CENTER FOR AI POLICY

US Open-Source AI Governance

Balancing Ideological and Geopolitical Considerations with China Competition

Claudia Wilson - Center for AI Policy
Emmie Hine - Digital Ethics Center, Yale University

February 2025

About us

Center for AI Policy

The Center for AI Policy's mission is to ensure safer AI now and going forwards. We are a nonpartisan research organization dedicated to mitigating the catastrophic risks of AI through policy development and advocacy. We're working with Congress and federal agencies to help them understand advanced AI development and effectively prepare for it. We share policy proposals, draft model legislation, and give feedback on others' policies.

Yale Digital Ethics Center

At the Yale Digital Ethics Center (DEC), we research the governance, ethical, legal, and social implications (GELSI) of digital innovation and technologies and their human, societal, and environmental impact. Through our work, we seek to design a better information society: critical, equitable, just, open, pluralistic, sustainable, and tolerant. We aim to identify and enhance the benefits of digital innovation and technologies while mitigating their risks and shortcomings.

Executive Summary

The proliferation of open-source artificial intelligence (AI) has triggered a contentious policy debate. Should open-source AI be considered for regulation as closed models have been? Two prevailing perspectives have emerged: one that focuses on geopolitical risk, particularly with respect to US-China competition, and one that is grounded in ideological values around open-source technology, such as innovation, transparency, and democracy. The former is broadly supportive of export controls and other regulations, while the latter opposes restrictions on open-source technology. While neither framing should be taken at face value, they do reflect legitimate tensions between promoting technological advancement and maintaining strategic advantage in an interconnected world.

Through its work with Congress, the Center for AI Policy (CAIP) has identified that US policymakers are grappling with how to reconcile these two perspectives, particularly in light of the highly advanced models released by Chinese startup DeepSeek in December 2024 and January 2025. Much public commentary takes a single perspective, perhaps with a throwaway comment acknowledging the other perspective, but there has been no attempt to consider both perspectives in a structured manner. This paper combines both perspectives into a single rubric with which to assess open-source AI policies. It then uses this rubric to analyze four different open-source AI policy proposals.

A rubric for open-source AI policies

Our rubric combines three ideological considerations and three geopolitical considerations. The three ideological considerations, as identified by existing literature, are *increased transparency*, *accelerated technological progress*, and *increased power distribution*. The three geopolitical considerations are *Chinese misuse of American open-source AI*, *backdoor risks* from the use of Chinese open-source AI, and *changes in global power dynamics* depending on which country dominates in open-source AI.

There are important nuances to these considerations.

- Open-source AI has been more relevant to certain forms of technological progress, such as specific-use applications, than other forms, such as frontier capabilities.
- Transparency, which is helpful for safety of models, can feasibly be achieved through external model audits rather than exclusively through open-sourcing.
- Misuse risk is heavily informed by marginal risk, which depends on whether Chinese actors already have access to equivalent capabilities. If proprietary models can easily be distilled to advance capabilities, as DeepSeek is suspected of doing, then open-sourcing models may have lower marginal risk.

- Even if there is limited marginal risk of *Chinese* misuse, there may still be significant marginal risk of misuse by other actors.
- Preoccupation with global power dynamics creates the risk of an unfettered AI arms race, which eschews the strategic value of developing safe and reliable models.
- There are tensions between the different geopolitical considerations: attempts to limit Chinese misuse of American open-source AI could undermine attempts to achieve US dominance in the global open-source landscape.

Context on Chinese and American open-source AI

We summarize the context that is relevant to geopolitical considerations.

- **Open models have been lagging closed model capabilities by roughly one year**. However, that gap has recently narrowed.
- **The Chinese open source community has exhibited a dependence on Llama** models to advance capabilities**,** citing access to computing power ("compute") as a bottleneck. This is at odds with the government's aspiration of "digital sovereignty."
- Regardless of dependence, **Chinese open models have demonstrated impressive capabilities**, sometimes exceeding the performance of American models.
- **DeepSeek's December 2024 and January 2025 models** demonstrated novel **algorithmic innovations** and **impressive performance**.
- Yet **these models do not necessarily signal a paradigm shift** in the performance of open models. They were inspired by closed models, and their breakthroughs may be further leveraged by closed models and augmented with compute.
- **Chinese open models were gaining in popularity even prior to DeepSeek.** In June 2024, there were eleven times more models derived from Llama than from Qwen. By December 2024, Llama had only 25% more derivative models.

Policy Analysis

We use this rubric to assess four policies: two seeking to address Chinese misuse of American open-source AI, and two seeking to address potential "backdoors" when using foreign open-source AI. The first policy is expansive export controls on open model components to China. The second policy is industry-led assessment of whether individual models should be made open-source, coupled with independent audits of those assessments. The third policy requires providers of government AI services and products to audit any model components that have been based upon external open-source model components. The fourth policy is government funding for an open-source repository of audits on commonly used open-source models and frameworks.

We find that blanket export controls on all open-source AI models would likely be sub-optimal and counterproductive. Requiring every user of every open model to undergo a know-your-customer (KYC) process would be highly disruptive to the development of specific-use applications, though it would have limited impact on frontier capabilities. It would also likely have limited efficacy in mitigating misuse risks by China**.** Furthermore, this policy would leave domestic misuse of open-source AI entirely unaddressed. There is also a genuine risk that export controls undermine US global power by introducing friction for other countries to use American technology. Given that the marginal risk of open-source AI is unclear, it may not be worth pursuing such a disruptive policy today.

A more reasonable alternative would be to require developers of foundation models to conduct a risk assessment of each model they intend to make open source. Developers could document their investigations and provide rationale for a decision around how they have released their model (e.g., all model components available without checks; requiring academics to provide an institutional email address to access full model components for chem-bio models). Like Meta's Frontier AI Framework (announced in January 2025), it would entail a structured examination of risk, but it would also be accompanied by independent assurance on risk assessments. A model-by-model approach, rather than blanket legislation, will likely be less disruptive to technological progress and could be more effective at mitigating risk. Furthermore, this policy would also mitigate misuse by domestic actors, unlike export controls, which focus exclusively on specific nation-states.

Regarding the "backdoor" risks of open-source AI, we find that **audits of government products** leveraging open-source AI could be a helpful mitigation. However, the information needed to conduct an audit could be unavailable, and in any case audit results would not typically be shared with the public. An alternative would be to create a public repository of audits of popular open-source AI models and frameworks. Such a resource would be a valuable public good and could create greater trust in open-source AI. Yet its success, as with government audits, depends on the ability to trace model components and the resources to conduct those audits.

Open-source AI is a dynamic technology and the policy space is nascent. **Going forward, policymakers should continue to monitor the performance of open models relative to closed models, as well as where algorithmic innovations are originating** to inform their assessment of marginal risk**.** For similar reasons, it would also be valuable to continue monitoring relative performance of Chinese and American models and develop more comprehensive comparative benchmarks. Future research should also include deeper investigation into open-source AI safety risks posed by non-state actors.

Contents

Introduction

The proliferation of open-source artificial intelligence (AI) has emerged as a critical national security concern. As advanced AI models—including "foundation models"—become freely available, governments face mounting challenges from their potential misuse in cyberattacks, bioweapons development, and military intelligence.[1] These risks are particularly salient in the context of United States (US)-China technological competition, where open-source AI sits at the intersection of innovation policy and national security strategy.

On top of national security concerns, the economic stakes are substantial; open-source software's demand-side value is estimated at $8.8 trillion,[2] and AI development frameworks and models are an increasingly important part of this. The intersecting security and economic implications—to say nothing of the other safety and ethical issues at play—make the regulation of open-source AI regulation a complex but vital question.

This report is structured as follows. Section 1 provides an overview of open-source AI, including a comparison of performance between open and closed models. Section 2 describes the interaction between the American and Chinese open-source AI landscapes and compares the performance and popularity of their models. Section 3 summarizes the open-source AI regulations of the US, California, the European Union (EU), and the United Kingdom (UK). Section 4 articulates the six considerations of our open-source AI policy rubric. Section 5 uses the rubric to assess four open-source AI policies. Section 6 discusses the limitations of this research and potential further work.

1. What is open-source AI?

1a. Definitions

One challenge when discussing open-source AI is the lack of consensus over its definition. The broader concept of "open source" has been defined by the Open Source Initiative (OSI) since 2006 and is widely accepted as the de facto definition of open source.[3] Their definition of what makes software "open-source" includes the source code being available with a license allowing for free distribution, regardless of who is using it and for what purpose.[4]

Open-source AI presents new definitional challenges because the ability to download, modify, and reproduce AI models is reliant on more than just source code.[1] Unlike other open-source software, it also requires access to training data and model weights, among other structural elements. Open-source AI has been defined by different stakeholders using different combinations of its characteristics, availability of specific model components, and proprietary features. This lack of definitional clarity is further exacerbated by the phenomenon of "open-washing," where companies frame their products as "open" to benefit from the positive connotations of "openness."[5,6]

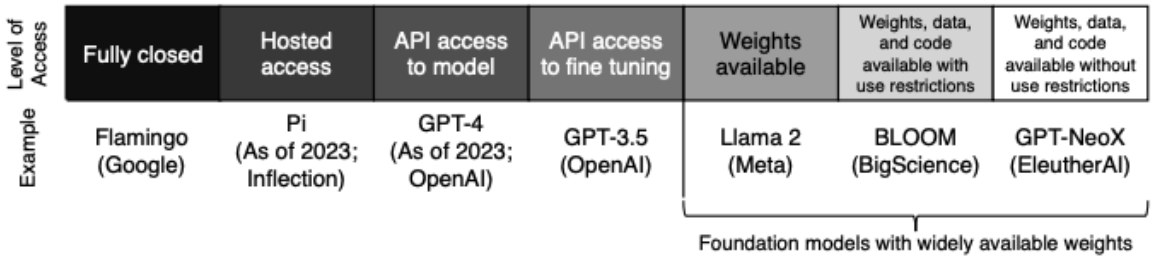
In 2024, the OSI consulted with a group of stakeholders to create a definition of "open-source AI." Under the OSI's definition,[7] for an AI system to be open-source, it must grant the freedom to a) use it for any purpose without permission, b) study how it works, c) modify it for any purpose, and d) share it with others. They also specify that to enable those freedoms, a provider must make available detailed information about training data, complete source code, and model parameters (e.g., weights and other configuration settings).

Despite engagement with a 70-person group of researchers, lawyers, industry experts, the OSI definition is contentious.[3] By this definition and as explicitly mentioned by the OSI,[3] Meta's Llama model would not be open-source because its training data and complete source code are not freely available.[8]

At a baseline level, open-source AI models provide greater and more liberal access to model components than closed models. It is far easier to build upon an open-source model compared to a closed-source model. One definition of open-source focuses on the *characteristics* associated with "openness"—transparency, reusability, and extensibility.[5] Although these characteristics provide color as to the purpose of open models, they lack sufficient specificity for easy categorization.

In practice, many in industry view openness as a spectrum depending on which specific *model components* are available. As seen in Figure 1 (republished with permission from the Stanford Institute for Human-Centered Artificial Intelligence), at minimum, the weights must be available for a model to be "open," but a model with unrestricted access to weights, data, and code would be "more open."[9]

Figure 1: Spectrum of open model components from Stanford Institute for Human-Centered Artificial Intelligence (HAI)[9]



However, definitions that focus exclusively on model components risk neglecting how proprietary features may influence the "open" characteristics of a model, particularly with respect to extensibility. For example, Llama 2's community agreement bans its use to train other language models and requires a special license if it is used in an app with more than 700 million monthly users.[10] This example is not intended to criticize Llama 2, but rather to highlight that it would be "less open" than a provider that freely allows any use of its model.

For the purposes of this paper, we will refer to models that make at least their model weights available as "**open**." When discussing the concept more generally, we will fall back on **open-source**.

1b. Open-source AI models versus frameworks

When discussing open-source AI, it is crucial to distinguish between the models themselves, like the aforementioned Llama 2, and the technical frameworks used to develop those models, like PyTorch. While a model refers to a specific algorithm trained on data to perform certain tasks, frameworks refer to a suite of tools that developers can use to build and train an AI model.[11,12] Both AI models and frameworks can be open-source, but a model built on an open-source framework does not itself have to be open-sourced.

1c. Differences between open and closed models

Due to the availability of model components, open models offer the ability to make more substantial changes to existing models than closed models do. Closed models are accessed via application programming interfaces (APIs) and wrappers like websites and apps and are less modifiable. There are similar APIs and wrappers for some open models, but they can also be directly downloaded and locally run. Open models can be modified and fine-tuned directly. While open models are free-of-charge, users still need to pay the cost of compute for fine-tuning and use, which can in some cases exceed the cost of closed-source models.[13] Meta's Llama 2, an open model, reportedly costs companies 50% to 100% more than OpenAI's GPT 3.5 Turbo.[13]

This reflects the different business models of open- and closed-source AI companies. Companies with closed models often have free versions and paid subscriptions, where a subscription unlocks more advanced models and capabilities.[14] Although these companies also often charge more for corporate subscriptions, they are yet to achieve profitability.[15] Unlike closed models, open models are offered free-of-charge, so for-profit companies need alternate revenue sources such as complementary proprietary products or hosted API services.[16] Some open companies charge businesses to build ready-to-use apps on top of these models, such as Mistral's partnership with Microsoft.[17] Similarly, Stability AI has a subscription fee for commercial use of certain models.[17] Likewise, Meta does not charge for access to its Llama models, but it may be earning money through partnerships with Microsoft and Amazon.[18] Companies also choose to release—or not—specific model components, with several companies refusing to release datasets for competitive purposes.[19]

Closed models currently perform better than most open models, but the gap may be narrowing.[20,21] As of 2024, the median performance advantage of closed models over open models was 24%, according to the Stanford Institute for Human-Centered Artificial Intelligence.[22,23] EpochAI also found that open models are reaching what were considered frontier capabilities one year later than closed-source models.[24] They found that open models lag by five months on the GPQA benchmark and by 16-25 months on the MMLU benchmark.[24]

The January 2025 release of Chinese open model DeepSeek-R1, which outperformed OpenAI's reasoning model o1 on several metrics, raises the possibility that open models could match the performance of closed models sooner than expected. However, there are several reasons that DeepSeek may not necessarily represent a paradigm shift. First, DeepSeek's uplift in capabilities was due to an algorithmic breakthrough (their combination of reinforcement learning with supervised fine-tuning),[25] and algorithmic breakthroughs

are inherently intermittent.[26] Thus, one breakthrough does not necessarily herald a continuing trend. Second, DeepSeek may have leveraged the progress made by closed models. DeepSeek-R1 was released several months after OpenAI debuted its o1 model.[27] It is likely that DeepSeek chose to focus on reasoning, because it had been "validated as effective" by OpenAI.[27] Further, OpenAI has said that it is "reviewing indications that DeepSeek may have inappropriately distilled" its models to obtain training data.[28] If this is true, then DeepSeek's algorithmic breakthroughs cannot be attributed to open models alone. Third, these algorithmic breakthroughs, now that they have been open-sourced, may be leveraged by companies building closed models, and if combined with larger amounts of compute, could lead to greater capabilities in closed models.

These differences between open and closed models take on particular significance in the context of US-China competition, where open-source AI has become a key battleground for technological leadership. Understanding how these two nations approach open-source AI development reveals both the opportunities and risks this technology presents, and is the topic of the next section.

2. American and Chinese open-source AI ecosystems

This section provides an overview of the open-source AI ecosystems in both the US and China and describes the interaction between these two ecosystems. Our intent is to emphasize that open-source AI does not align with strict national borders. Rather, it is a global ecosystem with many both contributing to and utilizing the innovation that is openly available. Compared to other goods and services, national contributions to open-source AI are intermingled, since the market is essentially freely available information. For the sake of scope, this section focuses exclusively on the Chinese and American components of the global open-source AI ecosystem.

2a. The US's open source history

The open-source movement predates AI. It began in the US as a wave of software innovation and evolved into a distinct community. In the nascent technology industry of the 1950s and 1960s, computer software was commonly co-developed by corporate researchers, including from IBM, and academics, and distributed for free.[29] However, as computers became widespread, a new business model relying on the copyrightability of software (established in a 1974 Third Circuit ruling) emerged.[30] Companies began distributing their software in machine code or binary, rather than human-readable source code.[29] In 1985, computer science researcher Richard Stallman founded the Free Software Foundation (FSF) to promote the development of free software and develop licenses for it.[29] Definitional debates around the meaning of "free"—in the sense of monetary or liberty—and the desire to distinguish "pragmatic, business-case grounds" from the FSF's "philosophically- and politically-focused label" led to the creation of the term "open source" and the launch of the OSI in 1998.[31] The open source movement also has links to the hacker culture of the early computing age in the US, which valued open sharing.[32] This cultural foundation is likely one of the reasons why the open-source ethos became so entrenched in the US.

Although the US's tech industry is primarily driven by tech companies developing proprietary software, open-source software plays a significant role in American innovation and corporations are active participants in its development. For example, the two most popular AI frameworks, PyTorch and TensorFlow, were developed by Meta and Google, respectively. While PyTorch is now under the auspices of the open-source software non-profit Linux Foundation, its lead maintainer is a Meta employee.[33] Google helps maintain TensorFlow under its Google Open Source program.[34] This pattern of corporate involvement in open source extends beyond AI frameworks: Microsoft owns GitHub, the

largest platform for sharing open-source code, and is a significant contributor to Linux kernel development,[35] while IBM acquired Red Hat, a company built entirely around open-source software, for $34 billion.[36] These investments reflect how major tech companies have found ways to build profitable business models around open-source technologies, whether through offering enterprise support, cloud services, or using the software to power their own products and services.

American companies have been active in developing open AI models, although many of the most advanced models are still proprietary. EleutherAI released GPT-J's weights in 2021,[37] and the BigScience research collaboration, organized in part by HuggingFace, released BLOOM, a fully open-source model, in 2022.[38] Big Tech followed shortly thereafter. Meta released the first version of Llama, an open-weight model, in February of 2023,[39] and Google followed with open-weight Gemma in February of 2024.[40]

2b. China's strategic shift to open-source AI

China does not have as long a history of open-source development as the US, with its period of active open-source development beginning in the early 2000s to mitigate reliance on US proprietary software.[41] Since then, China's open-source ecosystem has continued to develop. Until 2021, China was the second-highest ranked country in open-source GitHub contributions; although India has now eclipsed it, China still ranks third.[42] In recent years, Chinese companies and research institutions have been both producing and leveraging open-source AI. Alibaba has made many of its Qwen models open, while the state-sponsored Beijing Academy of Artificial Intelligence is also developing open models.[43] Recently, start-ups like DeepSeek and 01.AI have made previously closed models open.[44] As typical in the open source community, some of these models are based on other open models, such as Meta's Llama 2 and Llama 3.[41,45] However, recent releases have relied less on US open models. DeepSeek-V3, released in December 2024, was developed independently—at a much lower training cost than Llama (although DeepSeek's hardware expenditure was significant)[46]—and outperformed many US open and proprietary models on launch.[47]

A combination of top-down government strategy and organic business incentives has fueled China's strategic shift towards open-source AI. From the government's perspective, leveraging open-source AI is one way to advance indigenous innovation.[48,49] Even prior to the introduction of US export controls on semiconductor chips, the government has consistently identified open-source AI as critical to China's AI ambitions. China's primary AI strategy document, the *New Generation Artificial Intelligence Development Plan* (2017), committed to "encourage AI enterprises and research institutions to build open AI platforms for public open AI research and development."[50] Later in 2021, the *14th Five Year*

Plan, a whole-of-government strategy document, expressed the intent to "support the development of … digital technology open source communities."[51] A 2022 white paper by a state-affiliated think tank later wrote: "Open source is essentially an aggregation of talents and wisdom that can promote the rapid upgrading of AI frameworks."[52] Open source collaboration, which usually consists of a distributed, transnational community of developers, seems at odds with China's historic preference for centralized control,[53] but it is playing an increasingly important role in the Chinese technology sphere.

In recent years, the government has focused on more specific actions to support the open-source AI ecosystem. At China's 2024 annual legislative meeting, the Two Sessions, proposals by representatives for the Government's "AI+ initiative" included establishing multiple national-level open models and open sharing of high-quality training data.[45,54] In the same month, Shanghai city officials spoke on the need to develop national open source foundations.[45,55] Separately, at an Open Atom Foundation event, the Ministry of Industry and Information Technology (MIIT) expressed its support of the Foundation, while state media reported that the Ministry would strengthen nation-wide open-source AI organizations and support infrastructure construction for other open-source AI organizations.[56]

Outside of government imperatives, Chinese companies also have organic business incentives to pursue open-source AI. For companies that provide cloud services, such as Alibaba, their open models serve as a quasi-loss-leader, driving additional growth into their cloud services business.[44] Furthermore, American export controls have limited China's access to compute and thus incentivize open-source AI because building on an existing open model is less compute-intensive than training a model from scratch.[44]

2c. Interaction between Chinese and American open-source AI ecosystems

The open source landscape does not adhere to national boundaries because anyone can contribute to projects. As such, it is a global community. However, because many open AI models originate from large companies, they do have a clear national origin. Despite current US-China tensions and policies that attempt to bifurcate their technology supply chains, both American and Chinese researchers have used open models originating in the other country.

Chinese use of American open models

The Chinese AI ecosystem has relied heavily upon foreign—often American— and frameworks. Unlike foreign closed large language models (LLMs), which have yet to receive

license approvals by the Chinese Cyberspace Administration of China (CAC),[57–59] Chinese companies and individuals can currently access Western open models.[45,60] Indeed, there is an openly-acknowledged tendency for Chinese developers to rely on Meta's Llama models for capability improvements.[61] Both American and Chinese news outlets have reported on this dependence, with Chinese coverage citing the Beijing Academy of Artificial Intelligence (BigAI) and a presentation to Premier Li Qiang that highlighted the prevalence of Llama as a foundation for Chinese open models.[62–64] When interviewed about the Chinese AI ecosystem, DeepSeek's CEO also quipped that "most Chinese companies are used to following, not innovating."[65]

There are both commercial and military examples of Chinese developers leveraging American open models. One high-profile example of Chinese commercial use of American open models is 01.AI (另一万物), a Chinese start-up valued at over $1 billion.[66] Their flagship model, Yi-34B which debuted at the top of the Hugging Face leaderboard for open models in November 2023, built heavily upon Llama 2's framework.[66] In June 2024, a paper by six Chinese researchers detailed how they leveraged Llama to create an LLM that could gather and synthesize intelligence.[67] Two of the researchers were associated with the Academy of Military Science, a People's Liberation Army (PLA) research organization.[67] Although this use violated Meta's license, which at the time prohibited all military use of their products—now amended to exclude US government departments[68]—the license is essentially unenforceable since model components are freely available on Hugging Face. This garnered extensive US news coverage as one of the first confirmed examples of American open-source AI being used by the Chinese military.

American use of Chinese open models

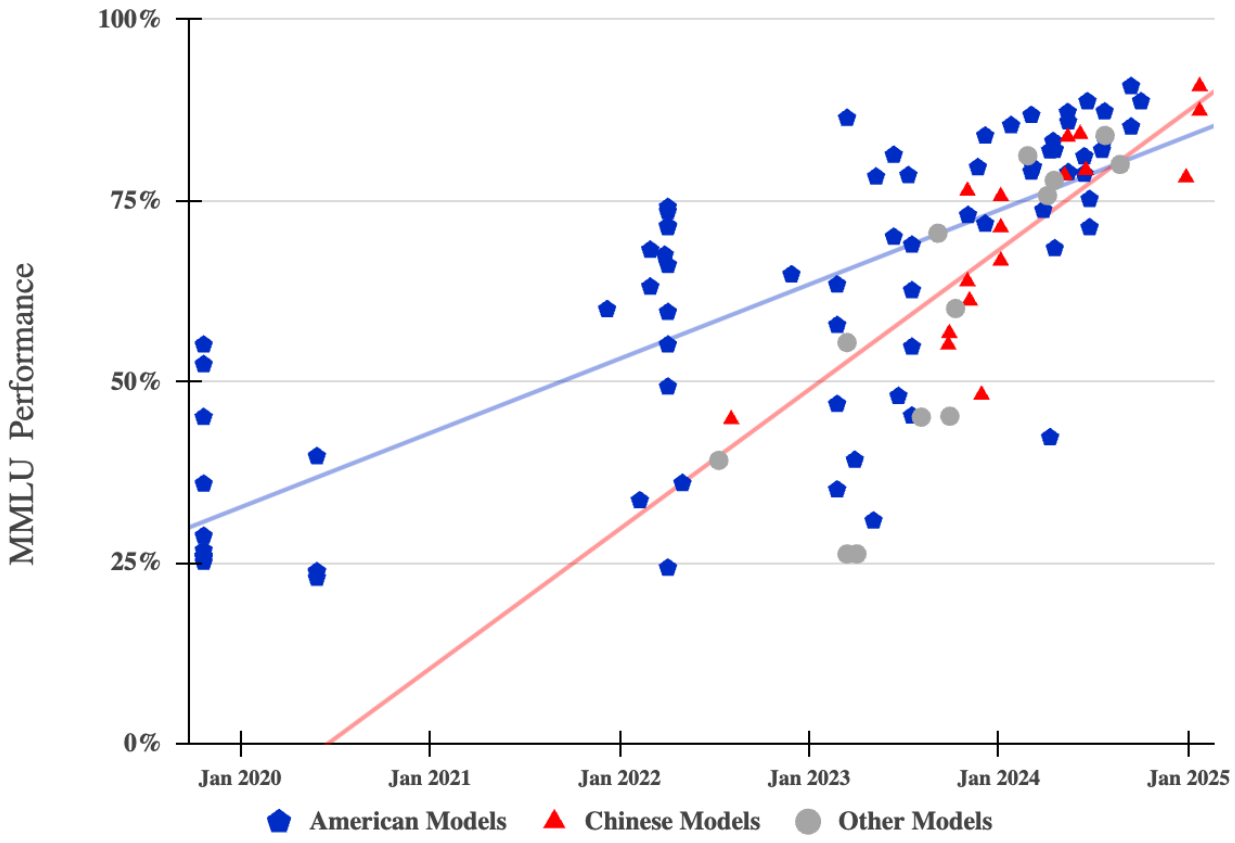
The integration of Chinese and American open models is a two-way street. Although the US does not seem to leverage Chinese models as much as China leverages US models, there are examples of American researchers using Chinese models. In 2024, an open model announced by researchers at Stanford University was revealed to be highly similar to a Chinese Llama derivative developed by Tsinghua University and Chinese company ModelBest.[69,70] The Stanford researchers did not originally credit the Chinese model, but apologized and did so after significant media attention.[69] In the same year, Abacus AI, an American startup, released Smaug-72-v0.1 and noted that it "is ultimately based on Qwen-72B," one of Alibaba's open models.[71,72] Later, Abacus AI also released a model Liberated-Qwen1.5-72B, an uncensored version of the Chinese Qwen1.5-72B.[73] After itself building on Llama, a modified version of an 01.Ai model later appeared in the US open-source AI ecosystem, demonstrating the highly integrated back-and-forth nature of open source innovation.[66] Going forward, we expect to see greater American use of Chinese open models, specifically of DeepSeek-V3 and R1.

2d. Comparison of Chinese and American Models

Performance

Although American models have historically demonstrated better performance than Chinese models, this gap appears to have narrowed as Chinese model performance has rapidly improved. Figure 2 shows the performance of frontier models against the Measuring Massive Multitask Language Understanding (MMLU) benchmark, which covers elementary mathematics, US history, computer science, law, and other topics. Prior to 2023, the dataset of frontier models had few Chinese models. Even as Chinese models improved in 2023, their performance against the MMLU benchmark lagged leading American models by over one year. Although a more comprehensive analysis would consider many benchmarks, the graph is consistent with the widely accepted trend that Chinese models are "catching up" to American models.[74–77]

Figure 2: MMLU performance of frontier models over time[24,25,78–81]



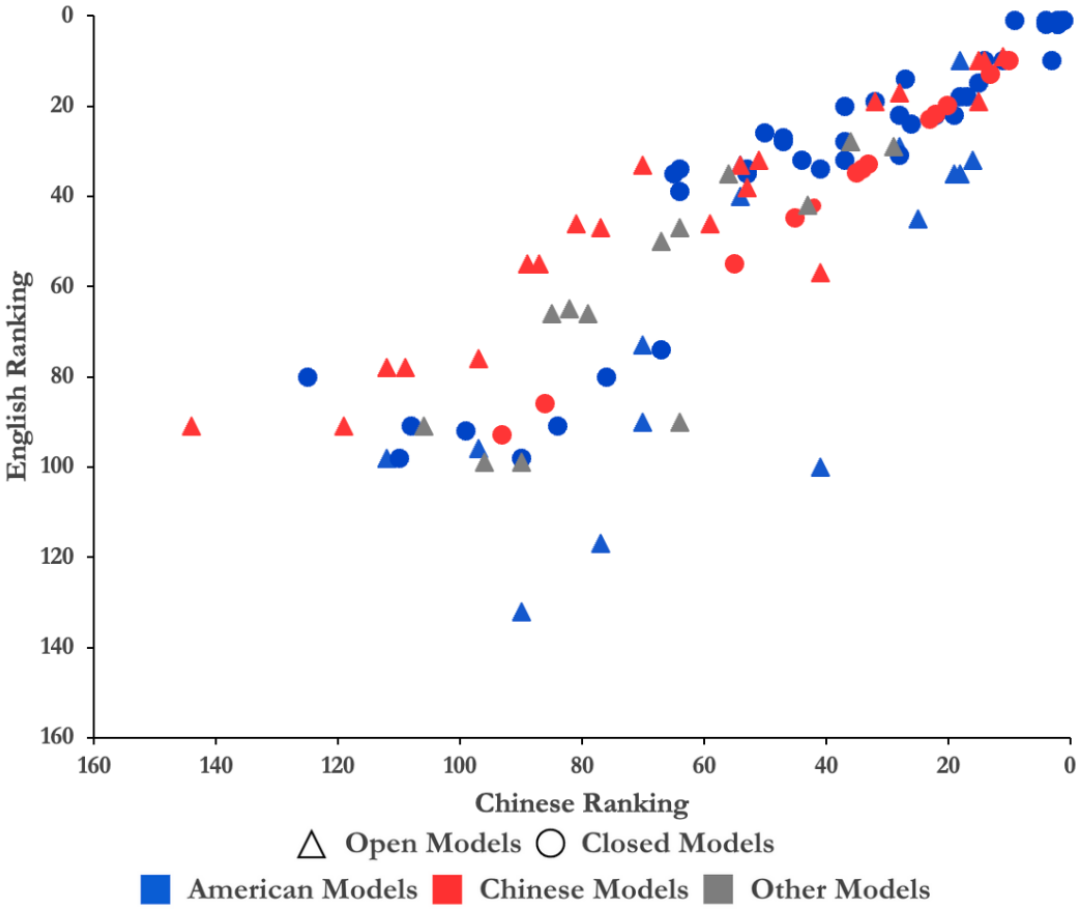
DeepSeek's latest model, DeepSeek-R1, is the most compelling example of this "catch-up." It took Chinese AI labs 12 to 18 months to catch up to GPT-4,[27] yet DeepSeek-R1 outperforms o1, the leading American model, on some metrics only three months after o1 was released.[25] On MMLU, DeepSeek-R1 is China's best model and is on par with OpenAI's o1-preview.[25,81] On other benchmarks, such as AIME and MATH, DeepSeek-R1 outperforms o1, while its coding abilities lag behind o1's.[27] It is unclear whether DeepSeek-R1 is an outlier or represents a broader improvement in Chinese innovation.

It is important to note that not all benchmarks give an unbiased ranking of Chinese and American models; some of them have certain language biases. One way to compare models in a more equitable manner would be to compare their rankings on both English tasks and Chinese tasks (Figure 3). These rankings come from Chatbot Arena, a site which allows users to make queries in the language of their choice and presents them with results from two anonymous chatbots, whereupon the user selects which answer they think is better.[82] Models in the top right corner received favorable rankings on both Chinese and English tasks, where models in the top left corner received more favorable rankings on Chinese tasks and models in the bottom right corner received more favorable rankings on English tasks. The top right is occupied by American closed models—primarily Google's Gemini family. Note that at time of data collection (February 3, 2025), DeepSeek-R1 tied for first in English, but had not accumulated enough ratings to be ranked in Chinese prompts. Outside of the Gemini family, perhaps unsurprisingly, models tend to perform better in the language of their country of origin. Models from other countries, primarily Canada, France, and Israel, tend to rank in the middle, with none threatening the top spots held by American and Chinese models.

A combination of factors may explain why Chinese models historically lagged American models prior to DeepSeek, including American export controls, Chinese censorship requirements, and Chinese tendency towards corporate espionage. First, American export controls have hindered Chinese access to compute, which has been responsible for 65% of capability improvements in AI models since 2014.[83] Chinese researchers have also blamed American restrictions and limitations on collaboration for holding back Chinese open innovation.[84] Second, China has onerous censorship requirements of AI models, which could potentially delay model release and increase fine-tuning costs.[85] Finally, China has a reputation for copying foreign innovations instead of developing them domestically.[76] There is a chance that old habits have resurfaced with AI models. This idea is substantiated by the DeepSeek CEO's characterization of the Chinese AI ecosystem below.

> *"We often say that there is a gap of one or two years between Chinese AI and the United States, but the real gap is the difference between originality and imitation."*
>
> **- Liang Wenfeng, DeepSeek CEO**[65]

Figure 3: Chatbot Arena Chinese and English Rankings



However, limited access to compute also seems to be spurring alternative forms of innovation. DeepSeek-V3 debuted an innovative architecture which required less compute, while DeepSeek-R1 mimicked o1's reasoning model and leveraged test-time compute, which is cheaper than pre-training compute.

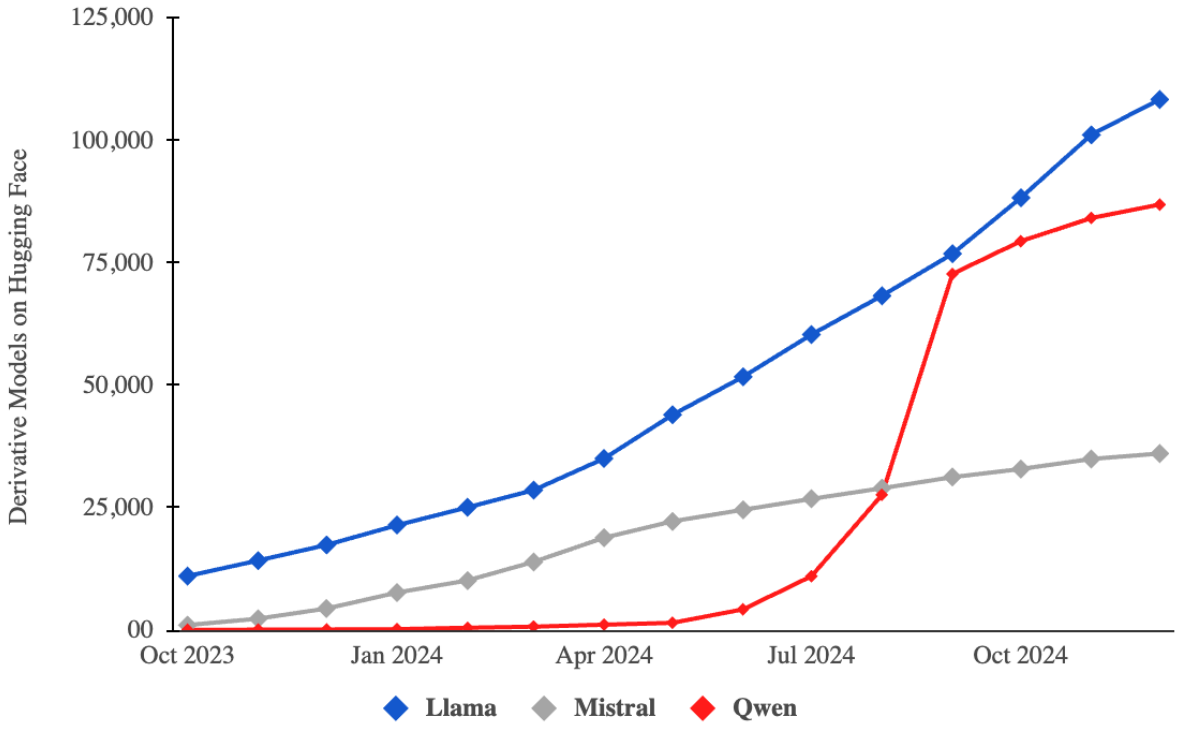
The government's role in promoting AI innovation is complex and evolving. China's defining regulatory tension is between social stability and economic development, but while the government initially seemed to be prioritizing social stability by passing multiple AI-related laws, there are signs that it is increasingly focusing on development.[86,87] For instance, between the draft and final versions of its generative AI provisions, it removed a requirement that AI outputs be "true and accurate."[87] The government also actively supports private-sector AI development.[53] As open models advance, their accessibility may threaten the stability-development paradigm, potentially requiring additional regulation.

Popularity

Although American models have historically been more popular, Chinese models lead in some popularity metrics. For example, the top three most "liked" natural language processing (NLP) models on Hugging Face are all Llama models, a potential indication that users enjoy building upon Llama models.[88] Conversely, it is a Qwen model that was the most downloaded on Hugging Face, responsible for 27% of downloads since 2022.[88] However, that model did not feature in the top ten most liked models, an indication that although many users downloaded the model, they may have been dissatisfied.

Chinese models demonstrated an impressive surge in popularity in late 2024. In June 2024, Llama had over eleven times more derivative models than Qwen. By December 2024, Llama had only 25% more derivative models. Figure 4, an updated version of the AI Technology Review graph on WeChat,[89,90] displays the trends in model popularity. Qwen's popularity drastically increased when Alibaba released over 100 Qwen 2.5 models in September 2024.[91] However, it is unclear whether this popularity is contained within Chinese borders or has been driven by international users of Qwen models. After September, Qwen's popularity plateaued and continues to lag Llama's popularity, though by far less than in early 2024. Since Qwen models are now outperformed by DeepSeek-V3 and DeepSeek-R1, the graph below may be understating the popularity of Chinese models.

Figure 4: Popularity of Llama, Qwen, and Mistral models[92]



2e. Tension between Chinese use of open-source AI and aspirations of self-reliance

China's reliance on foreign open models creates inherent tension with its typical strategy of digital sovereignty and Xi Jinping's "road of self-reliance,"[93] part of China's larger digital sovereignty strategy.[94,95] For example, use of foreign models seems antithetical to China's adamant pursuit of "internet sovereignty," which involves the "Great Firewall" and close control over its online space.[96] Similarly, dependence on American open-source AI also juxtaposes China's efforts to promote domestic semiconductor manufacturing.[94]

The Chinese government is likely aware of this tension and has demonstrated an intent to reduce dependence on foreign open-source AI. A pressing concern is that China's dependence on access to foreign open models and frameworks creates potential supply chain vulnerabilities. In 2022, a state-affiliated think tank remarked that "a considerable number of [China's] AI applications are built on the mainstream international AI frameworks."[52] Similarly, BigAI's 2024 presentation to Premier Li Qiang acknowledged that China "severely lacks autonomy" in the development of open-source AI.[63,97]

Several US actions have likely exacerbated Chinese concerns about continued access to American AI models and frameworks. For example, in February 2024, the National Telecommunications and Information Administration (NTIA) requested public input on a range of questions related to open models, including whether the US government should restrict availability of model weights.[98] The US's proposed Enhancing National Frameworks for Overseas Critical Exports (ENFORCE) Act[99] would give the president authority to restrict the export of "any software or hardware implementation of artificial intelligence, including artificial intelligence model weights and any numerical parameters associated with the artificial intelligence implementation." The latter clause seems to include open-weight models. Although the bill passed the Committee on Foreign Affairs, it failed to pass the House, but may be re-introduced in the 119th Congress.

In response to these American policy discussions, the Chinese government has encouraged companies to prioritize domestic AI frameworks, such as PaddlePaddle and MindSpore, over foreign ones such as PyTorch and TensorFlow.[45] Pre-emptively reducing reliance on foreign technologies could potentially temper the shock of any future export controls on open weight models.

However, there are several factors that could prolong Chinese dependence on both foreign AI models and frameworks. First, it is disruptive for developers to switch AI frameworks, and existing Chinese frameworks may not meet domestic needs.[45] Second, to enable a cutting-edge Chinese domestic open-source ecosystem, there likely needs to be a Chinese

organization that can provide a sufficiently powerful "foundation model" for other developers to build upon.

Since Chinese developers cannot easily scale compute due to existing export controls, to propel the performance of domestic AI without relying on foreign models, they will need to prioritize innovation and efficiency.[45] DeepSeek-V3 and ModelBest's MiniCPM model provide two examples of Chinese organizations successfully improving performance through innovations with limited computing power and could serve as key foundation models.[47,65,100]

> "*For many years, Chinese companies are used to others doing technological innovation, while we focused on application monetization — but this isn't inevitable.*"
>
> **- Liang Wenfeng, DeepSeek CEO**[65]

2f. Challenges with decoupling open source ecosystems

An American policy maker may hope to reap the benefits of openness while also maintaining a competitive edge in US-China competition. They may envision an ideal world as one in which the US has a robust open source ecosystem that delivers clear benefits to the US economy and strengthens US power, while Chinese actors are prevented from misusing such technology. American actors within this idealized ecosystem would be diligent about ensuring the provenance of the open models they leverage, so there would be minimal instances of Chinese "backdoor" access to American models. In other words, they may hope to "decouple" the Chinese and American open source ecosystems.

Since open-source AI by definition has no limitations on who can access it, such a decoupling would be difficult. Both nations would need to refuse to use open-source AI from the other nation. However, this goes against corporate interests, which incline towards leveraging potentially beneficial foreign innovations, and thus would require strict government action. Furthermore, as outlined in Section 5, attempts to fully "decouple" Chinese and American open source ecosystems through export controls may have unintended consequences. The inherent challenges in decoupling open-source ecosystems raise important questions about the feasibility and desirability of different regulatory approaches. The following section examines how, if at all, different jurisdictions have approached this challenge through regulation, revealing both commonalities and contrasts in their strategies.

3. How has open-source AI been considered in regulation?

States across the world are considering how to regulate AI. Approaches range from the EU's horizontal AI Act to China's vertical, sectoral laws to the US's fragmented, executive-led efforts.[87,101] However, open-source AI has generally been less addressed in regulation. Possible reasons for this include the performance gap between open and closed models leading to a perception that regulation is less needed, a lack of familiarity with open-source technologies, and the vocal nature of the open source community advocating against regulation. In this section, we detail how open-source AI is addressed, if at all, in five jurisdictions' AI governance efforts.

3a. United States

Concern about being eclipsed by China is central to much of the US's AI governance,[86,87] but to date the regulation of open-source AI has taken a backseat, characteristic of its light-touch approach to regulation. Neither Trump's 2019 Executive Order (EO) on Maintaining American Leadership in Artificial Intelligence[102] or Biden's 2023 EO on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence[103] mention open-source AI. The 2025 Framework for Artificial Intelligence Diffusion, which requires licenses to export model weights, explicitly excludes open model weights from the requirements.[104] The proposed 2024 ENFORCE Act would seemingly give the president authority to restrict the export of open-weight models, but how this would be accomplished in practice is uncertain. The bill passed out of committee with bipartisan support, whereupon the House Foreign Affairs Committee Chairman Michael McCaul delivered remarks on how the bill is aimed at preventing technology transfer to China and the CCP.[105] However, despite bipartisan consensus on the need to counter China, the bill did not pass in the 118th Congress.

3b. California

California's SB-1047,[106] the Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, was intended to prevent "critical harms" caused by advanced models, but faced controversy over its handling of open models. Its original requirement that developers be able to enact a "full shutdown" of models' training and use was heavily criticized for being unworkable for open models, which are out of their developer's control when downloaded by others. In a revised version, the shutdown provision was clarified to apply only to models and model derivatives "controlled by a developer," but backlash against the bill

from both industry and the open source community was intense enough that Governor Gavin Newsom vetoed it.[107]

3c. European Union

The EU's AI Act[108] is its flagship AI regulation. However, after complex negotiations and lobbying,[109,110] it explicitly excludes "free and open-source" AI from its purview (Art. 2), except for those that qualify as high-risk or prohibited systems or those subject to increased transparency requirements (Art. 2). The transparency exception thus includes AI designed to interact with "natural persons," generative AI, and emotion recognition and biometric categorization systems (Art. 50). Additionally, open-source general-purpose AI models that present "systemic risks" are still subject to documentation and authorized representative requirements. If models are monetized, the open-source exclusion no longer applies (Recital 103). Developers of open models are still encouraged to comply with "widely adopted documentation practices, such as model cards and data sheets, as a way to accelerate information sharing along the AI value chain" with the hope that this can lead to the "promotion of trustworthy AI systems in the Union" (Recital 89).

3d. United Kingdom

The UK's AI regulation is still developing. Its 2023 policy paper titled "A pro-innovation approach to AI regulation" mentions open-source AI in the context of how it can challenge life cycle accountability. Its proposed AI Regulation[111] does not explicitly mention open-source AI. However, the UK's intended "pro-innovation" approach to AI could indicate an overall lighter-touch approach to regulation similar to the US's.

3e. China

China's AI-related regulations do not explicitly exclude open-source AI, implying that they are intended to apply to open-source algorithms as well, regardless of technical feasibility. In particular, its generative AI regulation has created compliance concerns for open-source AI,[112] and its "deep synthesis" regulation[113,114] makes no mention of openness. However, in other areas it is more encouraging of openness, if not open source: its algorithm regulation[115,116] calls for following the principles of "openness and transparency" (Art. 4) and its provisions on generative AI[117,118] advocate for the "orderly opening of public data by type and grade" to expand "high-quality public training data resources" (Art. 6). As previously mentioned, China's government is prioritizing "self-reliance" in AI and technology. However, its lack of specific consideration of open-source AI in regulation may indicate a relative lack of preparedness to leverage it as a method of ensuring this sovereignty.

4. A rubric of open source considerations

The common exclusion of open-source AI from regulation is perhaps due to an uncertainty regarding the balance between its benefits and risks of openness. To analyze the implications of potential policies, we propose a rubric that fuses three ideological considerations with three geopolitical considerations. Our three ideological considerations are accelerated technological progress, increased transparency, and increased power distribution. The three geopolitical considerations are Chinese misuse of open models, "backdoor" risks from manipulated open-source projects, and global power dynamics. The ideological considerations are based on the convergence of ideas between a Centre for the Governance of AI[1] (GovAI) report and a policy brief published in *Science* (see Appendix 1).[9] The geopolitical considerations are based on arguments forwarded by the national security and geopolitical community.

Table 1: A rubric to assess open-source AI policies

Lens	Consideration	Implication for open source	Explanation
Ideological	**Accelerated technological progress**	Benefit	Open models can be widely implemented for different tasks, and allow for more contributors to push AI progress.
	Increased transparency	Benefit	Providing more information about models can enable better examination of model capabilities and risks, including through external auditing.
	Increased power distribution	Benefit	Open models allow for more people to provide input on the direction of AI and prevent single actors from exerting total control.
Geopolitical	**Chinese misuse**	Risk	American open models could be used by Chinese actors to harm US interests through traditional military use cases (e.g., improving intelligence capabilities, bioweapon design) or other threat vectors (e.g., disinformation, domestic surveillance).
	"Backdoor" risks	Risk	Open-source AI used by American corporations or governments may have been manipulated to include "backdoors" for malicious use by Chinese actors.
	Global power	Benefit	The proliferation of American open-source AI represents an opportunity to spread certain values (e.g., freedom from censorship) and build greater global influence if more countries are dependent on American technology. Conversely, China will reap these benefits if it is a primary provider of global open-source AI.

4.1. Accelerated technological progress

The first ideological consideration is accelerated technological progress. It refers to the idea that open models can drive innovation because they are easily customizable by many people.

Like the GovAI report, we divide accelerated technological progress into progress in *specific-use applications*, or how AI is used to achieve specific tasks, and progress in *frontier capabilities*, or overall model performance. *Specific-use applications* are clearly accelerated by open-sourcing AI models and frameworks, since a large focus of the open source community is building upon existing models and training open models on different datasets to serve a different purpose.[1,9] Given that it is often the application of technology to industry, rather than sheer innovation, that increases productivity, open-sourcing likely has positive impacts on economic growth.[119]

> *"The United States' advantage is decentralized and open innovation."*
>
> **- Mark Zuckerberg, Meta CEO**[120]

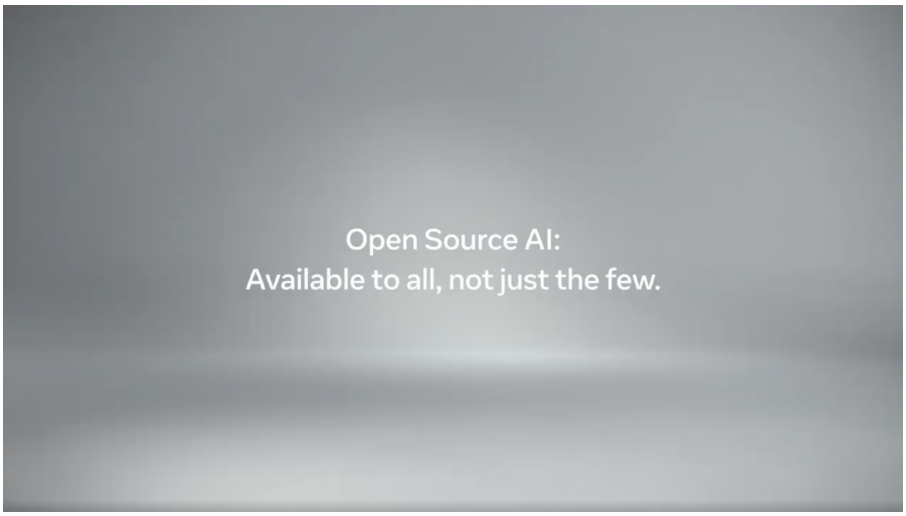
Conversely, open-sourcing seems less relevant to progress on *frontier capabilities*. To date, closed models have propelled frontier capabilities forwards, with open models still lagging by one year.[24] Moreover, open-sourcing has limited benefit for capabilities improvements due to "bottlenecks in the talent, compute, and data resources."[1] Given that 65% of capability improvements stem from increases in compute,[83] open-sourcing is unlikely to be a fundamental driver of capability improvements. Although open models have driven algorithmic innovations, many algorithmic innovations emerge from companies with closed models due to their ability to source high-quality talent.[121] Thus, we consider open source's impact on frontier capabilities to be limited. Of course, this may change; the aforementioned DeepSeek-V3 was trained with significantly less compute than models it outperforms on leaderboards.[47,78] If algorithmic innovations continue to emerge from the open-source AI community, then open-sourcing may become more relevant to driving frontier features and capabilities; Sam Altman credited DeepSeek-R1 for pushing OpenAI to plan to add chain-of-thought to its o1 model.[122]

The GovAI report also discusses how open-sourcing can advance progress in safety research by providing comprehensive model access, but that such model access can be achieved through other methods, such as external audits. Since there are feasible alternatives to achieving the same benefit, we exclude this benefit of progress on safety research from our analysis.

4.2. Increased transparency

The second ideological consideration is increased transparency. This consideration focuses on the transparency that sharing model components, such as model weights, source code, and datasets, provides. It also refers to the external evaluation process enabled by component disclosure, as identified by GovAI. Transparency is a clear benefit of open-source technology in general, but the degree to which it holds for open-source AI depends on the exact level of openness of the model. The *Science* paper focuses on a wide variety of transparency metrics based on the Foundation Model Transparency Index, which includes broader indicators like labor practice disclosures. These are not unique to companies that open model components, but companies that bill themselves as providing "open" models are more likely to engage in other transparent practices, like issuing detailed model cards.[123] Since our definition of open models is oriented around model components, we focus on the transparency benefits that open model components can enable, rather than those broader indicators.

Figure 5: Meta LinkedIn advertisement



4.3. Increased power distribution

The third ideological consideration is increased power distribution. It refers to the idea that open-sourcing AI models prevents concentration of power. Open models allow for more people to provide input on the direction of AI and prevent single actors from exerting total control. Meta's advertising campaign, which frames open-source AI as "available to all, not just the few" in Figure 5 channels the essence of this argument.[124]

This concept is discussed in a similar manner by the two reports. The *Science* article says that open AI models prevent the owners of closed models from unilaterally making

decisions about their use, and the GovAI report discusses how open models give more people influence over how AI is developed and used. Acknowledging that this is a form of "socioeconomic power" per the *Science* report, we derive the consideration of "increased power distribution."

4.4. Misuse of American open-source AI

The first geopolitical consideration is misuse of American open-source AI. It refers to the possibility that by making American models open-source, the US is equipping a strategic rival, likely China, with powerful technology for potential misuse.

Once Chinese developers, companies, and governments have access to open AI models, American developers can no longer influence how they use these models. Unlike closed models, open models have no oversight over usage or shutdown capability, safety measures can be removed, and models can be finetuned for potentially dangerous purposes.[6] For example, Stable Diffusion's filter was removed through a single line of code.[1] The lack of control extends to general use concerns; although Meta's acceptable use policy prohibits "military, warfare, nuclear industries or applications, [and] espionage" use,[125] it has no way to enforce this policy and was unable to prevent PLA-affiliated researchers from modifying its model.[67] Thus, Chinese actors could use American models in a manner that the US government would find undesirable—be it to improve Chinese intelligence capabilities, advance Chinese offensive cyber capabilities, or strengthen other Chinese military capabilities.

There are a variety of use cases that would be considered misuse by the US government, some of which are generic risks that could be posed by any actor, and some that are particularly salient to China. Theoretically, any malicious actor, state or non-state, could fine-tune a model to generate non-consensual intimate imagery (NCII), disinformation, launch cyberattacks, or try to design a new pathogen, though the latter would require access to a laboratory to execute. Other strategic competitors could leverage an open LLM to advance their intelligence capabilities. Similarly, image recognition tools could be used for surveillance amongst other authoritarian countries.

Still, China remains a greater strategic concern to the US government. On top of existing geopolitical competition issues, unlike other state or non-state actors, China has greater capabilities and has demonstrated more intent to pursue some of these use cases. For example, although any actor might be able to fine-tune a model to synthesize intelligence, they may not have the data to do so or the power to utilize that intelligence in a harmful manner. Conversely, China has a notoriously sophisticated surveillance state,[126,127] which is coupled with oppressive policies towards political dissidents and certain minorities, such

as the Uighur people.[128] There is also evidence that China is responsible for transnational repression, such as monitoring Chinese nationals on overseas university campuses, which could be augmented by greater intelligence synthesis.[129] Finally, and perhaps most pressingly for American policymakers, China may use such a model to inform military decisions or in military systems. Thus, China's access to open models may directly improve its military capabilities, which the US government would view as unacceptable given the current state of US-China competition.

Although these examples of Chinese misuse of American open-source AI are possible and significant, there are two limitations to this argument. These limitations do not indicate that these risks are irrelevant, but rather that the marginal impact of these risks may be overstated.

First, American open models may not be equipping China with significantly new capabilities, meaning that the *marginal risk* is relatively low.[9] The marginal risk of Chinese actors using open-source technologies can be intuited by comparing this use case to the capabilities Chinese actors already have access to—be it via Chinese closed models or other open-source capabilities. For example, when 01.AI's model, Yi-34B, leveraged Llama 2's architecture, *New York Times* coverage raised questions of geopolitics and national security.[64] However, EleutherAI, a non-profit open-source AI research group, characterized Yi-34B's use of Llama 2's model architecture as unremarkable and well within the bounds of common machine learning practices, because such model architecture was already available.[130] Using the logic of marginal risk, if leading Chinese models match or exceed the performance of Western models, then there is less marginal risk to releasing Western models as open-source.

Furthermore, as the Chinese AI ecosystem is already somewhat sophisticated, the marginal risk of access to open models may be greater for other actors, such as terrorist groups, who may not have the resources to train a model from scratch. While China does experience bottlenecks on compute, it is still capable of developing closed models from scratch, such as Baidu's Ernie Bot.[131] Crucially, both DeepSeek-R1 and DeepSeek-V3 outperformed Llama on some metrics.[25,78] While it is unclear whether these DeepSeek models signal a paradigm shift in the relative performance of Chinese models going forwards, if Chinese models trend closely to American models, then open models only increase marginal risk if they share algorithmic innovations previously unavailable to Chinese AI labs.

Second, these risks may receive disproportionate attention due to the anti-China sentiment that permeates the American political atmosphere. For example, there are other authoritarian countries that conduct domestic surveillance and could use open models to

augment such surveillance,[132,133] but those countries do not receive as much attention as Chinese examples. The preoccupation with Chinese misuse is in part due to current US-China competition, which heightens any fears about China having access to advanced technology. Certain American pundits quickly conflate Chinese military uses of AI with companies developing commercial models, which is not necessarily nefarious. However, distinguishing between the two is increasingly complex, given the expectation that Chinese companies may be collaborating with the Chinese military, as multiple American technology companies have done with the US security agencies.[134] Still, by assuming that any Chinese use of American AI is dangerous, even commercial uses, the US may find itself pursuing measures that explicitly aim to economically kneecap China, rather than measures which may have economic consequences but that are intended to address national security concerns.[135]

> *"The U.S. cannot allow Chinese Communist Party models such as DeepSeek to risk our national security and leverage our technology to advance their AI ambitions."*
>
> **- Rep. John Moolenaar (R- Michigan)**[136]

4.5 "Backdoor" risks

The second geopolitical consideration is "backdoor" risks. When American companies and government agencies use open-source AI, there is a chance that these models and toolkits are equipped with vulnerabilities that could be exploited by the Chinese government. Like with misuse risk, this risk is not unique to the Chinese government; both other state and non-state actors could plant a vulnerability or "backdoor" in open-source AI products, but it is still an important security concern.

The rate of security threats stemming from open-source AI tools and a history of vulnerabilities in other open-source software suggests that this is a non-negligible threat. In a survey of 1,000 corporate IT decision-makers, 10% of respondents revealed that their use of open-source AI tools had led to the accidental installation of malicious code, while 20% had experienced sensitive information exposure.[137] 30% of the respondents had seen their company accidentally expose security vulnerabilities when using open-source AI tools, and the majority of these incidents were classified as severe impact.[137]

Open-source technologies have created vulnerabilities with widespread impact, indicating that open-source AI could be a vehicle for large-scale threats. In 2024, a backdoor was discovered in XZ Utils, a data compression library available on most Linux installations.[138] The backdoor enabled anyone who had the encryption key, likely held by the entity who planted the backdoor, to upload and execute any code on the device with that version of XZ Utils.[138] Though the perpetrator is unknown, it is suspected to be a state actor based on

the sophistication of the attack and patterns of activity from the accounts associated with the backdoors.[139] The National Institute of Standards and Technology issued the backdoor a common vulnerability scoring system (CVSS) score of ten, the highest possible score and deemed it a "critical vulnerability."[140–142] Overall, 17% of vulnerabilities in open-source software are planted for malicious purposes,[143] which likely carries over to open-source AI as well.

Indeed, specific open-source AI models and toolkits have demonstrated vulnerability to compromise. In October 2024, Protect AI revealed that their bug bounty program discovered 34 vulnerabilities in open-source models and toolkits.[144] One vulnerability in Lunary, a production toolkit, would enable an attacker to update another user's prompt without authorization or to delete an external user record.[145,146]

Although many actors could utilize open-source AI as a backdoor, the Chinese government is likely to be a primary culprit of future attacks given its history of infiltrating US supply chains. CSIS has catalogued 224 examples of Chinese espionage since 2000, many of which are corporate espionage.[147] Salt Typhoon, a China-backed hacking group, is a recent example of a large-scale infiltration of American industry and critical infrastructure. The group successfully hacked at least eight American telecommunications providers to gain access to the communications of (then) President-elect Trump, Vice President-elect JD Vance, Vice President Kamala Harris, and State Department officials.[148,149]

4.6. Global power

The third geopolitical consideration is global power. It represents the view that the US and China need to compete for dominance in the global open-source AI ecosystem to support their economic and soft power.[150,151] By this logic, extensive use of Chinese open-source AI by third party countries would constitute a risk, whereas it is highly desirable for third-party countries to predominantly use American open-source AI. Below we will discuss three ways that leading in open-source AI translates to global power. It is important to note that the overall strength of a country's AI ecosystem—both open and closed—relates to its economic and thus global power as well; this is discussed in 4.1: Accelerated technological progress. Our discussion below focuses specifically on how the use of open-source AI directly affects geopolitical power.

First, open-source AI could be a vehicle for disseminating a Chinese perspective on the world.[150] For example, a Chinese open model may be politically censored or trained to share certain perspectives on the US and China; analysis suggests that this is the case for both open and closed models originating in China.[152,153] If users in other countries treat these models as a source of truth, there is a risk that China's view of history and the current state

of the world could become more prevalent. Although there is some anecdotal evidence that certain Chinese open models are uncensored when deployed locally,[154] the sheer volume of API users still means that many users would encounter model censorship.[155]

The Chinese government has also already demonstrated an inclination to appeal to other countries, particularly those in the Global South, and is pushing a narrative that exacerbates existing dissatisfaction with the existing world order and in some cases, contempt towards the United States.[156–159] Even within the guardrails of fact, models could tend towards portraying the United States in an unfavorable light by highlighting real atrocities committed by the US.

Second, if countries develop a reliance on either Chinese or American open software, this may provide a foundation for stronger alliances or relationships. This perspective associates technological dependence with global power. For example, if other countries are highly reliant on American open-source AI, then the US may build stronger ties with these countries or at minimum have greater leverage to exert. Taiwan's semiconductor foundries are a relevant example of technological leadership garnering geopolitical advantage; it is possible that American dependence on Taiwanese foundries may have influenced American foreign policy towards Taiwan.[160–163] Likewise, the dependence upon American open-source AI could strengthen American influence in other countries.

Third, proliferation of Chinese open models could aid Chinese intelligence efforts through collection of user data. While this is not a risk when models are downloaded and run locally, it is possible if users access the model through an API (including via a website or app).[164] If a user enters personal or sensitive information into a generative AI model prompt, it could be gathered by the company;[165] the terms of service of many AI companies, including OpenAI,[166] retain the right to store inputted information for later training use. Other potential future uses include the sale of data to other companies or its provision to governments.[167] Besides personal or sensitive data, there also is a chance that APIs are collecting data on user patterns,[167] which could be used to infer user attributes and behaviors.[168] Particularly in China, where the government retains broad rights to compel data disclosure from companies for nebulous "national security" purposes,[169] this data could be leveraged by the government to provide a strategic advantage.

DeepSeek-R1's popularity lends credence to these concerns. As of January 27, 2025, DeepSeek had 2.6 million downloads on the Apple and Google Play app stores alone, which excludes those that engage online, directly via the API, and those downloading the model weights.[170] Yet, even before DeepSeek-R1's release, there was evidence that Chinese open models were competitive with Western equivalents; Figure 4 above shows the impressive rise in Qwen's popularity through late 2024. Although the graph does not identify where

geographically derivative models were developed—many could be domestic users—it is a sign that Chinese models could match the popularity of American models.

DeepSeek-R1's release has catapulted this consideration into mainstream media.[171] For example, there has been extensive coverage of the possibility that DeepSeek is collecting personal data (e.g., proof of identity, passwords) and any prompt data, as permitted by its privacy policy.[164,172] Mainstream media outlets are also reporting on the censorship of DeepSeek,[173] as well as the potential for technological dependence. Although these are not novel concerns—OpenAI also has the right to collect user data and its models refuse to comment on certain political topics[174]—DeepSeek is a Chinese company. Thus, the concern is that DeepSeek enables adversarial surveillance[175] as opposed to the potential surveillance capitalism enabled by OpenAI.

Prior to the release of DeepSeek-R1, public discussions of open source as contributing to global power were mainly conducted by geopolitical analysts and some of the AI lab. For example, In November 2024, Nick Clegg, then Meta's Vice President for Global Affairs, published a blog espousing the national security benefits of American leadership in open-source AI.[176] His argument rested on the assertion that China and the US are competing to set the "global open source standard."[176]

> *"We believe it is in both America and the wider democratic world's interest for American open source models to excel and succeed over models from China and elsewhere."*
>
> **—Nick Clegg, VP Global Affairs, Meta**[176]

There are three important nuances to the argument that open-source AI is a vehicle for global power.

First, this argument can be in tension with concerns over Chinese misuse of open-source AI. Since creators of open models typically have limited control over *where* open models go, models released to third-party countries may easily be accessible by Chinese actors. Thus, it is challenging for the US to restrict Chinese access to open models while simultaneously seeking to dominate the global open-source AI ecosystem.

Second, some skepticism should be applied to corporate perspectives on this topic. That corporate interests are weighing in on matters of geopolitics could be interpreted as a signal of the strength of this argument, or instead a signal that AI companies have recognized and adapted to the current American political landscape, which views these companies as conduits of national technological power. China's days of designating big tech "national champions" seem to be behind it and the US has never had such an official

designation,[53] but big tech companies are still key to advancing economic power and the soft power that comes with having the most advanced models.

Third, there is a risk that preoccupation with open-source AI as national power could encourage an unfettered arms race. One prevalent perspective in the US, expressed below by US AI Czar David Sacks, is that safety measures slow innovation, lend an advantage to China, and thus should be deprioritized. However, there is limited evidence to suggest that ensuring the safety and reliability of frontier models would hinder the US's technological advantage.[177] For example, Chinese models are subject to censorship regulation and they have a licensing regime.[57] Despite these regulations, DeepSeek-V1 achieved impressive performance with its CEO citing limited access to compute as its key bottleneck,[65] rather than censorship requirements.

A deprioritization of safety forfeits strategic advantage and increases the likelihood of unintended harm. When engaging in competition, there is value in reliable AI models, particularly when used in a military context. If American and Chinese developers deprioritize the safety of frontier models, there is a chance that either nation's models could cause significant harm accidentally. Amidst current geopolitical tensions, an accidental incident could escalate significantly.

> *"DeepSeek R1 shows that the AI race will be very competitive and that President Trump was right to rescind the Biden EO, which hamstrung American AI companies without asking whether China would do the same. (Obviously not.) I'm confident in the U.S. but we can't be complacent."*
>
> **- David Sacks, US AI Czar**[178]

5. Analysis of specific policies

As the marginal risk of open-sourcing models is still unclear, the risks from misuse of open models may not warrant any policy intervention today. However, as capabilities advance, we expect the policy community to engage more with the misuse risks associated with open-source AI.

In anticipation of future consideration of open-source AI policies, here we discuss four open-source AI policies. We use our rubric from Section 4 to assess the potential implications of these policies. Two of these policies are designed to address the risk of misuse of American open-source AI by China, and two of these policies seek to mitigate the risks from using potentially Chinese-manipulated open-source AI.

Policy 1a: Export controls on powerful open models

The NTIA and ENFORCE Act have floated export controls that would affect open-source AI, but export controls on open model components are likely to be highly disruptive to the development of specific-use applications. If implemented without domestic safety measures on open models, they would have limited efficacy in mitigating misuse risks by China and would fail to address misuse by domestic actors. Export controls may also undermine US global power by introducing friction for other countries using American technology. For these reasons, we believe that there are more effective and less disruptive policies that address similar risks, but we explore the implications in more detail below.

In practice, export controls on model components would require developers of all open models to conduct "know your customer" (KYC) activities on any potential users of their models to ensure that they were not linked to Chinese actors. If American users pass KYC assessments, they could then gain full access to open model weights and other components. It is likely that these export controls would be limited to some definition of "riskier models," be it through considering compute as a proxy for model performance, the system aim, or the content of the training data. Theoretically, more model components could be safely shared with known customers, but it would still be developer-dependent, and there would be some delay in accessing these model components, since KYC could take some time. Ultimately, these models would no longer be considered "open models" since they would not be freely available online. We explore the implications of expansive export controls in more detail below.

Accelerated technological progress

Export controls on open-source AI would be highly disruptive to progress on specific-use applications, but its impact on frontier capabilities is less clear. As discussed in Section 4,

open-sourcing enables developers to customize existing models, apply them to tasks, and develop new applications of AI, so restricting model availability would certainly impact this.

Since all users, regardless of their nationality, would need to undergo a KYC process, anyone wanting to use covered open models would face a longer wait than before export controls. There are over one million registered users on Hugging Face and Meta's models have been downloaded 400 million times.[179,180] The sheer volume of open source users could equate to a higher volume of KYC requests and longer screening process than seen on export controls on products with fewer customers, such as semiconductors. This delay would likely affect the rate at which specific-use applications are developed, which in turn could have implications for economic growth.

Whether export controls would hinder frontier capabilities depends on how many algorithmic innovations emerge from the open source community. Historically, progress has been driven by closed models, but DeepSeek-V3 and R1 are interesting exceptions. Should open models be made closed through export controls, they could still inspire algorithmic innovations like how OpenAI's o1 inspired DeepSeek-R1, but it would require more work to develop similar capabilities.

Increased power distribution

Export controls would be somewhat disruptive to the distribution of power around AI. By design, export controls seek to concentrate the control of open-source AI away from potentially malicious actors. However, in theory, the only actors who would be prevented from accessing open models would be those linked to the Chinese government. Non-malicious civilian and corporate users of open-source AI should be able to continue using these models, albeit with a delay in the timing of their access. In practice, the delay in accessing open models may frustrate users and result in fewer people using open models. Whether people disengage from the open-source community depends on the extent of the delay and how burdensome KYC measures are. Essentially, the *opportunity* to distribute the control over AI remains, but with greater friction and potentially less uptake.

Increased transparency

Export controls would not be significantly disruptive to external model evaluation. As previously mentioned, there exist reasonable substitutions to the open-source community checking for bugs and researching safety issues, including third party auditors, red-team communities, and safety bounties.[1] Regarding technical transparency more broadly, while these controls could enable more components to be shared with fewer risks, this would still be developer-dependent, and if these controls discouraged the use of open-source AI, they would overall be detrimental to transparency.

Misuse of open-source AI

Export controls would be an imperfect mitigation for the risk of misuse of American open-source AI by the Chinese government. Although they will create friction for covered persons, they cannot guarantee that these users will not gain access to model components, particularly if there are no measures to mitigate the risks of domestic misuse of open models.

First, covered persons may exploit loopholes. For example, export controls on semiconductors originally focused exclusively on hardware, allowing Chinese government entities to legally access cloud computing.[181] Even after two updates to those export controls, loopholes remained.[182,183]

Second, even with no loopholes, there may be third party intermediaries who are willing to pass on information to the covered persons. In the case of semiconductor export controls, there have been examples of targeted entities gaining access to hardware smuggled through third party countries.[184] For example, when Meta made Llama model weights only accessible to academic researchers and other approved individuals, the weights were leaked within a week and became publicly available.[185] Unlike semiconductor chips, model components are information, meaning that they can cross borders far faster and easier than physical goods.

Model components can also be stolen more easily. While the AI developers sharing model components may have robust cybersecurity measures, it is unlikely that every user of an open model does. These users could be subject to cyberattacks by covered persons to gain access to model components. If export controls cannot prevent Chinese actors from obtaining semiconductor chips, they will be even less effective for open models.

Furthermore, export controls have an inherently narrow focus and neglect other risks associated with open-source AI. For example, there may be domestic actors that could seek to misuse open models. Export controls would also fail to mitigate any technical risks of open-source technology, such as models behaving in unexpected ways or acting in error when deployed in high-stakes situations (e.g., the maintenance of critical infrastructure, usage in military contexts). Alternate measures would be needed to mitigate these risks if open models capabilities reach the frontier and marginal risks are high.

"Backdoor" risks

Export controls could have a neutral to slightly negative impact on the risks from using Chinese manipulated open-source AI. Because export controls seek to address Chinese

misuse of American open-source AI, they have no positive impact on the "backdoor" access enabled by Chinese-manipulated models. However, there is a small chance that export controls indirectly increase this risk. Since export controls create friction in accessing American model components for all users, American companies and individuals may begin using other more accessible and less reliable open models. If those models tend to have more "backdoors," then the scale of this risk increases.

Global power

Export controls could undermine American global power, since they will likely limit the proliferation of American open-source AI. With KYC requirements introducing new friction on accessing American open models, companies and citizens of other countries could be drawn to more easily accessible models. Alternative models will likely be Chinese models, since these regions are the closest to tailing American dominance in open-source AI, as seen in Figures 2-4. If users in other countries shift away from American open-source AI, the US may forfeit a technological avenue to advance its influence. Furthermore, American policymakers may take issue with certain exclusions or framing of history and politics by Chinese models.

Weighing the implications

Export controls would be an imperfect mitigation of Chinese misuse risk and, when implemented in isolation, would leave domestic misuse risk entirely unaddressed. They would be highly disruptive to technological progress on specific-use applications and could potentially disrupt frontier capabilities if open-source AI becomes a core driver of algorithmic improvements. This policy also has potential consequences for global power, since it would restrict access to American open models.

Whether such a disruptive policy is a sensible approach partially depends on marginal risk of misuse. There is ongoing debate over the scale of the marginal risks enabled by open-source AI, particularly considering that DeepSeek-R1's innovations seemed more closely linked to a closed model—OpenAI's o1—than open models, although earlier iterations of DeepSeek likely leveraged Llama models.[27,28] Without unambiguous evidence that there is a high degree of marginal risk associated with open models, imperfect risk mitigation may not be worth the impact on innovation and geopolitics. For these reasons, we suggest that export controls are not an appropriate intervention given the current context. However, if the risks from open models significantly increases, such as if open models begin to outperform closed models, then this policy may be worth reconsidering

Table 3: The implications of export controls on American open models

Considerations		Implications
Ideological	**Accelerated technological progress**	• **Highly disruptive to specific-use applications** as export controls would create delays for those wanting to fine-tune models. • **Somewhat disruptive to frontier capabilities** since closed models would be largely unimpacted, but could limit algorithmic innovations coming out of the open-source community.
	Increased power distribution	• **Somewhat disruptive** as export controls introduce frictions for non-malicious users, who may disengage with open-source AI.
	Increased transparency	• **Less disruptive** as there are reasonable substitutes for open model evaluation (e.g., red-teaming).
Geopolitical	**Misuse of open-source AI**	• **Imperfect mitigation** as export controls introduce greater friction for Chinese actors in accessing these models, though difficulty in controlling information may limit the efficacy of the policy.
	"Backdoor" risks	• **Neutral to slightly negative impact** as export controls introduce friction on American models and could drive American users towards riskier models.
	Global power	• **Weakens American influence** as other countries are incentivized to use Chinese or other non-American models out of convenience.

Policy 1b: Industry-led assessments of model release

An alternative policy would require developers of foundation models to conduct an assessment of each model they intend to make open source. Developers would be required to assess the marginal risks associated with each model and provide rationale for why it is safe to release their model. A neutral third party such as an industry body or a government agency would review documentation. Areas of investigation would include comparing model performance to existing capabilities to understand the marginal capability uplift it would enable, as well as assessing whether the type of model is inherently dangerous. Keeping in mind that openness is a spectrum, developers could consider different model release options ranging from full open-source (where all model components and data are available to everyone) to somewhere in the middle (e.g., some model components available to everyone; gated access, hosted model access, or cloud-based access) to not releasing them.[185] For more dangerous models that also have clear benefits for research (e.g., certain chem-bio models), developers may consider providing full model access to only those with certain academic or institutional email addresses.

This policy shares some similarities with Meta's Frontier AI Framework, which proposes that evaluations take into account the deployment context when considering whether to release a model.[186] However, our inclusion of third party review prevents conflicts of interest from influencing the final decision on how to release a model.

We expect that this policy would be less disruptive than expansive export controls. The effect on both technological progress and global power would be smaller since it is a targeted policy that allows safer open models to remain fully open. It should also be more effective in mitigating misuse risks.

Accelerated technological progress

Compared to export controls, this policy is less disruptive to specific-use applications and could be more effective in mitigating misuse risks from open-source AI. Since the model-by-model assessment determines how models are released, less dangerous models would remain widely available, which would be less disruptive than the blanket KYC verification associated with export controls. As discussed in the export controls section, disruption to open model accessibility may have less of an impact on frontier capabilities, since most capabilities have been driven by closed models. Regardless, this model-by-model approach would have a smaller effect on frontier capabilities than export controls would.

Increased power distribution

There would be some disruption to power distribution, but this would be limited to the riskiest models. The targeted model-by-model evaluation would ideally enable democratic distribution of power for less dangerous models.

Increased transparency

As with export controls, this policy, to the extent that certain models would not be made publicly available, may limit transparency. However, there are alternative means to obtain transparency, such as through official external audits of limited-release models.

Misuse of open-source AI

The targeted nature of this policy could make it slightly more effective than export controls in mitigating misuse risks by China; if a model has a high likelihood of enabling dangerous misuse, then it would be harder to access. Conversely, export controls would allow all model components of a highly dangerous model to be available, but only to domestic actors. Unlike export controls, this policy would be far more effective in addressing non-Chinese misuse risks, such as those posed by other state or non-state actors. However, this policy has limitations for similar reasons to export controls.

Specifically, intermediaries willing to pass along model information and cyberattacks on developers could render it less effective as a mitigation.

"Backdoor" risks

This policy could slightly reduce "backdoor" risks, with any possible effect due to additional frictions on accessing specific models.

Global power

To the extent that specific models are not made freely available, this may also incentivize third party countries to use non-American models. However, the targeted nature of this regulation should limit this effect.

Table 4: The implications of industry-led assessments of model release

Considerations		Implications
Ideological	**Accelerated technological progress**	• **Somewhat disruptive to specific-use applications** as there are targeted decisions around what is not available as open source. • **Limited disruption** to frontier capabilities since closed models would be largely unimpacted and impact on open source is more targeted.
	Increased power distribution	• **Somewhat disruptive** as less risky open models can still be accessed by many, but riskier models harder to access.
	Increased transparency	• **Less disruptive** as there are reasonable substitutes for open model evaluation (e.g., red-teaming).
Geopolitical	**Misuse of open-source AI**	• **Imperfect mitigation** as export controls introduce greater friction for Chinese actors in accessing these models, though difficulty in controlling information may limit efficacy of the policy.
	"Backdoor" risks	• **Neutral to slightly negative effect**
	Global power	• **Slightly weakens American influence** as other countries are incentivized to use Chinese or other non-American models out of convenience.

Weighing the implications

This policy should be more effective and less disruptive than expansive export controls. Since it focuses on assessing how each model should be released, rather than exclusively considering geopolitical actors, it would mitigate both geopolitical and domestic misuse

risks. The targeted approach would likely be more effective in mitigating misuse, since highly dangerous model components would be harder to access, rather than being freely available to domestic actors as would be the case under export controls. It would be less disruptive to technological progress, since there would be no wait for less risky models, which could be made freely available online. Similarly, the targeted nature would also limit impacts on US power, as there would be less incentive for other countries to use non-American models.

Policy 2a: KYS for government use of open models

While outbound models create unique risks that must be addressed, as discussed, the open-source AI ecosystem is cross-pollinating. Thus, there are also salient backdoor risks associated with American use of open models, particularly in a US government context; the need to ensure the security of open-source software used by the federal government was acknowledged in two Biden executive orders,[187,188] as well as the National Cybersecurity Strategy.[189]

Since open-source software is also subject to backdoor risks, there are relevant best practices to minimizing those risks when using open-source AI. Specifically, the idea of "traceability," or knowing where all model components originated, is highly relevant to these risks.[190,191] Indeed, the US's National Cybersecurity Strategy acknowledges that open-source software supply chain risk mitigation is vital to national interests, and the US Cybersecurity and Infrastructure Security Agency (CISA) has been working to establish common standards for a "software bill of materials" (SBOM) to identify software components and their origins.[189,192]

This policy proposes implementing mandatory security verification procedures ("know your source," or KYS) for open AI models and certain open-source packages, such as model frameworks, used in government products or services. The aim would be to improve supply-chain security while maintaining an inflow of cutting-edge AI. The cornerstone of this policy is a comprehensive internal audit requirement for developers, be they agencies or contractors (including major technology companies that contract with the government[134]), that utilize open AI models in government projects or leverage open-source projects as key components in models for the government. For example, if a large technology company is selling a government agency a proprietary product that includes a fine-tuned external open model, then it would be required to examine that code. This concept of auditing the provenance of model components closely resembles the principle of traceability included in the Department of Defense's AI Ethical Principles.[193]

Auditing external open models, whether they are being used wholesale as part of a larger system or built upon for a new model, is necessary to ensure that those models are secure, as is auditing certain packages used in model development. This audit process would encompass technical assessment of model architecture and training procedures—which would vary based on how "open" the model is—along with thorough documentation of model lineage and development history. These audits should follow a standard format, potentially building on existing frameworks like the popular Open-Source Security Testing Methodology Manual[194] (more appropriate for non-AI packages) or the AI-specific AI Security Risk Assessment.[195] Audits should also incorporate analysis of open-source project commit histories and contributor patterns to identify potential anomalies and potentially also assess project leadership structures. Existing tools like the OpenSSF Scorecard,[196] recommended by several government agencies in a fact sheet on open-source software security,[197] could aid in this process. Covered organizations should also implement regular security vulnerability scanning and performance benchmarking against known attack vectors, with procedures updated continuously as new threats emerge.

A clear threshold for what models and other packages, likely including open-source AI frameworks, would be covered would need to be established to avoid overly burdening developers; having to audit every single open-source dependency would likely excessively stymie development for little security benefit.

This policy could be an effective risk mitigation for backdoors impacting government, though it requires some restriction on what open-source AI models government agencies can use. Its success heavily depends on whether model components such as code and information about training data are available. Since audits would not be public, this policy would not help companies and individuals navigate backdoor risks; an alternative is discussed as Policy 2b.

Accelerated technological progress

This policy could be somewhat disruptive to specific-use applications in a government context. While security verification could increase trust and adoption of open-source AI, these KYS audit requirements may slow the integration of AI into government departments. There is also a chance that advanced open models and/or frameworks are found to be insecure or cannot be effectively audited and government agencies elect not to use them.

This policy is unlikely to be disruptive to frontier capabilities since these are not typically advanced through government use of open-source products. It would also not disrupt specific-use applications outside of government products, so would have no effect on the broader open source community.

The defense contracting industry, accustomed to supply chain security measures, may not be excessively impacted, since they already have internal processes for the use of open-source software.[198] Conversely, newer government contractors and smaller players without the resources to devote to these audits, especially those that rely heavily on open-source software models and frameworks, may face higher barriers to entry. However, this may be considered the "cost of doing business" with the government, which has notoriously complex contracting processes.[199]

Increased power distribution

Contractor-led audits would be minimally impactful on the distribution of control of open-source AI, since they only apply to government uses.

Increased transparency

This policy would likely be beneficial for transparency, but only for government agencies. It is dedicated to external evaluations and thus would be beneficial by generating detailed documentation about model architectures and behaviors, while standardized assessments would create comparable metrics across models. It may encourage developers to focus more on the traceability of their model components so they can win government contracts. However, if this information on traceability is not made available to other model users, it has limited effect on broader transparency. Policy 2b will address these concerns.

Misuse of open-source AI

As a policy dedicated to "inbound" AI, this policy does not have direct bearing on the use of that AI and thus would have limited, if any, impact on the misuse of open-source AI by third-party actors.

"Backdoor" risks

This policy would mitigate backdoor risks, but imperfectly. The primary goal of this policy is to prevent backdoor risks in open-source AI; focusing on model traceability and security addresses a key vulnerability in the AI development pipeline. Mandating audits increases the likelihood that vulnerabilities will be detected, and regular screenings would help detect evolving threats. Regarding China, this policy would ideally catch any attempted state interference with open-source AI and thus mitigate attempts to increase state power through open-source AI interference with government agencies. However, it does not guarantee zero risk; the quality of the framework and capabilities of the auditor are key, particularly when it comes to sophisticated backdoors. This policy also has limited benefit to the individuals and companies outside the government contractor ecosystem that may be subject to backdoor risks.

Global power

This policy would have limited effect on the dominance of American open-source AI in the global landscape. It could potentially create strains with certain international government contractors, but this is unlikely, since many of the relevant AI companies are American.

Weighing the implications

While this policy would hopefully mitigate backdoor risks for the government, it may limit which open models and frameworks are considered secure enough to deploy in a US government setting, hampering specific-use applications. It also could place a greater compliance burden on smaller contractors leveraging open-source AI, potentially preventing the US government from using otherwise productive technology. How these considerations balance depends on whether policymakers are prioritizing mitigation of risk in government or government access to the full spectrum of available technologies. Since audits would not be public, this policy does not provide any benefit to companies or individuals who may be subject to backdoor risks.

Table 5: The implications of KYS for government use of open models

Considerations		Implications
Ideological	**Accelerated technological progress**	• **Somewhat disruptive** to specific-use applications in government if advanced open models are found to be insecure. • **Not disruptive** to frontier capabilities. • **Potential barriers to entry for** newer and smaller government contractors
	Increased power distribution	• **Minimally disruptive** and **potentially beneficial** by requiring more actors to get involved in auditing AI models.
	Increased transparency	• **Beneficial** by increasing the number of external evaluations, although transparency issues remain.
Geopolitical	**Misuse of open-source AI**	• **Neutral**
	"Backdoor" risks	• **Imperfectly mitigates** risks by mandating audits, but auditing is not a guarantee of catching backdoor risks.
	Global power	• **No effect** on global usage of American open-source AI

Policy 2b: Open-source audits

While Policy 2a could mitigate backdoor risks, it only makes audits available to those contractors and government agencies using the products, meaning that the public cannot benefit from these audits. This policy would address the need for greater public resources on managing open-source AI risks. CISA has recognized this as an issue, acknowledging the need to work with the open source community and develop frameworks to prioritize project risks.[200]

This policy would create an open and regularly updated repository of audits. Unlike Policy 2a, this repository does not necessarily need to be a government-led project. It could also be provided by an independent industry body, non-profit organization, or it could be community-maintained. Given that CISA has called for developing "a process to continuously assess threats to critical OSS dependencies,"[200] the federal government may have an interest in maintaining this repository.

Like Policy 2a, these audits would follow a specific format based on what kind of package is being audited. Any entity looking to use a covered model or key package could consult the repository, which would provide a comprehensive security audit and flag any identified security issues to take into consideration for their specific use cases. Ideally, identified vulnerabilities would also be reported to the CVE Project[201] to link the audit repository to the existing open-source software vulnerability identification and mitigation infrastructure. While the repository maintainer would have primary responsibility for performing the audits, other actors could contribute by flagging other suspected issues, and this approach would facilitate fixes by the open source community. Packages would be risk-ranked, with some threshold for when use is not recommended; this could be based on the existing Common Vulnerability Scoring System (CVSS).[202] If the package has not already been audited, the agency in charge of the repository would be responsible for auditing it in good time.

The success of this project is dependent on sufficient resources to maintain the repository and, as mentioned above, for models and frameworks to be sufficiently open to enable audits. As with Policy 2a, audits are only possible if all model components are available; companies may be less willing to make model components available for a public repository than under government contracts. Since it is simply a public resource and does not prevent users from leveraging unaudited products, it would be less disruptive to government applications, but could be less effective in mitigating misuse risk in a government context.

Accelerated technological progress

Compared to Policy 2a, Policy 2b would be less disruptive for government agencies. By removing the requirement for government agencies to have audited open models or frameworks, there is greater opportunity for agencies to use a wider array of products.

Outside of government agencies, a public repository of open-source audits could provide an important verification step for other users, similar to the CVE Project. By creating trust in reliable models, this could lead to greater uptake and potentially further innovation. However, if advanced open models are found to be insecure, it may lead to limited uptake and potentially disrupt progress on other specific-use applications.

Increased power distribution

Policy 2b would enable more users to engage with audits of open models and frameworks than Policy 2a, increasing participation in the open source community. While placing the repository under the control of a single entity would result in some concentration of power, it still allows for broader participation than contractor-led audits.

Increased transparency

Open-source audits would further increase transparency over Policy 2a by sharing the audit repository with the public. As mentioned in the discussion of accelerated technological progress, this transparency has clear benefits, meaning that the public repository is a form of public good. As with other public goods, there are open questions over maintenance and the risk of a "tragedy of the commons," but establishing a clear governance process would help avoid this.

Misuse of open-source AI

As a policy dedicated to assessing the risks of inbound AI, it would have limited impact on the misuse of open-source AI.

"Backdoor" risks

Like Policy 2a, Policy 2b would help mitigate risks from inbound open models, although it would not guarantee that risks would be caught. Open-sourcing the audits could make it more likely that vulnerabilities would be identified in non-government organizations. On the other hand, the open repository could provide adversaries with information on how to skirt detection. Furthermore, the repository is only beneficial if it remains uncompromised, and having a centralized resource would create a target for malicious actors.

Global power

This policy could drive global uptake of American open-source AI by increasing trust in the safety of these products. It could also enhance US influence by creating trusted standards for open-source auditing that other countries could adopt and build confidence in US-built AI technology, as many of the audits would be relevant to private-sector development as well.

Weighing the implications

Compared to Policy 2a, a repository of model audits would provide greater benefit for the public and grant government agencies with more discretion to use open models. However, this may be accompanied with greater likelihood of backdoor risks. The success of this policy depends on multiple factors: sufficient transparency to audit models, resources to maintain repository, and ability to prevent the repository becoming compromised.

Table 6: The implications of open-source audits

Considerations		Implications
Ideological	**Accelerated technological progress**	• **Somewhat disruptive** to specific-use applications if advanced open models are found to be insecure. • **Not disruptive** to frontier capabilities.
	Increased power distribution	• **Minimally disruptive** and **potentially beneficial** by requiring more actors to get involved in auditing AI models.
	Increased transparency	• **Beneficial** by increasing the number of external evaluations and making them open-source.
Geopolitical	**Misuse of open-source AI**	• **Neutral**
	"Backdoor" risks	• **Imperfectly mitigates** risks by mandating audits, but auditing is not a guarantee of catching backdoor risks.
	Global power	• **Slightly weakens** American power if overly burdensome on open-source development and use, but **less impactful** than non-open-sourced audits. • Could **benefit** American power if it fosters confidence in US-built open-source AI.

Limitations and further research

While this analysis reveals the relative strengths and weaknesses of different policy approaches, it also highlights several areas where our understanding remains incomplete. Further research and monitoring will be crucial for refining these policy recommendations and adapting them to evolving technological capabilities. Our assessment of each policy was partially informed by our current understanding of the marginal risk associated with open models. Over time, marginal risk could evolve, particularly if there are significant shifts in the performance of open models relative to closed models. We recommend that researchers and government bodies track the following areas over the coming years.

First, we recommend tracking how open models are performing relative to closed models. DeepSeek's success indicates that the estimated one-year capabilities gap between open and closed models[24] may be closing. We suggest that this analysis is updated periodically. If American open models begin to keep pace with the performance of closed models, then there may be a higher marginal risk associated with making American models open, depending on existing capabilities in China and among other potentially malicious actors. For example, if a new open model is released that has frontier capabilities unseen in closed models, then it could provide a substantial capability uplift to malicious actors. In that situation, closer examination of the benefits and disadvantages of export controls or other risk mitigations may be worthwhile.

Second, where algorithmic innovations are originating and what drives them should be monitored. There is some ambiguity whether export controls on open models would affect the development of frontier capabilities. Although capability improvements have been largely due to scaling and led by closed models, these closed models could also be leveraging algorithmic innovations from the open source community. Indeed, this may accelerate in the wake of DeepSeek-R1. Still, the more that algorithmic innovations stem from closed models, the less disruptive export controls on open models would be to American frontier capabilities. Furthermore, assessing the relative role of increased compute versus algorithmic innovation in capability uplifts would help assess the effectiveness of existing semiconductor chip export controls and inform policy for open models.

Third, track the popularity and performance of Chinese and American open models. Although there are existing attempts to compare the performance of models, there is a deficit of sufficiently objective benchmarks to compare Chinese and American models. Most existing benchmarks suffer from some degree of language bias. In other words, using an English-only benchmark like SuperGLUE[203] or a Chinese benchmark like SuperCLUE[204] to assess both Chinese and American models is not a level comparison. Thus, it would be

valuable to develop a more objective measure of performance for comparison and to track this over time, such as by expanding the Chatbot Arena and other metrics. Furthermore, it is worth tracking the popularity of Chinese models, either through the number of downloads or the number of derivative models, as data on which countries are leveraging American and Chinese open models would be a helpful proxy for global power and technological influence.

Finally, we highlight the need for greater research into technical safety mitigations for open models. While there have been some papers on anti-tamper safety training,[205] this field requires further study. Technical safety mitigations could offer a more precise form of risk mitigation compared to export controls, limiting disruption to innovation while more effectively mitigating the risk of misuse. We suggest greater investment into researching technical mitigations against misuse.

Conclusion

The debate over open-source AI regulation sits at the intersection of innovation, security, and great power competition. Our analysis reveals the complex trade-offs between preserving the benefits of open-source AI—promoting technological innovation, distributed power, and increasing transparency—and addressing legitimate national security concerns around potential misuse, backdoor risks, and global power dynamics.

Several key findings emerge from our examination of policy options. First, heavy-handed regulation through export controls would likely be counterproductive, creating substantial disruption to innovation while providing only imperfect protection against security risks. The global, information-based nature of open-source AI makes traditional export control frameworks particularly challenging to implement effectively. A more workable alternative would be independent pre-release audits addressing specific risks (Policy 1b).

Second, more targeted approaches also show promise addressing inbound risks. In particular, the creation of an open repository of security audits (Policy 2b) could help address security concerns while maintaining the spirit of open-source development. This approach balances security needs with innovation benefits better than individual contractor-led audits, though neither approach completely eliminates security risks.

Third, the current performance gap between open and closed models suggests that the marginal security risk from open-source AI may be lower than commonly assumed. However, this dynamic could shift as open-source capabilities advance, highlighting the importance of ongoing monitoring and assessment of relative capabilities.

Looking ahead, several areas require further attention from researchers and policymakers:

- Tracking the evolving performance gap between open and closed models.
- Understanding the sources and flow of algorithmic innovations.
- Developing more objective benchmarks to compare Chinese and American models.
- Advancing technical approaches to model security for durable risk mitigation.

The US approach to open-source AI will have lasting implications for both technological innovation and global influence. Rather than viewing this as a binary choice between complete openness and restriction, policymakers should pursue targeted interventions that preserve the benefits of open-source development while addressing specific security risks. Such an approach can help maintain US technological leadership while fostering responsible innovation in AI development.

Appendix: Derivation of ideological considerations

Our ideological considerations were derived by fusing concepts from two pieces of research. A white paper from the Centre for the Governance of AI (GovAI)[1] identifies "external model evaluation," "accelerate beneficial AI progress," and "distribute control over AI" as three key benefits. A policy brief published in *Science*[9] focused on three "fundamental societal objectives": "ensuring transparency," "catalyzing innovation," and "distributing power." Both publications have converged on three broad benefits, which we here classify as "increased transparency," "accelerated AI progress," and "increased power distribution." These "derived benefits" form our ideological considerations in Section 4.

Table 7: Traditional and derived benefits of open-source AI

GovAI benefit	*Science* benefit	Derived benefit	Explanation
Accelerate beneficial AI progress	Catalyzing innovation	**Accelerated technological progress**	Open AI models can be widely implemented for different tasks, and allow for more contributors to push AI progress.
External model evaluation	Ensuring transparency	**Increased transparency**	Providing more information about models allows for better examination of model capabilities and risks, including external auditing in some cases.
Distribute control over AI	Distributing power	**Increased power distribution**	Open AI models allow for more people to provide input on the direction of AI and prevent single actors from exerting total control.

Citations

1. Seger, E. *et al.* Open-Sourcing Highly Capable Foundation Models: An Evaluation of Risks, Benefits, and Alternative Methods for Pursuing Open-Source Objectives. *SSRN Electron. J.* (2023) doi:10.2139/ssrn.4596436.
2. Hoffmann, M., Nagle, F. & Zhou, Y. The Value of Open Source Software. *SSRN Electron. J.* (2024) doi:10.2139/ssrn.4693148.
3. Williams, R. & O'Donnell, J. We finally have a definition for open-source AI. *MIT Technology Review* https://www.technologyreview.com/2024/08/22/1097224/we-finally-have-a-definition-for-open-source-ai/ (2024).
4. The Open Source Definition. *Open Source Initiative* https://opensource.org/osd (2024).
5. Widder, D. G., West, S. & Whittaker, M. Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.4543807 (2023).
6. Botton, N. & Vermeulen, M. *Generative AI's Open Source Challenge*. (2024).
7. The Open Source AI Definition – 1.0. *Open Source Initiative* https://opensource.org/ai/open-source-ai-definition (2024).
8. Waters, R. Meta under fire for 'polluting' open-source. *Financial Times* (2024).
9. Bommasani, R. *et al.* Considerations for governing open foundation models. *Science* **386**, 151–153 (2024).
10. Vaughan-Nichols, S. Meta can call Llama 2 open source as much as it likes, but that doesn't mean it is. *The Register* https://www.theregister.com/2023/07/21/llama_is_not_open_source/ (2023).
11. Bonheur, K. Difference Between Architecture, Algorithm, and Model in AI. *Profolus* https://www.profolus.com/topics/difference-between-architecture-algorithm-and-model-in-ai/ (2024).
12. Mikulski, M. AI Libraries – What are the Differences? *Fingoweb* https://www.fingoweb.com/blog/what-are-the-differences-between-ai-libraries-and-how-to-utilize-them/ (2024).
13. Stephanie Palazzolo. Meta's Free AI Isn't Cheap to Use, Companies Say. *The Information* https://www.theinformation.com/articles/metas-free-ai-isnt-cheap-to-use-companies-say (2023).
14. OpenAI. ChatGPT Pricing. https://openai.com/chatgpt/pricing/.
15. Nolan, B. It will be awhile before OpenAI turns a profit. In the meantime, Microsoft gets a cut of revenue. *Business Insider* https://www.businessinsider.com/openai-profit-funding-ai-microsoft-chatgpt-revenue-2024-10 (2024).
16. Devansh. Understanding the Business of Open Source Software and AI. *Medium* https://machine-learning-made-simple.medium.com/understanding-the-business-of-open-source-software-and-ai-0aa43a480450 (2024).
17. Lin, B. Open-Source Companies Are Sharing Their AI Free. Can They Crack OpenAI's Dominance? *Wall Street Journal* (2024).
18. Vanian, J. Meta's unique approach to developing AI puzzles Wall Street, but techies love it. *CNBC* https://www.cnbc.com/2023/10/16/metas-open-source-approach-to-ai-puzzles-wall-street-techies-love-it.html (2023).
19. Tozzi, C. The importance and limitations of open source AI models. *TechTarget* https://www.techtarget.com/searchenterpriseai/tip/The-importance-and-limitations-of-open-source-AI-models (2024).
20. Pillay, T. The Gap Between Open and Closed AI Models Might Be Shrinking. Here's Why That Matters. *TIME* https://time.com/7171962/open-closed-ai-models-epoch/ (2024).
21. Nuñez, M. Open-source AI narrows gap with proprietary leaders, new benchmark reveals. *VentureBeat* https://venturebeat.com/ai/open-source-ai-narrows-gap-with-tech-giants-new-benchmark-reveals/ (2024).
22. *Artificial Intelligence Index Report* 2024. 16 https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf (2024).
23. Kalley Huang. Open-Source AI Struggles to Close Gap with Closed Rivals. *The Information* https://www.theinformation.com/articles/

open-source-ai-struggles-to-close-gap-with-closed-rivals (2024).
24. Cottier, B., You, J., Martemianova, N. & Owen, D. How Far Behind Are Open Models? *Epoch AI* https://epoch.ai/blog/open-models-report (2024).
25. DeepSeek-AI *et al.* DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. Preprint at https://doi.org/10.48550/arXiv.2501.12948 (2025).
26. Zachary Brown, Peter Slattery, & Haoran Lyu. What drives progress in AI? Trends in Algorithms. *MIT FutureTech* https://futuretech.mit.edu/news/what-drives-progress-in-ai-trends-in-algorithms (2024).
27. Peng, T. DeepSeek-R1 and Kimi k1.5: How Chinese AI Labs Are Closing the Gap with OpenAI's o1. *Recode China AI* https://recodechinaai.substack.com/p/deepseek-r1-and-kimi-k15-how-chinese?publication_id=302506&utm_campaign=email-post-title&r=49prte&utm_medium=email (2025).
28. Julia Shapero. OpenAI investigating whether DeepSeek improperly obtained data. *The Hill* https://thehill.com/policy/technology/5113470-openai-deepseek-data-theft/ (2025).
29. Maracke, C. Free and Open Source Software and FRAND-based patent licenses. *J. World Intellect. Prop.* **22**, 78–102 (2019).
30. Nussbaum, J. L. Apple Computer, Inc. v. Franklin Computer Corporation Puts the Byte Back into Copyright Protection for Computer Programs. *Gold. Gate Univ. Law Rev.* **14**, (1984).
31. History of the OSI. *Open Source Initiative* https://opensource.org/history (2018).
32. Ceraso, A. & Pruchnic, J. Introduction: Open Source Culture and Aesthetics. *Criticism* **53**, 337–375 (2011).
33. PyTorch Governance | Maintainers. *PyTorch 2.5 documentation* https://pytorch.org/docs/stable/community/persons_of_interest.html.
34. TensorFlow | Google Open Source Projects. *Google Open Source* https://opensource.google/projects/tensorflow.
35. Branscombe, M. What is Microsoft doing with Linux? Everything you need to know about its plans for open source. *TechRepublic* https://www.techrepublic.com/article/what-is-microsoft-doing-with-linux-everything-you-need-to-know-about-its-plans-for-open-source/ (2020).
36. IBM. IBM Completes Acquisition of Red Hat. *IBM Investor Relations* https://www.ibm.com/investor/news/ibm-completes-acquisition-of-red-hat (2019).
37. Vassilieva, N. Cerebras Makes It Easy to Harness the Predictive Power of GPT-J. *Cerebras* https://cerebras.ai/blog/cerebras-makes-it-easy-to-harness-the-predictive-power-of-gpt-j/ (2022).
38. Scao, T. L. *et al.* BLOOM: A 176B-Parameter Open-Access Multilingual Language Model. Preprint at https://doi.org/10.48550/arXiv.2211.05100 (2023).
39. Meta. Introducing LLaMA: A foundational, 65-billion-parameter language model. *Meta* https://ai.meta.com/blog/large-language-model-llama-meta-ai/ (2023).
40. Jeanine Banks & Tris Warkentin. Gemma: Introducing new state-of-the-art open models. *Google* https://blog.google/technology/developers/gemma-open-models/ (2024).
41. Arcesati, R. *China's AI Development Model in an Era of Technological Deglobalization.* https://merics.org/en/report/chinas-ai-development-model-era-technological-deglobalization (2024).
42. GitHub Staff. Octoverse: AI leads Python to top language as the number of global developers surges. *The GitHub Blog* https://github.blog/news-insights/octoverse/octoverse-2024/ (2024).
43. Liu, P. & Ding, Y. China debates pros and cons of open-source AI models. *ThinkChina* https://www.thinkchina.sg/technology/china-debates-pros-and-cons-open-source-ai-models (2024).
44. Yang, Z. Why Chinese companies are betting on open-source AI. *MIT Technology Review* https://www.technologyreview.com/2024/07/24/1095239/chinese-companies-open-source-ai/ (2024).
45. Triolo, P. & Schaefer, K. *China's Generative AI Ecosystem in 2024: Rising Investment and Expectations.* https://www.nbr.org/publication/chinas-generative-ai-ecosystem-in-2024-rising-investment-and-expectations/ (2024).
46. Patel, D., Kourabi, A., O'Laughlin, D. & Knuhtsen, R. DeepSeek Debates: Chinese Leadership On Cost, True Training Cost,

Closed Model Margin Impacts. *SemiAnalysis* https://semianalysis.com/2025/01/31/deepseek-debates/ (2025).
47. Sharma, S. DeepSeek-V3, ultra-large open-source AI, outperforms Llama and Qwen on launch. *VentureBeat* https://venturebeat.com/ai/deepseek-v3-ultra-large-open-source-ai-outperforms-llama-and-qwen-on-launch/ (2024).
48. Kaye, K. Will nationalism end global open-source AI collaboration? *Protocol* https://www.protocol.com/enterprise/china-us-ai-open-source (2022).
49. Ding, J. *Techno-Industrial Policy for New Infrastructure: China's Approach to Promoting Artificial Intelligence as a General Purpose Technology*.
50. Webster, G., Creemers, R., Kania, E. & Triolo, P. China's 'New Generation Artificial Intelligence Development Plan: Full Translation. *DigiChina* https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/ (2017).
51. Murphy, B. CSET Original Translation: China's 14th Five-Year Plan. *Center for Security and Emerging Technology* https://cset.georgetown.edu/publication/china-14th-five-year-plan/.
52. China Academy of Information and Communications Technology. White Paper on AI Framework Development (2022). *Center for Security and Emerging Technology* https://cset.georgetown.edu/publication/ai-frameworks-white-paper/ (2023).
53. Ding, J. *China's Uncharacteristic Approach to Artificial Intelligence* (AI) *Development*. https://ucigcc.org/blog/chinas-uncharacteristic-approach-to-artificial-intelligence-ai-development/ (2023).
54. Omaar, H. *How Innovative Is China in AI?* https://itif.org/publications/2024/08/26/how-innovative-is-china-in-ai/ (2024).
55. 关于AI和大模型，工信部、上海市、商汤科技等重磅发声. https://baijiahao.baidu.com/s?id=1794324841403169804&wfr=spider&for=pc.
56. 张兴华. 推动共研共创！加快开源基础设施建设 赋能数字经济发展_部门动态_中国政府网. https://www.gov.cn/lianbo/bumen/202409/content_6976505.htm (2024).
57. TechNode Feed. China now has over 180 LLMs approved for general use. *TechNode* http://technode.com/2024/08/13/china-now-has-over-180-llms-approved-for-general-use/ (2024).
58. Bajarin, T. Unpacking Apple's AI Roadblocks In The Chinese Market. *Forbes* https://www.forbes.com/sites/timbajarin/2024/11/27/unpacking-apples-ai-roadblocks-in-the-chinese-market/ (2024).
59. Ministry of Industry and Information Technology. 推进大模型赋能网络安全. *Weixin Official Accounts Platform* https://mp.weixin.qq.com/s/PEcMZzhzUYOaeaSL4iW1wg.
60. Peng, T. 🦙What Llama 3 Means to China, ERNIE Bot Hits 200 Million Users, and China Trails US in AI Models. *Recode China AI* https://recodechinaai.substack.com/p/what-llama-3-means-to-china-ernie (2024).
61. 张进. 别再说国产大模型技术突破要靠 Llama 3 开源了. 雷峰网 https://www.leiphone.com/category/ai/GLXVgbyWGiPmqf2o.html.
62. 科技时坛. 中国AI难以追赶美国的先进步伐？. 网易 https://www.163.com/dy/article/ITDQNUQG05562Z7A.html (2024).
63. Ye, J. How dependent is China on US artificial intelligence technology? *Reuters* (2024).
64. Mozur, P., Liu, J. & Metz, C. China's Rush to Dominate A.I. Comes With a Twist: It Depends on U.S. Technology. *The New York Times* (2024).
65. Schneider, J., Shen, A. & Zhang, I. Deepseek: The Quiet Giant Leading China's AI Race. *ChinaTalk* https://www.chinatalk.media/p/deepseek-ceo-interview-with-chinas (2023).
66. Knight, W. This Chinese Startup Is Winning the Open Source AI Race. *Wired* (2024).
67. Pomfret, J. & Pang, J. Exclusive: Chinese researchers develop AI model for military use on back of Meta's Llama. *Reuters* (2024).
68. Roth, E. Meta AI is ready for war. *The Verge* (2024).
69. Wu, J. Stanford AI project authors apologize for plagiarizing Chinese team. *TechNode* http://technode.com/2024/06/04/stanford-ai-project-authors-apologize-for-plagiarizing-chinese-large-language-model/ (2024).
70. openbmb. MiniCPM-Llama3-V-2_5. (2025).
71. Hugging Face. abacusai/Smaug-72B-v0.1. *Hugging Face* https://huggingface.co/abacusai/Smaug-72B-v0.1.

72. Nuñez, M. Meet 'Smaug-72B': The new king of open-source AI. *VentureBeat* https://venturebeat.com/ai/meet-smaug-72b-the-new-king-of-open-source-ai/ (2024).
73. Sharma, S. Meet 'Liberated Qwen', an uncensored LLM that strictly adheres to system prompts. *VentureBeat* https://venturebeat.com/ai/meet-liberated-qwen-an-uncensored-llm-that-strictly-adheres-to-system-prompts/ (2024).
74. China's AI industry has almost caught up with America's. *The Economist* (2025).
75. Apostoaie, E. Chinese AI Companies Are Catching Up Despite U.S. Restrictions. *The Wire China* https://www.thewirechina.com/2024/12/08/chinese-ai-companies-are-catching-up-despite-u-s-restrictions-chinas-ai-models/ (2024).
76. Thompson, C. How a Nation of Tech Copycats Transformed Into a Hub for Innovation. *Wired* (2015).
77. Booth, H. How China Is Advancing in AI Despite U.S. Chip Restrictions. *TIME* https://time.com/7204164/china-ai-advances-chips/ (2025).
78. DeepSeek-AI *et al.* DeepSeek-V3 Technical Report. Preprint at https://doi.org/10.48550/arXiv.2412.19437 (2024).
79. Anthropic. Model Card Addendum: Claude 3.5 Haiku and Upgraded Claude 3.5 Sonnet. Preprint at https://assets.anthropic.com/m/1cd9d098ac3e6467/original/Claude-3-Model-Card-October-Addendum.pdf.
80. Kimi Team. Kimi k1.5: Scaling Reinforcement Learning with LLMs. https://arxiv.org/html/2501.12599v1.
81. OpenAI. OpenAI o1-mini. https://openai.com/index/openai-o1-mini-advancing-cost-efficient-reasoning/ (2024).
82. Chiang, W.-L. *et al.* Chatbot Arena: An Open Platform for Evaluating LLMs by Human Preference. Preprint at https://doi.org/10.48550/arXiv.2403.04132 (2024).
83. Epoch AI. Machine Learning Trends. *Epoch AI* https://epoch.ai/trends (2024).
84. 大模型应用. Llama 3.1对我国AIGC产业发展的启示_llama开源. *CSDN* https://blog.csdn.net/y525698136/article/details/141437236 (2024).
85. Rosenberg, S. Censorship is slowing AI progress in China. *Axios* https://www.axios.com/2024/07/19/china-ai-race-government-censorship (2024).
86. Hine, E. & Floridi, L. Artificial intelligence with American values and Chinese characteristics: a comparative analysis of American and Chinese governmental AI policies. *AI Soc.* (2022) doi:10.1007/s00146-022-01499-8.
87. Hine, E. Governing Silicon Valley and Shenzhen: Assessing a New Era of Artificial Intelligence Governance in the United States and China. *Digit. Soc.* **3**, (2024).
88. Hugging Face. Most liked and downloaded models, from 2022 to today. *Hugging Face* https://huggingface-open-source-ai-year-in-review-2024.static.hf.space/index.html (2024).
89. Ding, J. Why is (Alibaba's) Tongyi the most popular open-source large model. *ChinAI* https://docs.google.com/document/d/1SqhEKdKg8MMPgS_KzKPKv8vi8pBKEzD9a5Pnq38bHjI/edit?tab=t.0&usp=embed_facebook.
90. 张进. 最受欢迎开源大模型，为什么是通义？. *Weixin Official Accounts Platform* https://mp.weixin.qq.com/s/6AvZLtydqnnuj5G7mGQ7ng (2024).
91. Kharpal, A. China's Alibaba launches over 100 new open-source AI models, releases text-to-video generation tool. *CNBC* https://www.cnbc.com/2024/09/19/alibaba-launches-over-100-new-ai-models-releases-text-to-video-generation.html (2024).
92. Hugging Face. Hub API Endpoints. *Hugging Fave* https://huggingface.co/docs/hub/en/api. Derivative models are models that mention the base model (e.g., llama) in their base model or model type.
93. de Soyres, F. & Moore, D. Assessing China's efforts to increase self-reliance. *CEPR* https://cepr.org/voxeu/columns/assessing-chinas-efforts-increase-self-reliance (2024).
94. Roberts, H., Hine, E. & Floridi, L. Digital Sovereignty, Digital Expansionism, and the Prospects for Global AI Governance. in *Quo Vadis, Sovereignty?* (eds. Timoteo, M., Verri, B. & Nanni, R.) (Springer, 2023).
95. Creemers, R. China's Conception of Cyber Sovereignty: Rhetoric and Reality. in *Governing Cyberspace: Behavior, Power and Diplomacy* (eds. Broeders, D. & Berg, B. van den) (Rowman & Littlefield, 2020).

96. Chander, A. & Sun, H. Sovereignty 2.0. *Georget. Law Fac. Publ. Works* (2021).
97. Kevin Xu [@kevinsxu]. Challenge 1: no self sufficiency in model architecture GPT-series is proprietary, and most Chinese models are built by leveraging open source LLaMA (This over-reliance on LLaMA is seen as a severe problem, not some shortcut to leapfrog the US) https://t.co/ayiKlorNSt. *Twitter* https://x.com/kevinsxu/status/1768365480509861888 (2024).
98. National Telecommunications and Information Administration. NTIA Receives More Than 300 Comments on Open Weight AI Models. *National Telecommunications and Information Administration* (2024).
99. Rep. McCaul, M. T. [R-T.-10. ENFORCE Act. (2024).
100. Ding, J. ChinAI #296: DeepSeek goes left, ModelBest goes right. *ChinAI Newsletter* https://chinai.substack.com/?utm_medium=email (2025).
101. O'Shaughnessy, M. & Sheehan, M. Lessons From the World's Two Experiments in AI Governance. *Carnegie Endowment for International Peace* https://carnegieendowment.org/posts/2023/02/lessons-from-the-worlds-two-experiments-in-ai-governance?lang=en (2023).
102. Executive Office of the President. Executive Order 13859: Maintaining American Leadership in Artificial Intelligence. *Federal Register* https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence (2019).
103. Executive Office of the President. Executive Order 14110 of October 30, 2023: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. (2023).
104. Bureau of Industry and Security, Department of Commerce. Framework for Artificial Intelligence Diffusion. *Federal Register* https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion (2025).
105. Committee on Foreign Affairs. Chairman McCaul's ENFORCE Act Passes Out of Committee with Broad Bipartisan Support 43-3. *Committee on Foreign Affairs* https://foreignaffairs.house.gov/press-release/chairman-mccauls-enforce-act-passes-out-of-committee-passes-with-broad-bipartisan-support-43-3/ (2024).
106. Wiener, S. *Safe and Secure Innovation for Frontier Artificial Intelligence Models Act*. SB 1047 (2024).
107. Kohler, S. All Eyes on Sacramento: SB 1047 and the AI Safety Debate. *Carnegie Endowment for International Peace* https://carnegieendowment.org/posts/2024/09/california-sb1047-ai-safety-regulation?lang=en (2024).
108. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA Relevance). OJ L*, 2024/1689 (2024).
109. Coulter, M., Mukherjee, S., Chee, F. Y., Mukherjee, S. & Chee, F. Y. EU's AI Act could exclude open-source models from regulation. *Reuters* (2023).
110. Bertuzzi, L. AI Act: EU policymakers nail down rules on AI models, butt heads on law enforcement. *Euractiv* https://www.euractiv.com/section/artificial-intelligence/news/ai-act-eu-policymakers-nail-down-rules-on-ai-models-butt-heads-on-law-enforcement/ (2023).
111. *Artificial Intelligence (Regulation) Bill [HL]*. (2023).
112. Sun, Y. & Zeng, J. Unveiling China's Generative AI Regulation. *https://fpf.org/* https://fpf.org/blog/unveiling-chinas-generative-ai-regulation/.
113. Ministry of Industry and Information Technology. *Hulianwang Xinxi Fuwu Shendu Hecheng Guanli Guiding [Provisions on the Administration of Deep Synthesis Internet Information Services]*. (2021).
114. China Law Translate. 互联网信息服务深度合成管理规定 （征求意见稿）- Provisions on the Administration of Deep Synthesis Internet Information Services (Draft for solicitation of comments). *China Law Translate* https://www.chinalawtranslate.com/deep-synthesis-draft/ (2022).
115. Ministry of Industry and Information Technology. *Hulianwang Xinxi Fuwu Suanfa Tuijian Guanli Guiding [Internet Information*

Service Algorithmic Recommendation Management Provisions]. (2021).
116. China Law Translate. 互联网信息服务算法推荐管理规定 （征求意见稿）- Provisions on the Administration of Internet Information Service Algorithmic Recommendation (Draft for Solicitation of Comments). *China Law Translate* https://www.chinalawtranslate.com/algorithm-regulation-draft/ (2021).
117. China Law Translate. Interim Measures for the Management of Generative Artificial Intelligence Services. *China Law Translate* https://www.chinalawtranslate.com/generative-ai-interim/ (2023).
118. Cyberspace Administration of China. *Shengchengshi Rengong Zhineng Fuwu Guanli Zanxing Banfa [Interim Measures for the Management of Generative Artificial Intelligence Services]*. (2023).
119. Ding, J. The diffusion deficit in scientific and technological power: re-assessing China's rise. *Rev. Int. Polit. Econ.* **31**, 173–198 (2024).
120. Mark Zuckerberg. Open Source AI is the Path Forward. *Meta* https://about.fb.com/news/2024/07/open-source-ai-is-the-path-forward/ (2024).
121. Pilz, K., Heim, L. & Brown, N. Increased Compute Efficiency and the Diffusion of AI Capabilities. Preprint at https://doi.org/10.48550/arXiv.2311.15377 (2024).
122. OpenAI. AMA with OpenAI's Sam Altman, Mark Chen, Kevin Weil, Srinivas Narayanan, Michelle Pokrass, and Hongyu Ren. *r/OpenAI* www.reddit.com/r/OpenAI/comments/1ieonxv/ama_with_openais_sam_altman_mark_chen_kevin_weil/ (2025).
123. Bommasani, R. *et al.* The Foundation Model Transparency Index. Preprint at https://doi.org/10.48550/arXiv.2310.12941 (2023).
124. *Meta TV Spot, 'Open Source AI: Available to All'*. (iSpot).
125. Meta AI. Llama 2 - Acceptable Use Policy. *Meta* https://ai.meta.com/llama-project/use-policy.
126. Maya Wang. China's Techno-Authoritarianism Has Gone Global. *Human Rights Watch* https://www.hrw.org/news/2021/04/08/chinas-techno-authoritarianism-has-gone-global (2021).
127. Lamensch, M. Evolving Surveillance Tech Whets the Authoritarian Impulse to See and Know All. *Centre for International Governance Innovation* https://www.cigionline.org/articles/evolving-surveillance-tech-whets-the-authoritarian-impulse-to-see-and-know-all/ (2023).
128. Human Rights Watch. *"Break Their Lineage, Break Their Roots": China's Crimes against Humanity Targeting Uyghurs and Other Turkic Muslims.* https://www.hrw.org/report/2021/04/19/break-their-lineage-break-their-roots/chinas-crimes-against-humanity-targeting (2021).
129. Gorokhovskaia, Y. & Linzer, I. *Transnational Repression | China Case Study.* https://freedomhouse.org/sites/default/files/2021-02/FH_TransnationalRepressionReport2021_rev020221_CaseStudy_China.pdf (2022).
130. Schoelkopf, H., Skowron, A. & Biderman, S. Yi-34B, Llama 2, and common practices in LLM training: a fact check of the New York Times. *EleutherAI Blog* https://blog.eleuther.ai/nyt-yi-34b-response/ (2024).
131. David, E. Baidu launches Ernie chatbot after Chinese government approval. *The Verge* https://www.theverge.com/2023/8/31/23853878/baidu-launch-ernie-ai-chatbot-china (2023).
132. Lamensch, M. Authoritarianism Has Been Reinvented for the Digital Age. *Centre for International Governance Innovation* https://www.cigionline.org/articles/authoritarianism-has-been-reinvented-for-the-digital-age/ (2021).
133. Tamara Kharroub. Mapping Digital Authoritarianism in the Arab World. *Arab Center Washington DC* https://arabcenterdc.org/resource/mapping-digital-authoritarianism-in-the-arab-world/ (2025).
134. Tremayne-Pengelly, A. Anthropic Joins A.I. Giants to Provide Models to US Defense Agencies. *Observer* https://observer.com/2024/11/openai-rival-anthropic-provide-ai-models-dod/ (2024).
135. Borges, C. & Palazzi, A. L. The U.S.-China Relationship amid China's Economic Woes. *CSIS* https://www.csis.org/blogs/perspectives-innovation/us-china-relationship-amid-chinas-economic-woes (2023).

136. Emmet Lyons. DeepSeek AI raises national security concerns, U.S. officials say. *CBS News* https://www.cbsnews.com/news/deepseek-ai-raises-national-security-concerns-trump/ (2025).
137. Anaconda. *The State of Enterprise Open-Source AI*. https://www.anaconda.com/lp/state-of-enterprise-open-source-ai (2024).
138. Goodin, D. The XZ Backdoor: Everything You Need to Know. *WIRED* https://www.wired.com/story/xz-backdoor-everything-you-need-to-know/ (2024).
139. Greenberg, A. The Mystery of 'Jia Tan,' the XZ Backdoor Mastermind. *Wired*.
140. Tatam, R. XZ Utils, The Backdoor Exploit & What We Can Learn About Risk. *Puppet* https://www.puppet.com/blog/xz-backdoor (2024).
141. NIST. CVE-2024-3094 Detail. *National Vulnerability Database* https://nvd.nist.gov/vuln/detail/CVE-2024-3094 (2024).
142. Yasar, K., Gillis, A. S. & Bacon, M. What is a Common Vulnerability Scoring System (CVSS)? *TechTarget* https://www.techtarget.com/searchsecurity/definition/CVSS-Common-Vulnerability-Scoring-System.
143. Tung, L. Open source: Almost one in five bugs are planted for malicious purposes. *ZDNET* https://www.zdnet.com/article/open-source-software-how-many-bugs-are-hidden-there-on-purpose/ (2020).
144. Dan McInerney & Marcello Salvati. Protect AI's October 2024 Vulnerability Report. *Protect AI* https://protectai.com/threat-research/2024-october-vulnerability-report (2024).
145. IDOR Vulnerability in Prompt Update Function. *Sightline* https://sightline.protectai.com/vulnerabilities/88d7a4f7-fb4f-40ff-8c32-276befb2dd78/assess (2024).
146. IDOR Vulnerability Allowing View/Delete of External Users. *Sightline* https://sightline.protectai.com/vulnerabilities/a8580293-cbec-4e97-8b6f-aec2c557f8ea/assess (2024).
147. CSIS Strategic Technologies Program. *Survey of Chinese Espionage in the United States Since 2000*. https://www.csis.org/programs/strategic-technologies-program/survey-chinese-espionage-united-states-2000 (2023).
148. Sakellariadis, J. The White House struggles to contain massive Chinese telco hack. *POLITICO* https://www.politico.com/news/2024/12/04/chinese-telco-hacks-white-house-00192714 (2024).
149. Nakashima, E. *et al.* Top senator calls Salt Typhoon 'worst telecom hack in our nation's history'. *Washington Post* (2024).
150. McBride, K. Open Source AI: The Overlooked National Security Imperative. *Just Security* https://www.justsecurity.org/96422/open-source-ai-the-overlooked-national-security-imperative/ (2024).
151. Brooks, B. & Fang, M. US leadership in AI requires open-source diplomacy. *The Hill* https://thehill.com/opinion/technology/5079721-china-ai-open-source-threat/ (2025).
152. Leonard Lin. An Analysis of Chinese LLM Censorship and Bias with Qwen 2 Instruct. *Hugging Face* https://huggingface.co/blog/leonardlin/chinese-llm-censorship-analysis (2024).
153. Buyl, M. *et al.* Large Language Models Reflect the Ideology of their Creators. Preprint at https://doi.org/10.48550/arXiv.2410.18417 (2024).
154. Xu, K. Was Zuck Right about Chinese AI Models? *Interconnected* https://substack.com/home/post/p-154697687 (2025).
155. Torbjørn Flensted. DeepSeek AI Statistics and Facts (2025). *SEO.AI* https://seo.ai/blog/deepseek-ai-statistics-and-facts (2025).
156. Lee, J. Understanding and Countering China's Global South Strategy in the Indo-Pacific. *Hudson Institute* https://www.hudson.org/economics/understanding-countering-chinas-global-south-strategy-indo-pacific-john-lee (2024).
157. Lim, K. China is Using the Global South's Agitations to its Advantage: It's Time for the World to Pay Attention. *Pacific Forum* https://pacforum.org/publications/yl-blog-77-china-is-using-the-global-souths-agitations-to-its-advantage-its-time-for-the-world-to-pay-attention/ (2024).
158. Jinping, X. Combining the Great Strength of the Global South to Build Together a Community with a Shared Future for Mankind. *Ministry of Foreign Affairs of the People's Republic of China*

https://www.mfa.gov.cn/eng/xw/zyxw/202410/t20241024_11515589.html (2024).
159. Ideas, L. “Self-Othering” and “Neighboring”: Understanding China’s Global South Strategy. *The China-Global South Project* https://chinaglobalsouth.com/analysis/self-othering-and-neighboring-understanding-chinas-global-south-strategy/ (2024).
160. Wu Jieh-min. Silicon Shield 2.0: A Taiwan Perspective. https://thediplomat.com/2024/09/silicon-shield-2-0-a-taiwan-perspective/ (2024).
161. Walters, R. Losing Taiwan’s Semiconductors Would Devastate the US Economy. *Hudson Institute* https://www.hudson.org/technology/losing-taiwan-semiconductor-would-devastate-us-economy-riley-walters (2024).
162. Jones, L., Krulikowski, S., Lotze, N. & Schreiber, S. *U.S. Exposure to the Taiwanese Semiconductor Industry*. (2024).
163. Kine, P. Biden leaves no doubt: ‘Strategic ambiguity’ toward Taiwan is dead. *POLITICO* https://www.politico.com/news/2022/09/19/biden-leaves-no-doubt-strategic-ambiguity-toward-taiwan-is-dead-00057658 (2022).
164. Burgess, M. & Newman, L. H. DeepSeek’s Popular AI App Is Explicitly Sending US Data to China. *Wired* (2025).
165. Daniels, J. How Generative AI Can Affect Your Business’ Data Privacy. *Forbes* https://www.forbes.com/councils/forbesbusinesscouncil/2023/05/01/how-generative-ai-can-affect-your-business-data-privacy/ (2023).
166. OpenAI Help Center. How your data is used to improve model performance. *OpenAI* https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance.
167. Jennifer King & Caroline Meinhardt. Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World. *Stanford HAI* https://hai.stanford.edu/white-paper-rethinking-privacy-ai-era-policy-provocations-data-centric-world.
168. Lee, H.-P., Yang, Y.-J., Davier, T. S. von, Forlizzi, J. & Das, S. Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks. Preprint at https://doi.org/10.48550/arXiv.2310.07879 (2024).
169. Mirasole, C. Understanding China’s Cybersecurity Law. *Lawfare* https://www.lawfaremedia.org/article/understanding-chinas-cybersecurity-law (2016).
170. Zeff, M. DeepSeek displaces ChatGPT as the App Store’s top app. *TechCrunch* https://techcrunch.com/2025/01/27/deepseek-displaces-chatgpt-as-the-app-stores-top-app/ (2025).
171. Chow, A. R. Why DeepSeek is Sparking TikTok-Like National Security Fears. *TIME* https://time.com/7210875/deepseek-national-security-threat-tiktok/ (2025).
172. Connatser, M. Chinese-made DeepSeek AI model records extensive online user data, stores it in China-based servers. *Tom’s Hardware* https://www.tomshardware.com/tech-industry/artificial-intelligence/chinese-made-deepseek-ai-model-collects-extensive-user-data-stores-it-on-china-based-servers (2025).
173. Wang, V. How Does DeepSeek’s A.I. Chatbot Navigate China’s Censors? Awkwardly. *The New York Times* (2025).
174. Navlakha, M. Google Bard, ChatGPT: Are AI chatbots suppressing information about Israel and Palestine? *Mashable* https://mashable.com/article/ai-chatbot-israel-palestine-chatgpt-google-bard (2023).
175. Patrick Tucker. How DeepSeek changed the future of AI—and what that means for national security. *Defense One* https://www.defenseone.com/technology/2025/01/how-deepseek-changed-future-aiand-what-means-national-security/402594/ (2025).
176. Clegg, N. Open Source AI Can Help America Lead in AI and Strengthen Global Security. *Meta* https://about.fb.com/news/2024/11/open-source-ai-america-global-security/ (2024).
177. Wilson, C. AI Safety and the US-China Arms Race. *Center for AI Policy* https://www.centeraipolicy.org/work/ai-safety-and-the-us-china-arms-race (2024).
178. David Sacks [@DavidSacks]. DeepSeek R1 shows that the AI race will be very competitive. *Twitter* https://x.com/DavidSacks/status/1883935713877782884 (2025).
179. Gillham, J. Hugging Face Statistics. *Originality.AI* https://originality.ai/blog/huggingface-statistics (2024).
180. Marshall, M. The enterprise verdict on AI models: Why open source will win.

VentureBeat https://venturebeat.com/ai/the-enterprise-verdict-on-ai-models-why-open-source-will-win/ (2024).
181. Olcott, E., Sevastopulo, D. & Liu, Q. Chinese AI groups use cloud services to evade US chip export controls. *Financial Times* (2023).
182. The Select Committee of the CCP. Moolenaar Urges Raimondo to Close Dangerous Loopholes in New Export Control Rules. http://selectcommitteeontheccp.house.gov/media/press-releases/moolenaar-urges-raimondo-close-dangerous-loopholes-new-export-control-rules (2024).
183. Benson, E. Updated October 7 Semiconductor Export Controls. *CSIS* https://www.csis.org/analysis/updated-october-7-semiconductor-export-controls (2023).
184. Fist, T., Heim, L. & Schneider, J. Chinese Firms Are Evading Chip Controls. *Foreign Policy* https://foreignpolicy.com/2023/06/21/china-united-states-semiconductor-chips-sanctions-evasion/ (2023).
185. Seger, E., Ovadya, A., Garfinkel, B., Siddarth, D. & Dafoe, A. Democratising AI: Multiple Meanings, Goals, and Methods. Preprint at https://doi.org/10.48550/arXiv.2303.12642 (2023).
186. Our Approach to Frontier AI. *Meta* https://about.fb.com/news/2025/02/meta-approach-frontier-ai/ (2025).
187. Executive Order 14028, Improving the Nation's Cybersecurity. *NIST* (2021).
188. Executive Office of the President. Strengthening and Promoting Innovation in the Nation's Cybersecurity. (2025).
189. The White House. National Cybersecurity Strategy. (2023).
190. Jack Cable. With Open Source Artificial Intelligence, Don't Forget the Lessons of Open Source Software. *CISA* https://www.cisa.gov/news-events/news/open-source-artificial-intelligence-dont-forget-lessons-open-source-software (2024).
191. Daniel Huynh & Jade Hardouin. Open Source is Crucial for AI Transparency but Needs More Tooling. *The Linux Foundation Projects* https://lfaidata.foundation/blog/2023/08/01/open-source-is-crucial-for-ai-transparency-but-needs-more-tooling/ (2023).
192. Tooling and Implementation Working Group hosted by CISA. Framing Software Component Transparency: Establishing a Common Software Bill of Materials (SBOM). (2024).
193. Kathleen Hicks. *Implementing Responsible Artificial Intelligence in the Department of Defense*. https://media.defense.gov/2021/May/27/2002730593/-1/-1/0/IMPLEMENTING-RESPONSIBL E-ARTIFICIAL-INTELLIGENCE-IN-THE-DEPARTMENT-O F-DEFENSE.PDF (2021).
194. Herzog, P. *Open-Source Security Testing Methodology Manual* 3. (ISECOM, 2010).
195. Pearce, W. *et al.* AI Security Risk Assessment. Preprint at https://github.com/Azure/AI-Security-Risk-Assessment/blob/main/AI_Risk_Assessment_v4.1.4.pdf (2021).
196. ossf/scorecard. Open Source Security Foundation (OpenSSF) (2025).
197. CISA, FBI, NSF, & US Department of the Treasury. Improving Security of Open Source Software in Operational Technology and Industrial Control Systems. (2023).
198. Chief Information Officer, Department of Defense. *Software Development and Open Source Software*. https://dodcio.defense.gov/portals/0/documents/library/softwaredev-opensource.pdf (2022).
199. Howard, R. Government Contracting Is Hard. *DoD Contract* https://www.dodcontract.com/blog/government-contracting-is-hard (2024).
200. CISA. CISA Open Source Software Security Roadmap. (2023).
201. CVE Project. https://www.cve.org/.
202. Common Vulnerability Scoring System SIG. *FIRST — Forum of Incident Response and Security Teams* https://www.first.org/cvss/.
203. SuperGLUE Benchmark. *SuperGLUE Benchmark* https://super.gluebenchmark.com/.
204. Xu, L. *et al.* SuperCLUE: A Comprehensive Chinese Large Language Model Benchmark. Preprint at https://doi.org/10.48550/arXiv.2307.15020 (2023).
205. Tamirisa, R. *et al.* Tamper-Resistant Safeguards for Open-Weight LLMs. Preprint at https://doi.org/10.48550/arXiv.2408.00761 (2024).