

BRIGHTSIDE CAPITAL

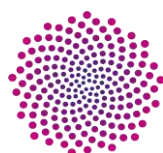
## THE RACE IS ON



*I'm waking up, I feel it in my bones  
Enough to make my system blow  
Welcome to the new age, to the new age*

*Radioactive – Imagine Dragons  
Night Visions, 2012*

Alla chiusura delle borse di venerdì scorso gli occhi degli investitori erano già puntati sulla pubblicazione dei risultati trimestrali di Meta e Microsoft (che insieme rappresentano il 9% della capitalizzazione di mercato dell'indice S&P500) in programma il 29/01 ed in concomitanza con un altro evento macro rilevante, ovvero la conferenza stampa del presidente della Federal Reserve Jerome Powell. Tuttavia, durante il weekend, è emersa una notizia "sorprendente" in grado di distogliere l'attenzione del mercato nei confronti degli eventi menzionati sopra: una startup cinese avrebbe creato un modello più efficiente e 10 volte più economico di ChatGPT, attirando il focus dei media sulla Silicon Valley e rischiando di innescare elevata volatilità su tutto il comparto azionario tecnologico.



## BRIGHTSIDE CAPITAL

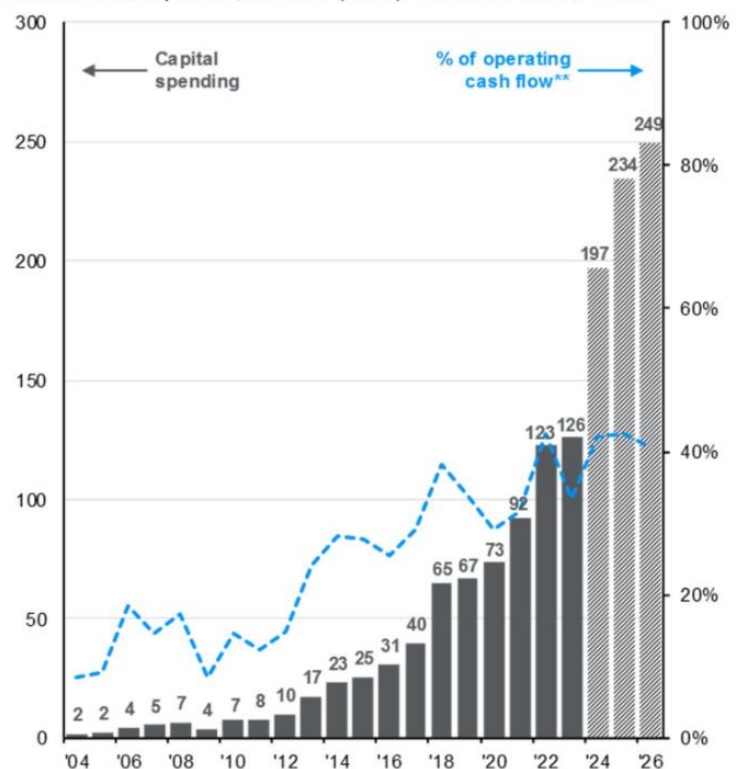
Il nome DeepSeek non era conosciuto ai più, ma gli addetti ai lavori avevano già menzionato l'operato dell'azienda cinese proprio pochi giorni prima durante il World Economic Forum di Davos, in Svizzera. Alexandr Wang, la cui azienda Scale AI fornisce dati di training ai principali attori dell'intelligenza artificiale, tra cui OpenAI, Google e Meta, dichiarava che la corsa all'AI tra Stati Uniti e Cina fosse una vera e propria "guerra dell'intelligenza artificiale". Wang, intervistato da CNBC nella giornata di giovedì 23, affermò che DeepSeek, il principale laboratorio di AI cinese, avesse rilasciato un "modello rivoluzionario" il giorno di Natale.

Pochi giorni prima, e poco dopo la sua inaugurazione alla Casa Bianca, il presidente Donald Trump aveva annunciato una joint venture con OpenAI, Oracle e SoftBank per investire miliardi di dollari nelle infrastrutture di intelligenza artificiale negli Stati Uniti. Il progetto, chiamato *Stargate*, è stato presentato alla Casa Bianca da Trump, dal CEO di SoftBank Masayoshi Son, dal co-fondatore di Oracle Larry Ellison e dal CEO di OpenAI Sam Altman. Tra i principali partner tecnologici iniziali figurano Microsoft, Nvidia e Oracle, oltre alla società di semiconduttori Arm. I promotori del progetto, criticato da Elon Musk per i dubbi nella capacità di Softbank di garantire i finanziamenti promessi<sup>1</sup>, hanno annunciato un investimento iniziale di 100 miliardi di dollari, con un piano di espansione fino a 500 miliardi di dollari nei prossimi quattro anni.

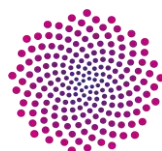
L'avvento di DeepSeek che, come vedremo nel corso dello scritto, non ha fatto altro che costruire il suo modello sopra delle fondamenta (americane e non solo) decisamente solide, ha posto il dubbio al mercato che le massicce spese in CAPEX (grafico a destra) da parte dei cosiddetti *Hyperscalers* (fornitori di servizi cloud su scala globale) fossero, in fin dei conti, poco giustificate. Se qualcuno può prendere un modello costoso creato da altri e costruirci sopra un prodotto così performante, quale sarebbe l'incentivo per chi sviluppa? In realtà, il costo per sviluppare nella sua interezza DeepSeek non è solo quello dichiarato di 6 milioni di

Capex from the major AI hyperscalers\*

USD billions; Alphabet, Amazon (AWS), Meta, Microsoft, Oracle



<sup>1</sup> <https://www.wsj.com/tech/musk-pours-cold-water-on-trump-backed-stargate-ai-project-53428d16>



## BRIGHTSIDE CAPITAL

dollari, ma comprende anche il costo di creazione dei modelli di base da cui è stato distillato, che essendo open source non ha rappresentato un esborso per DeepSeek. Quindi, la realizzazione della “Chat GPT cinese” è stata estremamente costosa nel suo complesso, ma i costi sono stati distribuiti in modo disomogeneo: Meta ha speso centinaia di miliardi, mentre DeepSeek ha investito solo pochi milioni.

Come spesso accade nei mercati, la prima reazione è istintiva e guidata dalla frenesia di vendere ed alleggerire il peso nei portafogli di azioni che, come nel caso di Nvidia, negli ultimi 2 anni hanno visto quintuplicare la capitalizzazione di mercato. Ed ecco che la news di DeepSeek, unita al fatto che (1) i titoli tecnologici scontano valutazioni relativamente alte (rispetto alla media storica) e che (2) il posizionamento degli investitori su questi titoli era estremamente rialzista (con pochissimi operatori di mercato posizionati “corti”), ha portato i titoli tecnologici americani a lasciare sul campo circa 1 trilioni di dollari durante la sessione di trading del successivo lunedì. A guidare il gruppo, Nvidia (-17%) con una riduzione della capitalizzazione di mercato di circa 600 miliardi di dollari, la più importante perdita giornaliera mai registrata nella storia.

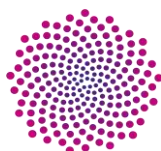
### Biggest Single-Day Market Cap Losses

	Firm	Biggest single day market cap losses	Date
1	Nvidia	-\$590B	1/27/25
2	Nvidia	-279	9/3/24
3	Meta	-251	2/3/22
4	Nvidia	-228	1/7/25
5	Nvidia	-212	4/19/24
6	Nvidia	-208	6/24/24
7	Amazon	-206	4/29/22
8	Nvidia	-205	7/17/24
9	Nvidia	-205	7/24/24
10	Nvidia	-197	8/29/24

Source: Bloomberg

## Le origini di DeepSeek e l'innovazione del suo modello

La startup è stata fondata nel 2023 da Liang Wenfeng, fondatore nel 2015 di High Flyer, un hedge fund cinese che si concentra sull'utilizzo di algoritmi di intelligenza artificiale per comprare e vendere azioni e che, ad oggi, vanta 8 miliardi di dollari in gestione. Nel 2019, High-Flyer ha istituito una divisione dedicata alla ricerca su algoritmi di intelligenza artificiale e alle loro applicazioni di base. Tuttavia Liang non era totalmente appagato e aveva iniziato ad elaborare l'idea di creare il suo modello di AI, senza abbandonare la gestione di High Flyer, e con l'obiettivo di poter competere con le grandi potenze asiatiche ByteDance e Alibaba. Così nel 2021 (prima che l'amministrazione Biden imponesse controlli sulle esportazioni di chip avanzati nell'ottobre del 2022) compra 10 mila chip Nvidia del modello H 800, seleziona un



## BRIGHTSIDE CAPITAL

team tra i migliori dipendenti del suo fondo, esperti nell'ottimizzare la potenza della GPU (o *Graphics Processing Unit*, che è un tipo di processore progettato per gestire calcoli complessi) e nel 2023 apre ufficialmente DeepSeek.

Il 20 Gennaio 2025, mentre il mondo era concentrato sul cambio di leadership americana con l'insediamento alla Casa Bianca del Presidente Trump, DeepSeek rilascia i suoi modelli **R1-Zero e R1**, focalizzati principalmente sul ragionamento matematico e sulla programmazione. Nel giro di una settimana, l'applicazione suscita un notevole interesse da parte degli utenti, arrivando addirittura ad occupare il gradino più alto del podio come applicazione più scaricata nell'App Store americano.

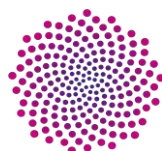
Come dichiarato nella documentazione del modello, DeepSeek R1-Zero, seppur promettente, soffre di limitazioni tra cui una compromessa leggibilità e un'eterogeneità linguistica, mescolando caratteri inglesi e cinesi. Il vero prodotto finale è invece il DeepSeek-R1 che, rispetto al suo predecessore, incorpora (1) cold start data, il che significa che il modello utilizza dati iniziali già disponibili e (2) una pipeline multi-stadio, che suddivide il training di un modello in diverse fasi sequenziali permettendo di gestire in maniera più efficiente la sua complessità. Per facilitare la comprensione, immaginiamo DeepSeek come l'indice di un libro. Anziché sfogliare pagina dopo pagina alla ricerca dell'informazione desiderata, **DeepSeek è in grado di individuare immediatamente la frase esatta, la parola chiave o il paragrafo che risponde alla nostra domanda.** È come avere un segnalibro intelligente che si sposta da solo, portando il lettore direttamente al cuore del contenuto. Con questo approccio, la startup cinese è riuscita ad ottenere costi per *token* notevolmente inferiori ai suoi rivali. La spesa in dollari per *token*, che rappresenta ciò che il modello deve analizzare o comprendere per contestualizzare una richiesta o un'istruzione, risulta addirittura 20-40 volte inferiore rispetto ai modelli della startup americana Open AI come GPT-4o oppure o1 (dettaglio nella tabella a lato).

EXHIBIT 5: **DeepSeek has priced their models 20-40x cheaper than OpenAI**

Model	\$/M Input Tokens	\$/M Output Tokens
Deepseek-chat (V3)	\$0.14	\$0.25
gpt-4o	\$2.50	\$10.00
Δ	-94%	-98%
Deepseek-Reasoner (R1)	\$0.55	\$2.19
o1	\$15.00	\$60.00
Δ	-96%	-96%

La seconda innovazione del modello di DeepSeek è la **distillazione**. Questo processo permette di trasferire le conoscenze dei modelli più grandi (i *teacher* o insegnanti) in un modello più piccolo ed efficiente (lo studente). Procedendo sempre per analogie, il termine "distillazione" ci rimanda al famoso processo per ottenere la grappa, con il quale condivide principi simili:

1. Entrambi i processi estraggono l'essenza più utile da una materia prima complessa, eliminando ciò che non serve per ottenere un prodotto più raffinato ed efficiente. Nella



## BRIGHTSIDE CAPITAL

grappa la base è la vinaccia, che contiene alcol, aromi e impurità, mentre nell'AI la base è un modello di grandi dimensioni (i.e. *Teacher Model*), addestrato su enormi quantità di dati, ma spesso troppo pesante per essere usato in modo efficiente.

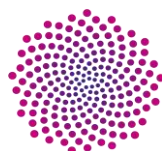
2. Le vinacce vengono riscaldate in un alambicco, trasformando alcol e aromi in vapore, così come il modello AI viene "riscaldato" attraverso un processo di *Knowledge Distillation*, in cui un modello più piccolo (i.e. *Student Model*) impara dai dati e dalle decisioni del modello grande.
3. Nel processo di separazione il distillatore elimina le frazioni indesiderate della grappa, come la testa e la coda, mantenendo solo il cuore, la parte più pregiata. Allo stesso modo, il modello AI distillato elimina le informazioni ridondanti, mantenendo solo le conoscenze più rilevanti e generalizzabili, riducendo il numero di parametri ma conservando le performance.
4. Dopo la distillazione la grappa può essere affinata in botti per migliorarne il gusto e la qualità, così come il modello distillato può essere ottimizzato ulteriormente con tecniche di *fine-tuning*, adattandolo a specifiche applicazioni.
5. Il risultato finale della distillazione della grappa è una bevanda raffinata, aromatica e più pura, mentre nella distillazione AI si ottiene un modello più leggero, veloce ed efficiente, pronto per essere usato in tempo reale su dispositivi meno potenti.

**Questa tecnica è molto comune nei modelli open-source**, che danno un elevato livello di flessibilità ai developer per esaminarli, modificarli e distribuirli a loro piacimento, contribuendo anche ad aumentare il network del progetto principale. D'altro canto, **Gemini (sviluppato da Google) e ChatGPT sono closed-source**: l'azienda fondatrice ne detiene le specifiche e la programmazione, in modo che il modello non venga ricreato da terzi e venduto. Uno dei motivi per cui le aziende scelgono strumenti closed-source, nonostante il costo aggiuntivo, è l'aspettativa da parte dell'utente di una manutenzione e di un supporto affidabili da parte della società madre e non da developer indipendenti.

Come riportato nel [white paper](#) pubblicato direttamente da DeepSeek (sezione 3.2), gli "insegnanti" del modello sono stati, come da tabella a lato, Qwen e Llama, rispettivamente l'AI open-source di Alibaba e di Meta. Sono, inoltre, state usate

Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64	pass@1	pass@1	pass@1	rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	1820
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633

Table 5 | Comparison of DeepSeek-R1 distilled models and other comparable models on reasoning-related benchmarks.



## BRIGHTSIDE CAPITAL

diverse varianti di questi due modelli, utilizzando come discriminante il numero di parametri preso in considerazione per ciascuna versione (come si evince in tabella).

In conclusione, utilizzando le fondamenta dei modelli open source di Alibaba e Meta (i quali hanno speso miliardi a loro volta per progettargli), DeepSeek è stata in grado di ottenere prodotti con ottime metriche di performance<sup>2</sup> e spendendo una frazione rispetto ai modelli rivali. Basti pensare che, per il modello V3 (introdotto il 26 dicembre del 2024), **la startup cinese ha dichiarato di aver speso, appunto, solamente 6 milioni di dollari per “allenarlo”**. Secondo il paper, DeepSeek ha utilizzato 2.048 GPU Nvidia H800 per un totale di 2.788.000 GPU hours, che rappresentano la quantità totale di calcolo usata, non il tempo effettivo trascorso. Ciò significa che se avessero usato una sola GPU, l’addestramento avrebbe richiesto 318 anni (ovvero 2.788.000 ore), ma grazie alle 2.048 GPU usate contemporaneamente, è stato completato in 57 giorni (2.788.000 diviso 2.048). Con un costo di \$2 per ora di GPU, il costo totale dell’addestramento è stato di **\$5.576 milioni**.

### I primi commenti dalla Silicon Valley e dal mondo tech

Proprio sul punto dei costi menzionato sopra e sull’utilizzo delle GPUs di Nvidia sono nati dei dubbi rappresentati dalle dichiarazioni del CEO di Scale AI che intervistato da CNBC afferma che DeepSeek abbia in realtà a disposizione 50.000 GPUs modello H100 (sottoposto al divieto di esportazione), più performanti rispetto alle H800 che invece DeepSeek afferma di aver utilizzato. Elon Musk sembra essere d’accordo con questa teoria, affermando perciò che in un modo o nell’altro, i cinesi siano riusciti ad aggirare le restrizioni.

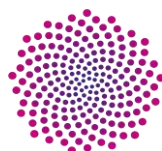
Poche ore dopo il tweet di Musk, Microsoft e OpenAi hanno lanciato un’indagine<sup>3</sup>, palesando il loro sospetto che DeepSeek abbia ottenuto in maniera illecita grandi quantità di dati direttamente da OpenAi. Teoria poi confermata dalla società di ChatGPT che, come riportato dal Financial Times<sup>4</sup>, accusa la startup cinese di aver usato il modello dell’azienda americana per allenare il loro software open-source. Nonostante questo, lo scorso giovedì Microsoft ha reso disponibile sulla piattaforma cloud “Azure” una nuova interfaccia per l’utilizzo della nuova AI cinese.

---

<sup>2</sup> <https://www.bracai.eu/post/DeepSeek-performance>

<sup>3</sup> ‘Microsoft probes if DeepSeek-linked group improperly obtained OpenAI data, Bloomberg News reports’, Reuters 29.01.25

<sup>4</sup> ‘OpenAI says it has evidence China’s DeepSeek used its model to train competitor’, Financial Times 29.01.25



## BRIGHTSIDE CAPITAL

Lasciando da parte il lato geopolitico per un momento (ce ne occupiamo nella sezione successiva) è innegabile che siamo davanti ad una rivoluzione nel modo di sviluppare l'intelligenza artificiale. **CNBC ha intervistato il CEO di [PerplexityAi](#) (motore di ricerca simile a ChatGPT), Arvid Srinivas, che descrive la necessità come la "madre" di tutte le invenzioni.** Srinivas articola questo concetto spiegando che, nonostante DeepSeek non sia, almeno ufficialmente, in possesso delle più avanzate componenti per lo sviluppo dell'intelligenza artificiale, sia comunque riuscita ad adattarsi in altri modi, creando un prodotto più efficiente e performante. DeepSeek, continua il CEO di Perplexity, è 10 volte più economico di ChatGPT e 15 volte in più di Sonnet (libreria di machine learning sviluppata da DeepMind).

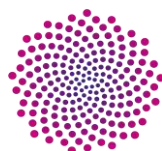
In conclusione, Srinivas dice di aver già iniziato l'implementazione di DeepSeek in Perplexity Ai, ciò permetterebbe a quest'ultima di realizzare più *task* ad un costo inferiore. Un punto critico però toccato dal giovane CEO è il tema dell'open-source: **[l'ampia adozione di DeepSeekR1 da parte degli sviluppatori americani potrebbe contribuire a consolidare la posizione della Cina come leader tecnologico globale.](#)**

Mentre in molti attendono di capirne di più su come effettivamente sia stato possibile per una startup cinese con circa 200 dipendenti arrivare ad una soluzione così efficiente ed economica, la sensazione generale è che in pochi si attendevano uno sviluppo così repentino. Nel 2023, Sam Altman, alla guida di OpenAI, il più grosso progetto di intelligenza artificiale generativa mai realizzato (con circa 18 miliardi di dollari raccolti dagli investitori), affermava che: "[è totalmente senza speranza competere con noi nell'addestramento di modelli fondamentali](#)"

**You can try, but not possible to build AI with \$10M: Sam Altman's old video goes viral after DeepSeek does it with \$5M**

### "DeepSeekology"

Guardando alla competizione tra potenze, il lancio dell'ultimo modello di intelligenza artificiale cinese DeepSeek ha certamente rappresentato uno shock. Le implicazioni geopolitiche, infatti, sono profonde e in sintesi le possiamo riassumere nei seguenti termini.



## BRIGHTSIDE CAPITAL

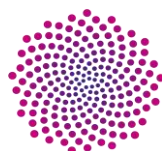
L'obiettivo principale di DeepSeek è sviluppare modelli di intelligenza artificiale che combinino la sovranità tecnologica riducendo, quindi, al minimo l'impatto delle sanzioni con la competitività globale, soprattutto nei mercati emergenti alleati della Cina, come Asia, Africa e America Latina. Finora, i decisori politici statunitensi hanno cercato di limitare le ambizioni tecnologiche della Cina attraverso i controlli sulle esportazioni e le sanzioni. Tuttavia, l'ascesa di DeepSeek suggerisce che queste misure **potrebbero esser inadeguate a fermare i progressi cinesi nello sviluppo dell'AI**. Rilasciando i suoi modelli con una licenza open source, DeepSeek ha dato potere agli sviluppatori e alle aziende di tutto il mondo per sfruttare le sue innovazioni. Questa mossa non solo accelera l'innovazione nell'ecosistema dell'AI, ma rappresenta anche una sfida diretta ai modelli proprietari che dominano il settore: a differenza di aziende occidentali come OpenAI o Google, che mantengono i loro modelli strettamente controllati, DeepSeek ha optato per trasparenza e accessibilità, incoraggiando una più ampia collaborazione all'interno dell'ecosistema.

In tal senso, gli Stati Uniti potrebbero dover **riconsiderare il loro approccio al mantenimento della leadership**. L'influente investitore VC Marc Andreessen ha definito tale evento un "**momento Sputnik**", riferendosi a quando gli Stati Uniti, scossi dal lancio del satellite sovietico, avevano investito miliardi in un modello di partnership tra settore pubblico e privato che ha contribuito a conquistare e sostenere il predominio tecnologico che avrebbe svolto un ruolo importante nella vittoria della Guerra Fredda.



L'equilibrio di potere nella corsa all'innovazione nel campo dell'AI sta cambiando, annunciando una nuova era di competizione e collaborazione nel mondo della tecnologia. Così, se una parte dell'amministrazione Trump, come ad esempio Marco Rubio, confermato Segretario di Stato, spinge per l'inasprimento dei dazi alla Cina, altri potrebbero avere opinioni divergenti sulla Cina. Elon Musk, ad esempio, ha notevoli interessi commerciali in Cina e ha trascorso anni a coltivare stretti legami con la leadership cinese in qualità di CEO di Tesla.

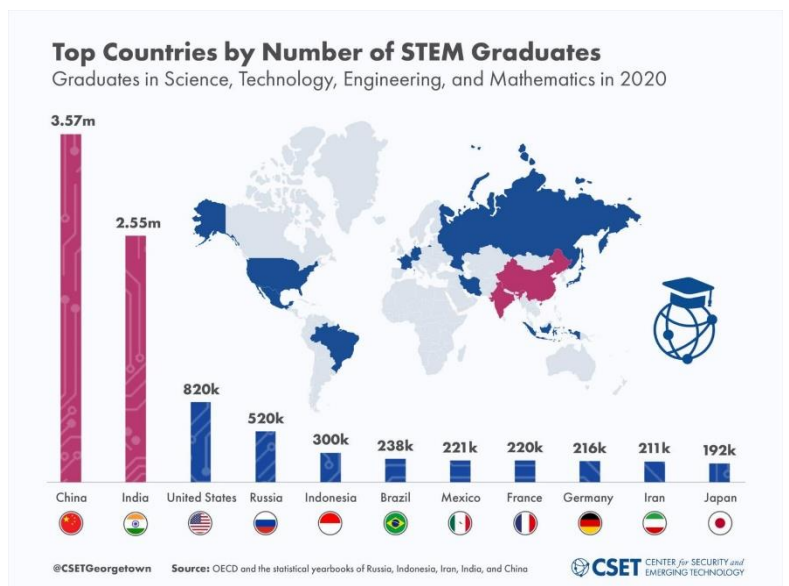
Come rileva Alessandro Aresu, studioso specializzato in scenari geopolitici legati all'industria tecnologica, di recente autore del saggio *Geopolitica dell'intelligenza artificiale* (ed. Feltrinelli), Trump ha fatto intendere di voler aprire un "grande negoziato" con la Cina, un G2 per



## BRIGHTSIDE CAPITAL

l'organizzazione di un mondo meno conflittuale<sup>5</sup>. Nel suo primo giorno di ritorno alla Casa Bianca, Trump ha inaspettatamente sospeso i dazi sulla Cina e non l'ha individuata come una minaccia, alimentando la prospettiva di un riavvicinamento. In un discorso pronunciato dopo il suo insediamento, il presidente degli Stati Uniti si è astenuto dal menzionare la Cina, suo ex avversario in una precedente guerra commerciale, pur affermando che i dazi avrebbero reso gli Stati Uniti "ricchi da morire", lasciando la porta aperta a nuovi negoziati con la seconda economia più grande del mondo<sup>6</sup>. Il presidente degli Stati Uniti ha anche rinviato il divieto di TikTok, suggerendo, allo stesso tempo, che gli Stati Uniti dovrebbero essere proprietari della metà delle attività statunitensi dell'app cinese in cambio del suo mantenimento in vita<sup>7</sup>. È ragionevole pensare che la Cina si prepari a questo negoziato mostrando tutte le sue carte, enfatizzando l'inutilità delle sanzioni oltre che la sua capacità di resilienza.

Infine, uno degli aspetti più importanti da riconsiderare è che il lancio dell'AI cinese evidenzia in tutta la sua gravità il declino non solo materiale ma anche umano che affligge l'industria americana e, più in generale, occidentale. Questione dettagliatamente trattata da Emmanuel Todd nel suo ultimo libro che spiega i fattori del declino occidentale, il problema centrale della competizione è che in occidente molti dei giovani più meritevoli hanno optato in modo sproporzionato per carriere universitarie in grado di garantire professioni più redditizie nella finanza e nei servizi, a scapito di percorsi formativi per i professionisti cosiddetti Stem (scienza, tecnologia, ingegneria e matematica) in grado di determinare innovazione in settori strategici come l'intelligenza artificiale<sup>8</sup>. Problema aggravato dal fatto che un numero crescente di scienziati cinesi sta lasciando gli Stati Uniti per incarichi all'estero, ultimo indicatore di come il peggioramento delle relazioni tra Stati Uniti e Cina negli ultimi anni ha complicato la collaborazione accademica e potrebbe ostacolare le ambizioni tecnologiche di Washington. Gli

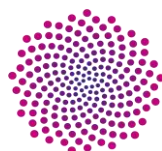


<sup>5</sup> Cfr. Muratore, A., *Alessandro Aresu: "DeepSeek, la sorpresa cinese che fa il botto sull'IA"*, in InsideOver, 27.01.2025.

<sup>6</sup> Cfr. Cash, J., Chen, X., *Trump raises prospects for a negotiated reset on US-China ties*, in Reuters, 21.01.2025.

<sup>7</sup> Cfr. Aa.Vv. *What does Trump's executive order mean for TikTok and who might buy it?*, in BBC, 22.01.2025.

<sup>8</sup> Cfr. Todd, E., *La sconfitta dell'Occidente*, Fazi Editore, Roma, 2024.



## BRIGHTSIDE CAPITAL

scienziati cinesi che vivono negli Stati Uniti hanno contribuito per decenni agli sforzi di ricerca che hanno guidato gli sviluppi nella tecnologia avanzata e nella scienza. Un numero crescente di loro cerca lavoro altrove, poiché il deterioramento delle relazioni geopolitiche alimenta un controllo extra dei ricercatori cinesi e Pechino intensifica gli sforzi per reclutare e trattenere i talenti<sup>9</sup>.

Qui ci avviciniamo alle radici del potere: il numero degli ingegneri rileva non solo per il predominio tecnologico o per la capacità economica reale di produrre beni e armi ma anche, strategicamente, per le capacità militari di un Paese. Ad esempio, la maggior parte degli ufficiali di un esercito moderno, soprattutto nei reparti tecnici di aeronautica e marina, sono ingegneri. Che gli Stati Uniti non siano in grado di formarne in numero sufficiente solleva dubbi sul reale potenziale dell'esercito americano in caso di conflitto.

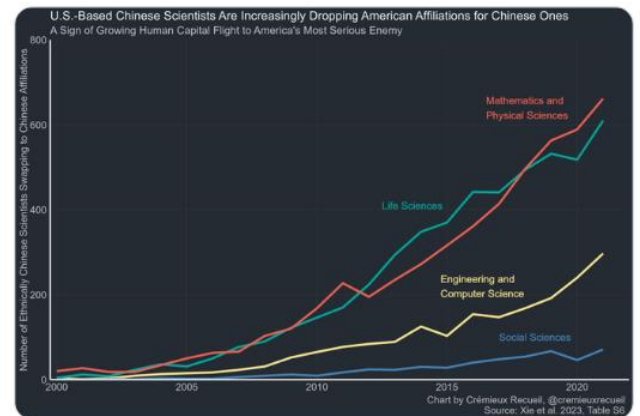


Crémieux @cremieuxrecueil · Oct 11, 2024

Ethnically **Chinese** scientists living in the U.S. are increasingly swapping their affiliations away from American universities **to Chinese** ones.



That's human capital flight **to** a state enemy.



169

325

2.2K

254K



**La fuga di cervelli verso le facoltà di diritto o di economia minaccia direttamente la potenza militare americana:** non si vince imponendo al nemico sanzioni economiche, questo ormai la guerra ucraina l'ha reso chiaro a tutti.

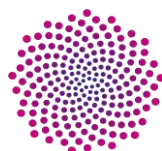
*“Non solo il livello ma la stessa sussistenza dell'industria dipendono dall'esistenza di centinaia di migliaia di menti qualificate e ben addestrate che dominano e fanno progredire incessantemente la tecnica. L'ingegnere è propriamente il silenzioso dominatore e il destino dell'industria meccanica. Il suo pensiero è come possibilità quello che la macchina è come realtà.”<sup>10</sup>*

### L'opportunità

La maggior parte degli investitori, analisti e commentatori di mercato sta focalizzando le proprie valutazioni su come DeepSeek potrebbe portare a una revisione al ribasso del CAPEX

<sup>9</sup> Cfr. Lu, C., *Chinese Scientists Are Leaving the United States*, in Foreign Policy, 13.07.2023.

<sup>10</sup> Spengler, O., *Il tramonto dell'Occidente. Lineamenti di una morfologia della storia mondiale*, ed. Longanesi, Milano, 1981.

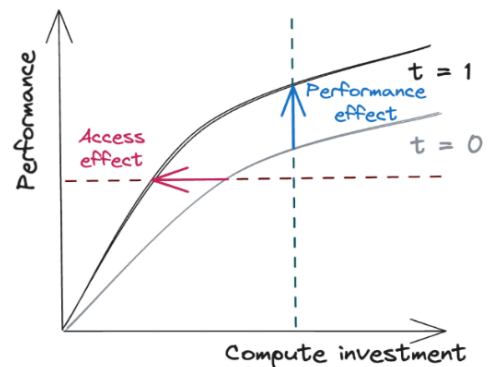


## BRIGHTSIDE CAPITAL

previsto dagli *hyperscaler*. L'alternativa a questa visione è che, come anticipato dal CEO di PerplexityAi, è che i principali sviluppatori di modelli LLM (o Large Language Model come GPT-4°, Gemini, Llama) stiano già studiando e integrando le efficienze architetturali di DeepSeek, ma con l'obiettivo di applicarle a un numero ancora maggiore di parametri e con una potenza computazionale superiore, cercando così di superare i risultati ottenuti dalla startup cinese.

Lennart Heim, ricercatore specializzato nella governance dell'intelligenza artificiale e Associate Information Scientist presso RAND Corporation, un think tank con sede a Washington, illustra la relazione tra investimento e performance: se DeepSeek è riuscito ad ottenere le stesse prestazioni con meno risorse computazionali (freccia rossa del grafico sotto), allora le stesse tecniche potrebbero permettere agli altri di raggiungere livelli di performance ancora più elevati con un maggiore investimento computazionale (freccia blu). Anche Sam Altman ha espresso un concetto simile in un tweet.

In conclusione, le dichiarazioni di Altman e i commenti del presidente Trump (*"L'introduzione di DeepSeek da parte di un'azienda cinese dovrebbe essere un campanello d'allarme per le nostre industrie, spingendoci a concentrarci con la massima attenzione sulla competizione"*) confermano che la corsa al progresso dell'intelligenza artificiale – qualunque forma essa assumerà – è ormai considerata un obiettivo chiave nazionale, paragonabile per importanza strategica forse solamente al Progetto Manhattan, che portò alla creazione della prima arma nucleare della storia durante la seconda guerra mondiale. **The race is on.**



Approfondimento a cura di **Nicola Lampis e Gilberto Moretti**

Lugano, 2 febbraio 2025

