CENTER FOR DEMOCRACY & TECHNOLOGY

March 14, 2025

To: Faisal D'Souza
NCO/NITRD
2415 Eisenhower Avenue
Alexandria, VA 22314

Re: AI Action Plan

This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in developing the AI Action Plan and associated documents without attribution.

I. Introduction

The Center for Democracy & Technology (CDT) respectfully submits these comments in response to the Networking and Information Technology Research and Development National Coordination Office's request for information on the highest priority actions that should be in the new AI Action Plan required under Executive Order 14179. CDT is a nonpartisan, nonprofit organization that works to advance civil rights and civil liberties in the digital age.

To solidify America's global leadership on AI, a strong governance plan is necessary to support broader AI adoption that is accountable to the American public. During his first term, President Trump established principles for AI governance that have been widely accepted on a bipartisan basis and that should serve as a foundation of the AI Action Plan. He initially issued Executive Order 13859, which stated that the U.S. must "drive development of appropriate technical standards and reduce barriers to safe testing and deployment of AI"[1] and "foster public trust and confidence in AI technologies and protect civil liberties, privacy, and American values in their application in order to fully realize the potential of AI technologies for the American people."[2] Executive Order 13960 subsequently specified that, in their own acquisition and use of AI, federal agencies should prioritize privacy, civil rights, and civil liberties; risk management; accuracy, reliability, and effectiveness; safety; transparency, understandability, and documentation; performance monitoring and corrective measures; and accountability.[3] These important criteria remain the building blocks of effective design and use of AI systems today.

[1] Exec. Order 13859, Maintaining American Leadership in Artificial Intelligence, Sec. 1(b), Feb. 14, 2019.
[2] *Id.* at Sec. 1(d). Sec. 2(c) of Executive Order 13859 also identified this principle as a core strategic objective for federal agencies.
[3] Exec. Order 13960, *Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*, Sec. 3, Dec. 3, 2020.

Following up on Executive Order 13859, the Office of Management and Budget's Memorandum M-21-06 provided guidance to federal agencies on the development of regulatory and non-regulatory approaches to private sector uses of AI.[4] That Memorandum advised agencies to carefully assess and address risks to people's privacy, individual rights and civil liberties, personal choice, civil rights, public health, safety, and security. It also counseled agencies to consider regulatory and non-regulatory approaches that incorporate public input, advance fairness and non-discrimination in outcomes and decisions produced by AI applications, and provide for transparency to enable the public to understand when AI is used in a given application and how it will impact members of the public. M-21-06 also discussed non-regulatory approaches that could be appropriate for agencies to take, such as developing and promoting sector-specific guidance, voluntary standards or frameworks, and pilot programs.

In the time since President Trump left office, the principles set forth in these foundational documents have achieved widespread acceptance in the broader ecosystem and served as the basis for industry and government agencies' programs to design, procure, and use AI. For example, the National Institute of Standards and Technology (NIST) developed the consensus-driven AI Risk Management Framework (AI RMF) in collaboration with industry, other agencies, and civil society, laying out the elements of trustworthy AI in a way that closely mirrors those from Executive Order 13960. The AI RMF lays out a voluntary risk management approach to achieve validity and reliability in AI systems, leading to greater trust and adoption and advancing innovation in the design, development, and deployment of AI in a wide range of sectors and use cases.[5]

Congress, on a bipartisan basis, also has endorsed similar principles. Last year the Bipartisan House Task Force on Artificial Intelligence gathered input from business leaders, government officials, technical experts, and other experts to develop a report that reiterated many of the same principles and identified them as key to promoting American AI innovation and leadership.[6] The Bipartisan Senate AI Working Group also developed a roadmap for AI

[4] Office of Management and Budget, M-21-06, Guidance for Regulation of Artificial Intelligence, Nov. 17, 2020.
[5] National Institute of Standards and Technology, AI Risk Management Framework (2023), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf [hereinafter NIST AI Risk Management Framework].
[6] Bipartisan House Task Force Report on Artificial Intelligence (2024), https://www.speaker.gov/wp-content/uploads/2024/12/AI-Task-Force-Report-FINAL.pdf.

innovation that similarly recognizes privacy, transparency and explainability, testing and evaluation, and safeguards against AI risks as key areas for attention.[7]

Companies have likewise voluntarily adopted similar principles as instrumental in their own AI governance commitments.[8] These shared values include privacy and security; transparency and explainability; safety; reliability; fairness; and human oversight, feedback mechanisms, and accountability.

Thus, the principles set forth in Executive Orders 13859 and 13960 and M-21-06 have become well-established, with bipartisan support and acceptance across the spectrum of stakeholders. They should form the basis of the AI Action Plan to help ensure that AI development and use by both businesses and government actors leads to effective and trustworthy AI systems, which will foster innovation, adoption, and continued American leadership.

Our comments discuss several elements that the AI Action Plan should include to help advance these goals:

- Continued work by NIST on the development of voluntary standards, evaluation and measurement methodologies, and other guidance.
- Ensuring safe, trustworthy, effective, and efficient government use of AI – from benefits and service delivery to law enforcement to national security – through the adoption of appropriate AI governance measures
- Promotion of an open AI ecosystem and investment in the National AI Research Resource
- Directing agencies to take regulatory and non-regulatory approaches to help shape private sector development and use of AI to engender trust and protect people's rights

---

[7] Bipartisan Senate AI Working Group, *Driving U.S. Innovation in Artificial Intelligence* (2024), https://www.schumer.senate.gov/imo/media/doc/Roadmap_Electronic1.32pm.pdf [hereinafter Bipartisan Senate AI Working Group Roadmap].

[8] *See e.g.*, Google, *AI Principles*, https://ai.google/responsibility/principles/; Microsoft, *Responsible AI: Principles and Approach*, https://www.microsoft.com/en-us/ai/principles-and-approach; McKinsey & Company, *Responsible AI (RAI) Principles*, https://www.mckinsey.com/capabilities/quantumblack/how-we-help-clients/generative-ai/responsible-ai-principles; Accenture, *Responsible AI: From Principles to Practice* (2021), https://www.accenture.com/content/dam/accenture/final/a-com-migration/manual/r3/pdf/pdf-149/Accenture-Responsible-AI-Final.pdf.

II. Continuing NIST's AI Work

NIST has a critical role in developing AI best practices, guidance, and research in areas such as effective testing and risk management of AI systems. NIST does not have regulatory authority, but rather develops voluntary standards or frameworks. The AI RMF, a seminal document widely used both here and abroad, describes how AI risks should be mapped, measured, and managed throughout the lifecycle of an AI application so that it is safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-protective, and fair. NIST has created companion resources to help it collaborate with other private and public sector entities in implementing the AI RMF, including through sector- and use-case-specific profiles, as well as crosswalks that describe how other frameworks map onto the AI RMF to facilitate its adoption in different contexts.

The AI Action Plan should direct NIST to continue building on the foundation it set with the AI RMF and subsequent work. NIST's voluntary technical standards are a vital tool for AI governance, and it is important that industry continue to receive guidance grounded in expertise in how AI systems technically work, how they can cause or contribute to risks, and how those risks can be mitigated. NIST provides unique technical expertise in the science around measurement and the complexities of AI systems, making it well-positioned to lead the development of robust standards that address the full spectrum of AI risks.[9]

These standards should continue to be developed through a multi-stakeholder process where NIST's expertise is combined with expertise from civil society, academia, government, and industry, and where input of all of these stakeholders is incorporated to balance the interests of protecting people's rights, ensuring safety and effectiveness, and advancing innovation.[10] Civil society perspectives and expertise are too often crowded out of multi-stakeholder processes, so the AI Action Plan should direct NIST to ensure this process meaningfully integrates these perspectives, which are necessary to spot and address issues directly affecting communities.[11] A truly multi-stakeholder process will also support a greater understanding of the relationship between an AI system's design and capabilities and its behavior and performance.

---

[9] Americans for Responsible Innovation et al, Coalition Letter to Senate Appropriations Committee, Apr. 23, 2024, https://responsibleinnovation.org/wp-content/uploads/2024/04/Coalition-Letter-to-Congress-In-Support-of-NIST-AI-Funding-SENATE.pdf.

[10] Miranda Bogen, *Ensuring NIST's AI Safety Institute Consortium Lives Up to its Potential*, Center for Democracy & Technology (Mar. 6, 2024), https://cdt.org/insights/ensuring-nists-ai-safety-institute-consortium-lives-up-to-its-potential/.

[11] *Id.*

The standards-development process should center not only the prospective security risks arising from capabilities related to chemical, biological, and radiological weapons and dual-use foundation models, but also the current, ongoing risks of AI such as privacy harms, ineffectiveness of the system, lack of fitness for purpose, and discrimination.[12] NIST's standards should also include a multifaceted approach for holistically and accurately measuring different qualities of an AI system, such as safety, efficacy, or fairness, and provide guidance on determining the validity and reliability of the measurements used.[13] There could be many ways to measure these qualities, or constructs, of an AI system, but different methods of measuring any given construct will lend themselves to different interpretations of the system's performance with respect to that construct.[14] Thus, NIST's help in such measurements is important, especially to help small businesses and other adopters of AI systems who benefit from external expertise on how to assess different AI tools.

III. Ensuring the Use of Trustworthy AI in the Federal Government

The use of AI by federal agencies holds promise as a tool for improving service delivery and enhancing operational efficiency. Indeed, AI systems can play a critical role in improving customer service, reducing administrative burden, preventing fraud and waste, and increasing the efficiency and effectiveness of benefits administration. Agencies are exploring these uses already: through AI use case inventories, federal agencies in 2024 publicly reported 2,133 AI use cases, a 200% increase from the previous year.[15] However, these potential benefits of AI in government also come with significant risks—ranging from wasted tax dollars on "snake oil" tools that are ineffective and expose an agency to legal liability, to the misuse of tools for sweeping program reforms rather than a nuanced analysis with appropriate process, expert input, transparency and review, to significant harms to people's lives and freedoms when their benefits are incorrectly denied or they are wrongly targeted by law enforcement.[16] Realizing the

---

[12] *Id.*

[13] Amy Winecoff & Miranda Bogen, *Trustworthy AI Needs Trustworthy Measurements*, Center for Democracy & Technology (Mar. 6, 2024), https://cdt.org/insights/trustworthy-ai-needs-trustworthy-measurements/.

[14] *Id.*

[15] Office of Management and Budget, 2024 Federal Agency AI Use Case Inventory (Jan. 23, 2025), https://github.com/ombegov/2024-Federal-AI-Use-Case-Inventory.

[16] Mary Keierleber, *Whistleblower: L.A. Schools' Chatbot Misused Student data as Tech Co. Crumbled*, The 74 (Jul. 1, 2024), https://www.the74million.org/article/whistleblower-l-a-schools-chatbot-misused-student-data-as-tech-co-crumbled/; Colin Lecher, *What Happens When an Algorithm Cuts Your Health Care*, The Verge (Mar 21, 2018), https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy; Colin Lecher, *NYC's AI Chatbot Tells Businesses to Break the Law*, The Markup (Mar. 29, 2024), https://themarkup.org/news/2024/03/29/nycs-ai-chatbot-tells-businesses-to-break-the-law.

potential of AI in government modernization requires guardrails and governance structures to mitigate such risks and advance safety, trustworthiness, and efficiency.

The first Trump Administration laid the foundation for promoting the transparent and effective use of AI in government with Executive Order 13960 on Promoting the Use of Trustworthy AI in the Federal Government. EO 13960 underscores that "agencies must [...] design, develop, acquire, and use AI in a manner that fosters public trust and confidence while protecting privacy, civil rights, civil liberties, and American values."[17] To achieve these goals, EO 13960 directed agencies to adhere to nine principles throughout the AI lifecycle: privacy, civil rights, and civil liberties; risk management; accuracy, reliability, and effectiveness; safety; transparency, understandability, and documentation; performance monitoring and corrective measures; and accountability for implementing safeguards.[18] Moreover, EO 13960 required agencies to publish AI use cases inventories publicly detailing their current and planned uses of AI, a requirement later codified into law with the enactment of the bipartisan *Advancing American AI Act*.[19]

Together, these measures establish a strong framework for federal agencies to deploy and govern AI systems safely. Moreover, the Bipartisan House Task Force on AI recently recognized the importance of these practices, recommending that federal agencies adopt AI governance standards and safeguards such as public notice and appeal.[20] Finally, many states across the country have since implemented similar requirements on AI systems used by state agencies.[21]

Drawing on widely accepted best practices for public sector AI governance, the AI Action Plan should center the following six best practices to guide federal agencies' use of AI for purposes such as benefits administration and law enforcement. The Administration should also ensure compliance with these principles in its existing use of AI, most significantly in DOGE efforts which appear to be leveraging AI without transparency or these other necessary guardrails in place.

---

[17] Exec. Order 13960.
[18] *Id.* at Sec. 3
[19] *Id.* at Sec. 5; Advancing American AI Act, Pub. L. 117–263, div. G, title LXXII, subtitle B, Dec. 23, 2022, 136 Stat. 3668.
[20] Bipartisan House Task Force Report on Artificial Intelligence, 1-22, *supra* note 6.
[21] Maddy Dwyer, *State Government Use of AI: The Opportunities of Executive Action in 2025*, Center for Democracy & Technology (Jan. 10, 2025), https://cdt.org/insights/state-government-use-of-ai-the-opportunities-of-executive-action-in-2025/; Quinn Anex-Ries, *Regulating Public Sector AI: Emerging Trends in State Legislation*, Center for Democracy & Technology (Jan. 10, 2025), https://cdt.org/insights/regulating-public-sector-ai-emerging-trends-in-state-legislation/.

A. *Risk Assessment and Mitigation*

Identifying and assessing the potential risks associated with specific AI use cases is a critical tool for agencies to proactively mitigate harms.[22] As federal agencies adopt a variety of AI use cases – ranging from tools to copy-edit reports to assisting with benefits decisions – determining the risks associated with each use case will aid agencies in tailoring their risk management practices and appropriately dedicating limited resources to the highest risk uses. In particular, agencies should identify high-risk use cases, including those that have a significant impact on individuals' privacy, safety, liberty, or legal rights, and adopt heightened risk mitigation measures for such systems.

B. *Testing and Evaluation*

Pre- and post-deployment testing enables public agencies to ensure that AI systems are fit-for-purpose, work effectively, and behave as expected in real-world environments prior to implementation, and to identify and address any errors or harms that may occur as systems are used.[23] A core component of such testing should include evaluating systems for potential biases based on protected characteristics to identify and prevent potential discrimination.[24] Avoiding such biases is a necessary predicate for trust and basic effectiveness. A biased AI system used in a law enforcement context, for example, could cause investigators to waste time and resources chasing down an incorrect lead or suspect and, in the worst case, lead to the wrongful deprivation of an individual's liberty.

C. *Centralized Governance and Oversight*

The AI Action Plan should maintain and expand existing agency and interagency AI governance structures. Federal agency chief artificial intelligence officers (CAIO), the interagency CAIO council, and agency AI governance boards are important mechanisms for ensuring that agencies have dedicated leaders focused on AI adoption and oversight, coordinating AI oversight and monitoring, promoting cross-agency collaboration, and enabling public-private partnerships.[25]

---

[22] NIST AI Risk Management Framework, *supra* note 5.
[23] Merlin Stein & Connor Dunlop, *Safe Beyond Sale: Post-Deployment Monitoring of AI*, Ada Lovelace Institute (Jun. 28, 2024), https://www.adalovelaceinstitute.org/blog/post-deployment-monitoring-of-ai/.
[24] National Institute of Standards and Technology, *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence* (2022), https://www.nist.gov/publications/towards-standard-identifying-and-managing-bias-artificial-intelligence.
[25] Bethany Abbate, *Chief AI Officers Must be Preserved in the Trump Administration*, FedScoop (Feb. 18 2025), https://fedscoop.com/chief-ai-officers-trump-administration/.

D. Privacy and Security

The adoption of AI systems in government can exacerbate existing privacy and security risks and can introduce new privacy and security harms.[26] These risks include the potential leakage of sensitive information and increasing the attack surface within government information systems by adding new, and often hard to assess, data flows. Given these risks, federal agencies should take steps to align all AI uses with existing privacy and cybersecurity requirements – such as requirements for agencies to conduct privacy impact assessments – and to proactively guard against novel privacy and security risks introduced by AI.

E. Public Engagement

Public engagement is a powerful tool for federal agencies to increase public trust and confidence in AI systems. By soliciting direct feedback from the public, federal agencies can ensure that the needs and interests of the American people are incorporated throughout the design, development, and use of AI systems.[27] Agencies should use formal and informal avenues to solicit such feedback including public hearings, notice and comment opportunities, and direct consultation with affected community groups.

F. Transparency and Disclosure

Public trust in the federal government's use and development of AI depends on transparency about when, why, and how federal agencies are using AI systems. The AI Action Plan can achieve this by building on agencies' existing use case inventories – a key channel for the public to learn information about how agencies are using and governing AI systems and for industry to understand AI needs within the public sector – and by requiring agencies to provide public notice and appeal when individuals are affected by AI systems in high-risk settings. In the specific context of law enforcement, the use of an AI system to generate evidence to be used in a criminal case, or leads in a criminal investigation, should be disclosed to a person who is accused of a crime based in part on such lead or evidence. This should be affirmatively embraced by the Department of Justice as an element of the due process rights of the accused.

---

[26] Center for Democracy & Technology, Comment to OMB on Federal Agencies' Use of Commercially Available Information (Dec. 16, 2024), https://cdt.org/insights/cdt-submits-comments-to-omb-on-federal-agencies-use-of-commercially-available-information/.

[27] Tina Park, *Stakeholder Engagement for Responsible AI*, Partnership on AI (Sept. 17, 2024), https://partnershiponai.org/stakeholder-engagement-for-responsible-ai-introducing-pais-guidelines-for-participatory-and-inclusive-ai/.

EO 13960 rightly established transparency as a key principle for the use of AI in government. Therefore, we note with deep alarm that the manner in which DOGE reportedly is using AI tools to make high-risk decisions across the federal government, including to determine layoffs and spending cuts, does not embody this principle.[28] The public and interested stakeholders lack basic information about privacy and security measures governing DOGE's access to the most sensitive data about Americans, and whether and how DOGE staff and agency leaders are using AI to drive high-stakes decisions. Therefore, we strongly urge that, through the AI Action Plan, DOGE website, AI inventories, and other means, the Administration clarify and proactively communicate how DOGE is using AI, and the relevant guardrails in place, in order to provide the transparency the first Trump Administration recognized is necessary to earn public trust.

IV. Aligning the Use of AI for National Security Purposes with Civil Liberties & the Constitution

Special concerns apply when using AI in national security settings. In this context, many use cases are high risk because life or liberty are at stake in the decision-making process involving AI, and some of those use cases cannot be made public on account of secrecy needs. But that secrecy can also shield the abuse and misuse of AI systems. This makes transparency, disclosure, and effective governance and oversight of AI systems in the national security context particularly important.

While classification may limit the extent of transparency in a national security setting, the AI Action Plan should build on transparency principles and practices developed both inside and outside national security agencies in the past decade, including by mandating declassification review of key documentation such as AI use-case inventories, impact assessments, and controlling legal memoranda about the use of AI, as well as regular and meaningful reporting to relevant Congressional committees and offices.

Because the Intelligence Community consists of 18 elements,[29] some level of centralized governance of the use of AI by each is essential. Each should already have a Chief AI Officer (CAIO) who can coordinate with other CAIOs for internal oversight purposes. Agencies should

[28] Hannah Natanson et al, *Elon Musk's DOGE is Feeding Sensitive Federal Data Into AI to Target Cuts*, Washington Post (Feb. 6, 2025), https://www.washingtonpost.com/nation/2025/02/06/elon-musk-doge-ai-department-education/; Makena Kelly, *DOGE is Working on Software that Automates the Firing of Government Workers*, Wired (Feb. 25, 2025), https://www.wired.com/story/doge-autorif-mass-firing-government-workers/.

[29] Office of the Director of National Intelligence, *Members of the IC,* https://www.dni.gov/index.php/what-we-do/members-of-the-ic (Last checked March 13, 2025).

clearly assign decision-making and internal oversight responsibilities, including requiring high-level officials' approval for procuring systems and for use cases that present particularly high risks, and mandating appropriate consultation for legal, civil rights, and privacy officials.

The AI Action Plan should recognize that independent external oversight is also critically important to promote safe, trustworthy, and efficient use of AI in the national security/intelligence arena. Many such uses will be classified and exposure of them could put national security at risk. At the same time, because the risk of abuse and misuse is high when such functions are kept secret, an oversight mechanism with expertise, independence and power to access relevant information (even if classified) should be established in the Executive Branch. CDT has recommended that Congress establish such a body,[30] and the AI Action Plan should support such an approach.

V. Advancing Competitiveness by Supporting Openness in the AI Ecosystem and Investing in the National AI Research Resource

A crucial element of the past few years of AI development has been the continued flourishing of open models at the AI frontier. Open models — models whose weights can be freely downloaded over the Internet — have several important benefits. They accelerate innovation and promote competition, by making it possible for startups and academics to innovate on top of cutting-edge AI tools. Open models also promote rapid AI diffusion and widespread AI adoption by making it possible for businesses to use AI systems more reliably and securely. Moreover, the availability of open models helps mitigate the pernicious and opaque centralization of power within a small number of large, unaccountable AI companies.

The AI Action Plan should set a course that ensures America remains a home for open model development. While the benefits of open models are increasingly widely recognized, some observers still advocate for restrictions on open model development (e.g., via restricting the export of open models, which would *de facto* amount to a ban on releasing open models), in an attempt to address perceived risks to public safety or national security. Such restrictions require close examination, and the evidence does not warrant such restrictions at this time. Restricting open model development now would *not* improve public safety or further national security — rather, it would sacrifice the considerable benefits associated with open models and cede

[30] Jake Laperruque, Testimony before the House Committee on Homeland Security, *Advancing Innovation (AI): Harnessing Artificial Intelligence To Defend and Secure the Homeland,* https://cdt.org/wp-content/uploads/2024/05/Jake-Laperruque-5-22-24-AI-Hearing-Written-Testimony.pdf (May 22, 2024), pp. 8-10.

leadership in the open model ecosystem to foreign adversaries. Rather than restricting open model development, the AI Action Plan should ensure that open models retain their central position in the American AI ecosystem, while promoting the development of voluntary standards to enable their safe and responsible development and use.

Over the past year, numerous observers have come to agree that policymakers should, rather than hastily suppressing open model development, focus their efforts on vigilantly monitoring developments in open models' capabilities and creating robust standards for identifying potential future risks. The Bipartisan House Task Force on Artificial Intelligence, for instance, found that "[o]pen AI models encourage innovation and competition," and in fact encouraged Congress to "bolster openness in AI development and use while continuing to ensure that models have appropriate safeguards."[31] Similarly, the National Telecommunications and Information Administration (NTIA), in its Report on Dual-Use Foundation Models with Widely Available Model Weights, found that open models "introduce a wide spectrum of benefits," such as "diversify[ing] and expand[ing] the array of actors [...] that participate in AI research and development," and that as such, "current evidence is not sufficient to definitively determine either that restrictions on [dual-use] open-weight models are warranted, or that restrictions will never be appropriate in the future."[32]

The most important benefit of open models is their potential to accelerate innovation in AI. The current flourishing of generative AI and foundation model technology would not have been possible without open research. For example, the technology that underlies today's LLMs — the neural network architecture known as the Transformer — has its origin in open research, with openly-published code and data.[33] Without this open research, as well as open-source ML development frameworks like PyTorch and TensorFlow, today's closed models would not exist.[34] Furthermore, open models enable a variety of AI research not enabled by closed foundation

[31] Bipartisan House Task Force Report on Artificial Intelligence, *supra* note 6, pp. 160–161.
[32] National Telecommunications and Information Administration, *Dual-Use Foundation Models with Widely Available Model Weights*, at 2 (2024), https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.
[33] Ashish Vaswani et al., *Attention Is All You Need*, *31st Conference on Neural Information Processing Systems (NIPS 2017)*, vol. 30, 2017, https://arxiv.org/abs/1706.03762.
[34] Max Langenkamp and Daniel N. Yue, *How Open Source Machine Learning Software Shapes AI*, *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22)*, 2022, https://dl.acm.org/doi/10.1145/3514094.3534167.

models,[35] including research around AI interpretability methods[36] and the public development of robust watermarking techniques.[37]

A second, related benefit of open models is their ability to facilitate the rapid adoption of AI by businesses. The history of open-source software should lead us to expect open models to enable faster, cheaper diffusion of foundation model technology to startups and other businesses large and small, as well as other developer and user communities around the world. So far, that is exactly what is occurring.[38] As the UK's Competition and Markets Authority has noted, "[a]t present a mix of open and closed-source foundation models are available and competing. This is allowing a range of firms to invest in and develop foundation models and as a result we are already seeing deployment of these foundation models in a growing range of applications across the economy."[39] Large companies such as Dell and Wells Fargo are starting to use open models to help with internal knowledge management and internal software coding, with Dell's SVP for AI Strategy noting "[a] lot of customer[s] are asking themselves: 'Wait a second, why am I paying for [a] super large model that knows very little about my business? Couldn't I just use one of these open-source models . . . ?'"[40]

Open models also mitigate the concentration of power associated with the closed AI model ecosystem. The market structure for closed foundation models tends toward concentration, including vertical integration, in large part due to the high costs of compute infrastructure for training.[41] Open models play a crucial role in mitigating the concentration of the AI ecosystem and promoting innovation and competition.[42] Moreover, a foundation model market dominated

---

[35] Sayash Kapoor et al., *On the Societal Impact of Open Foundation Models*, arXiv, (2024), https://arxiv.org/abs/2403.07918.
[36] Kevin Clark et al., *What Does BERT Look at? An Analysis of BERT's Attention*, arXiv, (2019), https://arxiv.org/abs/1906.04341.
[37] John Kirchenbauer et al., *A Watermark for Large Language Models*, arXiv, (2023), https://arxiv.org/abs/2301.10226.
[38] Competitions and Markets Authority, *AI Foundation Models: Initial Review*, GOV.UK, (2024), https://www.gov.uk/cma-cases/ai-foundation-models-initial-review.
[39] *Id.*
[40] Matt Marshall, *How Enterprises Are Using Open Source LLMs: 16 Examples*, VentureBeat (Feb. 2, 2024), https://venturebeat.com/ai/how-enterprises-are-using-open-source-llms-16-examples.
[41] Jai Vipra and Anton Korinek, *Market Concentration Implications of Foundation Models: The Invisible Hand of ChatGPT*, Center on Regulation and Markets at Brookings (Sept. 7, 2023), https://www.brookings.edu/articles/market-concentration-implications-of-foundation-models-the-invisible-hand-of-chatgpt; David Gray Widder, Sarah Myers West, and Meredith Whittaker, *Open (for Business): Big Tech, Concentrated Power, and the Political Economy of Open AI*, Social Science Research Network (2023), https://doi.org/10.2139/ssrn.4543807.
[42] Jake Denton, *The U.S. Shouldn't Go The Way Of Europe On AI*, Heritage Foundation (May 8, 2024), https://www.heritage.org/big-tech/commentary/the-us-shouldnt-go-the-way-europe-ai.

by a handful of closed systems carries other risks. One notable risk is that of an emergent "monoculture": a situation where many different decision-makers and service providers rely on the same closed systems, and where as a result, a handful of companies decide what knowledge and expression is allowed through this powerful new layer of information technology, raising the specter of undue power over politics and culture. The democratization of AI through open models creates a healthier ecosystem where power is distributed across many entities rather than concentrated in a few corporate gatekeepers, which ultimately benefits society through greater innovation, more diverse perspectives, and increased accountability.

Given how widely used open models are, it is crucial that the U.S. retains its lead in open model development and that American open models remain the base on which companies build their products. Recent developments, such as the release of the R1 model by the Chinese research lab DeepSeek, suggest that if the U.S. were to implement policies that stifle domestic open model development, open models developed by other countries would likely supplant American models as the basis for much of the AI economy.[43] Such a shift would be hazardous in multiple ways: not only do DeepSeek's models (along with some other Chinese-developed models) regurgitate government talking points, but they also may be built with secret "backdoors" that undermine the security of any system built atop them.[44] As a result, restricting open model development in the U.S. at this time may be actively detrimental to public safety and national security. Moreover, "for many countries open source AI represents the only opportunity to engage with the technology due to its prohibitively high development and training costs."[45] If America remains at the frontier of open model development, its models will likely become the basis for AI-based technologies in much of the world. But if the U.S. stifles domestic open model development, the basis for those technologies would likely be models developed by authoritarian governments.

Along similar lines, the Action Plan should recognize that at this time export restrictions on open model weights or other components would not further national security — rather, they would cede American leadership in a crucial component of the AI ecosystem. Insofar as the Administration may deem restrictions on the export of certain AI models necessary, such restrictions should exclude open models from their scope.

---

[43] Ben Brooks, *If China Shares AI, the US Can't Afford to Lock It Out*, The Hill (Feb. 6, 2025), https://thehill.com/opinion/technology/5123855-open-source-ai-deepseek/.
[44] Keegan McBride, *Open Source AI: The Overlooked National Security Imperative*, Just Security (June 6, 2024), https://www.justsecurity.org/96422/open-source-ai-the-overlooked-national-security-imperative/.
[45] Keegan McBride and Dean Ball, *The United States Must Win The Global Open Source Race*, Just Security (Nov. 7, 2024), https://www.justsecurity.org/104676/american-ai-leadership-requires-support-open-source/.

At the same time, open model development is not risk-free, especially as frontier AI models continue to become more capable. The government should begin creating the mechanisms necessary to better assess and monitor open models' capabilities to determine whether restrictions might be required in the future.[46] That includes consideration of what forms of information sharing and transparency from developers of open models may be necessary. The government also needs to invest in and support the ecosystem for testing of open models. In part, that includes providing NIST (and non-government entities) the resources to help establish clearer testing benchmarks for a range of foundation model risks and investment in research to develop technically feasible and effective testing and risk evaluation norms. The Action Plan should also call for the development of best practices and norms around the responsible release and use of open models. For example, some developers already publish materials providing transparency about their models and tools helpful to a deployer seeking to responsibly use the models.

The AI Action Plan should also recognize the critical role that the National AI Research Resource (NAIRR) should play in ensuring America's continued leadership in AI. The NAIRR would represent a vital investment in our nation's AI research infrastructure that can democratize access to the computational resources, data, and tools needed for cutting-edge AI development. Both the Bipartisan House Task Force on AI and the Bipartisan Senate AI Working Group have explicitly recognized the importance of NAIRR in strengthening America's AI research ecosystem and enabling participation from a wide range of researchers, startups, and businesses that might otherwise be excluded from frontier AI research due to resource constraints.[47] By asserting the implementation of the NAIRR as a priority within the AI Action Plan, the Administration can create a powerful platform that amplifies American innovation while providing the resources needed to address emerging AI challenges in a thoughtful, evidence-based manner.

VI. Shaping Private Sector Use of AI

Widespread adoption of AI requires businesses and individuals to trust that AI systems are effective, fit for purpose, and safe and that they will not undermine people's rights. As companies develop and incorporate AI systems, they need to adopt practical governance

[46] Dean Ball, *Open-Source AI And The Future*, Hyperdimensional (Jan. 23, 2025), https://www.hyperdimensional.co/p/invitations.

[47] Bipartisan House Task Force Report on Artificial Intelligence, *supra* note 6; Bipartisan Senate AI Working Group Roadmap, *supra* note 7, at 5.

measures to help them meet those goals, as well as comply with long-standing legal obligations. Agencies have the sector-specific expertise that can assist companies in this endeavor and protect against AI abuses that violate the laws they are dedicated to enforcing.

The AI Action Plan should direct agencies to take regulatory and non-regulatory approaches to, in the words of OMB Memorandum M-21-06 "ensure consistency and predictability of AI-related policies that advance American innovation and adoption of AI, while appropriately protecting privacy, civil liberties, national security, and American values and allowing sector- and application- specific approaches." For example, as discussed above, NIST should continue its work of developing voluntary standards, evaluation and measurement methodologies, and other frameworks. The FTC should use its Section 5 enforcement authority against providers of AI who develop and sell systems that are not fit for purpose or otherwise deceive customers about their capabilities. Sector-specific agencies can provide guidance about appropriate practices in their sectors, adapt their regulations as needed to reflect the use of AI, and bring enforcement actions when companies use AI in ways that result in violations of their regulatory obligations. The AI Action Plan should make clear that agencies should carry out these and other responsibilities, much as OMB previously did in M-21-06.

The AI Action Plan should also establish mechanisms for ongoing interagency coordination, such as through the Chief AI Officers Council. These mechanisms can help agencies effectively exercise their individual regulatory and enforcement authorities by providing a forum to exchange best practices and learnings, facilitate consistency, and otherwise coordinate, as well as provide the White House with an overview of how agencies are approaching AI to help inform its own policy development.

VII. Conclusion

The AI Action Plan should recognize that American leadership in AI requires the development and deployment of AI systems that are effective and fit-for-purpose, and that protect Americans' rights and safety. Only systems meeting those criteria will engender the trust needed to lead to adoption and use of AI, which in turn will fuel further investment and innovation. The Action Plan should include steps needed to ensure that both the government and the private sector pursue AI development that protects and advances fundamental American values.