DeepSeek R1 and DeepSeek V3: Technical Innovations and Impact

This paper was generated by OpenAI's Deep Research system (o1-pro) on the 9 February 2025.

1. Technical Overview

Architecture and Training – *DeepSeek-V3* is a large-scale **Mixture-of-Experts (MoE)** language model, boasting 671 billion total parameters with about 37 billion activated per token [huggingface.co [1]]. This MoE design means the model is split into specialized "expert" sub-models, and only the most relevant expert is engaged for a given input. Such an approach greatly saves computation since not all parameters fire for every query [datacamp.com [2]]. DeepSeek-V3 also introduces **Multi-Head Latent Attention (MLA)** and a novel **auxiliary-loss-free load balancing** mechanism to keep all experts utilized without the extra training loss term that MoE models usually need [huggingface.co [3]]. Another innovation is a **multi-token prediction objective**, allowing the model to predict multiple tokens per step during training [huggingface.co [4]] – a training tweak that boosts efficiency and performance. The model was pre-trained on an unprecedented 14.8 trillion tokens of high-quality text data [huggingface.co [5]], followed by supervised fine-tuning and reinforcement learning stages to refine its capabilities. Despite this massive scale, DeepSeek-V3's training was remarkably cost-effective: the company reports needing only about 2.788 million GPU-hours on Nvidia H800 chips [huggingface.co [6]], which corresponds to a **training budget of roughly $5.5 million** [lawfaremedia.org [7], lawfaremedia.org [8]]. This is orders of magnitude lower than the tens or hundreds of millions of dollars spent training comparably advanced models in the West. Such efficiency was achieved by intensive software and architectural optimizations – for example, combining known engineering tricks (custom GPU communication protocols, memory savings, etc.) to maximize output from limited hardware [wired.com [9], wired.com [10]]. In fact, DeepSeek's latest model is so computationally efficient that it **required only one-tenth the computing power of Meta's comparable Llama 3.1 model** to train [wired.com [11]]. The training process was also unusually stable (no irreversible loss spikes or restarts were needed) [huggingface.co [12]], demonstrating the robustness of DeepSeek's engineering.

DeepSeek-R1 is built on DeepSeek-V3 as a specialized *"reasoning"* model. Architecturally, R1 uses the same base network as V3 but is further trained (via reinforcement learning) to excel at complex, multi-step problem solving [datacamp.com [13], datacamp.com [14]]. During R1's fine-tuning, the model was allowed to generate multiple solution attempts for challenging tasks (like difficult coding or math problems), and a rule-based reward system reinforced correct reasoning paths [datacamp.com [15]]. In essence, R1 "learned how to think" by practicing on problems and receiving feedback. The result is a model that can engage in **chain-of-thought reasoning**: when prompted with a hard question, R1 internally works through a series of logical steps before producing a final answer [datacamp.com [16]]. Unlike V3, which is a straightforward next-word predictor, R1 will not answer immediately – it effectively **pauses to reason** through the problem and only outputs an answer after formulating a solution [datacamp.com [17]]. This makes R1 significantly slower in response time than V3 (often taking minutes for complex queries) [datacamp.com [18], datacamp.com [19]], but the trade-off is substantially improved performance on tasks requiring deep reasoning and planning. For example, DeepSeek reports that R1 matches or exceeds OpenAI's cutting-edge "o1" model on several reasoning-heavy benchmarks [wired.com [20]]. Indeed, R1 was designed as a direct competitor to OpenAI's latest

reasoning model, and it **demonstrably rivals those top-tier models in domains like advanced math, coding challenges, and logic puzzles** [wired.com [21], lawfaremedia.org [22]].

Cost, Performance and Novel Features – The excitement around DeepSeek-V3 and R1 stems from their **unprecedented combination of high performance and low cost**. Technically, these models introduced or refined several novel features: (1) the large-scale MoE architecture (with MLA) which allows *scaling parameters without proportional increases in inference cost*, (2) a multi-token training objective for efficiency, and (3) an integrated chain-of-thought reasoning capability in R1 through reinforcement learning. These innovations translated into **state-of-the-art capability at a fraction of the usual compute budget**. DeepSeek-V3 is *open-source* and its weights are freely available, unlike many Western counterparts [lawfaremedia.org [23]]. This openness and the reported training cost of ~$5.5M for V3 shocked the industry [lawfaremedia.org [24]] – for context, Meta's Llama or OpenAI's GPT series are estimated to cost on the order of tens to hundreds of millions to develop. DeepSeek's public technical report asserts that V3 *outperforms other open models and is on par with leading closed models* in many evaluations [huggingface.co [25]]. R1 further pushes the envelope by solving problems previously considered out of reach without massive "agent" systems. In short, **DeepSeek showed that clever architecture and training optimizations can achieve flagship performance without the enormous infrastructure traditionally deemed necessary** [wired.com [26], wired.com [27]]. This disruptive promise – **cutting AI development costs by an order of magnitude while delivering top performance** – is a key reason these models have generated such intense interest across both industry and academia.

In terms of *efficiency*, DeepSeek-V3's MoE design means it can deliver fast responses using only the needed subset of its network, making it highly efficient for deployment [datacamp.com [28], datacamp.com [29]]. R1, by contrast, sacrifices some speed for reasoning power, but it remains far more efficient than naive approaches to reasoning (which might require external tools or significantly larger models). Both V3 and R1 support very long context lengths (up to 64k tokens of input) [datacamp.com [30]], catering to tasks like lengthy document analysis or multi-turn dialogues – matching or exceeding the context windows of contemporary models (for instance, GPT-4 offers up to 32k tokens). **Compared to existing models**, DeepSeek-V3's closest analogs would be Google's earlier MoE research (e.g. Switch Transformers) and OpenAI's large dense models, but V3 distinguishes itself by *successfully marrying MoE with stability and high accuracy* at scale, something not seen in prior production LLMs. Meanwhile, DeepSeek-R1's closest counterpart is OpenAI's "reasoner" (the o1 model), but unlike OpenAI's closed model, R1's methodology and weights are open. This means researchers can inspect and build on R1, accelerating innovation. The key differentiator is that **DeepSeek delivered an OpenAI-level model through a lean startup approach**, emphasizing algorithmic ingenuity over brute-force expenditure [wired.com [31], wired.com [32]]. In summary, the DeepSeek models' architecture and training approach demonstrate a new paradigm: focusing on *smart design (Mixture-of-Experts, chain-of-thought training)* to compensate for limited hardware, thereby achieving a breakthrough in cost-performance ratio.

2. Comparison with Other Chinese Models

DeepSeek-R1 arrived amid a broader surge of AI model development in China. Around the same time, several Chinese tech companies and research labs introduced their own advanced large language models, many with similar aims of matching OpenAI's latest offerings. Below we compare DeepSeek R1/V3 with some notable Chinese models of late 2024–early 2025 and their intended applications:

- **Alibaba Qwen 2.5-Max** – In January 2025, Alibaba's cloud division rushed out *Qwen 2.5-Max*, an upgraded version of its AI model, claiming it **"outperforms… almost across the board GPT-4o, DeepSeek-V3 and Llama-3.1-405B"** on internal tests [reuters.com [33], reuters.com [34]]. This release was unusually timed during the Chinese New Year holidays, underscoring the pressure DeepSeek's meteoric rise put on incumbents [reuters.com [35]]. Qwen 2.5-Max is positioned as a general-purpose model (like DeepSeek-V3) for tasks from chat to content generation. While Alibaba has released smaller models openly before, it's unclear if Qwen 2.5-Max is open-source; however, Alibaba did feel compelled to **publicly benchmark against DeepSeek's performance**, signaling that DeepSeek set a new bar domestically. Alibaba's model ecosystem (the "Qwen" series) is aimed at enterprise and consumer applications, with multilingual support. The key difference is that Alibaba, as a tech giant, can train huge models, but DeepSeek's **leaner development and dramatically lower pricing forced Alibaba to react competitively** [reuters.com [36], reuters.com [37]].

- **ByteDance Doubao-1.5-Pro** – Just two days after DeepSeek-R1's debut, ByteDance (owner of TikTok) unveiled *Doubao-1.5-pro*, an update to its flagship AI model intended to challenge OpenAI's reasoning model as well [reuters.com [38]]. ByteDance claimed this model **outperformed OpenAI's o1 on a key benchmark (AIME) for complex instruction understanding** [reuters.com [39]]. Like DeepSeek-R1, Doubao-1.5-pro is a *"reasoning model"* designed for complex tasks. It also boasted aggressive pricing on ByteDance's cloud platform: as low as 2 yuan per million tokens for output on the 32k context version [reuters.com [40], reuters.com [41]]. This undercuts even DeepSeek's own low usage fees (DeepSeek-R1's API was offered around 16 yuan per million tokens) [reuters.com [42]]. ByteDance's entry shows the **race to build advanced reasoners** became truly competitive in China – multiple firms were benchmarking against OpenAI's best and against each other. While ByteDance has immense resources, it followed DeepSeek's lead in focusing on efficient, affordable access. One differentiator is that ByteDance's model served their ecosystem (e.g. integration with Volcano Engine cloud and potentially apps like Douyin/TikTok), whereas DeepSeek's models were released openly to the research community.

- **Baidu Ernie Bot (Wenxin)** – Baidu, the search giant, introduced *Ernie Bot* in March 2023 as **China's first equivalent to ChatGPT** [reuters.com [43]]. Although Ernie Bot's initial release disappointed investors (it was a relatively constrained demo) [reuters.com [44], reuters.com [45]], Baidu has continued iterating on it. By late 2024, Ernie's capabilities had improved, but Baidu was suddenly facing a new kind of competition. DeepSeek's open-source V3 model **outperformed Baidu and OpenAI models in some tests despite a smaller budget**, grabbing attention [reuters.com [46]]. Ernie Bot and DeepSeek-V3 serve similar purposes as general conversational AI for Chinese and English tasks, but DeepSeek's approach differed by being open and extremely low-cost. In response to models like DeepSeek-V2 and V3, Baidu (along with others) had to engage in a *price war*, slashing the costs of using their AI services [reuters.com [47]]. Ernie Bot remains a more closed system tied to Baidu's platforms, whereas DeepSeek's models can be freely adapted by anyone.

- **iFlyTek Spark** – iFlyTek, a leading Chinese AI firm, introduced its *Spark Desk* (星火大模型) in mid-2023 as another rival to ChatGPT. In early 2025, iFlyTek announced it had developed a **"reasoning" version of Spark** as well [reuters.com [48]], joining the wave of reasoner models. Spark is known for strong Chinese language capabilities and applications in education and office productivity. Compared to DeepSeek-R1, Spark's reasoner (details of which are limited in media) is likely proprietary and geared toward

domestic use cases (with the usual censorship of sensitive topics). DeepSeek-R1 still stood out for its open release and English proficiency, whereas many Chinese models like Spark primarily target Chinese-language tasks and comply with local content regulations.

- **Other Notable Models/Startups** – The late 2024 period saw a proliferation of Chinese LLM efforts. Startups like **MiniMax** and **Moonshot AI** also unveiled their own reasoning-capable models in the weeks following DeepSeek's release [reuters.com [49]]. MiniMax (a well-funded AI startup) has been working on general AI assistants, and reportedly its model "ABAB" was among those incorporating reasoning. **Zhipu AI**, another startup (spun off from Tsinghua University), earlier launched the **ChatGLM** series (with 130B parameters) and was drawn into the 2024 price war triggered by DeepSeek-V2 [cyber.fsi.stanford.edu [50]]. Zhipu's models are known for bilingual ability (Chinese-English) and were open-sourced as well, aligning with the open model movement DeepSeek champions. Even **Tencent** entered the fray with its "Hunyuan" foundation model (announced in 2023) and likely had to accelerate improvements after DeepSeek's advances [cyber.fsi.stanford.edu [51]]. In summary, **DeepSeek-R1 and V3 emerged not in isolation but as part of a dynamic ecosystem of Chinese AI models**. Many of these models share similar *capabilities* (natural language chat, coding assistance, content generation) and *ambitions* (rivaling US models), but they differ in strategy. DeepSeek's edge has been its openness and ultra-low cost focus, whereas giants like Alibaba and ByteDance initially relied on scale and integration into their own products. Now, however, we see even the big players adopting some of DeepSeek's playbook: open-sourcing more models (Alibaba open-sourced earlier Qwen models) [lawfaremedia.org [52]] and drastically cutting usage fees to stay competitive [reuters.com [53]]. All Chinese models must also operate under the country's AI regulations – for instance, **DeepSeek's AI assistant, like others, avoids certain sensitive political topics by design** [aiswiss.net [54]] to comply with government content rules. Yet, despite such constraints, the competition among these models has clearly accelerated China's AI capability. DeepSeek's rise essentially *raised the performance bar and lowered the price floor*, forcing its peers to quickly evolve their offerings.

3. Observations from the AI Research Community

The global AI research community has been both impressed and curious about DeepSeek R1 (and V3), leading to a flurry of technical evaluations and debates. **Initial Reception**: DeepSeek-R1 immediately became a hot topic among AI researchers worldwide – it was, notably, **the first publicly available model to match the performance of OpenAI's premier "reasoning" model (o1)** [lawfaremedia.org [55]]. This was significant because it showed that an open-source project from China could reach the frontier quality previously thought exclusive to firms like OpenAI, DeepMind, Anthropic, etc. In benchmark tests reported by DeepSeek (and later verified by some independent evaluations), R1 demonstrated **leading performance on complex math, coding, and logic challenges**, on par with or exceeding models like GPT-4 (often referred to as GPT-4o in comparisons) [wired.com [56], reuters.com [57]]. Researchers praised DeepSeek's technical report for its transparency – unlike some big labs, DeepSeek published detailed papers describing their methods [lawfaremedia.org [58]]. The AI community responded with considerable goodwill, especially since DeepSeek **open-sourced the model weights under an MIT license**, allowing anyone to study or deploy the model [lawfaremedia.org [59]]. This openness was highlighted as a positive example; as one expert noted, DeepSeek "pooled collective expertise" through open methods and proved that *cutting-edge models can be built using less resources, showing current norms leave plenty of room for optimization* [wired.com [60],

wired.com [61]]. In other words, researchers saw DeepSeek as validation that innovations in model architecture can yield huge efficiency gains. Many in the community were excited to experiment with R1 themselves, and within days of release, **R1 and its smaller distilled versions were being tested on various open leaderboards** (for tasks like coding, knowledge, and reasoning) where they achieved top rankings. The open-source AI crowd hailed DeepSeek's models as "darlings" of 2024, given their high quality and permissive usage [lawfaremedia.org [62]].

Technical Critiques and Evaluations: Alongside the praise, experts also offered critical analysis of DeepSeek's claims and identified some limitations. A major point of discussion was DeepSeek's **cost and efficiency claims**. The company's assertion that V3's full training cost was only $5.5M (with R1 being an extra fine-tune on top) drew skepticism from some AI researchers [vectara.com [63], cyber.fsi.stanford.edu [64]]. For instance, analysis by third parties suggested that while DeepSeek only counted *marginal GPU rental costs*, the true expenditure including hardware procurement and R&D could be much higher [lawfaremedia.org [65], cyber.fsi.stanford.edu [66]]. One estimate posited that earlier development stages likely used on the order of 50,000 GPUs and "could have cost north of $1 billion" if fully accounted [cyber.fsi.stanford.edu [67]]. DeepSeek's methodology of cost reporting – counting only the final run on rented GPUs – was debated in research forums as potentially understating the investment [cyber.fsi.stanford.edu [68]]. However, even skeptics conceded that **the team's engineering achievements are real**, enabling them to do a final training run in under 3 million GPU-hours, which is impressive regardless [cyber.fsi.stanford.edu [69]]. Another technical evaluation came from Vectara researchers, who examined **hallucination rates** of DeepSeek-R1. In a January 2025 blog analysis, they found that R1, despite its reasoning prowess, actually *hallucinates more frequently than DeepSeek-V3* in summarization tasks [vectara.com [70], vectara.com [71]]. Using an automated fact-checking benchmark, R1's responses contained unsupported statements about 14.3% of the time, versus around 3% for the base V3 model [vectara.com [72]]. This fourfold jump in hallucinations was observed consistently across evaluation methods, indicating that **R1's complex reasoning ability comes at the cost of increased fabrication of details** [vectara.com [73], vectara.com [74]]. Researchers theorize this may be because R1's chain-of-thought process can lead it down incorrect paths with high confidence, whereas V3 sticks more closely to learned knowledge. This finding has been a cautionary note: while R1 is powerful, it might be less reliable for factual queries than its predecessor, meaning more work is needed to align reasoning models to truthfulness.

There have also been discussions about *data sources and originality*. Some analysts noticed peculiar behaviors from R1 – for example, Lawfare noted R1's insistence in certain tests that it was "Microsoft's Copilot", hinting that **DeepSeek's training data likely included outputs or code from OpenAI/Microsoft systems** [cyber.fsi.stanford.edu [75]]. This raised questions about to what extent models like V3 absorbed content generated by Western models (a form of indirect knowledge transfer). However, no evidence of any illicit data use has surfaced; it's more an observation of how pervasive OpenAI's outputs have become in training corpora.

Global Research Community Discussion: Broadly, the release of DeepSeek-R1 sparked a healthy debate about AI development approaches. One camp of researchers sees it as proof that innovation and algorithmic efficiency can offset hardware advantages. They point to DeepSeek as *"a wake-up call"* to focus on smarter training, not just bigger models [cyber.fsi.stanford.edu [76]]. Another camp cautions against hyperbole: they stress that DeepSeek benefited from substantial (if cleverly utilized) compute and that **scaling resources still confers an edge** that shouldn't be discounted [lawfaremedia.org [77], cyber.fsi.stanford.edu [78]]. A Lawfare analysis piece concluded that, contrary to some alarmist takes, R1 does not mean U.S. leadership in AI has evaporated or that "compute doesn't matter" – those were deemed

mistaken impressions [lawfaremedia.org [79]]. Instead, the measured view is that DeepSeek represents an *important complementary path* in AI research. It has essentially confirmed that reinforcement learning-based reasoners (pioneered by OpenAI's o1) can be replicated by others with sufficient expertise [lawfaremedia.org [80], lawfaremedia.org [81]]. It also underscored the value of open science: because DeepSeek released their models openly, academics worldwide could analyze and validate them quickly. This stands in contrast to closed models where independent evaluation is harder. The community generally welcomed this transparency. Some researchers did raise security and safety questions – for instance, is it wise to open-source a model that can write code as effectively as R1 (which could include malware)? A Stanford Cyber Policy review noted one firm found DeepSeek-R1 was **four times more likely to generate insecure code or malware than OpenAI's model** in certain tests [cyber.fsi.stanford.edu [82]], likely because it has fewer safety guardrails. However, as of early 2025, experts assessed that even the most advanced models (OpenAI's o1 or DeepSeek's r1) are *not yet capable of truly dangerous autonomous actions* without human direction [lawfaremedia.org [83]]. Overall, the AI research community's reception of DeepSeek R1 has been a mix of *admiration* (for the technical feat and openness) and *prudent skepticism* (about some of the bolder claims and the potential pitfalls of the model's outputs). This balanced perspective recognizes DeepSeek-R1 as a milestone for open AI research, while also spurring discussions on how to mitigate issues like hallucination and to verify cost claims.

4. Financial and Geopolitical Discussions

The emergence of DeepSeek-V3 and R1 has had ripple effects beyond the lab – **shaking up financial markets and prompting geopolitical debate**. Upon the release of DeepSeek's free AI assistant (powered by V3/R1) in January 2025, global tech stocks saw a sharp selloff. Investors suddenly feared that Big Tech firms, which had been pouring billions into AI, might lose their competitive edge to a lean Chinese upstart. On Monday, Jan 27, U.S. markets reacted with what was described as a record-breaking wipeout of tech equity value [reuters.com [84]]. AI chipmaker Nvidia's stock fell about 17% in one day (erasing nearly $600 billion in market cap) and the shock spread across semiconductor and AI-exposed stocks, contributing to a **$1 trillion drop in tech sector value globally in a single day** [reuters.com [85], reuters.com [86]]. The trigger was widely reported as DeepSeek's new model announcement – a *"low-cost Chinese AI model"* that could threaten the dominance of U.S. rivals [reuters.com [87]]. DeepSeek drew attention by claiming its assistant needed far less data and ran at a fraction of the cost of services like OpenAI's [reuters.com [88]]. This led investors to question the sustainability of the current AI arms race spending: if a startup can achieve GPT-4-level results with millions instead of hundreds of millions, are the huge expenditures by Big Tech justified [aiswiss.net [89]]? Some analysts mused about an AI investment bubble, worrying that sky-high funding rounds might be based on outdated cost assumptions [aiswiss.net [90]].

In the days after, markets stabilized as skepticism set in about DeepSeek's claims (Wall Street wanted proof that the model was truly as cheap and powerful as touted) [reuters.com [91], reuters.com [92]]. Yet, the **financial narrative had shifted** – DeepSeek forced a conversation about efficiency in AI. Notably, Nvidia's rebound following the selloff was aided by a realization that even if models are more efficient, overall demand for AI compute will continue to grow as AI becomes more ubiquitous (an argument invoking Jevons' Paradox that greater efficiency can spur more consumption) [cyber.fsi.stanford.edu [93]]. Still, the immediate economic implication was clear: **companies like OpenAI, Google, Meta may need to justify their spending** if leaner alternatives can achieve similar results [aiswiss.net [94]]. There were even reports of prominent investors, like hedge fund managers, taking DeepSeek's development as a bullish sign for AI in the long run – Steven Cohen remarked that what DeepSeek did "advances the move to AI" and

could ultimately benefit the industry by lowering costs [reuters.com [95], reuters.com [96]]. Meanwhile, Chinese tech companies felt the pressure too. DeepSeek-V2 (an earlier model) had already sparked a domestic **price war in AI services in mid-2024**, when it offered usage at only 1 yuan per million tokens [reuters.com [97]]. DeepSeek's ultra-low pricing forced cloud providers like Alibaba to slash prices by up to 97% on their AI model offerings [reuters.com [98]]. With V3 and R1, DeepSeek again undercut others – R1's usage was priced at just ¥16 per million tokens ($2.20) versus OpenAI o1's ¥438 ($60) for the same volume [reuters.com [99]]. ByteDance then answered with even ¥2 per million for its model [reuters.com [100]]. Such **aggressive pricing is unprecedented** – good for consumers and enterprise adoption, but potentially squeezing profit margins for AI providers. Internationally, this raised strategic questions: will AI become a "race to the bottom" in cost, and how will companies recoup massive R&D investments if the expectation is set for cheap (or open-source) models? DeepSeek's success, backed by a private hedge fund, suggests alternative funding models (it wasn't seeking immediate profit, being funded as a long-term AGI research venture). In contrast, Big Tech firms have shareholders to satisfy. This dynamic – a free or low-cost model eroding the competitive moat of proprietary systems – was extensively discussed in financial media [reuters.com [101], reuters.com [102]]. Some commentators even likened DeepSeek's open-source disruption to the effect of open-source software on expensive enterprise software in earlier eras.

On the **geopolitical stage**, DeepSeek-R1's rise has been seen as a symbolic moment in the U.S.-China tech competition. That a Chinese startup could leapfrog and match American AI capabilities despite heavy U.S. export restrictions was notable. The model was developed *"despite the U.S. curbing chip exports to China three times in three years"*, as one report emphasized [aiswiss.net [103]]. Since 2019–2020, the U.S. government has tightened controls on advanced semiconductors (like Nvidia A100/H100 GPUs) being sold to China, precisely to slow China's progress in training frontier AI models. DeepSeek's accomplishment thus triggered debate about the efficacy of these controls. Policymakers and analysts asked if American export controls had failed or if "large-scale compute *matters at all* anymore" in light of R1[lawfaremedia.org [104]]. The consensus emerging is that **export controls have not been rendered moot, but they are not a silver bullet**. DeepSeek had to use older or restricted chips (it reportedly amassed 10,000 A100 GPUs before bans and used 50,000 slower H800 chips available in China) [cyber.fsi.stanford.edu [105]]. The fact that it still achieved a top-tier model indicates Chinese researchers found ways to mitigate hardware limitations through software ingenuity. U.S. analysts, like those at CSIS, noted that DeepSeek likely invested heavily (perhaps hundreds of millions) in infrastructure to compensate [cyber.fsi.stanford.edu [106], cyber.fsi.stanford.edu [107]], even if the marginal cost was low. In Washington, the political response was measured. President Donald Trump commented that DeepSeek was *"a wake-up call… we need to be laser focused on competing"*, but he also called it "a very positive development" in that it could reduce costs in AI research overall [cyber.fsi.stanford.edu [108]]. This suggests that U.S. leadership saw no immediate crisis, but rather an impetus to support domestic innovation (indeed, the administration had just announced a $500 billion "Stargate" AI infrastructure project with industry partners [cyber.fsi.stanford.edu [109]]). The strategic takeaway is that the U.S. will likely continue restricting hardware exports while also doubling down on its own AI investments, but it now recognizes that **smaller players can make leaps with creativity**, so monitoring global AI progress (and possibly cooperating on certain fronts) is crucial [cyber.fsi.stanford.edu [110]].

For China, DeepSeek's rise is a point of pride and a proof-of-concept. It aligns with the government's push for technological self-reliance amid foreign sanctions. Interestingly, DeepSeek appears to be a *private-sector triumph* rather than a state-directed project – observers noted that unlike some Chinese efforts (e.g. Huawei's chip development), DeepSeek's success does not seem state-planned [cyber.fsi.stanford.edu [111]]. It grew out of a

hedge fund's research arm and remained lean and independent. Chinese media and officials have nonetheless highlighted it as evidence that China's talent and "necessity-driven innovation" can yield world-class AI [wired.com [112]]. At the same time, because R1 is open source, it somewhat blurs national lines – researchers in the West are also benefiting from it. Some in the U.S. initially voiced suspicions: Was DeepSeek some kind of Chinese government or espionage front, given its sudden success? [lawfaremedia.org [113]]

But those theories have not been substantiated; DeepSeek operates transparently like a research lab (its core team is young Chinese PhDs and engineers) [reuters.com [114]]. The model's open availability also undercuts the notion of it being a covert weapon – if anything, it's a public good that everyone can inspect. Geopolitically, however, the *perception* of China catching up in AI is important. R1's debut prompted international media to discuss whether the balance of AI power was shifting. A **New York Times** piece (referenced in Lawfare) noted the broad tech stock selloff and framed R1 as forcing an "AI rethink" in the U.S. [lawfaremedia.org [115]]. Meanwhile, **Reuters** and **Wired** stressed the unintended consequences of the tech Cold War: U.S. chip bans arguably spurred Chinese innovators to find new ways to win, such as DeepSeek's focus on efficiency [wired.com [116]]. This underscores a paradox: while export controls slowed China's access to top chips, they may have catalyzed a breakthrough in software approaches that now challenges Western models on a different level.

Finally, discussions in international media have touched on the **future implications**. Economically, if models like DeepSeek-R1 can be developed cheaply, AI technology could diffuse more rapidly to new entrants and smaller countries, potentially leveling the playing field. Some compare this to how open-source Linux leveled operating systems. This raises questions about how American firms will maintain a competitive edge – likely through proprietary data, integration, and sheer brand/user base, since the algorithms are no longer exclusive. Geopolitically, the narrative of "AI supremacy" is being reevaluated: it's not just about having the most GPUs, but about the smartest use of whatever resources you have. DeepSeek's achievement, as one policy analyst put it, shows that *"necessity is the mother of invention"* – constrained by chip scarcity, Chinese researchers innovated in algorithms. Moving forward, we might see a bifurcation: U.S. companies continuing with massive scale (e.g. trillion-parameter dense models on state-of-the-art chips) and Chinese or open-source communities exploring radical efficiency (MoE, compression, etc.). Each approach could yield advantages, and in fact they might converge (Western firms may adopt some of DeepSeek's techniques to cut costs, while Chinese efforts may get more investment to scale up). In any case, DeepSeek R1 has become a **landmark case in the global AI discourse**, illustrating the interplay of technology, economics, and policy. It has provoked important conversations about the cost of innovation, the value of openness, and the resilience of national strategies in AI – ensuring that the impact of DeepSeek's models will be studied for years to come in both technical and policy circles.

References: DeepSeek-AI (2024), Reuters, Wired, Lawfare, Stanford Cyber Policy Center, CNBC, DataCamp, Vectara, etc. huggingface.co [117] , wired.com [118], reuters.com [119], reuters.com [120], lawfaremedia.org [121], vectara.com [122], reuters.com [123], aiswiss.net [124], cyber.fsi.stanford.edu [125] (Detailed citations are embedded above in the text.)

---

1 https://huggingface.co/papers/2412.19437#:~:text=We%20present%20DeepSeek,Tuning%20and%20Reinforcement

2 https://www.datacamp.com/blog/deepseek-r1-vs-v3

3 https://huggingface.co/papers/2412.19437#:~:text=We%20present%20DeepSeek,Tuning%20and%20Reinforcement

[4]https://huggingface.co/papers/2412.19437#:~:text=671B%20total%20parameters%20with%2037B,source%20models%20and%20achieves
[5]https://huggingface.co/papers/2412.19437#:~:text=strategy%20for%20load%20balancing%20and,any%20irrecoverable%20loss%20spikes%20or
[6] https://huggingface.co/papers/2412.19437#:~:text=reveal%20that%20DeepSeek,V3
[7] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=units%2C%20or%20GPUs%29,it%20remains%20an%20impressive%20figure
[8] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=and%20generating%20synthetic%20data%20,millions%20for%20American%20frontier%20models
[9] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=DeepSeek%20had%20to%20come%20up,"
[10] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=and%20Mixture,the%20research%20institution%20Epoch%20AI
[11] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=DeepSeek%20has%20also%20made%20significant,the%20research%20institution%20Epoch%20AI
[12] https://huggingface.co/papers/2412.19437#:~:text=reveal%20that%20DeepSeek,V3
[13] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=What%20sets%20DeepSeek,and%20develop%20reasoning%20paths%20autonomously
[14] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=One%20difference%20between%20V3%20and,it%20starts%20outputting%20the%20answer
[15] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=What%20sets%20DeepSeek,and%20develop%20reasoning%20paths%20autonomously
[16] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=One%20difference%20between%20V3%20and,it%20starts%20outputting%20the%20answer
[17] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=One%20difference%20between%20V3%20and,it%20starts%20outputting%20the%20answer
[18] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=,structure
[19] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=DeepSeek,interactions%20where%20speed%20is%20crucial
[20] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=On%20January%2020%2C%20DeepSeek%2C%20a,a%20run%20for%20their%20money
[21] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=On%20January%2020%2C%20DeepSeek%2C%20a,a%20run%20for%20their%20money
[22] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=language%20model%20called%20r1%2C%20and,Anthropic%2C%20Google%E2%80%99s%20DeepMind%2C%20and%20Meta
[23] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=prowess%2C%20r1%20is%20notable%20for,regarded%20models%20as%20open
[24] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=and%20generating%20synthetic%20data%20,millions%20for%20American%20frontier%20models
[25]https://huggingface.co/papers/2412.19437#:~:text=strategy%20for%20load%20balancing%20and,any%20irrecoverable%20loss%20spikes%20or
[26] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=and%20Mixture,the%20research%20institution%20Epoch%20AI
[27] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=AI%20companies%2C%20developing%20open%20source,%E2%80%9D
[28] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=One%20of%20DeepSeek,given%20task%2C%20saving%20on%20computational
[29] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=match%20at%20L250%20DeepSeek,interactions%20where%20speed%20is%20crucial
[30] https://www.datacamp.com/blog/deepseek-r1-vs-v3#:~:text=,speed%20comparison
[31] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=DeepSeek%E2%80%99s%20success%20points%20to%20an,using%20limited%20resources%20more%20efficiently
[32] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=DeepSeek%20had%20to%20come%20up,%E2%80%9D
[33] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=BEIJING%2C%20Jan%2029%20%28Reuters%29%20,V3
[34] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=%22Qwen%202.5,source%20AI%20models
[35] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=The%20unusual%20timing%20of%20the,but%20also%20its%20domestic%20competition

[36] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=The%20predecessor%20of%20DeepSeek%27s%20V3,it%20was%20released%20last%20May
[37] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=While%20large%20Chinese%20tech%20companies,students%20from%20top%20Chinese%20universities
[38] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=The%20company%20released%20Doubao,and%20respond%20to%20complex%20instructions
[39] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=The%20company%20released%20Doubao,and%20respond%20to%20complex%20instructions
[40] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=at%2016%20yuan%20%28%242,yuan%20for%20the%20same%20usage
[41] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=ByteDance%27s%20pricing%20is%20even%20more,ByteDance%27s%20cloud%20platform%20Volcano%20Engine
[42] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=DeepSeek%20proposed%20a%20cut,yuan%20for%20the%20same%20usage
[43] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=Other%20Chinese%20tech%20companies%20followed,HK%29%20%2C%20opens%20new%20tab
[44] https://www.reuters.com/technology/chinese-search-giant-baidu-introduces-ernie-bot-2023-03-16/#:~:text=URL%3A%20https%3A%2F%2Fwww.reuters.com%2Ftechnology%2Fchinese,main%20content
[45] https://www.reuters.com/technology/chinese-search-giant-baidu-introduces-ernie-bot-2023-03-16/#:~:text=
[46] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=DeepSeek%20drew%20widespread%20attention%20in,charge%20users%20a%20lot%20less
[47] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=The%20predecessor%20of%20DeepSeek%27s%20V3,it%20was%20released%20last%20May
[48] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=Other%20Chinese%20firms%20that%20have,Moonshot%20AI%2C%20Minimax%20and%20iFlyTek
[49] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=both%20performance%20metrics%20and%20fees,charged%20to%20users
[50] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=What%20makes%20DeepSeek%20particularly%20interesting,backed%2C%20it%20seems
[51] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=What%20makes%20DeepSeek%20particularly%20interesting,backed%2C%20it%20seems
[52] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=prowess%2C%20r1%20is%20notable%20for,regarded%20models%20as%20open
[53] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=DeepSeek%20proposed%20a%20cut,yuan%20for%20the%20same%20usage
[54] https://www.aiswiss.net/chinas-deepseek-ai-dethrones-chatgpt-on-app-store-heres-what-you-should-know/#:~:text=industry%20leaderboards%2C%20with%20users%20praising,of
[55] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=language%20model%20called%20r1%2C%20and,Anthropic%2C%20Google%E2%80%99s%20DeepMind%2C%20and%20Meta
[56] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=On%20January%2020%2C%20DeepSeek%2C%20a,a%20run%20for%20their%20money
[57] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=DeepSeek%20drew%20widespread%20attention%20in,charge%20users%20a%20lot%20less
[58] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=DeepSeek%E2%80%99s%20research%20papers%20and%20models,eyebrows%20with%20its%20exceptionally%20low
[59] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=prowess%2C%20r1%20is%20notable%20for,regarded%20models%20as%20open
[60] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=and%20Mixture,the%20research%20institution%20Epoch%20AI

[61] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=AI%20companies%2C%20developing%20open%20source,%E2%80%9D
[62] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=DeepSeek%E2%80%99s%20research%20papers%20and%20models,eyebrows%20with%20its%20exceptionally%20low
[63] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3#:~:text=According%20to%20Deepseek%2C%20the%20model,Deepseek%20with%20an%20MIT%20license
[64] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=total%20training%20cost%20could%20be,likely%20to%20have%20been%20used
[65] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=units%2C%20or%20GPUs%29,it%20remains%20an%20impressive%20figure
[66] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=total%20training%20cost%20could%20be,likely%20to%20have%20been%20used
[67] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=and%20experimentation,likely%20to%20have%20been%20used
[68] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=AI%20company%20familiar%20with%20training,likely%20to%20have%20been%20used
[69] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=While%20there%20is%20no%20current,with%20imprecise%20cost%20saving%20reporting
[70] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3#:~:text=part%20of%20our%20work%20to,it%20to%20Vectara%E2%80%99s%20HHEM%20leaderboard
[71] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3#:~:text=The%20results%20were%20surprising%3A%20Deepseek,V3
[72] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3#:~:text=DeepSeek%20R1%20DeepSeek%20V3%20Vectara%E2%80%99s,Pro%203.89%252.69
[73] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3#:~:text=DeepSeek%20R1%20DeepSeek%20V3%20Vectara%E2%80%99s,Pro%203.89%252.69
[74] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3#:~:text=Thus%20our%20surprise%3A%20consistently%20across,V3
[75] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=architecture,absorbed%20from%20OpenAI%20during%20training
[76] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=Back%20in%20the%20U,%E2%80%9D
[77] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=that%20Chinese%20AI%20companies%20would,chipmaker%20Nvidia%E2%80%99s%20stock%20falling%2017
[78] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=What%20will%20be%20the%20policy,chip%20export%20restrictions%20to%20China
[79] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=that%20Chinese%20AI%20companies%20would,chipmaker%20Nvidia%E2%80%99s%20stock%20falling%2017
[80] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=But%20the%20model%20that%20truly,however%2C%20was%20a%20mistaken%20assumption
[81] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=But%20the%20model%20that%20truly,however%2C%20was%20a%20mistaken%20assumption
[82] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=Center%20cyber,insecure%20code%20than%20OpenAI%27s
[83] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=anyone%20in%20the%20world%20to,regarded%20models%20as%20open
[84] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=day%27s%20record,rivals
[85] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=day%27s%20record,rivals
[86] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=On%20Monday%2C%20Nvidia%20lost%20about,shed%20more%20than%20%241%20trillion
[87] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=day%27s%20record,rivals
[88] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=Monday%27s%20selloff%2C%20which%20hammered%20many,lingered%20over%20its%20cost%20claims
[89] https://www.aiswiss.net/chinas-deepseek-ai-dethrones-chatgpt-on-app-store-heres-what-you-should-know/#:~:text=agree%20that%20the%20model%20was,on%20AIPerplexity%20AI%20revises%20Tiktok

[90] https://www.aiswiss.net/chinas-deepseek-ai-dethrones-chatgpt-on-app-store-heres-what-you-should-know/#:~:text=agree%20that%20the%20model%20was,on%20AIPerplexity%20AI%20revises%20Tiktok
[91] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=Monday%27s%20selloff%2C%20which%20hammered%20many,lingered%20over%20its%20cost%20claims
[92] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=,of%20tastytrade%20brokerage%20in%20Chicago
[93] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=In%20the%20long%20run%2C%20once,resources%2C%20from%20hardware%20to%20energy
[94] https://www.aiswiss.net/chinas-deepseek-ai-dethrones-chatgpt-on-app-store-heres-what-you-should-know/#:~:text=agree%20that%20the%20model%20was,on%20AIPerplexity%20AI%20revises%20Tiktok
[95] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=Point72%20Asset%20Management%27s%20founder%2C%20Steven,the%20move%20to%20artificial%20intelligence
[96] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/#:~:text=,of%20tastytrade%20brokerage%20in%20Chicago
[97] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=The%20predecessor%20of%20DeepSeek%27s%20V3,it%20was%20released%20last%20May
[98] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=The%20predecessor%20of%20DeepSeek%27s%20V3,it%20was%20released%20last%20May
[99] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=DeepSeek%20proposed%20a%20cut,yuan%20for%20the%20same%20usage
[100] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/#:~:text=at%2016%20yuan%20%28%242,yuan%20for%20the%20same%20usage
[101] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=The%20Jan,firms%20in%20the%20United%20States
[102] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=Valley%20and%20caused%20tech%20shares,firms%20in%20the%20United%20States
[103] https://www.aiswiss.net/chinas-deepseek-ai-dethrones-chatgpt-on-app-store-heres-what-you-should-know/#:~:text=lines%20of%20questioning,AI%20sector%20is%20awash%20with
[104] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=that%20Chinese%20AI%20companies%20would,chipmaker%20Nvidia%E2%80%99s%20stock%20falling%2017
[105] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=Based%20on%20reports%20from%20the,disclose%20its%20holding%20of%20any
[106] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=total%20training%20cost%20could%20be,likely%20to%20have%20been%20used
[107] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=Understandably%2C%20with%20the%20scant%20information,chips%2C%20and%20other%20infrastructure%20costs
[108] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=Back%20in%20the%20U,%E2%80%9D
[109] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=There%20is%20good%20reason%20for,of%20the%20different%20development%20approaches
[110] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=What%20will%20be%20the%20policy,chip%20export%20restrictions%20to%20China
[111] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock#:~:text=off%20a%20price%20war%20with,planned
[112] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=DeepSeek%E2%80%99s%20success%20points%20to%20an,using%20limited%20resources%20more%20efficiently
[113] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=that%20Chinese%20AI%20companies%20would,chipmaker%20Nvidia%E2%80%99s%20stock%20falling%2017
[114] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/#:~:text=While%20large%20Chinese%20tech%20companies,students%20from%20top%20Chinese%20universities

[115] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t#:~:text=that%20Chinese%20AI%20companies%20would,chipmaker%20Nvidia%E2%80%99s%20stock%20falling%2017
[116] https://www.wired.com/story/deepseek-china-model-ai/#:~:text=DeepSeek%E2%80%99s%20success%20points%20to%20an,using%20limited%20resources%20more%20efficiently
[117] https://huggingface.co/papers/2412.19437
[118] https://www.wired.com/story/deepseek-china-model-ai/
[119] https://www.reuters.com/technology/artificial-intelligence/alibaba-releases-ai-model-it-claims-surpasses-deepseek-v3-2025-01-29/
[120] https://www.reuters.com/technology/artificial-intelligence/tiktok-owner-bytedance-deepseek-lead-chinese-push-ai-reasoning-2025-01-22/
[121] https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t
[122] https://www.vectara.com/blog/deepseek-r1-hallucinates-more-than-deepseek-v3
[123] https://www.reuters.com/technology/tech-stock-selloff-deepens-deepseek-triggers-ai-rethink-2025-01-28/
[124] https://www.aiswiss.net/chinas-deepseek-ai-dethrones-chatgpt-on-app-store-heres-what-you-should-know/
[125] https://cyber.fsi.stanford.edu/publication/taking-stock-deepseek-shock