S. Hrg. 118–25

ARTIFICIAL INTELLIGENCE: RISKS AND OPPORTUNITIES

HEARING

BEFORE THE

COMMITTEE ON HOMELAND SECURITY AND GOVERNMENTAL AFFAIRS UNITED STATES SENATE

ONE HUNDRED EIGHTEENTH CONGRESS

FIRST SESSION

MARCH 8, 2023

Available via the World Wide Web: http://www.govinfo.gov

Printed for the use of the
Committee on Homeland Security and Governmental Affairs

U.S. GOVERNMENT PUBLISHING OFFICE
52–483 PDF WASHINGTON : 2023

COMMITTEE ON HOMELAND SECURITY AND GOVERNMENTAL AFFAIRS

GARY C. PETERS, Michigan, *Chairman*

THOMAS R. CARPER, Delaware	RAND PAUL, Kentucky
MAGGIE HASSAN, New Hampshire	RON JOHNSON, Wisconsin
KYRSTEN SINEMA, Arizona	JAMES LANKFORD, Oklahoma
JACKY ROSEN, Nevada	MITT ROMNEY, Utah
ALEX PADILLA, California	RICK SCOTT, Florida
JON OSSOFF, Georgia	JOSH HAWLEY, Missouri
RICHARD BLUMENTHAL, Connecticut	ROGER MARSHALL, Kansas

DAVID M. WEINBERG, *Staff Director*
ZACHARY I. SCHRAM, *Chief Counsel*
MICHELLE M. BENECKE, *Senior Counsel*
EVAN E. FREEMAN, *Counsel*
WILLIAM E. HENDERSON III, *Minority Staff Director*
CHRISTINA N. SALAZAR, *Minority Chief Counsel*
ANDREW J. HOPKINS, *Minority Counsel*
LAURA W. KILBRIDE, *Chief Clerk*
ASHLEY A. GONZALEZ, *Hearing Clerk*

CONTENTS

WITNESSES

WEDNESDAY, MARCH 8, 2023

ALPHABETICAL LIST OF WITNESSES

APPENDIX

ARTIFICIAL INTELLIGENCE: RISKS AND OPPORTUNITIES

Wednesday, March 8, 2023

U.S. SENATE,
COMMITTEE ON HOMELAND SECURITY
AND GOVERNMENTAL AFFAIRS,
Washington, DC.

The Committee met, pursuant to notice, at 10 a.m., in room SD–562, Dirksen Senate Office Building, Hon. Gary C. Peters, Chairman of the Committee, presiding.

Present: Senators Peters [presiding], Hassan, Sinema, Rosen, Padilla, Blumenthal, Johnson, and Scott.

OPENING STATEMENT OF SENATOR PETERS[1]

Chairman PETERS. The Committee will come to order.

Today's hearing will discuss both the potential risks as well as the opportunities associated with artificial intelligence (AI), examining how artificial intelligence affects our nation's competitiveness on a global stage, and discuss ways to ensure that these technologies are used both safely and responsibly.

The adoption of artificial intelligence in government, industry, and civil society has led to the rapid growth of advanced technology in virtually every sector, transforming millions of Americans' lives, millions of Americans all across our country.

From the development of lifesaving drugs and advanced manufacturing to helping businesses and governments better serve the public, to self-driving vehicles that will improve mobility and make our roads safer, artificial intelligence certainly holds great promise.

But this rapidly evolving technology also presents potential risks that could impact our safety, our privacy, and our economic and national security. We must ensure that the use of this technology becomes more widespread. We have to make sure that there are also the right safeguards in place to ensure it is being used appropriately.

One of the greatest challenges presented by artificial intelligence is the lack of transparency and accountability in how algorithms reach their results. Often, not even the scientists and the engineers who design the AI models fully understand how they arrive at the outputs that they produce. This lack of visibility into how AI systems make decisions creates challenges for building public trust in their use. AI models can also produce biased results that can have

[1] The prepared statement of Senator Peters appears in the Appendix on page 33.

unintended, but harmful consequences for the people interacting with those systems.

Some AI models, whether because of the data sets they are trained on or the way in which the algorithm is applied, are at risk of generating outputs that discriminate on the basis of race, sex, age, or disability.

Whether these systems are being used in criminal justice, college admissions, or even determining eligibility for a home loan, biased decisions and the lack of transparency surrounding them, can lead to adverse outcomes for people who may not be even aware that AI has played a role in the decisionmaking process. Building more transparency and accountability into these systems will help prevent any kind of bias that could undermine the utility of AI.

While many government organizations and businesses are working to build AI systems that enhance our daily lives, we must be open-eyed about the risks presented by bad actors and adversaries who may use AI to intentionally cause harm, or undermine our national interests.

Generative artificial intelligence like Chat Generative Pre-trained Transformer (ChatGPT) or deepfakes can be used to create convincing, but false information that can distort reality, undermine public trust, and even be used to cause widespread panic and fear in a worst-case scenario.

The risks from this kind of improper use also extend beyond our borders. Adversaries like the Chinese government are racing to be the world leaders in these technologies and to harness the economic advantages that dominance in artificial intelligence will certainly create. The United States must be at the forefront of developing our own AI systems and training people how to use them appropriately, to protect our global economic competitiveness.

If we do not, not only are we at risk of American entities having to purchase these mature technologies from an economic competitor like the Chinese government, there will be tools with little accountability that have been developed by an adversary that does not share our core American values, a serious national security risk.

Finally, artificial intelligence will have a significant impact on the future of work. There is no question that AI systems have the potential to disrupt the workplace as we currently know it. That is why it is essential as the United States develops these technologies, we are also developing a workforce that is ready to work alongside them. We must address concerns that AI tools could replace human workers and instead focus on how they can assist humans and enhance the workplace.

Our goal in today's hearing is to examine these types of risks and challenges and discuss what steps Congress should take to ensure that we are able to harness these benefits and opportunities with this technology. This includes ensuring that these technologies are used appropriately, and to protect the civil rights and civil liberties of all Americans.

Last Congress, I passed bipartisan laws that took steps to ensure the appropriate use of artificial intelligence by government, including through procurement safeguards and by boosting the knowledge of the acquisition workforce, to ensure they are properly trained to understand the risks and capabilities of these tech-

nologies. I look forward to building on those efforts this Congress, and working alongside my colleagues on the Committee to support the development of AI technologies, and ensure that they are being used both appropriately and effectively.

I hope that today's discussion will be the first of several on this important topic, and I am pleased to have our panel of witnesses with us today who are experts in the field of artificial intelligence and who can discuss the adoption of these systems and the broader impacts on industry, civil society, and government.

With that I would like to now turn it over to our acting Ranking Member, Senator Johnson.

OPENING STATEMENT OF SENATOR JOHNSON

Senator JOHNSON. Thank you, Mr. Chairman. I would like to start by asking consent to enter Dr. Jordan B. Peterson's testimony into the record.[1]

Chairman PETERS. Without objection.

Senator JOHNSON. Thank you. Now let me explain why I had to do that. Last Thursday, I was pretty late in the process, and I was asked by Ranking Member Paul to act as Ranking Member because he had a conflict with a pretty important hearing in Senate Foreign Relations Committee (SFRC). I was happy to do so because I have been very interested in the subject. Artificial intelligence has incredible impact, or will have incredible impact on our society and on individuals, and so I have been doing a fair amount of research on the topic. As a result I became aware of Jordan Peterson's interest in the topic as well.

As a matter of fact, two weekends ago I watched about an hour-and-a-half-long video of him interviewing Jim Keller and Jonathan, I think it is Pageau—I apologize if I am mispronouncing his name—on this topic. Again, they were thinking deeply about this subject and its impact on society.

First of all, who is Jordan B. Peterson? He is an author, a psychologist, an online educator, and professor emeritus at University of Toronto. For 20 years he taught some of the most highly regarded courses at Harvard and the University of Toronto, while publishing more than 100 well cited scientific papers and maintaining an active clinical and consulting practice. His international lecture tour has sold out more than 400 venues, and his best-selling books include 12 Rules for Life: An Anecdote for Chaos and Beyond Order: 12 More Rules for Life.

Unfortunately, the Chairman did not allow him to appear remotely, and we can talk about that a little bit later. But in lieu of an opening statement what I would like to do is read some of the key excerpts out of Dr. Peterson's testimony. We will see the insight and the thoughtfulness that we are missing by not having him here today.

He starts his testimony talking about the large language models, for example, like ChatGPT. He says, "Advanced large language models such as ChatGPT have burst onto the scene with a vengeance in the last six months. ChatGPT recently completed the standardized test (SAT) and scored 1020. A score of 1020 is equiva-

[1] The statement submitted by Dr. Peterson appears in the Appendix on page 65.

lent to an intelligent quotient (IQ) of about 110, which make ChatGPT more intelligence than 75 percent of people.

"The significance of all this should not be underestimated. We now have AI systems capable of engaging in genuine conversation, able to write, able to produce computer code, able to 'think,' and they will be much smarter very soon."

He goes on to talk about the rights given to the extended digital self. He writes, "For centuries we were also simple enough so that our name sufficed to identify us. Online, however, things are very different. Our digital identity is composed of the tools we use—the apps, programs, services, websites, et cetera—that we choose voluntarily to employ, as well as the records of our virtual behavior, our browsing patterns, our purchases, our records of travel, but the written communications and images we issue on platforms such as Instagram, Facebook, and more ominously, TikTok, which essentially operates under the control of the Chinese Communist Party (CCP). That extended digital self has very few rights, and our legal structure has not been able to adapt itself to the immense changes on the virtual front.

"The logical extension of such danger, and most likely outcome," in his estimation, "is the duplication in the West of something approximating the utter catastrophe of a so-called social credit system in China. Everything is tracked and controlled. The government can, with the stroke of a pen, seize the economic resources of any given individual or group." In parentheses he says, "Something that happened very ominously in Canada in the case of the truckers' convoy."

He goes on, "Developing AI capabilities will radically extend the surveillance State. China has about 400 cameras watching every 1,000 people. We could well be entering an era of authoritarian AI-mediated social shunning. The use of cameras should be banned. Machines should never be given the authority to ticket, try, punish, or limit the economic or practical activities with human beings." He goes on to talk about additional dangers. "In the next year, AI wizards will produce intelligence systems that will be able to produce representations of any person, doing anything that can be described, the so-called deepfakes. Imagine those being released on the eve of a critical election. Then imagine that happening everywhere, on every issue, thousands of times. Imagine being entirely unable to determine day-to-day what communication, from what person, photos, videos, auto recordings, writings is real and what is false. Then imagine that now, not in some distant future. That is where we are at. Steps must be taken on the legal front to make false digital representations of living persons not only illegal but seriously illegal."

He concludes, "The development of AI systems as intelligent as we are"—and I would add probably even more so—"is not some future possibility but a current actuality. The melding of AI-mediated intelligence systems with our capacity for monitoring and surveillance prepares the way for a tyranny so comprehensive that we can barely imagine it."

Now again, these are just excerpts from his testimony, and I wish Dr. Peterson could have been here remotely to offer that. But for whatever reason, even though we have the technology here, the

Chairman said he could not appear, we could not make it possible for him to appear remotely.

Now behind the scenes over the weekend, there were other reasons supplied. Talk about some book. That was all a ruse. It was a pretext for not allowing Dr. Peterson to testify, and I really cannot guess why. Some kind of ideological reason.

By the way, it was not unusual to get a witness pretty late in the process. As Chairman of this Committee over six years, it was very rare that I got testimony much more than the day before. Sometimes it could be hard to arrange witnesses. This was a somewhat unusual circumstance but not that unusual. So that should not be an excuse.

So blocking Dr. Peterson because we supposedly could not accommodate a remote witness is simply not credible. For whatever reason, the Chairman and his staff did not want to allow our witness. This is an action that is beyond unfortunate and something we will not condone, which is why no Republicans will attend this hearing.

I sincerely hope the Chairman will reconsider this partisan action and not repeat it in the future.

Chairman PETERS. Senator Johnson, if I could respond to that. I have been the Chair now, this is going into the third year. We have never blocked the minority from having a witness, and we are not blocking the minority from having this witness here now.

Senator JOHNSON. Yes, you are. He is not here.

Chairman PETERS. Let me go through the process. We started putting this together a month ago, one month, we would hope, that staff, in one month's time, could come up with witnesses. We did. We have three eminent witnesses that were presented to the Ranking Member. We go through interviews. All three of you had interviews with staff from both the majority and minority. It is what we do with every witness. For every single hearing we do that. We do not want to change that policy. That is a very important policy, so we have an understanding of who the witnesses are. We have an opportunity to prepare, to make this a good hearing.

A month ago we did that. We went through the process. We continually went to the Ranking Member and said, "Please provide your minority witness. We would like to move forward. We are excited about this hearing." We did not hear anything. We had to actually put a deadline. Please, by this deadline, last week, on Thursday, please provide a witness. We did not hear.

We finally got a witness, not from the Ranking Member but another Member, at 8 p.m. on Friday, with two business days prior to a hearing. There was a request for video. This is not a hybrid hearing. We have, for well over a year, everybody has appeared in person. I know maybe Senator Johnson likes——

Senator JOHNSON.—On the technology here.

Chairman PETERS. The witnesses appear in person. They have always appeared in person, for a long period of time. Perhaps Senator Johnson likes Coronavirus Disease 2019 (COVID–19) protocols. I am not sure. But we have had personal folks here, because I think it is important to have witnesses in person. Each and every one of you arranged your schedule to be here in person. You could have done video but you knew that was the rule of the Committee.

This is not a hybrid hearing. This is to be in person, and I think you have a much better hearing as a result of that.

We said that with this new person that came in at the end that we would need them to appear in person, just like each and every one of you took the time and trouble to get here, they would have to do the same thing. Perhaps if they had more time, if we actually heard from the minority in a normal time, they would have been able to make those arrangements to be here in person.

He was welcome to be here. If he wanted to sit here today we would have welcomed that. He would have had to go through the interview. It would have been short because we only had two business days to do this. We would have had to have an interview, like each and every one of you have done, and every single witness that comes before this Committee does it.

All we are asking Senator Johnson, is let us have the same process. I told you, or I told the Ranking Member, that your witnesses, we are going to have more AI hearings, he is welcome. If you want him to be your witness at a future hearing we would welcome him. He will be the minority witness. It was a time constraint.

Senator JOHNSON. OK. We will definitely take you up on that offer. But again, there were things happening behind the scenes, and again, I did not get brought into this process until late Thursday. We scrambled. We got him to agree to be a witness. We let you know it was going to be remote. The technology is obviously available.

But again, as Chairman of this Committee, I did not take it upon myself to vet your witnesses, the minority witnesses. That is your job. If you end up with somebody with troubling circumstances around his testimony, that is on you, not on the Committee. Dr. Peterson is eminently qualified. He has been talking about this. He put a lot of work into his testimony and not able to provide it.

Again, this situation, it is just not credible that we could not accommodate him remotely. It is not unusual that it is hard to sometimes find witnesses. I cannot speak for Senator Paul in terms of why he did not make the decision not to be Ranking Member, but I acted very expeditiously. I asked an eminently qualified individual to be a witness. He agreed. He put in the work. He provided insightful and thoughtful testimony. We should have allowed him to testify remotely, but we will take your offer for the next hearing and we will communicate that to Dr. Peterson.

Chairman PETERS. Senator, we want witnesses to be here in person. This is not a hybrid hearing. It was never noticed as a hybrid hearing.

Senator JOHNSON. That is fine. We just do not want——

Chairman PETERS. I understand.

Senator JOHNSON [continuing.] The majority blocking——

Chairman PETERS. We are not blocking.

Senator JOHNSON [continuing.] Or even vetting our minority witnesses. That is honestly not your job. The minority has a right to have witnesses appear before the Committee on the topic at hand, and to have you have veto power over that is not proper.

Chairman PETERS. Again, Senator Johnson, we can provide. We sent the letter to the Ranking Member, your witness can testify. They have to be in person, and they have to have an interview like

every other witness, and yet that did not happen, and the reason it did not happen was because it was such a short timeline. I get that. I know you were thrown this responsibility at the last moment.

Senator JOHNSON. I acted expeditiously, and I came up with an excellent witness, and it would have been great to have him appear remotely.

Chairman PETERS. We would have welcomed him.

Senator JOHNSON. Hopefully we will see him in person, as long as he is not too insulted by not being able to testify here today.

Chairman PETERS. Hopefully he is not insulted that he is being treated like everybody else. If he thinks that he should be treated differently than everybody else, well, in this Committee we treat everybody fairly. Everybody is treated the same way, and we believe that those rules should be followed.

We would hope that in the future, when you have a month to prepare for a hearing that you actually do the work and prepare for a hearing, and do not expect that everybody is just going to drop everything and change all the rules and do something different. Do the work. This is an important Committee. We have always worked on a consensus basis. You and I worked on a consensus basis.

Senator JOHNSON. That is right, but I never blocked any witnesses. But anyway, enough of this. Just get on with the hearing and we will attend the next one.

Chairman PETERS. Let us hope we can return to working in a bipartisan way and have folks do the work necessary so that these hearings go forward.

With that, let us get to the important business at hand.

It is the practice of the Homeland Security and Governmental Affairs Committee (HSGAC) to swear in witnesses, so if each of you will please stand and raise your right hand.

Do you swear that the testimony that you will give before this Committee will be the truth, the whole truth, and nothing but the truth, so help you, God?

Ms. GIVENS. I do.

Mr. VENKATASUBRAMANIAN. I do.

Mr. MATHENY. I do.

Chairman PETERS. Great. Thank you.

Today's first witness is Alexandra Reeve Givens. Ms. Givens is the President and Chief Executive Officer (CEO) of the Center for Democracy and Technology (CDT), whose mission is to ensure democracy and individual rights are at the center of the digital revolution. Previously, Ms. Givens served as the founding Executive Director of the Institute for Technology Law and Policy at Georgetown Law, and as Chief Counsel for Intellectual Property and Antitrust on the Senate Judiciary Committee. Ms. Givens has also served as an adjunct professor at Columbia University School of Law.

Ms. Givens, welcome to the Committee and thank you for appearing. You are recognized for your opening statement.

TESTIMONY OF ALEXANDRA REEVE GIVENS,[1] PRESIDENT AND CHIEF EXECUTIVE OFFICER, CENTER FOR DEMOCRACY AND TECHNOLOGY

Ms. GIVENS. Thank you very much, Senator Peters, and to Members of the Committee, thank you for inviting me to speak about the challenges and opportunities presented by AI.

The Center for Democracy and Technology, is a 28 year-old nonprofit, nonpartisan organization that works to protect civil rights, civil liberties, and democratic values in the digital age. CDT protects users' interests in areas ranging from commercial data practices to government surveillance to online content moderation to the use of technology in education and government services. AI is already transforming each one of these areas, so I am grateful for the Committee's focus on the topic today.

While AI 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 written testimony focuses on these risks in several areas that directly impact consumers. First, when AI or automated systems are used in decisions impacting people's access to economic opportunities, such as in employment, housing, and lending, and second, in the administration of government services, such as when AI or automated systems are used to detect fraud or determine benefits eligibility.

When AI systems are used in these high-risk settings without responsible design and accountability, 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 only by the people whose lives are upended by the decision but also by the businesses and government programs that are relying on these systems to work. Those businesses or government agencies are now bought into a system that is unfit for purpose, and may face legal, financial, and reputational consequences. That is why it benefits everyone to address the potential risks and limitations of AI.

My written testimony details harms that have already arisen in these contexts. For example, hiring tools that systematically downgraded women's resumes or an automated video interview system where a reporter gave answers in German and yet was still found to be a 73 percent match for a company.

In the government setting, the Michigan Integrated Data Automated System (MiDAS) in Michigan wrongfully classified up to 40,000 people's unemployment insurance applications as fraudulent based on design errors in the system. People who were already on the financial brink had their wages garnished, bank accounts levied, and were driven into bankruptcy. The State faced years of litigation and recently paid millions of dollars to victims.

Government programs in Europe, the United Kingdom (UK), and Australia have had similar problems.

When assessing these concerns, policymakers should consider several factors. First, poorly designed and governed AI systems can

[1] The prepared statement of Mr. Givens appears in the Appendix on page 35.

cause not just individual but systemic harm. In the hiring context, for example, an AI tool might replace the risk of a bad apple in human resource (HR), but it does so with a system that could be ineffective and discriminatory at scale. The resulting harms may impact an entire sector when a tool is used by multiple companies.

Second, harms do not just impact the people who are the subject of the decision but the businesses and agencies relying on those tools. That is why we need robust, specific guidance to help people navigate these issues and to enforce existing laws to ensure that developers take their obligations seriously.

Third, the subjects of AI decisionmaking often have no idea they are being assessed by an automated program, let alone how that tool may work, and neither do regulators.

Without increased transparency about when AI systems are being used and how they have been designed and tested, society will be hamstrung in its efforts to identify and address harms.

Fourth, AI systems need ongoing testing in their applied environment to make sure they are working as intended. But this is complicated because AI tools are often designed by one company and then deployed by many others in different settings. We need to work through the pathways of responsibility in this diffuse value chain.

Given these challenges, we need a cross-society effort for the responsible design, deployment, use, and governance of AI. My written testimony outlines several ways in which the government can lead in this work.

The first is to rapidly scale up guidance and resources to identify AI-related harms and mitigations. We need to help those non-expert businesses and agencies think about and address risk and when to say no to these tools altogether. The National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF) and the Blueprint for an AI Bill of Rights are good examples of this, but agencies across the Federal Government must lead in their respective sectors.

Second is that we must increase transparency, which is where legislation like the Algorithmic Accountability Act or similar models can be useful. It is time to normalize the idea that companies designing and deploying AI tools in high-risk settings must first analyze and document how they work, accounting for potential risks and steps they have taken to address them.

Third, as this Committee has well recognized, the Federal Government has an essential role to play in its own procurement, design, use, and funding of AI systems.

Congress directed Office of Management and Budget (OMB) to issue guidance and principles for the Federal acquisition and use of AI, which was boosted by Executive Orders (EO) from both the Trump and Biden administrations. This work must continue without delay and we must continue to support agencies in this work, such as through the National AI Initiative Office (NAIIO) that this Committee and Congress created.

I thank the Committee for its continued work in this and related areas, and I look forward to your questions.

Thank you.

Chairman PETERS. Thank you, Ms. Givens.

Our next witness is Dr. Venkatasubramanian who currently serves as Professor of Computer Science and Data Science at Brown University. His expertise includes data mining, machine learning (ML), algorithms, and computational geometry, specifically algorithmic fairness and their impacts on decisionmaking on society.

Previously Professor Venkatasubramanian served as the Assistant Director for Science and Justice in the White House Office of Science and Technology Policy.

Professor, welcome to the Committee, and thank you for appearing. You are recognized for your opening statement.

TESTIMONY OF SURESH VENKATASUBRAMANIAN,[1] Ph.D., PROFESSOR OF COMPUTER SCIENCE AND DATA SCIENCE, BROWN UNIVERSITY

Mr. VENKATASUBRAMANIAN. Thank you, Senator Peters and Members of the Homeland Security and Government Affairs Committee. I thank you for inviting me to testify at this important hearing on the risks and opportunities of AI. I am a professor of computer science and director of the Center for Technological Responsibility at Brown University.

I recently completed a stint as tech policy advisor in the White House and helped develop the Blueprint for an AI Bill of Rights.

I have spent the last decade studying and researching the impact of automated systems, and AI, on people's rights, opportunities, and access to services. I have also spent time advising State and local governments on sound approaches to governing the use of technology that impacts people's lives.

We are here today to talk about AI, a field of study trying to design systems that can sense, interact, reason, and behave in the way humans do, and in some cases even surpass us. People learn from the data we receive, and thus one sub-area of AI that is dominant right now, fueled by the collection of vast amounts of data, is machine learning, the design of systems that can incorporate historical data into the predictions they produce, and in some cases keep adapting as more data appears.

Virtually every sector of society is now touched by machine learning, and the most consequential decisions and experiences in our lives are mediated by algorithms—where we go to school, how we learn, how we get jobs, whether we can buy a house, what kind of loan we get, whether we get credit to start a small business, whether we are surveilled by law enforcement or incarcerated before a trial, how long a sentence for a convicted individual is, and whether we can get paroled.

The list goes on and on, and keeps expanding, with systems like GPT3, ChatGPT, and Bard, and many others that ingest extremely large amounts of data and huge compute power to create the plausibly realistic conversations that have caught our imagination over the past few months.

All these systems have something in common. They are algorithms for making algorithms. The distinctive feature of a machine learning system is that the output of the system is itself an algo-

[1] The prepared statement of Dr. Venkatassubramanian appears in the Appendix on page 53.

rithm that purports to solve an underlying problem, whether it is predicting your loan worthiness, searching for a face in a video stream, or even having a conversation with an individual.

As a consequence of the above, we do not actually know for sure whether and how these algorithms work and why they produce the output that they do. This might come as a surprise given how much we hear every day about the amazing and miraculous successes of AI. Yet AI systems fail.

They fail when the algorithms draw incorrect conclusions from data. They fail when they make predictions based on faulty or biased data. They fail when the results of one AI system are fed into another, or even the same one, amplifying errors along the way. They fail when they are so opaque that errors in how they function cannot even be detected.

The truth is AI systems are not magic. AI is technology, and like any other piece of technology that has benefited us—drugs, cars, planes—AI needs guardrails so we can be protected from the worst failures while still benefiting from the progress AI offers.

What should these guardrails look like? Any automated system that has meaningful impact on our rights, opportunities for advancement, and access to critical services should be tested so it works, and works well. It should not exhibit discriminatory behavior, be limited and careful in its use of our personal data, be transparent, and easily understandable, and be accompanied by human supervision for all the times that it fails. Moreover, all these protections should be documented and reported on clearly for independent scrutiny. Congress should enshrine these ideas in legislation, not just for government use of AI but for private sector uses of AI that have people-facing impact.

I am a computer scientist—a card-carrying computer scientist, I like to say—and my work is to imagine technological futures. There is a future in which automated technology is an assistant. It enables human freedom, liberty, and flourishing, where the technology we build is inclusive and helps all of us achieve our dreams and maximize our potential.

But there is another future in which we are at the mercy of technology, where the world is shaped by algorithms and we are forced to conform, in which those who have access to resources and power control the world and the rest of us are left behind. I know which future I want to imagine and work toward. I believe it is our job to lay down the rules of the road, the guardrails and the protections, so that we can achieve that future. I know we can do it if we try.

Thank you for giving me this opportunity to speak.

Chairman PETERS. Thank you, Professor.

Our next witness is Dr. Jason Matheny. Dr. Matheny currently serves as President of the RAND Corporation, a nonprofit institution that helps provide research and analysis to solve public policy challenges. Prior to his current role, Dr. Matheny led White House policy on technology and national security at the National Security Council (NSC) and the Office of Science and Technology Policy, and was the founding director of the Georgetown Center for Security and Emerging Technology. Dr. Matheny was congressionally ap-

pointed as a commissioner to the National Security Commission on Artificial Intelligence. Welcome to the Committee.

You may proceed with your opening statement.

TESTIMONY OF JASON MATHENY, Ph.D.,[1] PRESIDENT AND CHIEF EXECUTIVE OFFICER, RAND CORPORATION

Mr. MATHENY. Thank you, Chairman Peters and members of the Committee for the opportunity to testify today.

For the past 75 years, RAND has conducted nonpartisan policy research, and we currently manage four federally funded research and development (R&D) centers for the Federal Government, including one for the Department of Homeland Security and three for the Department of Defense (DOD). Today I will focus my comments on how AI affects national security and U.S. competitiveness.

Among a broad set of technologies, AI stands out both for its rate of progress and for its scope of applications.

AI holds the potential to broadly transform entire industries, including ones that are critical to our future competitiveness, such as medicine, manufacturing, and energy. Applications of AI also pose grave security challenges for which we are currently unprepared, including the development of novel cyber weapons, large-scale disinformation attacks, and the design of advanced biological weapons.

Threats from AI pose special challenges for national security for several reasons: the technologies are driven by commercial entities that are frequently outside our national security frameworks; the technologies are advancing quickly, typically outpacing policies and organizational reforms within government; assessments of the technologies require expertise that is concentrated in the private sector and that has rarely been used for national security; and the technologies lack conventional intelligence signatures that distinguish benign from malicious use, differentiate intentional from accidental misuse, or that permit attribution with confidence.

By most measures, the United States is currently the global leader in AI. However, this may change as the People's Republic of China seeks to become the world's primary AI innovation center by 2030, an explicit goal of China's AI national strategy. In addition, both China and Russia are pursuing militarized AI technologies, intensifying the challenges that I just outlined. In response, I will highlight eight actions that national security organizations, including the Department of Homeland Security (DHS), could take.

First, ensure that DHS cybersecurity strategies and cyber Red Team activities track developments in AI that are likely to affect cyber defense and cyber offense.

Second, within the National Institute of Standards and Technology industry stakeholders and U.S. allies and partners ensure that international standards for AI prioritize safety, security, and privacy, so that the technologies are less prone to misuse by surveillance States.

Third, consider creating a regulatory framework for AI that is informed by an evaluation of risks and benefits of AI to U.S. national security, civil liberties, and competitiveness.

---

[1] The prepared statement of Mr. Matheny appears in the Appendix on page 60.

Fourth, identify the high-performance computing hardware that is used for AI as critical infrastructure that can be stolen or subverted, and consider requirements for tracking where high-performance computing hardware goes and what it is being used for.

Fifth, work with the intelligence community (IC) to significantly expand the collection and analysis of information on key foreign public-and private-sector actors in adversary States involved in AI, and create new partnerships and information-sharing agreements among Federal, State, and local government agencies, the research community, and industry.

Sixth, leverage AI expertise in the private sector through short-term and part-time Federal appointments and security clearances for leading academic and industry AI experts who can advise the government on key technology developments, with appropriate checks on conflicts of interest.

Seventh, in Federal purchases and development of AI systems, include requirements for security, safety, and privacy measures that prevent AI systems from misbehaving due to accidents or adversaries, and require socially beneficial techniques, such as privacy-preserving machine learning and watermarking to detect generated text and deepfakes.

Eighth and last, increase our investments in biosecurity and biodefense, given the potential applications of AI to design pathogens that are much more destructive than those found in nature.

I thank the Committee for the opportunity to testify, and I look forward to questions.

Chairman PETERS. Thank you, Dr. Matheny.

Professor Venkatasubramanian, this question is for you. In your statement you describe the so-called black box of the AI systems, where developers themselves do not fully understand exactly what happened in that black box as it is making those decisions. You mentioned in your opening comments and your written comments some of those risks, but for the Committee's benefit could you tell us more about the risks that are associated when you have non-transparent algorithms?

Mr. VENKATASUBRAMANIAN. Thank you, Senator, and you can call me Professor V. That is fine. My students do that too.

Chairman PETERS. Professor V?

Mr. VENKATASUBRAMANIAN. Professor V is just fine.

Chairman PETERS. Thank you.

Mr. VENKATASUBRAMANIAN. To your question, when we do not know how an algorithm works or why it works, we also do not know how it fails and under what circumstances it fails, and that is where the biggest problem is. We do not even know how to tell whether it is failing or not.

If I use, for example, an algorithm to analyze a tissue scan, to determine whether a patient has cancer, such a failed algorithm could either falsely declare a patient free from cancer, which would be catastrophic, or falsely declare that they were positive for the test and therefore have to undergo harmful treatments that could be very harmful to them. We would not be able to tell the difference.

That is why safety testing, investigation, and transparency are so critical, because of the way in which machine learning algo-

rithms, and the fact that there are algorithms for generating algorithms, create these procedures that are very hard to understand. This comes up again with things like ChatGPT, where we do not know how they do what they do. They seem to be providing plausible answers, but as we have seen, it is very easy to get them to lie, or not lie but give answers that are false because we do not understand how they are working. That is where the lack of transparency is one of the biggest problems with understanding the effectiveness and whether these systems can work.

Chairman PETERS. I appreciate that.

Ms. Givens, you have done a lot of work in this area as well. I would certainly love to have your thoughts on the black box and accountability.

Ms. GIVENS. Thanks for the question. The thing that I think about is what is meaningful transparency, and the way to think about that is as somebody is deciding, as a small business, for example, whether to use one of these tools or even large and mid-sized businesses deciding right now whether they could integrate ChatGPT into some of their offerings.

What are the resources that will help them make an informed decision? Right now there are many different tests and approaches to safety measurement, to mitigating and measuring bias, but we really need to fast-track that conversation to make sure that we are talking about well-established, robust approaches to identifying and addressing risks.

We also need to think about a conversation of internal audits versus how we make that an external process that can have more accountability and visibility from the outside.

Then, of course, how to make guidance and disclosures that are useful for users. All of those are areas where there is nascent work now, but we need to turbocharge those efforts to actually make transparency have value.

Chairman PETERS. Thank you. Dr. V, we are talking about bias in these systems. As a computer scientist you have considerable expertise in this area. Could you tell the Committee how does bias actually get into these AI systems? We should know how it gets in so we can figure out how to deal with it.

Mr. VENKATASUBRAMANIAN. Thanks for that. There is a phrase in computer science that is called "garbage in, garbage out." It means that if you put bad data into an algorithm you will get a bad outcome. In machine learning, what we talk about now is "bias in, bias out." A machine learning algorithm that takes data that has hidden biases in it will invariably, almost certainly, detect and amplify those biases in its output.

We saw this happening when a company was training a system to predict who would be good people to hire. The system started picking up signals that the candidate was a woman, even if it was not explicitly mentioned—for example, a person whose curriculum vitae (CV) said that they went to Smith College—and then it started rejecting them. It turns out that in this case it was because the data being used to train the algorithm was itself biased. It was historical data on hiring from the company, and the company, as it turned out, had skewed and gender biased hiring practices.

One very important example where bias gets into an AI system is when the underlying data used to train the algorithm has biases coming from historical context.

Chairman PETERS. Can you mitigate that by having larger datasets? Is that one way to do it, or you still have to, in some way, examine those sometimes very large datasets that are training AI.

Mr. VENKATASUBRAMANIAN. Unfortunately, merely having more data does not actually solve the problem because if that more data continues to have those kinds of biases then you will just make the problem even worse. What is required is a collection of procedures, among them procedures that examine the sources of data, examine the biases in the data, even if it is a large dataset, and try to understand how those biases might be affecting what the algorithm would do.

Another set of procedures is to understand how the algorithm training is being done. There are certain best practices for how to train algorithms to try to mitigate these forms of bias, and they need to be put into place. When you do that you can mitigate a lot of these biases.

Similarly looking at, in context, how the algorithm is used and deployed and how the results are showing up and whether biases are showing up in the output as well.

In these three ways, if you have the appropriate practices put into place you could try to mitigate some of these biases. You may not remove all of them but you can definitely go a long way toward doing that.

Chairman PETERS. Thank you.

Ms. Givens, you have told us how public conversation about responsible AI has been evolving. Could you help us understand, what would a truly responsible AI system actually look like?

Ms. GIVENS. You have already started an important conversation around bias, but I think we also need to pull out the broader frame of are these systems working as intended. There is a functionality question to be had about are we actually able to rely on rational and predictable outcomes. Is the model structured in a way to actually allow people to have trust in the results that are being generated?

When NIST produced its AI Risk Management Framework they identified a number of characteristics of what makes a trustworthy AI system, and I think it is actually a very useful way to think about these issues. For them, the factors are is it valid and reliable; safe, secure, and resilient; accountable and transparent; explainable and interpretable; privacy enhanced; and fair with harmful bias managed. Really each of those elements is its own inquiry. We need our own bodies of work as to how to make sure each of those are being maintained. But I think that is an incredibly useful way of breaking down these different elements of what it is to develop responsible AI.

Then the final piece is we have to think about this through the entire lifecycle, so not just at the moment the tool is being designed in the first place but how and where it is being deployed, what that looks like in its contextual setting, and then because these tools, the whole way that they work is by learning over time, ongoing au-

diting and checks to make sure that they are still reliable, trustworthy, and have not brought in additional biases.

That is the way we need to think about a holistic approach to these questions.

Chairman PETERS. Thank you.

Senator Blumenthal, you are recognized for your questions.

OPENING STATEMENT OF SENATOR BLUMENTHAL

Senator BLUMENTHAL. Thank you very much, Mr. Chairman, and I want to thank you for having this hearing, and the panel that we have which is, as you referred to it as truly eminent, informed, very helpful, and I welcome your willingness to have additional hearings, which I think most certainly we will want to do.

Professor, I was interested in your reference to algorithms, quoting Princeton professor Narayanan, as "snake oil." For me the danger of that snake oil is not only the mistakes that can be made, that is, the failings, and you all have identified some of those failings, but sometimes how they work all too well, the algorithms which are essentially, for most people in this world, black boxes, driving content to children. I want to thank the Chairman for his support in the efforts that we are making to protect children better than we have before. But these algorithms that work all too well will identify an interest that a child has and then continue driving content to that child. The idea that artificial intelligence is something way off in the future I think is a little bit misleading because right now Google and others are using these algorithms to drive that content.

Could you describe whether they have control—and I will ask the other Members of the panel as well—whether they have and could exercise more control over what these algorithms do and whether they could make them more transparent.

Mr. VENKATASUBRAMANIAN. Thank you for the question, Senator. I should say up front I am not an expert on matters linked to children's safety online, but as an AI expert what I can tell you is that for all of these algorithms, and the kinds you mentioned, the things we have talked about today so far, the importance of governance, the importance of transparency of how these algorithms work, of having independent review and ongoing monitoring, are critically important to make sure that they do not have the consequences that we do not want them to have.

That idea of governance in AI, it is an important part of the process of determining what is it we want out of these algorithms we are deploying. Oftentimes we do not ask that question, and algorithms are used for engagement or for selling ads, and we do not ask the question of what impact they are having.

Having a broad framework, an overarching, comprehensive framework, where we can evaluate what these algorithms are, how they work, and what they are doing is a way, in general, that we can make sure that we can get the benefits of these systems and not get the harms.

But to your specific point about child online safety I will defer to others on the panel who have more expertise.

Senator BLUMENTHAL. Ms. Givens.

Ms. GIVENS. Senator, I know you have been a longtime leader on this issue, and we have worked for a long time with your staff on comprehensive privacy protections, not just for kids but for all consumers, frankly, engaging in these online platforms, where the hyper-targeting of content and of ads really can have harmful effects.

I agree that this is a priority area. We have heard policymakers across the country and internationally focused on these issues and thinking about what responses can look like.

Within my organization, one of the things we think about is how do you create the right incentives for companies to do well without creating adverse incentives that may end up, unfortunately, impacting kids, teens and their ability to access important information online. I think sometimes there can be questions about what are the right levers to push, how do we incentivize responsible design practices without creating a culture where, for example, it might be hard for teenagers online to access information about reproductive care or information that might be useful for them when they are exploring their gender identity or their family identity.

There is a balance to be struck here, but on the overall, making sure that companies are being responsible in this space is incredibly important.

Senator BLUMENTHAL. Yes, I thank you for the work that you have done in this area, and particularly with my office, I know you have been very positive and constructive, so I thank you.

I am hoping, to cut right to the chase, that we can move forward on the Kids Online Safety Act, which provides for more transparency and at the same time provides for tools and safeguards for children and parents to make judgments that give them, in effect, control back over their lives, which many feel now they are losing, and avoid the unintended consequences that you just referenced, unintended consequences that may involve constraints on free expression or other goals. I think there is a balance to be struck here. I think that is our goal. That is what the legislation has attempted to do.

I do not know whether anyone, whether you have any comments on this question. Mr. Matheny.

Mr. MATHENY. Thanks, Senator. The one thing I would add is just that the potential for misuse is grounds for considering an appropriate regulatory framework, and I think reason to be especially cautious about open sourcing large language models that could be misused.

Senator BLUMENTHAL. One of the goals of the legislation is, in fact, greater transparency, and open sourcing certainly is a way of addressing that issue.

Thank you, Mr. Chairman.

Chairman PETERS. Thank you. Senator Hassan, you are recognized for your questions.

OPENING STATEMENT OF SENATOR HASSAN

Senator HASSAN. Thank you, Mr. Chair, and I want to thank the panel for being here and for your work. I want to start with a question to you, Dr. Metheny.

I am Chair of the Subcommittee on Emerging Threats (ETSO), a Subcommittee of this Homeland Security Committee, and I focus there, among other things, on the risks that artificial intelligence could pose to the cybersecurity of critical infrastructure like electric grids and hospitals.

You talked a little bit about some of the risks that AI poses, but can you expand a little bit, how does AI impact the cybersecurity threat landscape and are there opportunities to utilize AI to counter these threats?

Mr. MATHENY. Thanks, Senator, for the question. The application that has probably gained the most public attention of these large language models is generating language that we are familiar with, natural language, so creating an English poem or an English short story. What is getting less attention, but could be more impactful on security, is the application of these large language models to be used for software generation, code generation, and computer programming languages rather than a natural language. Some of these applications are already fairly sophisticated, and an increasing fraction of new software engineering is taking place with the use, or assistance, of large language models.

If this trend continues, it is quite possible that the offense of cyber capabilities that today are accessible only to state-level actor offensive cyber programs could be accessible to a much larger number of actors, simply by having access to tools that are able to generate software at scale and requiring much less technical sophistication to do so. Those could pose risks then to critical infrastructure and other networks that are sensitive.

Senator HASSAN. Thank you for that. Are there capacities that could help counter that, that AI gives us?

Mr. MATHENY. The same tools can also be used to scale up cyber defense, and I think this will be a cat-and-mouse race to figure out, are the applications on the defensive side keeping up with the applications on the offensive side.

I do not know the answer to that question. I think it will be a continuous competition between offense and defense.

But we need to make sure that our cybersecurity organizations are keeping up with the trends in these large language models as they are applied.

Senator HASSAN. OK. Thank you. Another question for you, Dr. Matheny. AI capabilities will offer new opportunities for the intelligence community, theoretically at least, to improve national security. Are there ways you believe that AI can improve intelligence analysis?

Mr. MATHENY. Yes. I think that the application of AI systems, particularly in open source data, where the volumes of data exceed our ability to analyze using manual methods, is one of the most important areas for intelligence. We could be making use of a much broader range of open source imagery, open source text in order to understand what is happening in the world much faster, and be able to share it with the world much more quickly.

As we are seeing from the war in Ukraine, when we are able to share open source information we are able to change the way the world understands what is happening in a part of the world that we do not have direct access to.

Senator HASSAN. Thank you. Then another question for you, Dr. Matheny. We know that government initiatives generally involve a number of different Federal agencies, and one of the things I am interested in is how can the Federal Government ensure that their agencies are coordinating with one another on AI research and deployment for potential joint projects or initiatives?

Mr. MATHENY. Thanks, Senator. One of the things that RAND has been working on over the years is how investments by one organization within the Federal Government, say one of our R&D organizations, can be more broadly shared across the government faster and how we can harmonize different efforts so that we are not duplicating efforts in one area of research, so that tool that are created by one agency can be leveraged by another, and so that standards that are used by one agency, say for AI being used for a particular application, can be harmonized with those in another agency.

I think there are great gains in efficiency.

One of the ways of harmonizing this would be through Federal procurement and ensure that we are using a consistent set of standards. Another would be through agencies like the National Institute of Standards and Technology, that have a key role to play in creating test frameworks and testbeds where we can robustly evaluate the performance of these AI systems.

Senator HASSAN. Thank you.

Now a question for Professor V, as I will call you, and Ms. Givens. This is a question for both of you. There is growing concern among workers in many industries that AI could fundamentally change the nature of work in unpredictable ways. You have touched on this a little bit, but do you have recommendations for how the Federal Government should be addressing challenges that companies and employees face from the use of AI in the workplace? Dr. V, I will start with you.

Mr. Venkatasubramanian. Thank you for the question, Senator. I think there are two parts to this, to helping workers deal with displacement due to AI. One, of course, is training and skilling, and the Federal Government can invest effort and research into helping workers train for our science, technology, engineering and mathematics (STEM)-enabled world. I think the Federal Government is doing that, and we can do definitely a lot more on that.

I think it is even more important that we make sure that that training and that access to those skills is widely distributed and not just to those who have access to those already. That is one thing.

I think another component of this is when we talk about worker displacement due to AI. I fundamentally believe it is because of overpromising on the part of AI systems, that tends to not play out when these systems are deployed.

Systems are presented as being able to replace because of efficiencies, workers, but in fact they cause more problems than they deserve, and it is precisely because there is not governance, there is not the supervision, there is not the human supervision around these systems.

I would argue that rather than thinking about workers displaced by AI, if we put proper governance and structures in place we will

need more jobs for workers, in fact, to make sure that these systems, that are supposed to assist them, are not replacing them and doing it badly.

Senator HASSAN. Thank you. Ms. Givens.

Ms. GIVENS. One of the questions is to think not just about displacement but if we are striving for a goal of workers working alongside AI systems, what does that interaction actually look like? We are seeing this play out now. You can think about fulfillment centers, for example, where workers are actually tasked with extreme specificity to every motion that they take, in the name of efficiency. There are business reasons for doing that, but there are also very real human impacts on the workers who are micromanaged at that level and live in a far more surveilled environment than they did before. Delivery van drivers, there are many other examples of this.

There we need to think about things like workplace health and safety. The Occupational Safety and Health (OSHA) has a role to play. The Department of Labor (DOL) has a role to play. We need to think about enforcement, both of existing laws and how we create a movement for employers to understand what responsible practices look like and for workers to know and understand their rights.

Senator HASSAN. Thank you very much. Thank you, Mr. Chair.

Chairman PETERS. Thank you, Senator Hassan.

Senator Padilla, you are recognized for your questions.

OPENING STATEMENT OF SENATOR PADILLA

Senator PADILLA. Excited about the opportunities that advances in technology will offer to society. But as this conversation has already shown, with every disruptive technology there are risks that demand mitigation. For example, automated decisionmaking systems and tools risk actually exacerbating the many existing inequities in our society, and that actually leads me to my first question.

It is clear that investments in AI research and education have not been distributed equally across the nation's researchers and innovators. Racial and gender diversity in AI and computer science programs are severely lacking. This lack of diversity among students gives rise to the corresponding lack of diversity in the workforce. A lack of diversity in the workforce then contributes to the development of AI tools and approaches that either do not account for or actively perpetuate systemic bias and limits the breadth of ideas incorporated into AI innovation.

My question is for Dr. Venkatasubramanian. As an educator, how can we ensure our AI and computer science students and workforce reflect the diversity of our nation?

Mr. VENKATASUBRAMANIAN. Thank you, Senator, for that question. This is an issue that concerns me greatly, as you might imagine, as an educator. I see the students who come to me who are concerned about this, and more often than not the students who are most concerned about these issues are students who truly reflect the broad diversity in this country, which is in one way a very good thing, but it also shows where the gaps in our ability to deliver STEM education effectively to our population is.

I would say that in my experience when students are able to see themselves in the work that they do, and the topics they study, they are more engaged with it and they feel like technology, in this case, can speak to them. Thestudents who come to speak to me about concerns around bias and responsible AI come to me because they have literally said, "I finally see a place for myself in this tech ecosystem."

One of the reasons why I spend a lot of time talking about concerns about bias and inequities in technology isbecause it is only by speaking out loud about those issues and pointing to the ways in which we can use technology to mitigate those issues that we can actually bring in a population that feels like they are now being heard and that their concerns are being heard.

I view these as part of the same story, that by spending time recognizing the inequities of AI, by spending time recognizing the need to govern areas to take account of these inequities we are actually telling people, "We welcome you in this technology and in this technology-enabled world."

Senator PADILLA. Thank you. Ms. Givens, I would be remiss if I did not take the opportunity to ask the former Chief Intellectual Property Counsel for the Senate Judiciary Committee a question about intellectual property. AI is introducing novel questions about the extent of a creator's intellectual property rights, most notably in the world of copyrights. Do you have any advice for those of us on the Judiciary Committee as we enter this new era of internet protocol (IP) complexity?

Ms. GIVENS. I am afraid I do not have a solution for you on this incredibly complex issue, but I do think it is an area where much attention is needed. There are photographers and designers and artists out there who understandably are deeply worried about the erosion of their industry and the role that they can play with the creation of generative AI, and also that their work is being used to train those tools.

On the other hand, we have had a very long tradition of fair use principles, and uses for transformative works in the creative space. There has to be a healthy conversation around how we appreciate some of those concerns of creators without inhibiting what is, in itself, an expressive act, the creation of new and diverse and transformative works through these tools.

Senator PADILLA. Thank you. To be continued. Dr. Matheny, large language models are rapidly improving and generative AI can have many important and positive applications. However, as a former elections administrator, I want to share a specific concern that I have about the ease with which this technology could facilitate election disinformation campaigns. Generative AI could radically reduce the cost and time while increasing the impact of misinformation and disinformation and propaganda. Not only could someone make it seem like one of us on the dais said something that we did not say or endorsing something that we do not endorse, but also the ability of foreign actors to supercharge their efforts to interfere in our elections is absolutely clear.

Referencing back to the Judiciary Committee, we know from our law enforcement officials that it is actually domestic extremism and

white supremacy that pose the largest national security threats to the United States.

It is bad enough that Speaker McCarthy was willing to share with Tucker Carlson all the footage of January 6th, which is now being repackaged to make it seem like a whole different January 6, 2021, took place than what is reality.

That is using actual footage. Imagine AI-generated video and the power that it can have in reshaping people's perspectives and attempts to redefine the truth.

Doctor, in light of your testimony, how do you recommend that we prepare our elections infrastructure and political processes to address propaganda that is harder to detect?

Mr. MATHENY. Thanks for the question, Senator. For several years RAND has had a project on something we called "truth decay," which is the vulnerability of democracies to disinformation attacks and other attacks against norms of evidence used in policy debates. One concern that we have had for several years is that the application of AI to disinformation campaigns could, as you point out, radically reduce the costs and increase the scale and speed of text, and speech potentially, that is used in disinformation, in ways that are very difficult to distinguish from human-generated forms of text and speech.

I think one important area is in research on distinguishing generative model text and speech compared to ones that are authentic. First, how can we watermark the products of generative AI systems in ways that we can distinguish them, and second, for those systems that have not used watermarking, can we find other signatures that allow digital forensics to be able to distinguish that which is disinformation from that which is legitimate.

Senator PADILLA. Also to be continued. Thank you, Mr. Chairman.

Chairman PETERS. Thank you, Senator Padilla.

Senator Sinema, you will be recognized for your questions. The vote has been called. I am going to run to vote. If you could take the gavel while I vote and then come back, and then Senator Rosen, I will be back shortly.

OPENING STATEMENT OF SENATOR SINEMA

Senator Sinema. [presiding.] Sure. Thank you, Mr. Chairman. Thank you to our witnesses for being here today. AI has the potential to revolutionize Arizonans' lives in countless ways, both good and bad. As we continue to integrate AI into our society we must ensure that this technology is developed and deployed in an ethical, transparent, and responsible manner that safeguards our values, preserves our privacy, and protects our national security.

My first question is for Dr. Matheny. Generative AI is suddenly everywhere, including ChatGPT and deepfakes—those are the fake videos that make people appear to say or do things they did not actually say or do. I have some experience with that. The key to solving this challenge is transparency, and one of the most promising solutions is so called content provenance data. This allows digital creators to embed data in content that disclosures whether it is authentic, altered, or entirely synthetic.

Do you believe that increasing transparency around what content is original versus what is AI-generated should be a policy priority for policymakers, and if so, is promoting content provenance efforts one of the most promising ways to create that transparency?

Mr. MATHENY. Some work that RAND has done over the past few years has identified watermarking and other ways of asserting provenance for digital media as being an important countermeasure against deepfakes, other forms of generated media that could be malicious.

We also need to increase our ability to do forensics on media that may have been generated but does not leave as easy telltale signatures, either because the entities that generated that media have not participated in various kinds of regulatory efforts to introduce watermarking or provenance.

I think what is likely to be required are investments in each of these categories, some way of asserting provenance, some way of watermarking, and investments and research on forensics for those that do not participate in the other two.

Senator SINEMA. Thank you.

Ms. Givens, as Chair of this Committee's Government Operations Subcommittee I am committed to ensuring that the Federal Government serves as a role model for society when it comes to responsibly and ethically deploying AI. I also serve as the Chair of the Commerce Subcommittee that oversees NIST, which just released its first-ever AI Risk Management Framework.

What is your assessment of the Federal Government's current AI practices, particularly with respect to transparency, bias, accuracy, and effectiveness, and how can government better manage these risks when it deploys AI?

Ms. GIVENS. This Committee has taken some important steps to show the need for rigorous processes and how agencies think about their use, design, procurement of AI.

I think there is still quite a lot of room for growth. The AI Risk Management Framework released by NIST is an excellent starting point, but we really need to operationalize it. We need to make sure that it is useful for people in the sectors of applicability where they are working. NIST needs to keep up its work on measurement strategies and ways to actually identify bias and assess whether or not interventions are working and are appropriate.

Then one of the leading things that this Committee helped generate, and it is being bolstered by a number of Executive Orders, is the inventory of agency uses of AI and guidance coming from OMB, and those are still works in progress as far as I understand. One of the priorities, I think, needs to be expediting that work, for OMB to play its central coordinating role, helping guide acquisition and use principles, and then starting the cycle of agencies inventorying their uses and showing how they are going to comply with that guidance in a meaningful way.

Senator SINEMA. Thank you.

My next question is for Dr. Venkatasubramanian—I practiced that one—and Dr. Metheny.

I would like to continue on the topic of ethical AI but in the context of U.S.-China competition. As we compete against Beijing to win the AI race, America may lose if we focus solely on the size

of our datasets, since, frankly, China's authoritarian system lends itself to vacuuming up vast volumes of data with few privacy protections. In contrast, America's competitive advantage may be our values, if we can translate these values into developing AI that is transparent, efficient, and fair.

What advantages and disadvantages does our country face in the AI competition with China, and do you agree thatinstead of viewing our values as a liability in this competition America could and should view them as an asset?

Mr. VENKATASUBRAMANIAN. Thank you for the question, Senator. I completely agree with the idea that the United States has values that can be transmitted into the systems we build, and I would argue this is happening already, but unfortunately the United States is not leading on this. For example, in the European Union (EU), with the development of the AI Act and other legislation that is going to govern the use of technology, especially AI technology, there is an attempt to push forward on the kinds of responsible practices that I think have been, frankly, developed here in the United States but are now being used in Europe. I think the United States can take the innovative lead, on these practices and provide a model for, frankly, the rest of the world to follow in how we do AI that is innovative, as well as responsible, as well as ethical at the same time.

I think we should push forward on that, we should emphasize that, and we should prioritize investments in those directions by prioritizing it within Congress and within the Federal Government.

Mr. MATHENY. I think that the United States has a couple of asymmetric advantages compared to China in AI. The first is that we are a much more attractive destination for the world's computer scientists and engineers. The United States has only four percent of the global population.

China has only 18 percent. The other 78 percent is sort of up for grabs. The United States does a much better job of attracting scientists and engineers from overseas. Many of the scientists and engineers are attracted by our values, so I think those values are a deep part of our asymmetric advantage.

A second advantage that we have is our ability to work with allies and partners. The United States and China each are responsible for about 25 percent of global research and development spending. When you add the United States and its allies, and add China and its allies, China's percentage does not increase because it does not have alliances with strong technological powers. The United States increases from 25 percent to about 65 percent.

Again, this is a place where having friends who are attracted to our values, who share our commitment to privacy, democratic governance, and the rule of law works to our advantage.

Senator SINEMA. Thank you. Senator Rosen.

OPENING STATEMENT OF SENATOR ROSEN

Senator Rosen. Thank you, Senator Sinema, and thank you to the witnesses for testifying today. I want to really speak a lot about skilled workforce because it is challenging across all platforms, as we see. Everyone who comes to talk to me is challenged with find-

ing a skilled workforce, and our Federal agencies and the digital Workforce, no different.

The National Security Commission on Artificial Intelligence does warn, and I am going to quote here, "The human talent deficit is the government's most conspicuous AI deficit and the single greatest inhibitor to buying, building, and fielding AI-enabled technologies for national security purposes."

The government, of course, we cannot compete with private sector salaries. We suffer from recruitment and retention issues, and the sustained AI talent shortage at government agencies, everyone would argue, could undermine our competitiveness.

Dr. Matheny, what are the specific ways you think the Federal agencies can really work to improve and expand that AI talent pipeline, and how might academic partnerships and initiatives be leveraged right now, public-private sector, to fill some of these gaps perhaps?

Mr. MATHENY. Thanks so much for the question, Senator.

I think that one of our most important levers or tools like the Intergovernmental Personnel Act, which allows the Federal Government to leverage expertise that is in academia and that is in other parts of the private sector, to bring in technical experts for short-term appointments, where they can serve as subject matter experts within Federal agencies.

I think we also have roles, like special government experts, that allow those in the private sector to maintain their positions in the private sector while they still advise government.

We certainly need to buildup the expertise within our own Federal workforce, but we also need to find more agile ways of leveraging the expertise that is distributed throughout the private sector, and those are two, I think, of our most important authorities to do so.

Senator ROSEN. Thank you. Professor V, I am going to turn to you because what kind of research and development investments should we be making to do just this, to uphill, reskill, or some might say right-skill the folks that are out there that do want to work, giving them an onramp to these jobs that can continue to grow?

Mr. VENKATASUBRAMANIAN. Thank you, Senator. As I mentioned earlier in response to Senator Padilla, one of the reasons that animates students from across the spectrum to work in technology, especially those who have not been seen by technology or are not being represented by technology earlier, is a desire to do something in the public good, to do something to improve the way all of us get the benefits of technology.

I feel like the Federal Government is a place where a lot of these students, university students, come and say they want to work in the Federal Government. They do not want to work in the private sector because they want to do some good. I think that is where the Federal Government has a comparative advantage over the private sector, becausethe Federal Government can articulate a value of public good in working with technology. I think the Federal Government should advertise that, should focus on developing technology to help bring it to all in a responsible and ethical manner.

I think Congress should continue its work to grow Federal expertise through training and skilling programs in the Federal Government. I think Congress should bring back the Office of Technology Assessment to help Members of Congress, the legislature, get more expertise on these topics as well.

Senator ROSEN. I could not agree more, as a former software developer, and so I am going to continue on this vein as we think about AI, the application that we use it for, cybersecurity, that can help us in these hunt forward operations, highlighting, or flagging, if you will, things for then humans to discern what seems right. So AI technology is rapidly evolving, and like I said, we really have to work on this. The National Cybersecurity Commission calls for more AI funding for AI-enabled cyber defenses.

Again, Professor V, how do you think we can enable and use AI to detect malware, pattern recognition, the things that computers are really good at, on the defensive side, and how can we use that to harden our security against cyber threats?

Mr. VENKATASUBRAMANIAN. Senator Rosen, I think I will defer that question to Mr. Matheny here. He has much more expertise than I do on the national security side.

Mr. MATHENY. You are too kind. I am worried about the long run arms race between offense and defense on cyber. I think both sides are amplified in their abilities by applications of different kinds of AI approaches.

On the defense side, as you mentioned, pattern recognition for looking for network activity that could suggest that there is an attack in progress. Most attacks are discovered weeks after. It would be nice if we detected them while they were happening so that we could do something about them. I do think that AI offers some applications in this area and there are active projects at the Intelligence Advanced Research Projects Activity (IARPA) and Defense Advanced Research Projects Agency (DARPA) to apply AI to cyber defense.

On the offensive side, I think one concern is that we are going to see increased levels of sophistication among relatively moderately skilled programmers in developing code much more quickly that can be used offensively.

I think the main thing here is for Federal agencies to be aware of how AI is being applied both offensively and defensively so that we are not surprised.

Senator ROSEN. Yes, I think you are right about that.

I am going to continue in this vein about this national strategy because you spoke earlier about the EU publishing their coordinated plan on AI, and they are encouraging each of its member States to develop their own national strategies. Of course, last week the White House released our national cybersecurity strategy.

What do you think would be the potential value for the U.S. national artificial intelligence strategy, more broadly, and how can interagency collaboration on AI be improved so we can detect and respond to threats more rapidly?

Mr. MATHENY. First I think all agencies would benefit from being able to draw in greater expertise, and that need not just mean full-time employees. It can mean advisors, consultants. Second is hav-

ing a common frameworkfor AI standards that all Federal agencies can leverage.

Here I think there is a key role for NIST to serve in developing uniform guidance for standards, ensuring that we also participate in international standards like ISO, SC 42.

Then third, I think shared Federal procurement rules that allow agencies to be developing tools that are built toward common standards with a common test framework.

Senator ROSEN. Speaking as a former software developer, the word "common framework" is music to my ears, so I am just going to leave it at that.

Mr. Chairman, I yield back.

Chairman PETERS [presiding.] Thank you. Thank you, Senator Rosen.

Dr. Matheny, you have extensive experience investigating threats posed by AI and national security, which is why it is so wonderful to have you here today. You have also written in support of export bans on the Chinese government. Could you tell us more about the threats that AI poses in the hands of the Chinese government and its State-sponsored companies and why bans may be appropriate to look at?

Mr. MATHENY. Thanks, Mr. Chairman. One of the things that I worry about, and I am a bit of a Debbie Downer on this, is that AI can be used to accelerate the development of other technologies. We are seeing early forms of this, where tools like AlphaFold were used to solve a very hard problem in biochemistry, the protein folding problem.

The upside potential of this is enormous. We can imagine this being applied to medicine in a variety of beneficial ways. It can also be used, though, to develop novel pathogens, and States that have historically not hadas many taboos as democracies around the use of technologies such as biotechnology for malicious use, I worry deeplyabout how AI will be used to supercharge different research and development efforts.

The same goes for offensive cyber, and the same goes also for disinformation used both domestically within China's own population for human rights abuses, for surveillance applications in Xinjiang and elsewhere in China, and used to influence foreign populations.

Chairman PETERS. You talk about other uses, the dual use of this, and we know that AI has a great deal ofpotential to deal with diseases that we have been attempting to cure forever, diseases like cancer. But I am curious of your thoughts about AI systems being weaponized perhaps, to find biotoxins or chemical warfare agents. How concerned should we be about that?

Mr. MATHENY. Countries like China that have historically invested in biological weapons and that havedemonstrated an interest in ethnically targeted weapons greatly concern me. The use of AI for so-called genome-wideassociation studies to try to identify how one would ethnically target particular pathogens is one area of special concern. We know, from a variety of research efforts historically, that the most virulent or transmissible pathogens are not those that are found in nature but ones that can be constructed artificially. AI creates opportunity to enhance pathogens much more

quickly and perhaps in ways that deliver effects to specific populations that are vulnerable.

Chairman PETERS. Ms. Givens, you have talked about AI and privacy and how our privacy is in danger, and this actually picks up a little bit on this question about creating pathogens. Would you talk a little bit about the privacy risk associated with using AI in the context of biometric data? We are providing more biometric data in databases. What are some of the concerns that you have associated with that?

Ms. GIVENS. Absolutely. Biometric data is one of the most sensitive types of data we can have. If there is a data breach and my faceprint is taken—I am not changing my face any time soon and I do not have the capacity to do so—so this information is highly in need of protection.

That makes it challenging when we think about the use of biometric identifiers, for example, in the delivery of government services. An increasing focus in fraud detection, for example, uses face recognition technology, one-to-one matching. Of course, there are law enforcement uses that are underway in the United States as well. We really need to think long and hard about the security vulnerabilities that can be created through this technology.

In addition, there are real concerns about equity when these types of technologies are being used. When, for example, your ability to access government services is contingent on you being able to snap a good selfie on your phone, that can exclude a large number of people that do not have that technology on their phone. Government agencies need to think about responsible onramps, responsible transitions for others as well.

But the cybersecurity and privacy vulnerabilities are real, and that is why it is so important to come back tothis language we have been talking about around real procurement standards, real safeguards, to make sure that when the government is considering using this technology there is a weighing of pros and cons, and then making sure that risks are mitigated.

Chairman PETERS. Thank you.

Professor V, I have heard concerns about effective computing, which tries to discern someone's emotion from those facial expressions that Ms. Givens was just talking about. Could you tell the Committee more about effective computing and if you have concerns?

Mr. VENKATASUBRAMANIAN. Yes. Thank you for that.

The premise, or the stated premise of effective computing is that we can infer information about people's internal States, their emotions, their cognitive States, their affect, from external features, external features like facial recognition, external features like how they walk, what kind of microtargeted expressions on their face, wrinkles, frowns, and so on.

I have great concerns about this. The premise of effective computing is unfounded. It has no basis. AI systems cannot do this. They might claim they do but they cannot because there is no underlying science to back this up. There is no underlying science that says that you can, in fact, do this kind of inference of people's internal States from external features. It just does not work, and most claims are, pardon my expression, completely bogus.

Chairman PETERS. Professor V, based on your time at the Office of Science and Technology Policy and your contributions to the Blueprint for an AI Bill of Rights, could you paint a picture for us of what a truly accountable AI system would look like within a Federal agency?

Mr. VENKATASUBRAMANIAN. Yes. This Federal agency procures and wants to procure an AI system that would be used to impact people, it would start by consulting with advocates, community partners, and other stakeholders to ensure that any system it might want to procure truly benefits those being impacted, in an equitable manner.

The agency will lay out strict guidelines and specifications to make sure that only the specific task is being sold, and that the system is not being repurposed for other tasks as well. It will make sure that the procurement process incorporates information about testing and validation for the specific task, that the system, in fact, works, and that as appropriate, disparity mitigation has been performed and results of these disparity mitigations are presented to the agency before procurement. It would not hand over people's data to the vendor, and if necessary would only share data with the vendor in a very controlled environment, for development purposes only.

Any deployed algorithm, once the system is deployed, would be supervised by agency experts who have expertise in the domain of interest and can tell when the algorithm or the system might be generating inaccurate outputs. The system would be regularly re-evaluated on a standard, on a cadence, to make sure data shifts have not affected its behavior. The vendor would need to provide tools to explain the algorithm's behavior.

I think an agency that is doing deployment of accountable AI well would be doing all of these things.

Chairman PETERS. Professor, if Congress were to requirement all the practices that you mentioned, what government body do you think would be best suited to hold agencies accountable?

Mr. VENKATASUBRAMANIAN. I think it is helpful to maybe distinguish between private sector use cases and government use cases. For private sector use cases, the FTC and its new Office of Technology would be perhaps the best place to do this, and should be given the resources to do this kind of work. For government uses, using the National AI Office that Congress had created, and OMB would probably be the best place to have high-level guidance and supervision of these systems.

Chairman PETERS. Is there an example of an agency now that is using AI effectively and responsibly, in your opinion?

Mr. VENKATASUBRAMANIAN. The Department of Health and Human Services (HHS) has done an excellent job complying with congressional mandates around the inventory of AI, for example, and around executive orders around AI. They are being very careful, for example, in their Updates Rule 1557 and the development of guidelines together with the Food and Drug Administration (FDA) around the use of AI in diagnostics, and that is one agency I would definitely hold up as doing a good job in this space.

Chairman PETERS. Great. Dr. Metheny, this Committee has focused on laying some of the groundwork for responsible agency use

and acquisition of AI. In our legislation we require standards and safeguards for acquiring and deploying these technologies and ensuring that the Federal workforce is up to the task to do that.

Can you elaborate on what else we could be doing to make sure that government procures and uses AI effectively and responsibly?

Mr. MATHENY. Thank you, Mr. Chairman. I think the U.S. Government has a fair amount of purchasing power that it can leverage to require that procured technologies meet certain standards of safety, reliability, robustness, and those standards could be verified in compliance through a third-party audit. I think that is one important lever that the Federal Government has. It will still not be the primary purchaser, but the private sector, in order to comply with such standards, it would simply make business sense for them to ensure that their systems, on the whole, are compliant.

A second key area, I think, is ensuring that democracies—the United States, its allies, and partners ensure that the international standards for AI systems are ones that support democratic norms around privacy and self-determination. We have the opportunity, through the international standards processes such as SC 42 that I mentioned earlier, to make those standards be ones that are privacy preserving, that are compatible with encryption, for example, and I think that is an opportunity we should seize.

Chairman PETERS. Thank you. The last question before we wrap up this hearing I am going to pose to each of you.

I will start with you, Ms. Givens, and then we will just work down the dais there.

We have heard commentators and academics have warned about the risk of human-like artificial intelligence, or artificial general intelligence, and those tend to be a lot of apocalyptic, scary stories that people talk about. But my question to each of you is, what are the risks that artificial general intelligence pose, and realistically, how likely is that actually in the near future? What is your assessment of how fast this is going and when we may beconfronted with some of those even more challenging questions and issues?

I will start with you, Ms. Givens.

Ms. GIVENS. I will leave to some of my more technical colleagues to do the likelihood question. I never want to make a prediction on a congressional panel. But I will say that when we are talking about such sophisticated technology, it raises many of the issues that we are already facing now, but simply supercharged, which is why we have to get the fundamentals in front of us correct now. When we are thinking about, for example, rules-based systems and controls, we already have a hard enough time thinking about how to respond to machine learning models now. When we think about these advanced systems, the notion that those are going to evolve rapidly over time makes it even harder to contemplate.

We have to address these questions of competency, of responsible design practices from the beginning, and we have to get our fundamentals right now, in the opportunity before us immediately, the ways in which AI is harming people in their daily lives right now, and the lack of ability for government agencies right now to be able to meaningfully respond to it, for us to even begin to think about how we tackle the next generation of issues.

Chairman PETERS. I think it is an important point.

The technology we know, that we heard from the experts here, is advancing very rapidly. In the past we have tended to look at technology as it is developed and just be excited about the promise of it. It gets developed and then we start seeing some adverse consequences, and then we look at regulation or other types of ways of dealing with it.

In this case this is moving so fast that I am concerned that if it gets way ahead of us that we cannot use the model of the past, where we see how things work out and then we address it. We really have to be thinking ahead, thinking a few steps ahead, which is why I am asking this question about the probability of even more powerful systems.

Professor V, that is in your wheelhouse.

Mr. VENKATASUBRAMANIAN. Yes. People ask me what keeps me up at night. AGI does not keep me up at night.

The reason why it does not is because, as Ms. Givens mentioned, the problems we are likely to face with the apocalyptic visions of AGI are the same problems we are already facing right now with the systems that are alreadyin play.

I worry about people being sent to jail because of an error in an ML system. Whether you use some fancy AGI to do the same thing, it is the same problem, and we are seeing this problem right now.

I think that the Committee's time is well spent pondering the harms that we are facing right now from these systems, and I would say, again, it is hard to predict. I am a computer scientist so maybe I should predict. But I would say that my bet is that the harms we are going to see as these more powerful systems come online, even with ChatGPT, are no different from the harms we are seeing right now. If we focus our efforts and our energies on governance and regulation and guardrails to address the harms we are seeing right now, they will be able to adjust as the technology improves. I am not worried that what we have put in place today will be out of date or out of sync with the new tech. The new tech is like the old tech, just supercharged.

Chairman PETERS. Thank you. Dr. Matheny, you will have the last word.

Mr. MATHENY. As is typically my last words, I do not know, and I think it is a really hard question. I think whether or not artificial general intelligence proves to be nearer than thought or farther than thought, I think there are things that we can do today that are important in either case, including regulatory frameworks that include standards with third-party tests or audits. The governance of our hardware supply chain so that we understand where large amounts of computing is going, and we prevent large amounts of computing from going to places that do not have the same ethical standards that we and other democracies have. Increasing the overall level of awareness and capability within the policy community, as you are doing today.

Chairman PETERS. Great. Thank you.

I would like to thank our witnesses for joining us today, and certainly I am grateful for your contributions to this very important discussion. As you heard at the outset, this is not the end. We are going to have more hearings on this and continue to dig deeper into the subject matter and look forward to working with you on that.

We know that today, as has been pretty clearly outlined, that AI systems can write like humans, they can assess business outlooks for companies, and they can even, hopefully, help us cure cancer at some point in the future.

As we have heard, however, these new developments certainly bring new risks, and without responsible designs, the use of AI can be devastating and discriminatory. Biased AI systems can unfairly deny people job opportunities and open users to legal liability. AI can supercharge the privacy risks posed by biometric data collection.

We also have heard that advancements in AI pose new challenges for our global competitiveness and national security. China is challenging the United States for leadership in AI innovation, and both China and Russia are developing military applications for AI as well. AI developments can create entirely new types of cyber and biological threats, and we must prepare for this new AI—enhanced world.

As we have heard today, recent advancements in computing research and data collection and processing power means that now is the moment to act on artificial intelligence.

As Chairman of the Committee I am going to work to ensure the United States continues to lead on AI, and we can be leaders in both AI research and production and in responsible AI design. They are not mutually exclusive. We can do all of the above, and we must. Your testimony here today will help inform the Committee's future legislative activities and oversight actions on that issue, and we look forward to being continually engaged with each and every one of you.

The record for this hearing will remain open for 15 days, until 5 p.m. on March 23, 2023, for the submission of statements and questions for the record.

This hearing is now adjourned.

[Whereupon, at 11:30 a.m., the hearing was adjourned.]

APPENDIX

Chairman Peters Opening Statement As Prepared for Delivery
Full Committee Hearing: Artificial Intelligence: Risks and Opportunities
March 8, 2023

Today's hearing will discuss both the potential risks as well as the opportunities associated with artificial intelligence, examining how artificial intelligence affects our nation's competitiveness on a global stage, and discuss ways to ensure that these technologies are used both safely and responsibly.

The adoption of artificial intelligence in government, industry, and civil society, has led to the rapid growth of advanced technology in virtually every sector, transforming millions of Americans lives, all across our country.

From the development of lifesaving drugs and advanced manufacturing to helping businesses and governments better serve the public, to self-driving vehicles that will improve mobility and make our roads safer, artificial intelligence certainly holds great promise.

But this rapidly-evolving technology also presents potential risks that could impact our safety, our privacy, and our economic and national security. We must ensure that the use of this technology becomes more widespread, we have to make sure that they're also the right safeguards in place to ensure it is being used appropriately.

One of the greatest challenges presented by artificial intelligence is the lack of transparency and accountability in how algorithms reach their results. Often, not even the scientists and the engineers who design the AI models fully understand how they arrive at the outputs that they produce. This lack of visibility into how AI systems make decisions creates challenges for building public trust in their use.

AI models can also produce biased results that can have unintended, but harmful consequences for the people interacting with those systems.

Some AI models, whether because of the data sets they are trained on or the way in which the algorithm is applied are at risk of generating outputs that discriminate on the basis of race, sex, age, or disability.

Whether these systems are being used in criminal justice, college admissions, or even determining eligibility for a home loan, biased decisions and the lack of transparency surrounding them, can lead to adverse outcomes for people who may not be even aware that AI has played a role in the decision-making process. Building more transparency and accountability into these systems will help prevent any kind of bias that could undermine the utility of AI.

And while many government organizations and businesses are working to build AI systems that enhance our daily lives, we must be open-eyed about the risks presented by bad actors and adversaries who may use AI to intentionally cause harm, or undermine our national interests.

Generative artificial intelligence like ChatGPT or deepfakes can be used to create convincing, but false information that can distort reality, undermine public trust, and even be used to cause widespread panic and fear in a worst-case scenario.

The risks from this kind of improper use also extend beyond our borders.

Adversaries like the Chinese government are racing to be the world leaders in these technologies and to harness the economic advantages that dominance in artificial intelligence will certainly create. The United States must be at the forefront of developing our own AI systems and training people how to use them appropriately, to protect our global economic competitiveness.

If we do not, not only are we at risk of American entities having to purchase these mature technologies from an economic competitor like the Chinese government, they will be tools with little accountability that have been developed by an adversary that does not share our core American values, a serious national security risk.

Finally, artificial intelligence will have a significant impact on the future of work. There is no question that AI systems have the potential to disrupt the workplace as we currently know it.

And that's why it is essential as the United States develops these technologies, we are also developing a workforce that is ready to work alongside them. We must address concerns that AI tools could replace human workers and instead focus on how they can assist humans and enhance the workplace.

Our goal in today's hearing is to examine these types of risks and challenges and discuss what steps Congress should take to ensure that we are able to harness these benefits and opportunities with this technology.

That includes ensuring that these technologies are used appropriately, and to protect the civil rights and civil liberties of all Americans.

Last Congress, I passed bipartisan laws that took steps to ensure the appropriate use of artificial intelligence by government, including through procurement safeguards and by boosting the knowledge of the acquisition workforce, to ensure they are properly trained to understand the risks and capabilities of these technologies.

And I look forward to building on those efforts this Congress, and working alongside my colleagues on the Committee to support the development of AI technologies, and ensure that they are being used both appropriately and effectively.

I hope that today's discussion will be the first of several on this important topic. And I'm pleased to have our panel of witnesses with us today who are experts in the field of artificial intelligence and who can discuss the adoption of these systems and the broader impacts on industry, civil society, and government.

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,

CENTER FOR DEMOCRACY & TECHNOLOGY

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-tef7m7az3rgitacauazyw3u224-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).

The Senate Homeland Security and Government Affairs Committee Hearing on AI: Risks and Opportunities

Mar 8, 2023

Remarks delivered by Suresh Venkatasubramanian.

Chairman Peters, Ranking Member Paul, and members of the Homeland Security and Government Affairs Committee. I thank you for inviting me to testify at this important hearing on the risks and opportunities of AI. I'm a professor of computer science and director of the Center for Technological Responsibility at Brown University. I recently completed a stint as a White House tech policy advisor in the Biden Administration and included in my portfolio was developing the recently released Blueprint for an AI Bill of Rights.[1] I have spent the last decade studying and researching the impact of automated systems (and AI) on people's rights, opportunities, and ability to access services. I've also spent time working with civil society groups and advising state and local governments on sound approaches to governing the use of technology that impacts people's lives.

What is AI?

We are here today to talk about AI – artificial intelligence. As an academic discipline, AI seeks to design automated systems that can sense, interact, reason, and behave in the way that humans do, and in some cases even surpass us.

We learn from the data we receive. And thus, one sub-area of AI that is dominant right now, fueled by the collection of vast amounts of data, is *machine learning*[2] – the design of systems that can incorporate historical data into the predictions that they produce, and in some cases keep adapting as more data appears. Machine learning grew in part out of decades of work in statistics: this is important to bear in mind since many systems that say they are using AI are really using statistical techniques that were invented decades ago and that are now supercharged by data.

Virtually every sector of society is now touched by machine learning. Algorithms created via machine learning are used to incarcerate individuals before trial[3], decide what sentence they should get if convicted[4], and decide whether they should get parole, and under what conditions.[5] Algorithms created via machine learning are used to determine a detected

---

[1] https://www.whitehouse.gov/ostp/ai-bill-of-rights/
[2] Hal Daumé III. A course in machine learning. http://ciml.info/
[3] David G. Robinson and Logan Koepke. Civil Rights and Pretrial Risk Assessments. Upturn, Inc., Dec. 2019. https://www.upturn.org/static/files/Robinson-Koepke-Civil-Rights-Critical-Issue-Brief.pdf
[4] John Villasenor and Virginia Foggo. Algorithms and sentencing: what does due process require? Brookings Institute, Mar. 2019. https://www.brookings.edu/blog/techtank/2019/03/21/algorithms-and-sentencing-what-does-due-process-require/
[5] Casey et al., Using offender risk and needs assessment information at sentencing. Nat'l Center for State Courts, 2011. https://www.ncsc.org/__data/assets/pdf_file/0019/25174/rna-guide-final.pdf

sound could be a gunshot,[6] or whether a blurred partial picture of an individual matches a known suspect.[7] Algorithms are used to monitor children in school for risk of suicide;[8] they are used to predict learning outcomes, and likelihood of success in educational settings.[9]

Algorithms created via machine learning screen candidate resumes for employers, analyze the results of video interviews or online interactive tests, and provide "fit" scores when employers are making hiring decisions.[10]

Machine learning algorithms are used to determine whether applicants for benefits are legitimate or fraudulent, what kinds of benefits they are eligible for, and how much they should receive.[11] These same algorithms are used to assess whether children are at risk for neglect or abuse, and whether social workers should intervene in a family.[12] These algorithms decide whether individuals should get health care, and what kind of care.[13] They interpret the results of medical tests. They decide whether individuals should get insurance coverage, and what price they should pay for this coverage.[14] Algorithms decide whether a potential renter should be considered by a landlord,[15] and what price this tenant should pay.[16] They are used to estimate the market value for a house, and what mortgage rate an individual can be asked to pay.[17] Algorithms are used to decide whether someone is a good credit risk for a loan.[18]

---

[6] Ferguson et al. The Chicago police department's use of shotspotter technology. Office of the Inspector General, Chicago, Aug. 2021. https://igchicago.org/wp-content/uploads/2021/08/Chicago-Police-Departments-Use-of-ShotSpotter-Technology.pdf
[7] https://www.clearview.ai/law-enforcement
[8] https://www.gaggle.net/
[9] https://www.civitaslearning.com/
[10] Bogen and Rieke. Help Wanted: An examination of hiring algorithms, equity, and bias. Upturn, Dec 2018. https://apo.org.au/sites/default/files/resource-files/2018-12/apo-nid210071.pdf
[11] Angwin. The Seven-Year Struggle to Hold an Out-of-Control Algorithm to Account. The Markup, Oct. 2022. https://themarkup.org/newsletter/hello-world/the-seven-year-struggle-to-hold-an-out-of-control-algorithm-to-account
[12] Samant et al. Family Surveillance by Algorithm: The Rapidly Spreading Tools Few Have Heard Of. ACLU, Sep. 2021. https://www.aclu.org/news/womens-rights/family-surveillance-by-algorithm-the-rapidly-spreading-tools-few-have-heard-of
[13] Ziad Obermeyer, et al., Dissecting racial bias in an algorithm used to manage the health of populations, 366 Science (2019), https://www.science.org/doi/10.1126/science.aax2342.
[14] I. E. Kumar. Colorado DOI weighs in on how to prevent algorithmic discrimination in life insurance. Center for Tech Responsibility, Brown University, Mar 2023. https://cntr.substack.com/p/colorado-doi-weighs-in-on-how-to
[15] K. Waddell. How Tenant Screening Reports Make It Hard for People to Bounce Back From Tough Times. Consumer Reports, Mar 2021. https://www.consumerreports.org/algorithmic-bias/tenant-screening-reports-make-it-hard-to-bounce-back-from-tough-times-a2331058426/
[16] Vogell. Rent Going Up? One Company's Algorithm Could Be Why. ProPublica, Oct. 2022. https://www.propublica.org/article/yieldstar-rent-increase-realpage-rent
[17] Consumer Financial Protection Bureau. Consumer Financial Protection Bureau Outlines Options To Prevent Algorithmic Bias In Home Valuations. Feb. 2022. https://www.consumerfinance.gov/about-us/newsroom/cfpb-outlines-options-to-prevent-algorithmic-bias-in-home-valuations/
[18] I.E Kumar et al. Equalizing Credit Opportunity in Algorithms: Aligning Algorithmic Fairness Research with U.S. Fair Lending Regulation. Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22). https://doi.org/10.1145/3514094.3534154

The list goes on and on.

There are many ways to build such "learning" systems. One specific kind of system that has risen to great prominence, in part driven by the availability of cheap and powerful computing power, is deep learning.[19] Deep learning has proven to be most powerful when analyzing images, text, audio, or video. Deep learning algorithms are used in facial recognition systems, in systems that analyze brain scans for neurological disorders, in the cameras installed on cars with driver assist or some other form of autonomous driving, in systems that translate from one language to another, and in systems that convert speech to text and vice versa. This list has grown rapidly and will continue to grow as we develop the underlying technology.

A *transformer*[20] is a particular kind of deep learning system, and as the name suggests, learns how to transform inputs and generate new kinds of output. Transformers are most useful for generating new kinds of content, whether it be deepfakes, plausibly realistic video segments, and of course text dialogue systems like GPT3[21], ChatGPT[22], Bard[23], and many others. Transformers need to ingest extremely large amounts of data, and require huge compute power, to do what they do.

The Failures of AI

Whether the system being used is a standard machine learning system, or one using more specialized architectures like deep learning, or even a transformer, all these systems share some common features that are important for how we might govern them. These are not algorithms or computer programs like the software of the 80s and 90s, or even the 00s. They are "algorithms for making algorithms":[24] the distinctive feature of a machine learning system is that the output of the learning algorithm that is fed vast amounts of data *is itself an algorithm* that purports to solve the underlying problem, whether a prediction task, an image analysis, or a text-based interaction with a user.

As a consequence of the above, *we don't actually know for sure whether and how these algorithms work and why they produce the output that they do.* This might come as a surprise, given how much we hear every day about the amazing and miraculous successes of AI. And yet, we also hear every day about the failures of AI systems.

---

[19] The "deep" refers to a specific aspect of the design of these systems and is not a statement about the quality of the results produced.
[20] A. Vaswani et al. Attention is all you need. Advances in neural information processing systems 30 (2017). https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
[21] https://openai.com/blog/gpt-3-apps
[22] https://chat.openai.com/
[23] https://blog.google/technology/ai/bard-google-ai-search-updates/
[24] Venkatasubramanian. When an algorithm isn't. Medium, Oct 2015. https://medium.com/@geomblog/when-an-algorithm-isn-t-2b9fe01b9bb5

Princeton professor Arvind Narayanan has likened AI systems to Snake oil: "Much of what's being sold today as AI is snake oil: it does not, and cannot work".[25] Researchers Deb Raji, Lizzie Kumar, Aaron Horowitz, and Andrew Selbst have referred to this same problem as "The Fallacy of AI Functionality", asserting that "Deployed AI systems often do not work" and laying out a series of case studies illustrating the myriad, and different, ways in which AI systems fail.[26]

AI systems fail when the algorithms draw incorrect conclusions from data and penalize individuals subject to those conclusions. A company installed AI-powered cameras in its delivery vans to evaluate the road safety habits of its drivers, but the system incorrectly penalized drivers when other cars cut them off or when other events beyond their control took place on the road. As a result, drivers were incorrectly ineligible to receive a bonus.[27]

AI systems fail when they seek to make predictions based on faulty data: a system that tried to predict effectiveness of health interventions used historical data on the cost of health care that was racially biased and produced racially biased outcomes.[13] Another system ended up causing the IRS to audit Black taxpayers more often than other taxpayers, for no apparent reason.[28]

AI systems fail when they are built using data from one group of people, and then are applied to a different group of individuals. The National Disabled Law Students Association expressed concerns that individuals with disabilities were more likely to be flagged as potentially suspicious by remote proctoring AI systems because of their disability-specific access needs such as needing longer breaks or using screen readers or dictation software .[29]

AI systems fail when the results of one automated decision system are fed into another (or even the same one), causing any errors in the original system to be amplified. An algorithm used to deploy police was found to repeatedly send police to neighborhoods they regularly visit, even if those neighborhoods were not the ones with the highest crime rates. These

---

[25] A. Narayanan. How to recognize AI snake oil. Nov. 2019. https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf

[26] I. D. Raji et al. The Fallacy of AI Functionality. In 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). https://dl.acm.org/doi/fullHtml/10.1145/3531146.3533158

[27] Lauren Kaori Gurley. Amazon's AI Cameras Are Punishing Drivers for Mistakes They Didn't Make. Motherboard. Sep. 20, 2021. https://www.vice.com/en/article/88npjv/amazons-ai-cameras-are-punishing-drivers-for-mistakes-they-didnt-make

[28] Jim Tankersley. Black Americans Are Much More Likely to Face Tax Audits, Study Finds. New York Times, Jan. 31, 2023. https://www.nytimes.com/2023/01/31/us/politics/black-americans-irs-tax-audits.html

[29] See, e.g., National Disabled Law Students Association. Report on Concerns Regarding Online Administration of Bar Exams. Jul. 29, 2020. https://ndlsa.org/wp-content/uploads/2020/08/NDLSA_Online-Exam-Concerns-Report1.pdf; Lydia X. Z. Brown. How Automated Test Proctoring Software Discriminates Against Disabled Students. Center for Democracy and Technology. Nov. 16, 2020. https://cdt.org/insights/how-automated-test-proctoring-software-discriminates-against-disabled-students/

incorrect crime predictions were the result of a feedback loop generated from the reuse of data from previous arrests and algorithm predictions.[30]

AI systems fail when they are so opaque that errors in how they function cannot be detected. In one example, a system awarding benefits changed its criteria invisibly. Individuals were denied benefits due to data entry errors and other system flaws. These flaws were only revealed when an explanation of the system was demanded and produced.[31]

The truth is that AI systems are not magic, and nor are they, as some would have us believe, about to bring about the downfall of humanity. AI is technology, like so many others that have entered society before it. And like any other piece of "magical" technology – drugs, cars, planes – AI need guardrails so that we can be protected from the worst failures of the technology while still benefiting from the progress AI offers.

What we should be doing

Many proposals for guardrails exist. These include

- The Blueprint for an AI Bill of Rights[32] issued by the White House in October 2022, which lists five key principles that protect us when automated systems are deployed in ways that affect our rights, opportunities, and access to critical services. The Blueprint also provides a detailed set of expectations that systems should comply with in order to satisfy these principles;
- The AI Risk Management Framework[33] developed by the National Institute of Standards and Technology that will help those deploying and using AI systems to properly estimate risks associated with the use of the systems; and
- The AI accountability framework for Federal agencies and other entities[34] published by the General Accounting Office in 2021.

And Congress has already acted to provide some guidance, including passing

- The National AI Initiative Act and the AI in Government Act in the 116th Congress; and
- The AI Training Act and The Advancing American AI Act in the 117th Congress.

---

[30] Kristian Lum and William Isaac. To Predict and Serve? Significance. Vol. 13, No. 5, p. 14-19. Oct. 7, 2016. https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x; Aaron Sankin, Dhruv Mehrotra, Surya Mattu, and Annie Gilbertson. Crime Prediction Software Promised to Be Free of Biases. New Data Shows It Perpetuates Them. The Markup and Gizmodo. Dec. 2, 2021. https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them

[31] Jay Stanley. Pitfalls of Artificial Intelligence Decisionmaking Highlighted In Idaho ACLU Case. ACLU. Jun. 2, 2017. https://www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idaho-aclu-case

[32] https://www.whitehouse.gov/ostp/ai-bill-of-rights/

[33] https://www.nist.gov/itl/ai-risk-management-framework

[34] https://www.gao.gov/products/gao-21-519sp

But we need to do more if we want a society where we can enjoy the benefits of modern technology without so many of the harms. All the frameworks that I described have at their core a collection of ideas that should be the basis of legislation that places guardrails on the deployment of AI in society. These ideas are as follows:

- We should do rigorous and independent testing of automated systems to evaluate their safety, effectiveness, potential discriminatory outcomes, and other forms of impact. Evaluation should be performed before deployment, after deployment, and in an ongoing manner.
- There should be clear governance frameworks for any AI deployments that impact people, and there should be clear lines of responsibility and authority for overseeing these systems.
- Any deployment must come with a clear articulation of harms and risks in context, and a concrete focus on mitigation strategies.
- It should be very clear when algorithms are being used, and why individual decisions were made in the way they were, because without that none of the above is even possible.
- There should wherever reasonable be human alternative approaches to using automated systems, and ways for individuals to obtain human recourse when systems fail (because they will).[35]
- There should be clear and mandated reporting on all the above.

Some of these ideas have appeared in executive orders in both the Biden and Trump Administrations. In particular, the Biden Administration recently issued Executive Order 14091[36] that emphasizes a focus on equity in agencies when "designing, developing, acquiring, and using artificial intelligence", and asks agencies to remedy discrimination by "protecting the public from algorithmic discrimination".

Congress should enshrine these ideas in legislation and extend the scope of legislation not just to government uses of AI, but to private-sector uses of AI that have people-facing impact as well.

All the above examples of harms associated with the deployment of AI society were uncovered through civil advocacy, journalism, and *sociotechnical* research that brought scholars from technical disciplines, the social sciences, and the humanities together to study these "collisions" between technology and society. Such research is extremely

---

[35] This became a crucial issue recently. Individuals trying to obtain benefits from the government were required to use a third-party identity verification system. This system (based partially on facial recognition) failed to work (especially on individuals with darker skins) and there were no alternative pathways provided: in fact, people often had to wait for hours and hours on hold to reach a human operator because the system did not have appropriate means for human recourse. https://www.bloomberg.com/news/features/2022-01-20/cybersecurity-company-id-me-is-becoming-government-s-digital-gatekeeper

[36] https://www.federalregister.gov/documents/2023/02/22/2023-03779/further-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal

important and has been the most effective way to identify problems and propose concrete solutions, including all the ideas I mention above.

Congress should invest in innovative sociotechnical research that will continue to uncover and mitigate the harms that accrue as our "algorithmic society" expands.

Conclusion

I'm a computer scientist, and part of my work is to imagine technological futures. There's a future in which automated technology is an assistant: it enables human freedom, liberty, and flourishing. Where the technology we build is inclusive and helps us *all* achieve our dreams and maximize our potential.

But there's another future, in which we are at the mercy of technology, which the world is shaped by algorithms and we are forced to conform. In which those who have access to resources and power control that world and the rest of us are left behind.

I know which future I want to imagine and work towards. I believe that it is our job to lay down the rules of the road – the guardrails and protections – so that we can achieve that future. And I know we can do it if we try.

Thank you for giving me the opportunity to speak.

Testimony

Challenges to U.S. National Security and Competitiveness Posed by AI

Jason Matheny

CT-A2654-1
Testimony presented before the U.S. Senate Committee on Homeland Security and Governmental Affairs on March 8, 2023

For more information on this publication, visit www.rand.org/t/CTA2654-1.

Testimonies

RAND testimonies record testimony presented or submitted by RAND associates to federal, state, or local legislative committees; government-appointed commissions and panels; and private review and oversight bodies.

Published by the RAND Corporation, Santa Monica, Calif.

Challenges to U.S. National Security and Competitiveness Posed by AI

Testimony of Jason Matheny[1]
The RAND Corporation[2]

Before the Committee on Homeland Security and Governmental Affairs
United States Senate

March 8, 2023

Chairman Peters, Ranking Member Paul, and members of the committee: Good morning, and thank you for the opportunity to testify today. I'm the president and CEO of RAND, a nonprofit and nonpartisan research organization. Before RAND, I served in the White House National Security Council and Office of Science and Technology Policy, as a commissioner on the National Security Commission on Artificial Intelligence, as assistant director of national intelligence, and as director of the Intelligence Advanced Research Projects Activity, which develops advanced technologies for the U.S. intelligence community.

For the past 75 years, RAND has conducted research in support of U.S. national security, and we currently manage four federally funded research and development centers (FFRDCs) for the federal government: one for the Department of Homeland Security (DHS) and three for the Department of Defense. Today, I'll focus my comments on how artificial intelligence (AI) affects national security and U.S. competitiveness. Among a broad set of technologies, AI stands out for both its rate of progress and its scope of applications. AI holds the potential to broadly transform entire industries, including ones critical to our future economic competitiveness, such as medicine, manufacturing, and energy. Applications of AI also pose grave security challenges for which we are currently unprepared, including the development of novel cyber weapons,

[1] The opinions and conclusions expressed in this testimony are the author's alone and should not be interpreted as representing those of the RAND Corporation or any of the sponsors of its research.

[2] The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest. RAND's mission is enabled through its core values of quality and objectivity and its commitment to integrity and ethical behavior. RAND subjects its research publications to a robust and exacting quality-assurance process; avoids financial and other conflicts of interest through staff training, project screening, and a policy of mandatory disclosure; and pursues transparency through the open publication of research findings and recommendations, disclosure of the source of funding of published research, and policies to ensure intellectual independence. This testimony is not a research publication, but witnesses affiliated with RAND routinely draw on relevant research conducted in the organization.

large-scale disinformation attacks, and the design of advanced biological weapons. Threats from AI pose special challenges for national security for several reasons:

- The technologies are driven by commercial entities that are frequently outside our national security frameworks.
- The technologies are advancing quickly, typically outpacing policies and organizational reforms within government.
- Assessments of the technologies require expertise that is concentrated in the private sector and that has rarely been used for national security.
- The technologies lack conventional intelligence signatures that distinguish benign from malicious use, differentiate intentional from accidental misuse, or permit attribution with certainty.

The United States is currently the global leader in AI;[3] however, this may change as the People's Republic of China seeks to become the world's primary AI innovation center by 2030—an explicit goal of China's AI national strategy.[4] In addition, both China and Russia are pursuing militarized AI technologies,[5] intensifying the challenges I just outlined.

In response, I will highlight eight actions that national security organizations, including DHS, could take:

1. Ensure that DHS cybersecurity strategies and cyber Red Team activities track developments in AI that affect cyber defense and cyber offense.
2. With the National Institute of Standards and Technology, industry stakeholders, and U.S. allies and partners, ensure that international standards for AI prioritize privacy, security, and safety, so that the technologies are less prone to misuse by surveillance states.
3. Consider creating a regulatory framework for AI that is informed by an evaluation of risks and benefits of AI to U.S. national security, civil liberties, and competitiveness.
4. Identify the high-performance computing hardware used for AI as critical infrastructure that can be stolen or subverted. Consequently, consider requirements for tracking where high-performance computing hardware goes and what it is being used for.
5. Work with the intelligence community to significantly expand the collection and analysis of information on key foreign public- and private-sector actors in adversary states involved in AI, and create new partnerships and information-sharing agreements among federal, state, and local government agencies; the research community; and industry.
6. Leverage AI expertise in the private sector through short-term and part-time federal appointments and security clearances for leading private-sector AI experts who can advise the government on key technology developments.

---

[3] Although there are many ways to measure this, the Stanford Global AI Vibrancy Tool has consistently ranked the United States at the top. See Stanford University, "Global AI Vibrance Tool: Who's Leading the Global AI Race?" undated, https://aiindex.stanford.edu/vibrancy/.

[4] Graham Webster, Rogier Creemers, Elsa Kania, and Paul Triolo, "Full Translation: China's 'New Generation Artificial Intelligence Development Plan,'" DigiChina, August 1, 2017, https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

[5] Forrest E. Morgan, Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, and Derek Grossman, *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World*, RAND Corporation, RR-3139-AF, 2020, https://www.rand.org/pubs/research_reports/RR3139-1.html.

7. In federal purchases and development of AI systems, include requirements for security and safety measures that prevent AI systems from misbehaving due to accidents or adversaries. Also require socially beneficial techniques, such as privacy-preserving machine learning and watermarking to detect generated text and deepfakes.
8. Last, increase our investments in biosecurity and biodefense, given the potential applications of AI to design pathogens that are much more destructive than those found in nature.

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

AI and social policy

Dr. Jordan B. Peterson
March 06, 2023

Large Language Models and ChatGPT

Advanced Large Language Models such as ChatGPT have burst onto the scene with a vengeance in the last six months. These systems analyze the relationship between words, phrases and sentences by making reference to immense corpuses of written material, and essentially make sense of the world by extracting reliable patterns of repetition from human communication.

Thus, insofar as the world's structure, psychological, social and natural, is encoded in abstract verbal representation the LLM systems can model the world.

Systems such as ChatGPT are currently about as intelligent, by all appearances, as a low-average undergraduate at a decent American state university. I say that because ChatGPT recently completed the SAT and scored 1020. The average for the U California system undergraduate approximates 1300, which is three standard deviations above the ChatGPT 1020. SAT scores at the University of Kentucky are more in the ChatGPT range.

A SAT score of 1020 (ChatGPT's score) is equivalent to an IQ of 110, which is about 2/3 of a standard deviation above the population average. This makes ChatGPT more intelligent than 75% of people.

We will have LLM systems that are much more intelligent than average in short order. Users will be able to order verbal material (essays, etc.) written at a given IQ level within months or short years.

The significance of all this should not be underestimated. We now have AI systems capable of engaging in genuine conversation, able to write, able to produce computer code, and able to "think." And they will be much smarter very soon.

Rights to the Extended Digital Self

For centuries we were all simple enough so that our names sufficed to identify us, and to enable others to do so. That has changed dramatically in the last thirty years. We are all beset by the necessity to employ an ever-shifting plethora of usernames and passwords. Why? Because we are so complex in our extended digital selves that our old names are no longer stable or sophisticated enough to identity the many beings we have become in the virtual world.

We have a core biological and incorporated identity, and that is extended outward into intimate relationship, family, neighbourhood, community, city, state and country. The old

naming convention (Christian name, surname) served to identify each of us in that nexus. Online, however, things are very different.

Our digital identity is composed of the tools we use (the apps, programs, services, websites, etc., that we choose voluntarily to employ) as well as the record of our virtual behavior (our browsing patterns, our purchases, the records of our travel, the written communications and images we issue forward on platforms such as Instagram, Facebook and, more ominously, TikTok, which essentially operates under the control of the Chinese Communist Party). That extended digital self has very few rights, as our legal structure has not been able to adapt itself to the immense changes on the virtual front.

At present, the extended digital self--which is, increasingly, as much or even more of the self than the traditional embodied and socially-situated self—is owned not by the individuals who are extended in that manner but by the corporate entities that track, manipulate and sell the data comprising exactly that self. Fortunately—for now—most such entities are at least bound to the interests of those whom they track and trade by the desire to make money: the credit card company that tracks what you are doing so that its partners can sell you consumer goods more efficiently is at least trying to provide you with something you might hypothetically and voluntarily desire. In that manner, the "greed" of the corporation serves at least the whims of the individual, if nothing more.

When that capacity to track is turned to more explicitly ideological goals (think Environment, Social and Governance, for example, or Diversity, Inclusivity and Equity) then a cardinal danger emerges: the involuntary subjugation of the individual to the motivated ends of the propagandists pushing their political agenda. The logical extension of such danger (and the most likely outcome, in my estimation) is the duplication in the West of something approximating the utter catastrophe of the so-called Social Credit System in China. The CCP has exercised totalitarian control over the economic behavior of its citizens through the implementation of a centralized digital currency, which puts all of what was purely private in the hands of self-interested bureaucrats. Everything is tracked and controlled; the government can, with a stroke of the pen, seize the economic resources of any given individual or group (something that happened very ominously in Canada in the case of the Trucker's Convoy); can set expiry dates on money; can differentially tax any purchase (to further ESG, DIE or climate catastrophist goals); can lock people arbitrarily out of the entire economic and social system. The means to do all of this are already in place: the fact that the West unthinkingly and instantly mimicked the totalitarian response to the so-called Covid epidemic indicates precisely just how willing we are to lock down the population in the name of the greater good here in what were once free countries.

It is very much possible, by the way, that the only alternative to centralized digital currency systems and the terrible danger they pose is digital currencies that are distributed. BitCoin is the primary exemplar on that front. Serious consideration may have to be given soon to adopting currency alternatives like BitCoin, which have in principle been placed permanently outside the purview and control of centralized authorities.

Developing AI capacities will radically extend the surveillance state. There are already systems in place that can identify individuals in any crowd exposed to the ever-present gaze of surveillance cameras. China, which is of course in the lead of such developments, has about 400 cameras watching every 1000 people. The AI-enhanced software behind such cameras can identify people from their face, even when masked; from their gait, even when faces are hidden from view. We are at a point where everyone can be tracked all the time. This means that the extended digital self mentioned previously will now also comprise every trip ever taken and all meetings with other people. The latter is particularly dangerous: the Chinese "social credit" system is already set up such that if higher-scoring "citizens" of that state consort with lower scoring citizens (those who have broken whatever rules of conduct the scores of absolutist tyrants can generate) their own social credit scores, which enable access to even the most basic needs of life, will decrease.

We could well be entering an era of authoritarian AI-mediated social shunning.

The use of such cameras should be banned. Machines should never be given the authority to ticket, try, punish, or limit the economic or practical activities of human beings. The digital self should be treated legally as the logical extension of the corporal and psychological self into the virtual space, and the data comprising that extended self granted the protection of the intrinsic rights that already pertain to the traditional self. The government should be very wary of any forays into the domain of digital currency, as the ability to track all purchases, which will absolutely and immediately be gamed and perverted, poses an almost incalculable temptation to fear-mongering tyrants who want to remake the world to forestall their pet catastrophic emergency and who want to accrue to themselves all power so they can move forward with the dread efficiency they so ardently desire.

Additional Dangers

We are on the cusp of the ability to produce photorealistic video and audio representations using text alone. In the next year, the AI wizards will produce intelligent systems that will be able to produce representations of any person doing anything that can be described—the so-called deepfakes. This will mean that the body images of beautiful people, women in particular, will be stolen for use in personalized digital pornography. This is already happening. This will also mean that the personae of powerful people will be duplicated. Imagine, for example, a photo-realistic representation of Joe Biden discussing the financial gain he might accrue by passing money to his military-industrial friends as a consequence of using deception and treachery to further the expansion of the war in Russia/Ukraine. Imagine that generated in the kind of shaky and unprofessional manner that a clandestinely-wielded iPhone might produce: the "secret" recording of a "secret" off-the-record conversation. Imagine that released on the eve of a critical election.

Then imagine that happening everywhere, on every issue, thousands of times. Imagine being entirely unable to determine, day to day, what communication from what person (photos, videos, audio recordings, writings) is real and what is false.

Then: imagine that now, and not in some future. That's where we're at.

Steps must be taken on the legal front to make false digital representations of living persons not only illegal but seriously illegal. The most appropriate current analog might be the case of kidnapping. To kidnap someone and then to force them to confess to something political or personal, for example, on broadcast networks is properly regarded as a crime comparable only to rape, torture or murder. The same basic conceptual framework should obtain in the digital realm. The creation of a deepfake must soon be treated as among the most serious of crimes. It is almost impossible to overstate what a danger this technology poses. The material and psychological well-being of the citizenry and the very integrity of the state is dependent on trust. The ability to produce deepfakes that are indistinguishable from recorded reality compromises all of that.

Conclusion

The development of AI systems as intelligent as we are is not some future possibility, but a current actuality. The melding of AI-mediated intelligent systems with our capacity for monitoring and surveillance prepares the way for a tyranny so comprehensive that we can barely imagine it.

We each have an extended digital self which is in many ways as or even more real than our traditional embodied identities. None of us have the rights to that extended self. That opens up the door to a tyranny the thoroughness of which we can hardly imagine (envision North Korea run by Silicon Valley).

DATA&
SOCIETY

March 23, 2023

The Senate Homeland Security and Governmental Affairs Committee
Submitted via email to michelle_benecke@hsgac.senate.gov

Response to March 8 Hearing on AI: Risks and Opportunities

Dear Chair Peters, Ranking Member Paul, and members of the Homeland Security and Governmental Affairs Committee:

Data & Society Research Institute ("Data & Society" or "D&S") is pleased to submit testimony to the Committee regarding its March 8 hearing on the risks and opportunities of artificial intelligence ("AI"). Our organization is an independent, nonprofit research institute studying the social implications of data-centric technologies and automation. We produce empirical research that challenges the power asymmetries created and amplified by technology in society, including emerging technologies like AI.

Data & Society is pleased to see that the Committee is carefully examining the risks presented by AI. As AI systems become increasingly pervasive in many facets of society—impacting people's access to jobs, housing, credit, and more—this Committee and Congress should support AI research that is as much anchored to **sociotechnical research** as it is to technical knowledge.

What is Sociotechnical Research?

Sociotechnical research studies technologies in their social, political, economic, and cultural contexts. It recognizes that successful technological deployment is a result of integration with often-invisible human, material, and cultural infrastructures, and it seeks to make those infrastructures visible to better assess the use of technologies in new arenas.

A sociotechnical approach questions the expectation that technology's impact can be predicted from its technical properties alone. Moreover, it assumes that technical transformations to an existing process or function will have moral and political repercussions. Such an approach considers not simply how to best use a technology, but fundamentally *whether a given technology is appropriate in the first place*, and where it fits alongside non-technical means.

What are the methods of sociotechnical research?

- Sociotechnical research **draws from observations** gathered through quantitative, qualitative, or mixed methods approaches. It employs interview-based or ethnographic studies, computational analysis of logged data, sociological audits, case studies, and

DATA&
SOCIETY

historical analysis. Sociotechnical research may also propose theoretical framings that synthesize insights from observational studies or shape future studies.
- Methods of sociotechnical research **are inductive**, meaning they help to discover the unexpected when technology is deployed "in the real world."
- Methods of sociotechnical research capture **the viewpoint** of those who are impacted by a technology. These methods allow others to have a say in how technology is used and designed, and are a critical element in laying the groundwork for meaningful participation in AI governance.

Why is this critical for U.S. leadership in AI R&D?
Technical research absent broader engagement with experts on society, politics, economy, and culture is likely to reproduce patterns of incomplete, biased, and discriminatory solutions. Given the priority of an ecosystem built on trustworthy AI, the incorporation of sociotechnical research fully into R&D planning and spending is imperative. United States AI R&D should lead with an integrated approach that prioritizes both the social and the technical as elements of innovation and competitiveness in AI development.

American innovation must be directed towards AI solutions that perform successfully not only in highly controlled testbeds, but also in real-life use cases with social, political, cultural, and economic factors that will shape an AI system's impact in the world.

The United States must take a leading role to create an AI future that is both just and competitive. This is achievable with a balanced set of commitments across research and policymaking. The United States must lead AI innovation with an equity, safety, and trustworthiness framework, enacted through federal research commitments and funding that advance sociotechnical approaches. Recent AI initiatives, including the National Institute of Standards and Technology AI Risk Management Framework and the White House Office of Science and Technology Policy's Blueprint for an AI Bill of Rights, foreground sociotechnical research as a critical methodology to assess AI. Similarly, this Committee should assess the risks and opportunities of AI from a holistic viewpoint, centering sociotechnical research in Congressional legislation and oversight around AI, particularly within AI R&D priorities.

Thank you for the opportunity to provide written feedback for this hearing. We look forward to supporting the Committee as it continues to advance sound, just, and competitive AI governance.

Best,

Serena Oduro, Senior Policy Analyst

1212 New York Ave. NW
Suite 900
Washington, D.C. 20005
202-525-5717

Free Markets. Real Solutions.
www.rstreet.org

March 8, 2023

The Honorable Gary Peters
Chair
Homeland Security and Governmental
Affairs Committee
U.S. Senate
Washington, D.C. 20510

The Honorable Rand Paul
Ranking Member
Homeland Security and Governmental
Affairs Committee
U.S. Senate
Washington, D.C. 20510

Dear Chairman Peters, Ranking Member Paul and members of the Committee:

Thank you for your decision to hold a hearing on March 8, 2023 titled, "Artificial Intelligence: Risks and Opportunities". My name is Adam Thierer and I am a senior fellow at the R Street Institute. I also currently serve as a commissioner on the U.S. Chamber of Commerce's Commission on Artificial Intelligence Competitiveness, Inclusion, and Innovation, which will be releasing its final report tomorrow morning.[1]

It is essential that the United States be a leader in AI to ensure our continued global competitive standing and geopolitical security. The most important way to counter China, Europe and other nations attempting to overtake U.S. innovation on this front is to make sure we do not follow their lead in terms of heavy-handed control of digital systems. America's crucial advantage over other countries comes down to our uniquely agile and adaptive approach to technological governance.

As I noted in a recent piece for the R Street Institute:

> The European Union (EU) has implemented a wide variety of data collection mandates that have restricted innovation and competition across the continent. These regulatory burdens have left the EU with few homegrown information technology firms. As a result, the EU now mostly focuses on exporting its mandates globally…[2]

According to the Bureau of Economic Analysis, in 2021, the U.S. digital economy accounted for $3.7 trillion of gross output, $2.41 trillion of value added (or 10.3 percent of U.S. GDP), $1.24 trillion of compensation and 8 million jobs.[3] Globally, 18 of the world's top 25 digital tech

1212 New York Ave. NW
Suite 900
Washington, D.C. 20005
202-525-5717

companies by market capitalization are U.S.-based firms, and 46 of the top 100 firms with the most employees are U.S. companies.[4]

The American economic success story was driven by smart, bipartisan choices that Congress and the Clinton administration made in the 1990s. There are four key ingredients behind America's successful approach to digital innovation:

1. The first is freedom to innovate by default. Entrepreneurs were given a green light to experiment with bold new ideas without having to seek permission to innovate.
2. The second is world-class university programs and research labs. The United States is home to some of the world's leading technical educational programs that have produced much of the best talent in digital technology markets today.
3. The third factor is openness to global talent and investment. The United States opened its tech markets to skilled immigrants and global investors and they flocked here to enjoy the benefits of vibrant markets and our superior higher education institutions.
4. The fourth factor is the use of ongoing multi-stakeholder negotiations and flexible regulatory responses when concerns develop. The National Telecommunications and Information Administration (NTIA) and other agencies have brought together diverse stakeholders repeatedly to hammer out solutions to complicated technology problems.[5]

These ingredients are the secret sauce that have powered America's commanding lead in the internet and computing sectors. And now, they can help us lead the global AI race. The hard reality of AI governance is that it is going to be extremely difficult to establish any policy for algorithmic systems that is not quickly overtaken by fast-moving technological realities. There is no one-size-fits-all approach to AI that can preemptively plan for the challenges that we will face even a few months from now.

Government's role should be focused on helping to convene different stakeholders and working toward consensus on best practices on an ongoing basis.[6] In this regard, the National Institute of Standards and Technology (NIST) has taken important steps with its recently released *AI Risk Management Framework*.[7]

1212 New York Ave. NW
Suite 900
Washington, D.C. 20005
202-525-5717

Free Markets. Real Solutions.
www.rstreet.org

This NIST framework, which builds on previous multi-stakeholder efforts, is meant to help AI developers better understand how to identify and address various types of potential algorithmic risk. NIST notes it "is designed to address new risks as they emerge" instead of attempting to itemize them all in advance.[8] "This flexibility is particularly important where impacts are not easily foreseeable and applications are evolving," the agency explains.[9] Building on this, NIST and the NTIA can take the lead in extending their expertise in helping to convene ongoing multi-stakeholder efforts to bring diverse stakeholders to the table and hammer out consensus-driven best practices and solutions on the fly.

As this governance model for AI evolves, it should be guided by some key principles. Several of these recommendations are found in the U.S. Chamber of Commerce AI Commission report launching tomorrow.

First, AI governance should be risk-based and focus on system outcomes, instead of being preoccupied with system inputs or design. In other words, policy should concern itself more with actual algorithmic performance, not the underlying processes.[10] If policy is based on making AI perfectly transparent or explainable before anything launches, then innovation will suffer because of endless bureaucratic delays and paperwork compliance burdens.

Second, AI policy should utilize existing laws and remedies before adding new regulatory mandates. As noted, a vast array of laws and regulations already exist that can effectively govern algorithmic systems.

Third, AI policy should encourage the private sector to refine best practices and ethical guidelines continuously for algorithmic technologies. An extensive amount of work has already been done in this regard, but it will require constant vigilance and iteration to address emerging risks effectively.

Thank you for holding this hearing. I look forward to addressing your questions.

Sincerely,

1212 New York Ave. NW
Suite 900
Washington, D.C. 20005
202-525-5717

Free Markets. Real Solutions.
www.rstreet.org

/s/Adam Thierer
Senior Fellow
R Street Institute

---

[1] "Artificial Intelligence Commission: Preparing for the Future," U.S. Chamber of Commerce, last accessed March 3, 2023. https://www.uschamber.com/major-initiative/artificial-intelligence-commission.
[2] Adam Thierer, "Mapping the AI Policy Landscape Circa 2023: Seven Major Fault Lines," R Street Institute, Feb. 9, 2023. https://www.rstreet.org/commentary/mapping-the-ai-policy-landscape-circa-2023-seven-major-fault-lines.
[3] Tina Highfill and Christopher Surfield, "New and Revised Statistics of the U.S. Digital Economy, 2005–2021," Bureau of Economic Analysis, November 2022. https://www.bea.gov/system/files/2022-11/new-and-revised-statistics-of-the-us-digital-economy-2005-2021.pdf.
[4] "Largest tech companies by market cap," Companies Market Cap, last accessed March 4, 2023. https://companiesmarketcap.com/tech/largest-tech-companies-by-market-cap.
[5] Ryan Hagemann et al., "Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future," *Colorado Technology Law Journal* 17 (Feb. 5, 2018). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3118539.
[6] Lawrence E. Strickling and Jonah Force Hill, "Multi-stakeholder internet governance: successes and opportunities," *Journal of Cyber Policy* 2:3 (2017), pp. 298–99. https://www.tandfonline.com/doi/abs/10.1080/23738871.2017.1404619.
[7] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, U.S. Department of Commerce, January 2023. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.
[8] Ibid., p.4
[9] Ibid., p. 4.
[10] Daniel Castro, "Ten Principles for Regulation That Does Not Harm AI Innovation," Information Technology and Innovation Foundation, Feb. 8. 2023. https://itif.org/publications/2023/02/08/ten-principles-for-regulation-that-does-not-harm-ai-innovation.