

**13. Gennaio**

## **I paradossi dell'intelligenza artificiale in sanità**

*"E' un paradosso voler costruire la configurazione di uno spazio con coordinate di punti che non esistono."*

Tutti i sistemi sanitari hanno negli ultimi anni sottolineato la necessità di controllare i costi di gestione legati alla introduzione di tecnologie IA soprattutto perché l' inflazione non riesce a tenere il passo con le richieste operative.

Nel 2025, questi sforzi di contenimento dei costi persisteranno, e con una sfida in più: le spese per l'adozione di strumenti di intelligenza artificiale e le "babysitter" necessarie per supervisionarli, in altre parole i costi legati alla manodopera in intelligenza artificiale.

Al momento è un paradosso: mentre gli strumenti di intelligenza artificiale sono spesso pubblicizzati come il futuro dell'efficienza e dell'automazione, attualmente richiedono una **significativa supervisione e attenzione umana** per funzionare correttamente.

Nell'immaginario tutti pensano che l'intelligenza artificiale ci aiuterà con il nostro accesso e la nostra capacità e migliorerà l'assistenza e così via tutto ciò è potenzialmente vero, ma se aumenta il costo dell'assistenza del 20%, è fattibile?"

Inoltre l' intelligenza artificiale porta con sé una serie di sfide, tra cui il decadimento dell' algoritmo, in cui l'accuratezza predittiva diminuisce nel tempo. L'integrazione di questi strumenti nelle operazioni del sistema sanitario richiede un monitoraggio continuo e personale dedicato per garantire equità, affidabilità ed efficacia.

A Stanford, la verifica dell'affidabilità di due modelli di intelligenza artificiale ha richiesto da otto a dieci mesi e 115 ore di lavoro manuale. **La parte difficile non è determinare se gli strumenti funzionano, ma se continuano a funzionare.**

Il vero problema è che molti sistemi sanitari non hanno il talento e l'infrastruttura per soddisfare queste richieste.



**Robert Califf** commissario della FDA La FDA ha notato questo rischio all'inizio del 2024, con ha espresso preoccupazione per il fatto che i sistemi sanitari non hanno l'infrastruttura e gli strumenti per prendere le decisioni più importanti sull'efficacia di un'applicazione di intelligenza artificiale per i risultati *sanitari*.

*"Ho guardato in lungo e in largo", ha detto il dott. Califf a un recente panel dell'agenzia sull'IA. "Non credo che ci sia un singolo sistema sanitario, negli Stati Uniti, in grado di convalidare un algoritmo di IA che viene messo in atto in un sistema di assistenza clinica".*

I responsabili dei sistemi sanitari hanno anche espresso un sano scetticismo su quanta trasformazione sia possibile quando le soluzioni di intelligenza artificiale vengono aggiunte a sistemi frammentati o imperfetti, ampiamente diffusi nell'assistenza sanitaria.

paradossi

## I paradossi di oggi sono i pregiudizi di domani.

Anche se i paradossi creano stupore, tuttavia mettono in evidenza delle criticità che, una volta risolte, conducono a nuove e più profonde verità. I paradossi contribuiscono al progresso della conoscenza umana e rimettono insieme le cose che la logica divide

Ben lungi dall'essere uno strumento "imposta e dimentica" i controlli tecnici di routine rilevano che gli algoritmi decadono peggiora alcuni parametri quali la predittività. Molte istituzioni non monitorano regolarmente le prestazioni, : i sistemi di intelligenza artificiale richiedono un monitoraggio e un personale costanti per essere implementati e per continuare a funzionare correttamente.

In sostanza: servono persone e più macchine per garantire che i nuovi strumenti non creino problemi.

L'intelligenza artificiale è già diffusa nell'assistenza sanitaria. Gli algoritmi vengono utilizzati per prevedere il rischio di morte o peggioramento dei pazienti, per suggerire diagnosi o selezionare i pazienti, per registrare e riassumere le visite per risparmiare lavoro ai medici e per approvare le richieste di risarcimento assicurativo .

Se gli evangelisti della tecnologia hanno ragione, la tecnologia diventerà onnipresente e redditizia. *La società di investimenti Bessemer Venture Partners ha identificato circa 20 startup di intelligenza artificiale focalizzate sulla salute, sulla buona strada per realizzare 10 milioni di dollari di fatturato ciascuna in un anno. La FDA ha approvato quasi un migliaio di prodotti di intelligenza artificiale.*

Valutare se questi prodotti funzionano è una sfida. Valutare se continuano a funzionare, o se hanno sviluppato l'equivalente software di una guarnizione bruciata o di un motore che perde, è ancora più complicato.

Prendiamo uno studio recente della **Yale Medicine** che valuta sei "sistemi di allerta precoce", che avvisano i medici quando è probabile che i pazienti peggiorino rapidamente. Un supercomputer ha elaborato i dati per diversi giorni, ha affermato **Dana Edelson**, medico presso l'Università di Chicago e co-fondatore di un'azienda che ha fornito un algoritmo per lo studio. Il processo è stato fruttuoso, mostrando enormi differenze nelle prestazioni tra i sei prodotti.

Non è facile per ospedali e provider selezionare gli algoritmi migliori per le loro esigenze. Il medico medio non ha un supercomputer a disposizione e non esiste Consumer Reports per l'IA.



**"Non abbiamo standard"**, ha affermato **Jesse Ehrenfeld**, ex presidente immediato dell'American Medical Association. **"Non c'è nulla che io possa indicarvi oggi che sia uno standard su come valutare, monitorare, guardare le prestazioni di un modello di algoritmo, abilitato o meno dall'intelligenza artificiale, quando viene distribuito"**.

Forse il prodotto AI più comune negli studi medici è chiamato documentazione ambientale, un assistente abilitato dalla tecnologia che ascolta e riassume le visite dei pazienti, **"Non esiste uno standard in questo momento per confrontare l'output di questi strumenti"**.

Ed è un problema, quando anche piccoli errori possono essere devastanti. Un team della Stanford University ha provato a usare grandi modelli linguistici, la tecnologia alla base di popolari strumenti di intelligenza artificiale come ChatGPT, per riassumere la storia medica dei pazienti. Hanno confrontato i risultati con ciò che avrebbe scritto un medico.

***"Anche nel caso migliore, i modelli avevano un tasso di errore del 35. A volte le ragioni per cui gli algoritmi falliscono sono abbastanza logiche. Ad esempio, le modifiche ai dati sottostanti possono eroderne l'efficacia, come quando gli ospedali cambiano i fornitori di laboratorio.***

A volte, però, le insidie si aprono senza apparente motivo.



**Sandy Aronson**, dirigente tecnico del programma di medicina personalizzata

del Mass General Brigham di Boston, ha affermato che quando il suo team ha testato un'applicazione pensata per aiutare i consulenti genetici a reperire la letteratura pertinente sulle varianti del DNA, il prodotto ha manifestato **"non determinismo"**, ovvero quando veniva posta la stessa domanda più volte in un breve lasso di tempo, forniva risultati diversi.

Aronson è entusiasta delle potenzialità dei grandi modelli linguistici per riassumere le conoscenze per i consulenti genetici oberati di lavoro, ma "la tecnologia deve migliorare".

Se le metriche e gli standard sono scarsi e gli errori possono emergere per strane ragioni, cosa devono fare le istituzioni? Investire molte risorse.

A Stanford, ci sono voluti dagli otto ai dieci mesi e 115 ore-uomo solo per verificare due modelli per correttezza e affidabilità.

**KFF**  
Health  
News

Un team di esperti intervistati da **KFF Health News** hanno lanciato l'idea di un'intelligenza artificiale che monitori l'intelligenza artificiale, con qualche mago dei dati (umano) che monitori entrambi.

Tutti hanno riconosciuto che ciò richiederebbe alle organizzazioni di spendere ancora più soldi, una richiesta difficile data la realtà dei budget ospedalieri e la limitata offerta di specialisti in tecnologia AI.

***"È fantastico avere una visione in cui stiamo sciogliendo gli iceberg per avere un modello che monitora il loro modello", ha concluso il team, "Ma è davvero questo che volevamo. Di quante altre intelligenze naturali avremo bisogno per far funzionare quelle artificiali ?***