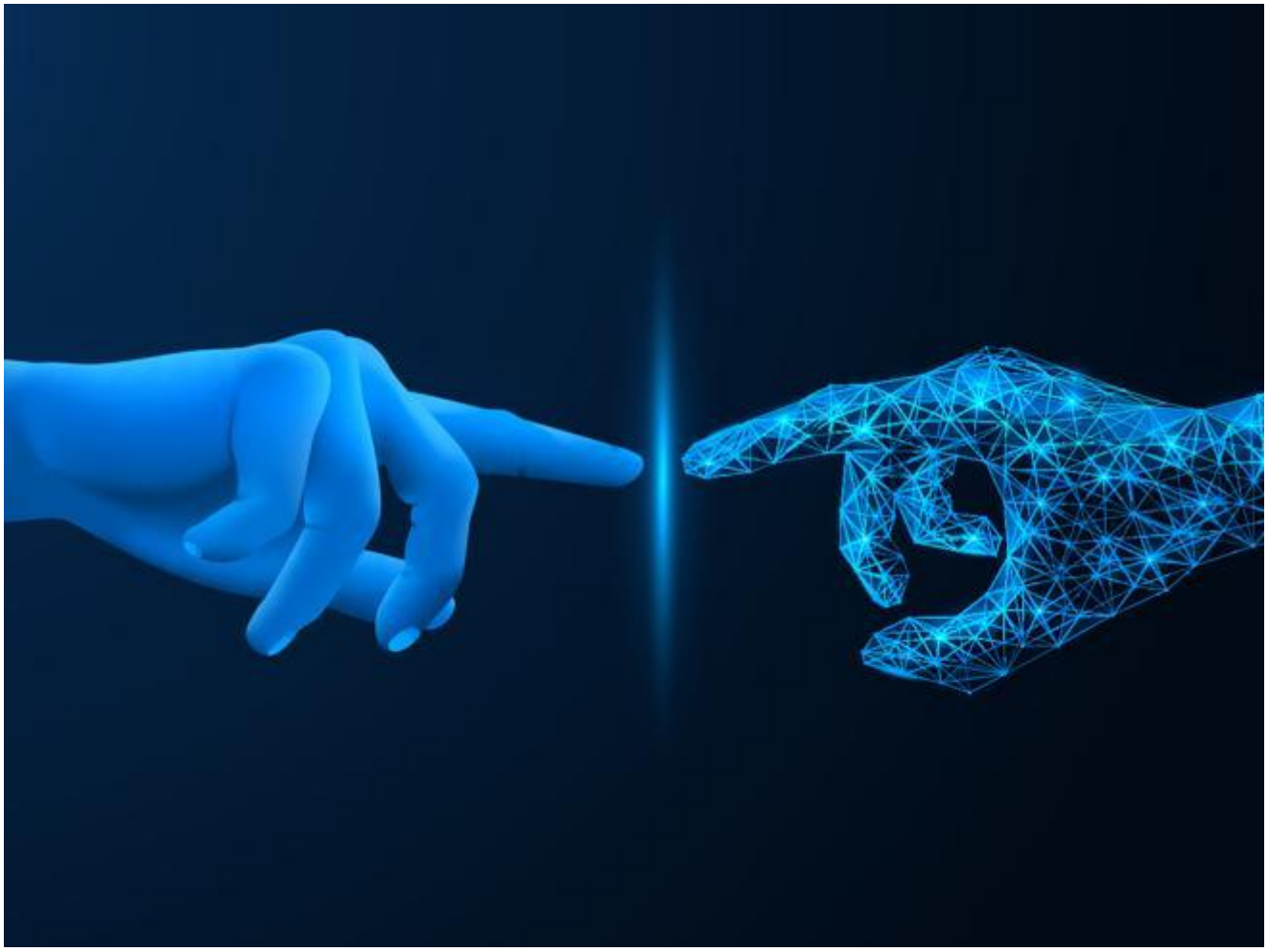
Geoffrey Hinton: «I militari hanno preso l’IA. Entro 20 anni esseri superintelligenti ci sostituiranno: rischiamo l’estinzione»

Il premio Nobel e l’intelligenza artificiale: «I leader delle aziende tecnologiche non hanno una posizione morale: che tristezza. Ora ascoltiamo il Papa». Venerdì 12 assieme ad altri scienziati discuterà con Leone XIV (Fonte: https://www.corriere.it/ 10 settembre 2025)



«La maggior parte dei principali ricercatori di intelligenza artificiale ritiene che molto probabilmente creeremo esseri molto più intelligenti di noi entro i prossimi 20 anni. La mia più grande preoccupazione è che questi esseri digitali superintelligenti semplicemente ci sostituiranno. Non avranno bisogno di noi. E poiché sono digitali, saranno anche immortali: voglio dire che **sarà possibile far risorgere una certa intelligenza artificiale con tutte le sue credenze e ricordi**. Al momento siamo ancora in grado di controllare quello che accade ma se ci sarà mai una competizione evolutiva tra intelligenze artificiali, penso che la specie umana sarà solo un ricordo del passato». Geoffrey Hinton è nella sua casa di Toronto. Alle sue spalle c’è una grande libreria a parete e lui — come al solito — è in piedi. Per via di un fastidioso problema alla schiena lo chiamano «l’uomo che non si siede mai» (ma in realtà, per brevi periodi e con alcuni accorgimenti, può farlo). Indossa una camicia blu che assieme ai capelli incanutiti e al sorriso dolce lo fanno assomigliare a quei personaggi che nei libri o nei film ci vengono a salvare da una minaccia terribile. Dicono che sia uno dei tre padri dell’intelligenza artificiale, ma visto che gli altri due (Yoshua Bengio e Yann LeCun) sono stati, in modi diversi, dei suoi allievi, se ci fosse un Olimpo

dell’intelligenza artificiale Geoff Hinton sarebbe Giove. E infatti lo scorso anno ha vinto il premio Nobel per la Fisica (con John Hopfield) «per scoperte e invenzioni fondamentali che consentono l’apprendimento automatico con reti neurali artificiali». Per semplificare: senza di lui, non avremmo Chat GPT.

Geoffrey Hinton, premio Nobel per la Fisica 2024

Quando l’ho invitato a far parte del gruppo di esperti di intelligenza artificiale che il 12 settembre incontrerà papa Leone XIV mi ha risposto: «Sono un ateo, sicuro che mi vogliano?». Poi però ha accettato perché evidentemente spera che il pontefice voglia giocare la partita della scienza: «Avrebbe una grande influenza». Molto più degli appelli e delle petizioni della comunità scientifica che in due anni si sono succeduti con lo scopo di rallentare lo sviluppo tecnologico in attesa di capirne meglio i rischi e della ragionevole certezza che la cosa non ci scappi di mano. Il primo a dare l’allarme è stato proprio Hinton alla fine di aprile del 2023 **quando improvvisamente lasciò Google, l’azienda per cui aveva lavorato gli ultimi dieci anni**, per poter parlare liberamente dei rischi dell’intelligenza artificiale. L’ultimo è stato sempre lui, un paio di mesi fa, quando ha firmato un appello, assieme al collega premio Nobel Giorgio Parisi, per chiedere a Open AI, l’azienda che ha sviluppato Chat GPT, maggior trasparenza («Stanno decidendo il nostro futuro a porte chiuse»). Partiamo da qui.

Avete avuto una certa risonanza e raccolto tremila firme autorevoli ma non penso che il vero scopo fosse solo quello.

«La questione è complicata dal fatto che c’è una causa legale tra Musk e Altman. Il motivo per cui Musk lo fa mentre allo stesso tempo sviluppa Grok è la competizione. Musk vuole avere meno concorrenza da OpenAI. Questo è ciò che credo, non ne ho la certezza, ma secondo me la sua motivazione nel voler mantenere OpenAI come “public benefit corporation” non è la stessa che abbiamo noi. La nostra motivazione è che OpenAI era stata creata per sviluppare l’IA in modo etico e invece sta cercando di sottrarsi a quell’impegno. Molti ricercatori, come Ilya Sutskever, se ne sono andati; credo che l’abbiano fatto perché non si dava abbastanza importanza alla sicurezza. L’obiettivo dell’appello era insomma fare pressione sul procuratore generale della California che può decidere se permettere o meno a OpenAI di diventare una società a scopo di lucro».

Quindi non vi aspettavate davvero una risposta da Sam Altman.

«No, certo che no».

Direi che la vera risposta all'appello è arrivata pochi giorni fa, alla Casa Bianca, quando il presidente degli Stati Uniti ha convocato tutti i ceo delle Big Tech, compresi OpenAI, Meta e Google…
«Il presidente ha cercato di impedire che per dieci anni negli Stati Uniti ci fossero regole per l'IA, ma ha fallito; i singoli Stati possono continuare a legiferare».

Che effetto le ha fatto vedere i leader delle grandi aziende tecnologiche così ossequiosi verso il potere politico? Forse è la prima volta nella storia che assistiamo a una concentrazione assoluta del potere.
«La mia reazione è stata di tristezza: vedere i leader di queste aziende non disposti a prendere nessuna posizione morale su nulla».
Quando arrivò la rivoluzione digitale, i sogni erano altri: un mondo più giusto, più equo, migliore per tutti.
«Non sono sicuro che fosse il sogno di tutti. Credo che per la maggior parte delle persone coinvolte, il sogno fosse diventare ricchi. Ma potevano convincersi che sì, si sarebbero arricchiti, ma allo stesso tempo avrebbero aiutato tutti. E in larga misura è stato così per il web: se mettiamo da parte le conseguenze sociali, guardando solo alla vita quotidiana, il web ha davvero aiutato le persone».
Prendiamo Tim Berners Lee: non è diventato ricco e ha dato i protocolli del web gratuitamente a tutti. Poi qualcosa è andato storto. Direi: il capitalismo.
«Io non sono totalmente contro il capitalismo: è stato molto efficace nel produrre nuove tecnologie, spesso con l'aiuto gratuito dei governi per la ricerca di base. Credo che, se ben regolato, il capitalismo vada bene. È accettabile che la gente cerchi di arricchirsi purché, arricchendosi, si aiutino anche gli altri. Il problema è quando la ricchezza e il potere vengono usati per eliminare le regole, così da fare profitti che danneggiano le persone».

È quello che sta accadendo adesso?
«Sì. La cosa che sembra importare loro di più è non pagare le tasse. E sostengono Trump perché abbassa le tasse: questa è la loro spinta primaria. Ma vogliono anche eliminare le regole che proteggono le persone per rendere più facile fare affari. Per questo assecondano Trump».

Come valuta quello che sta facendo l'Unione europea? Si dice che pensi troppo alla regolamentazione e troppo poco all'innovazione. Così restiamo indietro nella corsa all'intelligenza artificiale. Come sa, l'Unione europea ha approvato una legge importante, l'AI Act. Secondo lei è un buon punto di partenza o solo un ostacolo all'innovazione?
«Credo sia un punto di partenza accettabile ma ci sono parti che non mi piacciono: ad esempio, c'è una clausola che dice che nessuna di queste regole si applica agli usi militari dell'IA. Questo perché

diversi Paesi europei sono grandi produttori di armi e vogliono sviluppare sistemi letali e autonomi. Inoltre, per come l'ho capita io, la normativa europea è partita con un'enfasi particolare su discriminazione e bias, preoccupandosi più di questi aspetti che di altre minacce. Perciò mette molto l'accento sulla privacy. La privacy è importante, ma ci sono minacce più gravi».

Ha toccato un punto molto delicato. Le aziende della Silicon Valley erano nate con l'idea di non collaborare con il potere militare.

Quando DeepMind fu venduta a Google, una delle condizioni era che l'IA non sarebbe stata usata per scopi militari. Ma lo scorso gennaio quella condizione è stata rimossa. Ora tutte le aziende di intelligenza artificiale forniscono tecnologia al dipartimento della Difesa americano.

E pochi giorni fa Google ha firmato un accordo importante con Israele per fornire sistemi di intelligenza artificiale alle forze armate. Che ne pensa?

«C'è una regola semplice che usano gli scienziati politici: non guardare a ciò che la gente dice, ma a ciò che fa. Sono rimasto molto dispiaciuto quando Google ha eliminato l'impegno a non usare l'IA per scopi militari. È stata una delusione. E lo stesso quando hanno cancellato le politiche di inclusione solo perché era arrivato Trump».

E arriviamo a papa Leone XIV: appena insediato, ha detto di volersi occupare di intelligenza artificiale. Che ruolo pensa possa avere in questo dibattito?

«Beh, come saprà, ha circa un miliardo di fedeli. Se lui dicesse che è importante regolare l'IA, sarebbe un contrappeso alla narrativa dei leader delle Big Tech che ringraziano il presidente per non aver imposto regole. Io credo che il Papa abbia un'influenza politica reale, anche al di fuori del cattolicesimo. Molti leader religiosi, come il Dalai Lama, lo vedono come una voce morale. Non tutti i Papi sono stati così, ma questo sì, e anche il suo predecessore. Per molti non cattolici, le sue opinioni morali sono sensate, non infallibili, ma ascoltate. Se dirà che regolare l'IA è essenziale, avrà un impatto».

Cosa vorrebbe far sapere al Papa sui rischi dell'intelligenza artificiale?

«Credo che per il Pontefice sia inutile porre l'accento sui cosiddetti rischi esistenziali, come la sparizione della specie umana nel lungo periodo; e sia meglio concentrarsi sui rischi attuali. Non serve credere che l'IA sia un "essere" per capire che può portare disoccupazione di massa, corruzione delle democrazie, cyber-attacchi più facili ed efficaci, creazione di virus letali accessibili a chiunque. Sta già accadendo. Quindi su questi rischi a breve termine, quelli che chiamo "rischi dovuti a cattivi attori", possiamo trovare un accordo. Questi sono rischi che qualunque leader religioso o politico può capire, senza dover entrare nella questione metafisica se l'IA sia o meno un "essere"».

A volte sembra che i governi non capiscano fino in fondo la posta in gioco, che si fidino troppo di queste aziende.
«È proprio così. Spesso i governi non hanno abbastanza esperti interni e si affidano a consulenti che vengono… indovini da dove? Proprio dalle Big Tech. Così finiscono per sentire solo la voce delle aziende. È come se i regolatori di Wall Street fossero tutti ex banchieri che ancora pensano come banchieri».

È considerato un apocalittico. Lei come si sente? Ottimista o pessimista?
«Realista. Vedo entrambi i lati. Ma la storia ci dice che, senza regole, si abusa del potere. Perciò credo che, se non regoliamo, finirà male».

Crede che ci sia una lezione particolare che dovremmo ricordare ora?
«Sì. Pensiamo all'energia nucleare. Quando gli scienziati capirono che le loro scoperte potevano portare all'atomica, molti di loro provarono ad avvertire i governi. Ma alla fine furono i militari a decidere come usarla. Non fu la comunità scientifica a stabilire i limiti. Lo stesso rischio lo vedo oggi: se lasciamo che siano solo i governi o le aziende a decidere, l'IA sarà usata in modi molto pericolosi».

È davvero un momento unico nella storia dell'umanità o solo un ennesimo progresso?
«È un punto di svolta. Per la prima volta stiamo creando entità che possono essere più intelligenti di noi. Non è mai successo prima. Abbiamo creato macchine più forti, più veloci, ma mai più intelligenti. Questo cambia tutto».

Quindi cosa possiamo fare?
«Serve un movimento globale, non solo poche voci isolate. Dobbiamo creare un consenso ampio nella comunità scientifica, altrimenti i politici ci ignoreranno».
Dopo 90 minuti l'intervista è finita (questa è una versione molto condensata ma approvata dal professore). Resta il tempo per una domanda personale. Geoffrey Hinton viene da una famiglia eccezionale. L'Everest si chiama così per via di un suo prozio cartografo (e anche il suo secondo nome è Everest); e un altro è il padre della logica booleana (la base matematica su cui si basano i computer e i sistemi digitali). Come se non bastasse, **il padre gli ha sempre tenuto l'asticella altissima**. Gli diceva: «Se lavori il doppio di me, quando avrai il doppio della mia età, forse sarai bravo la metà».
Gli chiedo: ce l'ha fatta a realizzare l'obiettivo che le ha dato suo padre? Sorride: «No. Sicuramente almeno la metà».