National Artificial Intelligence Advisory Committee: Year-Two Insights Report

The National Artificial Intelligence Advisory Committee (NAIAC)

May 2024

NAIAC MEMBERS

Miriam Vogel, *Chair*
President and CEO of EqualAI

James Manyika, *Vice Chair*
Senior Vice President, Google, President for Research, Technology & Society

Amanda Ballantyne
Director of the AFL-CIO Technology Institute

Jack Clark
Co-founder of Anthropic

David Danks
Professor of Data Science and Philosophy at the University of California, San Diego

Victoria A. Espinel
President and CEO of BSA | The Software Alliance

Paula Goldman
Chief Ethical and Humane Use Officer at Salesforce

Susan Gonzales
Founder and CEO of AIandYou

Janet Haven
Executive Director of Data & Society

Daniel E. Ho
William Benjamin Scott and Luna M. Scott Professor of Law, Professor of Political Science, and Professor of Computer Science (by courtesy), Senior Fellow, Stanford Institute for Human-Centered AI, Stanford University

Ayanna Howard
Dean of Engineering at The Ohio State University

Jon Kleinberg
Professor in the Departments of Computer Science and Information Science at Cornell University

Ramayya Krishnan
W. W. Cooper and Ruth F. Cooper Professor of Management Science and Information Systems, Carnegie Mellon University

Ashley Llorens
Vice President, Distinguished Scientist, and Managing Director at Microsoft Research

Haniyeh Mahmoudian
Global AI Ethicist at DataRobot, Inc.

Christina Montgomery
Chief Privacy & Trust Officer and Vice President at IBM

Liz O'Sullivan
CEO of Vera

Fred Oswald
Professor and Herbert S. Autrey Chair in Social Sciences, Rice University

Trooper Sanders
CEO of Benefits Data Trust

Navrina Singh
Founder and CEO of Credo AI

Swami Sivasubramanian
Vice President for Data and Machine Learning Services at Amazon Web Services

Keith Strier
Vice President for Worldwide AI Initiatives at NVIDIA

Reggie Townsend
Vice President of Data Ethics at SAS Institute

NAIAC-LE SUBCOMMITTEE MEMBERS

Armando Aguilar
Assistant Chief of Police, Miami Police Department

Anthony Bak
Head of AI, Palantir

Amanda Ballantyne
Director of the AFL-CIO Technology Institute

Jane Bambauer
Director - Marion B. Brechner First Amendment Project, Brechner Eminent Scholar at the College of Journalism and Communications and at Levin College of Law, University of Florida

Esha Bhandari
Deputy Director of the American Civil Liberties Union's Speech, Privacy, and Technology Project

Jennifer Eberhardt
Professor of Organizational Behavior and Psychology, Stanford University

Farhang Heydari
Assistant Professor of Law, Vanderbilt Law School

Benji Hutchinson
Chief Revenue Officer of Rank One Computing

Rashawn Ray
Vice-President and Executive Director of the AIR Equity Initiative

Cynthia Rudin
Professor of Computer Science, Electrical and Computer Engineering, Statistical Science, Mathematics, Biostatistics & Bioinformatics at Duke University

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The rapid advancement of artificial intelligence (AI) technologies remains one of the most profound and complex developments today, in the United States and across the world. Each day, AI creates new opportunities, but also presents new risks for government and public stakeholders alike, requiring vigilant oversight and proactive governance. The U.S. government must meet this watershed moment with strong leadership and a clear vision to ensure that AI technology, rules, and standards reflect our American values. The National AI Advisory Committee ("NAIAC" or the "Committee") is tasked with supporting this critical and timely imperative.

In order to fulfill our mission, over our first year (May 2022 - May 2023), NAIAC prioritized learning about the current efforts and challenges surrounding AI use and oversight by the federal government. We engaged with key stakeholders inside the government and across the globe, to educate ourselves as a group and determine how best to harness our collective expertise in AI in offering guidance to our congressionally mandated audience, the President and the Executive Branch. Accordingly, we organized into five working groups aligned with statutory mandates — Trustworthy AI, Research & Development, Workforce, Competitiveness, International Collaboration — and had three public meetings to deliberate and prepare our Year 1 report.

Given our belief that our Committee — as well as the federal government — benefits greatly from hearing from the broadest array of expert and impacted community voices, we invited an array of experts to share their insights at our public meetings. Starting with our first Committee-run meeting in Palo Alto, California in October 2022 and since NAIAC released its Year 1 Report in May 2023, our Committee engaged leaders in government, industry, civil society, and academia, and technologists, researchers, policymakers, educators, and citizens to serve on these public panels. These discussions have enriched the public discourse on AI, and directly influenced the deliverables we present to the President and the wider public. Each session with a link to the recording is included in this report to facilitate access to viewing any session of interest.

The AI landscape changed during our second year of service, and NAIAC adapted its organizational structure and operational cadence to respond effectively to the corresponding dynamic needs of federal executive AI policy development.

We reconfigured our working groups and significantly increased the frequency of our public meetings, convening public sessions nearly every month. These sessions enabled increased public discourse and deliberations, and ultimately, deliverables.

Deliverables were direct outgrowths from our public sessions with stakeholders and the specific domain expertise of our Committee members in areas such as Generative AI, regulatory mechanisms and imperatives for AI policymaking, opportunities and challenges of AI in our workforce, and AI literacy.

We also welcomed the newly appointed members of the congressionally mandated Subcommittee on AI and Law Enforcement in our public meeting last October. This Subcommittee provides recommendations to NAIAC on AI use and oversight in the domain of federal law enforcement. And we have expanded the types of deliverables we generate in our second year. In addition to the standard yearly report, the Committee has published Findings, Recommendations, Statements, and other materials to provide timely and targeted responses to the rapidly evolving AI landscape.

This report encapsulates the activities and contributions of NAIAC throughout our second year. As a whole, our work reinforces the fact that AI technology and its use need to be frequently and continuously discussed, evaluated, and improved at the highest levels of government, as AI is used to provide opportunities and impacts every single American in all communities at an ever increasing pace. Our work stresses the importance of immediate and careful attention and action by the federal government to ensure AI is used to support American leadership and values, and that the benefits of AI technologies are leveraged to enhance the safety, security, and opportunities for all Americans.

INTRODUCTION

Since NAIAC published our inaugural report in May 2023, AI-related opportunities, risks, public policy, and innovation have all evolved at record speed.

The public policy landscape has shifted, most notably with the introduction of President Biden's Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (AI EO) on October 30, 2023. The impact of the AI EO was propelled further by the Office of Management and Budget's release of the Memorandum Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence on March 28, 2024. The Memorandum directs agencies to advance AI governance and innovation while managing risks from use of AI, particularly those affecting the rights and safety of the public. The Department of Commerce continued to advance AI policy and infrastructure development through each of its bureaus, as well as through the launch of the first-ever U.S. AI Safety Institute (AISI), which is housed at the National

Institute of Standards and Technology (NIST). This accompanies the creation of the U.S. AI Safety Institute Consortium (AISIC), to unite AI creators and users, academics, government and industry researchers, and civil society organizations in support of the development and deployment of safe and trustworthy AI. There are also various treaties and international bodies the Executive Branch convenes on AI policy coordination, including the U.S-EU Joint Trade and Technology Council ("TTC").

Meanwhile, the Department of State is carrying out multiple AI initiatives, including co-leadership of the TTC and the Hiroshima AI Process. The President explicitly addressed[1] the promises and perils of AI during his most recent State of the Union. And congressional leaders continue to hold numerous hearings and propose AI bills, both predating and following the AI Insight forums convened on the Hill last Fall.

Public policy beyond U.S. borders has likewise progressed at a rapid clip. EU lawmakers reached an agreement on the AI Act in December 2023, and the legislation was adopted in March 2024. Singapore updated its National AI Strategy and convened the Singapore Conference on AI. The UK held its influential AI Safety Summit, corresponding with the Bletchley Declaration by Countries Attending the AI Summit and the launch of the UK AI Safety Institute. This only scratches the surface of top-line developments in the many countries that have convened and participated in high-level AI dialogues and make progress on their respective national AI strategies.

International bodies have also made progress. The United Nations convened an AI Advisory Body that published its Interim Report, titled "Governing AI for Humanity," in December 2023. The Global Partnership on AI convened experts and policy makers and issued the Delhi declaration in December 2023. And G7 leaders endorsed the Hiroshima AI Process that has laid out 11 governing principles focused on risk mitigation, transparency, security, standards, and other key issues.

Amid these and other significant policy developments, AI technology itself has matured. For example, foundation models continue to expand in their power and capabilities, and AI systems are being deployed across an increasingly wide range of settings, creating an equally wide spectrum of opportunities and challenges.

Central to all these rapid changes in AI technologies and policies are the advancement and protection of the American people and their communities. There is a great deal more to learn about how AI technologies and policies affect our

[1] "Here at home, I've signed over 400 bipartisan bills. But there's more to pass my Unity Agenda... Harness — harness the promise of AI to protect us from peril. Ban AI voice impersonations and more." President Biden, 2024 State of the Union, The White House. https://www.whitehouse.gov/state-of-the-union-2024/.

people and our communities, including in the realms of education, work, health, family, and all other aspects of our lives.

While the landscape for AI policy and capabilities has changed dramatically since our term began, what remains constant is our focus at NAIAC. We are humbled by and steadfast in our mission. We understand the critical nature of our responsibilities at this time as we issue advice and recommendations to the President and the White House. Our goal is to advance policy that ensures AI is effective, trustworthy, innovative, and inclusive, and ultimately, that it benefits all Americans.

ADOPTIONS TO DATE

NAIAC is encouraged that the Biden Administration's AI EO and other recent executive actions have fulfilled and operationalized many priorities that were highlighted and recommended in the Committee's Year 1 Report. This section lists the Administrative action that corresponds with NAIAC guidance or recommendations. In addition to the below impacts, we are also encouraged to see our work cited as a foundation for the U.S. Senate's Future of AI Innovation Act.

NAIAC action

In the Year 1 Report, NAIAC proposed the creation of Chief AI Officer roles within federal agencies in order to organize and elevate AI leadership. We recommended these be senior positions tasked with capturing the benefits and promoting the adoption of trustworthy AI, both inside and outside of government. Specifically, we recommended that Chief AI Officers promote responsible AI innovation within their respective agencies; ensure compliance and manage risk; and ensure procurement of AI systems aligns with the agency's AI principles.

↳ **Corresponding White House action**

Section 10 of the executive order details "the requirement to designate at each agency [...] a Chief Artificial Intelligence Officer who shall hold primary responsibility in their agency, in coordination with other responsible officials, for coordinating their agency's use of AI, promoting AI innovation in their agency, and managing risks from their agency's use of AI."

NAIAC action

In the Year 1 Report, NAIAC proposed stronger AI leadership and coordination at each department and agency, in order to maintain global leadership and ensure all relevant stakeholders have a seat at the table. Specifically, we recommended the establishment of an Emerging Technology Council (ETC) to coordinate and drive technology policy across the U.S. government. The ETC would elevate interrelated key issues like domestic security, impacts to trade and labor, and mitigating algorithmic bias, and also would close relevant gaps between OSTP, NSC, NEC, OMB, departments, and agencies.

↳ **Corresponding White House action**

Section 12 of the executive order "establishe[s], within the Executive Office of the President, the White House Artificial Intelligence Council. The function of the White House AI Council is to coordinate the activities of agencies across the Federal Government to ensure the effective formulation, development, communication, industry engagement related to, and timely implementation of AI-related policies, including policies set forth in this order."

NAIAC action

In the Year 1 Report, NAIAC proposed supporting the public and private adoption of the NIST AI Risk Management Framework (AI RMF) in order to operationalize trustworthy AI governance. We cast the AI RMF as a key tool for effectively addressing risks in all phases of the AI lifecycle, and advocated that federal agencies leverage the AI RMF to address bias, discrimination, and other social harms and issues that arise when building, assessing, and governing AI systems. Specifically, we recommended the Administration launch a pilot program directing at least three agencies to implement the AI RMF and report on their lessons learned within one year.

↳ **Corresponding White House action**

The executive order calls for a range of federal agencies — including the Departments of Homeland Security and Commerce — to adopt the NIST AI Risk Management Framework.

NAIAC action

In the Year 1 Report, NAIAC proposed advancing the implementation plan from the National AI Research Resource (NAIRR) final report to create a large-scale national research resource. We noted that the AI R&D ecosystem in the U.S. is inaccessible to many individuals, groups, and organizations due to the vast data and computational requirements and costs. Students, non-profit organizations, local and tribal agencies, startups, and small businesses all struggle to contribute and compete in the advancement of trustworthy AI systems. We stressed that the NAIRR would directly address this problem, providing much-needed support and opportunities to under-resourced and underrepresented groups.

↳ **Corresponding White House action**

Section 5 of the executive order "launch[es] a pilot program implementing the National AI Research Resource (NAIRR), consistent with past recommendations of the NAIRR Task Force. The program shall pursue the infrastructure, governance mechanisms, and user interfaces to pilot an initial integration of distributed computational, data, model, and training resources to be made available to the research community in support of AI-related research and development."

NAIAC action

In the Year 1 Report, NAIAC proposed reforming immigration policies to attract and retain international tech talent. The current U.S. immigration system has remained unchanged for a decade and fails to respond to the needs of the U.S. economy adequately: More than half of the U.S. AI workforce and about 66 percent of U.S. AI

graduates were born abroad. This creates major obstacles for immigrants to remain in the U.S. after graduating from our first-class higher education institutions, thus depriving the nation of critical tech skills necessary for our AI economy to thrive. For example, H-1B work visas are selected through a lottery system and annually capped at 65,000 (plus an additional 20,000 for those with graduate degrees). And Green Cards for permanent residency have per-country quotas and often unworkable backlogs. We recommended the U.S. government lower these barriers so international tech talent can come, work, and stay in the country.

↳ **Corresponding White House action**

Section 5 of the executive order states "the Secretary of Homeland Security shall review and initiate any policy changes the Secretary determines necessary and appropriate to clarify and modernize immigration pathways for experts in AI and other critical and emerging technologies, including O-1A and EB-1 noncitizens of extraordinary ability; EB-2 advanced-degree holders and noncitizens of exceptional ability; and startup founders in AI and other critical and emerging technologies using the International Entrepreneur Rule."

NAIAC action

In the Year 1 Report, NAIAC proposed scaling an AI-capable federal workforce through various training programs and other strategic investments. Specifically, we recommended the U.S. government create a United States Digital Service Academy (an accredited, degree-granting university in the mold of the U.S. military service academies) and a Digital Service Academic Compact (which would allow Academy graduates to complete AI-relevant degrees at participating institutions). We also recommended the U.S. government invest in the incumbent federal workforce through initiatives like Office of Personnel Management-established AI career fields and upskilling courses. NAIAC also proposed the U.S. government boost short-term federal AI talent by strengthening programs like 18F, the Presidential Innovation Fellowship, the USDS, the GSA Centers of Excellence, and Intergovernmental Personnel Act assignments.

↳ **Corresponding White House action**

Section 10 of the executive order states the Director of OSTP, the Director of OMB, and others "identify priority mission areas for increased Federal Government AI talent, the types of talent that are highest priority to recruit and develop to ensure adequate implementation of this order and use of relevant enforcement and regulatory authorities to address AI risks, and accelerated hiring pathways."

NAIAC action

In the Year 1 Report, NAIAC proposed continued international collaboration and leadership on AI. We stressed that U.S. leadership on AI is central to a future that reflects strong democratic values around the world and where the U.S. maintains its competitive edge. Specifically, we recommended the U.S. government expand and deepen international alliances; internationalize the NIST AI RMF; fully fund the Bureau of Cyberspace and Digital Policy and Office of the Special Envoy for Critical and Emerging Technology; and establish the U.S.-based Multilateral AI Research Institute (MAIRI) to facilitate AI research and development.

↳ **Corresponding White House action**

The executive order calls for a range of international collaborations. For example, Section 11 calls on the Secretary of State and others to "lead efforts to establish a strong international framework for managing the risks and harnessing the benefits of AI, including by encouraging international allies and partners to support voluntary commitments similar to those that United States companies have made in pursuit of these objectives."

NAIAC action

In the Year 1 Report, NAIAC proposed establishing a U.S.-based multilateral coalition for international cooperation on accelerating AI for addressing global climate and sustainability goals. However, while progress in these fields is occurring rapidly, it is mostly in isolation within individual academic, government,

and private sector labs across nations. We stressed that the U.S. can more rapidly and effectively advance the development and deployment of important AI-supported climate innovations — such as Earth-scale Digital Twins and Sustainable Computing approaches — by working with international partners.

↳ **Corresponding White House action**

Section 5 of the executive order calls on the Secretary of Energy and others to “take steps to expand partnerships with industry, academia, other agencies, and international allies and partners to utilize the Department of Energy’s computing capabilities and AI testbeds to build foundation models that support new applications in science and energy, and for national security, including partnerships that increase community preparedness for climate-related risks, enable clean-energy deployment (including addressing delays in permitting reviews), and enhance grid reliability and resilience.”

NAIAC action

In the “Committee Member Perspectives” section of the Year 1 Report, four members of NAIAC highlighted the need for and utility of a rights-based approach to AI governance, including worker rights and participation.

↳**Corresponding White House action**

Section 2 of the executive order prioritizes equity, civil and human rights, worker rights, and worker participation in AI development. It reads: “Artificial Intelligence policies must be consistent with my Administration’s dedication to advancing equity and civil rights. My Administration cannot — and will not — tolerate the use of AI to disadvantage those who are already too often denied equal opportunity and justice." And: "In the workplace itself, AI should not be deployed in ways that undermine rights, worsen job quality, encourage undue worker surveillance, lessen market competition, introduce new health and safety risks, or cause harmful labor-force disruptions.”

Further, the March 28, 2024 Office of Management and Budget memo also prioritizes a "rights and safety" framework for addressing AI harms. The memo details instances in which AI systems are presumed to impact rights or safety, and requires federal agencies to undertake a series of evaluations and impact assessments which the system must pass before it is put into use.

YEAR 2 PRIORITIES, ACTIVITIES, AND DELIVERABLES

The following three subsections review NAIAC's Year 2 deliverables: the Committee's public briefings; the Committee's Findings, Recommendations, and other Materials; and the work of the Subcommittee on AI and Law Enforcement.

PUBLIC SESSIONS

These sessions are part of a series NAIAC held with various stakeholders to hear their concerns and hopes for AI, as well as specific recommendations they would like us to consider in our counsel to the President and the White House.

Briefing Title	Speakers	Date & Link
AI and Civil Rights	**Emily Chi**, Asian Americans Advancing Justice **Maria Town**, American Association for People with Disabilities **JudeAnne Heath**, Hispanic Technology and Telecommunications Partnership **Patrice Willoughby**, National Association for the Advancement of Colored People **Lisa Rice**, National Fair Housing Alliance **Frank Torres**, Leadership Conference on Civil and Human Rights	June 20, 2023
AI, Civil Rights, and Civil Liberties	**Olga Akselrod**, American Civil Liberties Union	June 22, 2023

	Tawana Petty, Algorithmic Justice League **Alexandra Givens**, Center for Democracy and Technology **Anna Tumadottir**, Creative Commons **Vinhcent Le**, Greenlining Institute **Quinn Anex-Rios**, Lawyers Committee for Civil Rights under Law **Amalea Smirniotopoulos**, National Association for the Advancement for Colored People **Hannah Sassaman**, Peoples Tech Project	
AI and Civil Rights	**Laura Montoya**, LatinX in AI **Yang Cheung**, Women in AI **Mason Grimshaw**, Indigenous in AI **Bhuva Shakti**, Women in AI **Arjun Subramonian**, Queer in AI **Gelyn Watkins**, Black in AI	**June 27, 2023**
AI and Education, Immigration, Labor, and Trade Leadership	**Erica Fein**, International Brotherhood of Electrical Workers **Brett Gibson**, American Federation of Labor and Congress of Industrial Organizations **Eric Gottwald**, American Federation of Labor and Congress of Industrial Organizations **Faraz Khan**, International Federation of Professional and Technical Engineers **Dan Mauer**, Communications Workers of America **Michael Wasser**, Department of Professional Employees **Rob Weil**, American Federation of Teachers	**June 27, 2023**
NAIAC Public Meeting	(No speakers)	**July 19, 2023**
Harnessing AI for the Benefit of Humanity	**Yoshua Bengio**, Université de Montreal **Francesca Rossi**, IBM	**Aug. 3, 2023**

	Stuart Russell, University of California, Berkeley	
Open-Source Generative AI	**Sam Gregory**, Witness **Abby Kukura**, Special Competitive Studies Project (SCSP) **Percy Liang**, Stanford University **Joelle Pineau**, Meta	**Aug. 3, 2023**
NAIAC Public Meeting	(No speakers)	**Sept. 12, 2023**
Identifying AI's Opportunities for Societal Benefit	**Priya Donti**, Climate Change AI **Salman Khan**, Khan Academy **Julie Kientz**, University of Washington **Isaac Kohane**, Harvard Medical School and Brigham and Women's Hospital	**Sept. 29, 2023**
How to Regulate Copyright Issues within the Context of Generative AI	**Aaron Cooper**, Business Software Alliance **Keith Kupferschmid**, Copyright Alliance **Catherine Stihler**, Creative Commons	**Sept. 29, 2023**
NAIAC Public Meeting	(No speakers)	**Oct. 19, 2023**
NAIAC Public Meeting	(No speakers)	**Nov. 15, 2023**
NAIAC Public Meeting	(No speakers)	**Dec. 13, 2023**
NAIAC Public Meeting	(No speakers)	**Jan. 19, 2024**
AI for Science	**Anima Anandkumar**, Caltech **Jeff Dean**, Google **Ece Kamar**, Microsoft **Rachel Mandelbaum**, Carnegie Mellon University **Kevin Murphy**, NASA	**Feb. 22, 2024**

Data Transparency	**Meghan Dierks**, Komodo Health **Jon Iwata**, Data & Trust Alliance **Yacine Jernite**, Hugging Face **Jeffery Smith**, Department of Health and Human Services	**Feb. 22, 2024**
AI in Latin America and the Caribbean	**Rodrigo Ferreira**, Rice University **Armando Guio Español**, Global Network of Internet & Society Centers, Harvard University **Lucía Trochez Ardila**, PIT Policy Lab **César Uribe**, Rice University	**Feb. 22, 2024**
Concepts in AI Safety: Scoping "Safety" in AI	**Deborah Raji**, Mozilla **Vincent Conitzer**, Carnegie Mellon University **Chris Meserole**, Frontier Model Forum **Arvind Narayanan**, Princeton University **Julia Angwin**, ProofNews **John C. Inglis**, Office of the National Cyber Director **Suresh Venkatasubramanian**, Brown University	**March 5, 2024**
Operationalizing AI Safety: Methodologies and Organizational Practice	**William Isaac**, DeepMind **Miranda Bogen**, Center for Democracy and Technology **Angela Jiang**, DeepMind **Tamara Kneese**, Data & Society **Joshua Kroll**, Naval Postgraduate School **Madhulika Srikumar**, Partnership on AI **Hoda Heidari**, Carnegie Mellon University **Yejin Choi**, University of Washington	**March 5, 2024**
Subcommittee on AI and Law	(No speakers)	**April 5, 2024**

Enforcement Public Meeting		
AI Transition for Workers	**Ifeoma Ajunwa**, Emory Law School **Justin Brown**, Oklahoma Department of Human Services **Michael Ellison**, CodePath	April 16, 2024
AI for Science	**Jean-Paul Chretien**, DARPA **Surya Ganguli**, Stanford University **Susan Gregurick**, NIGMS/NIH **Michael I. Jordan**, University of California, Berkeley **James Swanson**, Johnson & Johnson	April 16, 2024

FINDINGS, RECOMMENDATIONS, & MATERIALS

Below are Findings, Recommendations, and Materials deliberated on and published by NAIAC between May 2023 and May 2024.

Many of these documents are directly informed by the public briefings NAIAC organized throughout our Year 2. We held deep and far-ranging conversations on the intersection of AI and civil rights, privacy, education, immigration, trade, labor, healthcare, climate, creativity, science, and other key issues. NAIAC also published statements in response to major AI developments, including debates about the existential risks of AI and the introduction of the AI EO.

Each entry has a summary and link to the full document. We encourage readers to view the full documents.

FINDINGS:

Implementing the NIST AI RMF With a Rights-Respecting Approach (September 2023)
Implementation of the NIST AI RMF can help manage risks associated with AI and protect Americans' civil rights in the context of AI. (This Finding was a

precursor to the October 2023 Recommendation with the same name, which is listed below.)

Read the full finding document

The Potential Future Risks of AI (October 2023)

AI carries a number of positive developments and opportunities, but also several potential dangers if the technology is misused. These dangers include malicious objectives, unintended circumstances, and circumvention of safety measures; economic and societal risks; and catastrophic risks like destabilization of democracy.

Read the full finding document

Exploring the Impact of AI (November 2023)

As AI becomes more pervasive, communities in the U.S. are being impacted by this technology, including marginalized communities like American women, Asian Americans, Black Americans, Disabled Americans, Indigenous Americans, Latinx Americans, LGBTQ+ Americans, and others. Each of these communities — and the smaller communities within them — has a unique relationship with AI. But there are commonalities across groups. For example, AI has the potential to benefit marginalized communities through economic opportunities, innovation, and other means — but only if these communities have equal access to relevant resources. Alternatively, AI can also harm these communities, often by amplifying existing societal biases.

Read the full finding document

RECOMMENDATIONS:

Recommendation for International Emerging Economies (August 2023)

Cooperation with other countries is necessary to develop solutions on a global scale. An important part of this cooperation is dialogue with emerging

economies to understand the unique challenges and opportunities AI presents to other nations and to the global community. To address this, the federal government should ensure global gatherings or summits to increase global cooperation on AI include representatives from emerging economies. If such a gathering is not possible, we urge the federal government to proactively and directly engage with emerging economies.

Read the full recommendation document

Creating Institutional Structures to Support Safer AI Systems (October 2023)

The safety and reliability of AI is a critically necessary condition to engender trust and spur its widespread adoption and deployment. To address this, the federal government should establish a multi-agency-sponsored AI Lead Rapid Response Team (ALRT) to support advancing the safe and responsible development of AI. ALRT should focus on six core activities, including risk monitoring and collaboration with industry and academia.

Read the full recommendation document

AI's Procurement Challenge (October 2023)

The federal government is one of the largest purchasers of AI systems, and getting AI procurement right is also essential for the federal government to serve the American people in the 21st century. To address this, each federal agency should prioritize AI procurement for realizing its mission and include AI procurement within its Presidential transition plan.

Read the full recommendation document

Implementing the NIST AI RMF with a Rights-Respecting Approach (October 2023)

In its Year 1 Report, NAIAC recommended that the White House, in order to "operationalize trustworthy AI governance," "[s]upport public and private

adoption of the NIST AI Risk Management Framework." In doing so, federal agencies should: ensure AI cannot be used in ways that violate civil rights; define what "algorithmic discrimination" and other automated system harms look like within their jurisdiction; and develop the technical capacity to address and enforce against algorithmic discrimination and other AI harms.

Read the full recommendation document

Generative AI Away from the Frontier (October 2023)

It is important to have a clear understanding of the potential risks posed by non-frontier, more widely available Generative AI systems. To address this, the federal government should encourage companies that use these open-weight models to conduct voluntary risk assessments of Generative AI systems with more constrained access; and collaborate with diverse stakeholders to conduct risk assessments of Generative AI systems with unconstrained access.

Read the full recommendation document

Second Chance Skills and Opportunity Moonshot (October 2023)

Evidence suggests that Generative AI can improve the quality of output for those who are mid-range performers in professional tasks such as writing, perhaps even closing the gap with top performers. When coupled with the clear rise in skills-based hiring and stackable credentials, a wider range of workers have more pathways to attain meaningful employment and careers. To address this, the federal government should launch a moonshot to support adults in need of second chance skills and opportunities.

Read the full recommendation document

Improve Monitoring of Emerging Risks from AI through Adverse Event Reporting (November 2023)

A complete and accurate assessment of AI risks is essential to safeguard U.S. security, economic, and democratic interests. Yet the emerging risks of

advanced AI models are not well understood, posing significant challenges to regulation aimed at mitigating potential harms. To address this, the federal government should pilot an adverse event reporting system for AI that would allow developers, deployers, and users to report harmful post-deployment events stemming from AI systems.

Read the full recommendation document

Enhancing AI Literacy for the United States of America (November 2023)

As AI continues to dominate public discourse, more questions arise about the preparedness of the American public to understand, embrace, trust or adapt to an AI-infused world. To address this, the federal government should create a National AI Literacy Campaign that fosters national AI literacy; leverages the Biden Administration's digital equity campaign as a framework to create said National AI Literacy Campaign; invests in formal educational or existing learning frameworks to advance the AI literacy of the American population; and invests in informal learning opportunities such as stand-alone public sessions, social media campaigns, and public messaging efforts.

Read the full recommendation document

National Campaign on Lifelong AI Career Success (November 2023)

Later-in-life workers, veterans, and caregivers all face unique barriers to employment and career success, and the presence and growth of AI will create additional unique challenges. To address this, the federal government should mount a National Campaign on Lifelong AI Career Success to support these workers. This campaign will include breaking down myths about today's high-tech workers and high-tech jobs, while creating committed partnerships and opportunities with industry to develop viable career pipelines. Enhancing career mobility and career opportunities involving AI creates a more prosperous American citizenry and a stronger national economy.

Read the full recommendation document

Implementation of the NIST AI Safety Institute (December 2023)

Continued American leadership in AI development requires a commitment to proven and trusted methods, standards, and frameworks for the safety of AI models and systems. To address this, the federal government should implement a U.S. AI Safety Institute inside NIST that pioneers advances in AI measurement, evaluation, and assurance.

Read the full recommendation document

Note: This recommendation has now been implemented by the White House.

MATERIALS:

Rationales, Mechanisms, and Challenges to Regulating AI: A Concise Guide and Explanation (July 2023)

AI regulation is a deep and nuanced topic. This concise guide explains several rationales for potentially regulating AI; the main types of regulatory interventions; and some of the distinct challenges to effective AI regulation.

Read the full guide

FAQs on Foundation Models and Generative AI (August 2023)

Understanding foundation models and Generative AI is key to understanding the latest developments in the field at large. This concise guide provides an overview of this technology, along with its capabilities, uses, risks, and current guardrails.

Read the full FAQ

Statement on AI and Existential Risk (October 2023)

Amid growing public discourse about the existential risk of AI, NAIAC released a statement noting that "Arguments about existential risk from AI should not detract from the necessity of addressing existing risks of AI. Nor should arguments about existential risk from AI crowd out the consideration of opportunities that benefit society."

Read the full statement

FORTHCOMING:

Pending Findings and Recommendations (May 2024)
Further documents developed between January and May 2024 will be published shortly: FINDINGS: Enhancing AI's Positive Impact on Science and Medicine PROCEEDINGS: Towards Standards for Data Transparency for AI Models FINDINGS & RECOMMENDATIONS: AI Safety RECOMMENDATION: Data Challenges and Privacy Protections for Safeguarding Civil Rights in Government RECOMMENDATIONS: Harnessing AI for Scientific Progress RECOMMENDATION: Provide Authority and Resources to Promote Responsible Procurement Innovation for AI at Government Agencies

SUBCOMMITTEE ON AI AND LAW ENFORCEMENT

The use of AI technologies in the criminal justice system is an urgent issue, and when NAIAC was established, Congress directed the creation of a subcommittee to examine the topic.

The Subcommittee on AI and Law Enforcement identifies and makes recommendations to NAIAC about the legal, ethical, and responsible use of AI technologies if or when they are used:

- To influence a law enforcement action with respect to who, what, or where/when to investigate or engage law enforcement, including decisions related to the investigations of all crime (e.g., white collar crime, human trafficking, cybercrime, street crime, and violations of immigration/customs, as well as decisions related to pretrial detention, bail, corrections, and parole)

- To assess whether law enforcement has conducted its work effectively, and within its legal and ethical limits

- With the result of inducing the collection, combination, integration, or disclosure of data (including by private companies that collaborate or cooperate with law enforcement)

The Subcommittee on AI and Law Enforcement began this work in August 2023. In January, the subcommittee convened for in-person discussions in Miami, Florida. The subcommittee has also written a Year 1 Roadmap and three recommendations so far:

Law Enforcement Subcommittee: Year 1 Roadmap
Roadmap will be published to ai.gov/naiac shortly

Expand the AI Use Case Inventory by Limiting the 'Common Commercial Products' Exception
Recommendation will be published to ai.gov/naiac shortly

Expand the AI Use Case Inventory by Limiting the 'Sensitive Law Enforcement' Exception
Recommendation will be published to ai.gov/naiac shortly

Require Public Use Policies for High-Risk AI
Recommendation will be published to ai.gov/naiac shortly

LOOKING FORWARD

At our public session on May 2, 2024 and going forward, the Committee will continue to hold briefings and provide Recommendations and Findings on critical steps we believe the President should take to achieve the appropriate policy landscape and infrastructure to ensure the realization of trustworthy AI. We will also engage further

with key stakeholders and subject matter experts across the globe to inform our work and recommendations, as well as to share these insights with the White House and general public.

For the remainder of the Spring and Summer 2024, we will prioritize a few key areas of focus:

Education/Awareness:
This working group will map its prior recommendations regarding AI literacy to the recent executive order on AI, specifically sections 2, 4, 6, 7, 8, and 10 of the order. The working group will also seek to increase AI literacy within the context of the newly-created U.S. AI Safety Institute.

International Collaboration:
This working group will collaborate with the Subcommittee on AI and Law Enforcement to facilitate connections with international counterparts. It will support further education around and adoption of the NIST AI Risk Management Framework and other standards with international allies. It will coordinate the development of recommendations for capacity building in emerging economies with the Cyber Space and Digital Policy team at State, exploring ways to close the growing sovereign compute divide and to help foreign governments lock-in the right governance architecture for AI adoption. And it will examine how humanitarian groups — like USAID, the Red Cross, and others — can use responsible AI to further their missions.

AI Futures – Preparedness, Opportunities, and Competitiveness:
This working group will continue to engage with the forward-looking aspects of AI technology and its impacts, with a particular focus on three key areas: AI for science; AI hardware development; and measurement issues at the heart of understanding the impact AI will have on society and the economy. These topics will intersect with topics being addressed by other working groups (e.g., AI and workforce, AI and safety etc.) and we will collaborate across working groups to address these important issues most effectively and comprehensively.

Safety, Trust, and Rights:
This working group will continue to focus on the protection of rights — including worker rights — in the context of AI use, along with the advancement of the emerging science and methodologies of AI safety with the goal of earning public trust for AI systems through rigorous accountability structures. The working group will focus on data governance and transparency, which are integral to developing and deploying trustworthy AI. The group will also focus on the methodologies and approaches being used to ensure AI safety, and the implications for the AI Safety

Institute at NIST as well as other AI governance forums to support American innovation and leadership in ethical AI. Other topics the group will focus on include data collection privacy risks and the ability to conduct racial and gender disparity assessments in government; and the technology tools and infrastructure needed to build meaningful systems of AI accountability, such as through software, audits and impact assessments.

AI in Work and the Workforce:
This working group will engage the public and focus its efforts on exploring and developing new approaches to a just AI transition for American workers. Going beyond today's narrow reliance on reskilling as the solution to economic disruption and dislocation, the working group's efforts will take a much broader human-centered approach. The central goals are to expand workers' pathways to enriching jobs and economic mobility by (a) enhancing their skills, experiences, and career competitiveness while in the workforce, (b) supporting stronger job mobility options that improve careers and communities while fueling a competitive U.S. economy while (c) exploring policy options to enable workers to have access to financial resources and the time required to prepare for a changing economy. It will also explore new metrics that go beyond traditional measures such as GDP to reflect the economic well-being, quality of life, and mobility of the American people. These new metrics could be critically informative in the AI era to help understand where, when, and how economic output and labor and the livelihood of workers might diverge.

We will also use our May 2024 meeting to discuss and confirm plans for the remainder of 2024.

ACKNOWLEDGEMENTS

We want to extend our sincere thanks to the staff at our organizations who support our work and the tireless and comprehensive support from our partners at NIST who support this Committee so effectively, including Elham Tabassi, Melissa Taylor, Cheryl Gendron, Rachel Trello, and Alicia Chambers. We also want to express our gratitude to the individuals who generously shared their time and expertise at NAIAC stakeholder sessions, including:

Emily Chi, Asian Americans Advancing Justice

Maria Town, American Association for People with Disabilities

JudeAnne Heath, Hispanic Technology and Telecommunications Partnership

Patrice Willoughby, National Association for the Advancement of Colored People

Lisa Rice, National Fair Housing Alliance

Frank Torres, Leadership Conference on Civil and Human Rights

Olga Akselrod, American Civil Liberties Union

Tawana Petty, Algorithmic Justice League

Alexandra Givens, Center for Democracy and Technology

Anna Tumadottir, Creative Commons

Vinhcent Le, Greenlining Institute

Quinn Anex-Rios, Lawyers Committee for Civil Rights under Law

Amalea Smirniotopoulos, National Association for the Advancement for Colored People

Hannah Sassaman, Peoples Tech Project

Laura Montoya, LatinX in AI

Yang Cheung, Women in AI

Mason Grimshaw, Indigenous in AI

Bhuva Shakti, Women in AI

Arjun Subramonian, Queer in AI

Gelyn Watkins, Black in AI

Erica Fein, International Brotherhood of Electrical Workers

Brett Gibson, American Federation of Labor and Congress of Industrial Organizations

Eric Gottwald, American Federation of Labor and Congress of Industrial Organizations

Faraz Khan, International Federation of Professional and Technical Engineers

Dan Mauer, Communications Workers of America

Michael Wasser, Department of Professional Employees

Rob Weil, American Federation of Teachers

Yoshua Bengio, Université de Montreal

Francesca Rossi, IBM

Stuart Russell, University of California, Berkeley

Sam Gregory, Witness

Abby Kukura, Special Competitive Studies Project (SCSP)

Percy Liang, Stanford University

Joelle Pineau, Meta

Priya Donti, Climate Change AI

Salman Khan, Khan Academy

Julie Kientz, University of Washington

Isaac Kohane, Harvard Medical School and Brigham and Women's Hospital

Aaron Cooper, Business Software Alliance

Keith Kupferschmid, Copyright Alliance

Catherine Stihler, Creative Commons

Jeff Dean, Google

Ece Kamar, Microsoft

Rachel Mandelbaum, Carnegie Mellon University

Kevin Murphy, NASA

Meghan Dierks, Komodo Health

Jon Iwata, Data & Trust Alliance

Yacine Jernite, Hugging Face

Jeffery Smith, Department of Health and Human Services

Rodrigo Ferreira, Rice University

Armando Guio Español, Global Network of Internet & Society Centers, Harvard University

Lucía Trochez Ardila, PIT Policy Lab

César Uribe, Rice University

Deborah Raji, Mozilla

Vincent Conitzer, Carnegie Mellon University

Chris Meserole, Frontier Model Forum

Arvind Narayanan, Princeton University

Julia Angwin, ProofNews

John C. Inglis, Office of the National Cyber Director

Suresh Venkatasubramanian, Brown University

William Isaac, DeepMind

Miranda Bogen, Center for Democracy and Technology

Angela Jiang, DeepMind

Tamara Kneese, Data & Society

Joshua Kroll, Naval Postgraduate School

Madhulika Srikumar, Partnership on AI

Hoda Heidari, Carnegie Mellon University

Yejin Choi, University of Washington

Ifeoma Ajunwa, Emory Law School

Justin Brown, Oklahoma Department of Human Services

Michael Ellison, CodePath

Jean-Paul Chretien, DARPA

Surya Ganguli, Stanford University

Susan Gregurick, NIGMS/NIH

Michael I. Jordan, University of California, Berkeley

James Swanson, Johnson & Johnson

ABOUT NAIAC

The National Artificial Intelligence Advisory Committee (NAIAC) advises the President and the White House National AI Initiative Office (NAIIO) on the intersection of AI and innovation, competition, societal issues, the economy, law, international relations, and other areas that can and will be impacted by AI in the near and long term. Their work guides the U.S. government in leveraging AI in a uniquely American way — one that prioritizes democratic values and civil liberties, while also increasing opportunity.

NAIAC was established in April 2022 by the William M. (Mac) Thornberry National Defense Authorization Act. It first convened in May 2022. It consists of leading experts in AI across a wide range of domains, from industry to academia to civil society. https://www.ai.gov/naiac/

As of April 2024, the NAIAC working group structure is:

Working Group	Chairs	Members
AI Futures – Preparedness, Opportunities, and Competitiveness	Ramayya Krishnan, Haniyeh Mahmoudian	Jack Clark, David Danks, Ashley Llorens, and Swami Sivasubramanian
AI in Work and the Workforce	Amanda Ballantyne, Trooper Sanders	Fred Oswald and Reggie Townsend
AI Education & Awareness	Susan Gonzales, Reggie Townsend	Ayanna Howard and Jon Kleinberg
International Collaboration	Victoria Espinel, Navrina Singh	Christina Montgomery, Keith Strier, and David Danks
Safety, Trust, and Rights	Janet Haven, Daniel Ho, Christina Montgomery	Paula Goldman, Ashley Llorens, and Liz O'Sullivan

The Subcommittee on AI and Law Enforcement working group structure is:

Working Group	Chairs	Members
Processes	Farhang Heydari	Anthony Bak, Jane Bambauer, and Cynthia Rudin
Identification and Surveillance Set	Armando Aguilar	Jennifer Eberhardt, Benji Hutchinson, and Cynthia Rudin
Accountability AI	Jennifer Eberhardt	Armando Aguilar and Farhang Heydari
Officer Training	Benji Hutchinson	Anthony Bak, Jennifer Eberhardt, Farhang Heydari, and Rashawn Ray

###