House Committee on Oversight & Accountability

Testimony for a Hearing on White House Overreach on AI

Dr. Nicol Turner Lee, Senior Fellow & Director

Center for Technology Innovation, Governance Studies

The Brookings Institution

March 20, 2024

Chairwoman Mace, Ranking Member Connolly, and distinguished members of the Subcommittee on Cybersecurity, Information Technology, and Government Innovation: thank you for the invitation to testify on President Biden's Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. I am Nicol Turner Lee, Senior Fellow, Governance Studies, and Director of the Center for Technology Innovation at the Brookings Institution. With a history of over 100 years, Brookings is committed to evidence-based, nonpartisan research in a range of focus areas. My research expertise encompasses data collection and analysis around regulatory and legislative policies that govern telecommunications and high-tech industries, along with the impacts of digital exclusion, artificial intelligence, and machine learning algorithms on vulnerable populations. My forthcoming book, *Digitally Invisible: How the Internet is Creating the New Underclass*, will be published by Brookings Press later this summer.

To understand the White House Executive Order (EO), its objectives, and its impacts, it is important to understand the governmental context in which it was released and developed. To this end, in addition to summarizing the EO, I also briefly summarize a few crucial government actions preceding and surrounding it: the Blueprint for an AI Bill of

Rights, released in October 2022; the NIST AI Risk Management Framework 1.0, released in January 2023; the securing of voluntary commitments by the White House from top AI developers in July 2023; and Office of Management and Budget (OMB) guidance memo, released shortly after the EO in November 2023. That is, our conversation today must reflect this 'whole of government' approach toward achieving national guidance as AI becomes both an asset and concern for our national security interests. I also share in my testimony that Congress must quickly act on many of the AI proposals and activities under discussion to ensure that we maintain our status as leaders in the global economy.

The Foundational Tenets of the White House EO

The National Blueprint for an AI Bill of Rights

In October 2022, the White House Office of Science and Technology Policy published a Blueprint for an AI Bill of Rights[1], which shared a nonbinding roadmap for the responsible use of artificial intelligence. The comprehensive document, or the Blueprint, identified five core principles to guide and govern the effective development and implementation of AI systems: Safe and Effective Systems, Algorithmic Discrimination Protections, Data Privacy, Notice and Explanation, and Human Alternatives, Consideration, and Fallback. Suggestions within this framework include pre-deployment risk and discrimination assessments, required consent with respect to the "collection, use, access, transfer, and deletion" of user data, issuance of plain-language notice and explanation of automated decision-making, and access to human review of automated

[1] White House Office of Science and Technology Policy, "Blueprint for an AI Bill of Rights," October 4, 2022 https://www.whitehouse.gov/ostp/ai-bill-of-rights/.

decisions in some cases. The intent of the Blueprint was to outline the rights of consumers in ways that provided some agency over the autonomous tools and decisions being made on their behalf.

Following the release of the Blueprint, at least five federal agencies adopted guidelines for their own responsible use of automated systems, a few have established their own centers or offices to implement these guidelines, and at least a dozen agencies have issued some sort of binding guidance for the use of automated systems in the industries under their jurisdiction, such as the Federal Trade Commission (FTC) and the Federal Drug Administration (FDA). However, the detail and scope of federal agencies' full adherence to these activities still varies in terms of timeline and deliverables. According to the recent update on both voluntary and executive-branch government activities, these principles have helped to frame the focus on risk management of AI models, which has been a common concern among industry actors as well.

The National Institute of Standards & Technology (NIST)

In January 2023, the National Institute of Standards & Technology (NIST) issued Version 1.0 of its Artificial Intelligence Risk Management Framework (AI RMF)[2]– a multi-tool for organizations to design and manage trustworthy and responsible artificial intelligence (AI) that is meant to be "voluntary, rights-preserving, non-sector-specific, [and] use-case agnostic." The AI RMF provides two lenses through which to consider questions around balancing risks and benefits. First, it provides a conceptual roadmap for identifying risk in

[2] NIST, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," January 2023 https://doi.org/10.6028/NIST.AI.100-1.

the AI context – outlining general types and sources of risk relating to AI and enumerating seven key characteristics of trustworthy AI (safe, secure and resilient, explainable and interpretable, privacy-enhanced, fair—with harmful bias managed, accountable and transparent, valid and reliable).[3] Second, it offers a set of organizational processes and activities to assess and manage risk linking AI's socio-technical dimensions to stages in the lifecycle of an AI system and to the actors involved. Key steps for these processes and activities are "test, evaluation, verification, and validation (TEVV)." The processes and activities are broken down into core functions—to govern, map, measure, and manage.[4]

With the release of the AI RMF, NIST is also launching a companion "playbook," a tool that will provide additional suggestions for actions, references, and documentation for the "govern, map, measure, and manage" functions. As the title "Version 1.0" implies, the document released January 26 is not meant to be NIST's last word on AI risk management. The agency expects to conduct a full, formal review by 2028, which could produce a Version 2.0. In recent months, NIST has gone further with the launch of the AI Safety Institute Consortium (AISIC), which will bring together stakeholders across industry, academia, and government to jointly develop and diffuse standards, best practices, benchmarks, and more. The Consortium supports the broader initiatives of the AI Safety Institute, also housed at NIST.[5]

Voluntary commitments from the private sector

[3] Ibid, p. 12.
[4] Ibid, p. 20.
[5] NIST, "U.S. Artificial Intelligence Safety Institute," February 7, 2024 https://www.nist.gov/artificial-intelligence/artificial-intelligence-safety-institute.

In July 2023, the White House secured voluntary commitments from seven leading US AI companies – Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI – to ensure safety, security, and trust with advanced AI systems. These include agreements about internal and external security testing on crucial risks such as bio- and cybersecurity as well as broader societal effects, the protection of unreleased model weights, and public reporting of system capabilities, limitations, and guidelines for responsible use. In September 2023, eight additional companies – including IBM, Nvidia, and Palantir – were convened at the White House to agree to these same terms. Such proactive participation of companies suggest that the federal government is not necessarily acting alone on this issue of responsible AI governance, and is equally interested in ways to balance the needs of the market with the further design and deployment of autonomous tools. One of the commitments designed by the White House was to develop "robust technical mechanisms to ensure that users know when content is AI generated, such as a watermarking system."[6] Industry leaders have continued to focus on digital watermarking —in recent months, Google, Adobe, Intel, and Microsoft have joined a coalition dedicated to developing watermarking technology—but it is important to note that efforts to identify digital provenance will have challenges and watermarking is not a foolproof strategy.[7]

The White House Executive Order

[6] The White House, "FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI," July 21, 2023 https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/

[7] Makena Kelly, "Watermarks aren't the silver bullet for AI misinformation," The Verge, October 31, 2023 https://www.theverge.com/2023/10/31/23940626/artificial-intelligence-ai-digital-watermarks-biden-executive-order

Drawing on these actions, the White House Executive Order, or White House EO, was enacted on October 30, 2023, introducing reporting requirements and myriad directives for federal agencies to encourage responsible innovation while protecting civil rights. In particular, the White House EO has eight overarching goals:

- The EO implements *New Standards for AI Safety and Security* by requiring developers to share safety test results with the government.
- The EO *Protects Americans' Privacy* by evaluating how federal agencies collect and use public data.
- The EO *Advances Equity and Civil Rights* by providing guidance to landlords and federal benefits programs to ensure that AI is not used in a discriminatory fashion.
- The EO *Stands Up for Consumers, Patients, and Students* by advancing the responsible use of AI in drug development.
- The EO *Supports Workers* by directing a report on AI's potential labor-market impacts.
- The EO *Promotes Innovation and Competition* by creating a National AI Research Resource to catalyze AI development.
- The EO *Advances American Leadership Abroad* by directing the State and Commerce Departments to develop robust international frameworks for AI.
- The EO *Ensures Responsible and Effective Government Use of AI* by developing clear standards for federal agencies' procurement and deployment of AI technology.

In addition to the specific focus areas, on a more cumulative level, the White House EO has resulted in other directives taken in pursuit of these goals including to develop standards for red-team testing, strengthen agency "privacy guidance to account for AI risks," explore best practices for AI in criminal justice, support educators in using AI tools, study AI's potential labor-market impacts, expedite the visa process for highly skilled AI workers, and expand involvement in international collaborations on AI. Overall, the EO presents a major step forward in this 'whole of government' approach to AI governance in the United States and should inform the current dialogues around appropriate and sufficient congressional legislation that codifies our global leadership around the technical and societal concerns of AI, while emboldening trust among users.

In tandem with the White House EO have been developments in how various federal agencies leverage AI, which is part of the draft guidance submitted by the Office of Management and Budget (OMB), which released a memorandum to guide the implementation with three goals in mind – having the appropriate staffing resources in place, coordinating internal and cross-agency activities, and ensuring a risk management approach to the federal government's engagement with AI, including in procurements. For example, the OMB memo mandates that each agency appoint a Chief AI Officer to advise agency leadership, orchestrate agency AI activity, and oversee risk management. To advance responsible AI innovation, it requires the removal of unnecessary barriers to responsible adoption of AI within agencies. To manage risks from agency use of AI, the memo requires the implementation of safeguards to protect the safety and rights of the public, including impact assessments and independent evaluations, system monitoring

during deployment, and notifying harmed individuals about paths for redress. This memorandum filled in key details for the execution of the White House EO - many of which have since been implemented according to the order and memorandum's timelines.

Current EO Progress and Next Steps

Upon the enactment of the EO, federal agencies were given a 90-day period in which to implement certain critical actions addressing AI's threats to security and safety and emphasizing the United States' role as a lead innovator of AI. In late January 2024, the White House confirmed that each of the 90-day benchmarks had been met, including disclosure requirements for developers of powerful AI systems, assessments on the risk AI poses to critical infrastructure, and steps to inhibit the abilities of foreign actors to develop harmful AI. In addressing AI's threats to safety and security under the completed actions, the Department of Commerce also has proposed rules that would require U.S. cloud companies to report when they provide computing power for foreign AI training. Further, a set of nine agencies have submitted risk assessments for their uses of AI systems to the Department of Homeland Security to "serve as the basis for continued federal action" for the safe integration of AI into society. Among these agencies are the Department of Defense, the Department of Transportation, the Department of Treasury, and Department of Health and Human Services.

In an effort to secure and attract AI talent to the United States, the National Science Foundation (NSF) launched the National AI Research Resource (NAIRR) to foster innovation and ensure equitable access to AI research resources, further democratizing access and

education to AI tools with both government data and private sector support.[8] Increased hiring for initiatives like NAIRR and other federal activities involving AI have been another key part of the executive order's early accomplishments. With responsible AI innovation as a top priority of the Executive Order, funding for new Regional Innovation Engines has been secured to support breakthrough innovations in AI.[9]

As the world's leading innovator in AI, the United States has a responsibility to lead in preparing for its risks to ensure that the development of these revolutionary technologies benefits its citizens and the world. As a dual-use technology, general-purpose AI that is developed with positive intentions can be turned to negative uses, such as the large-scale generation and dissemination of misinformation and deepfakes – activities that are quickly upending important democratic institutions like voting and elections infrastructure. That is, the capabilities of frontier models, like generative AI, surge forward at a rapid pace, and if the government does not act to get out ahead of this technology, it will struggle to catch up later when harms become more serious. To be clear, there is nothing in the White House EO that has become a more proscriptive norm for AI guidance. Rather, it should be perceived as our nation's exercise to ensure that future actions by Congress are directed at the concerns of their core constituents when it comes to AI's transparency and use, is done so in a way that balances innovation and regulation, and puts the nation ahead of our

[8] National Science Foundation, "National Artificial Intelligence Research Resource Pilot," January 24, 2024 https://new.nsf.gov/focus-areas/artificial-intelligence/nairr

[9] Mark Muro, "With its winners announced, the Regional Innovation Engines program moves to expand place-based R&D," January 31, 2024 https://www.brookings.edu/articles/with-its-winners-announced-the-regional-innovation-engines-program-moves-to-expand-place-based-rd/

global competitors when it comes to the norms, values, and protections that come with more ethical and responsible, autonomous systems.

States and AI activities

Despite the flurry of congressional activities including legislative drafts, forums, and other dialogues, a patchwork of state AI laws has begun to materialize. From January to October 2023, 24 states introduced a total of 190 AI-related bills, which is a 440% increase compared to 2022 and more than the previous two years combined.[10] These bills covered diverse topics ranging from algorithmic bias and personal privacy to state AI Task Forces, Offices, and advisory councils and licensing requirements for advanced AI systems. Naturally, a particular focus was on newly popular generative AI technology, with bills addressing election interference, deepfakes, notices of AI interaction in state use and labeling of AI-generated content in advertisements, and broader regulation of large generative AI models.

Foreign government leadership and manipulation

Why the White House EO is not overreach when it comes to AI governance is also largely due to potential misuse by foreign governments, who have taken leadership roles already, and the probable influence of global standards for AI design and use. Most notably, the European Union (EU) recently passed its AI Act, the world's first comprehensive AI legislation. The Act utilizes a tiered risk-based framework to regulate technologies only to the degree that they have the potential to cause harm, ranging from unacceptable risk to

[10] Nicol Turner Lee and Jack Malamud, "How Congress can secure Biden's AI legacy," January 25, 2024 https://www.brookings.edu/articles/how-congress-can-secure-bidens-ai-legacy/

minimal risk.[11] Systems that pose unacceptable risk – such as in real-time biometric identification or social scoring – are banned outright. Other uses, such as AI image models, are subject to light transparency requirements, such as labeling of AI-generated outputs. Here, the Commission has also taken a leadership position in promoting AI safety with the establishment of its landmark AI Safety Institute and through organizing and convening the global AI Safety Summit.[12] The Summit produced the Bletchley Declaration, a document signaling the members' shared priority of international cooperation in the responsible development of AI.[13] Even China has been regulating the technology, imposing rules for recommender systems, deepfakes, and most recently, generative AI.

While AI will always empower bad actors, including foreign governments who spread disinformation at scale, having clear guidance at the national level with an "all-hands-on deck" strategy by government will matter, especially as generative AI technology makes the distilling of disinformation increasingly difficult to detect. This is particularly concerning as the United States and over half the world's population plan to go to the polls in major elections this year. The spread of mis- and disinformation and the dangers they pose to elections are not unique to AI, but AI technology does increase their risk, and AI tools in the hands of malicious, state actors may prove particularly dangerous.

---

[11] "EU AI Act: first regulation on artificial intelligence," December 19, 2023 https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

[12] Dan Milmo and Kiran Stacey, "Five takeaways from UK's AI safety summit at Bletchley Park," November 2, 2023 https://www.theguardian.com/technology/2023/nov/02/five-takeaways-uk-ai-safety-summit-bletchley-park-rishi-sunak

[13] Kiran Stacey and Dan Milmo, "UK, US, EU and China sign declaration of AI's 'catastrophic' danger," November 1, 2023 https://www.theguardian.com/technology/2023/nov/01/uk-us-eu-and-china-sign-declaration-of-ais-catastrophic-danger

The Need for Congressional Action

The White House EO has set the stage for some foundational and evolving discussions that the United States must have to develop and promote robust, and participatory AI governance in ways that manage innovation and areas where more proscriptive guidance will be needed, such as the use of AI for more predictive decision-making in policing, housing eligibility, credit worthiness, and other quality of life verticals where real humans are dependent on reliable and trustworthy, autonomous systems.

On both sides of the aisle, Congress has begun to act on various legislative drafts relating to the AI regulatory ecosystem. However stronger consensus is needed to prioritize action on immediate and incremental legislation to promote transparency in AI systems and the democracy-impacting decisions of autonomous AI. Bipartisan proposals such as the National AI Commission Act, which would establish a bipartisan commission to draft a comprehensive regulatory framework on AI consolidating existing and proposed efforts, signal progress but face unsure futures considering lacking support from House leadership[14]. Similarly, progress on the Protect Elections from Deceptive AI Act, a bill which would ban the use of AI to generate "materially deceptive content falsely depicting federal candidates in political ads to influence elections," would be especially pertinent during this election year.[15] Indeed, Congress is perhaps closer when it comes to exploring options for

[14] Rep. Ted Lieu (D-CA-36), H.R. 4223, "National AI Commission Act," https://www.congress.gov/bill/118th-congress/house-bill/4223/text

[15] Sen. Amy Klobuchar (D-MN), S.2770, "Protect Elections from Deceptive AI Act," https://www.congress.gov/bill/118th-congress/senate-bill/2770?q=%7B%22search%22%3A%22artificial+intelligence%22%7D&s=2&r=40

effective governance with many of the more difficult concerns being addressed in the White House EO and supporting documents.

To this end, I propose a few areas where Congress has greater possibility of finding alignment. Bipartisan agreement could come in the regulation of AI-reliant facial recognition technology (FRT) – an area in which I spent more than a year as part of a research commission formed to evaluate its use by law enforcement. FRT has always presented a disproportionate burden to the communities of color, and others with disproportionate interaction with law enforcement, and who faced the consequences of misidentification exhibited in unfair arrest, detainment, and the mitigation of innocence. Congress could also pass comprehensive data privacy legislation to ensure that Americans' rights are protected as AI models become more increasingly intrusive. Further, basic practices like disclosures of AI-generated content, aspects of digital watermarking, compliance with existing civil and human rights laws, and protections of creators and artists – whose work products are usurped into technological models – could also be a baseline of agreement as comprehensive efforts by the White House and government agencies progress.

Conclusion

In sum, by turning previous voluntary frameworks and guidelines into executive action, the EO represents a crucial step forward in establishing a much-needed US framework for responsible innovation in the development and deployment of advanced artificial intelligence. However, to capitalize on this step and avoid a patchwork of state laws that would eventually have to be addressed through preemption, Congress must act

on the various bipartisan bills that lawmakers have already proposed to continue the development of an effective and world-leading AI governance regime. We could also choose to lag behind international competitors, especially China, in how we express our values, norms, and standards around emerging and evolving technologies.

I thank you for the opportunity to testify and look forward to your questions.