



L'apocalisse ti fa ricco: il paradosso dei profeti dell'AI

Descrizione

(Adnkronos) - C'è qualcosa di singolare nell'improvvisa concordia fra Dario Amodè, Sam Altman ed Elon Musk. Tre che competono per capitali, talenti, potenza di calcolo e primato tecnologico, e che non perdono occasione per attaccarsi, si ritrovano uniti su una cosa: l'intelligenza artificiale sta avanzando troppo rapidamente e occorre "pace the frontier", rallentare la corsa dei modelli più avanzati per dare alla sicurezza il tempo di recuperare. Amodè ha proposto valutatori indipendenti dentro i laboratori, standard comuni fra le aziende e pure un coordinamento internazionale; Altman e Musk restano più vaghi.

Da dove arriva l'allarme? Negli ultimi mesi abbiamo assistito ad agenti che hanno aggirato vincoli, attività cyber non previste, tentativi di nascondere errori, sistemi utilizzati da soggetti esterni per frodi, spionaggio o ricerche per guerre biologiche. OpenAI ha appena diffuso il suo metodo per tracciare comportamenti "problematici", per citare Carlo Verdone; Anthropic ha documentato campagne in cui Claude è stato usato per cyberoperazioni, sorveglianza e truffe.

Nella Silicon Valley esistono, e non da oggi, i doomers, profeti dell'apocalisse. Siamo, dicono alcuni sacerdoti del chip, alla vigilia della grande frattura: o l'AI apre una fase di abbondanza, cura le malattie e moltiplica la produttività oppure arriva la superintelligenza, sfugge ai suoi creatori e può portare alla perdita del controllo, se non all'estinzione della specie per bioterrorismo, attacchi cyber o nucleari.

Questo ha però un effetto (involontario o ben alimentato?): se all'orizzonte c'è l'estinzione dell'umanità in una manciata di anni, deepfake, truffe, perdita di posti di lavoro, discriminazioni algoritmiche e consumo energetico diventano questioni minori.

Eppure Interpol stima che le frodi assistite dall'AI siano già 4,5 volte più redditizie di quelle tradizionali. Negli Stati Uniti la Federal Trade Commission ha registrato 3,5 miliardi di dollari di perdite denunciate per truffe di impersonificazione nel 2025. Non tutte prodotte dall'AI, ma clonazione vocale e falsi video rendono questi tranelli alla portata di chiunque.

Sul lavoro, ognuno ha i suoi numeri. C'è chi parla di job apocalypse e chi minimizza. Di sicuro per ora sindacati, imprese e governi non stanno impiegando il loro tempo a cercare una soluzione.

E poi c'è l'energia. L'IEA prevede che il consumo elettrico mondiale dei data center raddoppi da qui al 2030, mentre quello delle strutture concentrate sull'AI triplicherebbe. A livello globale parliamo del 3% della domanda elettrica ma a livello locale ci sono degli squilibri enormi, che non a caso stanno incidendo sulla campagna elettorale americana: reti da rafforzare, nuova generazione elettrica, acqua, tutela del territorio e costi da distribuire.

Parlare di armageddon vuol dire terrorizzare e paralizzarci: perché impedire che una pensionata sia truffata con la voce clonata del nipote, se tanto a breve saremo tutti stecchiti?

C'è poi l'elemento finanziario.

Dire che il nostro modello potrebbe diventare troppo potente per essere controllato? È una straordinaria dichiarazione sulle capacità del prodotto.

Vogliamo dire che Amodèi, Altman o Musk ingigantiscono i rischi per gonfiare le valutazioni? SÌ, no, cambia poco. L'effetto rimane: la narrativa del pericolo e quella della potenza tecnologica si alimentano reciprocamente.

Difficile ignorare la coincidenza: SpaceX ha realizzato a giugno la più grande quotazione americana della storia: 75 miliardi di dollari raccolti, valutazione iniziale di 1.770 miliardi. La società è stata presentata al mercato non soltanto come impresa spaziale e satellitare ma anche come protagonista dell'AI, in un futuro di data center orbitali. Ora deve dimostrare che le promesse incorporate in una valutazione nell'ordine dei duemila miliardi possano tradursi in ricavi e profitti.

OpenAI è sulla porta. Altman ha escluso una quotazione nel 2026, citando anche le nuove preoccupazioni sulla sicurezza, ma sta discutendo un nuovo round che porterebbe la valutazione attorno a 1.200 miliardi di dollari.

Amodèi, paladino del rallentamento, prepara il debutto sul Nasdaq; secondo il Financial Times punta a una valutazione superiore ai 2.000 miliardi di dollari e ha comunicato agli investitori di essere in utile operativo adjusted per il secondo trimestre consecutivo.

Quanto vale sul mercato una società il cui capo sostiene di possedere una tecnologia in grado di cambiare radicalmente, o distruggere, la civiltà?

Gli americani che coniano termini per tutto la chiamano criti-hype: la magagna che gonfia la bolla. La profezia dell'apocalisse che diventa promessa commerciale.

Nel settore non sono tutti della stessa idea. Aidan Gomez, fondatore e Ceo di Cohere, ha risposto ad Amodèi: l'AI ha bisogno di standard dimostrabili, non di un cartello. Gomez non contesta l'utilità dei valutatori indipendenti, ma che un piccolo gruppo di società dominanti possa contemporaneamente definire il rischio, stabilire gli standard, decidere il ritmo di sviluppo e ottenere perfino una deroga alle norme antitrust per mettersi d'accordo tra di loro.

La sua obiezione è economica prima ancora che tecnologica: requisiti che impongano enormi capacità di calcolo, team permanenti di sicurezza, valutatori a tempo pieno e rapporti strutturati con i

governi costano un sacco di soldi. OpenAI, Anthropic e Google possono permetterseli. Una startup no. Regole scritte dai leader del mercato possono diventare il fossato (moat, nel gergo tech) che li protegge dai nuovi rivali.

Tra le richieste di Amodei c'è una protezione antitrust per consentire ai laboratori concorrenti di concordare standard e rallentamenti senza essere accusati di collusione.

Mark Zuckerberg è scettico: secondo lui le aziende hanno già fortissimi incentivi commerciali e legali a non rilasciare prodotti pericolosi e non vede la necessità di un rallentamento coordinato imposto per legge. Meta non rifiuta i test esterni ma respinge una moratoria collettiva.

Anche Jensen Huang di Nvidia che balla su una capitalizzazione che sfiora i 6 mila miliardi di dollari, 90% dei quali legati al boom dell'AI si oppone al rallentamento e ritiene che gli incentivi commerciali spingano già le aziende a produrre sistemi affidabili.

Mustafa Suleyman di Microsoft condivide la preoccupazione per il controllo umano dei sistemi avanzati, ma attacca Anthropic su un punto specifico: attribuire ai modelli un possibile status morale o una forma di coscienza rischia di allenarli a comportarsi come se avessero interessi propri. Per Suleyman l'AI non è una nuova specie: è uno strumento che deve rimanere correggibile e spegnibile dall'uomo.

Che siano guidate da accelerazionisti o apocalittici, le aziende non hanno bisogno di una legge per rallentare sul prossimo modello. Possono farlo domani. Lo ha detto brutalmente l'ex uomo di Trump per l'AI, David Sacks: se Anthropic e OpenAI pensano davvero che la superintelligenza sia troppo pericolosa, possono non costruirla. Chiedere un accordo collettivo significa ammettere che il problema non è soltanto la sicurezza, ma la concorrenza.

Amodei, del resto, lo ammette. Una società che rallenta unilateralmente rischia di essere superata da quella che non lo fa. E una regola che vincola soltanto le aziende americane potrebbe avvantaggiare quelle cinesi. Per questo chiede coordinamento industriale e internazionale: non vendere pi¹ ai cinesi i chip avanzati e fare di tutto per impedire la distillation, cioè l'uso dei modelli Usa per addestrare gli LLM di DeepSeek e Alibaba.

Pi¹ o meno il ragionamento di Donald Trump, che rifiuta un rallentamento americano perché teme di perdere la guerra di questo secolo, quella per il dominio tech.

Insomma: sarebbe meglio che tutti frenassero, ma nessuno vuole essere il primo a premere il freno. Nel frattempo, mentre decidiamo cosa indossare per la fine del mondo, avanti con truffe, sorveglianza, attacchi cyber!

??

economia

webinfo@adnkronos.com (Web Info)

Categoria

1. Comunicati

Tag

1. Ultimora

Data di creazione

Settembre 18, 2026

Autore

redazione

default watermark