

# Chi controllerà intelligenze artificiali sempre più essenziali per il nostro futuro?

*L'AI non è nuova, e nemmeno i tentativi di regolarla. Ma i modelli generativi sono un grande balzo in avanti e c'è un crescente interesse per affrontarne gli impatti indesiderati.*

L'avvento di un' intelligenza artificiale (AI) sempre più performante ha aperto il terreno a una serie di questioni, ma due in particolare hanno suscitato ampi dibattiti. In che modo le nuove intelligenze artificiali cambieranno gli umani e la società? Metteranno in pericolo i posti di lavoro del futuro oppure copriranno solo quelli più meccanici e routinari liberando così la creatività delle persone? E come trovare il punto di equilibrio tra lo sviluppo dell'AI e la condizione umana? Va ricordato che gli algoritmi di apprendimento automatico sono **utilizzati con successo in diversi campi** per supportare gli umani. Tuttavia questi sistemi **non hanno neanche lontanamente un'intelligenza a livello umano** e non sono in grado prendere decisioni. Come ha spiegato **Yann LeCun**, Chief AI Scientist di Meta, "siamo portati a pensarlo perché quei sistemi sono molto avanzati nell'uso del linguaggio, ma la loro capacità di pensare per capire come funziona il mondo è molto limitata".

La preoccupazione di alcuni esperti per questo potere in crescita è incentrata però sul **problema dell'allineamento**: come garantire cioè che l'AI stia facendo (e farà) ciò che gli umani vogliono che faccia. "Il mio grande timore è che prima o poi qualcuno collegherà a loro (le AI, ndr) la capacità di creare i propri obiettivi secondari", ha detto **Goffrey Hinton**, uno dei pionieri del deep learning, aggiungendo: "Non credo che nessuno sappia davvero gestire questi problemi". Hinton ha lasciato Google all'inizio di maggio, in parte a causa dell'età, ma anche perché ha dichiarato di aver **cambiato idea sul rapporto tra umani e intelligenza digitale**. In un'intervista con il *New York Times*, Hinton ha affermato infatti che l' intelligenza artificiale generativa potrebbe diffondere disinformazione e, alla fine, minacciare l'umanità.

Non è il solo a sostenere questo scenario: centinaia di scienziati, sviluppatori e Chief executive hanno messo in guardia sui pericoli dell'AI in più lettere aperte: una pubblicata a maggio denunciava proprio il **rischio di estinzione per l'umanità**. "Mitigare il rischio di estinzione provocata dall'AI dovrebbe essere una priorità globale insieme ad altri rischi su scala sociale come le pandemie e la guerra nucleare", si legge nella breve dichiarazione. **Sam Altman**, amministratore delegato di OpenAI, creatore di ChatGPT, **Demis Hassabis**, amministratore delegato di Google DeepMind, e **Dario Amodei** di Anthropic hanno tutti messo la loro firma.

L'idea che, a un certo punto nel futuro, **la tecnologia diventerà così avanzata da superare l'intelligenza e il controllo umano** non è nuova. Era il 1983 quando **Vernor Vinge** coniava l'espressione "singolarità tecnologica". In pratica, si asseriva che qualora i computer avessero superato le capacità cognitive degli esseri umani, sarebbe emerso un nuovo livello di intelligenza irraggiungibile per le persone. Ai giorni nostri, la singolarità dell'intelligenza artificiale implica allora la creazione di un'AI così avanzata che può migliorare continuamente sé stessa a un ritmo più veloce di quanto facciano gli umani eseguendo lo stesso compito. Un momento che secondo i futuristi potrebbe essere non così lontano nel tempo, se è vero che Ray Kurzweil ha previsto l'arrivo della singolarità tecnologica entro il 2045.

La maggior parte degli esperti ritiene però che il punto di vista di Hinton e soci tenda a sovrastimare i rischi delle AI. Ad esempio, lo stesso LeCun usa toni completamente diversi: "Credo che **le macchine intelligenti inaugureranno un nuovo rinascimento per l'umanità**, una nuova era di illuminazione", afferma LeCun, che rifiuta l'idea che le macchine domineranno gli esseri umani "semplicemente perché sono più intelligenti, per non dire che distruggeranno gli esseri umani".

Poi c'è il tema del lavoro: alcuni tecnologi temono che l'intelligenza artificiale diventi così potente da essere in grado di svolgere la maggior parte dei lavori degli umani, portando alla disoccupazione di massa. Lavoratori da sussidiare magari con un reddito universale incondizionato o con gli stessi proventi delle macchine. Certo, è probabile che l' intelligenza artificiale sostituisca alcuni lavori: alcune attività di routine, ripetitive o pericolose, potranno essere automatizzate. Ma le nuove Ai creeranno anche nuovi posti di lavoro e aumenteranno la produttività. C'è un numero crescente di ricerche, raccolte per esempio dall'*Economist*, che evidenziano questi effetti benefici.

Pubblicità

<#? QUI-PUBBLICITA-SCORM1-[EL0790] ?#>

## Oligopolio

Il fatto che questi avvertimenti apocalittici siano stati lanciati dagli stessi soggetti che hanno contribuito a creare (e tutt'ora sviluppano) l' Ai ha fornito un ulteriore assist ai critici. E ha segnato una divisione più netta all'interno del campo delle intelligenze artificiali. Sicché molti osservatori hanno ricordato che **l' Ai è elaborata da tempo da una sorta di oligopolio**, il che prefigura che i mercati serviti da queste tecnologie saranno sempre più dominati da un piccolo numero di aziende.

Se guardiamo al panorama attuale, i modelli di Ai di grandi dimensioni, come ChatGpt-4 di Open Ai e Lamda di Google sono costosi da addestrare e impegnativi da sviluppare, cosa che rappresenta un'importante **barriera di entrata per altre aziende**. Di conseguenza, restano concentrati nelle mani di una manciata di gruppi, come Google, Microsoft, Baidu, e di startup ben finanziate come Open Ai.

Ma un punto cruciale nella regolazione delle nuove intelligenze artificiali riguarda anche **i dati che le alimentano, e che sono alla base delle loro consapevolezze**. L' Ai può fornire indicazioni che influiscono sull'ammissione di una persona a una scuola, sull'autorizzazione per un prestito bancario o sull'assunzione dopo un colloquio di lavoro. È abbastanza risaputo che i sistemi di intelligenza artificiale possono mostrare pregiudizi che derivano dalla loro programmazione e dalle fonti di dati; ad esempio, il software di apprendimento automatico potrebbe essere addestrato su un set di dati che sotto-rappresenta un particolare genere o gruppo etnico. Ecco perché oggi, avvertono gli esperti, è più che mai cruciale studiare i dataset su cui queste tecnologie formano le loro opinioni.

## I tre approcci

Le democrazie e le autocratie affrontano in modo diverso la sfida di controllare non solo gli algoritmi e i dati che li addestrano, ma anche e soprattutto le implicazioni etiche e sociali dell' intelligenza artificiale. **Tre principali approcci sono emersi** in tutto il mondo.

Sebbene la **Cina** abbia rilasciato i propri principi etici di intelligenza artificiale nel 2019, la tecnologia è rientrata **nell'orbita del governo** ed è stata sviluppata soprattutto per facilitare la sorveglianza di massa e controllare le minoranze.

L'approccio degli **Stati Uniti** all'intelligenza artificiale è stato storicamente assai leggero, *laissez-faire*, caratterizzato dall'idea che **le aziende debbano mantenere il controllo dello sviluppo industriale** e dei criteri legati alla governance. In altre parole, c'è la convinzione che norme e regolamenti statali troppo gravosi rappresentino degli ostacoli all'innovazione.

Ma sul fronte americano alcuni segnali potrebbero svelare un cambio di rotta. Di recente, infatti, Joe Biden ha annunciato nuove azioni per produrre una **"Ai responsabile"**. Ed è stato chiaro sul fatto che quando si tratta di intelligenza artificiale, "dobbiamo mettere al centro le persone e le comunità sostenendo un'innovazione responsabile che serva il bene pubblico, proteggendo al contempo la nostra società, sicurezza ed economia". E in questa cornice si può ricordare che la Casa Bianca ha annunciato lo stanziamento 140 milioni di dollari che saranno destinati al lancio di sette nuovi istituti nazionali di ricerca sull' Ai incentrati

sullo sviluppo di un'intelligenza artificiale "**etica e trasformativa per il bene pubblico**".

L'Unione europea è stata, per molti versi, **all'avanguardia nella regolamentazione** dei dati e nella governance dell'AI. Basti pensare al regolamento generale sulla protezione dei dati (GDPR), entrato in vigore nel 2018, che è diventato un modello per leggi simili in tutto il mondo. La continuità con quest'approccio è segnata dalle **nuove rigide regole** contenute nella proposta sull'AI Act che è stata votata a giugno dal Parlamento Ue. I deputati hanno **ampliato la classificazione delle aree ad alto rischio** per includere i danni alla salute, alla sicurezza, ai diritti fondamentali o all'ambiente delle persone. Hanno anche aggiunto a questo elenco i sistemi di AI **usati per influenzare gli elettori nelle campagne politiche** oppure suggerire contenuti mirati agli utenti in base ai loro comportamenti d'acquisto. I modelli generativi, come Chat Gpt, dovrebbero rispettare ulteriori requisiti di trasparenza, come rivelare che il contenuto è stato generato dall'intelligenza artificiale, progettare il modello per impedire che generi contenuto illegale e pubblicare riepiloghi dei dati protetti da copyright utilizzati per la formazione.

Anche dal successo di queste iniziative dipenderà l'affermazione o meno di un'**Ai umanistica**. Il principio di base di questo percorso è utilizzare l'intelligenza artificiale in prodotti che aumentano la capacità umana o **lavorano con gli umani per raggiungere i loro obiettivi**, evitando danni inutili e riducendo quelli collaterali. Nel convegno sull'etica e l'intelligenza artificiale promosso dal Cortile dei Gentili e FUTURAnetwork al recente Festival dello sviluppo sostenibile, la sociologa **Carla Collicelli**, ha evidenziato la necessità di trovare "il **senso del limite** e di lavorare **per il bene comune**". Per evitare, ha detto, "che le tecnologie vadano a vantaggio solo di alcune categorie o servano ad accumulare profitti senza tenere conto delle implicazioni etiche e sociali". Ma è importante farlo adesso.

*Andrea De Tommasi*

Fonte: Futuranetwork



Licenza Creative Commons

I contenuti presenti sul sito PuntoSicuro non possono essere utilizzati al fine di addestrare sistemi di intelligenza artificiale.

---

[www.puntosicuro.it](http://www.puntosicuro.it)