



Intelligenza artificiale, nel 2026 domanda elettrica alle stelle

Digitalizzazione. Secondo le previsioni della Iea per server farm, criptovalute e attività di Ai a livello mondiale i consumi di elettricità saranno compresi tra i 620 e i 1.050 TWh contro i 460 del 2022



È stato calcolato che una ricerca su ChatGPT consuma nove volte di più rispetto a una su Google

Sara Deganello

«Si prevede che entro il 2026 l'industria dell'intelligenza artificiale crescerà esponenzialmente arrivando a consumare almeno dieci volte la sua domanda nel 2023». La stima è dell'International Energy Agency (Iea), che nel report *Electricity 2024* indica come il fabbisogno elettrico dei data center dedicati all'AI nel mondo potrebbe arrivare a 100 TWh nel 2026. Ancora la Iea riporta il confronto tra la richiesta elettrica media di una ricerca su Google (0,3 Wh) e una su ChatGPT (2,9 Wh), con la seconda che consuma oltre 9 volte di più: con 9 miliardi di ricerche giornaliere, il fabbisogno aggiuntivo sarebbe di quasi 10 TWh annui. Sono comunque numeri approssimativi, complessi da definire. L'Iea tiene conto anche del paper pubblicato dal data scientist Alex de Vries su *Joule* alla fine del 2023 in cui, partendo dall'assunto che il 95% del mercato dei server AI è coperto da Nvidia, ha calcolato che se l'azienda ne ha consegnati nel 2023 100mila, insieme capaci di consumare tra i 5,7 e gli 8,9 TWh, ed entro il 2027 ne spedisce 1,5 milioni all'anno, a piena capacità richiederebbero tra gli 85,4 e i 134 TWh.

Sempre la Iea stima che data center, criptovalute e attività di AI abbiano consumato circa 460 TWh di

elettricità a livello mondiale nel 2022, quasi il 2% del totale. Prevede una forbice tra i 620 e i 1.050 TWh nel 2026: da 160 TWh a 590 TWh di domanda elettrica in più rispetto al 2022. «Vuol dire aggiungere almeno una Svezia, al massimo una Germania», chiosa l'agenzia nel report.

«Le attività di AI si basano su modelli molto sofisticati che hanno un paradigma di sviluppo e utilizzo radicalmente diverso rispetto ai tool digitali tradizionali. Questi modelli di AI sono creati in due fasi: una di "addestramento" dell'algoritmo offline e una di utilizzo predittivo online, entrambe svolte con il lavoro di potenti data center, che sono tra le infrastrutture più energivore esistenti. Il Dipartimento dell'Energia americano ne stima un consumo per metro quadro da 10 a 50 volte superiore rispetto a quello di un qualunque edificio commerciale, concentrando circa il 2% di tutto il consumo elettrico del Paese. Sviluppare strumenti di AI può essere quindi molto dispendioso a livello energetico. Ad esempio, per il Large Language Model (LLM) Falcon-40B sono serviti 2 mesi di addestramento, per un consumo energetico totale che può essere stimato in 100-200 MWh», spiega Roberto Ventura, managing director e partner di BCG. La diffusione delle attività di AI e il loro costo energetico è un valore che le aziende cominciano a valutare.

Il futuro non è, tuttavia, nella moltiplicazione di ChatGPT per general purpose: «Vedremo una specializzazione degli LLM: saranno molto più piccoli, altrimenti non sa-

ranno sostenibili», sostiene Nicola Gatti, direttore dell'Osservatorio Artificial Intelligence del Politecnico di Milano: «Inoltre ci saranno centri locali: le aziende italiane ed europee non possono permettersi di mettere i propri dati sul cloud: dovranno avere sistemi propri, costruendo il proprio LLM. E se calcoliamo che l'infrastruttura hardware di ChatGPT 3.5 non ha un costo inferiore ai 100 milioni di euro, è evidente che bisognerà ridurre». «È un tema che affrontiamo nel progetto nazionale di AI: riuscire ad avere sistemi che siano *energy aware* ed efficaci», conferma Paolo Traverso, direttore ICT della Fondazione Bruno Kessler di Trento.

L'AI stessa è uno strumento di efficientamento, come testimonia Luca Dal Fabbro, presidente del gruppo Iren: «Sono tanti i settori in cui l'AI potrà giocare un ruolo chiave: dalla gestione dei dati ESG alla valutazione della sostenibilità delle imprese e dei modelli di business, fino alla manutenzione preventiva in ambito idrico e ambientale per ottimizzare le performance e ridurre le dispersioni. Il gruppo Iren ha creato una unità dedicata allo



sviluppo dell'AI per migliorare il servizio ai clienti ma nello stesso tempo ha deciso di adottare un'intelligenza artificiale responsabile». Si tratta di un approccio in cui entrano valori come, appunto, responsabilità, equità, inclusione, trasparenza, privacy.

L'efficientamento energetico è già all'opera nelle infrastrutture AI. Il supercomputer Leonardo, gestito da Cineca a Bologna, utilizza una potenza di circa 5 MW al giorno. Non viene utilizzato soltanto per il training di intelligenza artificiale, ma per moltissimi altri scopi, dallo sviluppo di nuovi farmaci alla lotta alle epidemie. Cineca non solo ha collezionato protocolli Leed, certificazioni per prestazioni energeti-

che efficienti e titoli di efficienza energetica riconosciuti dal Gse: «Stiamo valutando la realizzazione di un parco di energia rinnovabile per i consumi di Leonardo nella zona del Tecnopolo: speriamo che veda la luce entro il 2024, poi seguiranno i tempi di realizzazione», racconta la direttrice generale di Cineca, Alessandra Poggiani. «Affinché la crescita dei consumi elettrici a livello mondiale, data anche dall'AI e dalla digitalizzazione, sia in linea con il percorso di decarbonizzazione è necessario che in parallelo cresca la produzione di energia rinnovabile. In Italia l'obiettivo al 2030 è raggiungere l'84% nel mix elettrico, portando la potenza degli impianti rinnovabili

a oltre 140 GW», aggiunge il presidente di Elettricità Futura Agostino Re Rebaudengo.

Davide Serra, ad di Algebris, recentemente a Davos ha indicato anche una fonte complementare: «È chiaro che ci sarà sempre bisogno di più energia nucleare, più reattori di terza e quarta generazione. La Cina ne produrrà 35 nei prossimi anni, Dubai e Qatar ne hanno già 6. Rinascita nucleare di terza e quarta generazione ed energia da fonti rinnovabili permetteranno lo sviluppo di AI in futuro, nel breve periodo è chiaro che ci sarà il rischio di un ulteriore aumento dei prezzi per l'energia, perché questa grande potenza di calcolo ne avrà molto bisogno».

© RIPRODUZIONE RISERVATA

CONSUMI

100

Terawattora

È la stima di Teja del fabbisogno elettrico globale dei data center AI nel 2026. Per Paolo Traverso della Fondazione Bruno Kessler «il problema del consumo energetico dovuto alle applicazioni di AI esiste».

5

Megawattora al giorno

Il supercomputer Leonardo, gestito da Cineca a Bologna, utilizza una potenza di circa 5 MW al giorno. È usato per il training di intelligenza artificiale ma anche per lo sviluppo di nuovi farmaci e la lotta alle epidemie. Si sta valutando la realizzazione di un parco per la produzione di elettricità da fonti rinnovabili nell'area del Tecnopolo

84%

IL MIX ELETTRICO

«Per essere in linea con il percorso di decarbonizzazione è necessario che in parallelo cresca la produzione di energia rinnovabile. In Italia l'obiettivo al

2030 è raggiungere l'84% nel mix elettrico, portando la potenza degli impianti rinnovabili a oltre 140 GW», ricorda il presidente di Elettricità Futura Agostino Re Rebaudengo.