

Geoffrey E. Hinton — Nobel per la Fisica 2024

“The Godfather of AI” — estratto dal discorso al banchetto della cerimonia

Contesto

Questo testo riassume (in traduzione) alcuni passaggi di un intervento di Geoffrey E. Hinton collegato al riconoscimento assegnato nel 2024 per i progressi legati a una nuova forma di intelligenza artificiale basata su reti neurali artificiali. Il tono è quello di un discorso pubblico: entusiasmo per i benefici, ma anche avvertimenti molto netti sui rischi a breve e lungo termine.

Spunti di riflessione

- *Quali benefici concreti promette l'intelligenza artificiale e a chi andranno realmente i vantaggi dell'aumento di produttività?*
- *In che modo strumenti progettati per “aiutare” possono già oggi amplificare polarizzazione, sorveglianza e manipolazione?*
- *Cosa significa, sul piano etico e politico, creare sistemi potenzialmente più intelligenti di noi senza sapere se potremo mantenerne il controllo?*

Estratto del discorso

Quest'anno i Comitati Nobel per la Fisica e per la Chimica hanno riconosciuto i progressi straordinari compiuti in una nuova forma di intelligenza artificiale che utilizza reti neurali artificiali per imparare a risolvere problemi computazionali complessi.

Questa nuova forma di intelligenza artificiale eccelle nel modellare l'intuizione umana più che il ragionamento umano e ci permetterà di creare assistenti altamente intelligenti e competenti, in grado di aumentare la produttività in quasi tutti i settori.

Se i benefici di questa maggiore produttività potranno essere condivisi equitamente, si tratterà di un enorme progresso per tutta l'umanità.

Purtroppo, il rapido sviluppo dell'intelligenza artificiale comporta anche molti rischi a breve termine. Ha già contribuito alla creazione di camere dell'eco divisive, offrendo alle persone contenuti che alimentano indignazione e polarizzazione.

È già utilizzata da governi autoritari per la sorveglianza di massa e da criminali informatici per attacchi di phishing.

In un futuro prossimo, l'intelligenza artificiale potrebbe essere impiegata per creare nuovi virus devastanti e armi letali spaventose, capaci di decidere autonomamente chi uccidere o mutilare.

Tutti questi rischi a breve termine richiedono un'attenzione urgente e decisa da parte dei governi e delle organizzazioni internazionali.

Esiste inoltre una minaccia esistenziale di più lungo periodo, che emergerà quando creeremo entità digitali più intelligenti di noi. Non sappiamo se saremo in grado di mantenerne il controllo, ma oggi abbiamo prove che, se verranno sviluppate da aziende guidate da profitti a breve termine, la nostra sicurezza non sarà la priorità.

Abbiamo urgente bisogno di ricerca su come impedire che queste nuove entità sviluppino il desiderio di prendere il controllo. Non si tratta più di fantascienza.

— Geoffrey Hinton

Idee chiave (per commento orale)

1. Benefici

Assistenti intelligenti e aumento della produttività in molti settori. Il punto “politico” implicito è la distribuzione equa dei benefici.

2. Rischi già attuali

Camere dell'eco e polarizzazione; sorveglianza di massa; phishing e attacchi informatici.

3. Rischio di lungo periodo

Perdita di controllo su entità digitali più intelligenti di noi. Servono ricerca e governance: la scienza non è “neutra”, è responsabilità.

Fonte (YouTube): **Banquet speech — Geoffrey Hinton (estratto)** — <https://www.youtube.com/watch?v=f5WQAk3dYo>