Written Testimony of Tim Fist

Director of Emerging Technology Policy, Institute for Progress (IFP)

Before the US House Committee on Science, Space, and Technology, Research & Technology Subcommittee, with topic: "DeepSeek: A Deep Dive"

Tuesday, April 8th, 2025

Subcommittee Chairman Obernolte, Committee Chairman Babin, Subcommittee Ranking Member Stevens, and other members of the subcommittee: Good morning, and thank you for the opportunity to testify today. I'm the Director of Emerging Technology Policy at the Institute for Progress (IFP), a nonprofit and nonpartisan research organization focused on innovation policy. Our organization is dedicated to accelerating the pace of scientific progress and steering the advancement of emerging technologies in a direction that is compatible with American values. As one example, we engage in a research partnership with the National Science Foundation to support its work on improving scientific grantmaking and ensuring the American R&D enterprise is operating effectively.[1]

I'll focus my comments today on two areas:

1. What can we learn from events earlier this year surrounding DeepSeek?
2. What actions should the federal government take to ensure the United States remains the global leader in AI? I focus on R&D, energy, and talent.

Many of my policy recommendations are drawn from IFP's response to the Office of Science and Technology Policy's request for information for an action plan for AI.[2]

Was DeepSeek a "Sputnik moment"?

In December and January, the Chinese AI company DeepSeek open-sourced two new models — V3 and R1 — and launched a free chatbot app that rapidly topped mobile app store rankings. In March alone, it was used over 800 million times.[3] For context, OpenAI's ChatGPT saw 5 billion uses that month — but the directional shift is clear: China is now distributing models close to the frontier globally, fast, and for free.

That prompted some to label DeepSeek a "Sputnik moment" for American AI.[4] The analogy is rhetorically powerful, but technically misleading. Where Sputnik represented a dramatic leap in capability, DeepSeek's most impressive advances are comparable to US models from mid-2024 — about eight months behind the state of the art.[5] Unlike Sputnik's closed technologies, DeepSeek's models are freely available. And while the Sputnik project was

designed and manufactured using Soviet technologies, DeepSeek heavily relies on American technologies.

From this, we should draw three lessons.

Chinese open-source AI is creating new security risks

The wide availability of cheap and capable Chinese models presents a security concern. China's civilian and military sectors are famously "fused," where the CCP can exert top-down control of companies and their products.[6] This risk is made more severe by the fact that AI models and applications can be designed to have backdoors that we have no good way of detecting, or designed with modifications to spread specific ideologies.[7]

Insofar as Chinese models are only used in chatbots, these risks are likely not an acute threat to national security. But some risks are clearly present: DeepSeek's chatbot application has demonstrated a tendency to suppress information on certain topics and amplify propaganda Beijing uses to discredit critics.[8] And there are reasons to think these risks will grow.

The new paradigm in AI development is "agents" — AI applications that can operate autonomously to achieve user-specified goals, like conducting research or maintaining software. If agents based on DeepSeek's models are deployed by American or allied organizations, this presents a clear vector for national security risks. A recent paper from the US AI developer Anthropic showed that models can be trained to act like "sleeper agents," behaving normally under certain conditions, but altering their behavior in pre-defined circumstances, such as after a certain date, and/or if the model detects it is deployed in an American organization or network.[9]

Late last year, the NSF launched a program on open-source security, whose due date for proposals is later this month.[10] Such programs should be expanded to tackle problems specific to open-source AI, including AI agents. I provide examples of where this research could focus below.

The US can't be caught off guard again

DeepSeek's research was publicly available. Its reliance on US chips was known. And yet its impact caught Washington off guard — much like Huawei's 2023 release of a smartphone powered by SMIC's 7nm chip, which revealed a blind spot in our understanding of Chinese manufacturing capacity.[11]

This means our initial reaction was informed only by DeepSeek's analysis of its capabilities, rather than our independent analysis. If the US is to outmaneuver China in a race for AI superiority, it needs to look further into the future and respond proactively.

The real issue is a lack of institutional technical capability. US policymakers need fast, expert analysis of foreign models, chips, and deployments — before they become headline events. That requires a dedicated team able to:

- Predict the capabilities of unreleased models, using weights, research, code, or indirect capability signals;
- Interpret foreign chip specs, firmware, and manufacturing disclosures;
- Integrate information from the intelligence community and top US model developers;
- Forecast national security implications across high-risk domains, particularly cyber and bio.
- Providing expert guidance on proactive policy decisions that require deep technical expertise, such as defining new technical parameters to use in export controls

This requires people who've worked on frontier-scale models. According to NSF data, the median age of federally employed scientists and engineers is over 45.[12] Past successful federal projects in emerging technologies look quite different. The median age of scientists and engineers working on the Manhattan Project was 25; on Apollo, 28.[13] The people building and stress-testing today's most capable models are often in their 20s, and almost entirely outside government.

One viable institutional home is NIST, which has flexible hiring authorities, relatively high pay scales, and technical credibility. While its core mission isn't national security, it could be tasked — via the Secretary of Commerce or National Security Council — to stand up or adapt a team (e.g., within the AI Safety Institute) focused on the above mission.

DeepSeek depends on American tech, and that gives us leverage

DeepSeek uses powerful computer chips developed by NVIDIA — an American firm — and its core algorithmic approach is based directly on breakthroughs made by American AI researchers.[14]

The fact that Chinese AI developers rely on American technologies gives Washington a powerful source of leverage in the form of export controls. DeepSeek's founder has openly acknowledged that "[their] problem has never been funding; it's the embargo on high-end chips."[15]

Although current export controls have gaps and flaws, DeepSeek's rise should not be taken as reason to abandon them; on the contrary, these export controls are the main — and perhaps only — obstacle to China achieving parity with the United States on AI capabilities.[16]

But while we should avoid handing the CCP technological advantages, export controls aren't a way to avoid competing altogether. For both open-source and proprietary models, America also needs to build the most capable, most reliable, and most secure AI, ensuring that US models are the models of choice for the world.

In doing so, the federal government should recognize it has an essential role to play in shaping the direction of technological development. In the past, the government has

shaped the direction of early internet and satellite technologies through DARPA, biomedical technology through the NIH, and genomic research through the Human Genome Project.

The full might of the American R&D engine has been a powerful force for aligning emerging technologies with American interests in the past, and it can be now for AI.

An R&D agenda for American leadership in AI

Within the field of AI, we have seen massive advances in fundamental capabilities, without equivalent advances in model robustness, interpretability, verification, and security. Private companies are less focused on these areas and more focused on discovering commercial applications. But the American public has a strong interest in ensuring that models are trustworthy in their application.

This also matters for American competitiveness and national security. If US models are more reliable than their foreign counterparts, it is more likely that American firms will be the provider of choice for the world, including in scientific applications that do not have strong commercial promise, but are important for soft power and advancing basic research.

The rapid deployment of AI systems for US military applications is also hindered by fundamental AI reliability challenges. Current systems lack transparency into their internal decision-making processes, exhibit unexpected behaviors when deployed in novel environments, and contain vulnerabilities across both software and hardware layers that could be exploited by sophisticated adversaries in contested environments.

Here, I'll lay out a range of promising R&D projects and funding mechanisms the federal government could use to solve these problems.

Research to support American open-source AI dominance

America's open-source AI ecosystem is strong, but not unchallenged. Based on publicly available information, DeepSeek's models have already been downloaded over 12 million times, not far behind Meta's, at 24 million.[17] Because many large organizations engage in lengthy processes to procure new software, the choice of an initial open-source model provider for internal use could be "sticky." The US must ensure it develops not only the best open-source chatbots but also the best open-source models used in areas such as science, robotics, and manufacturing.

Federal prize competitions are a proven tool for accelerating innovation and are already authorized under the America COMPETES Act. In 2024, federal R&D agencies under the jurisdiction of this committee offered over $60 million in prizes.[18] Prizes work particularly well for open-source AI:

- They can reward openness — by conditioning payouts on open-sourcing.
- They transfer downside risk — government only pays for successful outcomes.
- They encourage wide participation — essential in growing new fields.

- They can be targeted — restricted to American developers and focused on national-interest domains like drug discovery, materials science, and genomics, as well as reliability issues that hinder the adoption of American open-source models, such as methods to detect or remove backdoors, reduce misuse risk, or improve robustness in adversarial conditions.[19][20]

This model already works: the Vesuvius Challenge succeeded in building an open AI model that reads ancient carbonized scrolls — an impressive technical breakthrough, powered by a prize.[21]

A natural institutional home for such efforts is the proposed NIST Foundation, created under bipartisan legislation from this Committee.[22] A foundation could channel public, private, and philanthropic funds toward national goals, including open-source AI leadership.[23]

Finally, to boost adoption, the National AI Research Resource (NAIRR) should host American open-source models — subsidizing distribution for startups, researchers, and small firms.[24] Making US models easier to test, trust, and deploy is essential to securing long-term advantage.

A "Human Genome Project" for AI interpretability

AI interpretability research aims to develop a more concrete understanding of a model's predictions, decisions, or behavior.[25] Solving interpretability will allow for safer and more effective AI systems via more precise control, and the ability to detect and neutralize adversarial modifications such as hidden backdoors.

Early interpretability research suggests we may be on the cusp of meaningful theoretical breakthroughs.[26] However, the scale and urgency of this challenge demand a more ambitious approach than existing grant programs. A large-scale initiative — comparable in ambition to the Human Genome Project — could be instrumental in systematically understanding how today's AI models process information to exhibit particular capabilities.

Given the strong overlap of this work with defense interests (including increasing the reliability of AI models deployed in national security applications, and understanding the capabilities of adversary systems), this work could be supported with defense spending, targeting requirements set by the defense and intelligence communities.[27]

A "grand challenge" to develop new solutions could then be supported through proven, efficient funding mechanisms, such as prize competitions for novel interpretability research techniques, challenge-based acquisition programs, and advance market commitments, involving government commitments to purchase technical solutions that successfully meet certain criteria. These mechanisms powered Operation Warp Speed, which delivered a COVID vaccine in under a year, 10x faster than the normal speed.[28] The same model can work for interpretability — if we treat it with the urgency it deserves.

Ambitious investments in secure AI chips and data centers

Advanced AI systems depend on specialized chips whose integrity and security are essential, both for protecting our ability to develop AI models at home and for protecting critical infrastructure that will in the future depend on the reliability of AI systems, such as our energy and health systems.

Without robust protection of the chips on which AI models are built and deployed, America risks industrial espionage, sabotage, and a weakened ability to protect its supply chains. Today's AI chip security features, such as confidential computing, can already be used to protect sensitive data or to track whether a controlled AI chip has been smuggled after it has been exported overseas.[29] However, these technologies are vulnerable to physical attacks and to information leakage through "side channels", where an attacker can steal information from a chip by observing its electromagnetic emissions.

The US government is well-positioned to drive innovation in AI hardware security. Programs such as the National Semiconductor Technology Center (NSTC) run out of NIST, as well as NIST's existing hardware security standardization programs, can serve as focal points for accelerating research and implementation.[30]

To complement research and standardization work at the NSTC and NIST, the Department of Energy could pilot new, highly secure AI data centers. Existing secure data centers, such as those used for classified government operations, prioritize confidentiality and controlled access, but do not have strong performance and scale requirements. Advanced AI data centers operate at a much larger scale, with specialized hardware that is optimized for performance rather than security. This creates a security gap that must be addressed if America's economy is to increasingly rely on AI models.

The Department of Energy was an early adopter and enabler of GPUs for scientific computing applications. Today, with its expertise in building and operating cutting-edge computing infrastructure, it could also be the home for efforts to prototype *secure* infrastructure to enable the US economy to take advantage of the next industrial revolution.

Pre-deployment hardening for American critical infrastructure

A popular idea within AI policy communities is "pre-deployment evaluation" — running tests on AI models to ensure they are safe to release.[31] In a world where new AI capabilities are quickly open-sourced by developers across the world, relying on pre-deployment checks to prevent the proliferation of dangerous capabilities in areas like cyber and biological weapons is likely insufficient. Even if new dangerous capabilities are found, they are likely to proliferate regardless of the good intentions of firms and regulators.

With its police state and culture of heavy surveillance, China is likely better able to prevent its citizens from misusing open-source AI models. In the West, such an approach would undermine the values we are trying to protect. But we can use the innovativeness of our firms to solve this problem in a different way, and in doing so, improve our societal

resilience to a range of emerging risks. Because American firms are at the frontier of advances in AI, they can engage in "pre-deployment hardening" — proactively identifying the new risks that AI models might create and coordinating with governments to quickly roll out new protections.[32]

An urgent area for such activities is cybersecurity. Recent analysis from American frontier AI developers suggests that AI's cyberoffensive capabilities are rapidly improving — going from a 5 percent success rate at "capture the flag" tasks (where a model tries to find and exploit vulnerabilities in software) in 2023 to over 30 percent in 2024.[33]

Pre-deployment hardening for cybersecurity could involve US firms sharing information about newly discovered cyberoffensive capabilities in their models with the federal government and critical infrastructure providers, and then using their models to rapidly detect and patch vulnerabilities across the code bases used across the US economy and government. This should include open-source software libraries, which many critical applications depend upon, and for which Americans make up the lion's share of developers, five times more than any other country.[34]

This would be a nationwide undertaking, requiring coordination across both government and industry. In the near term, NIST, in collaboration with agencies such as CISA, could use its role as a convener of industry to create and standardize processes for pre-deployment hardening.

An energy agenda for American leadership in AI

Based on near-term trends, the computing infrastructure used to train frontier AI models will soon require multiple gigawatts of power – the equivalent of multiple nuclear power plants.[35]

To enable this, policymakers must unleash America's industrial capacity, reduce timelines for environmental permitting, and help developers take on the technical risks involved in an "all of the above" energy strategy, including scaling next-generation energy technologies such as small modular reactors and enhanced geothermal. Without this, even if the top AI labs are in the United States, they may be forced to either develop their models abroad, subject to foreign regulations and security risks, or not at all.

The Department of Energy (DOE) has an important role to play in this endeavor. As part of its responsibilities under January's Executive Order focused on accelerating the development of AI data centers within a 2-year time frame, DOE recently put out an RFI to industry focused on leasing out its land for AI data centers, creating what we have called "Special Compute Zones".[3637]

DOE should also use its authorities to ensure that AI data centers and associated energy infrastructure built on federal lands aren't held up by red tape and are built securely enough to protect critical American technologies from theft or sabotage. Specifically, the agency could:

- Identify existing energy assets (such as retired coal sites) that could be upgraded or repurposed under the Department of Energy's Loan Programs Authority (Section 1706 of Title XVII of the Energy Policy Act of 2005).
- Identify categorical exclusions to environmental permitting that can be adopted both by the DOE and by other agencies aiming to lease their land for AI infrastructure.[38] For example, DOE should establish categorical exclusions for early-stage project costs like design and site characterization, activities with minimal environmental impact, such as materials acquisition, and projects on previously disturbed lands.
- Create new security requirements for AI data centers focused on risks from nation-state-level actors, protecting American models from theft (which would drive China's own AI development), and protecting the US economy from industrial sabotage.

Strategic talent identification for US AI leadership

The US cannot lead in AI, semiconductors, or quantum if we lose the global race for scientific talent. A single defection can tilt an entire field: China's dominance in 5G stems, in part, from Huawei's use of polar codes — a breakthrough by MIT PhD Erdal Arikan, who returned to Turkey rather than being recruited to stay in the US.[39]

Meanwhile, China is no longer just trying to bring its diaspora home — it's now recruiting international scientists directly through programs like Qiming.[40] The US has no counterpart: recruitment is ad hoc, left to universities and firms, and largely blind to defense relevance.

This is a gap the Science Committee is well-positioned to address — not through immigration reform, but by ensuring that federally supported science agencies build robust talent identification capacity, particularly in AI and defense-relevant fields.

Congress could:

- Direct agencies to contract with FFRDCs to develop tools for mapping the global AI talent landscape, including predictive analytics to identify researchers whose work suggests outsized future impact;
- Incentivize principal investigators at National Labs, NSF-funded AI institutes, and affiliated universities to share regular updates on standout foreign researchers working in AI and robotics;
- Support the creation of a secure, centralized talent database, continuously updated and available to relevant executive agencies, that flags recruitment opportunities across high-priority technical domains.

Understanding where frontier AI talent resides — and how to reach it — should be seen as a basic responsibility of the US research enterprise. The Committee has the authority and oversight reach to make talent intelligence a core function of our national science infrastructure.

[1] National Science Foundation, "NSF Partners with the Institute for Progress to Test New Mechanisms for Funding Research and Innovation," September 28, 2023, https://www.nsf.gov/news/nsf-partners-institute-progress-test-new.
[2] Tim Fist et al., "An Action Plan for American Leadership in AI," Institute for Progress, March 17, 2025, https://ifp.org/an-action-plan-for-american-leadership-in-ai/.
[3] "Most Popular AI Tools," aitools.xyz, February 2025, https://aitools.xyz/popular-ai-tools/2025/february.
[4] "DeepSeek" "Sputnik" Google Search, accessed April 5, 2025, https://www.google.com/search?q=%22deepseek%22+%22sputnik%22.
[5] V3 is roughly equivalent in benchmark performance to US company Anthropic's Claude 3.5 Sonnet, which reportedly concluded training around 8 months before V3 ("LLM Leaderboard," LLM-Stats.com, accessed April 5, 2025, https://llm-stats.com/; Dario Amodei, "On DeepSeek and Export Controls," January 2025, https://darioamodei.com/on-deepseek-and-export-controls). It's likely true that V3 was cheaper to train than those models — DeepSeek's paper claims $5.6 million in cost of computer chips (DeepSeek AI, "DeepSeek-V3 Technical Report," arXiv, February 18, 2025, https://arxiv.org/pdf/2412.19437), which is about 10 times less than the cost of equivalent American models developed 8 months earlier. Claude 3.5 Sonnet has been estimated as using 3.65 x 10^25 FLOP. Using the same hardware available to DeepSeek — NVIDIA H800 GPUs — this would require around 20,000 GPUs running for 3 months. Running the same calculation for V3 yields around 2,000 GPUs running for 3 months (Epoch AI, "Notable AI Models," last updated April 4, 2025, https://epoch.ai/data/notable-ai-models?view=table; Weile Luo et al., "Benchmarking and Dissecting the Nvidia Hopper GPU Architecture," arXiv, February 21, 2024, https://arxiv.org/html/2402.13499v1). But we must remember that exponential cost reduction is the norm in the AI industry. Each year, as an average across the industry, it becomes around 4 times cheaper to develop a model at a fixed level of performance. See: Anson Ho et al., "Algorithmic Progress in Language Models," Epoch AI, March 12, 2024, https://epoch.ai/blog/algorithmic-progress-in-language-models; Marius Hobbahn, Lennart Heim, and Gökçe Aydos, "Trends in Machine Learning Hardware," Epoch AI, November 9, 2023, https://epoch.ai/blog/trends-in-machine-learning-hardware. This means that V3, while impressively efficient, wasn't too far off the overall cost-efficiency trend.
[6] Department of State, "Military-Civil Fusion and the People's Republic of China," accessed April 5, 2025, https://www.state.gov/wp-content/uploads/2020/05/What-is-MCF-One-Pager.pdf.
[7] Charles Q. Choi, "Undetectable Backdoors Plantable In Any Machine-Learning Algorithm," May 10, 2022, https://spectrum.ieee.org/machine-learningbackdoor; Evan Hubinger et al., "Sleeper Agents: Training Deceptive LLMs That Persist Through Safety Training," arXiv, January 17, 2024, https://arxiv.org/pdf/2401.05566.
[8] Donna Lu, "We tried out DeepSeek. It worked well, until we asked it about Tiananmen Square and Taiwan," *The Guardian*, January 28, 2025, https://www.theguardian.com/technology/2025/jan/28/we-tried-out-deepseek-it-works-well-until-we-asked-it-about-tiananmen-square-and-taiwan; Steven Lee Myers, "DeepSeek's Answers Include Chinese Propaganda, Researchers Say," *The New York Times,* January 31, 2025, https://www.nytimes.com/2025/01/31/technology/deepseek-chinese-propaganda.html.
[9] Anthropic, "Sleeper Agents: Training Deceptive LLMs that Persist Through Safety Training," January 14, 2024, https://www.anthropic.com/research/sleeper-agents-training-deceptive-llms-that-persist-through-safety-training.
[10] National Science Foundation, "Safety, Security, and Privacy of Open-Source Ecosystems (Safe-OSE)," accessed April 6, 2025, https://www.nsf.gov/funding/opportunities/safe-ose-safety-security-privacy-open-source-ecosystems.
[11] Jeff Pao, "SMIC Bypasses US Curbs to Make 7nm Chips," *Asia Times*, September 5, 2023, https://asiatimes.com/2023/09/smic-bypasses-us-curbs-to-make-7nm-chips/.

[12] Jaquelina C. Falkenheim, "Federal Scientists and Engineers," National Center for Science and Engineering Statistics, March 2022, https://ncses.nsf.gov/pubs/ncses22204/assets/federal-scientists-and-engineers/ncses22204.pdf.
[13] Pratyush Buddiga [@pratyushbuddiga], "Average age of engineers and scientists in the Manhattan Project was 25," X, February 3, 2025, https://x.com/pratyushbuddiga/status/1886472301811785903; Joe Carter, "5 Facts About the Apollo 11 Moon Landing," Acton Institute, July 18, 2019, https://rlo.acton.org/archives/110300-5-facts-about-the-apollo-11-moon-landing.html;
[14] Specifically, DeepSeek-V3 uses the "Transformer" architecture, developed by researchers at Google (Ashish Vaswani et al., "Attention Is All You Need," arXiv, submitted on June 12, 2017, https://arxiv.org/abs/1706.03762).
[15] Jordan Schneider et al., "DeepSeek: The Quiet Giant Leading China's AI Race," ChinaTalk, November 27, 2024, https://www.chinatalk.media/p/deepseek-ceo-interview-with-chinas
[16] Tim Fist et al., "An Action Plan for American Leadership in AI."
[17] "DeepSeek," Hugging Face, accessed April 5, 2025, https://huggingface.co/deepseek-ai; "Meta Llama," Hugging Face, accessed April 5, 2025, https://huggingface.co/meta-llama.
[18] General Services Administration, "Archived Challenges," Challenge.gov, https://www.challenge.gov/?state=archived.
[19] Miryam Naddaf, "AI Tool Diagnoses Diabetes, HIV and COVID from a Blood Sample," *Nature*, February 20, 2025, https://www.nature.com/articles/d41586-025-00528-y; Wei Chen et al., "Artificial Intelligence for Drug Discovery: Resources, Methods, and Applications," Molecular Therapy - Nucleic Acids, February 18, 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC10009646/; Kim Martineau, "Meet IBM's New Family of AI Models for Materials Discovery," IBM, December 20, 2024, https://research.ibm.com/blog/foundation-models-for-materials; Arc Institute, "AI Can Now Model and Design the Genetic Code for All Domains of Life with Evo 2," February 19, 2025, https://arcinstitute.org/news/blog/evo2.
[20] James Pomfret and Jessie Pang, "Exclusive: Chinese Researchers Develop AI Model for Military Use on Back of Meta's Llama," Reuters, November 1, 2024, https://www.reuters.com/technology/artificial-intelligence/chinese-researchers-develop-ai-model-military-use-back-metas-llama-2024-11-01/.
[21] Vesuvius Challenge, "Vesuvius Challenge 2023 Grand Prize awarded" February 5, 2024, https://scrollprize.org/grandprize.
[22] Madison Alder, "Bipartisan Bill to Foster Public-Private Partnerships for NIST Reintroduced in House, Senate," FedScoop, April 1, 2025, https://fedscoop.com/public-private-partnerships-bill-nist-house-senate/.
[23] Tim Fist, "The Case for a NIST Foundation: Four Ways a Non-Profit Foundation Could Supercharge NIST's Work on Emerging Technologies," Institute for Progress, June 18, 2024, https://ifp.org/nist-foundation/.
[24] "Recommendation 3: Host US Open-Source Models on the NAIRR" in Tim Fist et al., "An Action Plan for American Leadership in AI," Institute for Progress, March 17, 2025, https://arc.net/l/quote/dhgpeuxr.
[25] Lee Sharky and Bilal Chughtai et al., "Open Problems in Mechanistic Interpretability," arXiv, January 27, 2025, https://arxiv.org/pdf/2501.16496.
[26] Scott J. Mulligan, "Google DeepMind Has a New Way to Look Inside an AI's 'Mind'," MIT Technology Review, November 14, 2024, https://www.technologyreview.com/2024/11/14/1106871/google-deepmind-has-a-new-way-to-look-inside-an-ais-mind/; Kevin Roose, "A.I.'s Black Boxes Just Got a Little Less Mysterious", *The New York Times,* May 21, 2024, https://www.nytimes.com/2024/05/21/technology/ai-language-models-anthropic.html.
[27] Renaissance Philanthropy, "Target Product Profiles," accessed April 5, 2025, https://renaissancephilanthropy.org/playbooks/target-product-profile/.

[28] Gaurav Agrawal et al., "Fast-forward: Will the Speed of COVID-19 Vaccine Development Reset Industry Norms?," McKinsey & Company, May 13, 2021, https://www.mckinsey.com/industries/life-sciences/our-insights/fast-forward-will-the-speed-of-covid-19-vaccine-development-reset-industry-norms

[29] Tim Fist, Tao Burga, and Vivek Chilukuri, "Technology to Secure the AI Chip Supply Chain: A Working Paper," Center for a New American Security, December 11, 2024, https://www.cnas.org/publications/reports/technology-to-secure-the-ai-chip-supply-chain-a-primer

[30] This could include the NSTC coordinating its members to identify and standardize solutions to system-level vulnerabilities in AI chips and data centers, and using its role as a publicly subsidized consortium to prioritize making relevant intellectual property widely accessible to strengthen industry security. Such programs could also involve NIST, in collaboration with industry, collating existing hardware security standards and identifying and addressing gaps when applying them to AI chips and servers deployed in different operating environments.

[31] Toby Shevlane et al., "Model Evaluation for Extreme Risks," arXiv, May 24, 2023, https://cdn.governance.ai/Model_Evaulations_for_Extreme_Risks.pdf.

[32] This approach was also recently described as "adaptation buffers" (Helen Toner, "Nonproliferation Is the Wrong Approach to AI Misuse," Rising Tide (Substack), April 5, 2025, https://helentoner.substack.com/p/nonproliferation-is-the-wrong-approach.

[33] Anthropic, "Progress From Our Frontier Red Team," March 19, 2025, https://www.anthropic.com/news/strategic-warning-for-ai-risk-progress-and-insights-from-our-frontier-red-team.

[34] Johannes Wachs, "The Geography of Open Source Software: Evidence from GitHub," *Technological Forecasting and Social Change,* March 2022, https://www.sciencedirect.com/science/article/pii/S0040162522000105?via%3Dihub.

[35] "The AI Data Centers of the Future" in Tim Fist and Arnab Datta, "How to Build the Future of AI in the United States," Institute for Progress, October 23, 2024, https://ifp.org/future-of-ai-compute/#the-ai-data-centers-of-the-future.

[36] Department of Energy, "Request for Information on Artificial Intelligence Infrastructure on DOE Lands," April 3, 2025, https://www.energy.gov/sites/default/files/2025-04/RFI%20to%20Inform%20Public%20Bids%20to%20Construct%20AI%20Infrastructure%20%28website%20copy%29.pdf.

[37] Arnab Datta and Tim Fist, "Compute in America: A Policy Playbook," Institute for Progress, February 3, 2025, https://ifp.org/special-compute-zones.

[38] Following the 2023 Fiscal Responsibility Act, agencies can adopt categorical exclusions issued by other agencies.

[39] Graham Allison and Eric Schmidt, "The US Needs a Million Talents Program to Retain Technology Leadership," Foreign Policy, July 16, 2022, https://foreignpolicy.com/2022/07/16/immigration-us-technology-companies-work-visas-china-talent-competition-universities/.

[40] Julie Zhu et al., "Insight: China Quietly Recruits Overseas Chip Talent as US Tightens Curbs," Reuters, August 24, 2023, https://www.reuters.com/technology/china-quietly-recruits-overseas-chip-talent-us-tightens-curbs-2023-08-24/