

Testimony of

Kyle I. Chan, Ph.D.

Fellow in the John L. Thornton China Center  
at the Brookings Institution

Before the

U.S. House Select Committee on Strategic Competition Between the  
United States and the Chinese Communist Party  
hearing titled  
“China’s Campaign to Steal America’s AI Edge”

April 16, 2026

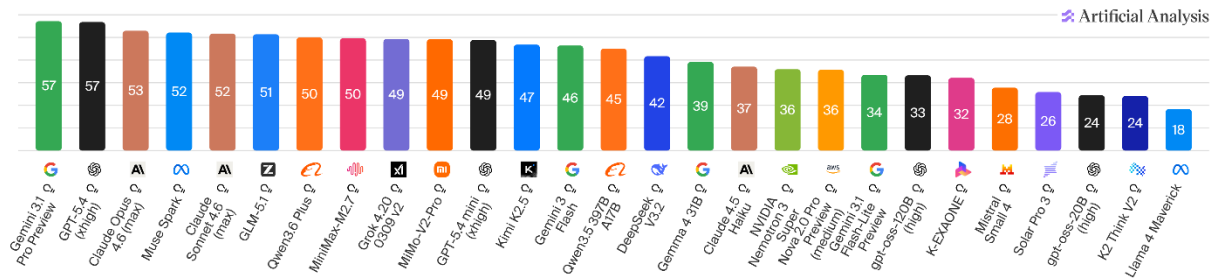
Chairman Moolenaar, Ranking Member Khanna, and distinguished Members of the Committee, thank you for the opportunity to testify on the U.S.-China AI competition.

China is pursuing a full-stack approach to AI development, from chips and compute infrastructure to foundation models and applications. The goal of Chinese policymakers is not to achieve AGI, or “artificial general intelligence,” but to leverage AI as a powerful, general-purpose technology that will turbocharge a wide range of sectors and services. National programs, such as China’s 2017 New Generation AI Development Plan and its more recent “AI Plus” Initiative, are geared towards integrating AI into manufacturing, healthcare, drug discovery, scientific research, education, and government services. China’s military and security agencies have long sought to use AI to improve both defensive and offensive capabilities.

China’s top AI models continue to lag behind American frontier models by several months or more. American AI models maintain a clear lead in overall performance across a wide range of industry benchmarks, from math and reasoning to code generation and long-horizon agentic tasks. Chinese AI labs, particularly startups, are constrained by access to compute, due to a combination of U.S. export controls on advanced AI chips and limited capital resources. Alibaba, one of China’s largest AI players, plans to invest over \$53 billion in AI over three years. In contrast, Microsoft spent approximately \$80 billion on AI capex in 2025 alone. America’s main hyperscalers—Alphabet, Amazon, Meta, and Microsoft—have plans to spend a total of \$650 billion just this year. American data centers are reaching gigawatt scales and deploying hundreds of thousands of AI accelerators. If everything boiled down to compute and the race to AGI, the United States would hold the decisive advantage.

## Artificial Analysis Intelligence Index

Artificial Analysis Intelligence Index v4.0 incorporates 10 evaluations: GDPval-AA, r<sup>2</sup>-Bench Telecom, Terminal-Bench Hard, SciCode, AA-LCR, AA-Omniscience, IFBench, Humanity's Last Exam, GPQA Diamond, CritPT



Source: [Artificial Analysis](#)

But China is pursuing a different approach to AI. While some Chinese AI companies, such as DeepSeek and Alibaba, also talk about trying to achieve AGI, Chinese policymakers and China's AI industry as a whole are more focused on running several different AI races. They are more focused on making progress in model efficiency, AI adoption, and the integration of AI into the physical world. China's focus on these dimensions of AI development is the result of several factors, including industry constraints—particularly access to large-scale compute and capital—as well as Beijing's policy priorities. In addition, China is aggressively pursuing semiconductor self-sufficiency, seeking to localize nearly every major segment of the semiconductor supply chain in the face of U.S.-led export controls.

## Efficiency

Chinese AI labs are focused on improving model efficiency, driving down the costs of deployment and trying to wring out greater performance from limited compute resources. Chinese AI companies, particularly startups, do not have access to the compute scale of their American competitors due to U.S. export controls on the cutting-edge AI chips, the lower performance and availability of Chinese domestic chip alternatives, and far less access to capital compared with the trillion-dollar valuations of their American peers. As a result, in an effort to keep pace with American AI labs, Chinese AI companies have had to resort to algorithmic and engineering solutions to compensate for their lower compute resources.

Chinese AI models have leaned heavily on mixture-of-experts (MoE) architectures, which activate only a subset of parameters for each token, reducing compute at inference while maintaining the capacity of much larger models. DeepSeek has developed a novel architecture called DeepSeek Sparse Attention to reduce the computational and memory costs of the original transformer attention mechanism, an approach that other Chinese AI labs such as Z.ai have adopted. Moonshot AI, the maker of the Kimi foundation models, has developed a hybrid linear attention architecture that can support context lengths up to one million tokens while dramatically reducing compute and memory costs.

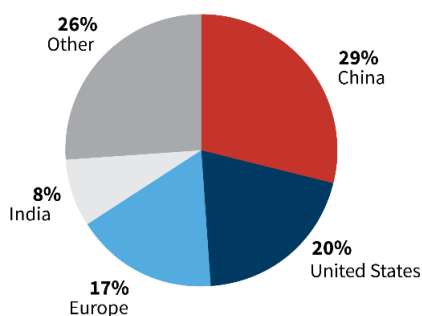
Chinese AI labs have also pushed the boundaries on more efficient engineering techniques, particularly in an area called quantization, which uses less precise data formats to compress models while minimizing performance loss. Alibaba has pushed aggressively in 4-bit quantization for its Qwen model series, although some research has found that extremely low-bit quantization does degrade model performance. Moonshot AI’s Kimi-K2-Thinking model is a natively INT4-quantized model, which greatly improves its deployment efficiency. American AI labs and chipmakers are also working in tandem to push the frontier of quantization, moving toward ultra-low-precision formats like FP4 to dramatically improve efficiency. But for Chinese AI labs, this push for greater efficiency is a much more central part of their research and development efforts, largely out of necessity.

Distillation is another way for Chinese AI labs to improve model performance in the face of compute constraints. Distillation involves the systematic use of outputs from more advanced models to train and improve less capable models. Anthropic recently released a report detailing large-scale distillation campaigns that they attribute to several Chinese AI labs, including DeepSeek. OpenAI and Google DeepMind have also documented similar distillation campaigns that operate through distributed networks of accounts to obscure their true origins. American AI leaders are now collaborating to prevent future distillation campaigns, particularly through information sharing.

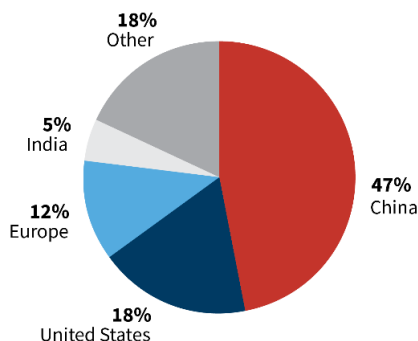
While these distillation campaigns are serious and should be prevented, they do not fully explain the progress made by Chinese AI labs in developing world-class foundation models. Distillation can only improve models to a limited extent, and Chinese AI labs have shown that they are making genuine innovations through published research papers and technical reports that are recognized by American AI researchers. Chinese AI researchers are represented in high numbers at top-tier AI research conferences, such as NeurIPS, reaching as much as half in 2022 (see chart below).

### Countries of Origin of Top-Tier AI Researchers, Based on Undergraduate Degree

2019



2022



Source: The Paulson Institute

Source: Matt Sheehan and Sophie Zhuang. 2025. "Have Top Chinese AI Researchers Stayed in the United States?" Carnegie Endowment for International Peace.

## Adoption

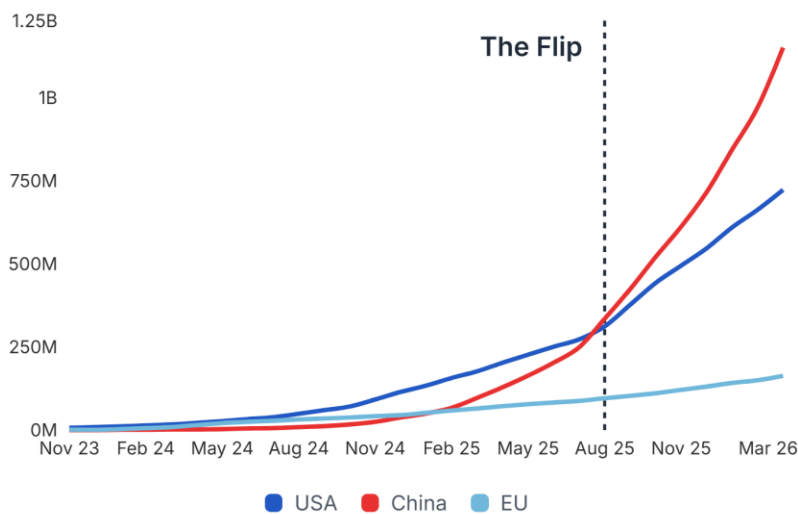
Chinese AI firms are prioritizing adoption, both domestically within China and around the world. In particular, many Chinese AI firms have pursued an open-source strategy, releasing models with open weights alongside detailed technical reports. This approach makes many top Chinese AI models not only free to use but also easy for developers around the world to download, adapt to their specific needs, and deploy across a range of platforms, including their own compute infrastructure. By contrast, most top American AI models remain closed and accessible only via paid subscriptions or API services. This has made American AI companies far more commercially successful in terms of direct revenues from model use. But China's open-source approach has made their AI models popular among developers who still seek strong model performance but are more cost-conscious.

This adoption-focused strategy is paying off. On platforms like Hugging Face, Chinese models have surpassed U.S. counterparts in total downloads, and derivative models built on Chinese foundations have outpaced those built on American ones. Meta's Llama models—once the industry standard for open models—has been overtaken in popularity by Alibaba's Qwen. At the same time, Chinese cloud providers such as Huawei, Alibaba, and Tencent are expanding their AI offerings abroad, particularly in emerging markets.

Developers globally are drawn to these models because they offer strong performance at little to no cost. Teams from Japan to Africa are building on models from Alibaba and DeepSeek, and adoption is growing even in Silicon Valley. Brian Chesky recently noted that Airbnb's customer service agent relies heavily on Qwen, describing it as fast, capable, and inexpensive. A growing number of Silicon Valley startups are similarly choosing to build on Chinese open models.

### Models Worldwide

Cumulative Downloads, 2023- present



The ATOM Project (updated 03/2026)

source: huggingface

Source: The ATOM Project

## Physical Integration

China is particularly focused on the integration of AI into the physical world. Chinese policymakers see physical applications of AI as an area with large-scale potential that plays to China's strengths in manufacturing and electronics supply chains. From autonomous vehicles and drone delivery systems to AI-powered wearable devices and agentic AI smartphones, Chinese firms and Chinese policymakers are aligned in striving to integrate AI into real-world applications.

Robotics and so-called "embodied AI" are a key priority for Chinese policymakers. Chinese robot maker Unitree announced that it had manufactured over 5,000 humanoid robots last year and is preparing for a major public listing in China this year. Chinese electric vehicle makers such as Xiaomi and NIO are experimenting with the deployment of humanoid robots on auto assembly lines. Local governments in China have helped to set up specialized labs for collecting robotics training data using human operators. China is hoping to leverage its advantages across a range of adjacent industries such as electric vehicles, smartphones, batteries, and sensors, which share an overlap with robotics supply chains. Embodied AI is a major priority in China's latest Five-Year Plan and a key target for its new \$138 billion national venture capital guidance fund.

To be sure, American tech firms and startups are also moving fast on robotics and physical AI. Google's Waymo and Tesla's FSD are widely regarded as the global industry leaders in autonomous driving technology. Top-tier American robotics firms such as Physical Intelligence and Google DeepMind are pioneering best-in-class robotics foundation models. But they may struggle to scale physical production like their Chinese counterparts given China's deep existing robotics supply chain.

## Semiconductor Self-Sufficiency

China is pursuing a Manhattan Project-like program to build a resilient, largely self-sufficient semiconductor supply chain. China's efforts to develop its domestic semiconductor industry stretch back decades, including the founding of China's two largest foundries, Hua Hong in the 1990s and SMIC in 2000. In 2014, China set out ambitious plans to become a global leader in all major segments of the semiconductor industry by 2030.

But it was two sets of U.S. policy actions that kicked China's semiconductor efforts into high gear. The first was a wave of tech controls under the first Trump administration that blocked Huawei from using TSMC for chip fabrication and prevented ASML's EUV lithography machines from being sold to China, among other actions. The second was the October 2022 and 2023 export controls under the Biden administration, which restricted the export of advanced AI chips and advanced chipmaking tools to China. Together, these U.S. actions sent shockwaves through China's Party-state system and caused Chinese policymakers to prioritize semiconductor self-sufficiency, not only in chip production but across the entire semiconductor supply chain.

China has made substantial progress overall in its semiconductor indigenization efforts but with varying outcomes across segments. On AI chips, Chinese domestic chips made up 41 percent of China's market in 2025 with approximately half of those sales coming from Huawei, according

Huawei's alternative platforms

Huawei	US / Global
Kirin 9000S	Qualcomm Snapdragon
Kirin X90	Intel Core processor
Kunpeng 920	x86-based server CPU
Ascend 910C	Nvidia H100 & H20
CloudMatrix 384	Nvidia GB200 NVL72
CANN	Nvidia CUDA
EulerOS	Other Linux-based server OS
MindSpore	PyTorch, TensorFlow
HarmonyOS Next	Android mobile OS
HarmonyOS PC	Windows OS
MetaERP	Oracle, SAP
GaussDB	Oracle SQL Developer

Kyle Chan. Source: media reports and company websites

to IDC. This contrasts with a 90 percent or more market share for Nvidia in the Chinese AI chip market before 2023. Huawei's latest Ascend 950PR chips are expected to scale production to 750,000 units this year and mark a notable shift toward more CUDA-compatible architecture, making it easier for Chinese AI developers to switch. Cambricon, another Chinese chipmaker, is planning to deliver 500,000 units of its AI accelerators in 2026, largely manufactured domestically according to reporting from Bloomberg.

At the single-chip level, Chinese AI chips are likely to remain significantly behind Nvidia's Blackwell and Rubin GPUs on key performance metrics, such as processing performance, memory capacity, and memory bandwidth. However, Chinese

AI chipmakers are seeking to make up for their lower single-chip performance by connecting clusters of chips together into more powerful hardware systems. For example, Huawei has developed AI hardware systems, such as the CloudMatrix 384 and the Atlas 950 SuperPoD, that offer strong compute performance across sets of chips, albeit with lower power efficiency than comparable U.S. offerings.

China's primary bottleneck in the near term is advanced-node fabrication capacity. Chinese foundries, particularly SMIC and Hua Hong, continue to face challenges in ramping up large volumes of leading-edge chips at sufficiently high yields due to U.S.-led restrictions on semiconductor manufacturing equipment. Because Chinese foundries are restricted from purchasing ASML's EUV lithography machines, they are forced to squeeze leading-edge fabrication out of older DUV lithography equipment through multi-patterning techniques, which severely reduce yield and drive up per-chip costs. Despite these constraints, SMIC, Hua Hong, and a network of foundries linked to Huawei are aiming to increase advanced chip production from 20,000 wafers per year to 100,000 wafers in one to two years, according to Nikkei.

To address these manufacturing bottlenecks, Chinese equipment makers such as SiCarrier, AMEC, NAURA, and SMEE are receiving government support to develop domestic chipmaking equipment, from etching and deposition to lithography. SMIC has begun testing China's first immersion DUV lithography machine from SiCarrier- and Huawei-linked Shanghai Yuliangsheng Technology. A lab in Shenzhen has built a prototype EUV lithography machine, according to Reuters, although operationalization and commercialization likely remain years away. More generally, Huawei has taken the lead role in coordinating an ecosystem of Chinese foundries and equipment makers. China's latest Five-Year Plan calls for "extraordinary measures" to break through U.S.-led export controls with a particular focus on semiconductors.

**Figure 1. China's AI Tech Stack and Industrial Policy**

Layer	Segment	Example Organizations			
Physical applications	Robotics	Estun, Inovance, Unitree	National and local robotics programs	National VC Guidance Fund "Embodied Intelligence"	National AI Industry Investment Fund
	Sensors and vision	Robosense, Hesai, SenseTime			
	Autonomous driving	Pony.AI, WeRide	Autonomous vehicle pilot zones		
	Drones	DJI, Autel, Ehang, AVIC	Local government procurement		
	Smart cities	Alibaba CityBrain			
	Biotechnology	WuXi AppTec, BGI Genomics, XtalPi			
Software applications	Agents	Manus, Ant Zhixiaobao	Local government open data platforms	China Internet Investment Fund	
	Healthcare	Baichuan, Tencent AIMIS			
	Finance	Ant Maxiaocai, Qifu			
	Education	Zuoyebang, Yuanfudao			
	Public services	iShenzhen, Hangzhou			
	Enterprise	Alibaba, ByteDance, Kingdee Cloud, Tencent			AI Pilot Zones
Data	Open data platforms	OpenI, NLJIC Shanghai			Local-government investment funds, industrial parks, talent incentives, startup programs
	Data exchanges	National Public Data Resource Platform			
Foundation models	Start-ups	DeepSeek, Minimax, Moonshot, Zhipu			Subsidized compute vouchers
	Big Tech	Qwen, ERNIE, Hunyuan			
	Public	Beijing Academy of AI WuDao			
Research and development	Commercial	Alibaba DAMO Academy	NSFC and provincial research funding	Promotion of open-source models and frameworks	State-backed AI labs
	Public	Shanghai AI Lab, Zhejiang Lab			
	Academic	Tsinghua, Zhejiang University			
Data centers	Commercial	Alibaba Cloud, Huawei Cloud	National Integrated Computing Power Network		
	Public	National Integrated Compute Network			
Servers	Operating systems	Huawei OpenEuler, Kylin OS, UOS			
	Hardware	Inspur, H3C, Huawei, Lenovo			
Software platforms	AI frameworks	Huawei MindSpore, Baidu PaddlePaddle	National Integrated Circuit Investment Fund		
	Parallel computing	Huawei CANN			
Semiconductors	Custom ASICs	Baidu Kunlun, Alibaba T-head			
	General-purpose AI	Huawei Ascend, Biren BR100, Cambricon			

NOTE: ASIC = application-specific integrated circuit; NSFC = National Natural Science Foundation of China; VC = venture capital.

Source: Kyle Chan, Gregory Smith, Jimmy Goodrich, Gerard DiPippo, Konstantin F. Pilz. 2025. Full Stack: China's Evolving Industrial Policy for AI. RAND Corporation.

# How America Can Win in AI

## Export Controls

Export controls are a critical policy tool for U.S. national security. The U.S. must not allow American technology to support China's military capabilities. And the U.S. must take action to maintain an edge in performance for our frontier AI models. We are already pursuing these two important policy goals today with strong restrictions on U.S. technology flows to Chinese entities linked to the People's Liberation Army, export controls on cutting-edge American semiconductors to China, and broader efforts to restrict China's access to advanced semiconductor manufacturing equipment, such as EUV lithography machines.

At the same time, export controls are only partial solutions with both benefits and costs. Export controls work most effectively when used in a selective and strategic manner, and they should not be treated as substitutes for a more comprehensive set of American AI policies. Current export controls on American AI chips to China have helped to slow China's AI development in the near term by making access to large-scale compute more difficult for Chinese AI labs. However, despite these restrictions, Chinese AI labs have managed to trail closely behind top American labs, likely through a combination of model efficiency improvements, chip smuggling, access to overseas compute resources, model distillation, and other factors.

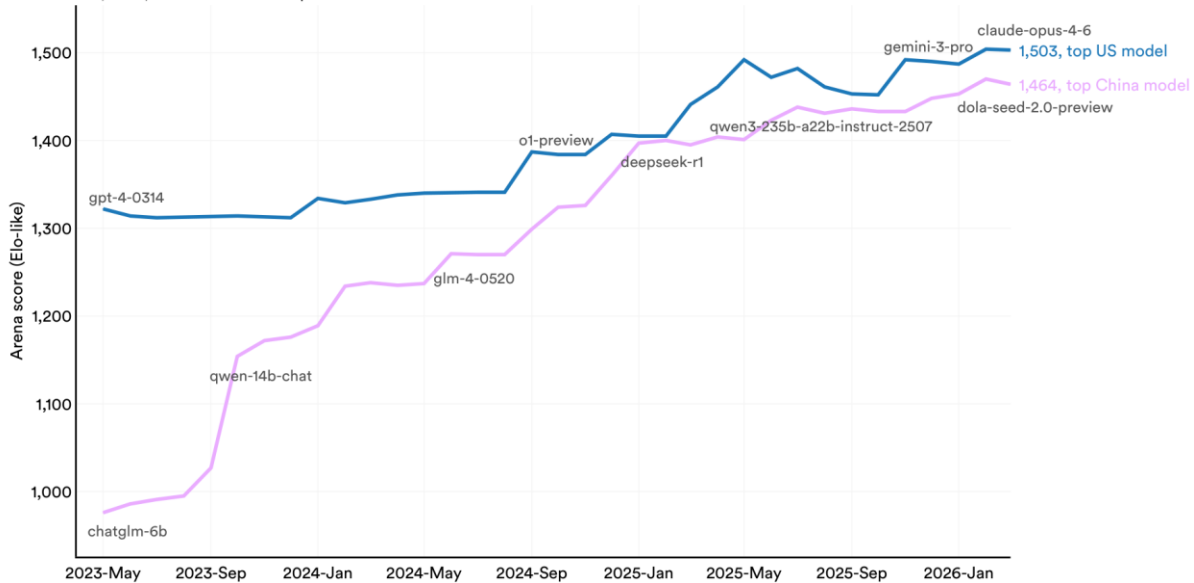
In addition, U.S.-led export controls on AI chips and chipmaking equipment have stimulated China's own semiconductor development efforts, prompting Chinese policymakers and industry participants to pivot to an accelerated, whole-of-nation effort to build a nearly complete domestic semiconductor supply chain aimed at being resistant to U.S.-led export controls. While export controls have slowed China's AI development in the near term, they may ultimately accelerate China's chip development efforts over the medium and long term. Chinese AI labs have been pushed to work with Chinese AI chipmakers, such as Huawei and Cambricon, to improve hardware-software integration and create a closed AI development loop that excludes U.S. technology.

While Chinese AI chips are not expected to close the performance gap with cutting-edge American AI chips for the foreseeable future on a single-chip basis, China's scale-up of AI chip production along with new chip designs that are increasingly performant and usable, such as the Huawei Ascend 950 series, will likely enable China's AI industry to continue making progress, even in the face of U.S.-led semiconductor export controls. In other words, export controls can slow China down in the near term but are unlikely to halt China's AI progress in the long run. Moreover, these measures have likely spurred China's efforts to build a resilient AI supply chain.

Given these trade-offs and limitations, the U.S. should continue to use export controls, sanctions, and other policy tools for the goals described earlier but shift policy focus toward other pressing aspects of AI development that demand greater attention. Export controls are important, but the bulk of our efforts should go toward ensuring American AI progress is world-leading, sustainable, and broadly beneficial to Americans.

### Performance of top United States vs. Chinese models on the Arena

Source: Arena, 2026 | Chart: 2026 AI Index report



Source: [Stanford HAI 2026 AI Index Report](#)

## Strengthening the American AI Stack

American private firms and startups have been leading the charge in many critical parts of the American AI stack. American AI companies create the world’s leading foundation models; American chipmakers produce the most powerful AI hardware; and American startups and enterprises are driving diffusion and adoption of AI across the economy. Yet, for all the successes of our private sector in driving American AI forward, there are glaring gaps in the American AI ecosystem that need support from U.S. policymakers and other actors.

The first gap is energy. American AI leaders have increasingly warned of a growing energy bottleneck for U.S. data center buildout. The [International Energy Agency](#) estimates that U.S. power demand for data centers will more than double from 2024 to 2030, reaching 426 terawatt-hours (TWh) or roughly 9 percent of total electricity demand. China has an advantage in this domain given its ability to quickly build and connect new power generation capacity. In 2025, China added over 540 gigawatts of new power capacity, about 80 percent of which was solar and wind. Over the past four years, China has built the equivalent of the entire U.S. power grid in terms of new power capacity. To support America’s data center buildout, U.S. policymakers at the federal and local levels must develop more streamlined permitting and licensing procedures, provide more resources for approving interconnection queue requests, and address supply chain bottlenecks in key electrical components, such as transformers and high-voltage transmission equipment.

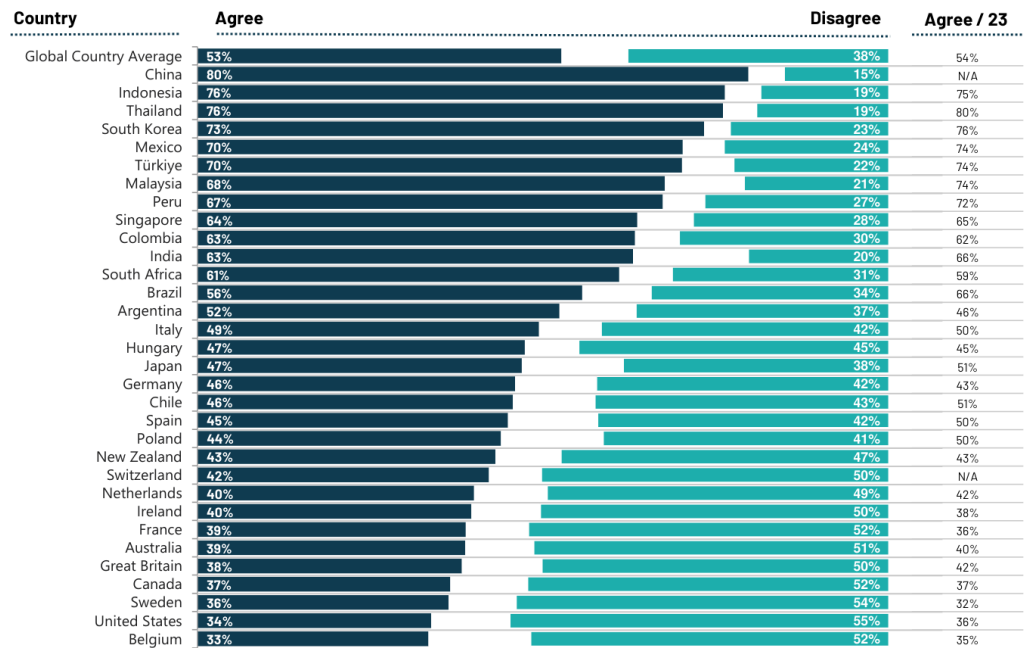
The second gap is open-source models. Many Chinese AI labs pursue an open-source strategy, releasing open model weights that users can freely download, customize, and deploy on their own preferred infrastructure. This open-source strategy has contributed significantly to the rapid

adoption of Chinese AI models around the world, including in the United States. Due to differing structural and commercial incentives, American AI labs are less motivated to release strong open-weight models with some exceptions, such as Google DeepMind’s Gemma series, OpenAI’s OSS models, and Nvidia’s Nemotron family. As the White House’s [AI Action Plan](#) rightly recognizes, the lack of robust open-source AI model offerings from the United States cedes a critical channel for global diffusion to Chinese models. U.S. policymakers might consider offering financial or other incentives to American AI labs to develop and support open-source foundation models alongside proprietary ones.

The third gap is compute for basic research. As the scale and costs of compute have grown, American private AI firms have pulled back from sharing cutting-edge research findings through academic exchanges. At the same time, non-commercial AI researchers at American universities do not have access to the same level of compute resources as their corporate counterparts. While American private firms may be more focused on scaling up technical improvements with greater near-term success, non-commercial AI researchers need greater access to compute resources for conducting longer-term AI research that may not have immediate commercial payoffs. While the federal government provides some fragmented support for compute access today, such as the National Science Foundation’s [ACCESS program](#), a more comprehensive program to equip academic and other non-commercial AI researchers with large-scale compute would strengthen America’s technological edge for the next generation of AI innovation.

How much do you agree or disagree with the following?  
**Products and services using artificial intelligence make me excited**

Base: 23,685 online adults under age 75 across 32 countries, interviewed April 19 – May 3, 2024



## AI Safety

As AI systems grow more powerful, so does their potential for misuse. From AI-enhanced cyberattacks to deepfakes to AI-enabled bioweapons, both the scope and scale of potential malicious uses of AI are expanding. So far, despite some efforts at the national and international level such as the United States' Center for AI Standards and Innovation (CAISI), regulation and policy around AI risks remains largely underdeveloped and left in the hands of the private sector. Yet, individual AI labs alone lack the capabilities and resources to comprehensively monitor—much less enforce—AI safety policies.

While significant attention has rightly focused on the security risks associated with China's growing AI capabilities, we should not overlook AI threats from non-state actors and third countries. There is a real possibility that some of these misuses of AI will draw on the combined deployment of American and Chinese AI models, just as many legitimate users already rely on both given the frontier performance of top American models and the cost-effectiveness of many Chinese ones. Moreover, a sophisticated actor may deliberately route API calls across multiple AI model providers—both American and Chinese—to exploit fragmented AI safety policies across companies and countries to evade detection and circumvent company-level safety mechanisms.

Even as the United States competes fiercely with China in AI and takes measures to protect Americans and American firms from risks related to China's AI development, both countries should consider developing common AI safety protocols and even communication channels where national interests overlap. Investigations by [CAISI](#) and others have found Chinese AI models lag behind American models in basic safety features, such as resistance to jailbreaking techniques. Meanwhile, Chinese AI labs are pursuing advancements in coding and long-horizon agentic tasks. Given the potential spillovers to the United States and the rest of the world, the U.S. has an interest in urging Chinese AI labs to strengthen their safety practices and in potentially forming channels for information-sharing, particularly for suspicious usage patterns. Given the current low levels of trust between the two countries, this will likely be a difficult uphill battle, but one that will be increasingly necessary as AI risks continue to escalate.

## Conclusion

Ultimately, the U.S.-China AI race is not a single contest but a competition across multiple dimensions: compute, models, adoption, integration, and deployment. The United States retains a clear lead at the technological frontier, particularly in compute scale and model performance, but China is advancing rapidly through efficiency gains, open-source diffusion, and deep integration of AI into the real economy. In the long run, the winner of the AI race will be determined not simply by who builds the most powerful models but by who can most effectively translate AI into broad-based economic and societal gains. For the United States, this means complementing export controls with sustained investment in infrastructure, research, and open ecosystems to ensure that American AI remains not only world-leading but also beneficial to Americans from all parts of society.