Stanford University
Human-Centered
Artificial Intelligence

DIGICHINA

Issue Brief
HAI Policy & Society
DigiChina
December 2025

Beyond DeepSeek: China's Diverse Open-Weight AI Ecosystem and Its Policy Implications

Caroline Meinhardt, Sabina Nong, Graham Webster, Tatsunori Hashimoto, Christopher D. Manning

Key Takeaways

After years of lagging behind, Chinese AI models — especially open-weight LLMs — seem to have caught up or even pulled ahead of their global counterparts in advanced AI model capabilities and adoption.

We profile and compare the capabilities and distinct features of four notable Chinese open-weight language model families, highlighting that China's ecosystem of open-weight LLMs is driven by a wide range of actors who are prioritizing the development of computationally efficient models optimized for flexible downstream deployment.

Diverse commercial strategies for translating open-weight model adoption into business success are emerging, yet their long-term viability remains uncertain.

The Chinese government's support of open-weight model development — while not the sole determinant of its success — has played a substantial role, though there is no guarantee it will continue.

The widespread global adoption of Chinese open-weight models may reshape global technology access and reliance patterns, and impact AI governance, safety, and competition. Policymakers should ground their policy actions in a granular understanding of real-world deployment.

Introduction

In 2022, China's AI developer community faced dual shocks from the United States. In October, the U.S. government imposed unilateral export controls on semiconductor manufacturing equipment and the most powerful chips for large language model (LLM) training. The following month, OpenAI brought state-of-the-art LLM technology to broad public attention with the launch of ChatGPT. Chinese commentators, noting the government launched a comprehensive plan for AI development five years earlier, asked why breakthroughs were not happening in China, and how Chinese developers could compete with the United States.

As Chinese labs scrambled to develop and release their own models, open-weight LLMs quickly took center stage. Initially, many labs built their models on top of Meta's open-weight Llama model and its architecture. But China's state

of the art would not be reliant on U.S.-trained models for long. A little more than two years after ChatGPT took many in China by surprise, a Hangzhou-based startup called DeepSeek released its R1 reasoning model, demonstrating capability and efficiency that shocked global investors and caused the AI chip leader Nvidia to suffer the biggest single-day loss of any company in U.S. stock market history. While keen analysts had monitored DeepSeek's early development, and some of its methods were known to AI scientists around the world, the broader U.S. discourse did not see the lab's success coming.

Now the strength of Chinese developers in open-weight language models is more widely recognized. DeepSeek has upgraded its models and promises further releases. Alibaba's Qwen models, which were in development long before DeepSeek's breakout moment, are widely used by developers around the world. Another Chinese tech giant, Baidu, which had been pursuing a closed-model strategy, has turned to releasing the weights of some of its flagship models openly. Today, Chinese-made open-weight models are unavoidable in the global competitive AI landscape.

This brief analyzes China's diverse open-model ecosystem, looking beyond DeepSeek. At a moment when Chinese models are increasingly being adopted around the world, including in the United States, and when policymakers are weighing measures to restrict who can build at AI's cutting edge, we pause to give context to the technical and commercial realities of leading Chinese AI labs against a constantly evolving geopolitical backdrop. We then dive deeper into the implications of Chinese open-weight model diffusion for technological development and policy interests in China, the United States, and the rest of the world — implications that raise a series of policy issues that policymakers, scholars, and AI developers might consider going forward. We focus on language models but note the strong performance of other kinds of open-weight AI such as speech, visual, or video models, which some developers integrate with their leading language models.

Today, Chinese-made open-weight models are unavoidable in the global competitive AI landscape.

Free to Deploy and Customize: Chinese Open-Weight Models Are Diffusing Around the World

Developers of products and services with LLMs can choose from dozens of models on which to build. Numerous factors enter the equation, but near the top of the list — alongside performance on relevant tasks — is whether the model is "open" or "closed."

Openness in the context of AI models can mean a lot of different things. Model release can be thought of on

a gradient ranging from models that are fully closed (i.e., not available to anyone outside of the developer organization) to those made publicly available for download under a variety of licensing terms. Many have argued that sharing training data and pre- and post-training code is necessary to achieve true "open-source" AI, but this level of openness is uncommon. The "open" or "open-weight" models we discuss in this brief refer to those whose weights — the parameters within the model that determine its output in response to an input — are free to download, use, and modify. This enables anyone to run the models autonomously outside the bounds of the developer's app or application programming interface (API) and to adapt them for new use cases. Developers from bedroom tinkerers to large global firms are able to modify and deploy models while maintaining ownership of their data environments.

Chinese models now appear to be pulling ahead of their U.S. counterparts when it comes to their downstream reach, though measuring model adoption is inherently hard.

For many applications, openness offers practical advantages (e.g., reproducibility and efficiency in low-resource environments) that can outweigh marginal gains in raw performance. Thus, when DeepSeek R1 and V3 were released, users found them appealing not because they were the best-performing models for every purpose, but because they were free to use and adapt, engineered for relatively low-cost deployment, and good enough for many use cases.

Chinese open models now appear to be pulling ahead of their U.S. counterparts when it comes to their downstream reach, though measuring model adoption is inherently hard. Because users can put open models to work on any sufficiently resourced infrastructure, no central data source has the full picture of their diffusion. Available indicators include model download counts and the number of derivative models published on platforms hosting open models like Hugging Face or GitHub, as well as usage statistics from cloud inference services such as OpenRouter.

For example, in September 2025 Alibaba's Qwen model family surpassed Llama to become the most downloaded LLM family on Hugging Face (*see* Figure 1). Between August 2024 and August 2025, Chinese open-model developers made up 17.1% of all Hugging Face downloads, slightly surpassing U.S. developers, who accounted for 15.8% of downloads. Since January 2025, derivative model uploads based on open models released by Alibaba and DeepSeek have been outpacing derivative models based on major U.S. and European models (*see* Figure 2). And in September 2025, Chinese fine-tuned or derivative models made up 63% of all new fine-tuned or derivative models released on Hugging Face (*see* Figure 3). Combined with anecdotal stories about adoption, these data points suggest a wide variety of contexts and geographies where Chinese models have been adopted.

Figure 1: Cumulative downloads of major open-weight LLMs on Hugging Face (November 2023 – October 2025)
Source: Hugging Face via the ATOM Project

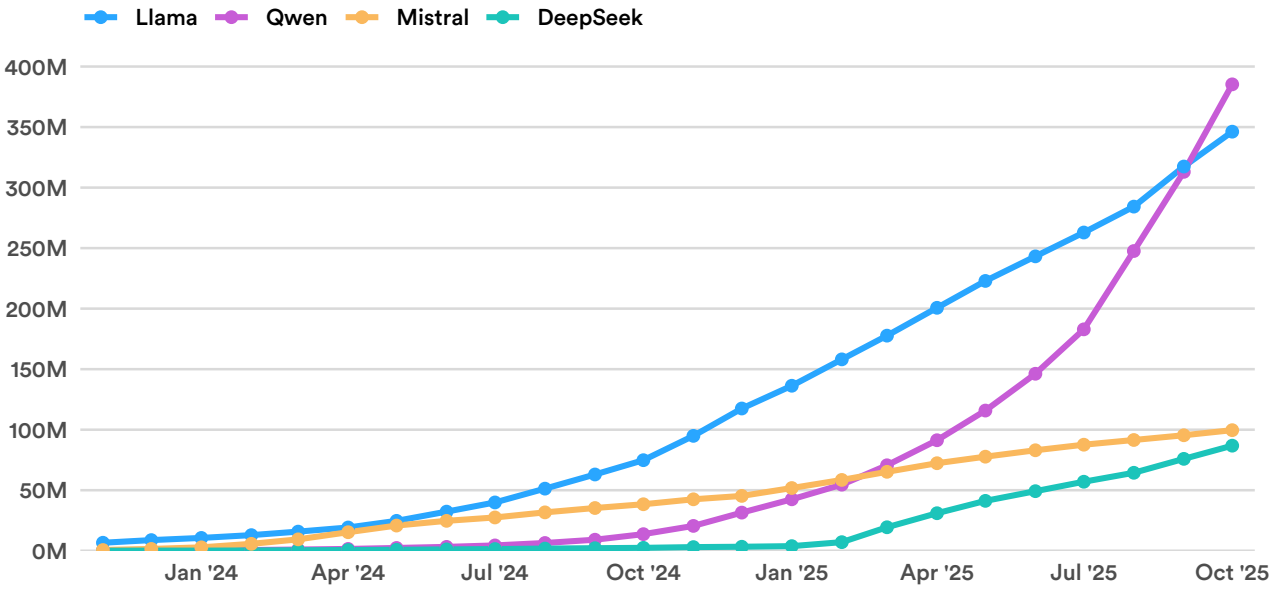


Figure 2: Monthly uploads of new fine-tuned or derivative models to Hugging Face, by base-model developer (September 2024 – September 2025)
Source: HuggingFace via the U.S. Center for AI Standards and Innovation

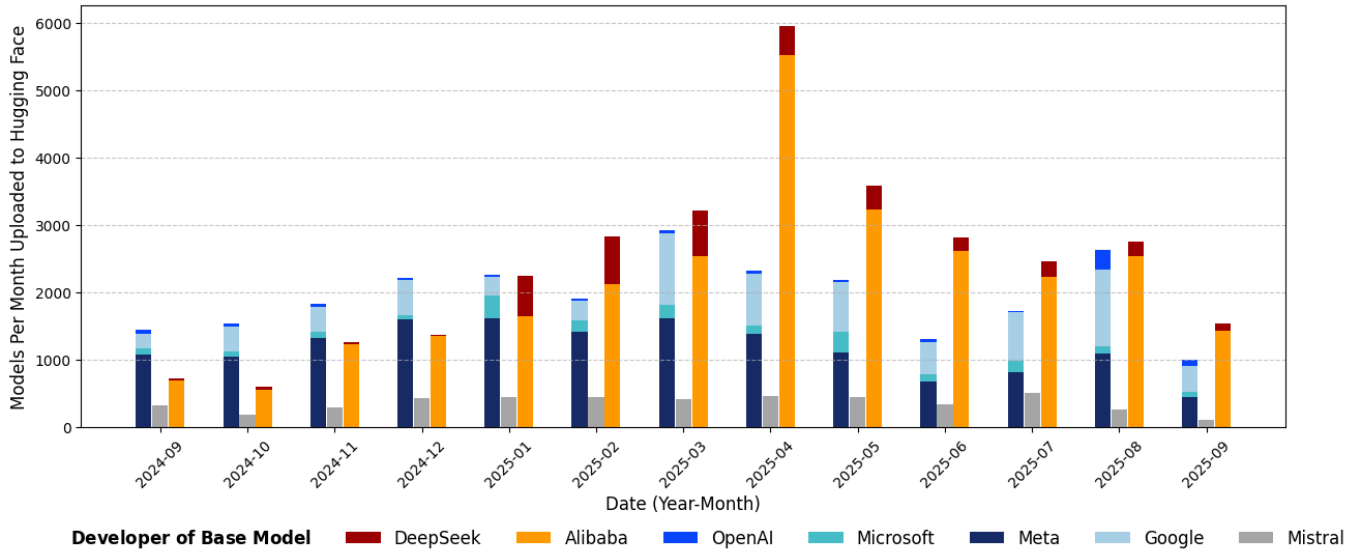
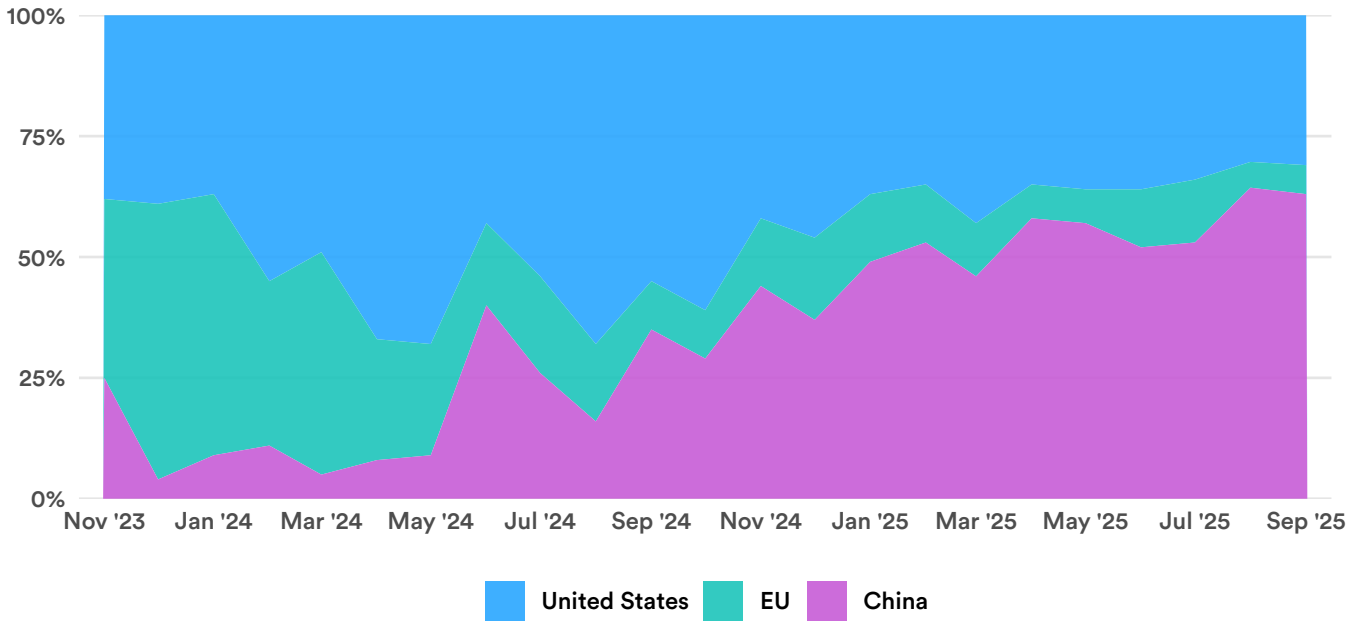


Figure 3: Breakdown of new fine-tuned or derivative models published on Hugging Face, by region of the base model (November 2023 — September 2025)
Source: Hugging Face via the ATOM Project



The Technical Realities of Chinese Open-Weight Models

After years of being seen as lagging behind in AI model development, Chinese open-weight developers are catching up or, in some instances, overtaking their U.S. counterparts in terms of model performance (*see* page 9). While a comprehensive mapping of China's open-model ecosystem is beyond the scope of this brief, we profile and compare the capabilities and distinct features of four handpicked models on page 11. Selected to provide a curated yet insightful snapshot of the ecosystem, these models represent a diverse range of model types, capabilities, and developers.

We draw several conclusions from our review of these four models and the broader open-weight model landscape in China.

First, China's open-model ecosystem is driven by a wide range of actors. While DeepSeek has been attracting outsized attention since 2024 and is often seen as leading Chinese development of open, cutting-edge AI models, in reality the field is much deeper. More than a dozen Chinese organizations have developed and continue to develop and release powerful AI models openly. The four model families we highlight, while among the most advanced and widely adopted in this current moment, are only examples. In a matter of months, the set of most advanced models could easily change.

More than a dozen Chinese organizations have developed and continue to develop and release powerful AI models openly.

In addition to the headline-grabbing startup DeepSeek and the well-known large tech company Alibaba, whose Qwen model family has become a household name in global AI circles, there is a deep bench of other Chinese developers pursuing open models. These include rapidly growing startups, including "tiger" AI unicorns (namely, Z.ai (formerly known as Zhipu AI), Moonshot AI, MiniMax, Baichuan AI, StepFun, and 01.AI); non-profit or university-based research labs (e.g., the Beijing Academy of Artificial Intelligence); and well-established large tech companies (e.g., Tencent, Baidu, Huawei, ByteDance).

Second, Chinese open-weight models now perform at near-state-of-the-art levels across major benchmarks and leaderboards, spanning general reasoning, coding, and tool use. On Chatbot Arena, a prominent crowdsourced leaderboard where users compare and rate outputs, not only do Chinese labs produce the top open-weight models, but the best of them barely trail the leading proprietary systems from U.S. firms. As of December 4, 2025, top closed models from Google DeepMind, xAI, and OpenAI edged out the best from China's Z.ai, Moonshot AI, Alibaba, and Baidu on ChatBot Arena, but their performances were so close that they are all listed as tied for first place, along with Anthropic's top offering. If we look only at open models, 22 releases from five Chinese labs bested the top-ranked open model from a U.S. lab, OpenAI's gpt-oss-120b, with only one other non-Chinese open model developed by French company Mistral featuring among the 25 top-scoring open models (*see* Table 1).

Table 1: Top-scoring open-weight models on ChatBot Arena

Source: OpenLM's ChatBot Arena ranking, as of December 4, 2025

Rank	Model	Lab	Country	Elo score*	Coding score	Vision score	AAII score†	MMLU-Pro score§
1	GLM-4.6	Z.ai	China	**1442**	1458		56	83.5
1	Kimi-K2-Thinking	Moonshot	China	**1438**	1450		66	84.8
2	GLM-4.5	Z.ai	China	**1430**	1448		54	83.5
2	Mistral Large 3	Mistral	France	**1429**	1450		40	81
2	Qwen3-VL-235B-A22B-Instruct	Alibaba	China	**1427**	1457	1246	49	82.8
2	DeepSeek-R1-0528	DeepSeek	China	**1426**	1436		57	84.9
2	DeepSeek-V3.2-Exp-Thinking	DeepSeek	China	**1421**	1438		58	85.1
2	LongCat-Flash-Chat	Meituan	China	**1420**	1460		49	82.7
2	Qwen3-235B-A22B-Instruct-2507	Alibaba	China	**1418**	1457		49	82.8
2	DeepSeek-V3.2-Exp	DeepSeek	China	**1418**	1431		47	83.6
2	DeepSeek-V3.1	DeepSeek	China	**1418**	1430		47	83.3
2	Qwen3-Next-80B-A3B-Instruct	Alibaba	China	**1417**	1456		57	82.4
2	DeepSeek-V3.1-Thinking	DeepSeek	China	**1417**	1437		58	85.1
2	Qwen3-235B-A22B-Think-ing-2507	Alibaba	China	**1416**	1442		62	84.3
3	Qwen3-VL-235B-A22B-Thinking	Alibaba	China	**1411**	1432	1215	62	84.3
3	GLM-4.5-Air	Z.ai	China	**1386**	1410		47	81.5
3	Qwen3-30B-A3B-Instruct-2507	Alibaba	China	**1382**	1425		44	77.7
3	Kimi-K2-0905-Preview	Moonshot	China	**1382**	1403		49	82.4
3	Qwen-VL-Max-2025-08-13	Alibaba	China	**1381**	1440	1213		
3	Kimi-K2-0711-Preview	Moonshot	China	**1380**	1402		47	82.4
3	DeepSeek-V3-0324	DeepSeek	China	**1377**	1391		42	81.9
3	DeepSeek-R1	DeepSeek	China	**1373**	1382		48	84.4
3	Qwen3-235B-A22B	Alibaba	China	**1369**	1394		46	82.8
3	gpt-oss-120b	OpenAI	US	**1368**	1398		59	80.8
3	Qwen3-Coder-480B-A35B-In-struct	Alibaba	China	**1358**	1406		43	78.8

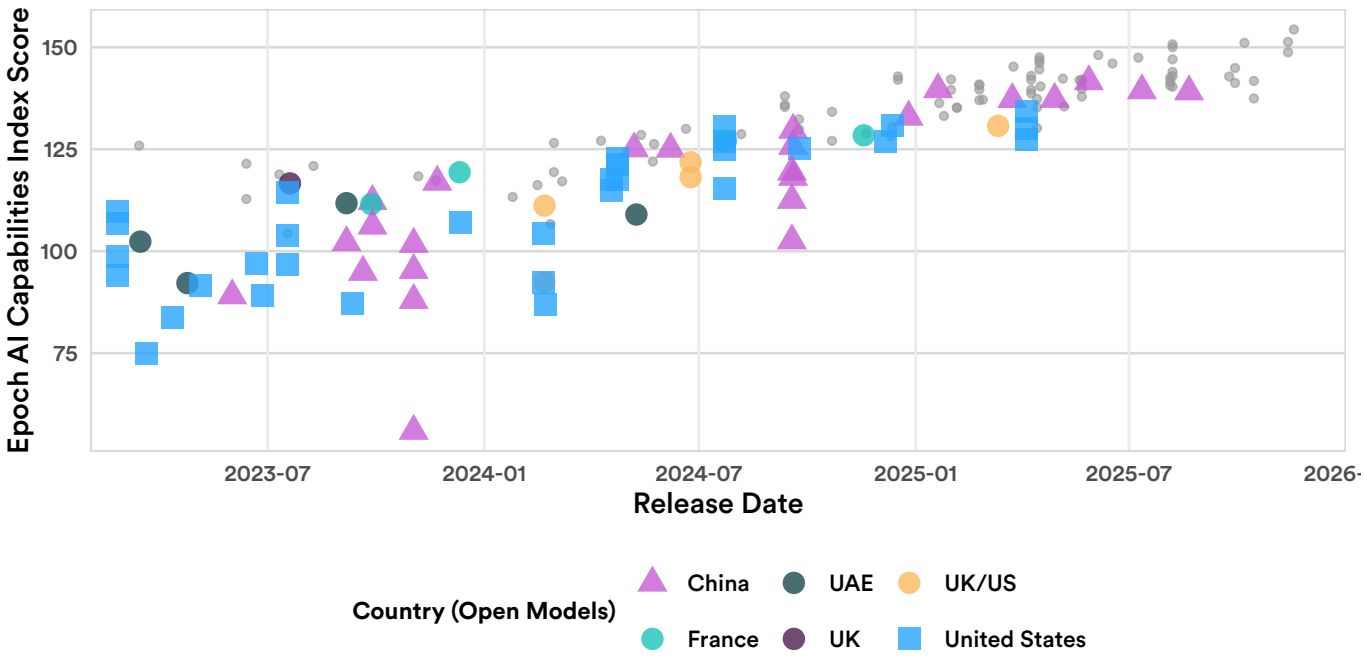
* Overarching score that determines a model's ChatBot Arena ranking. It is based on an Elo rating system that is computed using 5M+ crowdsourced votes.
† Score from the Artificial Analysis Intelligence Index, which incorporates the results of 10 model evaluations.
§ Score from the enhanced version of the MMLU benchmark for language understanding.

The ChatBot Arena scores in Table 1 also demonstrate that each model has distinct strengths: Alibaba's Qwen3 series boasts particularly strong multimodal and multilingual performance; DeepSeek's R1 models excel in step-by-step reasoning; Moonshot AI's Kimi K2 models specialize in coding and tool use; and Z.ai's GLM-4.5 models integrate these capabilities through multi-expert training to achieve balanced and generalist capabilities.

However, it is important to note that leaderboards such as Chatbot Arena are susceptible to leaderboard gaming and other hidden dynamics that can distort rankings. Moreover, the benchmark results used for model comparison often rely on self-reported performance by the developers and should therefore be interpreted with caution. For example, the Center for AI Standards and Innovation (CAISI) — the U.S. government's primary entity for AI testing and evaluation — found that many leading models perform differently under independent verification than in their own reported results, though the degree varies. More fundamentally, benchmarks vary widely in quality, given their lack of standardization. They often fail to measure exactly what they claim to measure, and developers' resulting interpretations of model capabilities can be misleading. Benchmark performance should therefore not be viewed as a definitive measurement of capability, yet it is still a helpful tool when comparing Chinese open-weight models to their global counterparts. In particular, indices that combine the results from many different benchmarks, such as the Epoch Capabilities Index and the Artificial Analysis Intelligence Index, can be helpful for visualizing and comparing model performance (*see* Figure 4 and Figure 5). Both of these show Chinese models catching up with their U.S. and other international counterparts.

Benchmark performance should [...] not be viewed as a definitive measurement of capability, yet it is still a helpful tool when comparing Chinese open-weight models to their global counterparts.

Figure 4: AI model capabilities over time, by country and openness (2023 – 2025)
Source: EpochAI. Open-weight models (colored by country) vs. closed models (gray)



Note: The Epoch Capabilities Index (ECI) combines the scores from 39 distinct AI benchmarks into one "general capability" scale.

Third, model architectures reveal that Chinese AI developers are prioritizing the development of computationally efficient open-weight models. Many Chinese developers are releasing open models that adopt the Mixture of Experts (MoE) architecture. MoE's appeal lies in its efficiency, squeezing better performance out of more limited computing resources (known as "compute") compared with earlier state-of-the-art methods. This greater efficiency makes it faster to both train and use models with less in light of the compute constraints Chinese developers face due to U.S. export controls on advanced AI chips. While MoE had already been employed in models outside of China, such as Mistral's Mixtral, DeepSeek's December 2024 and January 2025 releases introduced a particularly effective instantiation of a fine-grained, shared-experts architecture. Meanwhile, Chinese developers have employed other techniques to optimize for the use of the chips available in China.

Still, variance among Chinese open-weight model release strategies reflects underlying differences in organizational scale and infrastructure capacity. Smaller or more specialized developers such as DeepSeek, Moonshot AI, and Z.ai operate under tighter compute and resource constraints, and choose to focus their efforts on optimization within a single flagship model (such as DeepSeek-R1 or Kimi-K2) that maintains deployment stability at scale. In contrast, well-resourced companies such as

Alibaba, like their Big Tech counterparts in the United States, are supported by extensive in-house cloud ecosystems, and can sustain diversified model families across multiple sizes and modalities to serve varied downstream needs efficiently.

Fourth, Chinese open-weight developers are shifting toward more permissive licensing terms and diverse model variants, increasing flexibility for those who adopt their models. This shift is best demonstrated by comparing the terms of current releases with earlier releases from the same companies. For example, the license Alibaba attached to its smallest 3B and largest 72B variants of 2024's Qwen 2.5 limited its use to research applications, and DeepSeek's V3, released in December 2024, limited redistribution and large-scale commercial use. In contrast, 2025's Qwen3 and DeepSeek R1, are both more capable and were released with more permissive licenses (Apache 2.0 and MIT License), allowing broad use, modification, and redistribution.

From the perspective of the researchers at these labs, releasing models openly can be an important way to gain credibility in the global AI community. Many also argue that enabling wider access to models for fine-tuning and adoption is crucial for maintaining and growing a vibrant open-source community that can foster continued AI innovation and progress. The success of open models with highly permissive licensing has also created competitive pressure for developers to release openly. Even tech giant Baidu, whose CEO had been among the strongest voices in China to laud the advantages of proprietary models, made a U-turn in June 2025, when it released its Ernie 4.5 models openly.

At the same time, developers — especially big tech developers — are introducing diverse model variants optimized for specific application scenarios. One family might offer variants that are lightweight enough to run on a smartphone that does not rely on the cloud (e.g., those that have around 1B parameters) alongside flagship MoE models with more powerful capabilities but that require powerful computers to run (235B or even 1T parameters). In addition, Qwen3, Kimi K2, and GLM-4.5 all introduced dual-reasoning modes, enabling users to toggle between step-by-step reasoning and direct-response modes within or across model variants. This flexibility allows adopters to balance compute efficiency and task complexity according to their operational needs. Moreover, companies including Alibaba and Z.ai have introduced powerful multimodal or vision-reasoning branches, alongside text-based variants specialized for domains such as code and math, thereby broadening their functional reach and allowing for more adaptable downstream integration.

Together, these design and release choices feed an increasingly flexible deployment ecosystem, where model users can optimize for efficiency and functional specialization in addition to or instead of seeking out the best benchmark performance.

Table 2: Summarized overview of four prominent Chinese open-weight model families

Source: Authors' research. Information current as of December 4, 2025

Model Family	Qwen3	DeepSeek-R1	Kimi-K2	GLM-4.5
Overview	Broadest and most versatile Chinese open-weight family, with strong multilingual and multimodal support, and dual-reasoning modes.	Reasoning-first open-weight models with strong math and problem-solving performance and smaller distilled models.	Coding- and agentic-focused open-weight models, notable for efficient inference and fast non-thinking outputs.	Balanced, generalist open-weight models with strong performance across reasoning, coding, agentic, and visual tasks, with dual-reasoning modes.
Developer	Alibaba Cloud (阿里云) *Big Tech* Cloud computing subsidiary of publicly traded Alibaba Group, which has a current market cap of ~$380B.	DeepSeek (深度求索) *Startup* Private startup spun out from and funded by hedge fund High-Flyer. Speculatively valued at ~$15B, though estimates range from ~$1–$150B.	Moonshot AI (月之暗面) *Startup* Private startup funded in part by big tech (e.g., Alibaba, Tencent). Valued at ~$4B.	Z.ai (智谱) *Startup* Private startup with IPO plans spun out from Tsinghua University and funded by big tech (e.g., Alibaba, Tencent) and state entities. Valued at ~$5.6B.
Model Collection Details	Spans eight MoE[1] and dense[2] language models, ranging in size from 0.6–235B parameters. The largest has two efficient MoE variants featuring different active parameter sizes (3B and 22B active), and includes task-specialized versions fine-tuned for math, coding, multimodal reasoning, and ranking, supporting context lengths up to ~128K tokens.	Includes 671B-parameter MoE reasoning language models (37B active) in two variants: R1-Zero for raw, emergent reasoning and R1 for refined step-by-step reasoning. Also includes distilled Qwen, Qwen3, and LLaMA models ranging in size from 1.5–70B parameters, all supporting context lengths up to ~128K tokens.	Includes 1T-parameter MoE language models (32B active) with up to ~256K-token context, offered in Base and Instruct variants.	Includes the Air (106B, 12B active) and Base (355B, 32B active) MoE language model variants with up to ~128K-token context, alongside the multimodal GLM-4.5V for vision reasoning.
Release Timeline	Qwen3 First Release: April 29, 2025 *Notable lab releases:* Tongyi Qianwen (April 11, 2023) Qwen2 (June 7, 2024) Qwen2.5 (September 19, 2024) Qwen3-Omni (September 22, 2025)	DeepSeek-R1 First Release: January 20, 2025 *Notable lab releases:* DeepSeek-V2 (May 6, 2024) DeepSeek-V3 (December 26, 2024) DeepSeek-V3-0324 (March 24, 2025) DeepSeek-V3.2 (December 1, 2025)	Kimi-K2 First Release: July 11, 2025 *Notable lab releases:* Moonlight-A3B / Moonlight-16B-A3B (February 22, 2025) Kimi-Audio-7B (April 25, 2025) Kimi-VL-A3B (March 2025)	GLM-4.5 First Release: July 28, 2025 *Notable lab releases:* GLM-4 (June 5, 2024) CodeGeeX4 (July 2024) GLM-Edge (November 29, 2024) GLM-4.6 (September 30, 2025)
Model Openness	Open weights, architecture, code.	Open weights, architecture, code.	Open weights, architecture, code.	Open weights, architecture, code.
	Technical reports released publicly for Qwen3 (2 weeks after launch) and Qwen3-Omni (day of launch) featuring detailed benchmark performance.	Technical report released publicly (2 days after launch). Additional, peer-reviewed report published later that includes a 12-page safety report.	Technical report released publicly (2.5 weeks after launch), including a dedicated section on safety evaluations.	Technical report released publicly (1.5 weeks after launch) featuring detailed benchmark performance.
	License: Apache 2.0 *A frequently used, highly permissive license that includes patent grants.*[3]	License: MIT License *The most permissive license, frequently used but does not include patent grants.*	License: Modified MIT License *Permissive license, but includes attribution requirements for large-scale commercial deployments*[4] *that some argue counter the spirit of open source.*	License: MIT License *The most permissive license, frequently used but does not include patent grants.*
Notable Features and Capabilities	**Focus on multilingual & multimodal performance:** High self-reported scores, particularly in general knowledge tests, multilingual support (119 languages), and many audiovisual tasks (in its Omni variant). **Cost-efficient architecture:** Its MoE models achieve high performance while emphasizing efficient inference (e.g., the flagship model has a total of 235B parameters but only 22B activated ones per token). **Dual reasoning modes:** Supports both "thinking" (deliberative step-by-step) and "direct" (fast answer) modes in one model.	**Focus on math & reasoning performance:** High self-reported benchmark performance on competition-style math exams and complex reasoning tasks. **Reasoning-first design:** Trained to prioritize problem-solving, with outputs that expose chains-of-thought for transparency and distillation. **Distillation scalability:** R1's reasoning behavior was distilled into smaller, dense models (1.5B–70B parameters) built on Qwen 2.5 and Llama 3 base architectures that imitate R1-level reasoning while being cheaper to run.	**Focus on coding & agentic performance:** High self-reported results in software engineering and tool-use tasks, while also showing robust math reasoning. **Cost-efficient architecture:** Stabilized 1T-parameter training, while activating only 32B parameters per token aiming for fast, low-cost inference. **Agentic capabilities:** Strengthened for planning and tool use, making the model well-suited for software engineering and complex problem-solving. **Dual reasoning modes:** Offers fast, cost-efficient outputs by skipping explicit reasoning traces when they are unnecessary.	**Focus on generalist performance:** High self-reported scores across reasoning, coding, agentic, and visual benchmarks. **Dual-reasoning modes:** Supports both "thinking" (deliberative step-by-step) and "direct" (fast answer) modes in one model. **Reliable agentic capabilities:** Trained with synthetic tool-use and long-context data, aiming for more dependable function-calling and multistep planning. **Multistage expert training:** Trained specialized expert models for reasoning, coding, and agentic tasks, then unified them through self-distillation to create a balanced generalist model.

1 Mixture of Expert (MoE) models are machine learning models that are pretrained much faster with far less computing resources than their dense counterpart models. *See* https://huggingface.co/blog/moe.
2 Dense models are machine learning models that use all their parameters to process every input. They are easier to train and understand but computationally expensive. *See* https://maximilian-schwarzmueller.com/articles/understanding-mixture-of-experts-moe-llms/.
3 In an open-weight AI license, granting patents refers to the permission the licensor gives users to use any patented technology that is included in the model or its components without fear of being sued for patent infringement.
4 The license defines large-scale commercial deployment as having more than 100M monthly active users or generating a revenue of more than $20M per month. *See* https://huggingface.co/moonshotai/Kimi-K2-Instruct/blob/main/LICENSE.

Commercial Ambitions and Open Releases

The rapid adoption of Chinese open-weight models that are still, if only marginally, outperformed by closed models shows that user uptake does not depend solely on benchmarks. Factors including usage costs, the ability to deploy on existing infrastructure, and performance in specific relevant workflows also appear to drive user decisions. How open-weight model adoption may translate to business success for model developers remains uncertain, however, and different firms have different opportunities and strategies. The following examples illustrate this diversity.

As China's leading cloud provider, Alibaba markets the integration of its models into enterprise and government systems. The company promotes Qwen as an "AI operating system," a core intelligence layer it says can underpin a suite of business solutions, from content generation and recommendation to enterprise assistants. According to marketing materials, Alibaba promises to adapt open-weight models into modular components in an enterprise computing environment, citing organizations such as HP and AstraZeneca as their clients. Government adoption of Qwen models, such as the recent decision by Singapore's national AI program to build its flagship LLM on Qwen3, may also drive commercial traffic from across Southeast Asia to Alibaba's cloud platform.

DeepSeek and Z.ai, on the other hand, do not have their own cloud or large-scale compute infrastructure and appear to be pursuing collaborative deployment strategies for clients operating across different cloud and compute providers. For example, DeepSeek has reportedly provided on-premises deployments for Chinese public sector clients, catering to the specific security and compliance needs of these stakeholders. According to one March 2025 report, at least 72 local government agencies across the country had integrated localized versions of DeepSeek models into their governance systems. Chinese state media reported that these systems were deployed and fine-tuned with local data administrations and technical partners, while maintaining a feedback loop with DeepSeek's model team for model iterations.

It still remains to be seen just how lucrative these open-model commercial strategies will be. For now, like their Western counterparts, Chinese open-weight AI developers still rely on the indirect monetization of their models. By openly releasing models that are widely adopted around the world, they hope to create large and diverse user bases that can be directed to other paid products and services built on top of these models. Ultimately, these varying corporate strategies show that AI diffusion both within China and outward from China is taking place in a variety of ways and is not dominated by a single firm or business model.

The Chinese Government's Role in Open-Weight AI Development

Beyond the personal open-source inclinations of some prominent AI developers and commercial interests, there are also political drivers behind Chinese developers' prioritization of open-weight model releases. The Chinese government has shown longstanding top-down support for open technology development, framing "open source" as crucial to both national innovation strategy and geopolitical

goals. Support for open-source AI can be traced at least back to the government's 2017 New Generation Artificial Intelligence Development Plan, which championed "open source" and "open" collaboration as a way to aggregate domestic strengths for national advancement. The more specific goals are to unlock the collective power of academia, industry, and government and combine public data, infrastructure, and talent to support national development goals. The language used to frame these goals is diverse and often vague, but it signals support for open source software and, now, models. For example, a recent policy document that encourages AI adoption across industries encouraged academic organizations — which operate under state direction — to consider open source contributions alongside other scholarly performance metrics such as publications.

The Chinese government frames its model of open collaboration in implicit but clear contrast to U.S. export controls, supply chain restrictions, and closed models.

Internationally, the Chinese government's support of open-weight AI development is now part of its self-portrayal as a global advocate for equitable and inclusive technological development. The Ministry of Foreign Affairs' Global AI Governance Initiative, announced in October 2023 and the Global AI Governance Action Plan, released in July 2025, promote open-source AI and shared development across borders, emphasizing global "equal rights to develop and use AI." The Chinese government frames its model of open collaboration in implicit but clear contrast to U.S. export controls, supply chain restrictions, and closed models, which it portrays as barriers to the equitable distribution of the benefits of AI. This is part of China's broad diplomatic approach. For example, Beijing has framed open-source model sharing and AI infrastructure support as tools for joint development and local AI development in the context of the Forum on China–Africa Cooperation, a centerpiece of China's South-South diplomacy. Of course, for all the talk of openness, widespread adoption of Chinese open-weight models using Chinese-hosted infrastructure would also bring business opportunities and perhaps political influence.

Although open source is a watchword of the Chinese government's AI development initiatives, it is important to note that DeepSeek, once a low-profile AI lab that spun out of a private hedge fund, appears to have achieved its success with limited — if any — direct state support. It was only after the company had attracted widespread attention in the AI community for its V3 release that it received major political acknowledgment, with China's no. 2 leader, Premier Li Qiang, inviting DeepSeek founder Liang Wenfeng to join a high-level government symposium. While AI companies have received various forms of government support at the national and local level, researchers have noted that support for open-weight model development has primarily come in the form of incentivizing an enabling environment rather than directly subsidizing open-weight AI developers. However, some local governments are increasingly

providing tailored financial and policy support to AI organizations and projects that are specifically engaging with open-source communities. More broadly, it is also worth noting that the Chinese government's long-standing and targeted cultivation of top-notch engineering talent through education policies and public AI research funding created a deep, increasingly domestically educated AI talent pool. This cohort has become the bedrock of Chinese open-weight AI development while also being among the leading forces in pushing global AI research. The state has also supported the build-out of computing infrastructure that could be used for AI development.

It is by no means guaranteed that the Chinese government will continue to support open-weight AI model development. Government attention can be a double-edged sword, as evidenced by reports that DeepSeek has had some of its executives placed under travel restrictions. The "crackdown" on Chinese tech platforms that started in late 2020 shows that the Chinese government is constantly trying to balance its desire to nurture an environment that enables rapid innovation with its need to tightly control the private tech sector and ensure ideological alignment. While there are no imminent signs of a policy shift, it is possible that a major AI-related adverse event or national security threat could cause the government to quickly constrain open-weight AI development, especially amid policymakers' growing concerns over the misuse risks of open models.

Global Policy Implications

China's vibrant open-weight AI model ecosystem raises a variety of important policy questions. We highlight four areas that deserve particular attention below.

Global access and dependencies: The wide availability of high-performing Chinese AI models opens new pathways for organizations and individuals in less computationally resourced parts of the world to access advanced AI, thereby shaping global AI diffusion and cross-border technological reliance patterns. With model performance converging at the frontier, AI adopters with limited resources to build advanced models themselves, especially in low- and middle-income countries, may prioritize affordable and dependable access to enable industrial upgrading and other productivity gains. Chinese models that, at the very least, are "good enough," openly available under highly permissive licenses, and competitive or superior in terms of usage costs have certain advantages in meeting this demand. The logic also applies in well-resourced countries: U.S. companies, ranging from established large tech companies to some of the most hyped AI startups, are widely adopting Chinese open-weight models. The existence of open-weight Chinese models at the good-enough level may thus decrease global actors' reliance on U.S. companies providing models through APIs.

Open models do not necessarily mean users are free of dependency on the model maker. The history of Chinese IT infrastructure vendors such as Huawei and ZTE supplying connectivity and data-center projects in Global South markets gives them experience offering technology solutions around the world. Huawei, accordingly, has marketed DeepSeek integration as part of its cloud offerings in countries like South Africa and other parts of sub-Saharan Africa. For those working toward "sovereign AI," an open model provided through cloud services operated by a foreign firm still poses concerns about long-term dependencies. The extent of these dependencies is highly contingent on how many aspects of the tech stack (from the underlying hardware to the training

data and model architecture) are bundled together in AI offerings.

AI governance: The rapid diffusion of Chinese open-weight models around the world also raises concerns that adopting Chinese AI models could effectively spread Chinese political censorship or propaganda, threaten data security, or open users to cybersecurity risks. When Chinese developers produce models that can legally be used in China, content restrictions and domestic standards and regulations (e.g., for model registration) will shape their work — from collecting and cleaning training data through model pre-training and fine-tuning. Though various investigations have found that censorship guardrails can be removed fairly easily, and many adopters of open models, especially enterprise users, are not deterred by potential model censorship, the potential for Chinese state preferences to influence user experience around the world is real.

Although open-weight models by definition can be run outside of the creators' infrastructure, in reality, many will use the apps, APIs, and integrated solutions offered by DeepSeek, Alibaba, and others. This typically means user data is under the control of these companies and may physically travel to China, potentially exposing information to legal or extralegal access by the Chinese government or corporate competitors. This risk is significantly mitigated by running the model locally, on a trusted cloud provider, or via a trusted inference provider like Hugging Face, but for end users relying on Chinese cloud or infrastructure providers, similar concerns remain. Finally, while there is no verified evidence of deliberate backdoors for advanced AI systems, past cases involving Chinese technology, such as Hikvision equipment, will continue to fuel speculation that Chinese developers could be brought into the service of the Chinese government in novel ways.

AI safety: There is some evidence that Chinese developers may be less focused than their U.S. counterparts on the various risks under the rubric of "AI safety." An evaluation by CAISI, the U.S. government's AI testing center, found that DeepSeek models on average were 12 times more susceptible to jailbreaking attacks than comparable U.S. models. Other independent evaluations conducted by safety researchers also demonstrate that DeepSeek's guardrails can easily be bypassed.

An in-depth evaluation of the security risks of open models is beyond the scope of this paper, as is a comprehensive review of the wide range of other risks caused by the dual-use nature of open-weight (and closed-weight) models. However, in the context of far-reaching global diffusion, it is important to note that what happens — or does not happen — in China regarding AI safety will not necessarily stay in China. Policymakers and others who wish to understand risks or implement risk reduction measures cannot ignore and likely need to engage with Chinese labs and regulators.

Geopolitical competition and open-closed debates: Fears of Chinese AI "leadership" in AI remain front and center in geopolitical competition discussions. China's advances in open-weight AI model development are shaping how the United States approaches open-weight AI development and AI competition more broadly. Until relatively recently, major U.S. AI developers, from OpenAI to Google DeepMind, had prioritized releasing their most advanced language models as closed, proprietary models. Meanwhile, the U.S. policy community had been focused on debating the potential risks of openly released AI models. Yet following the global shock waves caused by the release of DeepSeek's R1 model, U.S. policy thinking changed. Newly inaugurated President

Donald Trump referred to R1 as "a wake-up call" to focus on AI competition, and his AI czar, David Sacks, cited it as justification for pursuing a deregulatory federal approach to AI development. Accordingly, in July 2025, the White House released America's AI Action Plan, which elevates open-weight models as a strategic asset for U.S. innovation and security, while also emphasizing the need to strengthen export controls on foreign adversaries including China to ensure the United States can "win" the "AI race." The following month, OpenAI — for the first time since its open release of GPT-2 nearly six years prior — released two open-weight models under the Apache 2.0 license. Several newer U.S. startups and academic labs have since fully committed to open-weight AI development.

It should be noted that some of these policy shifts were already underway. Yet the DeepSeek moment — and subsequent open-weight models released by other Chinese AI labs — likely played a pivotal role in expanding the willingness to accept the risks related to open-weight model development to compete more directly with China. It also helped foster an acute recognition that leadership in AI now depends not only on proprietary systems but on the reach, adoption, and normative influence of open-weight models worldwide.

Conclusion

Each of the policy implications discussed above shares the mark of uncertainty. But questions about who is using which technologies, in what ways, with what architectures of independent control or reliance are both answerable and urgent if policymakers, advocates, and technologists are to make better choices.

This calls for a continued effort to collect and share both quantitative and qualitative data about real-world deployments of open-weight AI.

To understand policy implications in areas ranging from data governance to economic competition to national security, therefore, policymakers and researchers will need to move beyond the relatively basic data currently available from open-source platforms such as Hugging Face and GitHub. Outcomes do not depend on downloads or derivative model development alone; they rely on architecture and control of the hardware and software systems that turn models into useful services. Moreover, China's current edge in open-weight AI adoption is only a year old and may not last forever. This calls for a continued effort to collect and share both quantitative and qualitative data about real-world deployments of open-weight AI. Such an interdisciplinary challenge will at times require significant attention to technical detail in various fields, as well as substantial industry and regional domain knowledge. As a first step, researchers and citizens alike can make geopolitical or industry narratives concrete and subject them to empirical scrutiny.

Inevitably, some policy actions will be taken (or foregone) in response to insufficient data and uncertainty about future techno-social trends. The evolving understanding of AI deployments and effects calls for similarly evolving policy approaches. Any policies designed to compete with China or proactively spread U.S. alternatives should be grounded in a maximally granular understanding of how LLMs are actually being deployed and what local needs and goals are driving AI adoption, and they should maintain flexibility as the data and technology evolve. Researchers will also need to look beyond language models, as the implications will vary, for example, in vision or code-assist models.

Finally, selective engagement with Chinese labs, academics, and policymakers should not be avoided or unnecessarily constrained, given the wide range of AI governance and safety challenges Chinese actors face, just as their U.S. and other global counterparts do. As some have argued, focusing engagement with China on nontechnical topics such as incident reporting and risk management frameworks without leaking sensitive technical insights is not only possible but urgent. It is also worth considering that the very nature of open-model releases enables better scrutiny of models. There is space for academic collaboration with Chinese counterparts to increase our understanding of the risks of open-weight AI models and the efficacy of various guardrails.

Stanford University's Institute for Human-Centered Artificial Intelligence (HAI) applies rigorous analysis and research to pressing policy questions on artificial intelligence. A pillar of HAI is to inform policymakers, industry leaders, and civil society by disseminating scholarship to a wide audience. HAI is a nonpartisan research institute, representing a range of voices.

The DigiChina Project, part of the Program on Geopolitics, Technology, and Governance at Stanford University's Center for International Security and Cooperation, is a collaborative effort to analyze Chinese technology policy through direct engagement with primary sources, providing analysis, context, translation, and expert opinion.

The views expressed in this policy brief reflect the views of the authors. For further information, please contact **HAI-Policy@stanford.edu.**

Caroline Meinhardt is a policy research manager at the Stanford Institute for Human-Centered Artificial Intelligence (HAI).

Sabina Nong is an AI governance researcher and a research assistant at Stanford HAI.

Graham Webster is a lecturer and research scholar in the Program on Geopolitics, Technology, and Governance at Stanford University and editor-in-chief of the DigiChina Project.

Tatsunori Hashimoto is an assistant professor of computer science at Stanford University and a faculty affiliate at Stanford HAI.

Christopher D. Manning is a professor of linguistics and of computer science at Stanford University and an associate director at Stanford HAI.

Acknowledgments

The authors thank Kendra Schaefer, Daniel Zhang, Elena Cryst, and participants of a closed-door roundtable on open-weight AI models and regulation convened at Stanford University for their valuable insights; and Irene Solaiman, Nathan Lambert, Florian Brand, and the ATOM Project for sharing data and analysis.

Stanford University
Human-Centered
Artificial Intelligence

Stanford HAI: 353 Jane Stanford Way, Stanford CA 94305-5008

T 650.725.4537 **F** 650.123.4567 **E** HAI-Policy@stanford.edu **hai.stanford.edu**