CORRIERE DELLA SERA

22-NOV-2026
da pag. 39 / foglio 1 / 2

Quotidiano - Dir. Resp.: Luciano Fontana
Tiratura: 159487 Diffusione: 194432 Lettori: 1595000 (Data Stampa 0006901)

DATA STAMPA
45° Anniversario

L'INTERVISTA PARLA L'ESPERTO DELL'ONU

Data Stampa 6901-Data Stampa 6901
Data Stampa 6901-Data Stampa 6901

Bengio: «L'umanità non è ancora pronta per l'AI, ma c'è tempo per cambiare strada»

Il guru dell'intelligenza artificiale: ecco i rischi da evitare

La corsa a vincere
Per alcuni governi è una gara da vincere a tutti i costi: ma rischiamo di perdere tutti

di **Riccardo Luna**

Yoshua Bengio, 62 anni, premio Turing nel 2019 per il contributo dato allo sviluppo delle reti neurali, presidente del comitato di esperti sull'intelligenza artificiale dell'Onu: anche lei ha comprato le azioni di Anthropic che hanno appena debuttato a Wall Street ad una quotazione stratosferica?

«Io? No, no. Non voglio proprio farlo in questo momento».

Si aspettava che avrebbero avuto tanto successo in Borsa?

«Gli investitori hanno capito che man mano che si progredisce verso macchine sempre più intelligenti, queste sostituiranno l'intera economia umana. A tappe. Si può facilmente immaginare di che profitto parliamo».

Perché il mercato premia una tecnologia che gli stessi proponenti giudicano pericolosa per l'umanità?

«Se sei un investitore le tue decisioni non sono realmente guidate da considerazioni etiche. Anzi, sei portato a pensare che comunque il tuo piccolo investimento non cambierà il mondo, il quale invece andrà in una direzione o in un'altra indipendentemente dal fatto che tu investa o meno. Ma così le forze del mercato vanno in una direzione che, a livello globale, finisce per essere negativa per tutti. Non è una novità. L'abbiamo visto con il clima e con l'inquinamento».

A settembre alle Nazioni Unite ha affermato che la corsa per una intelligenza artificiale sempre più potente «non è mica una legge di natura». Che intende?

«Ci sono governi che partecipano allo sviluppo dell'intelligenza artificiale come se fosse una gara da vincere a tutti i costi. Ma perdono di vista il fatto che è una gara in cui rischiamo di perdere tutti. Anche se qualcuno dovesse vincere, potremmo perdere tutti per una super intelligenza cattiva o perché qualcuno ha deciso di usare questi strumenti in maniera distruttiva per l'umanità intera. Questa è una partita che nessuno dovrebbe voler giocare; piuttosto dobbiamo cambiare gioco. Da un punto di vista tecnico vuol dire cambiare il modo in cui questi modelli vengono addestrati perché in futuro non ci sia la possibilità che qualcuno voglia tornare a giocare alla fine del mondo».

Venerdì il vertice delle Nazioni Unite di Ginevra sulle armi autonome è finito senza un accordo per l'opposizione di Stati Uniti, Russia e Israele, sebbene gli altri 95 Paesi fossero favorevoli. Cosa ci dice del mondo in cui viviamo?

«Che non abbiamo un adeguato regime di coordinamento internazionale, un sistema che funzioni per la maggioranza. In realtà, non funziona nemmeno per i russi o per gli americani. A livello planetario, siamo anche impegnati in una concorrenza malsana che dà origine a guerre; ma anche a pratiche di dominio economico che non sono eticamente accettabili. Quindi c'è qualcosa che non va nelle nostre istituzioni globali, e credo che i rischi che l'IA comporta a livello globale, di vario tipo, o ci distruggeranno oppure ci costringeranno a ricostruire seriamente le istituzioni globali in modo che funzionino per tutti. E se non troviamo una soluzione potrebbe andare davvero male. Ma sono ottimista. Penso a quanto velocemente le persone e i governi hanno cambiato rotta quando hanno compreso i rischi del Covid. Quando comprendiamo i rischi e la posta in gioco sappiamo cambiare le cose molto rapidamente».

Come presidente del comitato di esperti Onu, a settembre ha aperto la seduta del Consiglio di Sicurezza parlando prima di Sam Altman e Dario Amodei. Che potere hanno davvero le Nazioni Unite?

«L'Onu ha il potere di convocazione. Ha credibilità. E anche se la riunione sulle armi autonome letali non ha prodotto un accordo, è importante vedere quanti Paesi



ARTICOLO NON CEDIBILE AD ALTRI AD USO ESCLUSIVO DEL CLIENTE CHE LO RICEVE - DS6901 - S.30528 - L.1878 - T.1748

CORRIERE DELLA SERA

22-NOV-2026
da pag. 39 / foglio 2 / 2

Quotidiano - Dir. Resp.: Luciano Fontana
Tiratura: 159487 Diffusione: 194432 Lettori: 1595000 (Data Stampa 0006901)

DATA STAMPA
45° Anniversario

ritengono che sia necessario trovarlo e chi invece vi si oppone. C'è una cosa che conta sempre di più nel futuro che stiamo costruendo: l'opinione pubblica. E le Nazioni Unite possono avere un ruolo importante nel plasmarla e nell'unire Paesi che, singolarmente, potrebbero non avere la possibilità di far sentire la propria voce ma insieme possono farlo. Questo non significa che sia facile. Non significa che sia abbastanza. Ma dovremmo fare tutto il possibile per creare un mondo migliore».

Veniamo ai rischi. Recentemente ha detto che presto potremmo non accorgerci che i chatbot ci stanno ingannando. Quanto presto?

«Stiamo già notando che stanno diventando sempre più bravi a ingannare i sistemi di monitoraggio che abbiamo messo a sorvegliarli. Se continua così, saremo nei guai e loro saranno in grado di imbrogliare in un modo che non saremo nemmeno in grado di rilevare. Potremmo non sapere mai che stanno preparando cose davvero negative».

Lo scorso ottobre è stato ribattezzato «l'ottobre nero della sicurezza informatica». Novembre non sta andando meglio. Quanto è davvero pericoloso tutto ciò? Dobbiamo considerare questi allarmi la nuova normalità?

«Sì. E probabilmente andrà peggio. L'infrastruttura software nel mondo è semplicemente inadeguata: secondo gli attuali standard di sicurezza informatica, è in una situazione pessima. Il problema dell'AI è che consente di lanciare attacchi informatici anche a chi non ha competenze sofisticate in materia. E se pensiamo, per esempio, a un ospedale, non avrà mai il tempo e né le competenze per rimediare. Ci saranno delle vittime».

Quindi è d'accordo con quanto il primo ministro britannico ha detto ai suoi cittadini ad agosto, ovvero di prepararsi come se fossimo alla vigilia di una guerra. Siamo a questo punto?

«Dobbiamo essere pronti per l'eventualità che si verifichino eventi in cui Internet non funzioni; o che il sistema bancario vada fuori uso per un tempo prolungato. Dobbiamo essere in grado di andare avanti anche quando i computer non funzionano ed evitare conseguenze catastrofiche, tipo persone che non riescono a procurarsi cibo o acqua».

Cosa può fare le scienza?

«Sono un ricercatore nel campo dell'apprendimento automatico, quindi cerco di usare le mie competenze per capire cosa c'è che non va attualmente nel modo in cui stiamo addestrando l'AI e per progettare metodi diversi di addestramento che garantiscano che l'AI sia almeno onesta e non cerchi di danneggiarci. È su questo che sto lavorando da oltre un anno ormai. Il nostro obiettivo è dimostrare al mondo come si può cambiare il modo in cui si addestrano i sistemi, in modo che almeno non cerchino di prendere il controllo dell'umanità. Visti i primi risultati, sono molto ottimista».

© RIPRODUZIONE RISERVATA

I punti

- Il comitato di esperti sull'intelligenza artificiale dell'Onu invoca la governance globale e inclusiva per bilanciare opportunità di sviluppo e rischi sistemici, evitando che la tecnologia resti monopolio di pochi colossi o Stati
- Rifiutando una super-agenzia centralizzata, propone sette pilastri: un panel scientifico stile Ipcc (per il clima), un dialogo intergovernativo, standard tecnici e sui dati condivisi, più un fondo e una rete per sostenere i Paesi in via di sviluppo. Il fine è tutelare i diritti umani superando l'autoregolamentazione privata



Canadese Yoshua Bengio è un informatico e ricercatore canadese, riconosciuto come uno dei pionieri delle reti neurali e del deep learning. È professore ordinario a all'Università di Montréal, fondatore e scientific advisor di Mila dopo esserne stato direttore scientifico (foto Ap)

ARTICOLO NON CEDIBILE AD ALTRI AD USO ESCLUSIVO DEL CLIENTE CHE LO RICEVE - DS6901 - S.30528 - L.1878 - T.1748