ANTHROP\C

548 Market St., PMB 90375
San Francisco, CA 94104-5401

412-837-9797
anthropic.com

March 6, 2025

Faisal D'Souza, NCO
Office of Science and Technology Policy
Executive Office of the President
2415 Eisenhower Avenue
Alexandria, VA 22314

Submitted by email to ostp-ai-rfi@nitrd.gov

Re: Request for Information (RFI) on the Development of an Artificial Intelligence (AI) Action Plan ("Plan")

Introduction

Anthropic strongly supports OSTP's effort to define the priority policy actions needed to maintain and strengthen America's dominance and leadership in artificial intelligence. In this submission, we outline specific actions the administration should take to ensure that America maximally benefits from the potential of advanced AI systems.

Our recommendations encompass two categories: (1) national security imperatives to strengthen U.S. security by safeguarding vital technological infrastructure and intellectual assets from foreign threats; and (2) investments the U.S. government should make to cultivate a robust AI development and deployment ecosystem that fosters American prosperity and ensures that AI-driven economic benefits are widely shared across society.

Powerful AI technology will be built during this Administration. Given the rapid pace of development, it is imperative that this technology be treated as a critical national asset through a targeted AI Action Plan that strengthens American economic competitiveness while bolstering our national security.

About Anthropic

Anthropic is a leading frontier AI model developer working to build reliable, interpretable, and steerable artificial intelligence systems. Our flagship AI assistant, Claude, represents the state

of the art in Large Language Model (LLM) technology. As one of the ten most valuable private companies in the U.S., we conduct cutting-edge artificial intelligence research and deploy our research discoveries as products that serve millions of Americans. Our customers, ranging from Fortune 500 companies and U.S. government agencies to small businesses and consumers, use Claude as an AI co-pilot to enhance productivity on sophisticated tasks including software development, data analysis, and scientific research. In February 2025, we released Claude 3.7 Sonnet, which is by many performance benchmarks the most powerful and capable commercially-available AI system in the world.

What we mean by "powerful AI"

When we discuss "powerful AI", we are referring to systems that represent a major advancement beyond today's AI model capabilities. Powerful AI systems will have the following properties:

- Intellectual capabilities matching or exceeding that of Nobel Prize winners across most disciplines—including biology, computer science, mathematics, and engineering.
- The ability to navigate all interfaces available to a human doing digital work today, including the ability to process and generate text, audio, and video, the ability to autonomously control technology instruments like mice and keyboards, and the ability to access and browse the internet.
- The ability to autonomously reason through complex tasks over extended periods—hours, days, or even weeks—seeking clarification and feedback when needed, much like a highly capable employee would.
- The ability to interface with the physical world; controlling laboratory equipment, robotic systems, and manufacturing tools through digital connections.

A useful conceptual framework is to envision powerful AI as equivalent to "a country of geniuses in a datacenter"—a concentration of intellectual capability that fundamentally transforms our understanding of what is possible.

Based on current research trajectories, we anticipate that powerful AI systems could emerge as soon as late 2026 or 2027. Anthropic's CEO Dario Amodei explores these developments and their implications more thoroughly in his essay *Machines of Loving Grace*.[1]

Taking the opportunities and risks of powerful AI seriously

We expect that the economic and national security implications of this technology will be tremendous. As such, the U.S. government should be prepared to implement ambitious policy responses to effectively navigate this rapid technological transition.

[1] Dario Amodei, Machines of Loving Grace (Oct. 2024), *available at:* https://darioamodei.com/machines-of-loving-grace (accessed Mar. 6, 2025).

Anthropic recommends that the Administration focus on two strategic imperatives. First, it must strengthen national security frameworks both during the critical development phase of powerful AI and in the new technological landscape that emerges afterward. Second, it must pursue policies that enable these systems to generate broad-based American prosperity, ensuring that productivity gains and new capabilities benefit citizens across all segments of society. This dual focus will position the United States to lead in an era where computational intelligence becomes a defining element of national strength and societal wellbeing.

Enhancing American National Security

Build the federal government's capacity to test and evaluate powerful AI models for national security capabilities

As artificial intelligence systems grow increasingly sophisticated, they will present significant national security implications—on one hand, potentially enabling adversaries to pursue harmful objectives; on the other, providing the United States with strategic advantages when properly leveraged. **To optimize national security outcomes, the federal government must develop robust capabilities to rapidly assess any powerful AI system, foreign or domestic, for potential national security uses and misuses.** Establishing this evaluation infrastructure provides national security agencies with critical insight into emerging model capabilities, enabling the identification of potential threats before they materialize.

We anticipate dramatic capability advancements in frontier AI models over the next 2-4 years, particularly in domains with significant security implications including biological weapon and cybersecurity risks.[2] Ultimately, we believe if there is evidence that AI systems pose critical national security risks then developers like Anthropic should be required[3] to test their systems for these risks—but to be able to make that decision, we must first equip the U.S. government with the capacity to generate and deliberate about this evidence.

Our most recent system, Claude 3.7 Sonnet, demonstrates concerning improvements in its capacity to support aspects of biological weapons development[4]—insights we uncovered through our internal testing protocols and validated through voluntary security exercises conducted in partnership with the U.S. and U.K. AI Safety and Security Institutes.[5] This trajectory suggests, consistent with scaling laws research, that numerous AI systems will increasingly embody significant national security implications in the coming years. Thus, making

---

[2] Y. Bengio et al., "International AI Safety Report" (January 2025), *available at:* https://assets.publishing.service.gov.uk/media/679a0c48a77d250007d313ee/International_AI_Safety_Report_2025_accessible_f.pdf (accessed Mar. 6, 2025)

[3] The case for targeted regulation https://www.anthropic.com/news/the-case-for-targeted-regulation (accessed Mar. 6, 2025)

[4] Claude 3.7 Sonnet System Card (Feb. 24, 2025), *available at:* https://assets.anthropic.com/m/785e231869ea8b3b/original/claude-3-7-sonnet-system-card.pdf (accessed Mar. 6, 2025).

[5] Claude 3.5 Sonnet (Jun. 20, 2024), *available at:* https://www.anthropic.com/news/claude-3-5-sonnet (accessed Mar. 6, 2025).

comprehensive government awareness is imperative, particularly as China advances its efforts to build powerful dual-use AI systems.

The critical importance of robust evaluation capabilities was highlighted by the release of DeepSeek R1—a Chinese AI model freely distributed online—earlier this year. While DeepSeek itself does not demonstrate direct national security-relevant capabilities, early model evaluations conducted by Anthropic showed that R1 complied with answering most biological weaponization questions, even when formulated with a clearly malicious intent. This highlights the crucial importance of equipping the U.S. government with the capacity to rapidly evaluate whether future models—foreign or domestic—released onto the open internet possess security-relevant properties that merit national security attention.

To establish effective government testing and evaluation for capabilities of powerful AI systems, the administration should:

- Preserve the AI Safety Institute in the Department of Commerce and build on the MOUs it has signed with U.S. AI companies—including Anthropic—to advance the state of the art in third-party testing of AI systems for national security risks.
- Direct the National Institutes of Standards and Technology (NIST), in consultation with the Intelligence Community, Department of Defense, Department of Homeland Security, and other relevant agencies, to develop comprehensive national security evaluations for powerful AI models, in partnership with frontier AI developers, and develop a protocol for systematically testing powerful AI models for these vulnerabilities.
- Ensure that the federal government has access to the classified cloud and on-premises computing infrastructure needed to conduct thorough evaluations of powerful AI models.
- Build a team of interdisciplinary professionals within the federal government with national security knowledge and technical AI expertise to analyze potential security vulnerabilities and assess deployed systems.

Hardening export controls to widen the U.S. AI lead

Export controls on semiconductors, semiconductor tooling, and certain model weights represent critical policy instruments that will slow adversarial development of powerful AI and enable the United States to maintain a durable lead in frontier AI development. The semiconductor restrictions initiated during President Trump's first term and subsequently enhanced have effectively constrained asymmetric state competitors from achieving significant advances in AI development capabilities. This policy approach demonstrates the efficacy of precisely targeted export controls on critical technologies.

To prevent advanced AI models and AI infrastructure from being acquired by adversaries, **we strongly recommend the administration strengthen export controls on computational resources and implement appropriate export restrictions on certain model weights.** Amongst other things, we recommend the Administration consider:

- **Controlling the H20 chips.** Current export controls do not apply to the H20, a high-memory chip introduced in 2024 for sale to China that can be used to train and run powerful models. While these chips underperform H100s for initial training, they excel at text generation ("sampling")—a fundamental component of advanced reinforcement learning methodologies critical to current frontier model capability advancements. The Trump administration has an opportunity to close this loophole.

- **Requiring countries to sign government-to-government agreements outlining measures to prevent smuggling.** As a prerequisite for hosting data centers with more than 50,000 chips from U.S. companies, the U.S. should mandate that countries at high-risk for chip smuggling comply with a government-to-government agreement that 1) requires them to align their export control systems with the U.S., 2) takes security measures to address chip smuggling to China, and 3) stops their companies from working with the Chinese military. The Department of Commerce's January 2025 Interim Final Rule on the Framework for Artificial Intelligence Diffusion (the "Diffusion Rule") already contains the possibility for such agreements, laying a foundation for further policy development.

- **Closely examine and reduce the 1,700 H100 no-license required threshold for orders to Tier 2 countries in the Diffusion Rule.** Currently, the Diffusion Rule allows advanced chip orders from Tier 2 countries for less than 1,700 H100s —an approximately $40 million order—to proceed without review. These orders do not count against the Rule's caps, regardless of the purchaser. While these thresholds address legitimate commercial purposes, we believe that they also pose smuggling risks. We recommend that the Administration consider reducing the number of H100s that Tier 2 countries can purchase without review to further mitigate smuggling risks. Determining the optimal lower threshold would require comprehensive analysis balancing smuggling prevention against commercial facilitation. To determine a revised figure, we recommend the determination be made by the four members of the End-User Review Committee.

- **Increase funding for the Bureau of Industry and Security (BIS) for export enforcement.** Export controls are only effective with proper enforcement. A thorough assessment of BIS's current enforcement capabilities and the potential benefits of additional resources would significantly enhance the overall effectiveness of these controls.

Dramatically improve the security of U.S. frontier labs

As discussed earlier, the next generations of powerful AI models will be highly capable and economically valuable, and thus prime targets for misuse and abuse by bad actors. Today's individual AI systems already command valuations in the high billions of dollars, making them attractive targets for sophisticated threat actors. Moreover, these AI systems represent the product of computational resources that have been strategically denied to certain nations through export controls, doubling their appeal as valuable theft targets. Most critically, the

successful theft of even a single frontier model could significantly harm the entire export control regime, providing adversaries immediate access to the refined intellectual property derived from restricted computational resources. To mitigate these risks, the federal government should partner with industry leaders to **substantially enhance security protocols at frontier AI laboratories to prevent adversarial misuse and abuse of powerful AI technologies**.

To achieve this, we strongly recommend the Administration:

- Establish classified and unclassified communication channels between American frontier AI laboratories and the Intelligence Community for threat intelligence sharing, similar to Information Sharing and Analysis Centers used in critical infrastructure sectors. This should include both traditional cyber threat intelligence, as well as broader observations by industry or government of malicious use of models, especially by foreign actors.
- Create systematic collaboration between frontier AI companies and the Intelligence Community agencies, including Five Eyes partners, to monitor adversary capabilities.
- Elevate collection and analysis of adversarial AI development to a top intelligence priority, as to provide strategic warning and support export controls.
- Expedite security clearances for industry professionals to aid collaboration.
- Direct NIST to develop next-generation cyber and physical security standards specific to AI training and inference clusters.
- Direct NIST to develop technical standards for confidential computing technologies that protect model weights and user data through encryption even during active processing.
- Develop meaningful incentives for implementing enhanced security measures via procurement requirements for systems supporting federal government deployments.
- Direct DOE/DNI to conduct a study on advanced security requirements that may become appropriate to ensure sufficient control over and security of highly agentic models.

By implementing this comprehensive security framework, the federal government will substantially strengthen the defensive posture of American AI companies, and significantly impede the ability of bad actors to misappropriate cutting-edge American technology and weaponize it against U.S. interests.

Promoting American Prosperity

Securing and scaling up U.S. energy supply

We commend the Trump Administration's efforts to expand domestic energy supply as a Day 1 priority. We project that by 2027, training a single frontier AI model will require networked computing clusters drawing approximately five gigawatts of power. When multiplied across all leading AI developers, the United States will need to deploy tens of additional gigawatts of energy capacity within the next three years to maintain its competitive edge and support AI model development at home. Furthermore, beyond the energy to train AI models, substantial power will be needed to operate inference clusters that deploy these models for applications throughout business, government, and society.

The federal government should consider establishing an ambitious national target: **build 50 additional gigawatts of power dedicated to the AI industry by 2027.** To achieve this, the federal government should:

- Task federal agencies with streamlining permitting processes by accelerating reviews, enforcing timelines, and promoting inter-agency coordination to eliminate bureaucratic bottlenecks.
- Direct federal agencies to expedite transmission line approvals to rapidly connect new energy sources to new data centers.
- Work with state and local governments to reduce permitting burdens for new energy and data center construction.
- Consider allocation of existing federal funding toward strategic energy infrastructure projects.
- Explore opportunities to leverage federal real estate for co-locating power generation facilities and next-generation data centers.

Failing to address these energy requirements presents serious risks to America's technological leadership, as U.S. AI developers may be forced to relocate operations overseas or postpone crucial R&D activities until sufficient domestic capacity becomes available. This scenario could effectively transfer the infrastructure foundation of our AI economy to foreign competitors and risk that the world's leading AI technology will not represent American values. Indeed, some authoritarian regimes who do not share our country's democratic values and may pose security threats are already actively courting American AI companies with promises of abundant, low-cost energy.[6] If U.S. developers migrate model development or storing of model weights to these countries in order to access these energy sources, this could expose sensitive intellectual property to transfer or theft, enable the creation of AI systems without proper security protocols, and potentially subject valuable AI assets to disruption or coercion by foreign powers.

Promoting rapid AI procurement across the federal government

One of AI's most transformative impacts will be revolutionizing government operations—both in delivering benefits to Americans and in supporting U.S. national security imperatives. However, realizing these benefits requires ensuring government institutions can effectively implement and utilize these technologies. We propose an ambitious initiative: **across the whole of government, the Administration should systematically identify every instance where federal employees process text, images, audio, or video data, and augment these workflows with appropriate AI systems.**

---

[6] *See, e.g.*, Hannah Jo Uy, "How the Middle East is Emerging as a Data Center Powerhouse Amid Booming AI Demand," MIT Sloane Management Review (Oct. 23, 2024, *available at*: https://www.mitsloanme.com/article/how-the-middle-east-is-emerging-as-a-data-center-powerhouse-amid-booming-ai-demand/ (accessed Mar. 6, 2025); Marissa Newman, Mark Bergen, and Olivia Solon, "Race for AI Supremacy in Middle East is Measured in Data Centers," Bloomberg (Nov. 4, 2024), *available at*: https://www.bloomberg.com/news/articles/2024-04-11/race-for-ai-supremacy-in-middle-east-is-measured-in-data-centers?embedded-checkout=true (accessed Mar. 6, 2025).

If achieved, this initiative would effectively provide every government worker with an AI-powered assistant, dramatically increasing productivity and effectiveness. Starting by conducting a comprehensive inventory and initial implementation processes will naturally reveal the institutional barriers currently impeding effective AI adoption across federal agencies.

To execute this vision, the Administration should pursue several key actions:

- The White House should task the Office of Management and Budget (OMB) to work with Congress to rapidly address resource constraints, procurement limitations, and programmatic obstacles to federal AI adoption, incorporating provisions for substantial AI acquisitions in the President's Budget.
- Coordinate a cross-agency effort to identify and eliminate regulatory and procedural barriers to rapid AI deployment at the federal agencies, for both civilian and national security applications.
- Direct the Department of Defense and the Intelligence Community to use the full extent of their existing authorities to accelerate AI research, development, and procurement.
- Identify the largest programs in civilian agencies where AI automation or augmentation can deliver the most significant and tangible public benefits—such as streamlining tax processing at the Internal Revenue Service, enhancing healthcare delivery at the Department of Veterans Affairs, reducing delays due to documentation processing at Health and Human Services, or reducing backlogs at the Social Security Administration.

We also encourage the White House to leverage existing frameworks to enhance federal procurement for national security purposes, particularly the directives in the October 2024 National Security Memorandum (NSM) on Artificial Intelligence and the accompanying Framework to Advance AI Governance and Risk Management in National Security. These frameworks enable piloting frontier AI applications in government and the development of robust testing and evaluation capabilities to build confidence in these systems.

Additionally, we strongly advocate for the creation of a joint working group between the Department of Defense and the Office of the Director of National Intelligence to develop recommendations for the Federal Acquisition Regulatory Council (FARC) on accelerating procurement processes for AI systems while maintaining rigorous security and reliability standards. The FARC should then consider appropriate amendments to the Federal Acquisition Regulation based on these recommendations to create a procurement environment that balances innovation with responsible governance.

Monitoring the economic impacts of AI and preparing for major actions

Technology equivalent to a "country of geniuses inside a datacenter" will fundamentally transform our economy. To ensure Americans thrive during this transition, the government must **vigilantly monitor economic indicators and industry developments.**

Our recommendations draw from our own work. In February, we launched the Anthropic Economic Index in February, an initiative designed to assess AI's impact on labor markets by correlating our usage data with the Bureau of Labor Statistics (BLS) O*NET framework.[7] While acknowledging the inherent limitations of our dataset and methodology, this approach provides automated, granular insights into AI's evolving economic role and identifies early indicators of future impacts as these systems advance in capability. We anticipate that 2025 will likely mark the beginning of more visible, large-scale economic effects from AI technologies.

The Administration can enhance governmental data collection mechanisms to better capture AI adoption patterns and economic implications and prepare for the possibility of significant change by exploring the following ideas:

- The Census Bureau's American Time Use Survey should incorporate specific questions about AI usage, distinguishing between personal and professional applications while gathering detailed information about task types and systems employed.
- The Census Bureau's Annual Business Survey should include more detailed questions about the types of AI technology used by businesses, along with the tasks performed by or with the assistance of AI technologies, while also disaggregating between LLMs and other forms of AI products.
- The Annual Business Survey should increase attention to collection of data on the tasks performed by workers, making them compatible with O*NET, forming the basis of a more rigorous understanding of the distribution of tasks in the U.S. economy.
- We believe that computing power will become an increasingly significant driver of economic growth. Accordingly, the White House should track the relationship between investments in AI computational resources and economic performance to inform strategic investments in domestic infrastructure and related supply chains.
- The White House should engage with Congress on and task relevant agencies with examining how AI adoption might reshape the composition of the national tax base, and ensure the government maintains visibility into potential structural economic shifts.

This forward-looking approach to understanding the economic transition that AI will bring will help ensure the benefits of technological advancement are broadly shared across American society.

[7] Anthropic Economic Index (Feb. 10, 2025), *available at:* https://www.anthropic.com/news/the-anthropic-economic-index (accessed Mar. 6, 2025).

Conclusion

We believe that extremely powerful AI technology will be developed during this Administration. President Trump and this Administration have a tremendous opportunity to enable American dominance in AI by unlocking the benefits of this technology and ensuring that America remains dominant over China and our other adversaries.

Throughout this submission we have advocated for the federal government to have access to better and more information about the national security and economic impacts of AI, as well as the security practices of the AI companies themselves. We believe it is essential that this administration finds ways to encourage AI companies to share more information about how they develop and test their models, along with the security practices they use to prevent them being stolen. The more we can encourage companies to arrive at a common standard of information disclosure, the easier it will be for the federal government to have the information it needs to tackle the next few years—years which we believe will be pivotal for AI development.

It is imperative that we approach the possibility of rapid technological development with appropriate seriousness and implement strategic measures today to ensure the American people fully benefit from these advancements. Simultaneously, we must leverage the comprehensive capabilities of the federal government to establish robust security frameworks that safeguard these transformative technologies. By acting decisively now, we can position the United States to lead in this technological revolution, while maintaining the security essential to our national interests.