IFRI
MEMOS

Geopolitics of Technology Center

MARCH 26, 2025

A “DeepSeek Moment”?

Benjamin PAJOT

Key Takeaways

- DeepSeek, hailed as a champion of Chinese AI, represents less a revolution than a significant optimization of existing technologies.
- Doubts remain regarding the figures put forward by the start-up, inviting a more measured response to the media hype surrounding China’s technological catch-up.
- Nonetheless, DeepSeek signals the need to question an economic model based solely on the race for computational power.
- By betting on open innovation, Europe can carve out its own path in a competition that is far from being a zero-sum game.

Introduction

The release of the generative artificial intelligence (AI) model DeepSeek-R on January 20th, 2025, set off a wave of more or less sincere panic in Silicon Valley. With performances surpassing those of OpenAI’s ChatGPT4, the small Chinese start-up, 143 employees strong at the time, reaffirmed its talents, first glimpsed a month earlier with the publication of DeepSeek v3 in late 2024. Founded and led since 2023 by Liang Wenfeng, head of the High-Flyer hedge fund of which DeepSeek is a subsidiary, the company has benefited from substantial initial funding, allowing it to devote itself to research rather than chasing profitability. Its ties to Chinese authorities remain unclear, and its sudden success may have come as a pleasant surprise for Beijing, which has hailed it as a national champion in its technological competition with Washington.

Rather than a “Sputnik moment”, the United States (US) leaders of this rapidly changing industry are now experiencing a “DeepSeek moment”, as so-called “reasoning” models mark a significant step in the evolution of large language models (LLMs). DeepSeek-R1, drawing inspiration from its competitors, also reflects the growing popularity of distillation, a process whereby the capabilities and “knowledge” of another AI model can be extracted by training based on its responses, resulting in a credible iteration of the model. The freely accessible model reflects China’s strategic positioning in open-source AI, and challenges Europe’s role in a game that appears to be less rigid and more open.

Beyond the media frenzy

Serious contender, temporary market panic

DeepSeek-R1’s performance, after undergoing a series of benchmark tests, has made it the leading open model (ahead of Llama3, developed by Meta) and a direct competitor to closed models such as OpenAI’s 01 and Anthropic’s Claude 3.5 Sonnet, which until now were considered to be among the most advanced on the market. These results, together with the figures provided by the Chinese start-up showing the model’s low development costs and major efficiency gains, sent shockwaves through the sector in the US. They were interpreted as an indication not only of China catching up to US technologies, but also of a looming economic impasse for US players in the sector. This resulted in an unusually severe stock market panic, with chip manufacturer NVIDIA losing 17% of its valuation (600 billion dollars, $).

While there is still debate over development costs, the efficiency gains appear to be real. These are made possible by a selective activation method: instead of running the entire model each time a query is made, the model is divided into a number of “expert” sub-models, which are only called upon when the query falls within their area of expertise.

These gains should theoretically lead to a reduction in model development and operating costs, making them more accessible and therefore more marketable. However clever it may be, this kind of optimization certainly could be replicated, which explains why US competitors have devoted so much effort to dissecting DeepSeek-R1.

The stock market panic quickly subsided once the initial shock died down, and investors took stock of what was, on the whole, a relatively minor upset. The real concern was not so much the intellectual prowess and modest size of the Chinese team, but rather its nationality. This was evidenced by an open-source reasoning model developed by US researchers, which performed similarly to OpenAI and DeepSeek for less than $50, but was almost completely ignored.[1] Confronted once again with their fear of being technologically outdone, previously by Russia and Japan, and now by China, the United States seemed overcome with doubt. Confronted with a company anointed by Xi Jinping to embody AI "with Chinese characteristics" — a mix of state guidance, private sector ingenuity, and beneficial open-source collaboration — the US seemed to believe itself in a new space race.

> Confronted once again with the fear of being technologically outdone, the US seemed overcome with doubt

The superficial appeal of a "new Sputnik moment"

Many were quick to comment on the launch of the Chinese model, describing it as a "new Sputnik moment".[2] Tempting as it may be to compare it with the launch of the first Soviet satellite in 1957 — which sent shockwaves through US strategic circles and led to the creation of the National Aeronautics and Space Administration (NASA) — seems disproportionate. DeepSeek-R1 is neither a technological revolution nor a new technical paradigm, rather an optimization of existing technology. It does not, in other words, constitute a major technological or scientific breakthrough, comparable to the launch of the world's first satellite. It has not jeopardized US leadership, at a time when massive investments in AI demonstrate the country's long-term commitment in this competitive arena.

The strategic feigned surprise of US stakeholders when faced with a Chinese company capable of competing with their national ecosystem is more reminiscent of that which seemed to shake Silicon Valley when Huawei launched a line of smartphones with cutting-edge chips (reaching the 7-nanometer mark).[3] The US's recurring bouts of media frenzy need to be considered in their historical context. Washington regularly resorts to singling out an adversary, unilaterally or otherwise, to initiate a productive power struggle in order to fuel the growth of its technology sector.

1. M. Zeff, "Researchers Created an Open Rival to OpenAI's o1 'Reasoning' Model for under $50", *TechCrunch*, February 5, 2025.
2. S. Kauffmann, "Contrairement au 'moment Spoutnik' en 1957, il n'y aura probablement pas de moment DeepSeek aux États-Unis", *Le Monde*, January 29, 2025.
3. Q. Liu, "How Huawei Surprised the US with a Cutting-edge Chip Made in China", *Financial Times*, November 30, 2023.

A great leap forward for China?

Following DeepSeek's release, the model's integration into major Chinese platforms (WeChat, Baidu, Huawei Cloud), and its partnerships with various national industrial players (a dozen motorcycle and car manufacturers and several major ports) reveals a form of patriotic frenzy surrounding the start-up, which has been held up as a symbol of Chinese resurgence opposite the US.[4] This dynamic should not be underestimated, however, as the rise of Chinese stakeholders seems real and profound, in spite of US efforts to prevent their emergence.[5] Other high-performance models are currently being developed by China's leading technology firms: Qwen 2.5 (Alibaba), Ernie Bot (Baidu), Doubao 1.5 Pro (ByteDance) and Kimi k1.5 (Moonshot AI) claim to be as good as or better than US models in certain segments.[6]

The rise of Chinese stakeholders seems real, in spite of US efforts to prevent their emergence

This trend could accelerate with the launch of a new national investment fund, totaling 8.2 billion dollars, dedicated to the development of Chinese AI.[7] Beijing is also actively engaged in international governance initiatives in order to strengthen its influence on global AI standards and frameworks. Even so, the momentum generated by Chinese competition could ultimately benefit the US ecosystem, spurring it on to greater innovation.

The alleged failure of the US export control regime

The release of DeepSeek-R1 also prompted extensive commentary on the US's export control policy toward China. Aimed at preventing China from overtaking the US technologically due to national security concerns, these measures restrict access to innovations in the field of semiconductors, in particular NVIDIA's H100 chips, currently considered the most advanced on the market.[8] That DeepSeek was able to develop such an advanced model, supposedly without access to these chips, has given rise to several hypotheses. Either the Chinese start-up is telling the truth, and has therefore managed to overcome the technical limitations imposed by this regime, by optimizing its use of inferior H800 chips. In this case, DeepSeek-R1 would be an example of optimization and adaptation in response to US retaliatory measures. Alternatively, the company is concealing that it was able to previously acquire, or has ongoing access, to H100 chips,

4. T. Westbrook and S. Shen, "Nicknames Bloom in China's Tech Renaissance", Reuters, February 19, 2025.
5. H. Booth, "How China Is Advancing in AI Despite U.S. Chip Restrictions", *Time*, January 28, 2025.
6. Z. Soo, "DeepSeek Has Rattled the AI Industry: Here's a Look at Other Chinese AI Models", *Time*, January 28, 2025.
7. B. Jiang, "Tech War: China Creates US$8.2 Billion AI Investment Fund amid Tightened US Trade Controls", *South China Morning Post*, January 20, 2025.
8. M. Velliet, "From Non-proliferation to Strategic Competition: US Export Controls and China", *International Politics*, January 6, 2025.

which would mean that the smuggling networks making them available in China are more capable than has been estimated. In either case, this situation casts doubt on the effectiveness of the export control regime.

While one of the Biden administration's final measures was to work to reinforce it, many stakeholders – starting with NVIDIA, for whom the Chinese market is essential – are now trying to reduce its scope and are intensively lobbying US authorities to that end.[9] It should, however, be emphasized that this is not a zero-sum game: DeepSeek-R1's development was probably fueled by the desire to find practical solutions to overcome the US embargo, but it is equally true that the embargo may have limited how "revolutionary" the Chinese model could become. While demand for NVIDIA chips continues to surge in China, we could well see a tightening of the export control regime in the near future, in the name of Trump's America First doctrine.[10]

Doubts remain

Questions concerning the cost of training

Part of the reason DeepSeek-R1's cost of training caused such a stir is that it seemed to be undervalued. The Chinese start-up adopted an aggressive communications strategy, citing marginal training costs ($6 million), which did not factor in underlying infrastructure costs – DeepSeek owns its own data centers –, payroll, or preliminary research and development (R&D) costs. Suspicions were raised further when a specialized firm claimed that the Chinese company actually had 10,000 H100 chips at its disposal. In this case, DeepSeek's technical architecture would have required $1.6 billion in investment, a more plausible figure.[11] Such an allegation, if proven true, would effectively nullify some of DeepSeek-R1's supposed advances and challenge certain related claims (regarding China's technological rise, its economic threat, or the effectiveness of the export control regime). But it would not wipe the slate clean: DeepSeek-R1 would remain a locally available open-source model, requiring fewer resources and energy, supported by a composite architecture, with an innovative "reasoning" architecture.

Concerns surrounding practices

DeepSeek-R1 has drawn heavy criticism regarding its openness and security. The model has indeed been made public, but its training data and its adjustment algorithm (fine-tuning, specialization phase of the pre-trained model) are still inaccessible. Without

9. T. Hunnicutt et al., "Trump and Nvidia CEO Discuss DeepSeek, AI Chip Exports during Meeting, Source Says", Reuters, February 1, 2025.
10. F. Potkin and C. Pan, "Exclusive: Nvidia's H20 Chip Orders Jump as Chinese Firms Adopt DeepSeek's AI Models, Sources Say", Reuters, February 25, 2025.
11. D. Patel et al., "DeepSeek Debates: Chinese Leadership on Cost, True Training Cost, Closed Model Margin Impacts. H100 Pricing Soaring, Subsidized Inference Pricing, Export Controls, MLA", Semianalysis, January 31, 2025.

human supervision, it is likely that training will reinforce biases rather than rectify them. Having been developed with a Chinese audience in mind, the model also conforms to Beijing's expectations in terms of propaganda and revisionism, as its responses to queries concerning Tiananmen, Taiwan and Uyghurs have shown.

In terms of security, the DeepSeek app reportedly collects personal data (as is typical of such services, similarly to TikTok or Instagram), which is then hosted in China by digital giant ByteDance. This in itself could constitute a security issue, as Chinese companies are subject to the 2017 Intelligence Law (reinforced in 2023), which entitles Beijing to access their data. What's more, these data transfers to Chinese servers are not always encrypted, and according to experts, the application lacks basic security measures.[12] The US cybersecurity firm Wiz detected and reported a major data breach exposing DeepSeek users' queries and personal information.[13] DeepSeek-R1 is also reportedly susceptible to cyberattacks via prompt injections aimed at generating toxic content,[14] but also appears to have cyberoffensive capabilities, which are thought to be due to its "reasoning" abilities. Its vulnerability to jailbreaking (the circumvention of safeguards via malicious prompts) could turn it into a vector for malware development.[15] These concerns have consequently led to attempts to ban the application in various countries (Australia, South Korea, the US, Italy, Taiwan).

An economic bombshell

Challenging the "bigger is better" strategy

The real disruption caused by DeepSeek-R1's release is that it challenges the sector's dominant economic model. By reducing inference costs and, above all, the computing power needed to train and run its model, DeepSeek has challenged the "bigger is better" strategy embraced by the sector's US leaders. In practice, this amounted to developing ever larger AI models, with huge data and computing power requirements, and correspondingly high chip and energy demands. Such mammoth operations, which also served as barriers to entry for would-be newcomers to the industry, were until now financed mainly by fundraising and capital increases. In the absence of competition, the dominant players had not bothered to improve their energy and economic efficiency.

DeepSeek-R1's release disrupts the sector's dominant economic model

12. A. Hoog, "NowSecure Uncovers Multiple Security and Privacy Flaws in DeepSeek iOS Mobile App", NowSecure, February 6, 2025.
13. G. Nagali, "Wiz Research Uncovers Exposed DeepSeek Database Leaking Sensitive Information, Including Chat History", Wiz, January 29, 2025.
14. D. Goodin, "DeepSeek iOS App Sends Data Unencrypted to ByteDance-controlled Servers", Ars Technica, February 6, 2025.
15. N. Miles, "DeepSeek Deep Dive Part 1: Creating Malware, Including Keyloggers and Ransomware", Tenable, March 13, 2025.

Their potential profitability has suddenly come under threat, and shareholders will fear for their return on investment. DeepSeek's open-source distribution also paves the way for the development of low-cost services, opening up opportunities for less well-funded competitors, while forcing industry leaders to adapt.[16] As distillation becomes the new norm and developers appear to be able to duplicate proprietary models and release them under an open-source license, an entire economic model is being upended.[17] This could elicit a legal response from these models' proprietors (led by OpenAI, Anthropic and Google), with the support of US authorities.[18]

Big Tech may stand to gain

While DeepSeek may initially appear threatening to the Silicon Valley giants, who provide the majority of the hardware infrastructure necessary for AI training, the situation may not actually be quite so dire. If the Jevons paradox holds true, they could directly benefit from more widespread adoption of AI models, which would ultimately offset the decline in hardware demand resulting from the efficiency gains achieved by DeepSeek-R1.[19] High-performance computing needs could even grow proportionally to the commercialization of generative AI. NVIDIA, the main focus of the stock market panic, began to recover as this hypothesis appeared to gain traction in financial circles. The company should continue to dominate the sector as it can adapt to the latest developments (lower inference costs, chips tailored to the increasingly specific requirements of reasoning models) and meet increased demand. Similarly, Google, Meta and OpenAI already have the necessary infrastructure and expertise to adapt, in this case by integrating the optimization strategies implemented by DeepSeek.

> The Silicon Valley giants could be tempted to capitalize on the panic generated by DeepSeek

In addition, actors like Microsoft, Amazon and Apple, which have so far lagged behind in the field of generative AI, will also benefit from both lower development costs and the opportunities presented by the availability of open-source models. Apple, which designs its own unified memory chips, making it easier to run lighter models on its products, could even take center stage again by integrating consumer AI services into its products.[20] Though they should therefore only suffer minor setbacks from these latest developments, the Silicon Valley giants could nevertheless be tempted to capitalize on the panic generated by DeepSeek in the United States. Whether they advocate for the sector's

16. R. Waters, "Big Tech Is Moving on from the DeepSeek Shock", *Financial Times*, February 14, 2025.
17. B. Edwards, "Hugging Face Clones OpenAI's Deep Research in 24 Hours", Ars Technica, February 5, 2025.
18. M. Sweney and D. Milmo, "OpenAI 'Reviewing' Allegations That Its AI Models Were Used to Make DeepSeek", *The Guardian*, January 29, 2025.
19. Named after a 19th-century British economist, this paradox associates technical optimization with an increase in resource consumption.
20. E. Sayegh, "DeepSeek Hype Vs. Nvidia And Apple: Will Calmer Minds Prevail?", *Forbes*, January 28, 2025.

complete deregulation or call for it to be consolidated under them and to their exclusive benefit, whatever shape their lobbying may take is a legitimate cause for concern.[21]

Could the AI bubble burst?

This challenge to the "bigger is better" strategy implicitly undermines the rationale behind the spending the biggest players in AI have indulged in until now. Any such doubts surrounding the sector's economic viability are only exacerbated by these companies' significant operating deficits.[22] As semiconductor miniaturization gradually grows more difficult, LLMs – by their designers' own admission – come up against their limits, and synthetic or fragmented data makes their training increasingly fraught (hallucinations, self-intoxication, total inexplicability), concerns about the AI bubble bursting are being raised once more. What's more, AI systems' growing energy consumption, which efficiency gains will struggle to offset, not only drives up costs but could also constrain the operation of up to 40% of data centers due to power shortages.[23] In this respect, the White House's "Stargate" project is likely to exacerbate the issue and quickly run up against the reality of available resources, as competing energy needs already appear to be a major issue in concerned regions.[24] There is, however, no guarantee that the bubble will burst, as the massive investments which have been announced in the United States and Europe, including France, together with venture capital, should help to keep the sector afloat after a record year in 2024.[25]

Lessons for Europe

Putting an end to the "winner takes all" mentality

This series of events was a stark reminder that the "AI race" is a global phenomenon and is anything but set in stone, still less like a casino game where the winner walks away with the entire pot. This "winner takes all" mentality, which prevails in the US and in China, is in fact deliberately fostered in order to stifle all competition once a technical lead has been established. DeepSeek is proving once again that becoming the dominant innovator does not guarantee a lasting advantage: the ability to innovate is not solely dependent on capital strength or stock market hierarchy, and there are still many paths forward for AI. With development no longer simply a matter of accumulating computing power, and efficiency and accessibility also coming into play, European actors have more opportunities to

21. T. Wheeler, "DeepSeek Is Not a Good Reason for Big Tech to Become More Powerful", Brookings, February 11, 2025; B. Pajot, "Les risques de l'IA. Enjeux discursifs d'une technologie stratégique", *Études de l'Ifri*, Ifri, June 2024.
22. C. Weinberg, "OpenAI Projections Imply Losses Tripling to $14 Billion in 2026", *The Information*, October 9, 2024; D. Moyo, "Three Reasons Why AI's Momentum Could Stall in 2025", Project Syndicate, January 23, 2025.
23. "Gartner Predicts Power Shortages Will Restrict 40% of AI Data Centers By 2027", Gartner, November 12, 2024.
24. C. Albon, "Power Generation Challenges Could Overshadow Stargate AI Initiative", *Defense News*, January 24, 2025.
25. I. Lunden, "AI Investments Surged 62% to $110B in 2024 While Startup Funding Overall Declined 12%", TechCrunch, February 11, 2025.

compete with their US and Chinese counterparts. Many of them are, in fact, involved in the design of small-scale and specialized models: they could therefore stand to benefit from the economies of scale achieved through optimization.

That said, it would be foolish to suggest that investment in physical infrastructure is no longer a priority. Even with these efficiency gains, the sector remains dependent on advanced semiconductors. While distillation allows for the inexpensive reproduction of costlier models, it has not, as of yet, allowed us to break new ground and create models far superior to those already on the market. In its success, DeepSeek implicitly reaffirmed the value of operating its own data centers. The announcement of French and European infrastructure development plans at the Paris Summit is one attempt at a solution, even if it remains questionable from a financial, energetic and ecological point of view. There is a risk of falling into the same "bigger is better" trap we can see in the US, when instead the emphasis should be on pooling computing and research capabilities. But this would require an industrial planning strategy unencumbered by technological nationalism within the European Union (EU), the prospect of which still seems remote.

Investing more in open innovation

Following China's national strategy of favoring open-source software as a means of strengthening its technological self-sufficiency, it appears that Chinese actors today are pursuing the development of open-source AI for strategic purposes.[26] The aim is as much to attract talent as to encourage the adoption of Chinese models, facilitating data collection and investments, as well as exerting a form of soft power in the field of engineering, in particular towards emerging countries, in line with the Belt and Road Initiative.[27] This strategy aims to challenge the proprietary US ecosystem and circumvent the barriers to innovation it has created. Such a strategy, which promotes the idea of AI as a national resource or even a public good rather than a corporate asset, would make sense for Europe.

DeepSeek has positioned itself where European actors should have

DeepSeek has positioned itself where European players should have: a historic home and major contributor to open source. Europe, too, is capable of producing innovations of this magnitude. While open source "with Chinese characteristics" has strict boundaries –the government encouraging collaboration within a controlled system where state-backed companies and tech giants hold the reins – Europe can and must promote a different approach, based on its values and principles. With this in mind, the Paris AI Summit helped to identify some promising ideas that should be further explored.

26. A. Panier, "Software Power: The Economic and Geopolitical Implications of Open Source Software", *Études de l'Ifri*, Ifri, December 12, 2022.
27. P. Bloom, "DeepSeek: How China's Embrace of Open-source AI Caused a Geopolitical Earthquake", *The Conversation*, February 12, 2025.

The strengthening of the Open Euro LLM project and the creation of the CurrentAI foundation, dedicated to funding and developing open-source and frugal AI, are promising instruments for the promotion of open innovation. They invite us to think of AI as a public, collective good, closely tied to the digital commons.

Exploring new avenues

DeepSeek's greatest contribution has been to expand the realm of possibilities. There are now a number of approaches with the potential to influence the future of AI. In terms of hardware, a range of innovations, from photonic chips (which use light rather than electricity to transmit data) to neuromorphic processors (inspired by the structure of neurons), could reduce AI energy consumption. R&D investments in alternative semiconductor materials and advances in rare earth and metal recycling practices could reduce their environmental impact. Finally, designing data centers locally and distributing them more evenly across regions, so that the heat they produce can be reused and to mitigate water resource requirements for cooling, is another potential strategy.

DeepSeek's greatest contribution has been to expand the realm of possibilities

In terms of software, optimized computation techniques like "linear complexity multiplication" (L-Mul) could significantly improve energy efficiency. While "reasoning" models have given new momentum to a flagging LLM market, some actors, like Meta, are now busy developing "Large Concept Models" (LCMs). Supposedly capable of abstract thought, such models would be able to develop contextual understanding independently of the language used to train them or interact with them.

Conclusion

Neither insignificant nor a Sputnik moment, this "DeepSeek moment" should instead inspire Europe to rise to the challenge and assert itself on the world stage. The highly publicized rivalry between China and the US belies the fact that both countries are pursuing AI development strategies that are motivated less by innovation and concern for the public interest than by economic and political control. Regardless of the optimizations it has achieved, DeepSeek-R1 exemplifies an engineering philosophy that prioritizes capacity expansion (in terms of scale, speed, or efficiency) over the responsible management of AI's societal impacts.

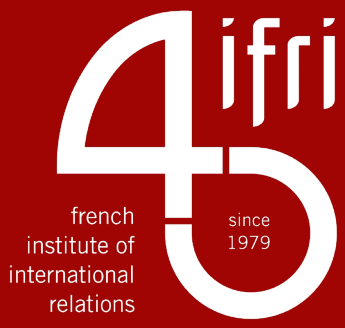
While the Chinese model's appearance has thrown the entire sector's economic viability into question, there is still a long way to go before AI is able to serve the public interest. Europe should seize this opportunity to try to guide the development of AI and harness this technology's potential. It can rely on the strength of its own ecosystem and on open innovation, which will be key to the emergence of AI "with European characteristics".

Benjamin Pajot *is an Associate Research Fellow at Ifri's Geopolitics of Technologies Center and a freelance researcher in digital geopolitics. He has previously worked as a Project Manager on digital and cyber issues at the French Ministry of Europe and Foreign Affairs' Strategy, Analysis and Policy Planning Centre (CAPS).*

How to cite this publication:

Benjamin Pajot, "A 'DeepSeek Moment'?", *Ifri Memos*, Ifri, March 26, 2025.

ISBN: 979-10-373-1050-7



27 rue de la Procession
75740 Paris cedex 15 – France

Ifri.org