

# ARTIFICIAL INTELLIGENCE IN GOVERNMENT

---

---

## HEARING

BEFORE THE

COMMITTEE ON  
HOMELAND SECURITY AND  
GOVERNMENTAL AFFAIRS  
UNITED STATES SENATE  
ONE HUNDRED EIGHTEENTH CONGRESS

FIRST SESSION

\_\_\_\_\_  
MAY 16, 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

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*

KENDAL B. TIGNER, *Minority Professional Staff Member*

LAURA W. KILBRIDE, *Chief Clerk*

ASHLEY A. GONZALEZ, *Hearing Clerk*

# CONTENTS

Opening statements:	Page
Senator Peters .....	1
Senator Paul .....	3
Senator Lankford .....	18
Senator Scott .....	20
Senator Hassan .....	23
Senator Rosen .....	25
Senator Padilla .....	28
Senator Ossoff .....	31
Prepared statements:	
Senator Peters .....	37
Senator Paul .....	39

## WITNESSES

TUESDAY, MAY 16, 2023

Lynne E. Parker, Ph.D., Associate Vice Chancellor and Director, AI for Tennessee Initiative, University of Tennessee .....	5
Taka Ariga, Chief Data Scientist, U.S. Government Accountability Office .....	7
Daniel E. Ho, Professor, Stanford Law School .....	8
Richard A. Eppink, of Counsel, American Civil Liberties Union of Idaho Foundation .....	10
Jacob Siegel, Writer .....	12

## ALPHABETICAL LIST OF WITNESSES

Ariga, Taka:	
Testimony .....	7
Prepared statement .....	48
Eppink, Richard A.:	
Testimony .....	10
Prepared statement .....	72
Ho, Daniel E.:	
Testimony .....	8
Prepared statement .....	66
Parker, Lynne E.:	
Testimony .....	5
Prepared statement .....	42
Siegel, Jacob:	
Testimony .....	12
Prepared statement .....	85

## APPENDIX

Statement submitted for the Record by Association for Computing Machinery	129
Responses to post-hearing questions for the Record:	
Mr. Siegel .....	141



# ARTIFICIAL INTELLIGENCE IN GOVERNMENT

---

Tuesday, May 16, 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 Peters, Chairman of the Committee, presiding.

Present: Senators Peters [presiding], Hassan, Sinema, Rosen, Padilla, Ossoff, Paul, Johnson, Lankford, Scott, Hawley, and Marshall.

## OPENING STATEMENT OF SENATOR PETERS<sup>1</sup>

Chairman PETERS. The Committee will come to order.

Today's hearing is the second in a series that I plan to convene on artificial intelligence (AI). At our first hearing in March, we discussed the transformative potential of AI, as well as the possible risks these technologies can pose. Today, we will be discussing how AI has the potential to help government better serve the American people, such as by improving the ways agencies deliver services and also what pitfalls we need to be aware of as government increasingly adopts AI.

The Federal Government is already using AI in an effort to provide more efficient services, assess potential security threats, and automate routine tasks to enhance the Federal workforce.

Earlier this month, the White House announced new efforts to invest in American leadership to develop AI technologies and promote the responsible use of AI within the Federal Government. Later this summer, the Office of Management and Budget (OMB) is expected to release new guidance on Federal Government use of AI, implementing legislation this Committee advanced in 2020, and was later signed into law in government funding legislation.

U.S. leadership in the development and use of AI systems, by both the private sector and government, is essential for our global economic competitiveness. We should work to ensure that government can adopt and deploy these tools to help improve American lives, but as we do so we must ensure we are also prepared to address the potential risks and harms that AI systems can present.

The potential for bias in AI applications can have serious consequences for Federal Government use. A recent study found that an algorithm used by the Internal Revenue Service (IRS) to deter-

---

<sup>1</sup>The prepared statement of Senator Peters appears in the Appendix on page 37.

mine who should be audited was erroneously more likely to recommend Black taxpayers than white taxpayers, and the government was not prepared with the data or training necessary to actually recognize this biased outcome.

As we heard in our last hearing, AI algorithms often lack transparency and accountability for how they arrive at certain outcomes. Even the engineers who design them do not always understand how they reach the conclusions that they reach.

In government applications, this can present serious risks to Americans who may unknowingly be interacting with an AI, and who may struggle to get answers about why an AI system made a certain determination.

For example, at least a dozen States deployed algorithms to decide eligibility for disability benefits, which resulted in denying thousands of recipients this critical assistance that helped them live independently, and left them with little opportunity to understand why the decision was made or how they could possibly appeal it.

The enormous amounts of data that can be collected as a result of using AI systems also presents concerns about privacy. Existing privacy laws do not envision these types of applications.

As agencies use more AI tools, they will need to ensure they are securing and appropriately using any data inputs to avoid accidental disclosures or unintended uses that harm Americans' civil rights or their civil liberties.

Finally, we must ensure our Federal workforce is ready to procure and oversee the use of AI systems to ensure they are benefiting Americans. Last Congress, I authored legislation that was signed into law requiring officials charged with procuring AI tools to be trained in both their capabilities to improve agency missions, and their potential risks, to ensure responsible use.

Last week, I introduced bipartisan legislation to build on that effort by requiring Federal agency supervisors and managers to receive similar training.

I am looking forward to today's discussion and to continuing to work with my colleagues on both sides of the aisle to advance solutions that will help encourage American development of AI, and ensure it is being used appropriately.

During today's hearing, we will be discussing some of those strategies, including the need to conduct inventories of current Federal Government AI applications, requiring ongoing audits to ensure the accuracy and effectiveness of AI systems, and considering responsible standards that need to be met as the Federal Government continues to acquire additional AI tools.

I am grateful to our expert witnesses for joining us today. We look forward to a fruitful discussion, and a discussion that will likely continue well beyond this hearing and be engaging for the foreseeable future.

I would now like to recognize Ranking Member Paul for his opening statement before we hear from our witnesses. Ranking Member Paul.

### OPENING STATEMENT OF SENATOR PAUL<sup>1</sup>

Senator PAUL. In 1975, the late Senator Frank Church said, “The United States government has perfected a technological capability that enables us to monitor the messages that go through the air. That capability at any time could be turned around on the American people, and no American would have any privacy left, such is the capability to monitor everything. There would be no place to hide.”

These words came as Senator Church led the Senate Select Committee to Study Governmental Operations with Respect to Intelligence Activities, better known as the Church Committee.

The Church Committee’s 1976 final report exposed numerous secret Federal programs that violated the constitutional rights of American citizens it deemed to be threats to existing social and political order. These programs surveilled and targeted individuals like Martin Luther King Jr. and domestic organizations like the Southern Christian Leadership Conference, as well as infiltrated movements to incite rivalries and discredit leaders.

Nearly 50 years later, Senator Church’s ominous warning that the government could weaponize technology against the American people reads more like a premonition. There is truly becoming “no place to hide.”

In recent decades, journalists and whistleblowers exposed examples of our government leveraging emerging technologies to violate the privacy and civil liberties of its citizens.

Intelligence agencies conducted surveillance of video game users, collecting data on the contents of communications between players.

The Department of Homeland Security (DHS) tracked the locations of individuals and groups participating in the Black Lives Matter (BLM) movement.

The Drug Enforcement Administration (DEA) conducted “covert surveillance” of people protesting the death of George Floyd.

It is only getting worse. Just last month, the American Civil Liberties Union (ACLU) acknowledged, “The Biden administration has been quietly deploying and expanding programs that surveil what people say on social media, using tools that allow agents and analysts to invisibly monitor the vast amount of protected speech that occurs online.”

How are they doing it? Using artificial intelligence.

For years, Federal agencies, including the Department of Homeland Security, State Department (DOS), National Science Foundation (NSF), and the Federal Bureau of Investigation (FBI) have been colluding with private organizations and social media companies to combat what they deemed to be “disinformation.” Jacob Siegel, in “Tablet,” wrote, “Disinformation is both the name of the crime and the means to covering it up, a weapon that doubles as a disguise.” I think that is an apt way of looking at disinformation. It is a tool for those who want to limit speech, but it also doubles as a disguise and a means of covering up what they are actually trying to do.

The purpose, so they claimed, was to combat foreign malign influence. But in reality, the government was not suppressing foreign

<sup>1</sup>The prepared statement of Senator Paul appears in the Appendix on page 39.

“misinformation.” It was working to censor domestic speech by Americans.

Since 2020, the Federal Government has awarded over 500 contracts and grants related to “misinformation” or “disinformation.” George Orwell would be proud. While the grant awardees and their proprietary AI and machine learning technologies differ, their goals are consistent: to “mine” the internet, identify conversations indicative of “harmful” narratives, track those “threats,” and develop countermeasures before messages go viral. One National Science Foundation-funded company’s mission statement claims that “social media is being manipulated and ideas are being spread uncontrollably online.”

The solution it provides? An automatic controversy detection algorithm to help identify things that are “potentially opinion-shifting” in order to make communication “more productive and less dangerous,” in other words, censorship.

During the Coronavirus Disease 2019 (COVID–19) pandemic, we witnessed the accelerated use of artificial intelligence technologies to monitor and suppress public debate on issues like natural immunity, masks, and the origin of the virus. Multiple Federal agencies, including the Department of Defense (DOD) and the State Department, funded automated disinformation-detection technologies designed to monitor and suppress public debate on issues like vaccines and the origins of COVID–19.

Writer Jacob Siegel, in a fantastic, yet haunting, narrative explaining the last decade of U.S. Government domestic censorship efforts said, “Disinformation, now and for all time, is whatever they say it is. That is not a sign that the concept is being misused or corrupted; it is the precise functioning of a totalitarian system.”

Make no mistake. The United States is engaging in the same activities we criticize other countries for. But unlike China and North Korea, the United States government attempts to conceal its involvement using private entities as front companies to do its dirty work.

But make no mistake. The intent is the same. Control the narrative, eliminate dissent, and retain power.

This should terrify all Americans. The government is using your hard-earned tax dollars to surveil and censor your protected speech. Artificial intelligence is only going to make it easier for the government to do this, and harder to detect.

This should not be a partisan issue. We must get to the bottom of how the Federal Government uses AI to violate the privacy and civil liberties of the American people before it is too late.

Chairman PETERS. Thank you, Ranking Member Paul.

It is the practice of the Homeland Security and Governmental Affairs Committee (HSGAC) to swear in witnesses, so if each of you would 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?

Mr. EPPINK. I do.

Mr. ARIGA. I do.

Dr. PARKER. I do.

Mr. HO. I do.

Mr. SIEGEL. I do.

Chairman PETERS. Thank you. You may be seated.

Our first witness is Dr. Lynn Parker. Dr. Parker is the Associate Vice Chancellor and Director of the AI Tennessee Initiative at the University of Tennessee (UT). Dr. Parker spent four years as Deputy United States Chief Technology Officer (CTO) and Director of the National AI Initiative Office (NAIIO) within the White House Office of Science and Technology Policy (OSTP).

Before joining OSTP in 2018, Dr. Parker served as the interim Dean of the University of Tennessee's Tickle College of Engineering. She has also served as the National Science Foundation's Division Director for Information and Intelligence Systems.

Dr. Parker, welcome to the Committee. We look forward to hearing your opening statement.

**TESTIMONY OF LYNNE E. PARKER, PH.D.,<sup>1</sup> ASSOCIATE VICE CHANCELLOR AND DIRECTOR, AI FOR TENNESSEE INITIATIVE, UNIVERSITY OF TENNESSEE**

Dr. PARKER. Thank you so much. Chairman Peters, Ranking Member Paul, and Members of the Committee, thank you for inviting me to testify at this hearing on AI in government. I am Associate Vice Chancellor at the University of Tennessee, Knoxville, and Director of the AI for Tennessee Initiative, where we are working to establish Tennessee as a leader in the data-driven knowledge economy.

My remarks today focus on ways that Federal agencies can better leverage and govern the responsible use of AI in advancing their missions and providing services to the American people.

Federal uses of AI are becoming increasingly transparent as agencies make available their AI use case inventories in compliance with Executive Order (EO) 13960 and the Advancing American AI Act. The extreme variety of Federal AI use cases creates challenges for developing a flexible approach to the responsible governance and use of AI by the Federal Government.

To help accelerate the responsible governance and use of AI in government, I offer the following recommendations.

First, as directed by the AI in Government Act of 2020, and Executive Order 13960, OMB should prioritize and adequately resource their work on creating Federal guidance for the use of AI in government. This guidance should address the wide diversity of use cases of AI across the Federal Government, encourage the responsible adoption of AI to improve public services while protecting privacy, civil rights, and civil liberties, and be operational for use by the agencies.

Second, Congress should require Federal agencies to use the National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF) during the design, development, procurement, use, and management of AI. Beginning with a standardized assessment of the risks posed by use cases of AI is a key step that can be taken now by all Federal agencies without needing to wait for additional OMB guidance.

<sup>1</sup>The prepared statement of Dr. Parker appears in the Appendix on page 42.

Third, Congress should require every Federal agency to have a current and regularly updated AI strategic plan that includes that agency's approach to the responsible adoption of AI.

Fourth, Congress should direct each agency to hire and resource a Chief AI Officer (CAIO) who is responsible for overseeing the development and regular update of the organization's AI strategy and use of AI.

Fifth, Congress should direct the creation of an interagency Chief AI Officers Council as an effective way to coordinate the governance and use of AI within the Federal Government.

Sixth, the proposed Chief AI Officer's Council should review the agency AI use case inventories for common application areas and identify dozens of key agency processes that could be transformed with AI, in a manner consistent with privacy, civil rights, and civil liberties.

Seventh, Congress should accelerate the responsible and innovative adoption of AI by providing agencies with AI innovation funds as part of their annual operating budgets.

Eighth, to help address AI workforce shortages in the Federal Government, and as directed by the AI in Government Act of 2020, the Office of Personnel Management (OPM) should prioritize and adequately resource their work on the AI occupational series so that Federal agencies will be better positioned to strengthen their AI workforces.

Ninth, Congress should direct the development of a National Initiative for AI Education Framework, analogous to the NIST National Initiative for Cybersecurity Education (NICE), Framework, that was developed in 2017, to provide a comprehensive and standardized approach to describing AI roles and the associated knowledge, skills, and abilities needed for those roles.

Finally, to help strengthen the breadth and diversity of talent in the nation's AI ecosystem, Congress should authorize and fund the National AI Research Resource, as proposed by the congressionally directed National AI Research Resource Task Force. Such a resource would help develop new AI talent, with some of this talent likely choosing to use their AI skills to support the Federal Government in its adoption and governance of responsible AI.

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

Chairman PETERS. Thank you. Thank you, Dr. Parker.

Our next witness is Taka Ariga. Mr. Ariga is the Chief Data Scientist and Director of the Innovation Lab at the U.S. Government Accountability Office (GAO). As an integral part of the Science, Technology, Assessment, and Analytics team, he helps GAO develop and implement advanced analytical capabilities for its auditing processes.

Prior to joining GAO, he held executive positions at Deloitte, Ernst & Young and Booz Allen Hamilton, where he worked with audit, compliance, legal, and regulated entities.

Welcome to the Committee, and we look forward to your testimony.

**TESTIMONY OF TAKA ARIGA,<sup>1</sup> CHIEF DATA SCIENTIST, U.S.  
GOVERNMENT ACCOUNTABILITY OFFICE**

Mr. ARIGA. Chairman Peters, Ranking Member Paul, and distinguished Members of the Committee, thank you for inviting me to participate in today's hearing on artificial intelligence in the Federal Government.

As GAO's Chief Data Scientist and Director of our Innovation Lab, I see AI's potential in action every day. But as you noted in your March 8th hearing, significant risks and challenges exist wherever AI is assisting or replacing discretionary decisionmaking.

AI is undeniably an integral part of a functioning digital fabric. However, the Federal Government is certainly not immune from consequences of this powerful technology. The need to promote responsible and accountable use of AI is even more striking now in the face of growing dangers from unfair, unintended, or misleading outcomes that carry cascading societal impacts.

Paradoxically, agencies continue to face acute short of Federal digital talent needed to implement accountability practices. We must address both challenges at pace with, and perhaps even anticipate, rapid advances in AI capabilities.

GAO has issued more than two dozen reports over the past several years to promote responsible and accountable use of AI. One of the most notable moments is the introduction of our AI Accountability Framework in summer of 2021. This first-of-its-kind blueprint moved beyond high-level aspirations and laid out 33 key implementation practices across four pillars for Federal agencies to consider as they navigate the AI development lifecycle. In this framework, we also stress the importance of taking a team sport approach that integrates perspectives from an ecosystem of stakeholders.

Beyond GAO's own use across our audit engagements, we are seeing adoption of the framework by the broader oversight community to conduct AI evaluations.

On the workforce front, GAO has steadfastly reported on mission-critical gaps for Federal expertise in science and technology (S&T) since 2001. The Federal Government, as a whole, continues to face barriers in hiring, managing, and retaining staff with advanced technical skills, the very skills needed to design, develop, deploy, and monitor AI systems.

In our November 2021 report, GAO gathered perspective from technology leaders across Federal, academic, and nonprofit entities to explore the concept of establishing a U.S. Digital Services Academy (USDSA). The aim is to improve the Federal pipeline of highly trained digital talent that can effectively and responsibly modernize government, including implementation of AI systems. Ultimately, having a robust cadre of a digital-ready, Federal workforce ensures humans can successfully remain in, and never out of the AI loop.

GAO remains committed to supporting Congress on the "trust but verify" part of the AI accountability equation. We have formed an internal AI community of practice that includes every GAO mission team. We have established internal training resources to en-

---

<sup>1</sup>The prepared statement of Mr. Ariga appears in the Appendix on page 48.

hance data literacy and data science. We have hired more data scientists. Our Innovation Lab is actively exploring a variety of impactful machine learning techniques to transform audit function. We remain engaged with a network of oversight partners, academic, and governmental entities at all levels to exchange insights. All of these efforts give GAO hands-on experience to stay at the forefront of AI technology, with which to offer technical assistance and strengthen oversight capacity.

Realizing accountable AI is a continuing journey that requires a whole-of-government approach, Federal agencies need more specific guidance on effective implementation of AI. At the same time, we need practical policy solutions that address interconnected challenges on privacy, civil liberties, and workforce readiness.

We know AI capabilities will evolve at an incredible speed, and the use of AI will continue to diffuse across facets of governmental functions. GAO believes that the Federal Government can, and must, simultaneously realize opportunities afforded by AI and be leaders in good governance, transparency, and compliance in this age of algorithmic renaissance.

Chairman Peters, Ranking Member Paul, and Members of the Committee, this concludes my prepared statement. I will be happy to answer any questions you may have.

Chairman PETERS. Thank you, Mr. Ariga.

Our next witness is Professor Daniel Ho. Professor Ho is the William Benjamin Scott and Luna M. Scott Professor of Law at Stanford Law School, Professor of Political Science, Senior Fellow at the Stanford Institute for Economic and Policy Research, Associate Director of the Stanford Institute for Human-Centered Artificial Intelligence (HAI), and Director of the Regulation, Evaluation, and Governance Lab. He also serves on the National Artificial Intelligence Advisory Commission and is a Senior Advisor on Responsible AI to the Department of Labor (DOL).

Professor, welcome to the Committee. We look forward to your opening remarks.

#### **TESTIMONY OF DANIEL E. HO,<sup>1</sup> PROFESSOR, STANFORD LAW SCHOOL**

Mr. HO. Chairman Peters, Ranking Member Paul, and Members of the Committee, it is an honor to speak with you today.

The youth government has an exceptional opportunity. It can seize this moment of AI innovation to modernize Federal programs, catalyze scientific advancements, and protect the rights and benefits of all Americans. Doing so will strengthen America, but strategic leadership, Federal workforce investments, digital infrastructure, and procurement modernization will be necessary. The Federal Government needs to go from having a few pockets of innovation to a culture of innovation.

Let me start with why AI in government matters so much. First, government should lead by example and demonstrate how responsible AI can modernize Federal programs. In a report we showed how early AI innovation in nearly half of the largest 140 Federal agencies can transform Veterans Benefits Administration (VBA),

<sup>1</sup>The prepared statement of Mr. Ho appears in the Appendix on page 66.

improve monitoring of adverse drug events, and help protect workers, consumers, and the environment.

Take the Social Security Administration (SSA), which can hear over half a million disability appeals per year. With great foresight, SSA began investing in data infrastructure and tools to modernize case adjudication in the 1990s. This culminated in an AI tool that allows judges to use natural language processing to check draft decisions for some 30 errors, accelerating and improving case processing, leaving some to call the official who pioneered these early investments the “Steve Jobs of the SSA.”

Second, government agencies are, of course, critical for effective regulation of the risks of AI, and striking the right balance between innovation and safeguards requires expertise in government. Getting technical talent into the Federal workforce is the single biggest obstacle for effective regulation. Government cannot govern AI if it does not understand AI.

While much progress has been made, including legislation from this Committee, we still have a long way to go. When our research team at Stanford examined the implementation of AI-related legal requirements, stemming from two Executive Orders and the AI in Government Act, we found a critical gap in leadership, strategic planning, and capacity. For instance, 88 percent of agencies failed to submit AI plans to identify regulatory authorities, and the implementation of a key transparency measure, agency disclosure of AI use cases, has been inconsistent, and the Office of Personnel Management has yet to release a required report, due July of last year, to forecast the AI employment needs and to create an AI hiring line.

This must change. The Federal workforce does herculean work but faces fundamental challenges developing teams that can design, implement, and regulate AI effectively and responsibly. Stanford’s HAI’s AI Index highlights that 65 percent of AI Ph.D.’s land in industry, 28 percent in academia, and less than two percent in government. Or in the words of one entrepreneur, “The best minds of my generation are thinking about how to make people click on ads.”

Strengthening the pipeline of technical talent into the public sector and providing career paths is urgent. As the National Security Commission on AI noted, it is not just compensation. “It is the perception, and too often the reality, that it is difficult for digital talent in government to perform meaningful work.” I have seen this firsthand. One Stanford AI Ph.D. student became so frustrated by an agency’s decades-old software stack and lack of computing resources that he gave up on his aspirations for a career in government and went back to work in industry.

Let me conclude with four recommendations.

First, strategic leadership from the Federal Government is required to coordinate and drive forward trustworthy AI innovation. Congress should borrow a page from the bipartisan Evidence Act and empower chief AI Officers to ensure that senior leadership within agencies is driving forward responsible AI innovation and oversight.

Second, Congress should establish new pathways and trajectories for technical talent in government. We need better models—build-

ing on the U.S. Digital Service (USDS), public-private partnerships, and academic-agency partnerships—to attract AI talent to public service, build cross-functional teams, and provide pathways for career advancement.

Third, government procurement is critical to capitalize on American innovation and spur developments of rights-preserving, privacy-enhancing technologies. We need to move toward more modular forms of contracting, which the Department of Defense has illustrated, that enables more effective development, acquisition, assessment, and auditing of AI systems.

Last, we have to invest in digital infrastructure, including the National AI Research Resource, endorsed by multiple Federal task forces, for secure access to administrative data and large-scale computing resources to level the playing field. Government data, which is higher quality, more representative, and reliable than web data that many models are trained on, is an important part of the solution. When the U.S. Geological Service (USGS) made Landsat satellite imagery free to researchers in 2008, it generated \$3 to \$4 billion in benefits annually, catalyzing discoveries in habitat modification, climate change, and poverty. That is the promise of getting public sector innovation right.

The U.S. Government should act expeditiously to foster responsible AI adoption.

I am looking forward to your questions.

Chairman PETERS. Thank you. Our next witness is Ritchie Eppink. Mr. Eppink serves as Counsel for the American Civil Liberties Union of Idaho Foundation. He was previously the Justice Architect for the Idaho Legal Aid Services, and before that a Fulbright Fellow.

Mr. Eppink, you are recognized for your opening statement.

**TESTIMONY OF RICHARD A. EPPINK,<sup>1</sup> OF COUNSEL,  
AMERICAN CIVIL LIBERTIES UNION OF IDAHO FOUNDATION**

Mr. EPPINK. Thank you, Chairman Peters, Ranking Member Paul, and Committee Members for your attention to artificial intelligence and automated decisionmaking in government programs.

I was invited here today because I have been working for over a decade with Idahoans with developmental disabilities and their families to challenge secret decisions made by computerized algorithmic systems. Only through litigation that I have helped these families pursue were they able to access the algorithms the State of Idaho uses to make decisions about the health care that they depend on day to day.

Once we opened the black box that concealed that automated system, we found that it was built out of corrupt data, relied on inputs that the State never validated, and produced results that even those who created it could not explain. A Federal court ruled that the system was unconstitutional.

Yet a decade after filing suit and over seven years since winning in court we are still litigating the case, battling for due process against still more black box secrecy. Decades-long class actions by indigent families are not a viable plan for AI governance in tax-

<sup>1</sup>The prepared statement of Mr. Eppink appears in the Appendix on page 72.

payer programs. We need Federal regulation and enforcement to protect basic fairness and constitutional rights in government programs that use AI in automated decisionmaking systems.

A little bit about the Idaho lawsuit. Although in the past our society shamefully confined people with developmental disabilities in State hospitals, today, through Medicaid, they can get services at home and in their communities instead, at a savings to the government and to taxpayers.

One of my clients was Christie Mathwig. She was a bulwark in her rural community of Troy, Idaho, a mother, and a leader in her church, bible studies, and Al-Anon. She was also diagnosed as a teenager with a rare neuromuscular disease and relied on workers in her community for health care and support throughout her day. Each year, the State assigned her a budget: not dollars that she would ever see, but reimbursement for her health care providers, based on assessment results plugged into an automated system.

Christie came to me when the State told her her budget would suddenly be cut by tens of thousands of dollars, more than 20 percent, and not enough for the 24-hour support she needed to survive. The State refused to provide an explanation, claiming that the system was a “trade secret.”

We filed suit, and a Federal court quickly ordered the agency to disclose the system to us. It turned out to be just a handful of formulas coded into a basic Microsoft Excel spreadsheet. As rudimentary as it was, it still took us many months, three experts, and over \$40,000 to reverse-engineer the system, catalog its flaws, and assess the harm that its results could wreak upon our clients. We presented our analysis to the court, and it ruled that the system arbitrarily deprives participants of their property rights and, hence, violates due process.

I want to point out three dangers with automated government decisionmaking that this Idaho litigation, which is known as the *K.W. v. Armstrong* lawsuit, illustrates, and then I want to share three solutions that Congress should enmesh across Federal programs like these.

First, the dangers. One, black boxes that conceal government use of AI. If my clients had not found a lawyer with the time and resources to help, they probably still would not know that Idaho was using an automated system to make decisions about them in the first place.

Two, black boxes concealing how these systems work, including bad data that they are trained on. Once a Federal court order put Idaho’s secret formulas into our hands, it took a mammoth effort to figure out all the many things that were wrong with it, including erroneous and corrupt data underlying it.

Three, black boxes that prevent accountability. Idaho’s Medicaid agency has been fighting again and again, including just last month, to ban my clients from accessing the very information they need to challenge the results of its automated system.

Now toward solutions. There are three that I want to stress today.

First, the people that these AI systems make decisions about should be integrally involved in their development, implementation, and evaluation. This is a cornerstone of the court-ordered set-

tlement agreement in the Idaho case, and it is the solution that could prevent the most dangers.

Two, government agencies must implement constitutional rights through regulation and enforcement specific to AI systems. Case-by-case litigation, which we know from Idaho is immensely resource intensive, is not a sustainable solution.

Three, transparency requirements and governance standards must apply to these systems from before they start until after they finish. Black boxes have plagued the Idaho system since 2012, and continue through to the present, in 2023. My clients have a right to the same information the government does to evaluate these systems and to challenge their results, and private contractors' proprietary interests can never be allowed to trump individual due process and equal protection rights.

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

Chairman PETERS. Thank you.

Our final witness is Jacob Siegel. Mr. Siegel is a writer and senior editor of News and The Scroll for Tablet magazine. He previously covered national security and digital culture for the Daily Beast. His writing has been published in the New York Times, Politico, the New York Daily News, Vice, Rolling Stone, and the National Endowment for the Humanities magazine.

Mr. Siegel, welcome. You may proceed with your opening comments.

#### **TESTIMONY OF JACOB SIEGEL,<sup>1</sup> WRITER**

Mr. SIEGEL. Good morning, Chairman Peters, Ranking Member Paul, and Members of the Committee. Thank you for the opportunity to testify. I am Jacob Siegel, a writer and senior editor at Tablet magazine as well as a former military intelligence officer and Iraq and Afghanistan veteran.

Throughout history, warfare has spurred the development of transformative new technologies. My experiences in the war on terror provided me with a glimpse of the AI revolution that is remaking America's political system and culture in ways that have already proved incompatible with our system of democracy and self-government, and may soon become irreversible.

I first encountered issues of artificial intelligence in governance when I was deployed to western Afghanistan in 2012. What I found there was that in the midst of a great deal of confusion and ambivalence about the U.S. mission and what we were still trying to accomplish after a decade at war, the military had turned to powerful new information technologies (IT) to fill the strategic void.

On critical fronts like defeating the Taliban and standing up the Afghan Security Forces, our success remained illusory. But in the face of this systemic failure the United States developed a special capacity for building databases.

The theory of data-driven warfare was that collecting enough information and marrying it with the proper algorithms into AIs that could perform predictive analysis opened a technical portal into the future. We could stop the next improvised explosive device (IED)

<sup>1</sup>The prepared statement of Mr. Siegel appears in the Appendix on page 85.

attack before it occurred, control events on the ground, and win over the Afghans to our cause.

It did not work out that way. Years before the United States withdrawal and the Taliban's return to power, I had come to see that the gap between our official metrics of success and the reality on the ground was not only a result of measuring the wrong things. By translating critical questions of politics and policy into the language of data, we had outsourced the most fundamental responsibility of statecraft to machines, while rendering the essential notions of war, victory, and peace, obscured to America's leaders.

AI moves us at an exponential rate from obscurity to the impenetrable darkness of the so-called black box. As the computer scientist, Stephen Wolfram noted when he testified before the Senate in 2019, "If we want to seriously use the power of computation and AI then inevitably there won't be a human explainable story about what is happening inside."

America was founded on the ideal that individual citizens, through their free and informed actions, should participate in their government. But for a free people to participate in the making of their own laws and the meaning of their own lives, they must be knowledgeable about the world around them. Centralized applications of AI that invisibly alter the architecture of perception and reality, for instance by performing mass censorship of certain phrases or narratives, makes such knowledge impossible.

Moreover, as the writer, James Poulos, has noted, "The AI approach to governance undermines individual's faith in their own capacity for reason to action because it is driven by a logic of seeing technology as better and stronger than humanity."

In other words, a technology that is intrinsically threatening to human interests, with a potential transformational power on the order of the printing press or the wheel, which is, at this moment, being funded and deployed by multiple government agencies, appears destined to move further away from human understanding the more it progresses. Yet there is no chance that the U.S. Government and U.S.-based corporations are going to abandon a technology this powerful, nor would such an outcome necessarily be desirable, given that it would cede the space to competitors like China.

We seem to be caught in a trap. There is a vital national interest in promoting the advancement of AI, yet at present the government's primary use of AI appears to be as a political weapon to censor information that it or its third-party partners deem harmful.

Example abound from recent years of this kind of AI-driven informational control system, which is deployed at every opportunity in the name of public safety and emergency. It is in the name of safety that government officials are now calling for even more control over AI.

Earlier this month, Jen Easterly, the Director of the Cybersecurity and Infrastructure Security Agency (CISA) called for more regulation of AI, warning that we need to be very mindful of making some of the same mistakes with artificial intelligence that we have made with technology. But regulating AI so that it becomes an even more powerful tool of censorship for enforcing party orthodoxies will increase neither our safety nor our security.

Easterly also recently argued that China has already established guardrails to ensure that AI represents Chinese values, and the United States should do the same. While emulation of the Chinese model of top-down, party-driven social control appears to be the direction that AI in governance are moving in the United States, I would submit, respectfully, that continuing in this direction will mean the end of our tradition of self-government and the American way of life.

Thank you.

Chairman PETERS. Thank you, Mr. Siegel.

Mr. Eppink, in your testimony you told us about how you are suing the Idaho Medicaid program for its failure to disclose the algorithm that it used to substantially cut recipients' health care services. But could you tell the Committee more about how your clients even learned about the use of this automated system?

Mr. EPPINK. Certainly, Chairman. The Idaho system, when we got started, I think like many other systems around the country probably still today, was not disclosed to anyone outside of the Idaho agency, so far as I know, and after about a half dozen families had contacted me, scared that their lives would be upended, I figured I would send a letter, which I did, and I got one back from the agency's lawyers saying that the system was a trade secret. Once we knew it was a secret, we had to go from there to file a lawsuit to find out more about what the system was.

Chairman PETERS. Right now we are hearing an awful lot about generative AI tools, like Chat Generative Pre-Trained Transformer (ChatGPT). It is all over the news. Everybody is chatting about it. It is a hot topic.

The system that you are talking about was relatively simple and an older, automated system compared to what we are seeing right now. What are your thoughts on that fact and how it should inform our efforts to set Federal AI policy?

Mr. EPPINK. Yes. I guess for starters, black boxes are black boxes, no matter how big they get and no matter what is inside of them. Even though the Idaho system was just an Excel spreadsheet, it still ended up taking us a significant amount of time—I think it was maybe months, possibly a matter of years—to get all of the information on which it was built.

Federal AI policy standards and enforcement that are governing use of automated systems and AI are not solving necessarily for the complexity of those systems. They are solving for the harms that they are causing.

I do not know that we need new standards for each new technology. I think the same principles that we can apply to protect and prevent some of the things that we have had to litigate in Idaho would apply to more complex systems as well.

Chairman PETERS. Clearly, as I listen to your testimony, you believe that there is simply just not enough transparency. What would you recommend when it comes to holding these platforms more accountable?

Mr. EPPINK. What is critical is standards and enforcement that are specific to AI and automated decisionmaking systems, which we have had to enforce in Idaho through the litigation. We have basic American principles like due process and equal protection

that are there. Where our courts struggle and where our agencies still have gaps as far as what they are overseeing is how to apply those in this new context to these automated systems. We know that the litigation in court, I think, is too unwieldy. Although we have done it in Idaho, it is not going to meet the proliferation of these systems.

Each agency, for instance, with the Medicare program in Idaho, the Centers for Medicare and Medicaid Services (CMS) should be overseeing programs that are using automated systems, like Idaho, and making sure that they are complying with standards that are specific to those programs. We already have a lot of jumping-off points that can be used to develop those standards. They just need to be put in place and then enforced by these agencies.

Chairman PETERS. Thank you.

Dr. Parker, I have two questions for you. First, some government uses of AI are clearly high risk, like when the Department of Homeland Security uses facial recognition technology. On the other hand, there is also some low-risk use of AI, like when the Forest Service classifies tree canopy coverage with machine learning tools.

The first question. Should testing, auditing, and procurement requirements be different, depending on the type of system, and two, how should the Federal Government decide what is actually high risk and what is low risk?

Dr. PARKER. Thank you for the question, Mr. Chair. There are certainly many types of AI systems, and I think what is getting attention a lot in the press these days are some of the extreme cases. But we want to also encourage the use of AI in the mundane cases because they can improve services for the American people and they can improve efficiencies.

NIST has come out recently, at the direction of Congress, with a NIST AI Risk Management Framework. I think evaluating the risks of each individual use case is important to have the right governance approach. We do not need onerous regulations and oversight of simple use cases that no one believes are going to harm anyone.

Having that overarching policy, that for every use case of AI within the Federal Government those agencies step through the NIST AI Risk Management Framework, to determine what is the risk level. If it is low, then the steps that are needed in order to make sure it is used safely are much more minimized. If it is more consequential, then, of course, we do need to have much more oversight. Certainly GAO's Accountability Framework or the Blueprint for the AI Bill of Rights are designed to try to inform how to address some of those high-risk uses.

Chairman PETERS. Thank you, Dr. Parker. You mentioned the Government Accountability Office, so this question is for you, Mr. Ariga. In your testimony you mentioned how the GAO has created a toolkit to audit the AI system used by government, and that you are piloting that toolkit right now with one agency.

My question for you is, can you tell us more about how the pilot is going, what roadblocks you may be facing, and at this time would it be feasible for all agencies to use the GAO Accountability Framework to audit their AI systems, and if not, why?

Mr. ARIGA. Thank you, Chairman Peters, for that question. GAO published our AI Accountability Framework back in the summer of 2021, and it is really a first-of-its-kind blueprint for agencies to consider as they navigate the AI development lifecycle. The basic premise is that if agencies can implement practices, entities like GAO can then come afterwards to identify and evaluate those, and point out any findings and recommendations there may be needed.

Certainly GAO has a number of ongoing audits using the AI Accountability Framework, and we are also hearing that other oversight entities are also using the framework itself to conduct their AI evaluations.

There are three challenges that particularly, I think, hamper the AI development within the Federal Government space. One is that while our framework looks at the AI development lifecycle, agencies can use more domain-specific guidance. For example, the ingredient list for transparency in medical diagnostics may look very different than facial recognition or even autonomous vehicles.

We also need policies around the notion of privacy, transparency, and compliance. And then last, we certainly need a digital-ready workforce to implement those practices in a way that keeps humans in the loop.

Chairman PETERS. Great. Thank you.

Ranking Member Paul, you are recognized for your questions.

Senator PAUL. When I first read 1984 I was concerned but I realized at the time we did not have the technology for 24/7 surveillance and we did not have two-way television sets. There was not this ability to abuse our rights in such a systematic fashion. We do have that technology now, so it is of more concern. But I still think it is a mistake to concentrate on the technology and not the concentration of power. I, for one, do not fear AI at all if the Bill of Rights were protected. I think it really is a question of whether or not we would allow AI to invade our rights as protected by the Bill of Rights.

Mr. Siegel, do you believe it is possible for the government to define and police disinformation without infringing on the First Amendment?

Mr. SIEGEL. I think, Ranking Member, that if it was a strictly defined category of disinformation that applied specifically to foreign actors that had strict limits on its application, that it might indeed be possible but then in practice that is, that is not what its going to do.

Senator PAUL. I think that makes a good point. The Bill of Rights does not necessarily apply to every person living in Libya or Sudan or any other different country. We surveil them all the time. We collect all of the phone information. I think in one month we collected every phone call in Italy. I do not know why Italy, but we collected every phone call in Italy. The only way you could possibly get through that is with some kind of AI program. We have like thousands and thousands of hours of audio. The only way you could get through that is with help.

But this is war we are talking about, when we are interacting with foreigners. When we are talking about us, we are supposed to have the Bill of Rights as protection.

I guess I would make the point that I do not think disinformation for the U.S. citizen, I do not think it can be defined. I think it is in the eyes of the beholder. If you are willing to police information you run a great risk of infringing the First Amendment. I do not know how you do define disinformation.

For example, in the Virality Project, which was from Twitter, but it also was Twitter working with the FBI, and the FBI paid Twitter for the information, they explicitly said that we need to take down things, even if they are true, if they might breed vaccine hesitancy. If someone did die from a vaccine and it is absolutely true, we are still going to take it down because the people are not smart enough to understand and have this information.

Do you think for domestically, maybe part of the solution would be saying that we restrict this type of technology from being used to limit or hinder protected speech?

Mr. SIEGEL. Yes, Ranking Member. I think that in the domestic context it is inherently if not unconstitutional then direct infringement on the Constitution, potentially, at least, and that it is also worth pointing out that from the origins of the modern U.S. counter-disinformation establishment in 2016, there are explicit references to the strategic difficulty in distinguishing between foreign and domestic actors.

The origins of the Global Engagement Center (GEC), for instance, in 2016, there is already talk about how the internet, from the leaders of the GEC, about how the globalized nature of the internet makes it so that what are referred to as old-fashioned privacy laws are now a hindrance on the proper collection of what is considered to be relevant information in surveillance purposes.

I think that in the domestic context not only is there an inherent difficulty in enforcing any standard of disinformation that is not inherently politicized, it also appears to be very difficult to make strict legal distinctions between foreign and domestic actors that are then enforced.

Senator PAUL. The simple way of narrowing it down is to put it in the context of the Bill of Rights, basically that if it is a right that is protective, AI should not be used in any way to define or limit that right. The forest canopy, I do not care if you use AI. Overclassification, we have 25 million records, and we want to run AI through it. My son writes the code for AI, so I am not against AI in any way, but it just needs to not infringe on speech.

On the medical example, you have to have a human. Even with the overclassification problem, if you have an AI trolling through there to point out what could be declassified, still some human is going to have to look at it to declassify it at some point, but it helps to organize things.

Really this is about protecting the Bill of Rights. But what disappoints me is in the current political atmosphere there is not really so much a bipartisan concern for this. I have heard almost nothing from the other side on the idea that the FBI was going to Twitter and saying, "Take down this information. We think this does not meet your policy," and Twitter is going, "Well hell, it is a lot of work. Will you pay us?" The FBI is saying, "Yes, we will pay you to take down this information." They say, "Is all of it disinformation, not true?" "Some of it is true. We just do not think

it is helpful for society to read this. It is not helpful for them, even if it is true.”

That should scare us all, on the right or the left, and yet the outrage seems to have been largely one-sided.

For example, I have a bill, and it simply would say this: nobody in government can meet with, collude with, or work with anybody in media of any form to limit or restrict protected speech. The Constitution, the courts have defined protected speech. It would largely keep them out of this. Some people are like, “Oh, people would say things are untrue.” Yes, but unfortunately the First Amendment allows bad speech, allows people to say terrible things, and occasionally would allow people to say something. But most of the time there is actually a debate, with evidence on both sides, and that is what the news media is about, trying to figure out what the truth is. It is not like one source has all the truth.

I think if we could get back to more discussing the Bill of Rights in the context of artificial intelligence I think we would have a better chance of getting to the solution. Sure, will there be some specific controls on artificial intelligence? Yes, but it is not so much that we should fear technology. We should fear the technology in the hands of people who would abuse our liberty.

But thank you all for appearing.

Chairman PETERS. Thank you, Senator Paul.

Senator Lankford, you are recognized for your questions.

#### **OPENING STATEMENT OF SENATOR LANKFORD**

Senator LANKFORD. Mr. Chairman, thank you. Thank you all for being here and being in the conversation. Obviously, this is an ongoing conversation. I want to drill down on a couple of comments that a couple of you made and a couple of you have implied on this.

Dr. Parker, you made the statement, “responsible use of AI.” Mr. Ho, you made the statement, “responsible AI and AI innovation.”

OK. Define for me what “responsible use of AI” is, and maybe that is the flip side of what irresponsible AI is.

Dr. PARKER. Thank you for the question, Senator. I think this is the million-dollar question, and this is the reason why I think guidance from OMB, for instance, as directed in the AI in Government Act and the Executive Order 13960 is so critical, because that process that agencies are expected to follow to ensure that their AI is used in a way that is upholding of our expectations for whatever responsible means is key. I think by defining the processes that agencies must abide by as they look at their use cases of AI, look at the risks of AI, that will inherently help us understand what is responsible.

But if you look sort of theoretically at what responsible AI means, there are a number of principles that around the world we have converged on as those characteristics of AI that we agree fall into either what we call “responsible” or “trustworthy” AI. These are things like safe and effective. They are the way that the systems actually work is consistent with their intended use. We have accountability. There are a number of these kinds of principles that taken as a whole can be reflective of our sense of whether or not the AI system is trustworthy or whether or not it is responsible.

To summarize, I think if we have AI systems that we agree, through a set of processes, that I think should be defined by OMB, that allows those AI systems to demonstrate those agreed-upon principles, then I think we would call that “responsible” use of AI and “trustworthy” AI.

Senator LANKFORD. OK. Mr. Ho, do you want to take a stab at that?

Mr. HO. Sure. I agree with much of what Dr. Parker said. I think there actually is a fair bit of agreement when you look at various frameworks, the Executive Order 13960 on trustworthy AI within government, the Blueprint for an AI Bill of Rights, the NIST AI Risk Management Framework, and the GAO framework, that Taka Ariga mentioned. There have been attempts to try to actually look at the agreement, and there actually is fairly large agreement across these in terms of privacy protections, nondiscrimination, safe and effective forms of AI, and human awareness of how they are being used.

I think the frontier of all of this is really how to take these principles and bring them into practice, and that is where having the requisite technical talent and leadership within government agencies is going to be so absolutely critical. Because we can all agree on principles in practice, but trying to actually drive them into operation I think is a really important frontier.

Senator LANKFORD. Much of the AI conversation really boils around not just a responsible use but the data that is behind it. If the data is not good, then the whole system is going to be bad.

Here is the fear that I hear from people is that if a wrong dataset ends up into the mix here and it affects you and your family, as you have dealt with in Idaho, and you have a false dataset or a false algorithm or something that has a glitch in it, you can be directly affected by that and you do not even know how to be able to reach into it.

All these different AI researchers now are all trying to get more access to government data, No. 1, because it is perceived to be free, when actually the taxpayers paid for this, but they want to get mass amounts of data, they want mass amounts of faces, they want mass amounts of information. The Consumer Financial Protection Bureau (CFPB) just came out with a new rule for banks, for loans they want 80 different datapoints, and they want to make those publicly available.

There is a huge amount of data that government is pulling and making it more and more available, where it is easier to connect the dots and to be able to identify individual people, where maybe that dataset does not identify that person, but it is not hard to get three different datasets, combine them together with their cellphone records that are publicly available for that data point, and to be able to identify this person and all of their habits, all of their locations, everything else.

There is a concern on the availability of data that is out there and the continual push by every entity to say, “We want our AI to be better. Our AI can be better, we want more data on every individual.” Set that aside with responsible use of AI and with privacy, because right now privacy seems to be losing the battle, and so people can do more with AI because they could if they had more data.

Where am I off?

Mr. HO. Senator, thank you for that exposition. I think privacy is absolutely paramount, and I think as the number of the panelists have noted, a national privacy legislation would be quite important here, where currently we have a system that is a kind of patchwork system.

I think the other dimension that I would point out is that, I think going back to what Ranking Member Paul noted at the beginning in terms of the conversations in the 1970s, Congress made a choice to enact the 1974 Privacy Act that really only reached sort of the data that government agencies have, and really did not reach kind of the private sector.

I think some of the places where we see the most acute concerns of hoovering up lots of data and identifying individuals is in the small number of technology companies that have scraped the entire World Wide Web and they are building models off this. That is why I think coupled with comprehensive privacy legislation, I think what is really needed is a kind of data strategy, and I think there are some really good blueprints for this in the National Secure Data Service that came out of the Evidence Act, that is trying to put in a series of kind of safeguards to really ensure that it is the right people with the right safeguards who have access to data. That is the same kind of blueprint that is being used for the National AI Research Resource, so that when folks have access to administrative data it is done in a secure and privacy-protecting way.

Senator LANKFORD. Does anybody else want to make a comment on that?

Mr. EPPINK. Two other things I might add, Senator, on data. One is to recognize that, as I have talked about a little bit, the costs and the time necessary to analyze this data for someone who these systems are making decisions about is inaccessible. In the data, often, especially when you are talking about my clients who are people with developmental disabilities, is corrupted not only potentially by the creation of the data in the first place but by years and years of discrimination and other effects that have biased the data in the first place.

Senator LANKFORD. Thank you, Mr. Chairman.

Chairman PETERS. Thank you.

Senator Scott, you are recognized for your questions.

#### **OPENING STATEMENT OF SENATOR SCOTT**

Senator SCOTT. First, thank you, Chairman Peters, for holding this hearing.

Artificial intelligence surely has productive uses, but it can also present great threats, especially to our children. Today I am introducing my AI Shield for Kids (ASK) Act, to prevent children from accessing artificial intelligence features on social media sites without the consent of a parent or guardian.

I have long been a supporter of doing more to keep our kids safe online. There is no doubt that we must do more to combat the emerging threats our children face each and every day on the Internet. Like probably some other people here, I have a bunch of grandkids, and I do not ever want to put them at risk.

Mr. Ariga, AI poses significant risks to Americans, particularly vulnerable Americans such as children and youth. There are threats to privacy, user manipulation, and safety concerns. We know that AI regulation is lagging behind the speed of AI development and use. What guardrails are needed to protect vulnerable Americans from threats from intrusive AI, such as Snapchat's chatbot?

Mr. ARIGA. Thank you, Senator Scott, for that question. For us at GAO, we certainly believe in the "trust but verify" part of that equation. We want to be able to assess implemented practices that agencies have adopted to make sure that they are in line with the internal control principles as well as in the areas of privacy protection and safeguards that you have mentioned.

Within our AI Accountability Framework we have actually laid out specific practices in the areas of governance, data, performance, and continuous monitoring to make sure that the agencies themselves have infrastructure as well as evidence of those implementations, and in areas where we find potential recommendations and areas of improvement we will certainly issue findings in our report as well.

Senator SCOTT. Snapchat admitted that their AI technology is experimental, so do you agree that it is alarming and problematic that Snapchat would force their users, which comprise almost 60 percent of American teens, to use their chatbot features unless they paid to disable it with a subscription service?

It is hard to believe. If you do not want it, you have to pay to get it off, if you want to keep using Snapchat.

Mr. ARIGA. Yes, I will answer that. GAO's role is to provide rigorous oversight. Should agencies decide to use any AI technology, including Snapchat, we certainly stand by, ready to do the rigorous programmatic assessments of those programs, in line with what we have laid out in our Accountability Framework.

Senator SCOTT. Do you think you ought to have to pay to get Snapchat off, The AI off your app that you want to use?

Mr. ARIGA. We will be interested in hearing the legal rationale and any other governance structures that arrive at that decision.

Senator SCOTT. How about parents? Do you think parents and guardians have the right to protect their children and revoke consent without being charged a fee?

Mr. ARIGA. Certainly GAO has done a number of work around education, around childhood development, and so we would sort of take that similar approach at looking at programmatic implementation to see how that aligns with what we have laid out in the Accountability Framework.

Senator SCOTT. Do you agree that you ought to get consent of parents for teenagers?

Mr. ARIGA. Again, I will go back to whether those legal rationales were done in a deliberately and sound way.

Senator SCOTT. Do you have kids?

Mr. ARIGA. Yes, I do, two.

Senator SCOTT. OK. Do you watch what they look at?

Mr. ARIGA. For sure. When my daughter looks at YouTube videos I certainly want to make sure that the content is appropriate for her age. As a parent I certainly share the same concerns that you

do have. But from a GAO perspective, we try to be specific with our oversight role to say as agencies implement these capabilities there are specific sort of guardrails and expectations from internal controls that GAO will be looking at.

Senator SCOTT. As a parent, do you think that you ought to have to pay to get AI off something that your teenager uses? Forget your job. Just as a parent, what do you think?

Mr. ARIGA. Fortunately, I do not have a teenage kid just yet. But eventually I think I will have to grapple with that reality, in addition to all the subscription fees that we have already paid for.

Senator SCOTT. It appears that unelected administrative officials in DHS and CISA and others in this Administration have urged censorship on disinformation. It is frustrating when people say the border is secure and the laptop was not real, stuff like that.

Mr. Siegel, as a writer, can you talk about the dangers of Big Government colluding with Big Tech and corporations to censor journalists like the New York Post on a story that was true?

Mr. SIEGEL. Yes, Senator, thank you. I can address that. I think the dangers cannot be overstated. I think that kind of collusion between government and Big Tech and interference in open discourse in the political system is, strictly speaking, incompatible with the continued practice of democracy and self-government. I do not think that you can have free and fair elections when there is mass censorship occurring at scale, when there is collusion between various intelligence agencies, Federal agencies, not only Big Tech corporations but also other third-party, nonprofit actors who are essentially operating in a para-governmental role.

I do not think that it is simply a question of censorship, though in a technical sense censorship is what is occurring and certainly that is bad enough. It seems to me to be a surreptitious form of government, that there is a form of governmental authority and control being exercised through these relationships, determining how Americans are conditioned to perceive various policies and acts of government, ranging from things related to public health and vaccines to foreign wars. That insofar as sovereignty has been secretly transferred from individual citizens to these new relationships of power, it seems to me incompatible with the traditional American system of democracy.

Senator SCOTT. What do you think about the dangers and risks of the government potentially funding and moving toward algorithms that use AI technology to censor Americans online?

Mr. SIEGEL. I think it is a related sense of risk, a related set of risks. I think that the greater risk, what we have not fully seen yet, is censorship that is effectively invisible because it uses AI to trap speech and narratives on the wire, as the phrase goes, meaning that rather than waiting for the New York Post to publish something, and then Twitter, Facebook, and other Silicon Valley companies operating under the direction of Federal Government agencies then censoring that information, potentially what we see in the future is AI being used to censor information before it is ever published. That could happen in forums where we have come to expect mass censorship and which we view as public forums like on Facebook or on Twitter, for instance. But it could also happen on what you might think of as the back end.

For instance, Google was censoring Google Docs during the pandemic. There was a white paper on hydroxychloroquine that was published to Google Docs, and I think most people's understanding of Google Docs is that it is a kind of semi-private document that Google has given to you as one of its suite of services. But that illusion of ownership or control is just that—it is an illusion. This paper, white paper, whatever one thinks of its contents, was deleted by Google without explanation, without reference to any formal policy violation.

It is those kinds of invisible censoring acts occurring at scale that I think AI could drive, and it would be an even more significant threat than what we have already seen.

Senator SCOTT. Thank you. Thank you, Chairman.

Chairman PETERS. Thank you, Senator Scott.

Senator Hassan, you are recognized for your questions.

#### **OPENING STATEMENT OF SENATOR HASSAN**

Senator HASSAN. Thank you very much, Mr. Chair, and I want to thank you and the Ranking Member for holding this hearing, and I really want to thank the witnesses for being here today. Thank you for sharing your expertise and perspectives with us.

I want to start with a question to you, Dr. Parker. Congress created the National AI Initiative Office to improve efficiency and help Federal and local governments, researchers, business leaders, and other stakeholders collaborate on AI issues. You were the first National AI Director, but that position has remained vacant since you left the post in August.

As the former Director, why is it important that we fill this position?

Dr. PARKER. Thank you, Senator, for that question. The National AI Initiative Act of 2020 set up three very important goals for our nation about us leading the world in AI research and development (R&D), about making sure that we also lead in the development of trustworthy AI and both the development and use of trustworthy AI in both the public and private sectors, and to make sure that we are educating and training our workforce to participate in these AI activities.

It also put in charge of these activities the National AI Initiative Office, and that is in addition to a coordination role among all the Federal agencies and what they are doing in the AI space.

Given the magnitude of the importance of the National AI Initiative, as set forth by Congress, I think it is imperative that we have leadership in the White House that is overseeing these activities, that is pushing forth the innovation and the research that we need to address many of the kinds of concerns that have been raised in this panel so far, to make sure we are training and educating our workforce and leading new initiatives across the government that could help advance these, and the governance issues that we are addressing today.

That leadership vacuum I think has contributed to a number of challenges that we have across the board, in terms of being able to implement these good policies that are in place.

Senator HASSAN. Thank you very much for that.

To Mr. Ariga, AI researchers have highlighted some of the potential public safety risks posed by AI systems. These safety concerns include AI systems providing dangerous information to bad actors or potentially acting in unpredictable ways that run counter to the intent of designers. Increasing the safety and predictability of AI systems requires more technical research into the methods used to create these systems. What can the Federal Government do to support or coordinate research that would help improve the safety of AI systems?

Mr. ARIGA. Thank you, Senator Hassan. In our AI Accountability Framework we laid out specific governance practices where agencies can adopt to really consider, for example, is AI even necessary in this particular use case? In fact, if it is, what are some of the organizational structures that are in place to assess the legality, the compliance factor, as well as driving some of these inherently governmental functions?

We go back to our framework itself that lays out not only governance but looking at data performance, as well as the important topic of continuous monitoring when it comes to AI.

Senator HASSAN. Thank you. Another question for Dr. Parker. Congress created the National Artificial Intelligence Strategy, as you noted, to establish goals, priorities, and metrics for guiding and evaluating interagency work on AI. That strategy focuses on things like AI research and investments in workforce development. However, the strategy does not currently require a strategic level focus on safeguards to prevent AI from being used in a manner that harms our country or society.

Can you speak to how a focus on AI safeguards could be incorporated into the National AI Strategy?

Dr. PARKER. Yes. Thank you, Senator. I think it is true. If you look at, say, the role of the National AI Initiative Office, it is to oversee the research and the education and the coordination, but it does not provide a role for that governance. We need an approach to be able to govern that responsible use of AI.

I think one thing that could be done is to have kind of a two-part approach. One is to have Chief AI Officers at every agency that are responsible for these kinds of activities within each individual agency. This has been a challenge because even in the Executive Order 13960, that requires agencies to identify a responsible AI official, that has not happened well. We do not have a single point of contact or a single responsible person at the agencies that oversee these activities.

But agencies all have different missions, and so in order to coordinate in a consistent way across all of the Federal agencies I think the creation of something like a Chief AI Officer's Council, that is led by OMB and OSTP's National AI Initiative Office, perhaps with representation from the General Service Administration (GSA's) Center of Excellence (COE) in AI as well as the Community of Practice would provide the expertise across the Federal agencies to coordinate these processes and provide the leadership for the government as a whole.

Senator HASSAN. I think that makes a lot of sense, and, when you think about emerging technologies generally we tend to focus on the potential of the technology without thinking about the nec-

essary safeguards, in my view, early enough. I think a lot of what you have heard around the dais this morning is concerns about the safety issues and the way that AI can impact our democracy, but we need to be thoughtful about how we can actually address those issues in a way that is consistent with our values. I appreciate your answer.

I do have another question for you, Dr. Parker, on deep fakes, which are obviously images and videos that are generated artificially, and they are becoming increasingly realistic.

Last year, malicious actors deployed a deep fake video of Ukrainian President Zelensky telling Ukrainian soldiers to lay down their arms and surrender to Russia, for example. As artificial intelligence advances, deep fakes like the Zelensky one will become harder to identify and debunk, and I am concerned that in the hands of our adversaries, deep fakes pose a really significant threat.

How do you assess the Federal Government's current efforts to identify and debunk deep fakes? How can the Federal Government prepare for a future with extremely realistic deep fakes?

Dr. PARKER. Thank you, Senator. I am a technologist, and so I think often of these challenges from a research perspective and what we can do in terms of coming up with new approaches that help us address these kinds of challenges. There are some activities in the research space to do things like watermarking, that can allow you to determine how a particular piece of data or an image or a video, where it came from, what we call the provenance of it, what its history is.

Those kinds of approaches, if we could watermark these kinds of images and content in a way that allows us to trace back its origins and where it came from, and is it real or is it not real, that is a step forward to giving us the technical ability to address these kinds of challenges.

On top of that, of course, are the governance approaches, and frankly, I did not do a lot of work myself in government work and the governance of deep fakes. But I think the technical approaches that I mentioned will help, and so certainly an increased attention to those kinds of technologies is helpful.

Senator HASSAN. Thank you, and I am over time. I will follow up with one of the questions I think we will face, Mr. Chair, is how does the government work with the private sector and academic to try to make sure that we are harnessing various ideas and approaches here. Thank you very much.

Chairman PETERS. Thank you, Senator Hassan, and actually, that topic is going to be a future hearing, because it is incredibly important. We will look forward to discussing it at length.

Senator Rosen, you are recognized for your questions.

#### **OPENING STATEMENT OF SENATOR ROSEN**

Senator ROSEN. Thank you, Chairman Peters. I really appreciate you holding this hearing today, and thank you to all the witnesses, for all you do, for your work and your thoughtfulness in this really important area.

Of course, we are talking about standards, and all of us are worried about technology moves faster than we can oftentimes even

adapt to it and figure out what we need to do. Standards are really important. We know that China has explicit plans to become a standards-issuing country, and part of its push to increase global influence, it coordinates national standards work across government and industry. China's strategy, they involve targeting emerging technologies like quantum computing, big data, 5G, artificial intelligence, where the global rules really have yet to be defined.

In order for the United States to remain a leader in AI and maintain a national security edge, our response must be one of leadership, coordination, and above all, cooperation, and this means working, like Senator Hassan said, with the private sector and academia, investing in R&D for emerging technologies, coordinating with relevant agencies, and engaging with international standards-setting bodies.

Mr. Ho, can you describe the importance and impact of U.S. participation in these international standards-setting bodies for the development of emerging technologies, including AI?

Mr. HO. Thank you for that question, Senator Rosen. I think you may have been one to introduce legislation that actually fosters international cooperation, and I think that is exactly the kind of effort that is required right now in this context of geopolitical competition, where if it is possible to have international cooperation schemes with like-minded countries, there is a way really to address this current question of the concentration of who really builds, owns, and guides these kinds of AI systems.

One proposal that is very much in line with what you proposed earlier is the proposal for the Multilateral AI Research Institute, to enable like-minded countries to collaborate together, to engage in this kind of standards-setting. I think that is an absolutely critical potential path forward for the future.

Senator ROSEN. Thank you for bringing that up because I do believe by doing nothing we are actually doing something. That leads me to my next question, to Dr. Parker, and again, to you, Mr. Ho.

Earlier this year, the National Institute of Standards and Technology did release an AI Management Framework, with the goal of improving trustworthiness of artificial intelligence, like we are talking about deep fakes and others. About three months later, the White House Office of Science and Technology Policy issued its Blueprint for an AI Bill of Rights.

How should the private sector view these two bills? Are they complementary? Do we need to merge something into one? What are both of your opinions on this.

Dr. PARKER. Thank you for that question, Senator. The NIST AI Risk Management Framework is a framework that applies, in my opinion, to any use case of AI. It gives you a standard approach to be able to consider any given use case of AI, and step through a number of areas as it relates to identifying risk and governing and managing them, and so forth.

The answer there may be that that particular use case of AI is low risk, so the additional steps that might be needed to govern it might be a few, or it could be much more substantial.

The Blueprint for the AI Bill of Rights that the White House OSTP issued is coming at the challenge from a very specific category of applications of AI, and these are applications of AI that

may harm individuals or community groups or society in terms of your civil rights and civil liberty and privacy, your access to resources, and that type of use case.

After applying the NIST AI Risk Management Framework, high-consequence risks were identified, and particularly that affect individuals or, again, communities or society, then the Blueprint for the AI Bill of Rights would be a way to think through what are the rights of an individual and what are those issues that need to be addressed.

Senator ROSEN. Those are complementary. Thank you. Mr. Ho.

Mr. HO. Yes, I think it is a really important question. There are a lot of commonalities in terms of the principles that the NIST AI Risk Management Framework, the Blueprint for an AI Bill of Rights, and Executive Order 13960 are trying to get at.

What I have seen in a number of agencies is a real struggle of how to actually bring that into practice when agencies are thinking about piloting, evaluating, and implementing these kinds of AI use cases. That is why I think what is really critical, and what I highlighted in my opening remarks, is to build pathways for technical talent into the public sector. It will not be possible to really do the kind of red-teaming, evaluation of these kinds of use cases unless we build on existing short-term programs like the Presidential Innovation Fellows (PIF), the GSA, things like the Intergovernment Personnel Act (IPA) mechanism, but also think about long-term pathways, things like the U.S. Digital Service Academy, and why it is so important that OPM actually create the AI hiring line that was due as of July of last year.

Senator ROSEN. You did not really know my next question, did you, because it really is on the workforce challenges, so you set me up just perfectly there. Our existing cyber workforce shortages, we know at Federal agencies they have a really significant impact on our national security, so the private sector trying to hire for AI, for cybersecurity, everything in the technology space. There is really this huge gap between talent and the jobs that we have to fill, so we have to continue to invest in our cyber workforce.

I am going to ask both of you again. Dr. Parker, how do you think we can use AI in the short term to overcome cybersecurity skills shortages across Federal agencies, knowing that I believe AI potentiates what humans can do and may help us do things faster, but in the end humans need to make those decisions. But what do you think we can do while we are building the pipeline?

Dr. PARKER. I do think that there are some uses of AI, some applications of AI that are increasing productivity of individuals. That is true for many different kinds of areas. You asked about cybersecurity, in particular, or just the cyber workforce in general. I think being able to use AI, again, in ways that are somewhat mundane, but there are ways that we can manage a lot of paperwork and be able to identify ways that we can more efficiently address the needs of the American consumer.

I do think it is challenging at the moment to say that we can use AI as a substitute for people. I think AI is not ready for that. I think AI is very much more a collaborative tool, and as a collaborative tool it can help, again, people to work more efficiently. At

the same time, we desperately need to work on getting more expertise into government.

One quick way of doing that, I think, is to leverage these programs like the Intergovernmental Personnel Act and the Presidential Innovation Fellows Program to get people from industry and from academia into government. I think that is a very quick way to leverage those programs more, to get more expertise in government.

Senator ROSEN. Thank you. I am over my time, so do you want me to let him finish or take it for the record later, Mr. Chairman.

Chairman PETERS. Be quick.

Mr. HO. I will be quick. One estimate has it that we need 40,000 positions in the public sector for cybersecurity. It is absolutely critical to figure out the pathways of bringing people in.

I think it is not just a matter of salary scales. As I mentioned in my opening remarks, the National Security Commission on AI really highlighted that what is really needed is also providing opportunities for technical talent to perform meaningful work within government.

The last thing that I will just say quickly, as to why really the way I tend to think about it is augmenting the existing Federal workforce rather than displacing them, is that you are always going to ultimately need a human in charge.

I will give you one example going back to 1983, when there was an automated missile detection system that the Union of Soviet Socialist Republics (USSR) was using, that indicated that there were multiple missiles being fired. There was one individual by the name of Petrov who went against strategic operating protocol because he had a hunch that the system was malfunctioning. He is often said to be the person who actually saved us from nuclear war.

We have to have humans in charge to understand the limitations of these kinds of systems, who know when humans should override them, in order to really work effectively and safely with these kinds of systems.

Senator ROSEN. Thank you. I could not agree more.

Chairman PETERS. Thank you, Senator Rosen.

Senator Padilla, you are recognized for your questions.

#### **OPENING STATEMENT OF SENATOR PADILLA**

Senator PADILLA. Thank you, Mr. Chair. I wanted to thank you also for your flexibility. I am back and forth with the Judiciary subcommittee on AI this morning as well.

Artificial intelligence is already transforming how government agencies serve the public. Federal agencies are leveraging AI in a number of important ways, including to support disaster response and emergency management efforts, to detect financial fraud and identity theft, to offer chatbots and virtual assistants that enhance customer service and engagement, as well as for environmental monitoring and conservation.

I lay this out because a lot of times we are trying to have these conversations about AI, the potential, some risks, and things to keep in mind, as if it is a far-off-into-the-future dynamic. No, this is a here-and-now conversation.

But clearly there are risks that we will need to address, especially the sensitive context of providing government services. Automated decisionmaking systems and tools risk exacerbating the many existing inequities in our society, as some of the testimony today reflects. Again, Mr. Chairman, thank you for calling this hearing. I look forward to hearing from our panel today on how we can ensure that fostering trustworthy, equitable, and accountable applications of AI in government can be achieved.

Now in reviewing the published inventory of AI use across the Federal Government—again, today, not into the future—I was pleasantly surprised to see that many agencies are building AI tools in-house. Some argue that it is more cost effective for our government to simply purchase AI products developed by the private sector, whether it is off the shelf or customized. However, many of the testimonies that we have heard today explicitly mentioned that government technologists and in-house tools are important to ensuring that they comply with the relevant regulations and have the focus of furthering agencies' mission, not always achieved when you contract out.

Mr. Eppink, first question. In your testimony you highlighted some of the tensions that arise when government relies on proprietary private sector tools that are not always available for public scrutiny. What are some of the factors that should go into agency decisions to build a tool internally as opposed to procuring it externally?

Mr. EPPINK. Thank you, Senator, for the question, and you are absolutely right. This is definitely a here-and-now problem, and we do have these frameworks that have been discussed today, but what we do not have, at least that I have observed, is the on-the-ground enforcement and policing of those, and that is especially true when it comes to these in-house versus proprietary systems.

This should be an easy question, I think. We cannot allow proprietary interests—when we are talking about government making decisions about individuals and their families—to hold due process rights hostage. To the extent that a government agency or a State-funded agency needs to use, or wants to use automated decision-making if they have gone through the process, a transparent process of deciding that that is appropriate in the first place, if there is going to be proprietary or private interest providing any of that, the price of admission is going to have to honor Americans' due process rights, equal protection rights, which would mean the transparency of those proprietary systems, so that my clients, for instance, in Idaho, can access the information about the data, the methods, and the processes involved so they can evaluate whether they need to challenge those government decisions, and challenge them and appeal them if they have to.

Senator PADILLA. Thank you. For the follow-up question I invite Dr. Parker and Professor Ho to chime in. What are some of the short-term steps Congress can take to help agencies hire and retain technical talent in this area? We heard from Senator Rosen about the field, in general, I think more in the private sector the need for some of the workforce challenges in this space. But on the government side, in particular.

Dr. PARKER. Thank you for that question. I think there are a number of barriers right now to having that type of expertise in government. One of the barriers, I think, is a salary barrier, and an understanding of the kinds of skills and knowledge that are needed in order to fill a particular AI role. I think if the occupational series that OPM is working on were developed, it would help us to identify what those skills and roles and knowledge are that are needed for certain jobs in AI in the Federal Government, and that would give us more ability to reach out to people that have those kinds of skills and to train to those skills through things like boot camps.

A few years ago there was a boot camp in cybersecurity that brought in people from across the Federal Government who were interested in learning cybersecurity skills but did not yet have them. The Chief Information Officers (CIOs) got together and trained these people, and they now have those additional cybersecurity skills. We could do something similar in AI.

The challenge is always how to scale it. Congress could fund these kinds of boot camps so that we could use the current workforce that we have but to provide those new skills and knowledge that is needed to participate in the AI space.

Senator PADILLA. Professor Ho, anything unique to add, because I have an additional question on a different topic for you as well.

Mr. HO. I will just say, quickly, as I see on this side of the coast, Senator, I will say quickly the other part is to build on Science, technology, engineering, and mathematics (STEM) education. I think for the longest period of time the United States has just been a magnet for top scientific and technical talent. But increasingly there are international students that are choosing to leave the country. The country of Canada, for instance, has had very specialized programs to actually attract top AI talent into the country, and that would be another mechanism really to ensure that we maintain U.S. leadership in AI.

Senator PADILLA. You frame it as many graduate students choosing to leave the country. In too many instances there is no choice, given our immigration system, the need for modernization, but that may be a topic for another day.

Professor Ho, I mentioned another topic I wanted to make sure to ask a question on. In the paper that co-authored on the use of artificial intelligence in Federal administrative agencies, you found that law enforcement applications were the most common use case for Federal agencies that adopted AI systems.

As you talk about building trust and ensuring accountability for automated decisionmaking, law enforcement activity is often the area in which the public has the least amount of insight, let alone oversight, yet it is also the area where government decisionmaking, including tools used to guide that decisionmaking, have tremendous consequences for the public, particularly for historically marginalized communities.

Professor Ho, can you speak to the unique challenges of guarding against bias and ensuring accountability and equity in the use of AI by law enforcement?

Mr. HO. Thank you, Senator, for that question, and thank you for noting the report that we conducted a number of years ago. I

should clarify that in that report we looked largely at the use of AI by civil agencies, and one of the challenges has actually been exactly to get transparency within the criminal justice system. That is reflected, as well, in some of the exemptions under, for instance, the AI Bill of Rights or the scope of requirements to file AI use case inventories. That is why I think these AI use case inventories are such a kind of critical tool of transparency. You cannot manage what you cannot measure, and so I think AI use case inventories are an important first step in that regard.

Let me loop back, though, to the earlier question that you had, just in terms of accountability, and give you one example of the sort of enforcement context that I think is very telling that connects this to the necessity to have human talent within government. Within the Securities and Exchange Commission (SEC) there was an internal team that built out a series of very innovative ways to scan filings, to look at the risk, for instance, of insider trading based on what folks had filed to the SEC, using natural language processing.

But what was very important in that use case was that it was internally developed first, and it was the interaction between the technical engineers and the line-level prosecutors that really ensured accountability. The line-level prosecutors within the SEC said, "I am not persuaded by your risk score. You need to explain to me why this system is actually identifying this case as something that I should prosecute, because ultimately I need to bring this in front of a judge." That is exactly the reason why technical talent and the ability of technical talent to work with the domain experts is going to be so critical to have these kinds of forms of internal accountability for AI systems.

Senator PADILLA. Thank you very much, and Mr. Chair, if I could just add one last note, not another question but just one last note. Tapping into my prior experience as Secretary of State of California, where we introduced the first chatbot in California State government to assist people navigating our website, also automating how business owners and entrepreneurs could file their necessary paperwork with the State of California, there is, at times, a personnel concern that through additional automation and efficiency that we do not need as many workers, and therefore it is a mechanism toward layoffs and staff reductions.

What we were able to do instead was actually free up people's valuable time from, "pushing paper" to improve customer service and dedicate that experienced frontline personnel to some of the more complex questions or troubleshooting that a lot of either individuals or customers of government sometimes have frustrations with, long wait times, et cetera. There can indeed be that win-win.

Thank you, Mr. Chair.

Chairman PETERS. Thank you, Senator Padilla.

Senator Ossoff, you are recognized for your questions.

#### **OPENING STATEMENT OF SENATOR OSSOFF**

Senator OSSOFF. Thank you, Mr. Chairman, and thank you to our panelists for joining us today.

Some of these questions I think will get at implementation within government and some will also touch upon how we think about

potential proposals for broader regulation of the technology, so please bear with me as we move through the discussion.

First is about definitions, and maybe Dr. Parker, I could start with you, and I would also like to hear from Professor Ho. But as these technologies become more ubiquitous and modular and they are incorporated in software suites, and they are plugged in as tools for various purposes throughout an IT infrastructure, how, both in terms of administering within government, their deployment, but also thinking about broader regulation do we fundamentally define what it is that we are regulating? As concisely as you can, how would you define the scope of the technologies that are the subject of our interests and require additional scrutiny?

Dr. PARKER. Certainly the definition is something that people do not agree on, in general. I like to focus on use cases and how the use of a technology is important in impacting the particular application domain.

I look at it in terms of systems that are typically data driven, they often learn over time and change their behavior over time, and they are often doing tasks that we frequently have attributed to require human intelligence in the past. That is my succinct definition.

Senator OSSOFF. OK. Thank you, Dr. Parker. Professor Ho, again, imagine you are legislative counsel and you are drafting legislative text, and your purpose is to define the technology subject to certain regulations. What text would you propose?

Mr. HO. I very much agree with Dr. Parker that the focus should be on use cases. I think a lot of regulations refer back to the National Defense Authorization Act (NDAA) definition, which is a relatively expansive one. One way to handle this, I actually think it would be to have further guidance and clarification coming out of the relevant offices that are working with agencies. For instance, the inconsistency in how AI use case inventories have been handled steps exactly from your question.

Senator OSSOFF. Is it about use case or is it about capability? The use case is going to be almost universal, I think, within a couple of years. Is it the purpose for which it is being used that defines whether it constitutes the regulated technology, if we are thinking about regulation, or is it the capability of the technology as the qualities of the software?

Mr. HO. I think, definitionally, we can go with the capabilities of the software, and that is very much, I think, I would agree with Dr. Parker that the kind of focus typically has been on machine learning systems that are able to learn and predict in the way that really simulates what humans could do.

But I think when it comes to the actual regulation, which I think you are also trying to get it, it is really important to look at the particular use cases to identify what kind of risks are being posed.

Senator OSSOFF. All right. Let us think about some of the constitutional and other questions that may arise. Mr. Eppink, from your standpoint, thinking about, for example, evidentiary predicate that justifies certain police actions or can be used to secure certain court orders. One of the things that I have been grappling with is given the massive datasets that are available in open source or that can be purchased by a state actor, the capacity for predictive

behavioral modeling is potentially very significant. Perhaps with a high degree of integrity some of this technology will be able to assign probabilities with respect to future conduct by individuals. The risk here, I think, is that prosecutors or law enforcement agencies may use such predictive modeling in order to justify forms of surveillance or to seek warrants or to take other action.

How do you think about that risk and how should Congress think about that risk?

Mr. EPPINK. Thank you, Senator. Yes, especially based on what we have heard today, we have, I think, at best, principles to start to think about how to build these systems in a safe way, in a democratic way, but I do not believe they are being built that way right now. Especially when we are talking about law enforcement and the risk of life and liberty, the technology may be developing but the governance of that technology is not yet there to be condoning use of AI in those contexts. I am not even sure, based on what we have seen in Idaho in the Medicaid context of comparatively very rudimentary automated system, that we are there yet, and I think the necessity of litigation there corroborates that.

We have to go beyond these jumping-off points that we have been discussing today, create clear governance, and include the people who these systems will be making decisions about in the process of selecting whether there will be automated decisionmaking in the first place and how those things will be crafted to ensure that the systems are built fairly and used fairly.

Senator OSSOFF. Dr. Parker, how do you think about the due process concerns, and those could be both in a criminal context or in a more seemingly mundane administrative context, like eligibility for certain forms of aid?

Dr. PARKER. I think this is a complex question, Senator, and thank you for that.

Senator OSSOFF. You have one minute and nine seconds.

Dr. PARKER. I am not going to answer it all. I am going to suggest that one way to move forward is for the Subcommittee on AI in Law Enforcement that was directed by Congress in the National AI Initiative Act, to actually be established. This is something that is a subcommittee of the National AI Advisory Committee, and that would be experts from across the different sectors of interest, whether it be private sector or—

Senator OSSOFF. OK. So more study needed.

Dr. PARKER [continuing]. A group of experts that can enable us to understand this more deeply.

Senator OSSOFF. OK. Big subject, seven minutes.

My final question for you, Dr. Parker—and you have probably thought about this a lot, now and in your past roles—in government there is a lot of focus on insider threats, whether in an intelligence context or otherwise. When we are thinking about autonomous actors, or actors who have a measure of autonomy, within public agencies, who are not humans, how do we think about the risk of cooptation, manipulation, exploitation—there is cybersecurity aspect here. There is a software design aspect here—that these tools themselves could pose certain insider threats for unauthorized disclosure or for using their network accesses to enable penetration of government systems, or otherwise.

Dr. PARKER. I think there is a question of what could be and what is true today. I do not work in the national security space myself, but I think what we can do with AI systems right now are there are technologies, if you are thinking about a human insider threat, there are technologies that can actually track behaviors and determine whether or not people are doing what they should be doing in the intelligence system space. I am talking about inside the trusted people. There are those kinds of technologies that can help us protect against the human insider threat.

But then there is also the bigger question of down the road if these more general AI systems have the capability to dig into our systems and what we do about that. That is not something that I think is the here and now. I think it is something that we need to think ahead on. But I do not believe those systems exist today.

Senator OSSOFF. Mr. Chairman, could I ask one more question? I am over time.

OK, please, as succinctly as you can, Dr. Parker, and then Mr. Eppink, if there were three State actions, things that governments do, where, referencing what Professor Ho said earlier, we not only must ensure that humans are in the loop, but let us say we wanted to take a provisional decision that at this time no such action should be permitted to be even influenced by this kind of technology. Like for example, decisions about when to use lethal force, decisions about when to conduct surveillance.

Are there certain governmental functions that given the level of risk right now you would nominate to rule out as being supplemented or guided or supported by these technologies, or are there not?

Dr. PARKER. The launching of nuclear weapons, for instance. I think DOD is overseeing a number of ethical uses of AI, but there are some cases like that, that I think strategically cannot be allowed to happen. That would be my No. 1 on the list.

Senator OSSOFF. Mr. Eppink.

Mr. EPPINK. Yes, I think we can look to the Constitution. We can look to life, liberty, property, privacy. To the extent that we have systems that are involved in making decisions on those things that affect individuals and families, we have to ensure that there is the transparency, the inclusion, the reliability, the independent auditing and testing of those programs before they should be deployed in government uses.

I think we can get there, in some instances, but we are not there now, and especially when you were talking about making decisions about use of lethal force or other decisions that could take a life or threaten a life. I think there is governance that is not yet in place to ensure those systems are fair.

Senator OSSOFF. Thank you. Thank you, Mr. Chairman.

Chairman PETERS. Thank you, Senator Ossoff.

We are ready to wrap up a great hearing. We have covered an awful lot of ground, a big topic. I would like to wrap up these kinds of hearings with just asking folks to kind of focus one of the issues.

I am going to be asking everybody on the panel—and I am going to start with you, Mr. Siegel, and then we will work down the panel—so in terms of AI policy going forward, what would you say, from a legislative standpoint, because we are legislators here, in

Congress, in terms of legislating, what would be the No. 1 item that we should prioritize in terms of thinking about future legislation?

I am going to start with you, Mr. Siegel. All of you have a lot of different issues that you are working with, but No. 1, for legislating. There are other things that we could be doing, regulations and other kinds of policies, or norms, et cetera, but from a legislative standpoint is there something that really stands out to each of you?

Mr. Siegel.

Mr. SIEGEL. Understood, Mr. Chairman. Transparency would be the No. 1 issue from a legislative point of view, enforcing transparency in the use of AI. I think that if you were to focus legislatively on transparency that would take care of a host of attendant issues such as collusion between the government and corporate sector, privacy issues, even potentially pointing the way toward some kind of private data ownership so that the uses of people's data was something that if they were not at least initially financially staked into, they could at least have some visibility on.

Transparency, I think, is where to begin, and from there that would lead into other related matters.

Chairman PETERS. Thank you, Mr. Siegel. Mr. Ho.

Mr. HO. Chairman, a number of months ago, Eric Schmidt was asked about how Congress had implemented the recommendations from the National Security Commission on AI, and he was generally quite favorable. But there was one thing he singled out, which is getting technical talent into the Federal workforce, which is the predicate for, I think, addressing a range of these issues that we have talked about today.

Take procurement. I think the AI Training Act is fantastic. But what I have seen at a number of agencies is that procurement officials say, "It is the business unit within the agency that has to make this decision," and the business unit says, "It is really up to the procurement official." I think what we really need is the kind of blended expertise that brings domain expertise and technical talent together to really be able to protect American values.

Chairman PETERS. Thank you. Dr. Parker.

Dr. PARKER. Thank you, Mr. Chairman. I would say that we are suffering right now from a lack of leadership and a lack of prioritization on these topics. One, I think, quick way, legislatively, that could be done to address this is to appoint those Chief AI Officers at each agency, where they are given the responsibility and the resources to oversee the uses of AI and to develop strategies for their use of AI within their agencies, with accountability for delivering those and updating them regularly.

To establish a coordination body, like a Chief AI Officers Council, that is responsible for coordinating these activities across all the Federal Government, with an intent of prioritizing this activity across the Federal Government, providing that leadership as well. Then they can also work on helping with the workforce issues, I believe, within their agencies, by identifying what the needs are and putting those priorities on getting that workforce into the Federal Government.

Chairman PETERS. Thank you. Mr. Ariga.

Mr. ARIGA. Thank you, Chairman. My No. 1 priority, and you might indulge me in the hyphenated response, one is on disclosure where discretionary decisionmaking is being impacted. The use of it, the process, and any redress. But fundamentally it is the digital-ready workforce that will make such disclosure effective. I will echo with Dr. Ho's recommendation in terms of developing a Federal digital-ready workforce.

Chairman PETERS. Thank you. Mr. Eppink.

Mr. EPPINK. Thank you, Chairman. The experts on these systems are the people that the systems make decisions about. I worked on this litigation in Idaho for 12 years now, almost, and I have worked with agency officials, I have communicated with Federal overseers, I have communicated with the courts. But it is time and time again my clients who have been able to spot the most important systemic problems with these systems.

I appreciate the opportunity to testify to the Committee today. The people that the Committee especially should hear from and the policies that the Congress should put in place should make sure that at each step in the development and selection of these systems, the people who they make decisions about are included it. My clients, people with disabilities in Idaho, often use the phrase, "nothing about us, without us," and that is critical in these automated decisionmaking systems and AI systems. Thank you.

Chairman PETERS. Thank you. I would certainly like to thank each of our witnesses for being here today, and I think, the entire Committee is grateful for your expertise and for your willingness to come forward to answer our questions.

I think as we heard today the use of automated systems to help government provide public service more efficiently is not new. However, as we enter the age of rapid development of advanced machine learning methods and other forms of artificial intelligence now—not waiting, now—is going to be the time to ensure that these systems that the government is using and will procure for use in the future do not have unintended, harmful consequences.

It is critical the Federal Government act quickly to set appropriate guardrails and oversight policies that protect the public. I think we all agree that Americans deserve a government that is modern, efficient, innovative, as well as one that is transparency, fair, trustworthy, and protects their privacy. As Chairman of this Committee I am going to continue to work to ensure that the government lives up to those key principles, and your testimony today will help inform the Committee's future legislative activities and oversight actions in the years ahead.

The record for this hearing will remain open for 15 days, until 5 p.m. on May 31, 2023, for the submission of statements and questions for the record.

This hearing is now adjourned.

[Whereupon, at 11:54 a.m., the hearing was adjourned.]

# A P P E N D I X

---

## **Chairman Peters Opening Statement As Prepared for Delivery Full Committee Hearing: Artificial Intelligence in Government May 16, 2023**

Today's hearing is the second in a series that I plan to convene on artificial intelligence. At our first hearing in March, we discussed the transformative potential of AI, as well as the possible risks these technologies can pose.

Today, we'll be discussing how AI has the potential to help government better serve the American people, such as by improving the ways agencies deliver services, and what pitfalls we need to be aware of as government increasingly adopts AI tools.

The federal government is already using AI in an effort to provide more efficient services, assess potential security threats, and automate routine tasks to enhance the federal workforce.

Earlier this month, the White House announced new efforts to invest in American leadership to develop AI technologies and promote the responsible use of AI within the federal government. Later this summer, the Office of Management and Budget is expected to release new guidance on federal government use of AI, implementing legislation this Committee advanced in 2020 and was later signed into law in government funding legislation.

U.S. leadership in the development and use of AI systems, by both the private sector and government, is essential for our global economic competitiveness. We should work to ensure that government can adopt and deploy these tools to help improve American lives, but as we do so, we must ensure we are also prepared to address the potential risks and harms AI systems can present.

The potential for bias in AI applications can have serious consequences for federal government use.

A recent study found that an algorithm used by the IRS to determine who should be audited was erroneously more likely to recommend Black taxpayers than white taxpayers, and the government wasn't prepared with the data or training necessary to recognize this biased outcome.

As we heard in our last hearing, AI algorithms often lack transparency and accountability for how they arrive at certain outcomes. Even the engineers who design them do not always understand how they reach certain conclusions.

In government applications, this can present serious risks to Americans who may unknowingly be interacting with an AI, and who may struggle to get answers about why an AI system made a certain determination.

For example, at least a dozen states deployed algorithms to decide eligibility for disability benefits, which resulted in denying thousands of recipients this critical assistance that helped

them live independently, and left them with little opportunity to understand why the decision was made or how they could appeal it.

The enormous amounts of data that can be collected as a result of using AI systems also presents concerns about privacy. Existing privacy laws did not envision these types of applications.

As agencies use more AI tools, they will need to ensure they are securing and appropriately using any data inputs to avoid accidental disclosures or unintended uses that harm Americans' civil rights and civil liberties.

Finally, we must ensure our federal workforce is ready to procure and oversee the use of AI systems to ensure they are benefitting Americans. Last Congress, I authored legislation that was signed into law requiring officials charged with procuring AI tools to be trained in both their capabilities to improve agency missions, and their potential risks, to ensure responsible use.

Last week, I introduced bipartisan legislation to build on that effort by requiring federal agency supervisors and managers to receive similar training.

I'm looking forward to today's discussion, and to continuing to work with my colleagues on both sides of the aisle to advance solutions that will help encourage American development of AI, and ensure its being used appropriately.

During today's hearing, we'll be discussing some of those strategies, including the need to conduct inventories of current federal government AI applications, requiring ongoing audits to ensure the accuracy and effectiveness of AI systems, and considering responsible standards that need to be met as the federal government continues to acquire additional AI tools.

I'm grateful to our expert witnesses for joining us to help tackle this rapidly evolving issue.

Opening Statement of Ranking Member Rand Paul

May 16, 2023

In 1975, the late Senator Frank Church said, “The United States government has perfected a technological capability that enables us to monitor the messages that go through the air.

That capability at any time could be turned around on the American people, and no American would have any privacy left, such is the capability to monitor everything. There would be no place to hide.”

These words came as Senator Church lead the Senate Select Committee to Study Governmental Operations with Respect to Intelligence Activities, better known as the Church Committee.

The Church Committee’s 1976 final report exposed numerous secret federal government programs that violated the constitutional rights of American citizens it deemed to be threats to existing social and political order.

These programs surveilled and targeted individuals like Martin Luther King Jr. and domestic organizations like the Southern Christian Leadership Conference, as well as infiltrated movements to incite rivalries and discredit leaders.

Nearly fifty years later, Senator Church’s ominous warning that the government could weaponize technology against the American people reads more like a premonition. There is truly becoming “no place to hide”.

In recent decades, journalists and whistleblowers exposed examples of our government leveraging emerging technologies to violate the privacy and civil liberties of its citizens.

Intelligence agencies conducted surveillance of video game users collecting data on the contents of communications between players.

The Department of Homeland Security tracked the live locations of individuals and groups participating in the Black Lives Matter movement.

The Drug Enforcement Administration conducted “covert surveillance” of people protesting the death of George Floyd.

And it’s only getting worse. Just last month, the ACLU acknowledged “The Biden administration has been quietly deploying and expanding programs that surveil what people say on social media, using tools that allow agents and analysts to invisibly monitor the vast amount of protected speech that occurs online.”

And how are they doing it? Using artificial intelligence.

For years, federal agencies including the Department of Homeland Security, State Department, National Science Foundation, and FBI have been colluding with private organizations and social media companies to combat what they deemed to be “disinformation”.

The purpose, so they claimed, was to combat foreign malign influence.

But in reality, the government wasn’t suppressing foreign “misinformation”. It was working to censor domestic speech by Americans.

Since 2020, the federal government has awarded over 500 contracts and grants related to "misinformation" or "disinformation."

While the grant awardees and their proprietary AI and machine learning technologies differ, their goals are consistent: to "mine" the internet, identify conversations indicative of “harmful” narratives, track those "threats," and develop countermeasures before messages go viral.

One NSF-funded company’s mission statement claims that “social media is being manipulated and ideas are being spread uncontrollably online.”

The solution it provides? An automatic controversy detection algorithm to help identify things that are “potentially opinion-shifting” in order to make communication “more productive and less dangerous”.

During the COVID-19 pandemic, we witnessed the accelerated use of AI technologies to monitor and suppress public debate on issues like natural immunity, masks, and the origins of the virus.

Multiple federal agencies, including DOD and the State Department, funded automated disinformation-detection technologies designed to monitor and suppress public debate on issues like vaccines and the origins of COVID-19.

Writer Jacob Siegel in a fantastic, yet haunting, narrative explaining the last decade of US government domestic censorship efforts said, “Disinformation, now and for all time, is whatever they say it is.

That is not a sign that the concept is being misused or corrupted; it is the precise functioning of a totalitarian system.”

Make no mistake, the United States is engaging in the same activities we criticize other countries for.

But unlike China and North Korea, the United States government attempts to conceal its involvement using private entities as front companies to do its dirty work.

But make no mistake, the intent is the same. Control the narrative, eliminate dissent, and retain power.

This should terrify all Americans. The government is using your hard-earned tax dollars to surveil and censor your protected speech.

Artificial intelligence is only going to make it easier for the government to do this, and harder to detect.

This should not be a partisan issue. We must get to the bottom of how the federal government uses AI to violate the privacy and civil liberties of the American people before it's too late.

Testimony of Lynne E. Parker<sup>1</sup>  
Associate Vice Chancellor and Director of the AI for Tennessee Initiative  
The University of Tennessee, Knoxville<sup>2</sup>

Before the United States Senate Committee on Homeland Security and Governmental Affairs  
Hearing Entitled “Artificial Intelligence in Government”

May 16, 2023

Chairman Peters, Ranking Member Paul, and members of the Committee, thank you for inviting me to testify at this hearing on artificial intelligence (AI) in government. I am an Associate Vice Chancellor at the University of Tennessee, Knoxville (UT), and Director of the AI for Tennessee Initiative. I am also a professor of computer science at UT and have had a 30+ year research career advancing the fields of AI and robotics. Prior to my current role, I served for four years across two administrations in the White House Office of Science and Technology Policy (OSTP) as Founding Director of the National AI Initiative Office, Deputy United States Chief Technology Officer, and assistant director of AI. My focus in these OSTP roles was on the development of AI policies bolstering research, governance, education and workforce training, international engagement, and the Federal government’s use of AI.

My remarks today focus on ways that U.S. Federal agencies can better leverage and govern the responsible use of AI in advancing their missions and providing services to the American people.

#### Current Use of AI by Federal Agencies

Federal agencies are currently using AI in many ways to enhance their operations, improve the quality of their services, train their workforce, and much more. These uses of AI are becoming increasingly transparent as agencies comply with Section 5 of Executive Order 13960,<sup>3</sup> “Promoting the Use of Trustworthy AI in the Federal Government.” This executive order requires agencies to conduct annual inventories of their AI use cases and share them with other government agencies and the public, to the extent practicable and in accordance with applicable laws and policies. Many federal agencies have now complied with this directive, making their inventories available on AI.gov.<sup>4</sup> As of May 10, 2023, twenty departments and agencies have catalogued their use cases, coordinated by the U.S. Chief Information Officers Council, which provided implementation guidance for creating these AI use case inventories.<sup>5</sup>

Much can be learned from these AI use case inventories, including an understanding of common ways that Federal agencies are leveraging AI and implications for ensuring the responsible use of AI across a wide range of missions. As is clear from a cursory review of these AI use case inventories, not all AI uses

---

<sup>1</sup> The opinions expressed in this testimony are solely those of the author and do not necessarily represent those of the University of Tennessee nor any other organization with which the author collaborates.

<sup>2</sup> The University of Tennessee, Knoxville, is Tennessee’s flagship university and premier public research institution. It is classified by the Carnegie Classification of Institutions of Higher Education as producing very high research activity (R1 category) and is a Carnegie Community Engaged university. The vision of UT is of a world enriched by our ideas, improved through our action, and inspired by the Volunteer spirit of service and leadership.

<sup>3</sup> Executive Order 13960, 85 FR 78939 (December 3, 2020).

<sup>4</sup> <https://www.ai.gov/ai-use-case-inventories/> (accessed on May 10, 2023).

<sup>5</sup> <https://www.cio.gov/policies-and-priorities/Executive-Order-13960-AI-Use-Case-Inventories-Reference/> (accessed on May 10, 2023).

require the same level of governance and oversight. For example, Federal agencies are currently using AI to process large amounts of paperwork, automate routine tasks, streamline processes for grant applications, improve customer service via chatbots, review solicitations for regulatory compliance, prevent or detect cyberattacks, secure access to sensitive facilities, and advance mission-specific goals such as predicting hurricane paths, building maps to monitor the Nation’s forest resources, and scheduling predictive maintenance. These use cases do not all raise the same level of concerns regarding the protection of privacy, civil rights, and civil liberties, since many uses do not directly impact an individual’s rights, opportunities, or access to critical resources or services—areas of emphasis for OSTP’s Blueprint for an AI Bill of Rights.<sup>6</sup> But the extreme variety of use cases does create challenges for developing a flexible approach to the responsible governance and use of AI by the Federal government.

In this context, a risk-based approach is needed for determining how best to govern each Federal use case of AI. As directed by the National AI Initiative Act of 2020,<sup>7</sup> the National Institute of Standards and Technology (NIST) published the AI Risk Management Framework (AI RMF 1.0) on January 26, 2023,<sup>8</sup> which now provides a comprehensive and flexible framework to manage the risks of AI to individuals, organizations, and society. **I believe that Congress should require all Federal agencies to use the NIST AI RMF during the design, development, procurement, use, and management of their use cases of AI, promoting the responsible adoption of AI.** Federal AI use cases that pose more than insignificant risks will, of course, require additional governance, as discussed in the next subsection. But beginning with a standardized assessment of the risks posed by each use case of AI is a key step that can be taken now by all Federal agencies, without needing to wait for additional guidance. As stated in the NIST AI RMF (page 1), “understanding and managing the risks of AI systems will help to enhance trustworthiness, and in turn, cultivate public trust.”

### Guidance for the Federal Use of AI

The Federal government can set an example for the private sector by demonstrating how AI can be leveraged to advance agency missions and enhance services for the American people, while at the same time protecting privacy, civil rights, and civil liberties. This demonstration would not only highlight the governments’ commitment to responsible AI adoption but also inspire similar responsible practices in the private sector.

The responsible adoption of AI by the Federal government, however, is currently being slowed by the lack of overarching guidance for the design, development, procurement, use, and management of AI by Federal agencies. The AI in Government Act of 2020 and executive order 13960 directed the Office of Management and Budget (OMB) to create guidance for the Federal agency use of AI. I am heartened by the May 4<sup>th</sup> White House announcement that OMB plans to draft policy guidance on the use of AI systems by the U.S. government and release the draft for public comment this summer.<sup>9</sup> Well considered and structured guidance that promotes the innovative and responsible use of AI technologies within the

<sup>6</sup> Blueprint for an AI Bill of Rights, <https://www.whitehouse.gov/ostp/ai-bill-of-rights/> (accessed on May 10, 2023).

<sup>7</sup> National Artificial Intelligence Initiative Act of 2020 (Pub.L. 116-283) § 5301.

<sup>8</sup> <https://www.nist.gov/itl/ai-risk-management-framework> (accessed on May 10, 2023).

<sup>9</sup> White House Fact Sheet: Biden-Harris Administration Announces New Actions to Promote Responsible AI Innovation that Protects American’s Rights and Safety, May 4, 2023, <https://www.whitehouse.gov/bricfing-room/statements-releases/2023/05/04/fact-sheet-biden-harris-administration-announces-new-actions-to-promote-responsible-ai-innovation-that-protects-americans-rights-and-safety/> (accessed on May 10, 2023).

Federal government while protecting privacy, civil rights, and civil liberties, would go a long way toward achieving the second purpose of the National AI Initiative Act of 2020—for the United States to “lead the world in the development and use of trustworthy artificial intelligence systems in the public and private sectors.” **OMB should prioritize and adequately resource their work on this guidance to ensure that it appropriately addresses the wide diversity of use cases of AI across the Federal government; encourages the responsible adoption of innovative AI to improve public services while protecting privacy, civil rights, and civil liberties; and can be operationalized for practical use by Federal agencies.**

### Individual Agency Responsibilities for Governance and Use of AI

While government-wide policies and implementation guidance for AI will apply to all agencies, individual agencies also have important responsibilities to ensure that they are appropriately implementing Federal guidance for their mission context. **I believe that Congress should require each Federal agency, department, or bureau to have a current and regularly updated AI strategic plan (made publicly available) that includes the agency’s approach to the responsible adoption of AI.** While many agencies already have AI strategies,<sup>10</sup> all agencies should develop these strategies and update them on a regular basis (e.g., every 3 years), to ensure that they remain current as AI technologies rapidly evolve. These strategies should cover all the agency’s activities in AI, whether they be research-oriented pursuits or operational use cases.

**Additionally, it is my belief that Congress should direct each agency to hire and resource a Chief AI Officer (CAIO) who is responsible for overseeing the development and regular update of the organization’s AI strategy, as well as for coordinating the responsible design, development, procurement, use, and management of AI within that organization.** While the CAIO role could be filled (in a dual-hat capacity) by the organization’s Chief Data Officer, Chief Information Officer, or Chief Technology Officer, the scope of required expertise in a Chief AI Officer is different from that of leaders who oversee IT, data, or technology in general. Flexibility should be provided to agencies to fill these roles as is most beneficial to each agency, while ensuring that the appropriate AI leadership is in place. The CAIO would be the “responsible AI official,” as required by executive order 13960, Sec. 8, and would serve as the primary point of contact for interagency coordination on the responsible use of AI within Federal agencies. The CAIO would also coordinate closely with that agency’s representative(s) to the interagency coordination committee described in the National AI Initiative Act of 2020, Sec. 5103, which oversees the research, education, workforce development, and outreach goals of the National AI Initiative.

### Coordination of Federal Agency Use of AI

To ensure the efficient and consistent implementation of AI guidance across Federal agencies, it is important to establish a single coordination body. This body would be responsible for regularly developing and updating AI use guidance, assisting agencies in implementing this guidance, overseeing the annual inventories of agency AI use cases, leveraging insights from the AI use case inventories, and handling other matters related to the practical governance and use of AI by Federal agencies. Previous Congressional action has led to the creation of: (1) the National AI Initiative Office (NAIIO) within OSTP, which coordinates key goals of the National AI Initiative around AI research, education, workforce

<sup>10</sup> See <https://www.ai.gov/strategy-documents/> for links to current national and agency-specific AI strategy documents (accessed on May 10, 2023).

development, outreach, and interagency coordination on these topics, and (2) the AI Center of Excellence within the General Services Administration (GSA), which, together with GSA's AI Community of Practice, facilitates the adoption of AI and improves the cohesion and competency in the adoption and use of AI within the Federal government. However, neither of these bodies is given the role of leading and coordinating Federal agencies in their practical governance and use of AI.

**I believe that Congress should direct the creation of an interagency Chief AI Officers Council (CAIOC) as an effective way to coordinate the governance and use of AI within the Federal government. This Council would be co-led by OMB and OSTP's NAIIO, with representation from the proposed agency Chief AI Officers as well as GSA's AI Center of Excellence and/or AI Community of Practice.** Given OMB's role for managing Federal agencies, including oversight of agency performance, procurement, financial management, and information technology,<sup>11</sup> it makes sense for them to be a co-leader of the CAIOC. Designating OSTP's NAIIO as a co-leader of the CAIOC would leverage OSTP's deep subject matter expertise and role in overseeing interagency activities in research, education, and workforce development; their ability to leverage a strong interagency coordination network; and their ability to convene external stakeholders to gather relevant input from non-government experts. Given the wide range of agency missions, it is important to include broad representation of the agencies. The proposed agency Chief AI Officers would be the experts at each agency who can speak to their agency-specific needs and approaches to the responsible governance and use of AI. Finally, GSA's interagency work via the AI Center of Excellence and AI Community of Practice provides the opportunity to inform and facilitate the development of cross-agency best practices and lessons learned on the responsible use of AI via working groups and related activities that help accelerate the government-wide implementation of responsible AI.

The CAIOC can also help increase efficiencies across agencies by further leveraging the AI use case inventories that agencies have compiled. **The proposed Chief AI Officers Council should review the AI use case inventories for common application areas and identify dozens of key agency processes that could be transformed with AI, in a manner consistent with privacy, civil rights, and civil liberties.** These efforts would further the goals of executive order 14058,<sup>12</sup> "Transforming Federal Customer Experience and Service Delivery to Rebuild Trust in Government," by enabling agencies to learn from each other and build efficient, common, and shared approaches to key processes that improve the delivery of services for the American people. **Congress can accelerate the responsible and innovative adoption of AI within Federal agencies by providing agencies with AI innovation funds as part of their annual operating budgets.**

### Workforce Challenges Around Federal Use of AI

Federal agencies face a shortage of AI workers who have the expertise to design, develop, procure, use, and manage responsible AI applications. The AI in Government Act of 2020 took action to help address this challenge by requiring the Office of Personnel Management (OPM) to establish or update an occupational series for AI. By creating an occupational series for AI, OPM could help agencies identify prospective employees with the skills necessary to promote and govern the responsible use of AI in the Federal government. This occupational series would be particularly helpful since only a few colleges and universities have degree programs in AI, making it challenging to identify those with the skills needed for

<sup>11</sup> <https://www.whitehouse.gov/omb/> (accessed on May 10, 2023).

<sup>12</sup> Executive Order 14058, 86 FR 71357 (December 13, 2021).

the Federal AI workforce. **OPM should prioritize and adequately resource their work on the AI occupational series, so that Federal agencies will be better positioned to strengthen their AI workforces.**

An additional action that could help in building up the Federal AI workforce, along with the Nation's AI workforce in general, would be the development of an AI Workforce Framework similar to the National Initiative for Cybersecurity Education (NICE) Framework that was initially published in 2017 and is updated regularly. The NICE Framework, developed via a multistakeholder process led by NIST, provides an understanding of the tasks, knowledge, and skills needed to perform cybersecurity work. Such a framework is helpful both for the students who want to develop cybersecurity skills, as well as the employers who want to understand the types of skills needed to perform cybersecurity work. **Congress should direct the development of a National Initiative for AI Education Framework, analogous to the NICE Framework, to provide a comprehensive and standardized approach to describing AI roles and the associated knowledge, skills, and abilities needed for those roles.**

An expanded pipeline of skilled AI workers would also help increase the number of potential Federal employees with expertise in AI. One strategy to cultivate AI experts is by promoting their participation in AI research, particularly drawing from underrepresented and under-resourced groups. As directed by Congress in the National AI Initiative Act of 2020,<sup>13</sup> the National AI Research Resource Task Force released their final report in January 2023.<sup>14</sup> This report provides an implementation plan to create widely accessible AI research cyberinfrastructure, including computational resources, data, testbeds, algorithms, software, services, networks, and expertise that would help democratize participation in AI research and development and increase the diversity of AI talent. **Congress should authorize and fund the National AI Research Resource to help strengthen the breadth and diversity of talent in the AI research ecosystem.** Some of this talent would likely choose to use their AI skills to support the Federal government in its adoption and governance of responsible AI.

#### Summary of Recommendations for AI in Government

The responsible adoption of AI by the Federal government can provide many benefits to the American people, if done in a manner that upholds privacy, civil rights, and civil liberties. The following list summarizes my recommendations for steps the Federal government could take to accelerate this process:

- 1) **Use NIST AI Risk Management Framework.** Congress should require Federal agencies to use the NIST AI Risk Management Framework during the design, development, procurement, use, and management of their use cases of AI, promoting the responsible use of AI.
- 2) **Develop Federal AI guidance.** OMB should prioritize and adequately resource their work on creating Federal guidance for the use of AI in government, ensuring that it appropriately addresses the wide diversity of use cases of AI across the Federal government; encourages the responsible adoption of innovative AI to improve public services while protecting privacy, civil rights, and civil liberties; and can be operationalized for practical use by the agencies.
- 3) **Create agency AI strategic plans.** Congress should require each Federal agency, department, and bureau to have a current and regularly updated AI strategic plan (made publicly available) that includes the agency's approach to the responsible adoption of AI.

<sup>13</sup> National Artificial Intelligence Initiative Act of 2020 (Pub.L. 116-283) § 5106.

<sup>14</sup> <https://www.ai.gov/wp-content/uploads/2023/01/NAIRR-TF-Final-Report-2023.pdf> (accessed May 10, 2023).

- 4) **Hire agency Chief AI Officers.** Congress should direct each agency to hire and resource a Chief AI Officer (CAIO) who is responsible for overseeing the development and regular update of the organization's AI strategy, as well as coordinating the responsible design, development, procurement, use, and management of AI within that organization.
- 5) **Create Chief AI Officers Council.** Congress should direct the creation of an interagency Chief AI Officers Council (CAIOC) as an effective way to coordinate the governance and use of AI within the Federal government. This Council would be co-led by OMB and OSTP's NAIHO, with representation from the agency Chief AI Officers, as well as GSA's AI Center of Excellence and/or AI Community of Practice.
- 6) **Leverage AI use case inventories.** The proposed Chief AI Officers Council should review the AI use case inventories for common application areas and identify dozens of key agency processes that could be transformed with AI, in a manner consistent with privacy, civil rights, and civil liberties.
- 7) **Supply AI innovation funds.** Congress can accelerate the responsible and innovative adoption of AI within Federal agencies by providing agencies with AI innovation funds as part of their annual operating budgets.
- 8) **Complete development of AI occupational series.** OPM should prioritize and adequately resource their work on the AI occupational series, so that Federal agencies will be better positioned to strengthen their AI workforces.
- 9) **Create National Initiative for AI Education Framework.** Congress should direct the development of a National Initiative for AI Education Framework, analogous to the NICE framework, to provide a comprehensive and standardized approach to describing AI roles and the associated knowledge, skills, and abilities needed for those roles.
- 10) **Fund National AI Research Resource.** Congress should authorize and fund the National AI Research Resource to strengthen the breadth and diversity of talent in the AI research ecosystem. Some of this talent would likely choose to use their AI skills to support the Federal government in its adoption and governance of responsible AI.

I thank the committee for the opportunity to testify on AI in government.

United States Government Accountability Office

---



Testimony  
Before the Committee on Homeland  
Security and Governmental Affairs,  
U.S. Senate

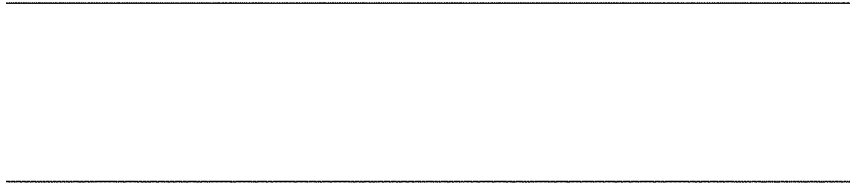
---

For Release on Delivery  
Expected at 10:00 a.m. ET  
Tuesday, May 16, 2023

## ARTIFICIAL INTELLIGENCE

### Key Practices to Help Ensure Accountability in Federal Use of AI

Taka Ariga, Chief Data Scientist, Science, Technology  
Assessment and Analytics



This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

# GAO Highlights

Highlights of GAO-23-106811, a testimony before the Committee on Homeland Security and Governmental Affairs, U.S. Senate

## Why GAO Did This Study

To help managers ensure accountability and the responsible use of AI in government programs and processes, GAO has developed an AI Accountability Framework. Separately, GAO has identified mission-critical gaps in federal workforce skills and expertise in science and technology as high-risk areas since 2001.

This testimony summarizes two related reports—GAO-22-105388 and GAO-21-519SP. The first report addresses the digital skills needed to modernize the federal government. The second report describes discussions by experts on the types of risks and challenges in applying AI systems in the public sector.

To develop the June 2021 AI Framework, GAO convened a Comptroller General Forum in September 2020 with AI experts from across the federal government, industry, and nonprofit sectors. The Framework was informed by an extensive literature review, and the key practices were independently validated by program officials and subject matter experts.

For the November 2021 report on digital workforce skills, GAO convened a roundtable discussion in October 2021 comprised of chief technology officers, chief data officers, and chief information officers, among others. Participants discussed ways to develop a dedicated talent pool to help meet the federal government's needs for digital expertise.

View GAO-23-106811. For more information, contact Taka Ariga, Chief Data Scientist, 202-512-6888, [ariga@gao.gov](mailto:ariga@gao.gov)

May 16, 2023

## ARTIFICIAL INTELLIGENCE

### Key Practices to Help Ensure Accountability in Federal Use of AI

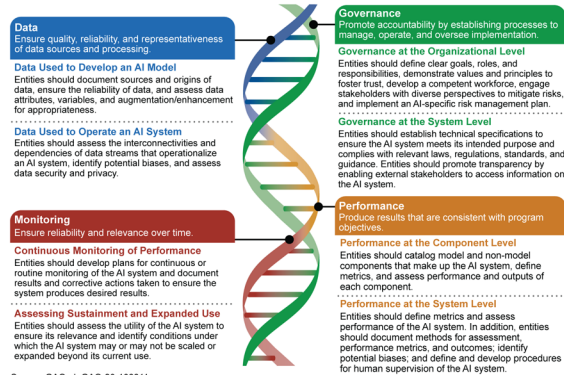
#### What GAO Found

Artificial intelligence (AI) is evolving at a rapid pace and the federal government cannot afford to be reactive to its complexities, risks, and societal consequences. Federal guidance has focused on ensuring AI is responsible, equitable, traceable, reliable, and governable. Third-party assessments and audits are important to achieving these goals. However, a critical mass of workforce expertise is needed to enable federal agencies to accelerate the delivery and adoption of AI.

Participants in an October 2021 roundtable convened by GAO discussed agencies' needs for digital services staff, the types of work that a more technical workforce could execute, in areas such as artificial intelligence, and challenges associated with current hiring methods. They noted such staff would require a variety of digital and government-related skills. Participants also discussed challenges associated with existing policies, infrastructure, laws, and regulations that may hinder agency recruitment and retention of digital services staff.

During a September 2020 Comptroller General Forum on AI, experts discussed approaches to ensure federal workers have the skills and expertise needed for AI implementation. Experts also discussed how principles and frameworks on the use of AI can be operationalized into practices for managers and supervisors of these systems, as well as third-party assessors. Following the forum, GAO developed an AI Accountability Framework of key practices to help ensure responsible AI use by federal agencies and other entities involved in AI systems. The Framework is organized around four complementary principles: governance, data, performance, and monitoring.

#### Artificial Intelligence (AI) Accountability Framework



Source: GAO. | GAO-23-106811

Chairman Peters, Ranking Member Paul, and Members of the Committee:

Thank you for the opportunity to discuss our work on artificial intelligence (AI). My testimony today summarizes two relevant GAO reports: our June 2021 Framework entitled *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*<sup>1</sup> and our November 2021 report on developing a pipeline of federal digital staff, entitled *Digital Services: Considerations for a Federal Academy to Develop a Pipeline of Digital Staff*.<sup>2</sup>

In our AI Accountability Framework, we highlighted that, given the rapid pace at which AI is evolving, the federal government cannot afford to be reactive to AI's complexities, risks, and societal consequences. GAO's objective was to identify key practices to help ensure accountability and responsible AI use by federal agencies and other entities.<sup>3</sup> Foundational to solving the AI accountability challenge is having a critical mass of digital expertise to help accelerate responsible delivery and adoption of AI capabilities. A talented and diverse cadre of digital-ready federal employees is essential to a government that can effectively design, develop, deploy, use, and monitor AI systems. In our Digital Services report, we noted that, as the federal government continues its modernization efforts, it faces a severe shortage of digital expertise, including in the field of AI. Each federal agency is individually coping with challenges in hiring, managing, and retaining staff with digital services skills because of a limited pipeline of candidates and bureaucratic processes.

Various federal guidance have attempted to guide responsible, equitable, traceable, reliable, and governable AI capabilities. At the same time, robust and independent audits are important to ensuring that these goals are achieved. However, as AI technology advances, responsible management of AI systems will be challenging if the skills necessary to successfully develop,

---

<sup>1</sup>GAO, *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*, [GAO-21-519SP](#) (Washington, D.C.: June 30, 2021).

<sup>2</sup>GAO, *Digital Services: Considerations for a Federal Academy to Develop a Pipeline of Digital Staff*, [GAO-22-105388](#) (Washington, D.C.: Nov. 19, 2021).

<sup>3</sup>The Framework is organized around four complementary principles that address governance, data, performance, and monitoring. For each principle, the Framework describes key practices for federal agencies and other entities that are considering, selecting, and implementing AI systems. Each practice includes a set of questions for entities, auditors, and third-party assessors to consider as well as procedures for auditors and third-party assessors.

buy, or use AI capabilities are lacking. In our AI Accountability Framework, we highlight the need to recruit, develop, and retain competent personnel to ensure accountability and responsible use of AI in government programs and processes.

Our AI Accountability Framework distills insights from cross-sectoral 23 experts convened during the Forum on Artificial Intelligence by the Comptroller General of the United States held on September 9 and 10, 2020. The work for the report also included an extensive literature review and independent validation of key practices from program officials and subject matter experts.<sup>4</sup>

For our Digital Services report, GAO convened a roundtable discussion on October 13, 2021 comprised of chief technology officers, chief data officers, chief information officers, and those in similar roles across the federal government, as well as knowledgeable representatives from academia and nonprofits. Additional information about our scope and methodology can be found in that report.

We performed the work on which this testimony is based in accordance with all applicable sections of GAO's Quality Assurance Framework.

## **Background**

### AI Life Cycle

The life cycle of an AI system involves four phases: design, development, deployment, and continuous monitoring.<sup>5</sup> As shown in figure 1, each phase includes considerations articulating the system's concepts, collecting and processing data, building one or more machine learning

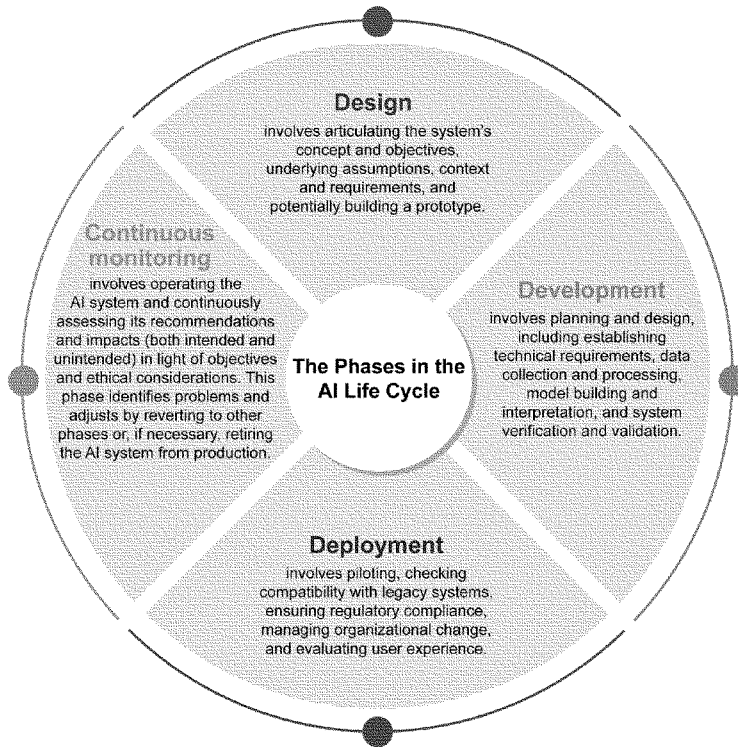
---

<sup>4</sup>GAO currently has work underway on federal agencies' efforts and plans related to AI and the Department of Homeland Security's use of AI. We expect to publish the former in fall 2023 and the latter in early 2024.

<sup>5</sup>See OECD, *Artificial Intelligence in Society* (OECD Publishing: Paris, France, revised Aug. 2019), accessed Apr. 4, 2021, <https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm>; Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update* (Washington, D.C.: June 2019); and GAO, *Artificial Intelligence in Health Care: Benefits and Challenges of Technologies to Augment Patient Care*, [GAO-21-7SP](#) (Washington, D.C.: Nov. 30, 2020).

models, validating the system, continuously assessing its impact and, if necessary, retiring an AI system from production.<sup>6</sup>

Figure 1: The Phases in the AI Life Cycle



Source: GAO. | GAO-23-106811

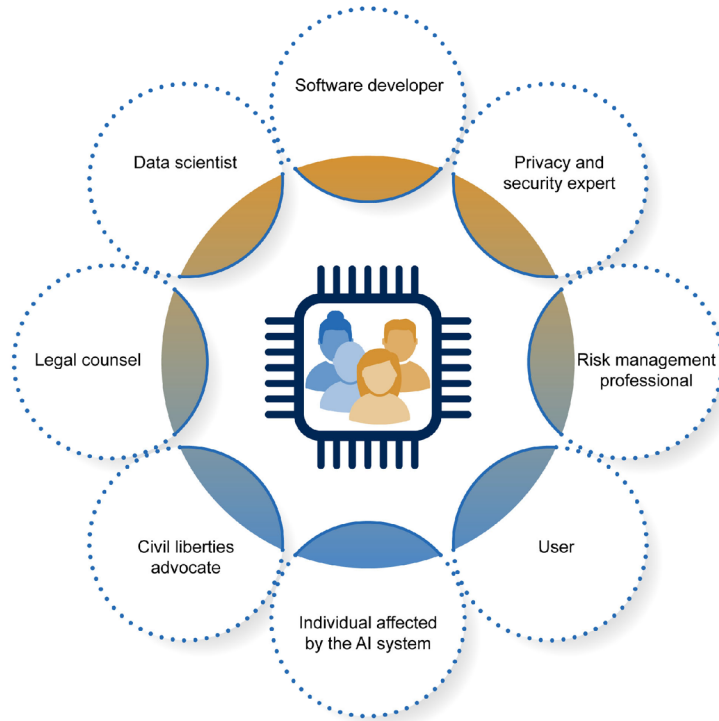
Technical and Societal Implications of AI

Implementing AI systems involves assessing technical performance, as well as identifying and mitigating any societal concerns. For example, to manage technical performance, AI technical

<sup>6</sup>OECD, *Artificial Intelligence in Society*.

stakeholders—data scientists, data engineers, developers, cybersecurity specialists, program managers, and others—will have to ensure that the AI system solves the problem initially identified; uses data sets appropriate for the problem; selects the most suitable learning algorithms; and evaluates and validates the system and its components to ensure it is functioning as intended. Without such assurances, AI systems may perform in unintended ways or otherwise not achieve the goals set out to achieve. As shown in figure 2, in addition to the AI technical stakeholders noted above, a broader community of participants—policy and legal experts, subject matter experts, and individuals using the AI system or impacted by its use, among others—should be engaged in AI development.

**Figure 2: Example of the Community of Stakeholders Engaged in AI Development**

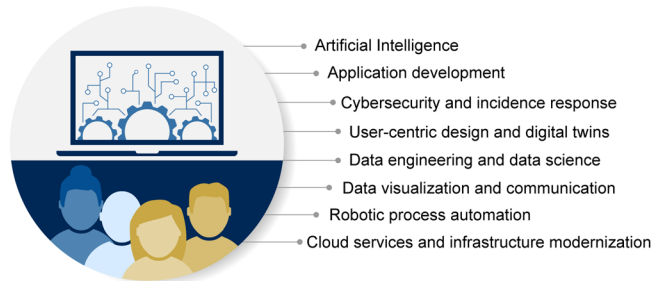


Source: GAO. | GAO-23-108811

### Federal Government Digital Services

Federal agencies rely on digital services to interact with the public and improve organizational performance. Such digital services, as defined by the Office of Management and Budget, include the delivery of digital information (e.g., data or content) and transactional services (e.g., online forms) across a variety of platforms, devices, and delivery mechanisms, such as websites, mobile applications, and social media. The digital services take a variety of forms (see fig. 3).

**Figure 3: Examples of Digital Services Skills, Expertise, and Disciplines**



Source: GAO analysis of roundtable participants' perspectives, and industry and government documents (text); GAO (icons). | GAO-23-106811

Individuals can obtain the necessary digital skills through a variety of pathways. For example, they can attend undergraduate and graduate degree programs, certification programs, and digital skills "boot camps," or they can access free online courses and learn on their own. Additionally, some employers provide on-the-job training in areas such as AI, data science, and cloud services. For example, one company we interviewed has established an academy to provide its new digital services employees with a multi-week, in-person training to enhance their skills.

### **Developing a Federal Digital Workforce Pipeline**

Effective use of AI to improve government operations requires a digitally-ready workforce. Since 2001, however, GAO has identified mission-critical gaps in federal workforce skills and expertise

in fields such as science, technology, engineering, and mathematics as high-risk areas.<sup>7</sup> Agencies' needs for digital services staff span varying degrees of urgency and roles.

During our October 2021 roundtable discussion, technology leaders and knowledgeable experts shared their perspectives on developing a pipeline of federal digital staff. The discussion included observations about agencies' immediate and long-term needs, key characteristics of a digital services academy, and agency and government-wide considerations around recruitment and retention of digital services staff.

#### Immediate and Long-Term Needs

Roundtable participants discussed agencies' immediate and long-term needs for digital services staff, the types of inherently governmental work that a digital-ready workforce could execute, and challenges associated with current hiring methods. For example, one roundtable participant noted that their agency had more than 2,000 open positions requiring digital skill sets, and another described numerous project backlogs. Such gaps may lead to cascading implementation challenges.

Additionally, participants said there is a long-term need for in-house talent across roles such as executives, program staff, product managers, software developers, and engineers who understand data architecture and algorithmic elements.

#### Key Characteristics of a Digital Services Academy

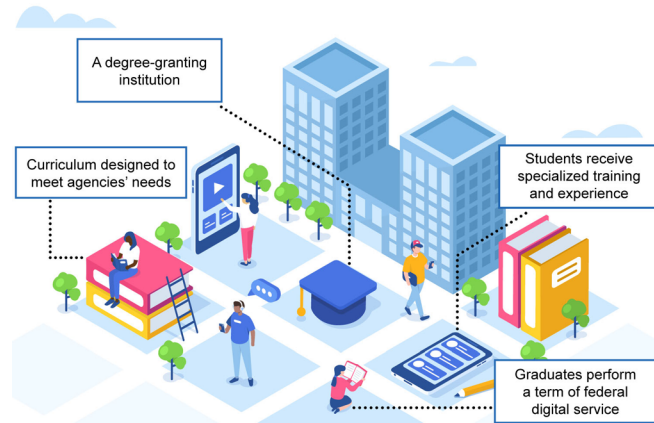
Multiple reports by national advisory groups have suggested that one solution to the lack of digital expertise is that the federal government establish a new service academy—similar to the military academies—to train future civil servants in the digital competencies needed to modernize government (see fig. 4).<sup>8</sup>

---

<sup>7</sup>GAO, *High-Risk Series: Dedicated Leadership Needed to Address Limited Progress in Most High-Risk Areas*, GAO-21-119SP (Washington, D.C.: Mar. 2, 2021).

<sup>8</sup>The National Security Commission on Artificial Intelligence, *Final Report* (Arlington, VA.: 2021) and The National Artificial Intelligence Advisory Committee, *Year 1 Report* (May 2023).

**Figure 4: Example of a Digital Services Academy Concept**



Source: GAO analysis of roundtable participants' perspectives and government documents (text) and Irina Strelnikova/stock.adobe.com (graphic). | GAO-23-106811

A digital services academy could help develop the pipeline of digital services workers to better meet the needs of the federal workforce, according to roundtable participants. Digital services staff could apply advanced technologies, such as AI in health care, or conduct investigative work using machine learning systems. Roundtable participants noted that digital services staff could also use newer technologies to develop services faster or at a lower cost.

Considerations for such an academy include the kinds of skills that would be taught and the composition and size of a graduating class. Digital services staff would require a variety of both digital and government-related skills to meet agencies' needs. Digital skills include application development, data engineering, and other core AI competencies. Government-related skills include knowing how to navigate the requirements of federal data governance and information assurance regimes. In addition, participants noted that a master's degree pipeline may be more appropriate than an undergraduate degree pipeline because agencies need staff with advanced skills in leading projects and programs, data curation, and digitalization.

A digital services academy composed of a diverse student body may further help address societal impacts. One participant noted that programs may attract a more diverse student body if they have a technical component and a social, mission-driven component. For example, a

course on “responsible data science” would likely attract students who are demographically diverse and interested in mission-driven work.

#### Agency and Government-wide Considerations

Agencies can prepare for a pipeline of qualified digital services staff by taking steps such as integrating mission needs into digital services projects, developing professional growth opportunities, cultivating institutional relationships, establishing support networks, and building a data-centric culture, according to roundtable participants. At the same time, participants discussed government-wide challenges associated with existing policies, infrastructure, laws, and regulations that may hinder agency recruitment and retention of digital services staff. For example:

- *Modernizing technological infrastructure.* Participants said a lack of modern technology infrastructure limits the ability of government agencies to leverage the skills of digital services staff.
- *Addressing compensation concerns.* Current salaries and compensation for federal digital services staff are not competitive with the private sector.
- *Streamlining the federal hiring process.* Without a more streamlined approach to onboarding staff, many digital services staff would likely not be willing to wait out the lengthy federal hiring process when the private sector can hire more quickly.

#### **Factors Affecting Oversight of AI in the Public Sector**

Our AI Accountability Framework emphasizes substantive approaches third-party assessors and auditors should take to develop credible assurance assessments of AI systems. Experts in our forum discussed how principles on the use of AI can be operationalized into practices for managers and supervisors of these systems, as well as third-party assessors. The forum included topics such as governance factors to consider in auditing AI systems, criteria auditors can use in assessing AI systems, issues and challenges in auditing AI systems in the public

sector, and evaluation of AI systems for bias and equity.<sup>9</sup> Participants also highlighted challenges that federal agencies are facing, such as having a need for technical expertise, a limited understanding of how AI makes its decisions, and limited access to key information due to commercial procurement of such systems.

Our AI Accountability Framework is organized around four complementary principles, which address governance, data, performance, and monitoring. For each principle, the framework describes key practices for federal agencies and other entities that are considering, selecting, and implementing AI systems. For example:

- **Governance.** This principle describes key practices to promote accountability by establishing processes to manage, operate, and oversee AI implementation. For example, *Workforce* highlights the importance of recruiting, developing, and retaining personnel with multidisciplinary skills and experience in design, development, deployment, assessment, and monitoring of AI systems.
- **Data.** This principle describes key practices to help entities use data that are appropriate for the intended use of each AI system. For example, *Reliability* emphasizes the need to ensure the reliability of the data used to develop the models.
- **Performance.** This principle describes key practices to help entities produce results that are consistent with program objectives. For example, *Bias* describes the necessity of identifying potential biases, inequities, and other societal concerns resulting from the AI system.
- **Monitoring.** This principle describes key practices to help entities ensure their AI systems remain reliable and relevant over time. For example, *Traceability* discusses how entities will need to document results of monitoring activities and any corrective actions taken to promote traceability and transparency.

---

<sup>9</sup>For more information on topics discussed at the CG Forum such as factors affecting oversight of AI, AI governance, sources of evidence, methods to assess implementation of AI systems, and identifying and mitigating potential bias and inequities, see Appendix II of the Framework.

Additionally, each practice includes a set of questions for entities, auditors, and third-party assessors to consider, as well as procedures for auditors and third-party assessors. For more information on the principles and key practices within the Framework, see Appendix I.

In summary, we noted in our AI Accountability Framework that AI is evolving at a pace at which we cannot afford to be reactive to its complexities, risks, and societal consequences. Auditors and the oversight community play a vital role in the “trust but verify” equation and need a blueprint to evaluate this changing technology.

More importantly, organizations that build, purchase, and deploy AI need a framework to understand how AI systems will be evaluated. In recent years, both foreign and domestic stakeholders have developed governance and auditing frameworks, in part, to address the technical and societal issues associated with using AI in the public sector.

GAO looks forward to seeing our Framework in use by federal agencies, and to working with the oversight community, researchers, industry, and the Congress to bring verifiable AI oversight to the cross-cutting work that GAO will continue to undertake.

Chairman Peters, Ranking Member Paul, and Members of the Committee, this completes my prepared statement. I would be pleased to respond to any questions that you may have at this time.

#### **GAO Contact and Staff Acknowledgments**

If you or your staff have any questions about this testimony, please contact Taka Ariga at (202) 512-6888 or [arigat@gao.gov](mailto:arigat@gao.gov). Contact points for our Office of Congressional Relations and Public Affairs may be found on the last page of this statement. GAO staff who made key contributions to this testimony are Farahnaaz Khakoo-Mausel (Assistant Director), Jon D. Menaster (Analyst-in-Charge), Lisa Gardner, Nicole Catanzarite, Louise Fickel, Ryan Han, Lisa Gardner, Stephanie Palmer, and Evonne Tang.

Appendix I: Summaries of Key Practices in GAO's AI Accountability Framework



## 1. Governance

To help entities promote accountability and responsible use of AI systems, GAO identified key practices for establishing governance structures and processes to manage, operate, and oversee the implementation of these systems.

### Key Practices

#### Governance at the Organizational Level



- 1.1 Clear goals:** Define clear goals and objectives for the AI system to ensure intended outcomes are achieved.
- 1.2 Roles and responsibilities:** Define clear roles, responsibilities, and delegation of authority for the AI system to ensure effective operations, timely corrections, and sustained oversight.
- 1.3 Values:** Demonstrate a commitment to values and principles established by the entity to foster public trust in responsible use of the AI system.
- 1.4 Workforce:** Recruit, develop, and retain personnel with multidisciplinary skills and experiences in design, development, deployment, assessment, and monitoring of AI systems.
- 1.5 Stakeholder involvement:** Include diverse perspectives from a community of stakeholders throughout the AI life cycle to mitigate risks.
- 1.6 Risk management:** Implement an AI-specific risk management plan to systematically identify, analyze, and mitigate risks.

#### Governance at the Systems Level

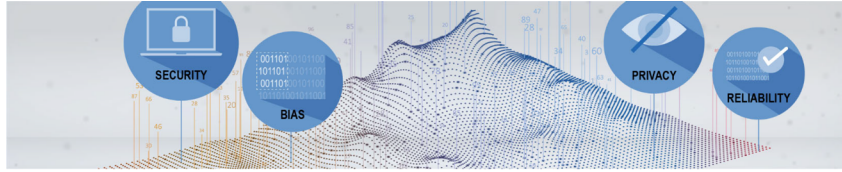
- 1.7 Specifications:** Establish and document technical specifications to ensure the AI system meets its intended purpose.
- 1.8 Compliance:** Ensure the AI system complies with relevant laws, regulations, standards, and guidance.
- 1.9 Transparency:** Promote transparency by enabling external stakeholders to access information on the design, operation, and limitations of the AI system.

#### Example of an AI Governance Structure

In 2020, the Department of Defense (DOD) established an AI Executive Steering Group, which was created as the senior governance body to provide coordination and oversight of DOD's AI policies and activities. The Executive Steering Committee oversees nine subcommittees, one of which is on ethics. That subcommittee is responsible for providing practical guidance on how to apply the ethical principles for AI adopted by DOD to the different phases of the AI life cycle.

#### Selected Discussion from the Comptroller General Forum

- Entities should implement governance structures for AI systems that incorporate organizational values, consider risks, assign clear roles and responsibilities, and involve multidisciplinary stakeholders.
- Entities should define a governance structure that includes clear goals and objectives, which translates into systems requirements and performance metrics.
- Entities should include diverse perspectives from technical and non-technical communities throughout the AI life cycle to anticipate and mitigate unintended consequences including potential bias and discrimination.



## 2. Data

To help entities use data that are appropriate for the intended use of each AI system, GAO identified key practices to ensure data are of high quality, reliable, and representative.

### Key Practices



#### Data used for Model Development

- 2.1 Sources:** Document sources and origins of data used to develop the models underpinning the AI system.
- 2.2 Reliability:** Assess reliability of data used to develop the models.
- 2.3 Categorization:** Assess attributes used to categorize data.
- 2.4 Variable selection:** Assess data variables used in the AI component models.
- 2.5 Enhancement:** Assess the use of synthetic, imputed, and/or augmented data.

#### Data Used for System Operation

- 2.6 Dependency:** Assess interconnectivities and dependencies of data streams that operationalize the AI system.
- 2.7 Bias:** Assess reliability, quality, and representativeness of all the data used in the system's operation, including any potential biases, inequities, and other societal concerns associated with the AI system's data.
- 2.8 Security and privacy:** Assess data security and privacy for the AI system.

#### Example of Data Reliability

In 2019, the European Union Agency for Fundamental Rights released the report *Data Quality and Artificial Intelligence – Mitigating Bias and Error to Protect Fundamental Rights*. The report emphasizes the need for high-quality data and algorithms in machine learning systems and AI, and how transparency about data used in AI systems may help to prevent rights violations. The report also explains how AI systems use data, provides examples of how biases could be introduced, and provides examples of how low-quality data might affect accuracy and outcomes. Criteria for assessing data quality listed in the report include completeness, accuracy, consistency, timeliness, duplication, validity, availability, and whether the data are fit for the purpose.

#### Selected Discussion from the Comptroller General Forum

- Entities should provide documentation describing how training and testing data have been acquired or collected, prepared, and updated to demonstrate data quality and reliability.
- Entities should test data used in AI systems for biases. Biases may be introduced unintentionally during data collection and labeling.
- Entities should monitor data after deploying AI systems to identify potential data drift, which can lead to unintended consequences.

Source: GAO, majcot/stock.adobe.com (header); GAO (illustration). | GAO-23-106811



### 3. Performance

To help entities ensure AI systems produce results that are consistent with program objectives, GAO identified key practices for ensuring that systems meets their intended purposes.

#### Key Practices

##### Performance at the Component Level

- 3.1 Documentation:** Catalog model and non-model components, along with operating specifications and parameters.
- 3.2 Metrics:** Define performance metrics that are precise, consistent, and reproducible.
- 3.3 Assessment:** Assess the performance of each component against defined metrics to ensure it functions as intended and is consistent with program goals and objectives.
- 3.4 Outputs:** Assess whether outputs of each component are appropriate for the operational context of the AI system.

##### Performance at the System-Level

- 3.5 Documentation:** Document the methods for assessment, performance metrics, and outcomes of the AI system to provide transparency over its performance.
- 3.6 Metrics:** Define performance metrics that are precise, consistent, and reproducible.
- 3.7 Assessment:** Assess performance against defined metrics to ensure the AI system functions as intended and is sufficiently robust.
- 3.8 Bias:** Identify potential biases, inequities, and other societal concerns resulting from the AI system.
- 3.9 Human supervision:** Define and develop procedures for human supervision of the AI system to ensure accountability.

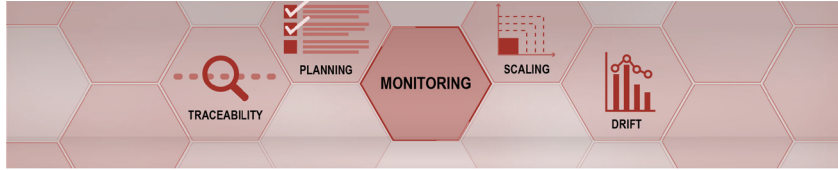


#### Example of Performance Documentation

Industry and nonprofit entities provided several examples of how entities can document performance by recording several aspects of AI systems, including intended use, specifications, testing methodology and test results, ethical considerations, and evaluation. Each of those examples includes questions or factors for consideration to guide entities in designing, developing, and deploying AI systems.

#### Selected Discussion from the Comptroller General Forum

- Entities should document requirements—including performance metrics—for the AI system throughout the life cycle.
- Entities should document methods to assess performance—which can include input-output tests, stress tests, and evaluations of model drift—to ensure AI systems meet their intended goals.
- Entities should provide access to performance test results, change logs, and other documentation describing updates and key design choices, and provide a copy of the model or algorithm code to third-party assessors of AI systems.



## 4. Monitoring

To help entities ensure reliability and relevance of AI systems over time, GAO identified key practices for monitoring performance and assessing sustainment and expanded use.

### Key Practices



#### Continuous Monitoring of Performance

- 4.1 Planning:** Develop plans for continuous or routine monitoring of the AI system to ensure it performs as intended.
- 4.2 Drift:** Establish the range of data and model drift that is acceptable to ensure the AI system produces desired results.
- 4.3 Traceability:** Document results of monitoring activities and any corrective actions taken to promote traceability and transparency.

#### Assessing Sustainment and Expanded Use

- 4.4 Ongoing assessment:** Assess the utility of the AI system to ensure its relevance to the current context.
- 4.5 Scaling:** Identify conditions, if any, under which the AI system may be scaled or expanded beyond its current use.

#### Example of Monitoring

In 2020, the World Economic Forum released the *Companion to the Model AI Governance Framework – Implementation and Self-Assessment Guide for Organizations*, which includes guidance on data monitoring and a discussion of ongoing monitoring, review, and tuning of AI algorithms and models. The guidance suggests updating AI systems based on changes in the operational environment, as well as documenting when and how the update took place, and the impact it had on the model outputs.

Source: GAO. | GAO-23-106811

#### Selected Discussion from the Comptroller General Forum

- Entities should continuously monitor and evaluate the AI system to ensure it addresses program objectives.
- Entities should monitor changes in the data and models to ensure relevance and appropriateness.
- Entities should continuously monitor the AI system to ensure the system is appropriate in its current operating context.

**Appendix II: Selected List of GAO Reports on Artificial Intelligence**

GAO, *Artificial Intelligence in Health Care: Benefits and Challenges of Machine Learning Technologies for Medical Diagnostics*, [GAO-22-104629](#) (Washington, D.C.: Sept. 29, 2022).

GAO, *Facial Recognition Technology: Federal Agencies Use and Related Privacy Protections*, [GAO-22-106100](#) (Washington, D.C.: June 29, 2022).

GAO, *Artificial Intelligence: DOD Should Improve Strategies, Inventory Process, and Collaboration Guidance*, [GAO-22-105834](#) (Washington, D.C.: Mar. 30, 2022).

GAO, *Artificial Intelligence: Status of Developing and Acquiring Capabilities for Weapon Systems*, [GAO-22-104765](#) (Washington, D.C.: Feb. 17, 2022).

GAO, *Digital Services: Considerations for a Federal Academy to Develop a Pipeline of Digital Staff*, [GAO-22-105388](#) (Washington, D.C.: Nov. 19, 2021).

GAO, *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*, [GAO-21-519SP](#) (Washington, D.C.: June 30, 2021).

GAO, *Facial Recognition Technology: Federal Law Enforcement Agencies Should Better Assess Privacy and Other Risks*, [GAO-21-518](#) (Washington, D.C.: June 3, 2021).

GAO, *Artificial Intelligence in Health Care: Benefits and Challenges of Technologies to Augment Patient Care*, [GAO-21-7SP](#) (Washington, D.C.: Nov. 30, 2020).

GAO, *Artificial Intelligence in Health Care: Benefits and Challenges of Machine Learning in Drug Development*, [GAO-20-215SP](#) (Washington, D.C.: Dec. 20, 2019, reissued Jan. 31, 2020).

GAO, *Technology Assessment: Artificial Intelligence: Emerging Opportunities, Challenges, and Implications*, [GAO-18-142SP](#) (Washington, D.C.: Mar. 28, 2018).

## Opportunities and Risks of Artificial Intelligence in the Public Sector

Daniel E. Ho<sup>1</sup>  
Stanford University

*Testimony presented to the U.S. Senate Committee on Homeland Security and Governmental Affairs on May 16, 2023*

Chairman Peters, Ranking Member Paul, and Members of the Committee, it is an honor to speak with you today.

I'm a professor at Stanford University, where I serve as Associate Director of Stanford's Institute for Human-Centered AI (or HAI) and Director of the Regulation, Evaluation, and Governance Lab (or RegLab). I also serve as a Member of the National AI Advisory Committee (NAIAC) and the RegLab works with numerous federal agencies, but I speak to you today in my personal capacity.

The U.S. government has an exceptional opportunity. It can seize this moment of AI innovation to modernize federal programs, make agencies more effective, catalyze scientific advancements, and protect civil rights and liberties for the benefit of all Americans. Doing so will strengthen America. But strategic leadership, critical investments in the federal workforce and digital infrastructure, and adapting procurement to AI are preconditions.

### I. The Importance of Public Sector AI

---

<sup>1</sup> William Benjamin Scott and Luna M. Scott Professor of Law; Professor of Political Science; Senior Fellow at Stanford Institute for Economic Policy Research, Stanford University; Associate Director, Stanford Institute for Human-Centered Artificial Intelligence (HAI); Faculty Director, Stanford Regulation, Evaluation, and Governance Lab (RegLab). All views expressed in this testimony are provided in an individual capacity and do not represent the views of any affiliated institution, agency, or the National AI Advisory Commission (NAIAC).

Let me start with what is possible if the government gets this right.

First, government should lead by example and demonstrate how responsible AI can modernize federal programs. In a report to the Administrative Conference of the United States that I co-authored with administrative law scholars, we showed how early AI innovation in nearly half of the largest 142 federal agencies can transform the administration of government benefits like veteran disability compensation, improve monitoring of public health risks and adverse drug effects, and help protect workers, consumers, and the environment.<sup>2</sup>

Take the Social Security Administration (SSA), which pays benefits to some 18 million Americans annually. The SSA's administrative judges can hear over a half million disability appeals per year.<sup>3</sup> With great foresight, SSA began investing in data infrastructure and tools to modernize case adjudication in the 1990s. In one pilot, SSA used AI to reorder cases by similarity, allowing adjudicators to learn complex areas of the law more effectively. The resulting "micro-specialization" increased the speed and accuracy of adjudicators in the pilot.<sup>4</sup> SSA's early investments culminated in an AI tool that allows judges to check draft decisions for some 30 errors.<sup>5</sup> Such innovations can expedite and improve agency decision-making to better serve American citizens and some have called the official who pioneered these early investments the "Steve Jobs of the SSA."<sup>6</sup>

---

<sup>2</sup> Engstrom, David Freeman, Daniel E. Ho, Catherine Sharkey, and Mariano-Florentino Cuéllar. 2020. "Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies." Administrative Conference of the United States.

<sup>3</sup> Ames, David, Cassandra Handan-Nader, Daniel E. Ho, and David Marcus. 2020. "Due Process and Mass Adjudication: Crisis and Reform." *Stanford Law Review* 72:1.

<sup>4</sup> "SSA reported 12% reduction in case processing time and 7.5% reduction in returns from administrative appeal judges to attorneys." Engstrom, David Freeman, Daniel E. Ho, Catherine Sharkey, and Mariano-Florentino Cuéllar. 2020. "Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies." Administrative Conference of the United States.

<sup>5</sup> Glaze, Kurt, Daniel E. Ho, Gerald K. Ray, and Christine Tsang. 2022. "Artificial Intelligence for Adjudication: The Social Security Administration and AI Governance." In *The Oxford Handbook of AI Governance*. Oxford University Press, Handbook on AI Governance.

<sup>6</sup> See Administrative Conference of the United States, Recommendation 2021-10, "Quality Assurance Systems in Agency Adjudication," 87 Federal Register 1722 (2022).

Second, government agencies are of course also critical for regulation, be it for FDA's approval of AI medical devices or the Department of Transportation's approach to self-driving cars.<sup>7</sup> The right regulatory approach – striking the appropriate balance between innovation and safeguards – requires expertise in government. Government cannot govern AI if it does not understand AI.<sup>8</sup> Getting technical talent into the federal workforce is the biggest obstacle to the U.S. government's internal adoption of AI, effective regulation of its risks, and successful implementation of critical AI policy recommendations from the National Security Commission on AI, NAIAC, and others.<sup>9</sup>

## II. Challenges

While much progress has been made, including the Advancing American AI Act, AI Training for the Acquisition Workforce Act, and other important legislation from this committee, we still have a long way to go.

When our research team at Stanford examined the implementation of AI-related legal requirements that EO 13,960, 13,859, and the AI in Government Act placed on federal agencies, we found a critical gap in leadership, strategic planning, and capacity.<sup>10</sup> For example:

<sup>7</sup> See, e.g., FDA, Artificial Intelligence and Machine Learning in Software as a Medical Device, <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>; DOT, USDOT Automated Vehicles Activities, <https://www.transportation.gov/AV>.

<sup>8</sup> As an example, FDA approvals of AI medical devices appear to be based on a limited number of hospital sites, and the performance of an algorithm that performs well in hospital A can degrade significantly in hospital B. Technical expertise and domain knowledge are required to ensure that the device approval process is adapted to accounts such complexities with AI systems. See Wu, Eric, Kevin Wu, Roxana Daneshjou, David Ouyang, Daniel E. Ho, and James Zou. 2021. "How Medical AI Devices Are Evaluated: Limitations and Recommendations from an Analysis of FDA Approvals." *Nature Medicine* 27 (4): 582–84.

<sup>9</sup> I am not alone in this concern. When Eric Schmidt was asked a few months ago about his biggest concern on Congress's implementation of the National Security Commission on AI's recommendations, he singled out one thing: technical talent. Eric Schmidt, Testimony to House Subcommittee on Cybersecurity, Information Technology, and Government Innovation, March 8, 2023 ("The area that I'm most focused on right now is basically the training problem and I just don't see the progress in the government to reform the way it hires and promotes technical people.").

<sup>10</sup> Lawrence, Christie, Isaac Cui, and Daniel E. Ho. 2022. "Implementation Challenges to Three Pillars of America's AI Strategy." Stanford HAI-RegLab White Paper. [https://dho.stanford.edu/wp-content/uploads/AI\\_Implementation.pdf](https://dho.stanford.edu/wp-content/uploads/AI_Implementation.pdf)

- 88 percent of agencies failed to submit AI Plans to identify regulatory authorities and mechanisms to promote responsible AI and protect Americans' rights and safety;<sup>11</sup>
- The implementation of a key transparency measure – agency disclosure of its AI use cases through published inventories – has been inconsistent;<sup>12</sup> and
- The Office of Personnel Management has yet to release a required report – due July of last year – to forecast employment needs and to create an AI hiring line.<sup>13</sup>

This must change. The federal workforce does Herculean work, but faces fundamental challenges developing teams that can design, implement, and regulate AI effectively and responsibly. The most recent AI Index report by Stanford HAI highlights that 65% of AI PhDs land in industry, 28% in academia, and less than 2% in government.<sup>14</sup> Or, in the words of one entrepreneur: “The best minds of my generation are thinking about how to make people click ads.”<sup>15</sup>

Strengthening the pipeline of technical talent into the public sector is paramount. So too is ensuring that agencies have the right digital infrastructure, technical directives, and career paths to nurture and grow technical talent. As the National Security Commission on AI noted, it's not just compensation: “It is the perception, and too often the reality, that it is difficult for digital talent in government to perform meaningful work.”<sup>16</sup> I've seen firsthand how the government is failing to recruit and retain technical experts. One Stanford AI Ph.D. student, for instance, became so frustrated by an agency's decades-old software stack and lack of advanced (GPU) computing that he gave up on government and went back to work in industry.

### III. Recommendations

<sup>11</sup> See Section 6(c) of EO 13859, OMB M-21-06, and Lawrence, Cai, and Ho, *supra*, note 10.

<sup>12</sup> See Section 5(e) of EO 13960 and Lawrence, Cai, and Ho, *supra*, note 10.

<sup>13</sup> Section 105 of the AI in Government Act directed the Office of Personnel Management to create an AI occupational series and estimate AI-related workforce needs in each federal agency by July 2022.

<sup>14</sup> Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, “The AI Index 2023 Annual Report,” AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

<sup>15</sup> Ashlee Vance, “This Tech Bubble Is Different,” *Bloomberg* (Apr. 14, 2011).

<sup>16</sup> Schmidt, Eric, Bob Work, Safra Catz, Steve Chien, Chris Darby, Kenneth Ford, Jose-Marie Griffiths et al. “National Security Commission on Artificial Intelligence (AI), Final Report.” 2021.

Let me conclude with four recommendations necessary for the U.S. government to maintain American leadership in AI innovation and responsible AI.

First, strategic leadership from the federal government is required to coordinate and drive forward AI innovation and trustworthy adoption. Congress should borrow a page from the bipartisan Evidence Act, which required the appointment of officials responsible for data and evaluation, and empower Chief AI Officers to ensure that senior leadership within agencies is driving forward responsible AI innovation.<sup>17</sup> The White House, too, must also be organized and staffed to rise to this challenge.<sup>18</sup>

Second, Congress should establish new pathways and trajectories for technical talent in government. We need better models – building on the US Digital Service, public-private partnerships, and academic-agency partnerships – to attract AI talent to public service and build cross-functional teams. Retaining AI talent requires giving them meaningful positions related to their expertise.

Third, an effective procurement system should capitalize on American innovation and spur developments of rights-preserving, privacy-enhancing technologies. The AI Training Act is a fantastic step in the right direction, but we cannot rely on procurement officials alone. We need upskilling of business units and to enable more modular forms of contracting – which DOD has illustrated – that enables more effective development, acquisition, assessment, and auditing of AI systems.<sup>19</sup>

Last, we have to invest in digital infrastructure. The federal task force proposal for a National AI Research Resource would give AI researchers, students, and small businesses secure access to

---

<sup>17</sup> This recommendation is made by the National AI Advisory Commission's Year 1 Report.

<https://www.ai.gov/wp-content/uploads/2023/05/NAIAC-Report-Year1.pdf>

<sup>18</sup> See National AI Advisory Commission's Year 1 Report: Schmidt, Eric, Bob Work, Safra Catz, Steve Chien, Chris Darby, Kenneth Ford, Jose-Marie Griffiths et al. "National Security Commission on Artificial Intelligence (AI), Final Report." 2021.

<sup>19</sup> Raji, Inioluwa Deborah, Peggy Xu, Colleen Honigsberg, and Daniel Ho. 2022. "Outsider Oversight: Designing a Third Party Audit Ecosystem for Ai Governance." In Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society, 557–71.

high-quality administrative data and computing resources to level the playing field between industry, academia, and government.<sup>20</sup> AI requires high-fidelity data and many of the negative impacts of AI we've observed stem from training large models on anything on the web, including unverified and harmful information. Government data, which is higher-quality, more representative, and more reliable is an important part of the solution. When the U.S. Geological Service made Landsat satellite imagery free to researchers in 2008, it generated 3 to 4 billion dollars in benefits annually, catalyzing discoveries in habitat modification, climate change, and poverty.<sup>21</sup> That is the promise of getting the public sector innovation infrastructure right.

The U.S. government should act expeditiously to foster responsible AI adoption. AI that embodies American values and helps agencies better serve Americans equitably can also build public trust and confidence. Thank you for the opportunity to speak. I'm looking forward to your questions.

---

<sup>20</sup> National Artificial Intelligence Research Resource Task Force, "Strengthening and Democratizing the U.S. Artificial Intelligence Innovation Ecosystem: An Implementation Plan for a National Artificial Intelligence Research Resource," Jan. 2023, <https://www.ai.gov/wp-content/uploads/2023/01/NAIRR-TF-Final-Report-2023.pdf>. Ho, Daniel, Jennifer King, Russell Wald, and Christopher Wan. 2021. "Building a National AI Research Resource: A Blueprint for the National Research Cloud." Stanford HAI White Paper. [https://hai.stanford.edu/sites/default/files/2022-01/HAI\\_NRCR\\_v17.pdf](https://hai.stanford.edu/sites/default/files/2022-01/HAI_NRCR_v17.pdf).

<sup>21</sup> Straub, Crista L., Stephen R. Koontz, and John B. Loomis. "Economic Valuation of Landsat Imagery." 2019.



## AI in Government

United States Senate Committee on  
Homeland Security & Government Affairs

May 16, 2023

### Testimony of Ritchie Eppink

Chair Peters, Ranking Member Paul, and members of the committee: thank you for your attention to artificial intelligence and automated decision-making in government programs, and for inviting me to testify about it at this important hearing. I was invited here today because I have been working for over a decade with Idahoans with developmental disabilities and their families to address black boxes around automated decision-making hidden in a federally funded program. Only through the litigation that I've helped these families pursue were they able to access the secret computerized algorithms the State of Idaho uses to make decisions about the health care they depend on day to day. Once we opened the black box that concealed that automated system, we found that Idaho built the system out of corrupt data, relied on inputs that the State never validated, and produced unfair results that even those who created it could not explain. A federal court ruled that the system was unconstitutional. Yet a decade after filing suit and over seven years since winning in court we are still litigating the case, battling for due process against still more black box secrecy. The lesson is this: Decades-long class actions by indigent families are not a viable plan for AI governance in federal programs. We need federal regulation and enforcement to protect basic fairness and constitutional rights and establish clear guardrails for whether and how government programs use AI and other automated decision-making systems.

I am Of Counsel with the American Civil Liberties Union of Idaho, where I was previously the Legal Director for nine years. Before that I held the title of Justice Architect at Idaho Legal Aid Services, a federally funded civil legal services program where the case I'm here to testify about today originated. I have also held a Fulbright Fellowship, during which I studied strategies for community education about legal rights and responsibilities. My principal practice currently is as an attorney and co-founder of Wrest Collective, a community-funded, sliding scale nonprofit law firm in Boise, Idaho.

### The *K.W. v. Armstrong* Lawsuit

The State of Idaho, through its federally funded Medicaid program, provides health care for some of its residents experiencing drastic poverty. Among those eligible for Medicaid in Idaho are certain people with developmental and intellectual disabilities. Although in the past the government would have confined these people in state hospitals and other institutional facilities, through Medicaid they can choose to get services at home and in their communities instead—at a savings to the government and taxpayers. Those who are eligible must present themselves annually for an assessment. In Idaho’s system, assessors complete two assessment instruments for each person. The assessor plugs these assessment results into an automated system that converts them into a dollar amount for each person. That amount is called the person’s “budget.” It’s the presumptive maximum amount the person can use for the Medicaid services related to their developmental or intellectual disabilities over the coming year. Medicaid does not pay the person their budget amount. Rather, the person’s support providers get reimbursed from the budget for services actually rendered.

The lawsuit I’m here to testify about, known as *K.W. v. Armstrong*, began in 2012. People across Idaho started getting notices, after their annual assessments, that their automated budgets had dropped dramatically. Some of them saw their individualized budgets suddenly drop by tens of thousands of dollars and more than 30%. For one of my clients, the late Christie Mathwig, the automated system suddenly cut her budget by 21% with no explanation. Living in rural Idaho where wildfires and winter storms are common, she could not evacuate in an emergency or administer her own medications without help. But under her reduced budget, she would lose the 24-hour assistance she needed to survive. Another client, who I’ll call by her initials “A.L.” as used in the lawsuit, was institutionalized when she was 9 years old. She remained in institutions and hospitals and incarcerated until she had access to Medicaid services for a community placement in Idaho. Though she



*Christie Mathwig in her rural Idaho home*

previously had access to a \$42 thousand budget to pay for Medicaid services she needed, in 2012 Idaho's automated system cut her budget by nearly 20% to \$34 thousand with no explanation. Another of my clients, "Matthew S.," a then 40 year-old man living in his community with the help of developmental therapy, saw his budget drop from \$52 thousand to \$34 thousand, a 35% cut. Even the State's own assessor noted that a reduction in his services would result in reduced independence and loss of skills. But, as with Christie and A.L., the State gave no explanation for the cut. For each of them, these sudden and severe cuts put their independence, their safety, and their liberty at immediate risk.

Christie, A.L., Matthew, and others reached out to me for legal help. I naively sent a letter asking the State to explain how it translated assessment results into individual budget amounts. The State's Department of Health and Welfare, which administers the Medicaid program in Idaho, refused to explain it. Its attorneys replied that the system was a "trade secret."

After that response, I helped thirteen people with severe budget cuts file a joint lawsuit. We asked a federal court to order the Department to disclose its system. Within a few weeks of filing suit, we got that order. Then we got the system. It was a set of formulas in a fairly basic Microsoft Excel spreadsheet. The Department's assessors enter annual assessment results into a copy of the spreadsheet for each person. The spreadsheet, in hidden cells, computes the person's budget amount.

Now that we had the formulas, we started trying to make sense of them. We got what little documentation there was about them. We learned that the Department concocted the formulas in-house. It also created one of the two assessment instruments that feed the inputs to those formulas. Department staff had just brainstormed the assessment questions. They never validated, standardized, or audited the instrument.

We took the testimony of the state agency employee who devised the formulas and learned that he used statistical predictions to select them. Specifically, he used statistical software to predict how much Medicaid users would spend in the future, extrapolating based on their assessment results and how much they had spent in the past. We also took testimony from the head of the Department's developmental disabilities bureau, the Department worker who supervised review of individuals' Medicaid service plans, the assessment supervisor, and others. They acknowledged that my clients' budgets dropped suddenly after the Department re-ran the statistical modeling and revised its formulas. But none of them could explain why the new statistical model cut some individual budgets so severely. Each time we asked who could explain it, the witness deposed that day would point the finger towards another Department

bureaucrat. By the end of the depositions, the finger pointing had gone around in circles.

We had to hire three experts to make sense of the State's formulas, the data the Department based them on, the statistical modeling it derived them from, and the impact they had on our clients. The experts found serious problems with both the data and the modeling, and therefore the formulas themselves. When we presented our evidence to the court, the court agreed. The court ruled that the Department's formulas were so unreliable that they arbitrarily deprived people of their Medicaid budgets and violated the due process guaranteed in the United States Constitution. To reach its ruling, the court found that the data the Department fed into the statistical modeling was incomplete and probably full of errors to start with. The Department statistician who developed the system, the court noted, later discovered the data harbored geographic bias as well. The resulting formulas would never compute an adequate budget for about 15% of people, the court concluded; and when human Department workers reviewed the automated budgets, they increased more than 60% of them. But the Department did not explain why a person's budget went down, partly because nobody even within the Department could meaningfully understand the automated cuts. The agency also had no written criteria for increasing budgets after human review. Many who contested their budget cuts would need help from a skilled advocate to handle the complex appeal process, anyhow. In short, the automated system was unconstitutional. And the human review available to correct bad automated decisions was not enough to save it. Idaho's system, rudimentary compared to some of the complex AI available today, highlights how critical mistakes in statistical models can rapidly accumulate even in what seem on the surface to be simple automations.

The problems did not end there, either. Though one of the two assessment instruments that Idaho used was its own unvalidated tool, the other was a standardized instrument developed by a private company. The company, with the State's help, fought to ban people with developmental disabilities from accessing the assessment booklet that directs assessors in completing the tool and details each person's individual scores. But, as the court had to point out, without that booklet, people relying on these Medicaid services could not effectively cross-examine the assessors or challenge errors in their automated budgets.

The court ordered the Department to overhaul its automated system. The court also ordered the Department to ensure access to all parts of the assessment booklet needed to fully challenge a budget cut, regularly test the new system, set out criteria for what a person has to show to get their automated budget increased, and make certain that everyone in the Medicaid program has a

“suitable representative” to help pursue a higher budget through the Department’s appeal process.<sup>1</sup>

Just to reach that point in the case, we had to spend over \$40 thousand on experts to analyze all the problems with the automated system. Plus, our team had to put in more than 2,000 attorney and paralegal hours to vindicate our clients’ constitutional rights and secure a settlement agreement after the court’s ruling.

Unfortunately, the case continues today as we contend against a proposed new system that repeats some of the same problems the old system had. After the court approved a settlement agreement, the Department selected a new assessment instrument on which it began building a new automated budget system. Despite the court’s prior rulings—and a subsequent meeting where stakeholders made clear that transparency was essential for the new instrument—we learned that the Department had chosen another black box assessment instrument. As it did before, the State helped this instrument’s private publisher fight to ban people with developmental disabilities and their advocates from accessing the manual that assessors must follow to properly complete the assessment. Once again, the publisher claimed this vital information was proprietary. But without the manual, people relying on these Medicaid services cannot effectively cross-examine assessors or challenge errors in the automated budgets the State would assign to them. We argued over this secrecy in court again this spring. But following the argument and before the court could issue a decision, the private publisher told the State it would not let its instrument be used in Idaho if indigent Medicaid recipients might get to access the manual, which it sells for \$130 on its website.

Transparency may not be the only problem with the new system, either. Preliminary analysis of the new system—including by the firm that helped the State develop it—suggests that the new system is biased against some groups of people who use this Medicaid program. The pre-implementation data showed, for example, that the system could produce budgets that would only “somewhat” meet the needs of 23% of people in one group and “not at all” meet the needs of another 23% in that group.

Although this long-running litigation has kept these recurring problems with automated government decision-making systems in check, it should not require a class action lawsuit to ensure that these systems meet constitutional minimums. Federal policies, regulations, and enforcement are appropriate and

<sup>1</sup> *K.W. v. Armstrong*, 180 F. Supp. 3d 703 (D. Idaho 2016); see also *K.W. v. Armstrong*, 789 F.3d 962 (9th Cir. 2015).

essential to govern automated decision-making systems like these, and to safeguard against the life-upending harm they can cause when not implemented reliably and fairly.

### Dangers this Lawsuit Flags

With governments increasingly relying on automated decision-making systems, scholars, advocates, judges, and policymakers alike have begun to realize these systems' sweeping effects and complex risks. The *K.W. v. Armstrong* case is among the most instructive litigation challenges to one of these systems. The White House's October 2022 *Blueprint for an AI Bill of Rights* references this litigation,<sup>2</sup> and the case continues to be featured in leading scholarship on civil and human rights in this context.<sup>3</sup> It is probably so often cited because it

<sup>2</sup> White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People* 42 & n.86 (Oct. 2022), <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>.

<sup>3</sup> See, e.g., Rashida Richardson, *Defining and Demystifying Automated Decision Systems*, 81 MD. L. REV. 785, 800 & n.70 (2022); Sarah Brown, *Promulgating Poverty: How AI Technology Exacerbates Poverty Issues in Public Programs*, 49 N. KY. L. REV. 267, 277–281 (2022); Chris Chambers Goodman, *AI, Can You Hear Me? Promoting Procedural Due Process in Government Use of Artificial Intelligence Technologies*, 28 RICH. J.L. & TECH. 700, 718 n.81, 720 & nn. 90–91 (2022); Francesca Bignami, *Artificial Intelligence Accountability of Public Administration*, 70 AM. J. COMP. L. 312, 333 n.65 (2022); Ryan Calo & Danielle Keats Citron, *The Automated Administrative State: A Crisis of Legitimacy*, 70 EMORY L.J. 797, 823 (2021); Kristen E. Egger, *Artificial Intelligence in the Workplace: Exploring Liability Under the Americans with Disabilities Act and Regulatory Solutions*, 60 WASHBURN L.J. 527, 542 (2021); Cary Coglianese & Lavi M. Ben Dor, *AI in Adjudication and Administration*, 86 BROOK. L. REV. 791, 834 (2021); Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265, 1279 (2020); Sarah Valentine, *Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control*, 46 FORDHAM URB. L.J. 364, 414 (2019); see also Sarah Brown, *Promulgating Poverty: How AI Technology Exacerbates Poverty Issues in Public Programs*, 49 N. KY. L. REV. 267, 277 (2022); Charles Tait Graves & Sonia K. Katyal, *From Trade Secrecy to Seclusion*, 109 GEO. L.J. 1337, 1379 nn.207–208 (2021); Noah Bunnell, *Remedying Public-Sector Algorithmic Harms: The Case for Local and State Regulation Via Independent Agency*, 54 COLUM. J.L. & SOC. PROBS. 261, 303 n.83 (2021); Christopher Slobogin, *Preventive Justice: How Algorithms, Parole Boards, and Limiting Retributivism Could End Mass Incarceration*, 56 WAKE FOREST L. REV. 97, 166 n.357 (2021); Frank Pasquale, *Normative Dimensions of Consensual*

illustrates several of the many ways that automated decision-making systems can go wrong. These illustrations are especially poignant considering how rudimentary Idaho's system was: if formulas on a basic Excel spreadsheet can present so many constitutional problems, governance to safeguard against these dangers in today's more complex AI systems are all the more critical. We discovered, and had to litigate, each of the following failures of automated decision-making as part of this case:

- **Black boxes concealing that the government is using an automated system**

Although Idaho had been using an automated budgeting system for several years before the *K.W. v. Armstrong* lawsuit began, until then the State did not tell Medicaid recipients anything about it, including that it was using an automated system at all. Only after I sent a letter asking the state Medicaid agency to explain how it computed the sudden 2012 budget reductions did my clients learn that the State considered the system a “trade secret”—and only after we filed a federal lawsuit did we learn that secret was a handful of formulas coded into an Excel spreadsheet. It was a lucky coincidence both that my clients connected with me when I had time, as a busy legal aid lawyer, to take on their case, and that I was naive enough to demand to know how the State had come up with my clients' budget amounts.

Undoubtedly, other states, federal programs, and maybe even other programs in Idaho are currently using AI or automated systems that remain undisclosed to the people those systems make decisions about. This raises core due process concerns. People are unaware that automated systems are being used, let alone how they work or that the systems may be making erroneous or

*Application of Black Box Artificial Intelligence in Administrative Adjudication of Benefits Claims*, LAW & CONTEMP. PROBS., 2021, at 35, 38 n.14 (2021); Cary Coglianese & Erik Lampmann, *Contracting for Algorithmic Accountability*, 6 ADMIN. L. REV. ACCORD 175, 185 n.48 (2021); Aziz Z. Huq, *Constitutional Rights in the Machine-Learning State*, 105 CORNELL L. REV. 1875, 1951 (2020); Jacqueline G. Schafer, *Harnessing AI Innovation for Struggling Families*, U. ILL. J.L. TECH. & POL'Y, Fall 2020, at 411, 442 n.193 (2020); Michael Arkfeld, *A Call to Action: Litigating and Judging Artificial Intelligence Cases*, JUDGES' J., Winter 2020, at 6, 9; Kate Crawford & Jason Schultz, *AI Systems As State Actors*, 119 COLUM. L. REV. 1941, 1949 n.44 (2019); Vera Eidelman, *The First Amendment Case for Public Access to Secret Algorithms Used in Criminal Trials*, 34 GA. ST. U. L. REV. 915, 944 n.11 (2018); Frederick Schauer, *How (and If) Law Matters*, 129 HARV. L. REV. F. 350, 353 n.15 (2016).

discriminatory decisions. This severely curtails their ability to seek legal redress to protect their rights.

- **Black boxes concealing flaws in the automated system**

Once a federal court order got Idaho's secret formulas into our hands, we had to figure out how they worked and how Idaho had developed them. We learned, through this process, that the state agency's cursory internal documentation of its system could not sufficiently explain it. It took months of litigation discovery to gather the evidence necessary to evaluate the system, figure out how it worked, and identify the assumptions and data it relied on. Agency officials pointed the finger at each other when asked to explain the system's basic functions, and ultimately none could meaningfully explain it at all. We had to spend over \$40 thousand on a statistician, a Medicaid resource allocation specialist, and a developmental disability expert to reverse engineer the system, catalog its flaws, and assess the harm its results could wreak upon our clients.

Notably, the lack of transparency compounded this system's harms. Not only did Idaho create and deploy a tool that made erroneous determinations, but it did so without understanding how the tool worked and without adequate processes to detect the errors.

- **Corrupt and discriminatory data**

As our experts evaluated the automated system after we got our hands on it, they discovered that it was built from corrupt and biased data. The inputs that drove the system's formulas, for starters, came partly from a homegrown assessment instrument that the State had never validated or standardized. And the State had done nothing to ensure consistency across different assessors or with the same assessor over time.

Then, out of the data the State compiled to compute the statistical formulas at the heart of its system, more than two thirds of the records were either plainly erroneous, mismatched with the agency's systems, or contained incomplete or unbelievable information.

Further analysis then revealed that data from one of Idaho's most populous regions was underrepresented in the data sample the Department used. This oversight made a substantial difference, unjustifiably biasing the system's operation.

- **Black boxes preventing accountability**

Compounding all these problems, Idaho has also repeatedly fought to withhold from Medicaid recipients the very information they need to challenge the automated system's results. During the initial round of litigation in this case, the State tried to ban my clients from using the assessment booklet to challenge their budgets. The State lost in court, with the judge ruling that such a ban would violate constitutional due process guarantees. Automated systems must preserve the rights of those they make decisions about to challenge errors and cross-examine assessors and others whose determinations impact the system's results.

This spring, yet again, the State fought to withhold critical information about the new assessment it planned to implement. Taking up the cause of its private contractor, rather than the constitutional rights of Idahoans relying on Medicaid, the State tried to ban my clients from accessing the manual that sets out what assessors must do for their assessments to be valid. While we awaited a court ruling on such a ban, the private contractor pulled out of its relationship with the State. When governments spend taxpayer funds on contractors who claim proprietary protections in public programs that make decisions about individual lives, private interests can hold constitutional rights hostage. Federal policy and federal agencies supervising federal programs like Medicaid, and not just federal courts, should prohibit this. As another court put it in a similar case: "When a public agency adopts a policy of making high stakes . . . decisions based on secret algorithms incompatible with minimum due process, the proper remedy is to overturn the policy, while leaving the trade secrets intact."<sup>4</sup>

## **Solutions**

Since the *K.W. v. Armstrong* lawsuit was filed in 2012, there has been an explosion of attention on AI and automated decision-making by policymakers, scholars, and scientists. At each stage of the case, we've had more and better resources to turn to as we try to remedy the problems with Idaho's system. The lawsuit has served as a laboratory as the court and the parties have explored the complexities of automated decision-making. Some of the most salient and important solutions these lessons point to fall into three categories:

<sup>4</sup> *Houston Federation of Teachers v. Houston Independent School Dist.*, 251 F. Supp. 3d 1168, 1179 (S.D. Tex. 2017).

- **The people who AI and automated systems make decisions about must be integrally involved with the development, implementation, and assessment of those systems**

This solution is the most important, because if earnestly implemented it will be the most effective at preventing the dangers AI poses. In my own work for over a decade litigating automated decision-making in Idaho’s Medicaid system, I have failed again and again to spot systemic problems that my clients experience as everyday realities. Only by cultivating long-term dialogue with people using Medicaid services across Idaho have I discovered some of the system’s many veiled flaws. Bureaucrats, advocates, courts, computer scientists, and other technical experts do not have the experience necessary to assess automated decision-making systems and understand their impacts at the depth that those who these systems make decisions about can.

A cornerstone of the court-ordered settlement agreement in the *K.W. v. Armstrong* case addressed this, requiring the Idaho state agency to “encourage engagement and active involvement of class members, their guardians, and other community stakeholders” as it developed a new system, and to do so “throughout the development of the new [system].”<sup>5</sup> To ensure that involvement would be meaningful, the settlement agreement also requires that communications engaging with my clients about the new system’s development must “use clear language and layout, appropriate to the circumstances of the class members and their guardians.”<sup>6</sup>

Policies and regulations governing uses of AI and automated decision-making in government programs should include the same requirements; and appropriate agencies should enforce those requirements. This is consistent with the very first principal of the *Blueprint for an AI Bill of Rights*, which begins by instructing that “[a]utomated systems should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system.”<sup>7</sup>

<sup>5</sup> Class Action Settlement Agreement, *K.W. v. Armstrong*, Doc. no. 306-1, at 10 (D. Idaho Sept. 15, 2016).

<sup>6</sup> *Id.*

<sup>7</sup> *Blueprint for an AI Bill of Rights* at 5 (full citation above at note 2).

- **Federal and state agencies must protect constitutional rights to due process, transparency, and equal protection through regulation and enforcement specific to AI and automated systems**

For public programs that individuals and families depend on for their health and safety, like Medicaid, their constitutional and statutory rights to due process, transparency, and equal protection are, largely, already well established. But, as the *K.W. v. Armstrong* litigation continues to demonstrate, those rights must be carefully guarded, and courts are not best suited to make sure that governments respect civil and constitutional rights when they implement automated systems. Litigation is immensely expensive and time-intensive: this one lawsuit about one automated system in one bureau within the Medicaid program in one state with a population of just 2 million people is still going after 11 years and hundreds of thousands of dollars in costs and attorney fees. Case-by-case litigation over constitutional rights in programs like this will never meet the rising tide of these AI systems, nor can it fully undo the harms that the application of these systems can cause to impacted individuals and communities.

Rather, agencies, such as the Centers for Medicare and Medicaid Services that administers the federal Medicaid program, must prescribe regulations to govern the use, implementation, and testing of automated systems used within particular programs—and must also closely monitor and enforce compliance with those regulations. These regulations should address not only the newest, most cutting-edge versions of automated-decision systems such as generative AI but must likewise address the simpler but no less impactful algorithms and automated systems that are already used to make critical decisions throughout the government.<sup>8</sup>

As our litigation demonstrates, agencies should play a driving role. Agencies must set standards for assessing whether and how automated decision-making should be used at all, along with standards for independent pre- and post-deployment audits that center civil rights and civil liberties as well as safety and effectiveness. Agencies must also erect and enforce guardrails for determining when mitigation or decommission measures are appropriate. Congress and federal agencies are especially well positioned as stewards of

<sup>8</sup> David Freeman Engstrom et al., *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies* (Feb. 2020), <https://www.acus.gov/sites/default/files/documents/Government%20by%20Algorithm.pdf> (finding nearly half of the 142 largest administrative agencies use artificial intelligence for diverse purposes including regulatory analysis, enforcement, and adjudication).

federal tax dollars to lead in implementing such standards in federally funded programs like Medicaid.

In addition to assessments and audits, programs considering or already using automated systems must implement proper notice for the people against whom these systems are deployed, plus robust opt-out and human review processes. Especially in programs like the one in *K.W. v. Armstrong*, which my clients rely on for their day-to-day safety and survival, the risk of harm from erroneous decisions is extraordinary. Plus, my clients' disabilities and indigency make conventional administrative appeal processes inaccessible and often impossible for them to use. When using automated systems, government agencies must make it easy and accessible for people to "raise their hand" to opt out or have an informed human review automated results to check their validity. Those who these systems make decisions about should also always have private rights of action so they can access judicial review if agency-level protections fail.

- **Government AI and automated systems must be fully transparent and subject to clear standards from before they start until after they finish**

Black boxes have plagued the Idaho system challenged in the *K.W. v. Armstrong* case from its filing in 2012 through to the present, in 2023. The system has been shrouded in black boxes around its use and implementation, its functioning, and critical information to challenge its validity and results. The state agency's involvement with private contractors claiming proprietary interests in decision-making methods has exacerbated these black box problems. Most recently, one of those contractors has refused to cooperate with the agency rather than risk the transparency required by constitutional due process.

Beginning when government programs first consider using a new automated system, and then throughout each system's design and implementation, as well as during testing and post-use evaluation, government AI and automated systems must be fully transparent. Those who these systems make decisions about have due process rights to challenge those decisions and equal protection rights against bias. They and the public must always have access to the same information that the government does about whether automated systems are in use, how these systems work, the data they were trained on, any algorithms and instructions they follow, and all testing and analysis results. Moreover, it is critical that the auditing regime include external audits and community review. Private interests in trade secrets or proprietary methods and materials can never be allowed to trump individual due process rights, and taxpayers must retain robust access to oversee these public systems.

\* \* \* \* \*

I thank the Committee for its attention to the litigation in Idaho and the use of AI and automated decision-making in government programs generally. Thank you for the opportunity to testify. I look forward to your questions.

Jacob Siegel -- HSGAC -- May 16, 2023

**NEWS**

# A Guide to Understanding the Hoax of the Century

Thirteen ways of looking at disinformation

BY JACOB SIEGEL

MARCH 28, 2023

---

## **PROLOGUE: THE INFORMATION WAR**

In 1950, Sen. Joseph McCarthy claimed that he had proof of a communist spy ring operating inside the government. Overnight, the explosive accusations blew up in the national press, but the details kept changing. Initially, McCarthy said he had a list with the names of 205 communists in the State Department; the next day he revised it to 57. Since he kept the list a secret, the inconsistencies were beside the point. The point was the power of the accusation, which made McCarthy's name synonymous with the politics of the era.

For more than half a century, McCarthyism stood as a defining chapter in the worldview of American liberals: a warning about the dangerous allure of blacklists, witch hunts, and demagogues.

Until 2017, that is, when another list of alleged Russian agents roiled the American press and political class. A new outfit called Hamilton 68 claimed to have discovered hundreds of Russian-affiliated accounts that had infiltrated Twitter to sow chaos and help Donald Trump win the election. Russia stood accused of

hacking social media platforms, the new centers of power, and using them to covertly direct events inside the United States.

None of it was true. After reviewing Hamilton 68's secret list, Twitter's safety officer, Yoel Roth, privately admitted that his company was allowing "real people" to be "unilaterally labeled Russian stooges without evidence or recourse."

The Hamilton 68 episode played out as a nearly shot-for-shot remake of the McCarthy affair, with one important difference: McCarthy faced some resistance from leading journalists as well as from the U.S. intelligence agencies and his fellow members of Congress. In our time, those same groups lined up to support the new secret lists and attack anyone who questioned them.

When proof emerged earlier this year that Hamilton 68 was a high-level hoax perpetrated against the American people, it was met with a great wall of silence in the national press. The disinterest was so profound, it suggested a matter of principle rather than convenience for the standard-bearers of American liberalism who had lost faith in the promise of freedom and embraced a new ideal.

In his last days in office, President Barack Obama made the decision to set the country on a new course. On Dec. 23, 2016, he signed into law the Countering Foreign Propaganda and Disinformation Act, which used the language of defending the homeland to launch an open-ended, offensive information war.

Something in the looming specter of Donald Trump and the populist movements of 2016 reawakened sleeping monsters in the West. Disinformation, a half-forgotten relic of the Cold War, was newly spoken of as an urgent, existential threat. Russia was said to have exploited the vulnerabilities of the open internet to bypass U.S. strategic defenses by infiltrating private citizens' phones and laptops. The Kremlin's endgame was to colonize the minds of its targets, a tactic cyber warfare specialists call "cognitive hacking."

Defeating this specter was treated as a matter of national survival. “The U.S. Is Losing at Influence Warfare,” warned a December 2016 article in the defense industry journal, *Defense One*. The article quoted two government insiders arguing that laws written to protect U.S. citizens from state spying were jeopardizing national security. According to Rand Waltzman, a former program manager at the Defense Advanced Research Projects Agency, America’s adversaries enjoyed a “significant advantage” as the result of “legal and organizational constraints that we are subject to and they are not.”

The point was echoed by Michael Lumpkin, who headed the State Department’s Global Engagement Center (GEC), the agency Obama designated to run the U.S. counter-disinformation campaign. Lumpkin singled out the Privacy Act of 1974, a post-Watergate law protecting U.S. citizens from having their data collected by the government, as antiquated. “The 1974 act was created to make sure that we aren’t collecting data on U.S. citizens. Well, ... by definition the World Wide Web is worldwide. There is no passport that goes with it. If it’s a Tunisian citizen in the United States or a U.S. citizen in Tunisia, I don’t have the ability to discern that ... If I had more ability to work with that [personally identifiable information] and had access ... I could do more targeting, more definitively, to make sure I could hit the right message to the right audience at the right time.”

The message from the U.S. defense establishment was clear: To win the information war—an existential conflict taking place in the borderless dimensions of cyberspace—the government needed to dispense with outdated legal distinctions between foreign terrorists and American citizens.

Since 2016, the federal government has spent billions of dollars on turning the counter-disinformation complex into one of the most powerful forces in the modern world: a sprawling leviathan with tentacles reaching into both the public and private sector, which the government uses to direct a “whole of society” effort that aims to seize total control over the internet and achieve nothing less than the eradication of human error.

Step one in the national mobilization to defeat disinfo fused the U.S. national security infrastructure with the social media platforms, where the war was being fought. The government's lead counter-disinformation agency, the GEC, declared that its mission entailed "seeking out and engaging the best talent within the technology sector." To that end, the government started deputizing tech executives as de facto wartime information commissars.

At companies like Facebook, Twitter, Google, and Amazon, the upper management levels had always included veterans of the national security establishment. But with the new alliance between U.S. national security and social media, the former spooks and intelligence agency officials grew into a dominant bloc inside those companies; what had been a career ladder by which people stepped up from their government experience to reach private tech-sector jobs turned into an ouroboros that molded the two together. With the D.C.-Silicon Valley fusion, the federal bureaucracies could rely on informal social connections to push their agenda inside the tech companies.

In the fall of 2017, the FBI opened its Foreign Influence Task Force for the express purpose of monitoring social media to flag accounts trying to "discredit U.S. individuals and institutions." The Department of Homeland Security took on a similar role.

At around the same time, Hamilton 68 blew up. Publicly, Twitter's algorithms turned the Russian-influence-exposing "dashboard" into a major news story. Behind the scenes, Twitter executives quickly figured out that it was a scam. When Twitter reverse-engineered the secret list, it found, according to the journalist Matt Taibbi, that "instead of tracking how Russia influenced American attitudes, Hamilton 68 simply collected a handful of mostly real, mostly American accounts and described their organic conversations as Russian scheming." The discovery prompted Twitter's head of trust and safety, Yoel Roth, to suggest in an October

2017 email that the company take action to expose the hoax and “call this out on the bullshit it is.”

In the end, neither Roth nor anyone else said a word. Instead, they let a purveyor of industrial-grade bullshit—the old-fashioned term for *disinformation*—continue dumping its contents directly into the news stream.

It was not enough for a few powerful agencies to combat disinformation. The strategy of national mobilization called for “not only the whole-of-government, but also whole-of-society” approach, according to a document released by the GEC in 2018. “To counter propaganda and disinformation,” the agency stated, “will require leveraging expertise from across government, tech and marketing sectors, academia, and NGOs.”

This is how the government-created “war against disinformation” became the great moral crusade of its time. CIA officers at Langley came to share a cause with hip young journalists in Brooklyn, progressive nonprofits in D.C., George Soros-funded think tanks in Prague, racial equity consultants, private equity consultants, tech company staffers in Silicon Valley, Ivy League researchers, and failed British royals. Never Trump Republicans joined forces with the Democratic National Committee, which declared online disinformation “a whole-of-society problem that requires a whole-of-society response.”

Even trenchant critics of the phenomenon—including Taibbi and the *Columbia Journalism Review*'s Jeff Gerth, who recently published a dissection of the press's role in promoting false Trump-Russia collusion claims—have focused on the media's failures, a framing largely shared by conservative publications, which treat disinformation as an issue of partisan censorship bias. But while there's no question that the media has utterly disgraced itself, it's also a convenient fall guy—by far the weakest player in the counter-disinformation complex. The American

press, once the guardian of democracy, was hollowed out to the point that it could be worn like a hand puppet by the U.S. security agencies and party operatives. It would be nice to call what has taken place a tragedy, but an audience is meant to learn something from a tragedy. As a nation, America not only has learned nothing, it has been deliberately prevented from learning anything while being made to chase after shadows. This is not because Americans are stupid; it's because what has taken place is not a tragedy but something closer to a crime. Disinformation is both the name of the crime and the means of covering it up; a weapon that doubles as a disguise.

The crime is the information war itself, which was launched under false pretenses and by its nature destroys the essential boundaries between the public and private and between the foreign and domestic, on which peace and democracy depend. By conflating the anti-establishment politics of domestic populists with acts of war by foreign enemies, it justified turning weapons of war against Americans citizens. It turned the public arenas where social and political life take place into surveillance traps and targets for mass psychological operations. The crime is the routine violation of Americans' rights by unelected officials who secretly control what individuals can think and say.

What we are seeing now, in the revelations exposing the inner workings of the state-corporate censorship regime, is only the end of the beginning. The United States is still in the earliest stages of a mass mobilization that aims to harness every sector of society under a singular technocratic rule. The mobilization, which began as a response to the supposedly urgent menace of Russian interference, now evolves into a regime of total information control that has arrogated to itself the mission of eradicating abstract dangers such as error, injustice, and harm—a goal worthy only of leaders who believe themselves to be infallible, or comic-book supervillains.

The first phase of the information war was marked by distinctively human displays of incompetence and brute-force intimidation. But the next stage, already underway, is being carried out through both scalable processes of artificial

intelligence and algorithmic pre-censorship that are invisibly encoded into the infrastructure of the internet, where they can alter the perceptions of billions of people.

Something monstrous is taking shape in America. Formally, it exhibits the synergy of state and corporate power in service of a tribal zeal that is the hallmark of fascism. Yet anyone who spends time in America and is not a brainwashed zealot can tell that it is not a fascist country. What is coming into being is a new form of government and social organization that is as different from mid-twentieth century liberal democracy as the early American republic was from the British monarchism that it grew out of and eventually supplanted. A state organized on the principle that it exists to protect the sovereign rights of individuals, is being replaced by a digital leviathan that wields power through opaque algorithms and the manipulation of digital swarms. It resembles the Chinese system of social credit and one-party state control, and yet that, too, misses the distinctively American and providential character of the control system. In the time we lose trying to name it, the thing itself may disappear back into the bureaucratic shadows, covering up any trace of it with automated deletions from the top-secret data centers of Amazon Web Services, “the trusted cloud for government.”

When the blackbird flew out of sight,  
It marked the edge  
Of one of many circles.

In a technical or structural sense, the censorship regime’s aim is not to censor or to oppress, but to rule. That’s why the authorities can never be labeled as guilty of disinformation. Not when they lied about Hunter Biden’s laptops, not when they claimed that the lab leak was a racist conspiracy, not when they said that vaccines stopped transmission of the novel coronavirus. Disinformation, now and for all time, is whatever they say it is. That is not a sign that the concept is being misused or corrupted; it is the precise functioning of a totalitarian system.

If the underlying philosophy of the war against disinformation can be expressed in a single claim, it is this: You cannot be trusted with your own mind. What follows

is an attempt to see how this philosophy has manifested in reality. It approaches the subject of disinformation from 13 angles—like the “Thirteen Ways of Looking at a Blackbird,” Wallace Stevens’ 1917 poem—with the aim that the composite of these partial views will provide a useful impression of disinformation’s true shape and ultimate design.

---

## **CONTENTS**

I. Russophobia Returns, Unexpectedly: The Origins of Contemporary “Disinformation”

II. Trump’s Election: “It’s Facebook’s Fault”

III. Why Do We Need All This Data About People?

IV. The Internet: From Darling to Demon

V. Russiagate! Russiagate! Russiagate!

VI. Why the Post-9/11 “War on Terror” Never Ended

VII. The Rise of “Domestic Extremists”

VIII. The NGO Borg

IX. COVID-19

X. Hunter’s Laptops: The Exception to the Rule

XI. The New One-Party State

XII. The End of Censorship

XIII. After Democracy

Appendix: The Disinfo Dictionary

*Have insider information on the counter-disinformation complex? Email [jacobsiegel@protonmail.com](mailto:jacobsiegel@protonmail.com) or contact him or contact him on Twitter @jacob\_\_siegel.*

## **I. Russophobia Returns, Unexpectedly: The Origins of Contemporary “Disinformation”**

The foundations of the current information war were laid in response to a sequence of events that took place in 2014. First Russia tried to suppress the U.S.-backed Euromaidan movement in Ukraine; a few months later Russia invaded Crimea; and several months after that the Islamic State captured the city of Mosul in northern Iraq and declared it the capital of a new caliphate. In three separate conflicts, an enemy or rival power of the United States was seen to have successfully used not just military might but also social media messaging campaigns designed to confuse and demoralize its enemies—a combination known as “hybrid warfare.” These conflicts convinced U.S. and NATO security officials that the power of social media to shape public perceptions had evolved to the point where it could decide the outcome of modern wars—outcomes that might be counter to those the United States wanted. They concluded that the state had to acquire the means to take control over digital communications so that they could present reality as they wanted it to be, and prevent reality from becoming anything else.

Technically, *hybrid warfare* refers to an approach that combines military and non-military means—overt and covert operations mixed with cyberwarfare and influence operations—to both confuse and weaken a target while avoiding direct, full-scale conventional war. In practice, it is notoriously vague. “The term now covers every type of discernible Russian activity, from propaganda to conventional warfare, and most that exists in between,” wrote Russia analyst Michael Kofman in March 2016.

Over the past decade, Russia has indeed repeatedly employed tactics associated with hybrid warfare, including a push to target Western audiences with messaging on channels like RT and Sputnik News and with cyber operations such as the use of “troll” accounts. But this was not new even in 2014, and it was something the United States, as well as every other major power, engaged in as well. As early as 2011, the United States was building its own “troll armies” online by developing software to “secretly manipulate social media sites by using fake online personas to influence internet conversations and spread pro-American propaganda.”

“If you torture hybrid warfare long enough, it will tell you anything,” Kofman had admonished, which is precisely what began happening a few months later when Trump critics popularized the idea that a hidden Russian hand was the puppeteer of political developments inside the United States.

The leading voice promoting that claim was a former FBI officer and counterterrorism analyst named Clint Watts. In an article from August 2016, “How Russia Dominates Your Twitter Feed to Promote Lies (And, Trump, Too),” Watts and his co-author, Andrew Weisburd, described how Russia had revived its Cold War-era “Active Measures” campaign, using propaganda and disinformation to influence foreign audiences. As a result, according to the article, Trump voters and Russian propagandists were promoting the same stories on social media that were intended to make America look weak and incompetent. The authors made the extraordinary claim that the “melding of Russian-friendly accounts and Trumpkins has been going on for some time.” If that was true, it meant that anyone expressing support for Donald Trump might be an agent of the Russian government, whether

or not the person intended to play that role. It meant that the people they called “Trumpkins,” who made up half the country, were attacking America from within. It meant that politics was now war, as it is in many parts of the world, and tens of millions of Americans were the enemy.

Watts made his name as a counterterrorism analyst by studying the social media strategies used by ISIS, but with articles like this, he became the media’s go-to expert on Russian trolls and Kremlin disinformation campaigns. It seems he also had powerful backers.

In his book *The Assault on Intelligence*, retired CIA chief Michael Hayden called Watts “the one man, who more than any other was trying to ring the alarm more than two years before the 2016 elections.”

Hayden credited Watts in his book with teaching him the power of social media: “Watts pointed out to me that Twitter makes falsehoods seem more believable through sheer repetition and volume. He labeled it a kind of ‘computational propaganda.’ Twitter in turn drives mainstream media.”

A false story algorithmically amplified by Twitter and disseminated by the media—it’s no coincidence that this perfectly describes the “bullshit” spread on Twitter about Russian influence operations: In 2017, it was Watts who came up with the idea for the Hamilton 68 dashboard and helped spearhead the initiative.

## II. Trump’s Election: “It’s Facebook’s Fault”

No one thought Trump was a normal politician. Being an ogre, Trump horrified millions of Americans who felt a personal betrayal in the possibility that he would occupy the same office held by George Washington and Abe Lincoln. Trump also threatened the business interests of the most powerful sectors of society. It was the

latter offense, rather than his putative racism or flagrant un-presidentialness, that sent the ruling class into a state of apoplexy.

Given his focus in office on lowering the corporate tax rate, it's easy to forget that Republican officials and the party's donor class saw Trump as a dangerous radical who threatened their business ties with China, their access to cheap imported labor, and the lucrative business of constant war. But, indeed, that is how they saw him, as reflected in the unprecedented response to Trump's candidacy recorded by *The Wall Street Journal* in September 2016: "No chief executive at the nation's 100 largest companies had donated to Republican Donald Trump's presidential campaign through August, a sharp reversal from 2012, when nearly a third of the CEOs of Fortune 100 companies supported GOP nominee Mitt Romney."

The phenomenon was not unique to Trump. Bernie Sanders, the left-wing populist candidate in 2016, was also seen as a dangerous threat by the ruling class. But whereas the Democrats successfully sabotaged Sanders, Trump made it past his party's gatekeepers, which meant that he had to be dealt with by other means.

Two days after Trump took office, a smirking Senator Chuck Schumer told MSNBC's Rachel Maddow that it was "really dumb" of the new president to get on the bad side of the security agencies that were supposed to work for him: "Let me tell you, you take on the intelligence community, they have six ways from Sunday of getting back at you."

Trump had used sites like Twitter to bypass his party's elites and connect directly with his supporters. Therefore, to cripple the new president and ensure that no one like him could ever come to power again, the intel agencies had to break the independence of the social media platforms. Conveniently, it was the same lesson that many intelligence and defense officials had drawn from the ISIS and Russian campaigns of 2014—namely, that social media was too powerful to be left outside of state control—only applied to domestic politics, which meant the agencies would now have help from politicians who stood to benefit from the effort.

Immediately after the election, Hillary Clinton started blaming Facebook for her loss. Until this point, Facebook and Twitter had tried to remain above the political fray, fearful of jeopardizing potential profits by alienating either party. But now a profound change occurred, as the operation behind the Clinton campaign reoriented itself not simply to reform the social media platforms, but to conquer them. The lesson they took from Trump's victory was that Facebook and Twitter—more than Michigan and Florida—were the critical battlegrounds where political contests were won or lost. “Many of us are beginning to talk about what a big problem this is,” Clinton's chief digital strategist Teddy Goff told Politico the week after the election, referring to Facebook's alleged role in boosting Russian disinformation that helped Trump. “Both from the campaign and from the administration, and just sort of broader Obama orbit...this is one of the things we would like to take on post-election,” Goff said.

The press repeated that message so often that it gave the political strategy the appearance of objective validity:

“Donald Trump Won Because of Facebook”; *New York Magazine*, Nov. 9, 2016.

“Facebook, in Cross Hairs After Election, Is Said to Question Its Influence”; *The New York Times*, Nov. 12, 2016.

“Russian propaganda effort helped spread ‘fake news’ during election, experts say”; *The Washington Post*, Nov. 24, 2016.

“Disinformation, Not Fake News, Got Trump Elected, and It Is Not Stopping”; *The Intercept*, Dec. 6, 2016.

And on it went in countless articles that dominated the news cycle for the next two years.

At first, Facebook's CEO Mark Zuckerberg dismissed the charge that fake news posted on his platform had influenced the outcome of the election as “pretty crazy.” But Zuckerberg faced an intense pressure campaign in which every sector of the

American ruling class, including his own employees, blamed him for putting a Putin agent in the White House, effectively accusing him of high treason. The final straw came a few weeks after the election when Obama himself “publicly denounced the spread of fake news on Facebook.” Two days later, Zuckerberg folded: “Facebook announces new push against fake news after Obama comments.” The false yet foundational claim that Russia hacked the 2016 election provided a justification—just like the claims about weapons of mass destruction that triggered the Iraq War—to plunge America into a wartime state of exception. With the normal rules of constitutional democracy suspended, a coterie of party operatives and security officials then installed a vast, largely invisible new architecture of social control on the backend of the internet’s biggest platforms.

Though there was never a public order given, the U.S. government began enforcing martial law online.

### **III. Why Do We Need All This Data About People?**

The American doctrine of counterinsurgency (COIN) warfare famously calls for “winning hearts and minds.” The idea is that victory against insurgent groups depends on gaining the support of the local population, which cannot be accomplished by brute force alone. In places like Vietnam and Iraq, support was secured through a combination of nation-building and appealing to locals by providing them with goods they were presumed to value: money and jobs, for instance, or stability.

Because cultural values vary and what is prized by an Afghan villager may appear worthless to a Swedish accountant, successful counterinsurgents must learn what

makes the native population tick. To win over a mind, first you have to get inside it to understand its wants and fears. When that fails, there is another approach in the modern military arsenal to take its place: counterterrorism. Where counterinsurgency tries to win local support, counterterrorism tries to hunt down and kill designated enemies.

Despite the apparent tension in their contrasting approaches, the two strategies have often been used in tandem. Both rely on extensive surveillance networks to gather intelligence on their targets, whether that is figuring out where to dig wells or locating terrorists in order to kill them. But the counterinsurgent in particular imagines that if he can learn enough about a population, it will be possible to reengineer its society. Obtaining answers is just a matter of using the right resources: a combination of surveillance tools and social scientific methods, the joint output of which feeds into all-powerful centralized databases that are believed to contain the totality of the war.

I have observed, reflecting on my experiences as a U.S. Army intelligence officer in Afghanistan, how, “data analytics tools at the fingertips of anyone with access to an operations center or situation room seemed to promise the imminent convergence of map and territory,” but ended up becoming a trap as “U.S. forces could measure thousands of different things that we couldn’t understand.” We tried to cover for that deficit by acquiring even more data. If only we could gather enough information and harmonize it with the correct algorithms, we believed, the database would divine the future.

Not only is that framework foundational in modern American counterinsurgency doctrine, but also it was part of the original impetus for building the internet. The Pentagon built the proto-internet known as ARPANET in 1969 because it needed a decentralized communications infrastructure that could survive nuclear war—but that was not the only goal. The internet, writes Yasha Levine in his history of the subject, *Surveillance Valley*, was also “an attempt to build computer systems that could collect and share intelligence, watch the world in real time, and study and analyze people and political movements with the ultimate goal of predicting and preventing social upheaval. Some even dreamed of creating a sort of early

warning radar for human societies: a networked computer system that watched for social and political threats and intercepted them in much the same way that traditional radar did for hostile aircraft.”

In the days of the internet “freedom agenda,” the popular mythology of Silicon Valley depicted it as a laboratory of freaks, self-starters, free thinkers, and libertarian tinkerers who just wanted to make cool things without the government slowing them down. The alternative history, outlined in Levine’s book, highlights that the internet “always had a dual-use nature rooted in intelligence gathering and war.” There is truth in both versions, but after 2001 the distinction disappeared.

As Shoshana Zuboff writes in *The Age of Surveillance Capitalism*, at the start of the war on terror “the elective affinity between public intelligence agencies and the fledgling surveillance capitalist Google blossomed in the heat of emergency to produce a unique historical deformity: surveillance exceptionalism.”

In Afghanistan, the military had to employ costly drones and “Human Terrain Teams” staffed with adventurous academics to survey the local population and extract their relevant sociological data. But with Americans spending hours a day voluntarily feeding their every thought directly into data monopolies connected to the defense sector, it must have seemed trivially easy for anyone with control of the databases to manipulate the sentiments of the population at home.

More than a decade ago, the Pentagon began funding the development of a host of tools for detecting and countering terrorist messaging on social media. Some were part of a broader “memetic warfare” initiative inside the military that included proposals to weaponize memes to “defeat an enemy ideology and win over the masses of undecided noncombatants.” But most of the programs, launched in response to the rise of ISIS and the jihadist group’s adept use of social media, focused on scaling up automated means of detecting and censoring terrorist messaging online. Those efforts culminated in January 2016 with the State Department’s announcement that it would be opening the aforementioned Global Engagement Center, headed by Michael Lumpkin. Just a few months later, President Obama put the GEC in charge of the new war against disinformation.

On the same day that the GEC was announced, Obama and “various high-ranking members of the national security establishment met with representatives from Facebook, Twitter, YouTube, and other Internet powerhouses to discuss how the United States can fight ISIS messaging via social media.”

In the wake of the populist upheavals of 2016, leading figures in America’s ruling party seized upon the feedback loop of surveillance and control refined through the war on terror as a method for maintaining power inside the United States. Weapons created to fight ISIS and al-Qaeda were turned against Americans who entertained incorrect thoughts about the president or vaccine boosters or gender pronouns or the war in Ukraine.

Former State Department official Mike Benz, who now runs an organization called the Foundation for Freedom Online that bills itself as a digital free-speech watchdog, describes how a company called Graphika, which is “essentially a U.S. Department of Defense-funded censorship consortium” that was created to fight terrorists, was repurposed to censor political speech in America. The company, “initially funded to help do social media counterinsurgency work effectively in conflict zones for the U.S. military,” was then “redeployed domestically both on Covid censorship and political censorship,” Benz told an interviewer. “Graphika was deployed to monitor social media discourse about Covid and Covid origins, Covid conspiracies, or Covid sorts of issues.”

The fight against ISIS morphed into the fight against Trump and “Russian collusion,” which morphed into the fight against disinformation. But those were just branding changes; the underlying technological infrastructure and ruling-class philosophy, which claimed the right to remake the world based on a religious sense of expertise, remained unchanged. The human art of politics, which would have required real negotiation and compromise with Trump supporters, was abandoned in favor of a specious science of top-down social engineering that aimed to produce a totally administered society.

For the American ruling class, COIN replaced politics as the proper means of dealing with the natives.

## IV. The Internet: From Darling to Demon

Once upon a time, the internet was going to save the world. The first dot-com boom in the 1990s popularized the idea of the internet as a technology for maximizing human potential and spreading democracy. The Clinton administration's 1997 "A Framework for Global Electronic Commerce" put forth the vision: "The Internet is a medium that has tremendous potential for promoting individual freedom and individual empowerment" and "[t]herefore, where possible, the individual should be left in control of the way in which he or she uses this medium." The smart people in the West mocked the naive efforts in other parts of the world to control the flow of information. In 2000, President Clinton scoffed that China's internet crackdown was "like trying to nail Jell-O to the wall." The hype continued through the Bush administration, when internet companies were seen as crucial partners in the state's mass surveillance program and its plan to bring democracy to the Middle East.

But the hype really went into overdrive when President Obama was elected through a "big data"-driven campaign that prioritized social media outreach. There appeared to be a genuine philosophical alignment between Obama's political style as the "Hope" and "Change" president whose guiding principle in foreign policy was "Don't do dumb shit" and the internet search company whose original motto was "Do no evil." There were also deep personal ties connecting the two powers, with 252 cases over the course of Obama's presidency of people moving between jobs at the White House and Google. From 2009 to 2015, White House and Google employees were meeting, on average, more than once a week.

As Obama's secretary of state, Hillary Clinton led the government's "Internet freedom" agenda, which aimed to "promote online communications as a tool for opening up closed societies." In a speech from 2010, Clinton issued a warning about the spread of digital censorship in authoritarian regimes: "A new

information curtain is descending across much of the world,” she said. “And beyond this partition, viral videos and blog posts are becoming the samizdat of our day.”

It is a supreme irony that the very people who a decade ago led the freedom agenda for other countries have since pushed the United States to implement one of the largest and most powerful censorship machines in existence under the guise of fighting disinformation.

Or perhaps *irony* is not the right word to capture the difference between the freedom-loving Clinton of a decade ago and the pro-censorship activist of today, but it gets at what appears to be the about-face done by a class of people who were public standard-bearers for radically different ideas barely 10 years earlier. These people—politicians, first and foremost—saw (and presented) internet freedom as a positive force for humanity when it empowered them and served their interests, but as something demonic when it broke down those hierarchies of power and benefited their opponents. That’s how to bridge the gap between the Hillary Clinton of 2013 and the Clinton of 2023: Both see the internet as an immensely powerful tool for driving political processes and effecting regime change.

Which is why, in the Clinton and Obama worlds, the rise of Donald Trump looked like a profound betrayal—because, as they saw it, Silicon Valley could have stopped it but didn’t. As heads of the government’s internet policy, they had helped the tech companies build their fortunes on mass surveillance and evangelized the internet as a beacon of freedom and progress while turning a blind eye to their flagrant violations of antitrust statutes. In return, the tech companies had done the unthinkable—not because they had allowed Russia to “hack the election,” which was a desperate accusation thrown out to mask the stench of failure, but because they refused to intervene to prevent Donald Trump from winning.

In his book *Who Owns the Future?*, tech pioneer Jaron Lanier writes, “The primary business of digital networking has come to be the creation of ultra-secret mega-dossiers about what others are doing, and using this information to concentrate

money and power.” Because digital economies produce ever-greater concentrations of data and power, the inevitable happened: The tech companies got too powerful. What could the leaders of the ruling party do? They had two options. They could use the government’s regulatory power to counter-attack: Break up the data monopolies and restructure the social contract underwriting the internet so that individuals retained ownership of their data instead of having it ripped off every time they clicked into a public commons. Or, they could preserve the tech companies’ power while forcing them to drop the pretense of neutrality and instead line up behind the ruling party—a tempting prospect, given what they could do with all that power.

They chose option B.

Declaring the platforms guilty of electing Trump—a candidate every bit as loathsome to the highly educated elites in Silicon Valley as he was to the highly educated elites in New York and D.C.—provided the club that the media and the political class used to beat the tech companies into becoming more powerful and more obedient.

## **V. Russiagate! Russiagate! Russiagate!**

If one imagines that the American ruling class faced a problem—Donald Trump appeared to threaten their institutional survival—then the Russia investigation didn’t just provide the means to unite the various branches of that class, in and out of government, against a common foe. It also gave them the ultimate form of leverage over the most powerful non-aligned sector of society: the tech industry. The coordination necessary to carry out the Russian collusion frame-up was the vehicle, fusing (1) the political goals of the Democratic Party, (2) the institutional

agenda of the intelligence and security agencies, and (3) the narrative power and moral fervor of the media with (4) the tech companies' surveillance architecture. The secret FISA warrant that allowed U.S. security agencies to begin spying on the Trump campaign was based on the Steele dossier, a partisan hatchet job paid for by Hillary Clinton's team that consisted of provably false reports alleging a working relationship between Donald Trump and the Russian government. While a powerful short-term weapon against Trump, the dossier was also obvious bullshit, which suggested it might eventually become a liability.

Disinformation solved that problem while placing a nuclear-grade weapon in the arsenal of the anti-Trump resistance. In the beginning, disinformation had been only one among a half-dozen talking points coming from the anti-Trump camp. It won out over the others because it was capable of explaining anything and everything yet simultaneously remained so ambiguous it could not be disproved. Defensively, it provided a means to attack and discredit anyone who questioned the dossier or the larger claim that Trump colluded with Russia.

All the old McCarthyite tricks were new again. *The Washington Post* aggressively trumpeted the claim that disinformation swung the 2016 election, a crusade that began within days of Trump's victory, with the article "Russian propaganda effort helped spread 'fake news' during election, experts say." (The lead expert quoted in the article: Clint Watts.)

A steady flow of leaks from intelligence officials to national security reporters had already established the false narrative that there was credible evidence of collusion between the Trump campaign and the Kremlin. When Trump won in spite of those reports, the senior officials responsible for spreading them, most notably CIA chief John Brennan, doubled down on their claims. Two weeks before Trump took office, the Obama administration released a declassified version of an intelligence community assessment, known as an ICA, on "Russian Activities and Intentions in Recent Elections," which asserted that "Putin and the Russian government developed a clear preference for President-elect Trump."

The ICA was presented as the objective, nonpolitical consensus reached by multiple intelligence agencies. In the *Columbia Journalism Review*, Jeff Gerth writes that the assessment received “massive, and largely uncritical coverage” in the press. But, in fact, the ICA was just the opposite: a selectively curated political document that deliberately omitted contrary evidence to create the impression that the collusion narrative was not a widely disputed rumor, but an objective fact.

A classified report by the House Intelligence Committee on the creation of the ICA detailed just how unusual and nakedly political it was. “It wasn’t 17 agencies, and it wasn’t even a dozen analysts from the three agencies who wrote the assessment,” a senior intelligence official who read a draft version of the House report told the journalist Paul Sperry. “It was just five officers of the CIA who wrote it, and Brennan handpicked all five. And the lead writer was a good friend of Brennan’s.” An Obama appointee, Brennan had broken with precedent by weighing in on politics while serving as CIA director. That set the stage for his post-government career as an MSNBC analyst and “resistance” figure who made headlines by accusing Trump of treason.

Mike Pompeo, who succeeded Brennan at the CIA, said that as the agency’s director, he learned that “senior analysts who had been working on Russia for nearly their entire careers were made bystanders” when the ICA was being written. According to Sperry, Brennan “excluded conflicting evidence about Putin’s motives from the report, despite objections from some intelligence analysts who argued Putin counted on Clinton winning the election and viewed Trump as a ‘wild card.’” (Brennan was also the one who overrode the objections of other agencies to include the Steele dossier as part of the official assessment.)

Despite its irregularities, the ICA worked as intended: Trump began his presidency under a cloud of suspicion that he was never able to dispel. Just as Schumer promised, the intelligence officials wasted no time in taking their revenge.

And not only revenge, but also forward-planning action. The claim that Russia hacked the 2016 vote allowed federal agencies to implement the new public-private

ensorship machinery under the pretext of ensuring “election integrity.” People who expressed true and constitutionally protected opinions about the 2016 election (and later about issues like COVID-19 and the U.S. withdrawal from Afghanistan) were labeled un-American, racists, conspiracists, and stooges of Vladimir Putin and systematically removed from the digital public square to prevent their ideas from spreading disinformation. By an extremely conservative estimate based on public reporting, there have been tens of millions of such cases of censorship since Trump’s election.

And here’s the climax of this particular entry: On Jan. 6, 2017—the same day that Brennan’s ICA report lent institutional backing to the false claim that Putin helped Trump—Jeh Johnson, the outgoing Obama-appointed secretary of the Department of Homeland Security, announced that, in response to Russian electoral interference, he had designated U.S. election systems as “critical national infrastructure.” The move placed the property of 8,000 election jurisdictions across the country under the control of the DHS. It was a coup that Johnson had been attempting to pull off since the summer of 2016, but that, as he explained in a later speech, was blocked by local stakeholders who told him “that running elections in this country was the sovereign and exclusive responsibility of the states, and they did not want federal intrusion, a federal takeover, or federal regulation of that process.” So Johnson found a work-around by unilaterally rushing the measure through in his last days in office.

It’s clear now why Johnson was in such a rush: Within a few years, all of the claims used to justify the extraordinary federal seizure of the country’s electoral system would fall apart. In July 2019 the Mueller report concluded that Donald Trump did not collude with the Russian government—the same conclusion reached by the inspector general’s report into the origins of the Trump-Russia probe, released later that year. Finally, on Jan. 9, 2023, *The Washington Post* quietly published an addendum in its cybersecurity newsletter about New York University’s Center for Social Media and Politics study. Its conclusion: “Russian trolls on Twitter had little influence on 2016 voters.”

But by then it didn't matter. In the final two weeks of the Obama administration, the new counter-disinformation apparatus scored one of its most significant victories: the power to directly oversee federal elections that would have profound consequences for the 2020 contest between Trump and Joe Biden.

## **VI. Why the Post-9/11 “War on Terror” Never Ended**

Clint Watts, who headed up the Hamilton 68 initiative, and Michael Hayden, the former Air Force general, CIA chief, and NSA director who championed Watts, are both veterans of the U.S. counterterrorism establishment. Hayden ranks among the most senior intelligence officers the United States has ever produced and was a principal architect of the post-9/11 mass surveillance system. Indeed, an astounding percentage of the key figures in the counter-disinformation complex cut their teeth in the worlds of counterterrorism and counterinsurgency warfare.

Michael Lumpkin, who headed the GEC, the State Department agency that served as the first command center in the war against disinformation, is a former Navy SEAL with a counterterrorism background. The GEC itself grew out of the Center for Strategic Counterterrorism Communications before being repurposed to fight disinformation.

Twitter had the chance to stop the Hamilton 68 hoax before it got out of hand, yet chose not to. Why? The answer can be seen in the emails sent by a Twitter executive named Emily Horne, who advised against calling out the scam. Twitter had a smoking gun showing that the Alliance for Securing Democracy, the neoliberal think tank behind the Hamilton 68 initiative, was guilty of exactly the charge it made against others: peddling disinformation that inflamed domestic

political divisions and undermined the legitimacy of democratic institutions. But that had to be weighed against other factors, Horne suggested, such as the need to stay on the good side of a powerful organization. “We have to be careful in how much we push back on ASD publicly,” she wrote in February 2018.

The ASD was lucky to have someone like Horne on the inside of Twitter. Then again, maybe it wasn't luck. Horne had previously worked at the State Department, handling the “digital media and think tank outreach” portfolio. According to her [LinkedIn](#), she “worked closely with foreign policy reporters covering [ISIS] ... and executed communications plans relating to Counter-[ISIS] Coalition activities.” Put another way, she had a background in counterterrorism operations similar to Watts' but with more of an emphasis on spinning the press and civil society groups. From there she became the director for strategic communications for Obama's National Security Council, only leaving to join Twitter in June 2017. Sharpen the focus on that timeline, and here's what it shows: Horne joined Twitter one month before the launch of ASD, just in time to advocate for protecting a group run by the kind of power brokers who held the keys to her professional future.

It is no coincidence that the war against disinformation began at the very moment the Global War on Terror (GWOT) finally appeared to be coming to an end. Over two decades, the GWOT fulfilled President Dwight Eisenhower's warnings about the rise of a military-industrial complex with “unwarranted influence.” It evolved into a self-interested, self-justifying industry that employed thousands of people in and out of government who operated without clear oversight or strategic utility. It might have been possible for the U.S. security establishment to declare victory and move from a permanent war footing to a peacetime posture, but as one former White House national security official explained to me, that was unlikely. “If you work in counterterrorism,” the former official said, “there's no incentive to ever say that you're winning, kicking their ass, and they're a bunch of losers. It's all about hyping a threat.” He described “huge incentives to inflate the threat” that have been internalized in the culture of the U.S. defense establishment and are “of a nature that they don't require one to be particularly craven or intellectually dishonest.”

“This huge machinery was built around the war on terror,” the official said. “A massive infrastructure that includes the intelligence world, all the elements of DoD, including the combatant commands, CIA and FBI and all the other agencies. And then there are all the private contractors and the demand in think tanks. I mean, there are billions and billions of dollars at stake.”

The seamless transition from the war on terror to the war on disinformation was thus, in large measure, simply a matter of professional self-preservation. But it was not enough to sustain the previous system; to survive, it needed to continually raise the threat level.

In the months after the attacks of Sept. 11, 2001, George W. Bush promised to drain the swamps of radicalism in the Middle East. Only by making the region safe for democracy, Bush said, could he ensure that it would stop producing violent jihadists like Osama bin Laden.

Today, to keep America safe, it is no longer enough to invade the Middle East and bring its people democracy. According to the Biden White House and the army of disinformation experts, the threat is now coming from within. A network of right-wing domestic extremists, QAnon fanatics, and white nationalists is supported by a far larger population of some 70 million Trump voters whose political sympathies amount to a fifth column within the United States. But how did these people get radicalized into accepting the bitter and destructive white jihad of Trumpist ideology? Through the internet, of course, where the tech companies, by refusing to “do more” to combat the scourge of hate speech and fake news, allowed toxic disinformation to poison users’ minds.

After 9/11, the threat of terrorism was used to justify measures like the Patriot Act that suspended constitutional rights and placed millions of Americans under a shadow of mass surveillance. Those policies were once controversial but have come to be accepted as the natural prerogatives of state power. As journalist Glenn Greenwald observed, George W. Bush’s “with-us-or-with-the-terrorists’ directive

provoked a fair amount of outrage at the time but is now the prevailing mentality within U.S. liberalism and the broader Democratic Party.”

The war on terror was a dismal failure that ended with the Taliban returning to power in Afghanistan. It also became deeply unpopular with the public. Why, then, would Americans choose to empower the leaders and sages of that war to be the stewards of an even more expansive war against disinformation? It is possible to venture a guess: Americans did not choose them. Americans are no longer presumed to have the right to choose their own leaders or to question decisions made in the name of national security. Anyone who says otherwise can be labeled a domestic extremist.

## **VII. The Rise of “Domestic Extremists”**

A few weeks after Trump supporters rioted in the U.S. Capitol on Jan. 6, 2021, former director of the CIA’s Counterterrorism Center Robert Grenier wrote an article for *The New York Times* advocating for the United States to wage a “comprehensive counterinsurgency program” against its own citizens.

Counterinsurgency, as Grenier would know, is not a limited, surgical operation but a broad effort conducted across an entire society that inevitably involves collateral destruction. Targeting only the most violent extremists who attacked law enforcement officers at the Capitol would not be enough to defeat the insurgency. Victory would require winning the hearts and minds of the natives—in this case, the Christian dead-enders and rural populists radicalized by their grievances into embracing the Bin Laden-like cult of MAGA. Lucky for the government, there is a cadre of experts who are available to deal with this difficult problem: people like

Grenier, who now works as a consultant in the private-sector counterterrorism industry, where he has been employed since leaving the CIA.

Of course there are violent extremists in America, as there have always been. However, if anything, the problem is less severe now than it was in the 1960s and 1970s, when political violence was more common. Exaggerated claims about a new breed of domestic extremism so dangerous it cannot be handled through existing laws, including domestic terrorism statutes, is itself a product of the U.S.-led information war, which has effaced the difference between speech and action.

“Civil wars don’t start with gunshots. They start with words,” Clint Watts proclaimed in 2017 when he testified before Congress. “America’s war with itself has already begun. We all must act now on the social media battlefield to quell information rebellions that can quickly lead to violent confrontations.” Watts is a career veteran of military and government service who seems to share the belief, common among his colleagues, that once the internet entered its populist stage and threatened entrenched hierarchies, it became a grave danger to civilization. But this was a fearful response, informed by beliefs widely, and no doubt sincerely, shared in the Beltway that mistook an equally sincere populist backlash termed “the revolt of the public” by former CIA analyst Martin Gurri for an act of war. The standard Watts and others introduced, which quickly became the elite consensus, treats tweets and memes—the primary weapons of disinformation—as acts of war.

Using the hazy category of disinformation allowed security experts to conflate racist memes with mass shootings in Pittsburgh and Buffalo and with violent protests like the one that took place at the Capitol. It was a rubric for catastrophizing speech and maintaining a permanent state of fear and emergency. And it received the full backing of the Pentagon, the intelligence community, and President Biden, all of whom, notes Glenn Greenwald, have declared that “the gravest menace to American national security” is not Russia, ISIS, China, Iran, or North Korea, but “domestic extremists’ in general—and far-right white supremacist groups in particular.”

The Biden administration has steadily expanded domestic terrorism and counter-extremism programs. In February 2021, DHS officials announced that they had received additional funding to boost department-wide efforts at “preventing domestic terrorism,” including an initiative to counter the spread of disinformation online, which uses an approach seemingly borrowed from the Soviet handbook, called “attitudinal inoculation.”

## VIII. The NGO Borg

In November 2018, Harvard Kennedy School’s Shorenstein Center on Media Politics and Public Policy published a study titled “The Fight Against Disinformation in the U.S.: A Landscape Analysis.” The scope of the paper is comprehensive, but its authors are especially focused on the centrality of philanthropically funded nonprofit organizations and their relationship to the media. The Shorenstein Center is a key node in the complex the paper describes, giving the authors’ observations an insider’s perspective.

“In this landscape analysis, it became apparent that a number of key advocates swooping in to save journalism are not corporations or platforms or the U.S. government, but rather foundations and philanthropists who fear the loss of a free press and the underpinning of a healthy society. ... With none of the authoritative players—the government and platforms who push the content—stepping up to solve the problem quickly enough, the onus has fallen on a collective effort by newsrooms, universities, and foundations to flag what is authentic and what is not.”

To save journalism, to save democracy itself, Americans should count on the foundations and philanthropists—people like eBay founder Pierre Omidyar, Open Society Foundations’ George Soros, and internet entrepreneur and Democratic

Party fundraiser Reid Hoffman. In other words, Americans were being asked to rely on private billionaires who were pumping billions of dollars into civic organizations—through which they would influence the American political process.

There is no reason to question the motivations of the staffers at these NGOs, most of whom were no doubt perfectly sincere in the conviction that their work was restoring the “underpinning of a healthy society.” But certain observations can be made about the nature of that work. First, it placed them in a position below the billionaire philanthropists but above hundreds of millions of Americans whom they would guide and instruct as a new information clerisy by separating truth from falsehood, as wheat from chaff. Second, this mandate, and the enormous funding behind it, opened up thousands of new jobs for information regulators at a moment when traditional journalism was collapsing. Third, the first two points placed the immediate self-interest of the NGO staffers perfectly in line with the imperatives of the American ruling party and security state. In effect, a concept taken from the worlds of espionage and warfare—disinformation—was seeded into academic and nonprofit spaces, where it ballooned into a pseudoscience that was used as an instrument of partisan warfare.

Virtually overnight, the “whole of society” national mobilization to defeat disinformation that Obama initiated led to the creation and credentialing of a whole new class of experts and regulators.

The modern “fact-checking” industry, for instance, which impersonates a well-established scientific field, is in reality a nakedly partisan cadre of compliance officers for the Democratic Party. Its leading organization, the International Fact-Checking Network, was established in 2015 by the Poynter Institute, a central hub in the counter-disinformation complex.

Everywhere one looks now, there is a disinformation expert. They are found at every major media publication, in every branch of government, and in academic departments, crowding each other out on cable news programs, and of course staffing the NGOs. There is enough money coming from the counter-

disinformation mobilization to both fund new organizations and convince established ones like the Anti-Defamation League to parrot the new slogans and get in on the action.

How is it that so many people could suddenly become experts in a field —“disinformation”—that not 1 in 10,000 of them could have defined in 2014? Because expertise in disinformation involves ideological orientation, not technical knowledge. For proof, look no further than the arc traced by Prince Harry and Meghan Markle, who pivoted from being failed podcast hosts to joining the Aspen Institute’s Commission on Information Disorder. Such initiatives flourished in the years after Trump and Brexit.

But it went beyond celebrities. According to former State Department official Mike Benz, “To create a ‘whole of society’ consensus on the censorship of political opinions online that were ‘casting doubt’ ahead of the 2020 election, DHS organized ‘disinformation’ conferences to bring together tech companies, civil society groups, and news media to all build consensus—with DHS prodding (which is meaningful: many partners receive government funds through grants or contracts, or fear government regulatory or retaliatory threats)—on expanding social media censorship policies.”

A DHS memo, first made public by journalist Lee Fang, describes a DHS official’s comment “during an internal strategy discussion, that the agency should use third-party nonprofits as a “clearing house for information to avoid the appearance of government propaganda.”

It is not unusual that a government agency would want to work with private corporations and civil society groups, but in this case the result was to break the independence of organizations that should have been critically investigating the government’s efforts. The institutions that claim to act as watchdogs on government power rented themselves out as vehicles for manufacturing consensus.

Perhaps it is not a coincidence that the fields that have been most aggressive in cheerleading the war against disinformation and calling for greater censorship—

counterterrorism, journalism, epidemiology—share a public record of spectacular failure in recent years. The new information regulators failed to win over vaccine skeptics, convince MAGA diehards that the 2020 election was legitimate, or prevent the public from inquiring into the origins of the COVID-19 pandemic, as they tried desperately to do.

But they succeeded in galvanizing a wildly lucrative whole-of-society effort, providing thousands of new careers and a renewed mandate of heaven to the institutionalists who saw populism as the end of civilization.

## IX. COVID-19

By 2020, the counter-disinformation machine had grown into one of the most powerful forces in American society. Then the COVID-19 pandemic dumped jet fuel into its engine. In addition to fighting foreign threats and deterring domestic extremists, censoring “deadly disinformation” became an urgent need. To take just one example, Google’s censorship, which applied to its subsidiary sites like YouTube, called for “removing information that is problematic” and “anything that would go against World Health Organization recommendations”—a category that at different points in the constantly evolving narrative would have included wearing masks, implementing travel bans, saying that the virus is highly contagious, and suggesting it might have come from a laboratory.

President Biden publicly accused social media companies of “killing people” by not censoring enough vaccine disinformation. Using its new powers and direct channels inside the tech companies, the White House began sending lists of people it wanted banned, such as journalist Alex Berenson. Berenson was kicked off Twitter after tweeting that mRNA vaccines don’t “stop infection. Or transmission.” As it turned out, that was a true statement. The health authorities at the time were

either misinformed or lying about the vaccines' ability to prevent the spread of the virus. In fact, despite claims from the health authorities and political officials, the people in charge of the vaccine knew this all along. In the record of a meeting in December 2020, Food and Drug Administration adviser Dr. Patrick Moore stated, "Pfizer has presented no evidence in its data today that the vaccine has any effect on virus carriage or shedding, which is the fundamental basis for herd immunity." Dystopian in principle, the response to the pandemic was also totalitarian in practice. In the United States, the DHS produced a video in 2021 encouraging "children to report their own family members to Facebook for 'disinformation' if they challenge US government narratives on Covid-19."

"Due to both the pandemic and the disinformation about the election, there are increasing numbers of what extremism experts call 'vulnerable individuals' who could be radicalized," warned Elizabeth Neumann, former assistant secretary of Homeland Security for Counterterrorism and Threat Reduction, on the one-year anniversary of the Capitol riots.

Klaus Schwab, head of the World Economic Forum and *capo di tutti capi* of the global expert class, saw the pandemic as an opportunity to implement a "Great Reset" that could advance the cause of planetary information control: "The containment of the coronavirus pandemic will necessitate a global surveillance network capable of identifying new outbreaks as soon as they arise."

## **X. Hunter's Laptops: The Exception to the Rule**

The laptops are real. The FBI has known this since 2019, when it first took possession of them. When the *New York Post* attempted to report on them, dozens

of the most senior national security officials in the United States lied to the public, claiming the laptops were likely part of a Russian “disinformation” plot. Twitter, Facebook, and Google, operating as fully integrated branches of the state security infrastructure, carried out the government’s censorship orders based on that lie. The press swallowed the lie and cheered on the censorship.

The story of the laptops has been framed as many things, but the most fundamental truth about it is that it was the successful culmination of the yearslong effort to create a shadow regulatory bureaucracy built specifically to prevent a repeat of Trump’s 2016 victory.

It may be impossible to know exactly what effect the ban on reporting about Hunter Biden’s laptops had on the 2020 vote, but the story was clearly seen as threatening enough to warrant an openly authoritarian attack on the independence of the press. The damage to the country’s underlying social fabric, in which paranoia and conspiracy have been normalized, is incalculable. As recently as February, Rep. Alexandria Ocasio-Cortez referred to the scandal as the “half-fake laptop story” and as “an embarrassment,” months after even the Bidens had been forced to acknowledge that the story is authentic.

While the laptop is the best-known case of the ruling party’s intervention in the Trump-Biden race, its brazenness was an exception. The vast majority of the interference in the election was invisible to the public and took place through censorship mechanisms carried out under the auspices of “election integrity.” The legal framework for this had been put in place shortly after Trump took office, when the outgoing DHS chief Jeh Johnson passed an 11th-hour rule—over the vehement objections of local stakeholders—declaring election systems to be critical national infrastructure, thereby placing them under the supervision of the agency. Many observers had expected that the act would be repealed by Johnson’s successor, Trump-appointed John Kelly, but curiously it was left in place.

In 2018, Congress created a new agency inside of the DHS called the Cybersecurity and Infrastructure Security Agency (CISA) that was tasked with defending America’s infrastructure—now including its election systems—from foreign

attacks. In 2019, the DHS added another agency, the Foreign Influence and Interference Branch, which was focused on countering foreign disinformation. As if by design, the two roles merged. Russian hacking and other malign foreign-information attacks were said to threaten U.S. elections. But, of course, none of the officials in charge of these departments could say with certainty whether a particular claim was foreign disinformation, simply wrong, or merely inconvenient. Nina Jankowicz, the pick to lead the DHS's short-lived Disinformation Governance Board, lamented the problem in her book *How to Lose the Information War: Russia, Fake News and the Future of Conflict*. "What makes this information war so difficult to win," she wrote, "is not just the online tools that amplify and target its messages or the adversary that is sending them; it's the fact that those messages are often unwittingly delivered not by trolls or bots, but by authentic local voices."

The latitude inherent in the concept of disinformation enabled the claim that preventing electoral sabotage required censoring Americans' political views, lest an idea be shared in public that was originally planted by foreign agents.

In January 2021, CISA "transitioned its Countering Foreign Influence Task Force to promote more flexibility to focus on general MDM [ed. note: an acronym for *misinformation, disinformation, and malinformation*]," according to an August 2022 report from the DHS's Office of Inspector General. After the pretense of fighting a foreign threat fell away, what was left was the core mission to enforce a narrative monopoly over truth.

The new domestic-focused task force was staffed by 15 employees dedicated to finding "all types of disinformation"—but specifically that which related to "elections and critical infrastructure"—and being "responsive to current events," a euphemism for promoting the official line of divisive issues, as was the case with the "COVID-19 Disinformation Toolkit" released to "raise awareness related to the pandemic."

Kept a secret from the public, the switch was "plotted on DHS's own livestreams and internal documents," according to Mike Benz. "DHS insiders' collective

justification, without uttering a peep about the switch's revolutionary implications, was that 'domestic disinformation' was now a greater 'cyber threat to elections' than falsehoods flowing from foreign interference."

Just like that, without any public announcements or black helicopters flying in formation to herald the change, America had its own ministry of truth.

Together they operated an industrial-scale censorship machine in which the government and NGOs sent tickets to the tech companies that flagged objectionable content they wanted scrubbed. That structure allowed the DHS to outsource its work to the Election Integrity Project (EIP), a consortium of four groups: the Stanford Internet Observatory; private anti-disinformation company Graphika (which had formerly been employed by the Defense Department against groups like ISIS in the war on terror); Washington University's Center for an Informed Public; and the Atlantic Council's Digital Forensics Research Lab. Founded in 2020 in partnership with the DHS, the EIP served as the government's "deputized domestic disinformation flagger," according to [congressional testimony](#) from journalist Michael Shellenberger, who notes that the EIP claims it classified more than 20 million unique "misinformation incidents" between Aug. 15 and Dec. 12, 2020. As EIP head Alex Stamos explained, this was a work-around for the problem that the government "lacked both kinda the funding and the legal authorizations."

Looking at the censorship figures that the DHS's own partners reported for the 2020 election cycle in their internal audits, the [Foundation for Freedom Online](#) summarized the scope of the censorship campaign in seven bullet points:

- [22 million](#) tweets labeled "misinformation" on Twitter;
- [859 million](#) tweets collected in databases for "misinformation" analysis;
- [120](#) analysts monitoring social media "misinformation" in up to [20-hour](#) shifts;
- [15](#) tech platforms monitored for "misinformation," often in real-time;

- <1 hour average response time between government partners and tech platforms;
- Dozens of “misinformation narratives” targeted for platform-wide throttling; and
- Hundreds of millions of individual Facebook posts, YouTube videos, TikToks, and tweets impacted due to “misinformation” Terms of Service policy changes, an effort DHS partners openly plotted and bragged that tech companies would never have done without DHS partner insistence and “huge regulatory pressure” from government.

## XI. The New One-Party State

In February 2021, a long article in *Time* magazine by journalist Molly Ball celebrated the “Shadow Campaign That Saved the 2020 Election.” Biden’s victory, wrote Ball, was the result of a “conspiracy unfolding behind the scenes” that drew together “a vast, cross-partisan campaign to protect the election” in an “extraordinary shadow effort.” Among the many accomplishments of the heroic conspirators, Ball notes, they “successfully pressured social media companies to take a harder line against disinformation and used data-driven strategies to fight viral smears.” It is an incredible article, like an entry from the crime blotter that somehow got slipped into the society pages, a paean to the saviors of democracy that describes in detail how they dismembered it.

Not so long ago, talk of a “deep state” was enough to mark a person as a dangerous conspiracy theorist to be summarily flagged for monitoring and censorship. But language and attitudes evolve, and today the term has been cheekily reappropriated by supporters of the deep state. For instance, a new book, *American Resistance*, by

neoliberal national security analyst David Rothkopf, is subtitled *The Inside Story of How the Deep State Saved the Nation*.

The deep state refers to the power wielded by unelected government functionaries and their paragovernmental adjuncts who have administrative power to override the official, legal procedures of a government. But a ruling class describes a social group whose members are bound together by something deeper than institutional position: their shared values and instincts. While the term is often used loosely and sometimes as a pejorative rather than a descriptive label, in fact the American ruling class can be simply and straightforwardly defined.

Two criteria define membership in the ruling class. First, as Michael Lind has written, it is made up of people who belong to a “homogeneous national oligarchy, with the same accent, manners, values, and educational backgrounds from Boston to Austin and San Francisco to New York and Atlanta.” America has always had regional elites; what is unique about the present is the consolidation of a single, national ruling class.

Second, to be a member of the ruling class is to believe that only other members of your class can be allowed to lead the country. That is to say, members of the ruling class refuse to submit to the authority of anyone outside the group, whom they disqualify from eligibility by casting them as in some way illegitimate.

Faced with an external threat in the form of Trumpism, the natural cohesion and self-organizing dynamics of the social class were fortified by new top-down structures of coordination that were the goal and the result of Obama’s national mobilization. In the run-up to the 2020 election, according to reporting by Lee Fang and Ken Klippenstein for *The Intercept*, “tech companies including Twitter, Facebook, Reddit, Discord, Wikipedia, Microsoft, LinkedIn, and Verizon Media met on a monthly basis with the FBI, CISA, and other government representatives ... to discuss how firms would handle misinformation during the election.”

Historian Angelo Codevilla, who popularized the concept of an American “ruling class” in a 2010 essay and then became its primary chronicler, saw the new, national

aristocracy as an outgrowth of the opaque power acquired by the U.S. security agencies. “The bipartisan ruling class that grew in the Cold War, who imagined themselves and who managed to be regarded as entitled by expertise to conduct America’s business of war and peace, protected its status against a public from which it continued to diverge by translating the commonsense business of war and peace into a private, pseudo-technical language impenetrable to the uninitiated,” he wrote in his 2014 book, *To Make and Keep Peace Among Ourselves and with All Nations*.

What do the members of the ruling class believe? They believe, I argue, “in informational and management solutions to existential problems” and in their “own providential destiny and that of people like them to rule, regardless of their failures.” As a class, their highest principle is that they alone can wield power. If any other group were to rule, all progress and hope would be lost, and the dark forces of fascism and barbarism would at once sweep back over the earth. While technically an opposition party is still permitted to exist in the United States, the last time it attempted to govern nationally, it was subjected to a yearslong coup. In effect, any challenge to the authority of the ruling party, which represents the interests of the ruling class, is depicted as an existential threat to civilization.

An admirably direct articulation of this outlook was provided recently by famous atheist Sam Harris. Throughout the 2010s, Harris’ higher-level rationalism made him a star on YouTube, where thousands of videos showcased him “owning” and “pwning” religious opponents in debates. Then Trump arrived. Harris, like so many others who saw in the former president a threat to all that was good in the world, abandoned his principled commitment to the truth and became a defender of propaganda.

In a podcast appearance last year, Harris acknowledged the politically motivated censorship of reporting related to Hunter Biden’s laptops and admitted “a left-wing conspiracy to deny the presidency to Donald Trump.” But, echoing Ball, he declared this a good thing.

“I don’t care what’s in the Hunter Biden laptop. ... Hunter Biden could have had corpses of children in his basement, and I would not have cared,” Harris told his interviewers. He could overlook the murdered children because an even greater danger lurked in the possibility of Trump’s reelection, which Harris compared to “an asteroid hurtling toward Earth.”

With an asteroid hurtling toward Earth, even the most principled rationalists might end up asking for safety over truth. But an asteroid has been falling toward Earth every week for years now. The pattern in these cases is that the ruling class justifies taking liberties with the law to save the planet but ends up violating the Constitution to hide the truth and protect itself.

## **XII. The End of Censorship**

The public’s glimpses into the early stages of the transformation of America from democracy to digital leviathan are the result of lawsuits and FOIAs—information that had to be pried from the security state—and one lucky fluke. If Elon Musk had not decided to purchase Twitter, many of the crucial details in the history of American politics in the Trump era would have remained secret, possibly forever.

But the system reflected in those disclosures may well be on its way out. It is already possible to see how the kind of mass censorship practiced by the EIP, which requires considerable human labor and leaves behind plenty of evidence, could be replaced by artificial intelligence programs that use the information about targets accumulated in behavioral surveillance dossiers to manage their perceptions. The ultimate goal would be to recalibrate people’s experiences online through subtle manipulations of what they see in their search results and on their feed. The aim of such a scenario might be to prevent censor-worthy material from being produced in the first place.

In fact, that sounds rather similar to what [Google is already doing in Germany](#), where the company recently unveiled a new campaign to expand its “prebunking” initiative “that aims to make people more resilient to the corrosive effects of online misinformation,” according to the Associated Press. The announcement closely followed Microsoft founder Bill Gates’ appearance on a German podcast, during which he called for using [artificial intelligence to combat](#) “conspiracy theories” and “political polarization.” Meta has its own prebunking program. In a statement to the website [Just The News](#), Mike Benz called prebunking “a form of narrative censorship integrated into social media algorithms to stop citizens from forming specific social and political belief systems” and compared it to the “pre-crime” featured in dystopian science-fiction movie *Minority Report*.

Meanwhile, the military is developing weaponized AI technology to dominate the information space. According to [USASpending.gov](#), an official government website, the two largest contracts related to disinformation came from the Department of Defense to fund technologies for automatically detecting and defending against large-scale disinformation attacks. The first, for \$11.9 million, was awarded in June 2020 to PAR Government Systems Corporation, a defense contractor in upstate New York. The second, issued in July 2020 for \$10.9 million, went to a company called SRI International.

SRI International was originally connected to Stanford University before splitting off in the 1970s, a relevant detail considering that the Stanford Internet Observatory, an institution still directly connected to the school, led 2020’s EIP, which might well have been the largest mass censorship event in world history—a capstone of sorts to the record of pre-AI censorship.

Then there is the work going on at the National Science Foundation, a government agency that funds research in universities and private institutions. The NSF has its own program called the Convergence Accelerator Track F, which is helping to incubate a dozen automated disinformation-detection technologies explicitly designed to monitor issues like “vaccine hesitancy and electoral skepticism.”

“One of the most disturbing aspects” of the program, according to Benz, “is how similar they are to military-grade social media network censorship and monitoring tools developed by the Pentagon for the counterinsurgency and counterterrorism contexts abroad.”

In March, the NSF’s chief information officer, Dorothy Aronson, announced that the agency was “building a set of use cases” to explore how it could employ ChatGPT, the AI language model capable of a reasonable simulation of human speech, to further automate the production and dissemination of state propaganda.

The first great battles of the information war are over. They were waged by a class of journalists, retired generals, spies, Democratic Party bosses, party apparatchiks, and counterterrorism experts against the remnant of the American people who refused to submit to their authority.

Future battles fought through AI technologies will be harder to see.

### **XIII. After Democracy**

Less than three weeks before the 2020 presidential election, *The New York Times* published an important article titled “The First Amendment in the age of disinformation.” The essay’s author, *Times* staff writer and Yale Law School graduate Emily Bazelon, argued that the United States was “in the midst of an information crisis caused by the spread of viral disinformation” that she compares to the “catastrophic” health effects of the novel coronavirus. She quotes from a book by Yale philosopher Jason Stanley and linguist David Beaver: “Free speech threatens democracy as much as it also provides for its flourishing.”

So the problem of disinformation is also a problem of democracy itself—specifically, that there’s too much of it. To save liberal democracy, the experts prescribed two critical steps: America must become less free and less democratic. This necessary evolution will mean shutting out the voices of certain rabble-rousers in the online crowd who have forfeited the privilege of speaking freely. It will require following the wisdom of disinformation experts and outgrowing our parochial attachment to the Bill of Rights. This view may be jarring to people who are still attached to the American heritage of liberty and self-government, but it has become the official policy of the country’s ruling party and much of the American intelligentsia.

Former Clinton Labor Secretary Robert Reich responded to the news that Elon Musk was purchasing Twitter by declaring that preserving free speech online was “Musk’s dream. And Trump’s. And Putin’s. And the dream of every dictator, strongman, demagogue, and modern-day robber baron on Earth. For the rest of us, it would be a brave new nightmare.” According to Reich, censorship is “necessary to protect American democracy.”

To a ruling class that had already grown tired of democracy’s demand that freedom be granted to its subjects, disinformation provided a regulatory framework to replace the U.S. Constitution. By aiming at the impossible, the elimination of all error and deviation from party orthodoxy, the ruling class ensures that it will always be able to point to a looming threat from extremists—a threat that justifies its own iron grip on power.

A siren song calls on those of us alive at the dawn of the digital age to submit to the authority of machines that promise to optimize our lives and make us safer. Faced with the apocalyptic threat of the “infodemic,” we are led to believe that only superintelligent algorithms can protect us from the crushingly inhuman scale of the digital information assault. The old human arts of conversation, disagreement, and irony, on which democracy and much else depend, are subjected to a withering machinery of military-grade surveillance—surveillance that nothing can withstand and that aims to make us fearful of our capacity for reason.

*If you work in the “disinformation” or “misinformation” fields for the government or in the private sector, and are interested in discussing your experiences, you can contact me securely at [jacobsiegel@protonmail.com](mailto:jacobsiegel@protonmail.com) or on Twitter @jacob\_siegel. Source confidentiality is guaranteed.*

Jacob Siegel is senior editor of News and The Scroll, Tablet's daily afternoon news digest, which you can subscribe to [here](#).

---

#CENSORSHIP #FIRST AMENDMENT #FAKE NEWS #SOCIAL MEDIA  
#U.S. INTELLIGENCE



May 12, 2023

**By Electronic Mail**

Hon. Gary Peters, Chair  
Homeland Security and Governmental  
Affairs Committee  
United States Senate  
340 Dirksen Senate Office Building  
Washington, DC 20510

Hon. Rand Paul, Ranking Member  
Homeland Security and Governmental  
Affairs Committee  
United States Senate  
442 Hart Senate Office Building  
Washington, DC 20510

**Re: Available Apolitical, Non-Lobbying Expertise on Artificial Intelligence**

Dear Chairman Peters and Ranking Member Paul:

ACM, the Association for Computing Machinery, is the world's largest and longest established association of computing professionals, representing approximately 50,000 individuals in the United States and more than 100,000 worldwide. **ACM is a *non-profit, non-lobbying and non-political* organization whose U.S. Technology Policy Committee ("USTPC") is charged with providing policy and law makers throughout government with timely, substantive, and apolitical input on computing technology**, and the legal and social issues to which it gives rise. We recently marked our 30th year doing so.

As the Committee prepares to address pressing issues raised by the proliferation of generative artificial intelligence technologies, USTPC would be pleased to substantively assist you, your members, and their staffs in whatever manner would be most useful and efficient. Whether by responding in writing to specific inquiries, technically assessing draft legislation, or providing briefings by one or more of our expert members, USTPC stands ready to support you and the Committee's members with science-based information to meet Congress' challenge to understand and provide effective guardrails for guiding the evolution of these exciting but challenging technologies.

To those ends please contact Adam Eisgrau, ACM's Washington-based Director of Global Policy, at any time to tap our thousands of members' technical expertise. In the interim, please find attached our most recent recommendations and proposed principles bearing on artificial intelligence: a January 2023 ACM TechBrief on [Safer Algorithmic Systems](#); and our late 2022 [Joint Statement on Principles for Responsible Algorithmic Systems](#).

ACM U.S. Technology Policy Committee  
1701 Pennsylvania Ave NW, Suite 200  
Washington, DC 20006

+1 202.580.6555  
acmpo@acm.org  
[www.acm.org/public-policy/ustpc](http://www.acm.org/public-policy/ustpc)

Thank you for your time and consideration. We look forward to contributing to the Committee's and Congress' understanding of these vital issues.

Sincerely,


A handwritten signature in blue ink, appearing to read "Jeremy J. Epstein". The signature is fluid and cursive, with a large initial "J" and "E".

Jeremy J. Epstein, Chair

Attachments

Association for Computing Machinery

**acm** Technology Policy Council **TechBriefs**



**SAFER ALGORITHMIC SYSTEMS**  
First in a series on systems and trust

**Ben Sheiderman, Lead Author**

**PROBLEM**

*The ubiquity of algorithmic systems creates serious risks that are not being adequately addressed.*

**POLICY IMPLICATIONS**

- Enabling safer algorithmic systems must be a high research and policy priority of governments and all stakeholders.
- Organizational safety cultures must be broadly embraced and routinely woven into algorithmic system development and operation.
- Safer algorithmic systems will require multiple forms of sustained internal and independent oversight.

**SAFER ALGORITHMIC SYSTEMS: BY THE NUMBERS**

<b>2,000</b>	Minimum number of reports in the Artificial Intelligence Incident Database. <sup>1</sup>
<b>346</b>	Number of deaths in Boeing 737 MAX crashes in 2018 and 2019 caused by algorithmic system failure. <sup>2</sup>
<b>392</b>	Number of crashes involving vehicles with advanced or automated driving systems in the U.S. in 2021. <sup>3</sup>
<b>44</b>	Percentage of U.S. adults who think widespread use of driverless cars would be a bad idea. <sup>4</sup>
<b>80</b>	Projected value in billions of U.S. dollars of the global autonomous driving market by 2030. <sup>5</sup>
<b>83</b>	Percentage of U.S. adults who favor tougher standards for testing brain computer implants. <sup>6</sup>
<b>26,000</b>	Number of Dutch parents mistakenly identified by algorithms and heavily fined for tax fraud in 2020. <sup>7</sup>
<b>11,000</b>	Number of those parents singled out for government scrutiny based on their ethnicity. <sup>7</sup>

ILLUSTRATION: SHUTTERSTOCK

ACM's Technology Policy Committees provide cutting-edge, apolitical, non-lobbying scientific information about all aspects of computing to policy makers in the United States and Europe. To tap the deep expertise of ACM's 100,000 members worldwide, contact ACM's Global Policy Office at [acmpo@acm.org](mailto:acmpo@acm.org) or +1 202.580.6555.



### Overview

Embedded in immense numbers of products and processes, computing is ubiquitous in modern life. Algorithms are the underlying operating rules that control computers. While incredibly useful and generally benign, when deployed in complex systems algorithms can cause a variety of profound harms to individuals and to society, threatening opportunity, liberty, and even life itself.<sup>8</sup> For example, algorithmic systems underlying e-commerce, search engines, and online leisure activities can exhibit bias. Risks grow as flawed algorithmic decisions become more consequential (e.g., in hiring, judicial sentencing, or lending) and increase to unacceptable levels in life-critical applications (e.g., in medicine, transportation, or the military).

### **Reducing risks from algorithmic systems will require commitment by all stakeholders.**

Reducing risks from algorithmic systems will require commitment by all stakeholders<sup>9</sup> to more safety-oriented approaches sensitive to organizational and cultural considerations as well as technological ones. Such commitment must inform the development of algorithms and their operating environments from the outset, and of the larger software-driven systems into which algorithms are integrated. Drawing on human-centered social systems scholarship,<sup>10</sup> safety research and policy making must foster adoption of a safety culture within relevant organizations as well as internal and independent oversight mechanisms.

### **All stakeholders must prioritize research-driven algorithmic safety policy making**

Governments around the world, particularly in economically developed nations, have responded to the perils and promise of algorithmic systems by studying and adopting practices, policies, and laws to mitigate their risks in all sectors of society. Although perfectly safe systems are not possible, safer systems are.

In the United States, the White House Office of Science and Technology Policy released in late 2022 the Blueprint for an AI Bill of Rights.<sup>11</sup> The first of its five policy principles identified “safe and effective automated systems” as a national priority.<sup>12</sup> A parallel effort by the National Institute of Standards and Technology, as required by Congress, is underway to publish a voluntary AI Risk Management Framework<sup>13</sup> in early 2023.<sup>14</sup>

Similarly, the European Union’s Artificial Intelligence Act,<sup>15</sup> building on the work of the European Commission’s High-Level Expert Group on AI,<sup>16</sup> aims to make algorithmic

systems safer. The act would impose transparency and other legal requirements on an algorithmic system’s design according to the level of risk associated with its use. The U.S. and EU also have acted jointly, issuing an “AI Joint Roadmap” through the transatlantic Trade and Technology Council “to advance shared terminologies and taxonomies, but also to inform...approaches to AI risk management and trustworthy AI on both sides of the Atlantic.”<sup>17</sup> The U.K. is addressing these matters too, though it favors a sector-by-sector approach to risk management and ultimate regulation.<sup>18</sup> Outside Europe, the government of India, for example, has published a series of discussion papers, most recently one titled “Responsible AI.”<sup>19</sup>

International organizations and public interest organizations, as well as commercial enterprises and their associations, also have an important role to play in supporting research on safer algorithmic systems and establishing safety research and policy making priorities.<sup>20</sup> Several emerging legal frameworks for the governance of AI that establish impact assessment requirements also will promote safer algorithmic systems.<sup>21</sup> These global efforts must respond to the highly dynamic technological and sociopolitical environment of evolving algorithmic systems safety design, implementation, and deployment.

### **Safer algorithmic systems require an organizational safety culture**

The appropriate foundation for safer algorithmic systems is evidence-based technical and human-centered research from case studies, hazard analysis, ethnographic observations, data collection, and controlled empirical studies of alternative approaches.<sup>22</sup>

### **Perfectly safe algorithmic systems are not possible; safer systems are.**

Lessons from medical and aviation safety culture may be useful guides for improving algorithmic systems safety. In both fields, senior management typically provides clear vision statements and substantial financial and human resources. Both industries carefully track performance, recording failures, and much more common near misses. These provide valuable early lessons both about possible failures and the strategies that employees used to prevent near misses from becoming failures. Importantly, both industries conduct simulations of mass casualty or other disasters.<sup>1</sup>

A safety culture that embraces human factors engineering must be woven into algorithmic system design.<sup>23</sup> Safer algorithmic systems will be advanced by interpretable designs that give stakeholders a clear

understanding of what the algorithms will do.<sup>24</sup> Software engineering requirements should include improved algorithm design and application of user interface design guidelines to build comprehensible, predictable, and controllable systems.

### ***Human factors engineering must be woven into algorithmic system design.***

Organizations that prioritize safety often conduct adversarial “red team” tests in which expert users are asked to try to break the system. They also offer “bug bounties” to users who report omissions and errors capable of leading to major failures. These methods have proved to be effective in cybersecurity, and they should likewise be useful in making safer algorithmic systems.

Embrace of an algorithmic systems safety culture could provide corporate competitive advantage (and may become a legal requirement). This would lead to shifts in hiring, training, performance indicators, and reporting requirements, thus giving safety the priority it deserves in consequential and life-critical applications.<sup>25</sup>

### **Safer algorithmic systems require continuous attention and robust oversight**

Safer algorithmic systems also require continuous human-centered attention during system use, especially as contexts, users, data characteristics, and expectations change. Efforts undertaken to date in several spheres, while imperfect, can be instructive.

Performance logs, like the flight data recorders in civil aviation, provide diagnostic audit trails so that, when failures occur, investigators can conduct a retrospective analysis to determine what went wrong to ensure similar failures will not recur. Trusted third parties serve as clearinghouses for industry reporting and analysis of near misses and other issues,<sup>26</sup> preserving confidentiality

of, and thereby encouraging, unvarnished submissions while sharing aggregate results.

Incident reporting websites are another important potential source of guidance for designers and software engineers as they enable stakeholders to report on problems they faced and near misses that raised concerns.<sup>27</sup> Such sites have proved their value in civil aviation<sup>28</sup> as well as with medications, medical devices, cosmetics, food, and other products.<sup>29</sup> The AI Incident Database,<sup>30</sup> run by the Responsible AI Collaborative, has received over 2,000 reports since it launched in late 2020. The data provided are a valuable resource for understanding problems and suggesting improvements.

### ***Safer algorithmic systems require continuous human-centered attention.***

Robust oversight by both system developers and independent experts is also critical to enabling safer algorithmic systems. Government policy makers and corporate managers are moving rapidly to increase the diversity of stakeholder participants in design processes. They are also creating internal review boards to assess planned implementations, monitor existing applications, and study failures and near misses.

Internal review processes can ensure that these practices align with industry standard practices, which are emerging from professional societies such as ACM,<sup>31</sup> IEEE,<sup>32</sup> and the Association for Advancing Automation (formerly the Robotics Industries Association).<sup>33</sup> Additional work has come from other nongovernmental organizations, including UL Research Institutes<sup>34</sup> and the International Organization for Standardization.<sup>35</sup> In addition, there is growing awareness of the benefits of independent oversight, or what the U.S. AI Bill of Rights calls independent evaluation and reporting, with the results made public whenever possible.

## KEY CONCLUSIONS

- To promote safer algorithmic systems, research is needed on both human-centered and technical software development methods, improved testing, audit trails, and monitoring mechanisms, as well as training and governance.
- Building organizational safety cultures requires management leadership, focus in hiring and training, adoption of safety-related practices, and continuous attention.
- Internal and independent human-centered oversight mechanisms, both within government and organizations, are necessary to promote safer algorithmic systems.

## NOTES AND SOURCES

- See generally <https://incidentdatabase.ai/> and, e.g., "Incident 439: Detroit Police Wrongfully Arrested Black Man Due to Faulty Facial Recognition" (accessed on January 10, 2023), <https://incidentdatabase.ai/cite/439>.
- "Boeing 737 Max Crashes," Plane Crash Rates by Model, AirSafe, last modified March 22, 2022, <http://www.airsafe.com/events/models/b737.htm>.
- National Highway Traffic Safety Administration, *Summary Report: Standing General Order on Crash Reporting for Automated Driving Systems*, June 2022, <https://www.nhtsa.gov/sites/nhtsa.gov/files/2022-06/AIADS-SG-Report-June-2022.pdf>.
- Lee Rainie, Cary Funk, Monica Anderson, and Alec Tyson, "Americans Cautious about the Deployment of Driverless Cars," Pew Research Center, March 17, 2022, <https://www.pewresearch.org/internet/2022/03/17/americans-cautious-about-the-deployment-of-driverless-cars>.
- Johannes Doehmann, Eric Ebel, Kersten Heinke, Ruth Heras, Martin Kellner, and Fabian Seiner, "Autonomous Driving's Future: Convenient and Connected," McKinsey Center for Future Mobility (January 6, 2023) <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/autonomous-drivings-future-convenient-and-connected>.
- Rainie et al., "Public Cautious About Enhancing Cognitive Function Using Computer Chip Implants in the Brain," Pew Research Center, March 17, 2022, <https://www.pewresearch.org/internet/2022/03/17/public-cautious-about-enhancing-cognitive-function-using-computer-chip-implants-in-the-brain>.
- European Union Agency for Fundamental Rights, *Bias in Algorithms, Artificial Intelligence and Discrimination*, 2022, [https://fra.europa.eu/sites/default/files/fra\\_uploads/fra-2022-bias-in-algorithms\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf). See also John Henley, "Dutch Government Resigns over Child Benefits Scandal," *The Guardian* (January 15, 2021), <https://www.theguardian.com/world/2021/jan/15/dutch-government-resigns-over-child-benefits-scandal>.
- Many examples have been documented. See, e.g., Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Broadway Books, 2016), and ACM's long-running RISKS Forum, <https://www.acm.org/about/acm/risks-forum>.
- Stakeholders include but may not be limited to systems and software developers, public and private sector governance authorities, system users of all kinds, and all populations affected by a system's use.
- Such work draws on research across multiple disciplines, including sociology, psychology, and organizational behavior.
- White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights: Making Automated Systems for the American People*, 2022, <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>.
- This principle echoes the familiar expectation that pharmaceutical producers should ensure safe and effective medications and treatments. However, researchers, designers, implementers, and managers in both domains face substantial challenges in advancing safety because there is no single path to success.
- National Institute of Standards and Technology, "AI Risk Management Framework," <https://www.nist.gov/itl/ai-risk-management-framework>.
- NIST, *AI Risk Management Framework Playbook*, <https://pages.nist.gov/AIRMF/>.
- European Commission, "Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts," 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC3206>.
- European Commission, "Building Trust in Human-Centric Artificial Intelligence," August 4, 2019, [https://ec.europa.eu/jrc/communities/sites/default/files/ec\\_ai\\_ethics\\_communication\\_e\\_april\\_2019.pdf](https://ec.europa.eu/jrc/communities/sites/default/files/ec_ai_ethics_communication_e_april_2019.pdf).
- European Commission, "ETC Joint Roadmap for Trustworthy AI and Risk Management," December 1, 2022, <https://digital-strategy.ec.europa.eu/en/library/etc-joint-roadmap-trustworthy-ai-and-risk-management>.
- Gov.UK, "UK sets out proposals for new AI rulebook to unleash innovation and boost public trust in the technology," press release, July 18, 2022, <https://www.gov.uk/government/news/uk-sets-out-proposals-for-new-ai-rulebook-to-unleash-innovation-and-boost-public-trust-in-the-technology>.
- See NTTAYO, "Responsible AI for All—Adopting the Framework: A Use Case Approach on Facial Recognition Technology," November 2022, [https://www.ntt.ai/gov/sites/default/files/2022-11/AI\\_for\\_All\\_2022\\_02112022\\_0.pdf](https://www.ntt.ai/gov/sites/default/files/2022-11/AI_for_All_2022_02112022_0.pdf). Several Chinese institutions also have promulgated relevant principles and guidance. See Matt Sheehan, "China's New AI Governance Initiatives Shouldn't Be Ignored," January 4, 2022, <https://carnegeieduendowment.org/2022/01/04/china-s-new-ai-governance-initiatives-shouldnt-be-ignored/>.
- See, e.g., respectively: Organization for Economic Cooperation and Development, "OECD AI Principles Overview," May 2019, <https://oecd.ai/en/ai-principles>; Alexandra Circiari, "People, Risk and the Unique Requirements of AI," Ada Lovelace Institute, March 31, 2022, <https://www.adalovelaceinstitute.org/policy-briefings/ei-ai-act>; European Digital Rights, *Recommendations for a Fundamental Rights Based Artificial Intelligence Regulation*, June 4, 2020, [https://edri.org/wp-content/uploads/2020/06/AI\\_EDRIRRecommendations.pdf](https://edri.org/wp-content/uploads/2020/06/AI_EDRIRRecommendations.pdf); and Information Technology & Innovation Foundation, "Where Should US Policy Be Headed Next?," webinar, February 23, 2023, <https://itif.org/events/2023/01/17/where-should-us-ai-policy-be-headed-next/>.
- See Marc Rotenberg (2022), *Fair AI Practices*, Communications of the ACM Blog. Specifically, the UNESCO Recommendation on the Ethics of AI proposes an Ethical Impact Assessment (<https://www.unesco.org/en/articles/unesco-member-states-adopt-first-ever-global-agreement-ethics-artificial-intelligence>); the EU AI Act will require a detailed Conformity Assessment for high risk AI systems (<https://www.citp.org/resources/ei-ai-act/>); and the European Law Institute has proposed Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration ([https://www.europeanlawinstitute.eu/filedadmin/user\\_upload/p\\_ei/Publications/ELI\\_Model\\_Rules\\_on\\_Impact\\_Assessment\\_of\\_ADMSS\\_Used\\_by\\_Public\\_Administration.pdf](https://www.europeanlawinstitute.eu/filedadmin/user_upload/p_ei/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSS_Used_by_Public_Administration.pdf)). All of these policy instruments anticipate that assessments will be conducted prior to the deployment of an AI system and throughout the system life cycle.
- See Laura Wedinger et al., "Ethical and Social Risks of Harm from Language Models," Cornell University, arXiv preprint arXiv:2112.04359 (<https://arxiv.org/abs/2112.04359>). The Blueprint for the AI Bill of Rights reports that the National Science Foundation, for example, already funds extensive research to foster the development of automated systems that are safe, secure, and effective.
- As with all systems that have diverse stakeholders, accessibility must be designed in from the start and tested with users with disabilities.
- Cynthia Rudin, "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead," *Nature Machine Intelligence* 1 (2019): 206–15, <https://doi.org/10.1038/s42256-019-0040-8>.
- Nancy G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety* (Cambridge, MA: The MIT Press, 2016), <https://doi.org/10.7551/mitpress/8179.001.0001>.
- The Aviation Safety Information Analysis and Sharing program, operated on behalf of the Federal Aviation Administration (FAA) and the aviation community by the MTRF Corporation, for example, brings together a variety of data, including proprietary, in de-identified form, <https://portal.asias.aero/overview>.
- See Ben Shneiderman, *Human-Centered AI* (Oxford: Oxford University Press, 2022), <https://global.oup.com/academic/product/human-centered-ai-9780192845290>.
- NASA and the FAA run an Aviation Safety Reporting System, a public website that enables pilots, air traffic controllers, and others to report problems they have witnessed or even contributed to, <https://asr.arc.nasa.gov>.
- The U.S. Food and Drug Administration maintains an Adverse Event Reporting System that enables web-based public reporting by health care professionals, consumers, and manufacturers of possible safety issues, <https://open.fda.gov/data/fiers/>.
- See AI Incident Database, <https://incidentdatabase.ai>.
- Association for Computing Machinery Technology Policy Council, "Statement on Principles for Responsible Algorithmic Systems," October 26, 2022, <https://www.acm.org/binaries/content/assets/public-policy/final-joint-statement-update.pdf>.
- IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems," March 31, 2019: 1–294, <https://www.ieee.org/servlet/opac?punumber=9398611>.
- See *American National Standard for Industrial Robots and Robot Systems: Safety Requirements, ANSI/RIA R15.06-2012*, among other standards collected by the Association for Advancing Automation (AA), <https://www.automate.org/a3-content/robot-safety-standard-document>.
- See "Standards and Engagement" section, IJL Research Institutes, <https://ijl.org/initiatives>.
- See, e.g., *Quality management systems—Requirements*, ISO 9001:2015, International Organization for Standardization, <https://www.iso.org/popular-standards.html>.

## ADDITIONAL INFORMATION

With 100,000 members in 190 countries, the nonprofit Association for Computing Machinery is the world's largest and longest-established organization of professionals involved in all aspects of computing. Under the auspices of the global ACM Technology Policy Council, policy committees in the U.S. and Europe provide cutting-edge, apolitical, non-lobbying information about computing and its social impacts to policy makers at all levels of government in many forms, including briefings, testimony, consultation and rulemaking comments, reports, and analyses.

To tap the deep expertise of ACM's global membership, please contact ACM's Global Policy Office at [acmpo@acm.org](mailto:acmpo@acm.org) or +1 202.580.6555.

To receive ACM TechBriefs quarterly, in the body of a one-line email send [subscribe-acm-tpc-tech-briefs](mailto:subscribe-acm-tpc-tech-briefs) followed by your first and last names to [listserv@listserv.acm.org](mailto:listserv@listserv.acm.org).

## AUTHORSHIP &amp; ACKNOWLEDGMENTS

Ben Shneiderman is distinguished university professor emeritus in the Department of Computer Science at the University of Maryland and founding director of its Human-Computer Interaction Laboratory. He is a fellow of the American Association for the Advancement of Science, ACM, the Institute of Electrical and Electronics Engineers, and the National Academy of Inventors and a member of the National Academy of Engineering. He is the author or coauthor of more than 20 books, most recently *Human-Centered AI* (Oxford: Oxford University Press, 2022).

This TechBrief was also produced with guidance from: Kristian Hammond (Bill and Cathy Osborn Professor of Computer Science at Northwestern University in Evanston, Illinois); Larry Medsker (research professor in the Graduate School of Education and Human Development at the George Washington University in Washington, DC); and Marc Rotenberg (founder, Center for AI and Digital Policy in Washington, DC).



October 26, 2022

## STATEMENT ON PRINCIPLES FOR RESPONSIBLE ALGORITHMIC SYSTEMS<sup>1</sup>

Algorithmic systems, often based on artificial intelligence (AI),<sup>2</sup> are increasingly being used by governments and companies to make or recommend decisions that have far-reaching effects on individuals, organizations, and society. Many decisions in employment, credit, access to education, health care, and even criminal justice are made by machines, often without further substantive review by humans. While algorithmic systems hold the promise of making society more equitable, inclusive, and efficient, those results do not automatically flow from automation. Like decisions made by humans, machine-made ones can also fail to respect the rights of individuals and result in harmful discrimination and other negative effects. It is imperative, therefore, that algorithmic systems comply fully with established legal, ethical, and scientific norms and that the risks of their use be proportional to the specific problems being addressed.

An algorithm is a self-contained step-by-step set of operations used to perform calculations, data processing, and automated reasoning tasks. Many AI algorithms are based on statistical models that are “learned” or “trained” from datasets by using machine learning (ML). Others are driven by analytics: the discovery, interpretation, and communication of meaningful patterns in data.

Algorithms and other underlying mechanisms used by AI/ML systems to make specific decisions can be opaque, rendering them less understandable and making it more difficult to determine whether their outputs are biased or erroneous. ,

<sup>1</sup> This document updates and builds upon the ACM Europe and US Technology Policy Committees’ 2017 joint [Statement on Algorithmic Transparency and Accountability](#). The lead authors of this document were Ricardo Baeza-Yates and Jeanna Matthews. Important contributions were made by Vijay Chidambaram, Simson Garfinkel, Carlos E. Jimenez-Gomez, Bran Knowles, Arnon Rosenthal, Ben Schneiderman, Stuart Shapiro, and Alejandro Saucedo. Comments and other assistance also were provided by: Michel Beaudouin-Lafon, Jean Camp, Cansu Canca, Brian Dean, Jeremy Epstein, Oliver Grau, Chris Hankin, Jim Hendler, Harry Hochheiser, Lorena Jaume-Palasi, Lorraine Kisselburgh, Marc Rotenberg, Gerhard Schimpf, Jonathan Smith, Gurkan Solmaz and Alec Yasinsac.

<sup>2</sup> AI as used here refers to systems that employ machine learning (ML), including deep learning, reinforcement learning, statistical inference, or other algorithmic approaches from this field. Our recommendations also apply to algorithmic systems more broadly, including those not employing AI according to this definition.

Factors that make these systems opaque may be:

- informational (the data to train models and create analytics are used without the data subject's knowledge or explicit consent);
- technical (the algorithm may not lend itself to easy interpretation);
- economic (the cost of providing transparency may be excessive);
- competitive (transparency may compromise trade secrets or allow gaming/manipulation of decision boundaries); and/or
- social (revealing input may violate privacy expectations).

Even well-engineered algorithmic systems can produce unexplained outcomes or errors. They may contain bugs, or the training data used may not have been appropriate for the intended use. The conditions of their use also may have changed, thereby invalidating assumptions on which the design of such systems was based.

Further, simply using a widely representative dataset does not guarantee that the system will be free from bias. The way the data are processed, the user feedback loop, and how the system is deployed can all introduce problems. To mitigate the risks of bias or inaccuracy inherent in the use of automated decision-making systems:

- System builders and operators should adhere to the same standards in selecting inputs or architecting systems to which humans are held when making equivalent decisions;
- AI system developers should undertake extensive impact assessments prior to the deployment of AI systems;
- Policy makers should mandate that audit trails be used to achieve higher standards of accuracy, transparency, and fairness; and
- Operators of AI systems should be held responsible for the decisions they make using the system regardless of whether algorithmic tools are used.

The following instrumental principles, consistent with the ACM Code of Ethics,<sup>3</sup> are intended to foster fair, accurate, and beneficial algorithmic decision-making.

---

<sup>3</sup> The ACM Code of Ethics and Professional Conduct is designed to inspire and guide the ethical conduct of all computing professionals, including current and aspiring practitioners, instructors, students, influencers, and anyone who uses computing technology in an impactful way. The Code includes principles formulated as statements of responsibility, based on the understanding that the public good is always the primary consideration. Each principle is supplemented by guidelines, which provide explanations to assist computing professionals in understanding and applying the principle. See <https://www.acm.org/code-of-ethics>.

## PRINCIPLES FOR RESPONSIBLE ALGORITHMIC SYSTEMS

- 1. Legitimacy and competency:** Designers of algorithmic systems should have the management competence and explicit authorization to build and deploy such systems. They also need to have expertise in the application domain, a scientific basis for the systems' intended use, and be widely regarded as socially legitimate by stakeholders impacted by the system.<sup>4</sup> Legal and ethical assessments must be conducted to confirm that any risks introduced by the systems will be proportional to the problems being addressed, and that any benefit-harm trade-offs are understood by all relevant stakeholders.
- 2. Minimizing harm:** Managers, designers, developers, users, and other stakeholders of algorithmic systems should be aware of the possible errors and biases involved in their design, implementation, and use, and the potential harm that a system can cause to individuals and society. Organizations should routinely perform impact assessments on systems they employ to determine whether the system could generate harm, especially discriminatory harm, and to apply appropriate mitigations. When possible, they should learn from measures of actual performance, not solely patterns of past decisions that may themselves have been discriminatory.
- 3. Security and privacy:** Risk from malicious parties can be mitigated by introducing security and privacy best practices across every phase of the systems' lifecycles, including robust controls to mitigate new vulnerabilities that arise in the context of algorithmic systems.
- 4. Transparency:** System developers are encouraged to clearly document the way in which specific datasets, variables, and models were selected for development, training, validation, and testing, as well as the specific measures that were used to guarantee data and output quality. Systems should indicate their level of confidence in each output and humans should intervene when confidence is low. Developers also should document the approaches that were used to explore for potential biases. For systems with critical impact on life and well-being, independent verification and validation procedures should be required. Public scrutiny of the data and models provides maximum opportunity for correction. Developers thus should facilitate third-party testing in the public interest.<sup>5</sup>

---

<sup>4</sup> Projects with no clear scientific basis (e.g., inferring personality traits from facial images) should not be deployed.

<sup>5</sup> For example, by providing access and APIs for this purpose and removing terms of service clauses that discourage publication of results.

5. **Interpretability and explainability:** Managers of algorithmic systems are encouraged to produce information regarding both the procedures that the employed algorithms follow (interpretability) and the specific decisions that they make (explainability). Explainability may be just as important as accuracy, especially in public policy contexts or any environment in which there are concerns about how algorithms could be skewed to benefit one group over another without acknowledgement. It is important to distinguish between explanations and after-the-fact rationalizations that do not reflect the evidence or the decision-making process used to reach the conclusion being explained.
6. **Maintainability:** Evidence of all algorithmic systems' soundness should be collected throughout their life cycles, including documentation of system requirements, the design or implementation of changes, test cases and results, and a log of errors found and fixed.<sup>6</sup> Proper maintenance may require retraining systems with new training data and/or replacing the models employed.
7. **Contestability and auditability:** Regulators should encourage the adoption of mechanisms that enable individuals and groups to question outcomes and seek redress for adverse effects resulting from algorithmically informed decisions. Managers should ensure that data, models, algorithms, and decisions are recorded so that they can be audited and results replicated in cases where harm is suspected or alleged. Auditing strategies should be made public to enable individuals, public interest organizations, and researchers to review and recommend improvements.
8. **Accountability and responsibility:** Public and private bodies should be held accountable for decisions made by algorithms they use, even if it is not feasible to explain in detail how those algorithms produced their results. Such bodies should be responsible for entire systems as deployed in their specific contexts, not just for the individual parts that make up a given system. When problems in automated systems are detected, organizations responsible for deploying those systems should document the specific actions that they will take to remediate the problem and under what circumstances the use of such technologies should be suspended or terminated.
9. **Limiting environmental impacts:** Algorithmic systems should be engineered to report estimates of environmental impacts, including carbon emissions from both training and operational computations. AI systems should be designed to ensure that their carbon emissions are reasonable given the degree of accuracy required by the context in which they are deployed.

---

<sup>6</sup> Otherwise, the system may become less appropriate as inputs drift from those originally anticipated, or if the underlying real-world conditions change (*e.g.*, facial recognition systems are used on a wider or different demographic than was present in the training data).

## APPLICATION OF THE PRINCIPLES: GOVERNANCE AND TRADE-OFFS

The first principle of legitimacy and competency needs to be considered before implementing an algorithmic system. That is, the deploying body should have a clear governance process for deciding when to design and deploy an algorithmic system. The second principle of minimizing harm, especially discriminatory harm, is a core value of ethics and for that reason also informs other principles. It, and the remaining principles, should be addressed during every phase of system development and deployment to the extent necessary to minimize potential harms. These principles are most important for algorithmic systems that directly affect individuals and where there is little opportunity for humans to intervene.

The degree of transparency demanded of an algorithmic system should be consistent with the system's impact. We recommend identifying impact tiers such that higher requirements for transparency are applied to systems with higher levels of impact (*e.g.*, systems with risk to human life or systems in regulated areas such as hiring, housing, credit, and allocation of public resources). Similarly, the level of maintenance required should be commensurate with the impact of the system.

Professionals responsible for applying these principles must decide on necessary trade-offs based on their domain knowledge and consultation with stakeholders. Examples of such trade-offs include:

- Solutions should be proportionate to the problem being solved, even if that affects complexity or cost (*e.g.*, rejecting the use of public video surveillance for a simple prediction task).
- A wide variety of performance metrics should be considered and may be weighted differently based on the application domain. For example, in some health care applications the effects of false negatives can be much worse than false positives, while in criminal justice the consequences of false positives (*e.g.*, imprisoning an innocent person) can be much worse than false negatives. The most desirable operational system setup is rarely the one with maximum accuracy.
- Concerns over privacy, protecting trade secrets, or revelation of analytics that might allow malicious actors to game the system can justify restricting access to qualified individuals, but they should not be used to justify limiting third-party scrutiny or to excuse developers from the obligation to acknowledge and repair errors.
- Transparency must be paired with processes for accountability that enable stakeholders impacted by an algorithmic system to seek meaningful redress for harms done. Transparency should not be used to legitimize a system or to transfer responsibility to other parties.

- When a system's impact is high, a more explainable system may be preferable. In many cases, there is no trade-off between explainability and accuracy. In some contexts, however, incorrect explanations may be even worse than no explanation (*e.g.*, in health systems, a symptom may correspond to many possible illnesses, not just one).

Public policy is important. It is difficult to expect market forces to incentivize private companies to balance trade-offs that involve risks to individuals and to society where such companies' own interests are different. Public policies thus are necessary to require, or at least encourage, impact assessments and levels of explainability and auditability for different classes of systems. Public policies that clarify where audit trails are recorded and who has access to them will encourage designers and developers to consider failure modes and increase trust from users, stakeholders, and oversight bodies.

The foregoing recommendations focus on the responsible<sup>7</sup> design, development, and use of algorithmic systems; liability must be determined by law and public policy. The increasing power of algorithmic systems and their use in life-critical and consequential applications means that great care must be exercised in using them. These nine instrumental principles are meant to be inspirational in launching discussions, initiating research, and developing governance methods to bring benefits to a wide range of users, while promoting reliability, safety, and responsibility. In the end, it is the specific context that defines the correct design and use of an algorithmic system in collaboration with representatives of all impacted stakeholders.

---

<sup>7</sup> Designers and developers are urged to produce sufficient evidence of the reliability of a system so that it can be used responsibly, rather than putting the burden on the user to trust systems without sufficient evidence (*e.g.*, as in trustworthy AI).

**Post-Hearing Questions for the Record  
Submitted to Mr. Jacob Siegel  
From Ranking Member Rand Paul**

**“Artificial Intelligence in Government”  
May 16, 2023**

1. Mr. Siegel, in your essay, you argue that the concept of disinformation has been co-opted by the U.S. government, big tech, and the media to maintain control over society. You say that the recently exposed state-corporate censorship regime is only the beginning of the information war. Can you elaborate on what the second phase will look like?
2. Mr. Siegel, you note that the underlying philosophy of the war against disinformation is that individuals cannot be trusted with their own mind. What does this say about how the federal government views the ability of the American people to think for themselves? What are the consequences of allowing the government to control the flow of information and manipulate public discourse?
3. Mr. Siegel, you discuss the idea of “pre-bunking” which is a form of censorship that is built into social media algorithms to stop citizens from forming specific social and political belief systems. Where is “pre-bunking” currently being employed and what are the consequences of this kind of censorship?
4. Mr. Siegel, you mention the establishment of the Global Engagement Center (GEC) as the government's lead counter-disinformation agency. Can you discuss the potential ethical and societal implications of the GEC's collaboration with tech executives and their influence on information control?
5. Mr. Siegel, four days ago the Pentagon posted a job advertisement for a Senior Advisor for its newly established Influence and Perception Management Office. ODNI recently stood up a Foreign Malign Influence Center. What is your reaction to the expanding number of federal offices focused on disinformation?
  - a. Do you think it is appropriate for the Defense Department and Intelligence Community to be involved in surveilling and censoring Americans protected speech?
6. Mr. Siegel, are you concerned with the Director of Research of NSA's recent comments suggesting that the intelligence community should rely on commercially available AI models from companies like Meta, Google, and Microsoft for national security purposes?

**Mr. Siegel did not respond by time of printing. If a response is received, it will be on file for public inspection in the committee offices.**