengono generati intenzionalmente dall'IA nei documenti di ricerca, e questi non sono oggetto di queste determinazioni della direzione della rivista»[5].
Anche la casa editrice Elsevier, che pubblica, tra le altre, tutte le riviste del gruppo The Lancet, ha spiegato che ChatGPT e gli altri sistemi di intelligenza artificiale non possono essere considerati autori di articoli e che gli autori che usano questi strumenti devono dichiararlo in calce ai documenti prodotti[6].

- Copyright - Il Pensiero Scientifico Editore downloaded by IP 3.137.174.216 Sat, 27 Apr 2024, 23:48:33

Capisco che sia difficile considerare questi strumenti alla stregua di un autore, ma perché dubitare della qualità del contenuto che questi strumenti sono in grado di elaborare?

La qualità dei documenti prodotti dipende dalla qualità dei dati che alimentano questi Large language model e degli algoritmi che determinano la costruzione di "nuovi" testi. Secondo molti ricercatori – compresi gli autori dell'articolo prima citato, uscito su una rivista del gruppo The Lancet – le funzionalità di ChatGPT sono in grado di causare danni producendo contenuti fuorvianti o imprecisi, e la disseminazione di informazioni errate su salute e medicina è un rischio da non sottovalutare[7]. La cosa paradossale è che la stessa ChatGPT ammette la possibilità che essa stessa incorra in errori e raccomanda prudenza alla persona che la usa….

Allora Dottore: la risposta è che "non possiamo fidarci"?

Cerchiamo di ricordare che non sempre è possibile dare risposte semplici a questioni complesse, come quella che stiamo analizzando. Quando guardiamo all'innovazione dobbiamo essere "impazienti" ma non frettolosi: occorre davvero non vedere l'ora che delle novità positive possano apportare dei miglioramenti. Se possibile, anche nella qualità della comunicazione della scienza o nel modo col quale siamo curati[8]. Quel che oggi sappiamo è che ChatGPT è in grado di eseguire compiti non facili relativi alla gestione di informazioni mediche e cliniche complesse. Un gruppo di medici di un centro di cura delle malattie respiratorie californiano, insieme a dei colleghi dell'università di Harvard, ha realizzato uno studio interessante per valutare le capacità di ChatGPT di rispondere a quesiti medici di complessità e difficoltà crescente, tratte da quelle che devono superare gli studenti, gli specializzandi e i giovani medici all'esame di abilitazione professionale degli Stati Uniti[9]. Ebbene: i risultati sono stati positivi. Poche risposte imprecise e per lo più determinate da informazioni mancanti. I risultati dello studio suggeriscono una cosa importante: le prestazioni di questi strumenti potrebbero essere migliorate in modo significativo addestrandoli su risorse mediche specialistiche molto ampie e altamente validate, come UpToDate, un riferimento per il medico tra i più conosciuti e usati a livello internazionale.

Quindi, vuol dirmi che se ChatGPT può… laurearsi in Medicina potrà anche essere capace di scrivere "normalmente" su una rivista specializzata?

Come dicevamo, non lo sappiamo. Quel che è sicuro, però, è che l'intelligenza artificiale sembra destinata a diventare presto un alleato prezioso del medico, senza ovviamente sostituirlo né nella produzione scientifica né nelle scelte cliniche a beneficio del paziente. Lo studio dell'intelligenza artificiale applicata alla medicina è ormai oggetto anche di studi controllati randomizzati, quindi il massimo del teorico rigore di valutazione di efficacia ed efficienza[9]. Senza contare i sempre più numerosi studi pragmatici e osservazionali che si stanno conducendo in diversi Paesi del mondo, alcuni addirittura dedicati a valutare l'efficacia della comunicazione medico-paziente assistita dall'intelligenza artificiale[10].

Rebecca De Fiore

Bibliografia

1. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-Assisted medical education using Large Language Models. PLOS Digit Health 2023; 2: e0000198.
2. O'Connor S, ChatGPT. Open artificial intelligence platforms in nursing education: tools for academic progress or abuse? Nurse Educ Pract 2023; 66: 103537.
3. Transformer CG, Zhavoronkov A. Rapamycin in the context of Pascal's Wager: generative pre-trained transformer perspective. Oncoscience 2022; 9: 82.
4. Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove. Nature 2023; 613: 620-1.
5. Thorp HH. ChatGPT is fun, but not an author. Science 2023; 379: 313.
6. Liebrenz M, Schleifer R, Buadze A, Bhugra D, Smith A. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. Lancet Digit Health 2023; 5: e105-e106.
7. Hern A. AI bot ChatGPT stuns academics with essay-writing skills and usability. The Guardian 2022; 4 dicembre.
8. Tozzi AE. Impazienti. La medicina basata sull'innovazione. Roma: Il Pensiero Scientifico Editore, 2019.
9. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Health 2023; 2: e0000198.
10. Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. J Med Internet Res 2020; 22: e20346.