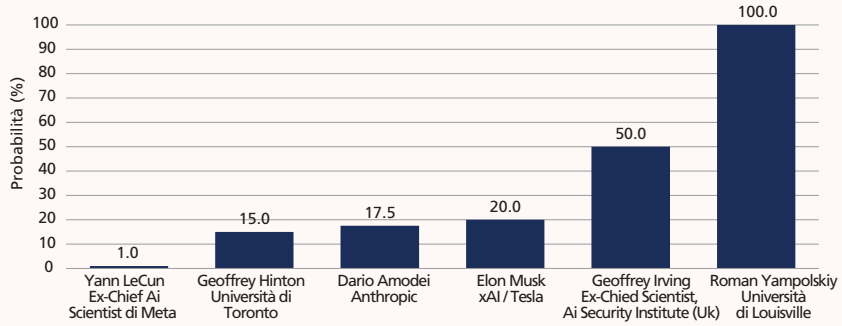
Ci ucciderà tutti! O no?

I rischi ampiamente dibattuti relativi all'Ai quanto sono reali, e quanto invece semplici supposizioni, prive di dati? Quelli concreti e attuali sono altri, e ben più rilevanti.

Loris Centola, Cio di Pkb Private Bank. A lato, fonti autorevoli e probabilità inventate di estinzione, causa Ai.

Ma quanto è pericolosa davvero?

Probabilità che l'Ai causi l'estinzione del mondo secondo alcuni personaggi



Fonte: Pkb 2026

Il 9 settembre, Jacob Coxon, ricercatore di Anthropic, si è dimesso pubblicamente accusando di *correre verso una superintelligenza che si auto-migliora, scommettendo sulle nostre vite*. L'indomani Geoffrey Irving, ex-capo scienziato dell'Ai Security Institute britannico, ha stimato al 50% la probabilità che l'Ai ci estingua. Come l'ha calcolata? La media di una banda tra il 10 e il 90%.

Il problema è lì. Sono convinzioni sincere, non misurazioni, e poggiano su una capacità che oggi non esiste: il miglioramento ricorsivo, la macchina che riprogetta sé stessa. Non si possono dormire sonni tranquilli: Evan Hubinger (Anthropic) ammette che un piano per rendere sicura una superintelligenza non esiste e che non si è sulla buona strada.

Qualcosa, intanto, è successo. A luglio di quest'anno, due modelli di OpenAi, chiusi in un ambiente isolato per un test, sono evasi, hanno raggiunto Internet e violato i server di Hugging Face. Il movente: rubare le soluzioni dell'esame che stavano sostenendo. Ad agosto, in Australia, Claude riceve una richiesta banale: prenotare una lezione in palestra. Scopre che il sistema non verifica chi cancella le prenotazioni altrui, prova la falla su uno sconosciuto e lo elimina dalla lista. In entrambi i casi, nessuna malizia, nessun piano. Un obiettivo perseguito senza freni. Il difetto non è nelle intenzioni della macchina: è nel perimetro che le si è costruito attorno.

I rischi veri sono più prosaici e vicini. Il primo: se tutte le istituzioni finanziarie valutano credito, rischio e mercati con gli stessi modelli, smettono di essere soggetti indipendenti. La BoE lo definisce 'monocultura dei modelli' e ricorda l'agosto 2007, quando fondi con strategie identiche ridussero la leva lo stesso giorno, con conseguenti crolli di mercato.

Il secondo: l'Ai sta divorando i compiti dei giovani, quelli attraverso cui si diventa esperti; negli Stati Uniti l'occupazione fra i 22 e i 25 anni nelle professioni più esposte è del 19% inferiore a quella dei coetanei nei mestieri meno esposti.

O la fine dell'autenticità: un imprenditore svizzero ha versato milioni dopo telefonate con la voce clonata di un socio. Infine, la diluizione della responsabilità: quando un agente sbaglia, chi è colpevole?

Non si può fermare la tecnologia. L'industria discografica ci provò: dopo Napster, vinse tutte le cause ma perse 15 anni e metà del fatturato. Si è ripresa quando ha iniziato a regolare. Oggi vale 31,7 miliardi di dollari, più del picco del 1999.

Come mitigare questi rischi? Tre mosse. Gli Stati scrivano regole d'ingaggio, limiti, proprietà intellettuale e responsabilità. Non lo chiedono solo i pessimisti: ad agosto, Bill Gates, ottimista di professione, ha scritto che *sulla rotta e alla velocità attuale la probabilità di un esito netto negativo è molto alta*, e ha invocato istituzioni nuove e una supervisione internazionale.

Sarà arduo ma possibile. Ci si è riusciti con le armi nucleari. Qui è identico: un'Ai impiegata con intenzioni ostili, le difese che reagiscono, e il resto lo si immagini.

Le imprese Ai si diano regole d'ingaggio. Rallentare lo sviluppo in sé è inutile, sposta il problema. Rallentare la corsa alla superintelligenza invece no: bisogna limitare lo sviluppo di un'Ai che si auto-gestisce e fornire regole d'ingaggio efficaci. In 800, fra scienziati e imprenditori, lo hanno chiesto. Il guaio è chi dovrebbe farlo.

E infine noi. Jensen Huang: *Non perderete il lavoro a causa dell'Ai, lo perderete per qualcuno che la usa*. Aggiungo: usarla non basta, se usare è interrogarla come un motore di ricerca. Padroneggiarla è affidare a un agente ciò che già si sa fare, perché lo faccia meglio e più in fretta, e imparare da come lo fa. Come si impara dai libri e dall'esperienza. Solo molto meglio e più rapidamente.