CENTER FOR DEMOCRACY & TECHNOLOGY

Testimony of Alexandra Reeve Givens
President & CEO, Center for Democracy & Technology

For the U.S. Senate Committee on Homeland Security and Government Affairs,
Hearing Entitled "Artificial Intelligence: Risks and Opportunities"

March 8, 2023

Members of the Committee, thank you for inviting me to speak about the challenges and opportunities presented by artificial intelligence. I am the President & CEO of the Center for Democracy & Technology (CDT), a 28-year old nonprofit, nonpartisan organization that works to protect users' civil rights, civil liberties and democratic values in the digital age.

CDT fights for policies and practices that protect users' interests — in areas ranging from commercial data practices, to government surveillance technology, to online content moderation, to the use of technology in education and the delivery of government services. Artificial intelligence is already transforming each of these areas, so I am grateful for the Committee's focus on this important topic today.

While artificial intelligence has the potential to generate new insights and make processes more efficient, it also poses risks of being unreliable, biased, and hard to explain or hold accountable. My testimony focuses on these risks in several areas that directly impact consumers: (i) when AI or automated systems are used in decisions impacting people's access to economic opportunities, such as in employment, housing, and lending; and (ii) in the administration of government services, such as when AI or automated systems are used to detect fraud or determine eligibility for public benefits programs.

When AI systems are deployed in these high-risk settings without responsible design and accountability measures, it can devastate people's lives. A person may be unfairly rejected from a job, be denied or unable to find housing, or be wrongly accused of fraud and stripped of the benefits they need to support their family. When this happens, the harm is felt not just by the people whose lives are upended by the decision, but also by the businesses or government programs that rely on those systems to work. Those businesses are now bought into a system that is unfit for purpose, and may face legal, financial, and reputational consequences. This is why it benefits *everyone* to have upfront, realistic conversations about the potential risks in certain AI uses – and why we need a cross-society effort to improve the responsible design, deployment, use and governance of AI.

The public conversation about responsible AI has matured significantly in recent years. There is now a robust research literature and many documented examples illustrating the potential risks of harm in various settings that affect consumers and workers.[1] Large companies are acknowledging these risks,[2] and there are high-profile government, multi-stakeholder and industry efforts focused on principles for the responsible use and governance of AI.[3] But we find ourselves at an inflection point. It is time to move beyond simply describing the potential risks

[1]See, e.g. annual proceedings of the ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT) and the AAAI/ACM Conference on Artificial Intelligence, Ethics & Society (Aies); tracks within the annual conferences of the Association for the Advancement of Artificial Intelligence, International Conference on Machine Learning, and Conference on Neural Information Processing Systems, among others.
[2] See, e.g., Microsoft's Responsible AI principles and resource center, https://www.microsoft.com/en-us/ai/responsible-ai; IBM AI ethics principles and resource center, https://www.ibm.com/artificial-intelligence/ethics; Google AI principles https://ai.google/principles/; Intel Responsible AI Pillars https://www.intel.com/content/www/us/en/artificial-intelligence/responsible-ai.html.
[3] See, e.g., OECD Principles on Artificial Intelligence (adopted May 21, 2019), available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449; Global Partnership on AI, https://gpai.ai/projects/responsible-ai/; in addition to G7 and G20 initiatives. Within the U.S., the National Institute for Standards & Technology recently released its Congressionally-mandated AI Risk Management Framework, and the National Science Foundation has issued various funding opportunities that focus on responsible AI (for an overview of U.S. government-backed efforts, see https://www.ai.gov/strategic-pillars/advancing-trustworthy-ai/). For multistakeholder and industry initiatives, see, e.g, IEEE Global Initiative On Ethics Of Autonomous And Intelligent Systems, https://standards.ieee.org/industry-connections/ec/autonomous-systems/; ISO work on artificial intelligence https://www.iso.org/committee/6794475/x/catalogue/;
Business Software Alliance Framework to Build Trust in AI, https://ai.bsa.org/; Business Roundtable Roadmap for Responsible AI, https://www.businessroundtable.org/policy-perspectives/technology/ai.

of AI systems and articulating high-level principles. We need a cross-society effort to meaningfully and concretely address those risks—protecting consumers and workers, guiding businesses, and shaping innovation to ensure that America's global AI leadership is grounded in a true commitment to trust, fairness, and democratic values.

As this Committee has recognized, the federal government can be a leader in modeling the responsible design, procurement, use and governance of AI, as well as in training responsible AI leaders, and ensuring federal research dollars focus not just on AI innovation, but on measuring and addressing potential harms. This Committee has already taken several important steps in this regard, passing the AI in Government Act, the Advancing American AI Innovation Act, the Artificial Intelligence Training for the Acquisition Workforce Act, the NAIRR Task Force Act, reporting out the Government Ownership and Oversight of Data in Artificial Intelligence Act, and more.

CDT hopes the Committee builds on this progress in the years ahead, and encourages Committees of other jurisdictions and appropriate federal agencies to do the same.

I. AI and Economic Opportunities

Increasingly, AI-driven tools are being used to inform decisions about employment, lending, insurance, tenant screening and in other settings that impact people's access to economic opportunities.[4] Today, I will focus on the use of AI in employment as an illustrative example, because it demonstrates the types of harm that can arise from poor design and governance, and

[4] Examples of these use cases are well described in the technical companion to the White House Office of Science & Technology Policy's Blueprint for an AI Bill of Rights (2022), https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf and NIST Special Publication 1270, *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence* (2022), https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf. The testimony of Prof. Suresh Venkatasubramanian also sets forth several examples in further detail.

how the breadth of stakeholders involved in using AI tools complicates the task of "responsible AI."

In the employment context, an increasing number of businesses are using AI and other automated systems to recruit, hire, evaluate, manage, and even terminate workers.[5] In hiring, these tools include resume screening programs that analyze the words used in candidates' resumes, tools that analyze video interviews, and computer games or quizzes that purport to measure a candidate's personality traits and use them to predict that candidate's "fit" for a particular job.[6]

In many cases, these tools are created by analyzing "successful" employees to identify traits for which future candidates are then assessed.[7] The risks in this approach are obvious: if the data used to train the AI system is not representative of wider society or reflects historical patterns of discrimination, it can reinforce existing bias and lack of representation in the workplace.[8] In one notorious example of this phenomenon, a resume screening tool was found to score candidates higher if their name was "Jared" and the word "lacrosse" appeared on their resume, even though those factors have no impact on job performance.[9] Similarly, Amazon famously discovered that a resume screening tool it was developing penalized female job applicants by assigning lesser value to resumes that referenced women's colleges or women's sports teams (they scrapped the

[5] Society of Human Resource Managers, "Fresh SHRM Research Explores Use of Automation and AI in HR" (Apr. 13, 2022), https://www.shrm.org/about-shrm/press-room/press-releases/pages/fresh-shrm-research-explores-use-of-automation-and-ai-in-hr.aspx ("nearly 1 in 4 organizations report using automation or artificial intelligence to support HR-related activities, including recruitment and hiring").
[6] See generally Ifeoma Ajunwa, *Protecting Workers' Civil Rights in the Digital Age*, 21 N.C.J.L & Tech. 1 (2020); see, also, e.g., Oracle: AI in Human Resources: The Time is Now (2019), available at https://www.oracle.com/a/ocom/docs/applications/hcm/oracle-ai-in-hr-wp.pdf.
[7] See, e.g., Keith E. Sonderling, Bradford J. Kelley, and Lance Casimir, *The Promise and The Peril: Artificial Intelligence and Employment Discrimination*, 77 U. MIA L. Rev. 1 (2022), https://repository.law.miami.edu/umlr/vol77/iss1/3.
[8] *Id.*
[9] Dave Gershon, "Companies are on the hook if their hiring algorithms are biased," Quartz, Oct. 22, 2018, https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased.

project).[10] In both cases, the AI tool was biased in ways that reflected larger systemic inequalities, and unfit because it was not accurately assessing the candidates most suited to the job.

Other types of hiring tools rate candidates based on how they perform in online games or answer quizzes, assessing candidates for qualities like "empathy," "humility", and "emotional stability."[11] Researchers have questioned reliance on such subjective and abstract traits, as well as whether the tools even measure what they claim to.[12] In one article published in the *MIT Technology Review*, a researcher conducted her portion of an English-language automated video interview *in German*, and yet was still determined to be a 73% personality match for the job.[13] When asked about the result, a psychologist working with the company said that the algorithm "pulled personality traits from her voice."[14] This raises significant questions about the tool's effectiveness, transparency in what it was measuring, and the risk of illegal discrimination because voice intonation can vary based on age, gender, nationality, disability, and other protected characteristics.

Both Republican- and Democrat-appointed members of the Equal Employment Opportunity Commission have sounded the alarm about these and other uses of AI in employment, as have members of Congress and the White House.[15]

---

[10] J. Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," Reuters, Oct. 2018. https://www.reuters.com/article/ us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.
[11] See, e.g., Aaron Konopasky, *Pre-Employment Tests of Fit Unde the Americans With Disabilities Act*, 30 S. Cal. Rev. L. & Soc. Just. 209 (2021), https://gould.usc.edu/students/journals/rlsj/issues/assets/docs/volume30/spring2021/Konopasky.pdf.
[12] *See, e.g.,* Alene Rhea, Kelsey Markey, Lauren D'Arinzo, Hilke Schellmann, Mona Sloane, Paul Squires, Julia Stoyanovich, *Resume Format, LinkedIn URLs and Other Unexpected Influences on AI Personality Prediction in Hiring: Results of an Audit* (AIES 2022), available at https://nyuscholars.nyu.edu/en/publications/resume-format-linkedin-urls-and-other-unexpected-influences-on-ai.
[13] Sheridan Wall and Hilke Schellmann, "We Tested AI Interview Tools. Here's What We Found," MIT Tech. Rev., Jul. 7, 2021, https://www.technologyreview.com/2021/07/07/1027916/we-tested-ai-interview-tools/.
[14] *Id.*
[15] Equal Employment Opportunity Commission, "EEOC Launches Initiative on Artificial Intelligence and Algorithmic Fairness," Oct. 28, 2021,

Lessons to be learned

The hiring example is illustrative of several core concerns about how AI systems can impact people and businesses alike: concerns this Committee and others should keep in mind as they consider the risks and opportunities of AI.

First, poorly designed and governed AI systems can cause not just individual, but systemic harm. In the context of employment, an AI tool replaces the risk of a "bad apple" human reviewer with a system that could perpetuate ineffectiveness and discrimination at scale, under the veil of data-based "objectivity." The resulting harms may not be limited to a single company, but across an entire sector when AI tools are repurposed for multiple companies.

Second, harms do not just impact the people who are the subject of a decision, but also the businesses that rely on these tools to work. In the hiring context, employers are understandably intrigued by AI's promised efficiencies, often without knowing the risks or having meaningful tools or standards by which to judge the products being sold. As a result, employers may buy products that are unfit for purpose and expose them to legal, financial and reputational liability. Some vendors have responded by publishing statements about their product testing, which upon closer examination fall far short.[16] We need to improve the availability of robust, use-specific

https://www.eeoc.gov/newsroom/eeoc-launches-initiative-artificial-intelligence-and-algorithmic-fairness; Keith E. Sonderling, "Op-Ed: Artificial Intelligence is Changing How HR is Handled at Companies. But Do Robots Care About Your Civil Rights?," Chicago Tribune, Sep. 20, 2021, https://www.chicagotribune.com/opinion/commentary/ct-opinion-robots-ai-civil-rights-amazon-20210920-tef7m7az3rgjtacauazvw3u224-story.html; Blueprint for an AI Bill of Rights (2022); "Bennet, Colleagues Call on EEOC to Clarify Authority to Investigate Bias in AI-Driven Hiring Technologies," Dec. 8, 2020, https://www.bennet.senate.gov/public/index.cfm/2020/12/bennet-colleagues-call-on-eeoc-to-clarify-authority-to-investigate-bias-in-ai-driven-hiring-technologies.

[16] *See, e.g.*, Matthew Scherer, "HireVue "AI Explainability Statement" Mostly Fails to Explain What it Does," Sep. 8 2022, https://cdt.org/insights/hirevue-ai-explainability-statement-mostly-fails-to-explain-what-it-does/?utm_source=rss&utm_medium=rss&utm_campaign=hirevue-ai-explainability-statement-mostly-fails-to-explain-what-it-does (noting how the competencies that one vendors' assessments claim to measure "are not moored to the actual responsibilities and functions of specific jobs"); Alexandra Givens, "How Algorithmic Bias Hurts People With Disabilities," (Slate, Feb. 6, 2020),

guidance to help businesses understand the risks and limitations of AI tools, and meaningfully assess whether they and their vendors have addressed them.[17]

Third, the people who are the subject of decision making by AI tools are often at an extreme information disadvantage, as are regulators and advocates trying to identify and address potential harms. In the hiring context, job applicants often have little insight into whether an AI tool is being used to assess their candidacy, let alone how that tool may work.[18] Without increased transparency about when AI systems are being used and how they have been designed and are being tested, society will be hamstrung in its efforts to identify and address harms.[19]

Fourth, AI tools are often designed by one company and then deployed by many others in diverse settings, creating challenges for the ongoing testing that is necessary to ensure AI systems work as intended. Because AI tools learn and adapt from their real-time use, they must be audited in the environments where they are being deployed, on a recurring basis. This is complicated when tools are designed by vendors and sold to businesses who use them in their

https://slate.com/technology/2020/02/algorithmic-bias-people-with-disabilities.html (observing that some vendors now test their hiring tools to evaluate whether they discriminate against women, people of color, or other marginalized groups, but those assessments do not work for disability discrimination).

[17] In the hiring context, CDT and a coalition of civil rights organizations recently published Civil Rights Standards to support employers, legal counsel, vendors and workers evaluating these tools. *Civil Rights Standards for 21st Century Employee Selection Procedures* (CDT et al., 2022), available at https://cdt.org/insights/civil-rights-standards-for-21st-century-employment-selection-procedures/. We have also advocated for the EEOC to issue more sector-specific guidance, as well as enforcing existing employment discrimination laws (*CDT Comments on EEOC Strategic Enforcement Plan 2023-2027*, Feb. 8, 2023, https://cdt.org/wp-content/uploads/2023/02/CDT-Comments-on-EEOC-Strategic-Enforcement-Plan-FY2023-2027.pdf.)

[18] See, e.g. *Essential Work: Analyzing the Hiring Technologies of Large Hourly Employers* (Upturn, 2021), https://www.upturn.org/work/essential-work/ ("It is simply impossible to fully assess employers' digital hiring practices from the outside.")

[19] See, e.g., Ifeoma Ajunwa, *An Auditing Imperative for Automated Hiring*, 34 Harv. J.L. & Tech. 1 (2021)., https://ssrn.com/abstract=3437631.

own contextual setting.[20] We need to work through pathways of responsibility in this diffuse value chain.

These four areas illustrate the pressing need for increased guidance, resources and accountability measures to shape how the private sector understands and responds to the potential harms of AI in high-risk settings, as I explain in Section iii below.

II. Use of AI in the Administration of Government Services

Another area where AI and automated systems can impact people's economic and social wellbeing is in the administration of government services.[21] Over the past two decades there have been multiple instances of agencies using such systems in public benefits programs. This includes 1-1 facial image matching for identity verification, and the use of AI systems to detect fraud and to determine applicants' eligibility for benefits programs. Several of these uses have resulted in significant harm.

Identity Verification. In the context of identity verification, AI-driven biometric tools have been used to verify individuals' identities in order to ensure that benefits and services are being provided to the correct recipient.[22] This includes fingerprint readers to access school lunches

[20] See, e.g., Jacqui Ayling & Adriene Chapman, *Putting AI ethics to work: are the tools fit for purpose?*, AI Ethics 2, 405–429 (2022). https://doi.org/10.1007/s43681-021-00084-x. ("A third of the Impact Assessment tools focus on Procurement processes for AI systems from 3rd-party vendors, indicating the need for not only producers of AI products to engage with ethical assessment, but also the customers for these products, who will be the ones deploying the products.")

[21] My testimony does not address the use of AI or automated and predictive systems by law enforcement, which raises significant risks of harm. See, e.g. Statement of over 40 civil society organizations, Civil Rights Concerns Regarding Law Enforcement Use of Face Recognition Technology (June 2021), https://www.brennancenter.org/our-work/research-reports/coalition-statement-highlights-major-civil-rights-concerns-face.

[22] Hannah Quay-de la Vallee, "Public Agencies' Use of Biometrics to Prevent Fraud and Abuse: Risks and Alternatives", Center for Democracy & Technology, June 7, 2022, https://cdt.org/insights/public-agencies-use-of-biometrics-to-prevent-fraud-and-abuse-risks-and-alternatives/.

and 1-1 facial image matching to access a government website.[23] While biometric systems can theoretically provide functionality such as ease of use (though this depends heavily on implementation), they also raise concerns with respect to privacy and equity. From a privacy standpoint, biometric data is incredibly sensitive and cannot be changed. Consequently, the large-scale collection of this information exposes individuals to significant harm if that data is breached, or if it is re-purposed in a different context such as for law enforcement uses.

Use of biometric data also raises equity concerns. Some biometric-based systems do not perform equally well for different populations of users, placing a disproportionate burden on certain communities based on race, disability, or economic status.[24] Additionally, biometric-based systems assume a certain level of technology access and comfort. For example, systems employed by several states that used facial recognition to match a selfie against a DMV photo failed for users who were unfamiliar with how to take a sufficiently "good" selfie or who did not have access to sufficiently advanced smartphones, causing people to wait days or weeks until their identity could be verified by a human representative.[25]

---

[23] Id., *see also, e.g.* Bayometric, Biometric Solutions For Schools, https://www.bayometric.com/biometric-solution-schools-fingerprint-lunch-line/ (last visited March 5, 2023).

[24] See, e.g., Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification* (Proceedings of the Conference on Fairness, Accountability & Transparency in Machine Learning 81:77-91, 2018), https://proceedings.mlr.press/v81/buolamwini18a.html#:~:text=%26%20Gebru%2C%20T..,%2Fv81%2Fbuolamwini18a.html. Although research has shown improvements in the accuracy of face recognition technology for some systems, and the 1:1 matching used in identity verification raises different accuracy concerns than classification systems or 1:many matching, the risk of different accuracy levels for protected classes must nevertheless be directly tested for and addressed. NIST operates an ongoing Fairness Verification Testing Program, available at https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing.

[25] Todd Feathers, "Facial Recognition Failures Are Locking People Out of Unemployment Systems," Vice, June 18, 2021, https://www.vice.com/en/article/5dbywn/facial-recognition-failures-are-locking-people-out-of-unemployment-systems ("In California, 1.4 million unemployment beneficiary accounts were abruptly suspended on New Year's Eve and the beneficiaries were required to re-verify their identity using ID.me, a process which many found difficult and resulted in them waiting for weeks to reactivate their accounts while they struggled to make ends meet… The story is similar in Florida, North Carolina, Pennsylvania, Arizona, and many other states.")

Program managers must be aware of these challenges and guard against them, such as by providing efficient alternative methods for people to prove their identity, implementing robust safeguards to protect users' data, and developing clear standards for procuring and auditing third-party solutions.[26] They should also consider less individually invasive approaches, such as robust cybersecurity protections to prevent the large-scale, organized fraud attacks that many states saw during the pandemic.[27]

Fraud detection. Some state and national governments have used AI systems to search for fraud in government benefits applications. One egregious example was the MiDAS system used by Michigan's Unemployment Insurance Agency from 2013-2015, which wrongly classified between 20,000 and 40,000 people's applications as fraudulent based on errors in database linkage, among other factors.[28] In many cases, these errors destroyed applicants' credit and financial security, with low-income applicants incorrectly having their wages garnished, bank accounts levied, and being driven into bankruptcy. Government programs in the Netherlands, UK and Australia have encountered similar problems, with disastrous human consequences.[29]

---

[26] Center for Democracy & Technology, *Report: Digital Identity Verification: Best Practices for Public Agencies* (2023), available at https://cdt.org/insights/digital-identity-verification-best-practices-for-public-agencies/.
[27] Hannah Quay de la Vallee, "Combatting Identify Fraud in Government Benefits Programs," Center for Democracy & Technology, Jan. 7 2022, available at https://cdt.org/insights/combatting-identify-fraud-in-government-benefits-programs-government-agencies-tackling-identity-fraud-should-look-to-cybersecurity-methods-avoid-ai-driven-approaches-that-can-penalize-real-applicant/.
[28] Alejandro de la Garza, "States' Automated Systems Are Trapping Citizens in Bureaucratic Nightmares With Their Lives on the Line," Time, May 28, 2020, https://time.com/5840609/algorithm-unemployment/; see also Robert Charette, *Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold,* IEEE Spectrum 18, 3 (2018), https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold.
[29] Robert Booth, "Computer says no: the people trapped in universal credit's 'black hole'", The Guardian, Oct. 14, 2019, https://www.theguardian.com/society/2019/oct/14/computer-says-no-the-people-trapped-in-universal-credits-black-hole; report of the U.N. Special Rapporteur on Extreme Poverty and Human Rights (Report A/74/493, Oct. 17, 2019), https://www.ohchr.org/en/press-releases/2019/10/world-stumbling-zombie-digital-welfare-dystopia-warns-un-human-rights-expert.

Failures in these programs not only harm the program participants, but have tied up agencies in litigation for violating users' due process rights and administrative procedure obligations, among other charges.[30] In response, advocates have called for greater transparency and public accountability in how these tools are developed, used, and monitored; procurement reforms; and reasonable safeguards such as providing rapid human appeal *before* a person faces wage garnishment or other repercussions for suspected fraud.[31]

Benefits eligibility. States are increasingly turning to data-driven tools to determine applicants' eligibility for benefits, or the amount of benefits they receive under a given program. Billed as a way to increase efficiency and root out fraud, these algorithm-driven tools have been implemented without much public debate, and have also given rise to litigation about lack of fairness and transparency.[32] A report by my organization explored rulings from courts in Idaho, Arkansas, Oregon and West Virginia, finding that programs adopted to administer Home- and Community-Based Services under the Medicaid Waiver Program violated beneficiaries' due process rights because of errors in the tools' design, lack of explainability, and lack of human review and appeal.[33] The harms were severe, with people losing funds for essential in-home care they needed to live independently. As with other AI systems, advocates are calling for greater

[30] For example, the State of Michigan recently announced a $20 million settlement in a class action suit arising out of the MIDAS controversy following seven years of litigation, https://www.michigan.gov/ag/news/press-releases/2022/10/20/som-settlement-of-civil-rights-class-action-alleging-false-accusations-of-unemployment-fraud.
[31] See, e.g., the Benefits Tech Advocacy Hub, a website maintained by Upturn, Legal Aid of Arkansas, and the National Health Law Program, https://www.upturn.org/work/benefits-tech-advocacy-hub/.
[32] *See* Lydia Brown, Michelle Richardson, Ridhi Shetty, Andrew Crawford et al, *Challenging the Use of Algorithm-driven Decision-making in Benefits Determinations Affecting People with Disabilities* (Center for Democracy & Technology, 2020), https://cdt.org/wp-content/uploads/2020/10/2020-10-21-Challenging-the-Use-of-Algorithm-driven-Decision-making-in-Benefits-Determinations-Affecting-People-with-Disabilities.pdf.
[33] Id.

transparency and public accountability in how these tools are developed, used, and monitored, as well as procurement reforms and reasonable safeguards for human interventions.[34]

III. A Cross-Society Effort to Mitigate Harms

The examples I have highlighted today illustrate the potential harms AI can cause in certain high-risk settings. While there are many uses of AI, and many conversations about AI regulation and best practices to be had, these types of applications directly impacting people's rights and access to opportunity require attention now. While solutions should not rest with government alone, there are numerous steps the federal government can take to advance such work, and through so doing, improve the United States' leadership in advancing trustworthy, responsible AI.

Guidance, Resources, & Enforcement for the Private Sector.

Policymakers have an important platform from which to educate developers, deployers and users of AI about potential risks and the need to identify, measure, and mitigate against them. One valuable contribution is the National Institute of Standards and Technology's AI Risk Management Framework (AI RMF), which Congress directed NIST to create as a voluntary resource for organizations to promote trustworthy and responsible AI development.[35] The NIST Framework provides detailed recommendations about how companies can map, measure, and manage risk presented by different uses of AI, including defining the characteristics of trustworthy AI for which companies should assess their systems, and who should be included in that process.[36] Additionally, the Office of Science and Technology Policy's Blueprint for an AI

[34] Id., see also Benefits Tech Advocacy Hub (fn 31); *Challenging the Use of Algorithm-driven Decision-Making* (fn 32) at 22-23; Erin McCormick, "What Happened When a 'Wildly Irrational' Algorithm Made Crucial Healthcare Decisions," The Guardian, Jul. 2, 2021, https://www.theguardian.com/us-news/2021/jul/02/algorithm-crucial-healthcare-decisions.
[35] *See* National AI Initiative Act of 2020, P.L. 116-283.
[36] Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST, Jan. 2023), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf. ("Characteristics of trustworthy AI systems include: valid

Bill of Rights includes concrete examples of policies and practices that can mitigate harms in high-risk AI settings that impact people's rights.[37]

These efforts provide important frameworks to guide industry conduct. However, more work is needed to give guidance at the sector-specific level, and to reach into the communities of businesses and start-ups where tools are being designed, deployed and used. NIST can build on the AI RMF by developing further guidance on specific questions such as explainable AI and measuring risk, and by facilitating the creation of "profiles" and case studies that adapt the AI RMF to particular circumstances.[38] But this work will also need to take place at a sectoral level, relying on the appropriate agencies of jurisdiction such as the Equal Employment Opportunity Commission, Department of Housing and Urban Development, Department of Education, and more.[39] Those agencies know their jurisdictional sectors, receive direct complaints from consumers, and have investigative and research powers, positioning them well to issue guidance, technical assistance and resources to educate businesses about their responsibilities, and consumers about their rights.

Federal agencies also have an important role to play in enforcing existing laws, and they should use those powers even when faced with novel fact patterns. When an AI system is sold without accurately representing its effectiveness and limitations, that may be an unfair and deceptive trade practice; similarly, when an AI system has a disparate impact on protected classes, it may

and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-enhanced, and fair with harmful bias managed.")

[37] Blueprint for an AI Bill of Rights: Algorithmic Discrimination Protections, White House Office of Science & Technology Policy (2022), https://www.whitehouse.gov/ostp/ai-bill-of-rights/algorithmic-discrimination-protections-2/.

[38] NIST identifies some of these next steps in the Roadmap for the NIST Artificial Intelligence Risk Management Framework (NIST, Jan. 2023), https://nist.gov/itl/ai-risk-management-framework/roadmap-nist-artificial-intelligence-risk-management-framework-ai.

[39] The Biden Administration identified a number of these possibilities in the Fact Sheet companion to the Blueprint for an AI Bill of Rights, which listed actions by various federal agencies. Efforts should not be restricted to those listed in the Fact Sheet, since many agencies could play an important role issuing guidance to their regulated sectors.

violate long-standing civil rights laws. Federal agencies can help to educate businesses about how existing laws apply to new factual applications, as some are already.[40] Enforcement actions can ensure businesses are paying attention.

Increasing transparency and risk management processes.

At this critical moment, policymakers should prioritize efforts to increase transparency and accountability in how AI systems are designed and used — while also fostering the creation of robust methodologies for measuring and addressing AI harms.

Several legislative proposals have been introduced with the goal of transparency and accountability in mind, including the Algorithmic Accountability Act, and the algorithmic impact assessment provision of the bipartisan American Data Privacy & Protection Act, the comprehensive federal privacy bill that last year received a near-unanimous vote in the House Committee on Energy & Commerce and is expected to be reintroduced this year.

While not a solve-all, these approaches establish important norms: they ask the developers of AI systems in high-risk settings to disclose how their tools are designed, to test them, and to share the analysis of those tests with an outside regulator. The effect of these bills would be to increase transparency about when and where high-risk AI systems are being used, and to normalize the principle that companies designing and deploying AI tools in high-risk settings must first analyze and document how they work, accounting for the potential risks and steps they have

[40] See Federal Trade Commission blogpost, "Keep Your AI Claims in Check," Feb. 27, 2023, https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check; EEOC/Dep't of Justice Technical Assistance Document, "The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees," May 12, 2022, https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence.

taken to mitigate those risks. Such a risk management process should be part of any normal business process, as NIST's AI Risk Management Framework helps show.

At the same time as policymakers consider the need to mandate algorithmic impact assessments or algorithmic audits in high-risk settings, businesses and consumers alike will benefit from increased focus on how to measure AI harms and assess the effectiveness of harm mitigations. As noted above, a business owner deciding whether to purchase and use an AI hiring tool must currently do their own analysis of its effectiveness or rely on assertions from the vendor, which can be woefully insufficient, potentially placing that business owner at legal risk. Businesses and consumers will benefit from more robust, well-vetted approaches to assessing harms, and the government can help advance this conversation.

NIST's AI-RMF Roadmap calls for NIST to work with the broader community to "develop tools, benchmarks, testbeds, and standardized methodologies for evaluating risks in AI and system trustworthiness, including from a socio-technical lens." This work is critical to help distill the varying approaches to risk measurement that are being explored by researchers and industry, and to move towards reliable standards that non-expert businesses and consumers can trust. Meaningful engagement on such work will also ensure the U.S. can contribute to ongoing international conversations on AI risk measurement and standards, an essential step for U.S. thought leadership on AI.[41] While NIST has an essential role to play in this endeavor, the work will also benefit from increased investment and prioritization by the National Science Foundation, and by federal government agencies leading by example in the government's own assessments when procuring, developing and funding AI tools.

[41] See U.S.-EU Joint Roadmap on AI Evaluation and Measurement Tools, Dec. 1, 2022, https://www.nist.gov/system/files/documents/2022/12/04/Joint_TTC_Roadmap_Dec2022_Final.pdf; National Institute for Standards & Technology, "U.S. Leadership In AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools", Aug. 9, 2019, https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf.

Of course, increased transparency and improved methods for risk measurement will only go so far: for some uses of AI, enforcement of existing laws and even further legislation will be needed to protect consumers and workers and to prevent other harms. But this work is an important step, and one the government can ramp up now to expedite trustworthiness in private and public uses of AI.

Leading through the federal government's use and funding of AI

As this Committee has recognized, the federal government has an essential role to play in its own responsible procurement, design, deployment, use and funding of AI systems. The Committee has already passed multiple bills with this goal in mind. The AI in Government Act of 2020 included important provisions for the Office of Management and Budget (OMB) to issue a memorandum to federal agencies that provides guidance and principles for the federal acquisition and use of AI, including for assessing and mitigating bias and avoiding unintended consequences. Coupled with Executive Order 13,859 and Executive Order 13,960, these mandates create an important framework for OMB to guide federal agencies, for federal agencies to inventory their uses of AI and publish plans to comply with OMB's guidance, and for this work to be completed annually going forward.[42] This important work should continue without delay.

As the federal government considers its path forthward, it can and should also consider how the NIST AI Risk Management Framework, the Blueprint for an AI Bill of Rights, and the principles

[42] Executive Order 14091 (Feb 16, 2023), includes further directives as to how federal agencies shall consider equity when designing, developing, acquiring and using AI, and requires consultation with agencies' civil rights offices. See Executive Order 14091, https://www.federalregister.gov/documents/2023/02/22/2023-03779/further-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal.

set forth in the relevant Executive Orders can be leveraged in this process to guide agency actions and assessments. Urged by bipartisan members of this Committee,[43] the National AI Research Resource (NAIRR) Task Force has already shown one way in which responsible AI frameworks can guide federal research efforts, recommending that the NAIRR "should set the standard for responsible AI research through the design and implementation of its governance processes," and "develop[] criteria and mechanisms for evaluating proposed research and resources for inclusion in the NAIRR from a privacy, civil rights, and civil liberties perspective" that "draw from the expectations. . . described in the Blueprint for an AI Bill of Rights as well as best practices defined in the AI Risk Management Framework."[44]

The Administration (and this Committee) can also consider ways to further support agencies' efforts to pursue responsible AI. A key step would be further supporting and resourcing the National Artificial Intelligence Initiative Office that Congress created in the National AI Initiative Act, to ensure it can reach its potential as an effective resource to "promote access to technologies, innovations, best practices, and expertise to agency missions and systems across the Federal Government."[45] The National AI Initiative Office has an additional important mandate to "conduct regular public outreach to diverse stakeholders, including through the convening of conferences and educational events", which requires resources and support to achieve. Further work could also be done to amplify other shared agency resources within the Federal Government, including the work of the General Services Administration and its AI

[43] Letter from Senators Portman, Heinrich, Reps. Gonzalez, Eshoo, to the Office of Science & Technology Policy and National Science Foundation regarding the National AI Research Resource, Jan. 27, 2022, https://www.hsgac.senate.gov/media/minority-media/portman-heinrich-gonzalez-eshoo-send-bipartisan-bicameral-letter-supporting-the-national-ai-research-resource/ ("In reiterating the congressional intent undergirding the NAIRR Task Force, we encourage you to expand your ongoing efforts related to developing and deploying safe and ethical AI, and urge you to use the NAIRR Task Force as a valuable tool in those efforts.")

[44] Report: Strengthening and Democratizing the U.S. Artificial Intelligence Innovation Ecosystem: An Implementation Plan for a National Artificial Intelligence Research Resource (National Artificial Intelligence Research Resource Task Force, Jan. 2023), at vi, 24-25 https://www.ai.gov/wp-content/uploads/2023/01/NAIRR-TF-Final-Report-2023.pdf.

[45] About - National Artificial Intelligence Initiative Office, https://www.ai.gov/about/#NAIIO_-_National_Artificial_Intelligence_Initiative_Office (last visited Mar. 5, 2023).

Center of Excellence,[46] the United States Digital Service, and the work of the Administrative Conference of the United States to ensure agencies comply with due process obligations and other administrative law requirements when procuring, designing, developing or using AI.[47]

This non-exhaustive list captures some of the diverse ways in which federal agencies and the Executive Office of the President, Congress, and this Committee can continue to address some of the potential risks of AI that directly impact the American people.

I thank the Committee for its continued attention to this important work. Only with attention to these and related issues can we be confident that the U.S. is leading in *responsible* innovation, protecting its citizens, and helping businesses and government agencies know when they can trust and responsibly use emerging AI tools.

[46] General Services Administration, AI Center of Excellence, https://coe.gsa.gov/coe/artificial-intelligence.html (last visited Mar. 5, 2023).
[47] Administrative Conference of the United States AI resources, https://www.acus.gov/ai (last visited Mar. 5, 2023).