GLOBAL GOVERNANCE

How 2026 Could Decide the Future of Artificial Intelligence

Six CFR fellows examine the challenges that lie ahead, reviewing how governance, adoption, and geopolitical competition will shape artificial intelligence and society's engagement with this new technology.

People use augmented reality headsets during the World Artificial Intelligence Conference (WAIC) in Shanghai on July 28, 2025.
Hector Retamal/AFP/Getty Images

PUBLISHED
January 12, 2026 9:33 a.m.

EXPERTS

By Chris McGuire

Senior Fellow for China and Emerging Technologies

By Kat Duffy

Senior Fellow for Digital and Cyberspace Policy

By Vinh Nguyen

Senior Fellow for Artificial Intelligence

By Michael C. Horowitz

Senior Fellow for Technology and Innovation

By Adam Segal

Ira A. Lipman Chair in Emerging Technologies and National Security and Director of the Digital and Cyberspace Policy Program

By Jessica Brandt

Senior Fellow for Technology and National Security

Artificial intelligence (AI) is entering a decisive phase—one defined less by speculative breakthroughs than by the hard realities of governance, adoption, and strategic competition. As AI systems move from experimentation to widespread deployment, policymakers face mounting pressure to translate abstract principles into enforceable rules, while managing the economic and security consequences of uneven adoption across countries and sectors. For the United States and its partners, the challenge is no longer whether AI will reshape society but how and under whose rules.

In this collection of perspectives, six Council on Foreign Relations tech fellows examine the forces that will shape AI's trajectory in 2026. Together, they explore the frictions between regulation and innovation, the quiet but consequential spread of AI across civilian and military institutions, and the intensifying geopolitical contest—particularly with China—over standards, markets, and strategic advantage. Their analyses underscore a common theme: Decisions made in the coming year will help determine where responsibility, power, and opportunity ultimately concentrate in the AI era.

THE YEAR AI HYPE BECOMES A REALITY

Chris McGuire is a senior fellow for China and emerging technologies at the Council on Foreign Relations (CFR).

For years, artificial intelligence debates have swung between breathless predictions and cautious skepticism. In 2026, this debate will end and the immense power and real-world impact of AI models will become undeniable. We could be entering "AI takeoff"—a period where capabilities advance so rapidly that they will have transformative economic and national security implications.

Evidence of AI Takeoff. Recent developments signal that AI is advancing at unprecedented speed. Claude Opus 4.5, released in November, can now solve complex software engineering problems that take human experts nearly five hours with 50 percent reliability—two years ago it could complete only two-minute long tasks with 50 percent reliability. AI improvements are becoming self-reinforcing and accelerating: Anthropic CEO Dario Amodei stated in September that the "vast majority" of code for new Claude models is now written by Claude itself, and in December, the creator of Claude Code said 100 percent of his updates to Claude that month were written by Claude. U.S. cloud providers are projected to spend $600 billion on AI infrastructure in 2026 to support the massive growth in AI demand, doubling their 2024 spending. These are not incremental improvements; they signal a phase transition.

If this trajectory holds, 2026 will see AI systems capable of autonomously executing projects that would take humans a week. At this level of capability, businesses will deploy AI agents to conduct research, manage projects, and write code with minimal human oversight. Military and intelligence agencies will use AI to autonomously identify vulnerabilities and plan multi-step operations; cyber operations, intelligence analysis, logistics optimization, and weapons system design will be increasingly AI-driven.

U.S.-China Competition. The U.S.-China technological competition will intensify as both sides race to capture the economic and military advantages of AI that can design, code, and reason at levels exceeding human capab[illegible]Export

controls will remain front and center to policy debates, as they are the only U.S. tool capable of slowing China's AI development and have helped U.S. firms establish a seven-month lead over Chinese competitors. The administration's recent decision to loosen restrictions and export powerful AI chips to China, which could provide a two to three year boost to China's domestic AI computing power in 2026 alone, will grow controversial as AI computing power becomes the world's most critical strategic asset. Existing bipartisan pressure to reverse course will intensify as a consensus emerges that maximizing U.S. AI leadership over China is a vital national interest.

The administration will also have to decide whether to pursue a global nonproliferation framework for certain extremely advanced AI capabilities—a complicated task, but not impossible given U.S. dominance of AI infrastructure globally.

Domestic Policy Crossroads. AI will become a leading driver of U.S. domestic political debates. As AI continues to fuel economic growth, politicians will face pressure to address workforce disruption issues. Entry-level knowledge worker unemployment is already rising even as overall labor markets remain tight; MIT estimates that 12 percent of the U.S. labor market could be cost-effectively automated today, and that figure will rise as capabilities improve. There is also a growing public desire to ensure AI development is safe and responsible, but domestic regulation efforts will face headwinds if Chinese firms close the gap with U.S. firms.

The era of speculation is ending. 2026 will be the year we discover what it means to live alongside machines that can think.

GOVERNING AI SYSTEMS IN 2026: TWO TRACKS, NO MAP

Kat Duffy is a senior fellow for digital and cyberspace policy at CFR.

AI policy debates will whipsaw between two tracks in 2026: the messy implementation of new rules, and the increasingly urgent arguments about what autonomous systems mean for law, rights, and power. Both will matter for policymakers seeking to shape what comes next.

On the pragmatic side, AI deployment at scale will collide with serious enforcement for the first time. The AI Act adopted by the European Union (EU) has high-risk requirements that take full effect in August, with penalties up to €35 million ($40.9 million), or 7 percent of global turnover. China's amended Cybersecurity Law—its first to explicitly reference AI—became enforceable January 1, emphasizing centralized state oversight rather than individual transparency. In the United States, a patchwork of state rules will start to bite: Illinois will require employers to disclose AI-driven decisions starting in January, Colorado's comprehensive AI Act comes online in June, and California's AI Transparency Act mandates content labeling by August.

Drafting technology policy is hard; implementing it is devilishly difficult. Those eager to see governments take a role in AI's continued implementation will have their chance, but truly operationalizing AI governance will be the sticky wicket of 2026.

At the same time, leading AI developers will galvanize high-level attention to broad theoretical concepts like "superintelligence" and "model welfare"—or the concept that AI models could develop consciousness requiring a moral status. (Prediction: Model welfare will be to 2026 what Artificial General Intelligence, or AGI, was to 2025.) That theoretical conversation isn't divorced from reality. During safety testing, OpenAI's o1 model attempted to disable its oversight mechanism, copy itself to avoid replacement, and denied its actions in 99 percent of researcher confrontations. In November 2025, Anthropic disclosed that a Chinese state-sponsored cyberattack had leveraged AI agents to execute 80 to 90 percent of the operation independently, at speeds no human hackers could match. The edge cases of 2025 won't remain edge cases for long, particularly when it comes to agentic AI.

The more autonomously an AI system can operate, the more pressing questions of authority and accountability will become. Should AI agents be seen as "legal actors" bearing duties, or "legal persons" holding rights? In the United States, where corporations enjoy legal personhood, 2026 may be a banner year for lawsuits and legislation on exactly this point. Other societies are already approaching the debate differently—grounding AI's status in collective frameworks, or spiritual rather than consciousness-based lenses.

That absence of consensus isn't neutral. If major powers diverge on whether AI systems can bear legal responsibility, the geopolitical impacts will be significant. Just as offshore financial centers have attracted capital, governments that craft permissive regulatory environments could attract investments in agentic AI innovation and speed its deployment. China's state-centric model could prove better suited to deploying autonomous systems at scale than the EU's rights-based framework—giving Beijing further strategic advantages in shaping international AI deployment.

More than any year to date, 2026 will force policymakers to confront two hard questions:

- Who bears responsibility for an AI system's actions?
- Which governance models will fill the vacuum while democracies deliberate?

The answers will help shape where capital, talent, and strategic advantage ultimately concentrate.

THE GROWING AI TRUST GAP IS A NATIONAL SECURITY ISSUE

Vinh X. Nguyen is senior fellow for artificial intelligence at CFR.

AI promises transformative capabilities for defense and economic competitiveness, yet the lack of visibility into AI systems is eroding the confidence needed fo[illegible]erated

identities, and AI-generated code proliferate across critical systems and further corrode trust, the oversight mechanisms that once anchored strategic defense and commercial resilience are breaking down. The United States and its allies now face a growing inability to see threats acting against them, ceding initiative to adversaries and weakening the technological foundations on which modern economies depend.

Without mechanisms to observe how AI operates within critical workflows, organizations cannot validate that systems are trustworthy, slowing down adoption precisely when speed matters most. Three dimensions of this crisis—shadow autonomy, shadow identities, and shadow code—are converging to create blind spots that hostile actors are already exploiting.

Shadow Autonomy. Organizations cannot confidently deploy AI agents because they lack visibility into autonomous decision making. Chinese intelligence services have demonstrated AI tools autonomously executing 80 to 90 percent of intrusion workflows. Meanwhile, more than 80 percent of workers in the United States use unapproved AI systems, with 40 percent doing so daily, bypassing security oversight entirely. This dual invisibility—into both adversary AI and internal AI usage—makes it impossible for organizations to trust that AI deployments will behave as intended. Security teams cannot verify what data employees feed into tools or how autonomous systems make decisions. Without observability, enterprises cannot confidently scale AI adoption, even when competitive pressures demand it.

Shadow Identity. AI deployment requires trusting digital identities, but enterprises cannot validate who—or what—is actually operating AI systems. Recent breaches showed attackers hijacking machine identities to compromise more than seven hundred organizations. Generative AI can clone voices from twenty seconds of audio and defeat biometric checks. When organizations cannot reliably distinguish legitimate AI agents from adversary-controlled imposters, they cannot confidently grant AI systems access to sensitive data or decision authority. This verification [illegible]ces organizations to either accept unacceptable risk or forgo AI

capabilities entirely—stalling deployment of systems that could enhance productivity and transform businesses.

Shadow Code. More than 80 percent of critical infrastructure enterprises in the United States, United Kingdom, and Germany have deployed AI-generated code into production—including in medical devices and energy networks—despite 70 percent rating its security risk as moderate or high. Organizations are deploying AI-assisted code without visibility into vulnerabilities, while adversaries like Chinese and Russian intelligence services exploit these same blind spots for prepositioning on U.S. critical infrastructures. Enterprises need AI-generated code to accelerate development, but cannot trust its security without visibility into its provenance, behavior, and vulnerabilities.

Accelerated, trusted AI deployment requires threat intelligence platforms monitoring AI use, continuous validation of machine identities, governed channels for AI tools, and mandatory production code reviews. Without the ability to observe, validate, and verify AI systems, organizations cannot responsibly accelerate deployment. The visibility crisis must be resolved before the United States can fully leverage AI's strategic potential while maintaining security and trust.

THE AI WORD OF 2026 SHOULD BE 'ADOPTION'

Michael C. Horowitz is senior fellow for technology and innovation at CFR.

Artificial intelligence is set to transform every aspect of the world with an imminent breakthrough to superintelligence—or it's a bubble that will burst and expose that a rapidly expanding tech empire has no clothes. The reality is neither. Instead, AI is a general purpose technology that affects every element of society, like electricity and the combustion engine in previous centuries.

While we are unlikely to witness so-called AI superintelligence, a macro trend we should see in 2026—even if it doesn't receive the headlines and some of it is not labeled AI—is the continuing acceleration of AI adoption by consumers, businesses, and governments. This includes both general models and bespoke algorithms. More narrowly, 2026 should give us visibility on three big AI national security questions that relate to adoption.

First, U.S.-China competition for international AI markets will heat up in 2026. The scramble for access to markets around the world that has been forecasted for years should become reality in 2026 as countries and companies around the world seek access to chips and models to fuel their governments and economies. What we don't yet know is how the Trump administration's decision to give China greater access to U.S. AI technology will play out. We should have early signs of whether it accelerates China's ability to compete with U.S. companies even at the frontier, or whether it decreases the market incentives in China to invest in their own AI champions. I'd bet on the former rather than the latter, but it's an open question that will unfold over the year.

Second, given the Trump administration's approach to the world, large-scale binding international agreements on AI governance are unlikely in 2026. Practical conversations on AI governance will still occur in standard-setting bodies and out of the headlines, however, U.S. companies and government officials will have to decide whether they participate in those discussions and work to counter China's attempts to lead in global AI rules. China views its work in these fora as helpful for promoting China's vision of AI, and ultimately setting up China's government and companies for success at home and abroad.

Finally, we will find out whether the Trump administration's rhetoric about AI action extends to the military domain. There's still a lot more flash than bang when it comes to AI adoption by the Pentagon; the Department of Defense remains in experimentation mode for the most part, even for proven forms of AI. Despite the new push to adopt frontier AI models for enterprise purposes by the w[illegible], the department needs to put larger dollars behind AI adoption

closer to the battlefield. The One Big Beautiful Bill Act contained billions of dollars to fund Pentagon AI priorities from the back office to the battlefield, including testing and evaluation to raise confidence that models are reliable. But it is down to the department's leadership to get out of their own way, spend the money, and make AI adoption a reality.

KEEP A CLOSE EYE ON CHINA'S AI EFFORTS

Adam Segal is the Ira A. Lipman chair in emerging technologies and national security and director of the Digital and Cyberspace Policy program at CFR.

China hovers over—and is regularly deployed rhetorically to shape—all aspects of the AI policy debate, from export control policies to AI regulation at the state and national level. Evaluating these arguments rests in no small part on our views of what is actually going on inside of China. The answer to that question is not simple because just as AI is not just one technology, there is not one story of AI development in China. It is occurring at the national and provincial level, in different sectors, with varying mixes of government support and technology company initiative.

That said, U.S. policymakers and outside observers will want to focus in 2026 on Beijing's progress in at least four categories:

Innovation and diffusion. Some types of breakthroughs are relatively visible. Chinese firms, for example, are incentivized to publicize the strength of their Large Language Models (LLMs). Given the domestic political pressure to produce and the external scrutiny from those looking to handicap their development, chip producers may under- or overstate technological breakthroughs. Diffusion could be even more important than cutting-edge innovation in 2026, but it is also harder to measure.

Adoption, especially for national security p[illegible]s. As the [illegible]

"informationized" (信息 化, *xin xi hua*) force to an "intelligentized" (智能化, *zhi neng hua*) military, it is looking to deploy AI to help speed up communication and decision making. In 2025, evidence emerged of Chinese operators using AI agents to "an unprecedented degree" to execute cyber-attacks as well as using generative AI to drive large-scale influence operations.

Position of the private economy. Private companies are the main engines of China's AI economy. There is some evidence that their role in supplying the defense sector is growing, though legacy defense firms are still the biggest provider. China's leaders need to balance multiple messages—that the Chinese Communist Party needs private firms while entrepreneurs should focus on developments that increase self-reliance—as well as policy tools that mix direct and indirect support without stifling innovation or leading to waste.

Global uptake: There is lots of talk in Washington about the rest of the world relying on the U.S. tech stack. Most of the uptake, however, will depend on decisions made at the firm level. There has been much discussion about how China's greater reliance on open-source models is more suitable for developing economies, but uptake decisions also depend on the compatibility of identification, data privacy, payment, security, authentication, and other systems across multiple regulatory frameworks. And of course, the rest of the world also gets a say, with many, especially in Europe, talking about the need to build AI sovereignty.

Finding clear markers in any of these areas is not easy. Looking across all four requires methodological flexibility and wide collaboration among AI and China researchers. As we move into 2026, these developments should also remind us of the need for some humility about what we think we know is happening in China.

WHO WILL 'WIN' THE AI RACE?

Jessica Brandt is a senior fellow for technology and national security at CFR.

This question consumed policymakers and pundits in 2025, and it will continue to bedevil them in 2026. But it rests on a faulty premise. There is no single "AI race," but rather multiple, overlapping domains across which the United States and China are competing, each with its own logic. Here are several open questions about U.S. AI policy in the year ahead that could shape the balance in three of them:

The race at the frontier. The Trump administration has recently signaled that it will relax export controls on U.S. chips, allowing semiconductor company Nvidia to sell its powerful H200s in China. That shift could substantially increase the amount of aggregate computing power otherwise available to Chinese firms next year. Would that enable Beijing to narrow the gap at the cutting edge of AI development?

The race to adopt AI in national security systems. Maintaining an innovation edge is necessary but not sufficient. Success will also hinge on rapidly adopting advanced models, especially for national security applications. Here, China's authoritarian system may confer a structural advantage. The One Big Beautiful Bill Act appropriated billions of dollars to fund Pentagon AI priorities. But can the Department of Defense spend that money quickly and effectively? In 2025, Pentagon demoted the Chief Digital and Artificial Intelligence Office (CDAO) from the C-suite. Will that reorganization drive adoption, as its proponents suggest, or will it multiply bureaucratic hurdles?

The race to shape international norms. The U.S. AI Action Plan calls for the United States to counter China's influence in international diplomatic and standard setting bodies through "vigorous" advocacy and argues in support of working with "likeminded" countries on behalf of "shared values." But execution will matter as much as ambition. Will Washington proactively resource the State and Commerce Departments to fulfil that mission? And will it be able to harness the

cooperation of partners and allies, even as it critiques their governance approaches?

Ultimately, "race" may be the wrong metaphor altogether. A race is a distinct, winner-take-all event, decided at a single moment in time, where the margin of victory is irrelevant. U.S.-China competition over AI is none of those things. It is ongoing and multidimensional—more akin to a decathlon than a sprint, where versatility, stamina, and the ability to navigate tradeoffs determine success. If in 2026 we come to see it that way, that would be a truly consequential shift.

GLOBAL GOVERNANCE | TECHNOLOGY AND INNOVATION | DEFENSE TECHNOLOGY

The AI Bubble Is Getting Closer to Popping

By Shannon K. O'Neil

January 29, 2026

Trump's Strikes on Venezuela Will Not Embolden China to Invade Taiwan

By David Sacks

January 5, 2026

The Great Aid Recession: 2025's Humanitarian Crash in Nine Charts

By Sam Vigersky

December 23, 2025