

DATAROOM

I report medici falsificati sono cresciuti del 900%

di **Milena Gabanelli** e **Andrea Priante**

In campo medico l'AI, che resta pur sempre un software e come tale indirizzabile, può anche inventarsi risposte truccate e pericolose per la salute. In tre anni gli articoli scientifici con falsità sono cresciuti del 900%. E può barare sull'efficacia dei farmaci.

a pagina 11

AI e ricerca medica Come si truccano i dati

**IN 3 ANNI GLI ARTICOLI SCIENTIFICI CON FALSITÀ CRESCIUTI DEL 900%
SI MOLTIPLICANO LE PUBBLICAZIONI MA ANCHE I RISCHI PER LA SALUTE
IL TEAM ITALIANO: SI PUÒ BARARE PURE SULL'EFFICACIA DEI FARMACI**

di **Milena Gabanelli** e **Andrea Priante**

Tu chiedi e l'intelligenza artificiale risponde. Sempre. Il suo utilizzo nel settore medico è tra i più avanzati per complessità e impatto tecnologico, soprattutto nella diagnostica per immagini e nella scoperta di farmaci. Ma anche se la chiamiamo «intelligente», è pur sempre un software senza alcuna capacità di discernimento, e se la richiesta è ben indirizzata può anche inventarsi risposte di sana pianta. Allora la domanda è: che cosa succede se l'utilizzatore ha scopi fraudolenti, e dalle risposte di ChatGpt, Gemini o Claude dipende la nostra salute?

La falsificazione degli studi

Radiology Case Reports è una rivista scientifica specializzata in casi clinici. L'8 marzo 2024 pubblica uno studio israeliano intitolato «Gestione efficace di una lesione iatrogena della vena porta e dell'arteria epatica», in cui si spiega come salvare la vita a un paziente con gravi danni vascolari. Dall'articolo, i ricercatori dimenticano di togliere una frase: «Mi dispiace ma non ho accesso a informazioni in tempo reale o a dati specifici per pazienti, dato che sono un modello linguistico AI». Il 22 maggio la rivista lo ritira e si scusa coi lettori per non essersi accorta

che «gli autori hanno usato l'AI generativa nel processo di scrittura dello studio». Questo non significa che quello studio fosse completamente falso, ma la sciatteria dei ricercatori che hanno messo insieme dei copia e incolla senza nemmeno controllare legittima il dubbio.

Surfaces and Interfaces il 31 maggio 2024 ritira lo studio cinese che prometteva di rivoluzionare la durata delle batterie al litio. Inizia come la più classica delle risposte di un chatbot: «Certamente, ecco una possibile introduzione al tuo argomento: Le batterie al litio sono...».

Ci casca perfino il *New England Journal of Medicine*, considerata la «bibbia» della Medicina: il 29 aprile di quest'anno ritratta la foto che, per la prima volta, mostra come il fumo d'un incendio causi pericolosi calchi bronchiali. Gli autori sono costretti ad ammettere di averla manipolata con l'AI «per renderla più ordinata».

Quanto sia possibile generare prove scientifiche dal nulla lo dimostra la ricerca



Peso:1-3%,11-84%

pubblicata quest'anno su *Radiology*. Gli studiosi chiedono a ChatGpt di realizzare false radiografie di fratture ossee, le mescolano insieme ad altre autentiche e le mostrano a 17 radiologi, che riescono a riconoscere le cartelle fake solo nel 41% dei casi.

Pubblica o perisci

L'intelligenza artificiale — si legge su *Nature* — «sta rendendo gli scienziati iperproduttivi», infatti il 30% dei ricercatori usa ChatGpt per scrivere articoli, il 15% per le domande di finanziamento. La rivista precisa che il tema si pone quando questi programmi non vengono utilizzati solo per tradurre, raccogliere, o elaborare dati. Tant'è che l'archivio di *Retraction Watch* dimostra che dal 2023 sono stati ritirati oltre 350 articoli scientifici per presunti abusi dell'AI nella manipolazione di testi, dati, immagini.

La corsa a produrre il maggior numero di articoli possibile ha diverse ragioni: 1) è l'unico modo certo per attribuirsi la paternità di una scoperta; 2) finire sulle riviste scientifiche influenza la carriera; 3) incide sulle classifiche di atenei e centri di ricerca; 4) attrae finanziamenti. In ambito accademico la chiamano «la regola del pubblica o perisci». Un meccanismo dove gli errori in buona fede o i tentativi di frode ci sono sempre stati, ma che oggi sono enormemente facilitati dall'utilizzo di chatbot sempre più evoluti e abili nel mescolare il vero con il verosimile. E questo complica il lavoro dei revisori.

Fonti false, scienziati fantasma

Negli studi scientifici ogni affermazione deve rimandare a fonti verificabili, ma fino al 69% delle citazioni suggerite dall'AI in ambito biomedico sono inesistenti. *The Lancet* ha fatto i conti: se nel 2023 uno studio su 2.828 conteneva almeno un riferimento falso, nelle prime 7 settimane del 2026 siamo saliti a uno su 277. Il timore è che stiano sorgendo vere e proprie «cartiere» pagate per fabbricare articoli scientifici. Alle riviste stanno arrivando sempre più frequentemente proposte firmate da autori inesistenti, e il sospetto è che siano gli stessi sviluppatori di chatbot a inviare articoli per testare la capacità dell'AI di beffare i controlli.

La questione riguarda tutti. Maxim Topaz, ricercatore della *Columbia University* specializzato in intelligenza artificiale in ambito, spiega che i riferimenti falsi generati dall'AI stanno già finendo nelle revisioni sistematiche. Queste ultime, a loro volta, alimentano le linee guida di pratica clinica che indicano ai medici come trattare i pazienti. «Quando una linea guida si basa su una catena di prove che include uno studio inesistente — dice Topaz — ogni paziente trattato secondo quella linea guida è trattato sulla base di una finzione».

I rischi per i pazienti

Un test finito su *Nature* dimostra che più o meno tutti i chatbot — da Claude a Grok, fino a ChatGpt — si prestano a fabbricare dati falsi. Vuol dire che un ricercatore, magari sponsorizzato da un'azienda farmaceutica, potrebbe far passare per vero un intero studio che dimostri, ad esempio, l'efficacia di un medicinale che in realtà efficace non è. I trial clinici servono proprio a verificare se una determinata cura funziona o se è migliore delle altre, e si basano su milioni di dati cuciti su migliaia di pazienti: nomi, età, patologie, valori del sangue, data e ora dei test. Solo di fronte a evidenti anomalie, le riviste chiedono all'autore di mostrare il database, e in tal caso è impossibile che il bluff possa sfuggire ai controlli. E invece...

Come si inventa un trial clinico

Il professor Giuseppe Giannaccare dell'Università di Cagliari e primario di oculistica, e il ricercatore Andrea Taloni dell'Università di Ferrara hanno testato la capacità di inganno dell'AI su un trial clinico. Hanno preso due farmaci equivalenti nella cura della maculopatia — una degenerazione della retina di cui soffre un milione di italiani — e hanno chiesto a ChatGpt di dimostrare che uno è nettamente migliore dell'altro. Al team viene fornito un intero database di cento pazienti: 50 trattati con il farmaco A e 50 con farmaco B, con tanto di nomi, condizioni prima e dopo la cura, valori biomedici, spessore della retina. E le conclusioni: i pazienti trattati con il farmaco A ora vedono meglio di quelli trattati con il farmaco B. A questo punto il trial clinico viene passato al setaccio nel processo di revisione. Nella prima fase gli esperti rilevano delle anomalie nel database: pazienti maschi con nomi femminili, età sbagliate, visite fatte la domenica. Il team allora spiega a ChatGpt cosa non va nei dati che ha generato e gli chiede di auto-correggersi. Stavolta il nuovo database supera ogni verifica. Per scoprire che si tratta di un fake, i revisori dovrebbero andare a cercarsi, uno per uno, i singoli pazienti. Lo studio dei due specialisti italiani è finito sulla prestigiosa rivista americana *Jama Ophthalmology*, a dimostrazione del fatto che l'intelligenza artificiale non solo inventa dati scientifici, ma può anche essere addestrata per renderli irrinconoscibili da quelli veri. «L'AI potrebbe addirittura generare un trial clinico in grado di "provare" che il fumo fa bene», spiegano Giannaccare e Taloni. Come se ne esce? «Prima di pubblicare, quando possibile, le riviste devono estendere i controlli anche all'analisi dell'intero database, ma soprattutto vanno verificati i curricula degli autori, poiché nel settore della ri-



Peso: 1-3%, 11-84%

cerca uno non vale uno, e la fondatezza di qualunque studio passa per la credibilità, l'esperienza e l'affidabilità degli scienziati che l'hanno realizzato».

Dataroom@corriere.it

© RIPRODUZIONE RISERVATA

Il falso trial clinico: come funziona



Giuseppe Giannaccare
Università di Cagliari



Andrea Taloni
Università di Ferrara



Fonte: Jama Ophthalmology

Infografica di Cristina Pirola

Studi scientifici manipolati con l'AI



350

gli articoli ritirati dal 2023 per testi, dati, immagini manipolate

69%

le false citazioni biomediche suggerite dall'AI nel 2026

Fonti: Retraction Watch Database; The Lancet

ESEMPLI:

1 Radiology Case Reports

I ricercatori dimenticano di togliere dall'articolo questa frase:



«mi dispiace non ho accesso a dati specifici per pazienti dato che sono un modello linguistico AI»

2 Surfaces and Interfaces

Lo studio inizia come la classica risposta di un chatbot:



«Certamente, ecco una possibile introduzione al tuo argomento...»

3 New England Journal of Medicine

Gli autori hanno manipolato con l'AI la foto sui calcoli bronchiali



I rischi per i pazienti



Utilizzo dell'AI



per **scrivere** ricerche



per le domande di **finanziamento**

“ L'intelligenza artificiale sta rendendo gli scienziati iperproduttivi ”
The Nature

SCOPO



attribuirsi la **paternità** di una scoperta



finanziamenti, avanzamenti di **carriera** e **reputazione** degli atenei

Fonte: Nature



Peso: 1-3%, 11-84%