National Artificial Intelligence Advisory Committee Briefing Minutes October 19, 2023

The National Artificial Intelligence Advisory Committee (NAIAC) held a hybrid in-person/virtual public briefing on Thursday, October 19, 2023. The briefing was recorded and is <u>available</u> <u>online</u>.

NAIAC Members

- Amanda Ballantyne
- Jack Clark
- David Danks
- Victoria Espinel
- Paula Goldman
- Susan Gonzalez
- Janet Haven
- Dan Ho
- Jon Kleinberg
- Ramayya Krishnan
- Ashley Llorens

- Haniyeh Mahmoudian
- James Manyika (Vice-Chair)
- Christina Montgomery
- Liz O'Sullivan
- Fred Oswald
- Trooper Sanders
- Navrina Singh
 - Swami Sivasubramanian
- Keith Strier
- Reggie Townsend
- Miriam Vogel (Chair)

NAIAC-LE Members

- Armondo Aguilar
- Anthony Bak
- Benji Hutchinson

NIST Staff Members

- Melissa Taylor, NIST
- Alicia Chambers, Designated Federal Officer (DFO), NIST

Meeting Minutes

Welcome Remarks

- Chambers called the meeting to order at 10:09 AM Eastern Standard Time and confirmed that the committee is operating under the Federal Advisory Committee Act and accessible to the public both in-person and via livestream. Time was reserved at the conclusion of the meeting for public questions. The Committee also received one written public comment.
- Taylor welcomed NAIAC members and thanked them for their service advising the President and the National AI Initiative Office while also informing the public. Taylor thanked members of the public for their participation and shared the NAIAC <u>email</u> and <u>mailing list</u>.

- Vogel extended warm thanks to NIST staff for their support and thanked NAIAC members for their work advising the White House and President about the risks and opportunities posed by AI technologies. Manyika also thanked NIST staff and NAIAC members for their time and effort.
- NAIAC published its <u>Year 1 Report</u> in May 2023, in fulfillment of its mandate. Taylor provided a brief update on the implementation status of NAIAC Year 1 recommendations. NIST staff will continue to track progress toward the implementation of NAIAC recommendations and provide updates to the Committee.
- In its second year, NAIAC reorganized its working group structure to increase its nimbleness and responsiveness. Working groups have recently published three additional deliverables, which can be accessed on the front page of the NAIAC <u>website</u>.
 - FAQs on Foundation Models and Generative AI: This document summarizes the current landscape of foundation models and generative AI to provide context for future recommendations. The document surveys the distinctive features of foundation models, outlines possible use cases, benefits, and risks, and identifies risk mitigation strategies and technical guardrails.
 - Implementing the NIST AI RMF With a Rights-Respecting Approach: This document connects the implementation of NIST's AI Risk Management Framework (RMF) in the federal government with the federal commitment to protecting civil rights. The document advances three findings: (1) NAIAC's Year 1 report recommended that the federal government adopt the RMF; (2) the RMF calls for AI actors to define their tolerance for AI risks in light of existing laws, values, and norms; (3) civil rights are fundamental to American citizenship, so the federal government must ensure the protection of American civil rights if it implements the RMF.
 - <u>Rationales, Mechanisms, and Challenges to Regulating AI</u>: This explainer summarizes emerging issues in AI regulation, surveying (1) rationales for regulating AI that address the technology's benefits and risks; (2) types of regulatory interventions that have been proposed; and (3) distinct regulatory challenges posed by AI.

Approval of Agenda by Vote

- Vogel outlined the meeting agenda: (1) deliberation and vote on seven working group deliverables; (2) remarks from Secretary of Commerce Gina Raimondo; (3) break; (4) introduction of reorganized working groups and working group agendas; (5) vote on a statement proposed by a working group; (6) update from NAIAC-LE Subcommittee; (7) public comments.
- The NAIAC Chair motioned a vote on the agenda. NAIAC members approved the agenda by voice vote.

Deliberation and Vote on Draft Recommendations and Findings

AI Procurement Working Group Recommendations

Lead: Dan Ho, Procurement of AI Systems WG

- Ho introduced the recommendation, and Sanders and Oswald offered additional context. In 2022, the federal government purchased \$3.3 billion in AI-related goods and over half of the AI used in the federal government was procured from commercial vendors. Procurement is a critical lever for the government to promote the responsible use of AI. It can ensure that the government captures the benefits of AI to better serve Americans and it can influence the broader AI ecosystem by promoting the widespread adoption of trustworthy AI practices. The recommendation, based on a series of briefings with government officials, presents emerging best practices for AI procurement, methods of addressing the AI and technical expertise gap in the federal procurement workforce, and strategies for identifying vendors.
- The NAIAC Chair invited NAIAC members to discuss *AI Procurement Working Group Recommendations*. No comments were forthcoming, so the Chair motioned a vote on the recommendation.
- In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.

• Miriam Vogel - A	Approve
--------------------	---------

- James Manyika Recuse
- Amanda Ballantyne Abstain
- Sayan Chakraborty Not present
- Jack Clark No Response
- David Danks Approve
- Victoria Espinel Recuse
- Paula Goldman Recuse
- Susan Gonzales Approve
- Janet Haven Approve
- Daniel Ho Approve
- Ayanna Howard Not present
- Jon Kleinberg Approve

- Ramayya Krishnan Approve
- Ashley Llorens Recuse
- Haniyeh Mahmoudian Recuse
- Christina Montgomery Recuse
- Liz O'Sullivan Recuse
- Fred Oswald Approve
- Frank Pasquale Not present
- Trooper Sanders Approve
- Navrina Singh Recuse
- Swami Sivasubramanian Recuse
- Keith Strier Not Present
- Reggie Townsend Recuse

Approve: 9 | Recuse = 10 | Abstain = 1 | Not present/no response: 5

Implementing the NIST AI RMF With a Rights-Respecting Approach

Leads: Janet Haven and Liz O'Sullivan, Rights-Respecting AI WG

- Haven introduced the recommendation, which builds on the findings presented earlier. The recommendation presents three steps for implementing the NIST AI RMF in alignment with the protection of civil rights and rights-based governance: (1) federal agencies that work on areas implicating civil rights should define their risk tolerance at an appropriate level to prioritize civil rights protections; (2) federal agencies should define what algorithmic discrimination comprises in their jurisdiction; and (3) federal regulators should establish the technical capacity to address and enforce against AI discrimination and other harms.
- Vogel invited NAIAC members to discuss the recommendation.
 - Vogel and Danks thanked the Rights-Respecting WG for developing a recommendation that demonstrates how rights-based and risk-based approaches to AI risk assessment and governance complement each other.
 - O'Sullivan applauded the recommendation for underscoring the importance of agency decision-making to establish appropriate AI risk tolerance levels. By clarifying how agencies might operationalize a rights-respecting approach to risk tolerance assessment, the recommendation facilitates AI adoption and helps identify circumstances under which AI ought not be used because it constitutes too great a risk to rights.
- The NAIAC Chair motioned a vote on *Implementing the NIST AI RMF With a Rights-Respecting Approach.* In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.

Miriam Vogel - Approve	Ramayya Krishnan - Approve
James Manyika - Approve	Ashley Llorens - Approve
Amanda Ballantyne - Approve	Haniyeh Mahmoudian - Approve
Sayan Chakraborty – Not present	Christina Montgomery - Approve
Jack Clark - Approve	Liz O'Sullivan - Approve
David Danks - Approve	• Fred Oswald – Approve
Victoria Espinel - Approve	• Frank Pasquale - Not present
Paula Goldman - Approve	Trooper Sanders - Approve
Susan Gonzales - Approve	Navrina Singh - Approve
Janet Haven - Approve	Swami Sivasubramanian - Approve
Daniel Ho - Approve	• Keith Strier – Not Present
Ayanna Howard – Not present	Reggie Townsend - Approve
Jon Kleinberg - Approve	

Approve: 21 | Not present/no response: 4

Generative AI Away from the Frontier

Leads: Paula Goldman and David Danks, Generative and NextGen AI: Safety and Assurance WG

- Danks introduced the recommendation by noting that, whereas cutting-edge generative Al systems currently receive significant attention, the risks posed by off-frontier generative Al systems are largely unknown. The recommendation outlines two complementary mechanisms for investigating the risks posed by off-frontier generative Al systems: (1) for systems with constrained access, companies should be encouraged to extend voluntary commitments to support independent testing, risk identification, and information sharing about risks; (2) for systems with unconstrained access, NIST should collaborate with stakeholders across industry, academia, civil society, and government to develop test and analysis environments and other measures to assess potential risks. Analyzing the risks of off-frontier generative AI models will provide the basis for effective, responsible regulation and governance strategies.
- Vogel invited NAIAC members to discuss the two recommendations.
 - Members agreed that the recommendation could more precisely specify the offfrontier models it seeks to include in voluntary commitments by (1) omitting reference to "previous-generation" generative AI systems and (2) proposing that commitments be driven by risk-based assessments.
 - The recommendation proposes distinct approaches to analyzing the risks of constrained- and unconstrained-access systems. These approaches reflect the distinct types of assessment afforded by system accessibility: *multistakeholder approaches* are well-suited to the wide accessibility of unconstrained systems; *voluntary testing commitments* for less accessible constrained systems leverage a mechanism already employed by the Biden Administration to guide safe corporate development of AI.
 - Danks edited the recommendation text to indicate that the multistakeholder approach for assessing unconstrained-access systems will be used to develop tools for testing all generative AI systems, not just open-source generative AI systems.
 - The Committee will consider developing future recommendations that propose benchmarking and oversight methods to track companies' compliance with the voluntary commitments secured by the Biden Administration.
- Kleinberg motioned a vote on the edited text of *Generative AI Away from the Frontier*. In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.
 - Miriam Vogel Approve
 - James Manyika App. w/ edits
 - Amanda Ballantyne App. w/ edits
 - Sayan Chakraborty Not present
 - Jack Clark App. w/ edits

- David Danks Approve
- Victoria Espinel Approve
- Paula Goldman Approve
- Susan Gonzales Approve
- Janet Haven Approve

- Daniel Ho Approve
- Ayanna Howard Not present
- Jon Kleinberg App. w/ edits
- Ramayya Krishnan Approve
- Ashley Llorens App. w/ edits
- Haniyeh Mahmoudian Approve
- Christina Montgomery App. w/ edits
- Liz O'Sullivan Approve

- Fred Oswald Approve
- Frank Pasquale Not present
- Trooper Sanders Approve
- Navrina Singh App. w/ edits
- Swami Sivasubramanian Recuse
- Keith Strier Not Present
- Reggie Townsend Reject

Approve: 12 | Approve with edits: 7 | Recuse: 1 | Reject: 1 | Not present/no response: 4

National Campaign on Lifelong AI Career Success

Lead: Trooper Sanders, AI in Work and the Workforce WG

- Sanders introduced the recommendation, which aims to address barriers to entering the AI/tech workforce faced by later-in-life workers, including caregivers, veterans, and military spouses. The recommendation proposes four measures: (1) correcting misperceptions about the suitability of later-in-life workers for high-tech jobs; (2) promoting the value of later-in-life workers for national economic success; (3) targeting outreach to later-in-life workers to provide skills-based transition support; and (4) generating commitments from employers to target, hire, and train later-in-life workers.
- Vogel invited NAIAC members to discuss the recommendation.
 - Members observed that expanding the AI workforce will require engagement with structural social challenges such as age and gender discrimination.
 - Members emphasized that later-in-life workers contribute valuable experience to the American workforce; members highlighted the need to make skill development and job readiness efforts more age-inclusive by expanding them beyond the traditional education pipeline.
 - Members proposed modifying the recommendation to emphasize regional inclusiveness, outreach to underrepresented communities, and transitions to highquality jobs that promote worker dignity rather than precarious or algorithmicallymanaged work.
 - Members suggested augmenting the recommendation with examples of existing job transition infrastructure (e.g., Department of Labor Workforce Hubs) that might be leveraged to operationalize the recommendation.
- Sanders will edit the text of the recommendation to include the four suggested modifications on underrepresented communities, regional inclusiveness, existing workforce transition infrastructure, and job quality and dignity. NAIAC members will vote on the edited recommendation at the next NAIAC public meeting.

Second Chance Skills and Opportunity Moonshot

Lead: Trooper Sanders, AI in Work and the Workforce WG

- Sanders introduced the recommendation, which notes that tens of millions of Americans are excluded from new AI-related career opportunities because they lack basic literacy, numeracy, and digital problem-solving skills. The recommendation calls for a national goal to make the United States a top-performing country in basic literacy, numeracy, and digital problem-solving skills. The recommendation proposes an all-of-society approach to this effort, expanding beyond traditional education to establish pathways for basic skill acquisition at all stages of life. The recommendation also calls for the development of measures to evaluate progress toward this goal.
- Vogel invited NAIAC members to discuss the recommendation.
 - Oswald noted that the recommendation dovetails with the Biden Administration's National Cyber Workforce and Education Strategy: basic skills provide a crucial foundation for the development of cyber and technical skills.
- Vogel motioned a vote on *Second Chance Skills and Opportunity Moonshot*. In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.
 - Miriam Vogel Approve
 - James Manyika Approve
 - Amanda Ballantyne Approve
 - Sayan Chakraborty Not present
 - Jack Clark Approve
 - David Danks Approve
 - Victoria Espinel Approve
 - Paula Goldman Approve
 - Susan Gonzales Approve
 - Janet Haven Approve
 - Daniel Ho Approve
 - Ayanna Howard Not present
 - Jon Kleinberg Approve

- Ramayya Krishnan Approve
- Ashley Llorens Approve
- Haniyeh Mahmoudian Approve
- Christina Montgomery Approve
- Liz O'Sullivan Approve
- Fred Oswald Approve
- Frank Pasquale Not present
- Trooper Sanders Approve
- Navrina Singh Approve
- Swami Sivasubramanian Approve
- Keith Strier Approve
- Reggie Townsend Approve

Approve: 22 | Not present/no response: 3

Institutional Structures to Create Safer AI Systems

Leads: Ramayya Krishnan and Haniyeh Mahmoudian, AI Futures: Sustaining Innovation in Next Gen AI WG

- Krishnan introduced the recommendation with the observation that AI safety and reliability are necessary to engender trust and widespread AI adoption and deployment. Cybersecurity initiatives such as the Computer Emergency Response Team (CERT) have demonstrated the efficacy of establishing institutional structures to manage and support the post-deployment tracking and resolution of software system vulnerabilities. The recommendation suggests a similar institutional infrastructure for AI to catalog incidents, test and verify models, and identify best practices, as well as create actionable response frameworks for AI threats and vulnerabilities that affect .mil, .gov, and .com domains. The recommendation suggests that this multi-agency AI Lead Response Team (ALRT) be housed within an existing federally-funded R&D center (FFRDC) and convene industry, government, academic, and civil society partners.
- Vogel invited NAIAC members to discuss the recommendation.
 - Members asked Krishnan to clarify how the ALRT's function differs from that of existing efforts to coordinate government technology policy and monitor AI incidents. Krishnan clarified that the ALRT would fill the need for a government institution to track and mitigate *post-deployment* AI incidents, failures, and vulnerabilities across .gov, .com, and .mil domains and share outcomes with AI developers and deployers.
 - Members suggested that the recommendation allow for a broader range of implementation pathways by (1) removing the requirement that the ALRT be established within an existing FFRDC and (2) including additional organizations that might contribute expertise to the ALRT, including university-affiliated research centers and new FFRDCs. The recommendation was edited to reflect these suggestions.
 - Members noted the breadth of expertise required to address technical, sociotechnical, and application-based risks and questioned whether this work should be conducted in a single, centralized institution or distributed across a federated set of institutions, each contributing distinct domain expertise. Llorens suggested that further stakeholder engagement was needed to inform the specification of the institutional structure.
 - Ho observed that the recommendation calls for *voluntary* reporting of adverse events but suggested that NAIAC consider future recommendations to *mandate* adverse event reporting. In cases of information asymmetry between private actors and the government, mandated reporting of adverse events can support effective regulatory action.
 - Townsend asked Vogel and Manyika to clarify whether recommendations should primarily signal goals or whether they should specify implementation plans as well.
 Vogel and Manyika noted that, because the scope of each recommendation will vary, their intended use and impact should be clearly stated. The committee will

need to effect a balance between outlining goals and proposing implementation details.

• Vogel motioned a vote on the edited text of *Institutional Structures to Create Safer AI systems.* In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.

Approve: 12 | Approve with edits: 8 | Not present/no response: 4

Findings on the Potential Future Risks of AI

Leads: Ramayya Krishnan and Haniyeh Mahmoudian, AI Futures: Sustaining Innovation in Next Gen AI WG

- Mahmoudian introduced the finding, which synthesizes public briefings by experts on AI risk. Experts outlined three primary risk categories: malicious objectives and unintended consequences, economic and societal risk, and catastrophic risk. The finding also notes ongoing debate among experts about the priority and likelihood of various types of risk. Finally, the finding records experts' proposed risk mitigation strategies, including adaptive regulation, research investment, multistakeholder approaches, and international coordination.
- Vogel invited NAIAC members to discuss the recommendation.
 - Members made suggestions about how the finding should attribute the expert perspectives it records. Mahmoudian will edit the text of the finding to reflect these suggestions: (1) a footnote citing OpenAI CEO Sam Altman's testimony before the Senate Judiciary Committee will be omitted so as not to suggest particular endorsement of one business leader's perspective; (2) normative statements in the

"Potential Solutions and Mitigation Strategies" section will be clearly labelled as expert suggestions rather than formal NAIAC recommendations; and (3) the introduction will state that the finding derives from the public hearing on AI risks with three invited experts.

- Members appreciated that the finding's risk mitigation strategies include adaptive regulation, as this measure is critical to addressing the rapid pace of technological change.
- Townsend asked whether the section on economic and societal risk might address law enforcement topics. Krishnan and Mahmoudian noted that the finding reports primarily on AI risks identified during the group's public briefings, which did not address law enforcement topics.
- Vogel motioned a vote on the edited text of *Findings on the Future Risks of AI*. In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.

Miriam Vogel – App. w/ edits	Ramayya Krishnan - Approve
James Manyika – Approve	• Ashley Llorens – App. w/ edits
Amanda Ballantyne – Not present	Haniyeh Mahmoudian – Approve
Sayan Chakraborty – Not present	• Christina Montgomery – App. w/ edi
Jack Clark – Approve	Liz O'Sullivan - Approve
David Danks – Approve	• Fred Oswald – Approve
Victoria Espinel – App. w/ edits	• Frank Pasquale - Not present
Paula Goldman - Approve	• Trooper Sanders - Approve
Susan Gonzales - Approve	Navrina Singh – Approve
Janet Haven – App. w/ edits	• Swami Sivasubramanian – App. w/
Daniel Ho - Approve	edits
Ayanna Howard – Not present	• Keith Strier – App. w/ edits
Jon Kleinberg – Approve	• Reggie Townsend – App. w/ edits

Approve: 13 | Approve with edits: 8 | Not present/no response: 4

Remarks from Secretary of Commerce Raimondo

- Vogel introduced Secretary of Commerce Raimondo, emphasizing her many roles as a businesswoman, lawyer, former governor of Rhode Island, and now champion of work on artificial intelligence within the Cabinet.
- Raimondo thanked NAIAC and NAIAC-LE members for their service as the speed of technological innovation increases the importance of work on AI. Raimondo outlined two simultaneous goals: (1) to ensure that the United States continues to lead the world in AI innovation and (2) to manage the risks of AI development. Meeting those goals will require coordination with global allies and industry partners to develop guiding principles and

policies that promote trust, transparency, and accountability. No country has yet instituted AI legislation. Raimondo highlighted the opportunity to develop policies, principles, risk frameworks, regulations, and statutes that enable AI innovation, protect privacy and civil rights, and support a wide range of stakeholders, including the public, workers, consumers, and students. NAIAC's work and expertise will be central to these efforts.

Law Enforcement Subcommittee Roadmap Presentation

- NAIAC-LE will make recommendations to NAIAC about the legal, ethical, and responsible use of AI technologies in law enforcement contexts. The NAIAC-LE Year 1 Roadmap outlines the Subcommittee's scope, organizing principles, and workplan.
- The Roadmap outlines the Subcommittee's two primary workstreams: (1) **law enforcement uses of AI** (10 use cases identified) and (2) **legal and ethical issues** raised by law enforcement applications of AI (13 legal/ethical issues organized into 3 categories). These workstreams intersect: each legal and ethical issue pertains to each use case. In addition, the Subcommittee plans to conduct a special technology study on facial recognition.
- The Subcommittee is forming four initial WGs to address an initial set of priority topics drawn from its use case and legal/ethical taxonomies. Two WGs focus on use cases and two focus on legal/ethical issues:
 - **Use Case WG #1**: Person identification algorithms, facial recognition, AI-assisted surveillance.
 - **Use Case WG #2**: Dot-connecting methods, resource allocation decisions, accountability algorithms, robotics.
 - **Legal/Ethical WG #1**: Performance evaluation and monitoring, bias, potential risks and harms of AI underutilization.
 - **Legal/Ethical WG #2**: Transparency and public input, pre-adoption process, postadoption procedures and audits, impact on law enforcement employees.
- NAIAC-LE will finalize its Year 1 Roadmap during November 2023 and develop its first recommendations for NAIAC during Q1 of 2024.

Questions and Comments

- NAIAC members asked whether certain topics were part of NAIAC-LE's intended scope.
 - Haven asked whether the study on facial recognition would include recommendations for government action in addition to descriptions of the technology's law enforcement applications and the legal/ethical concerns it raises. NAIAC-LE members noted that the Subcommittee must begin work on this project before it can provide a definitive account of intended deliverables. NAIAC-LE members are also mindful of the quantity of existing research on facial recognition and hope to develop novel work that advances research on this topic.
 - Haven noted that NAIAC-LE need not restrict its recommendations to those that advance the research frontier. She emphasized the significant value of sharing evidence-based recommendations grounded in NAIAC-LE's expert synthesis of existing work.
 - Kleinberg asked whether NAIAC-LE plans to study AI explainability—the provision of explanations for algorithmic decision-making—in law enforcement. Bak noted that

this topic will be addressed as it arises in use case analyses. However, the Subcommittee will consider explicitly addressing it as a distinct legal/ethical issue.

- O'Sullivan asked that the robotics use case include analysis of the application of force using robotics.
- Townsend asked whether the Subcommittee's exclusion of administrative law enforcement from its scope implied exclusion of algorithmic applications in judicial processes. Bak clarified that the Subcommittee would address algorithms used for judicial processes such as bail, sentencing, and pretrial detention. The exclusion of administrative law enforcement applies instead to cases such as Environmental Protection Agency enforcement.
- Townsend asked whether the Subcommittee would address individual consent for data use. Bak confirmed that the Subcommittee would address this under its data privacy workstream.
- Manyika asked whether the Subcommittee would address demographic disparities in law enforcement data collection, noting the biased predictions that result from such disparities. Bak confirmed that NAIAC-LE will address this issue in two intersecting ways: (1) each use case analysis will engage with data collection and (2) the Subcommittee will pursue a legal/ethical workstream on bias.
- Manyika asked whether the Subcommittee will address procurement of AI for law enforcement, observing that procurement methods have been implicated in failures of algorithmically driven law enforcement. Bak confirmed that the Subcommittee would address several procurement-related policies and procedures including acceptable use policies, trial periods, and experimental design to validate assessments of use.
- Llorens noted the potential for intersection between the work of NAIAC and that of NAIAC-LE, highlighting workplace considerations (e.g., performance monitoring) as one prominent example. Bak agreed, noting that NAIAC-LE has a workstream on the law enforcement workforce and that Ballantyne serves as the liaison between NAIAC and NAIAC-LE. Vogel emphasized the importance of frequent, consistent meetings in fostering collaboration between the Committee and Subcommittee.

Current Working Groups

- NAIAC has reorganized into five WGs, consolidating its workstreams to facilitate member engagement.
- Due to time constraints, comprehensive introductions of the reorganized WGs were deferred until the following NAIAC public meeting. However, Vogel shared the titles of the five current WGs: (1) AI Futures: Preparedness, Opportunities, and Competitiveness; (2) AI in Work and the Workforce; (3) Education/Awareness; (4) International Collaboration; (5) Safety, Trust, and Rights.

Public Questions and Comments

• A member of the public referred NAIAC members to an <u>article</u> by Mustafa Suleyman and Eric Schmidt that calls for an AI equivalent of the Intergovernmental Panel on Climate Change (IPCC) to develop an international analytical effort to examine the effects of AI.

Trust, Safety, & Rights WG Statement on AI Risk

Vote 1

- Ho introduced a statement on AI risk proposed by the Safety, Trust, and Rights WG: "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."
- Ho explained that this statement contributes to current debate about how AI regulation should prioritize and address AI risks. The statement counters the view that AI regulation should focus exclusively on existential risk.
- Ho motioned for a vote on the preceding statement. In the presence of quorum, each Member of the Committee was polled and decided by majority vote to advance the recommendation.
 - Miriam Vogel Approve
 - James Manyika Approve
 - Amanda Ballantyne Not present
 - Sayan Chakraborty Not present
 - Jack Clark Approve
 - David Danks Approve
 - Victoria Espinel Approve
 - Paula Goldman Approve
 - Susan Gonzales Approve
 - Janet Haven Approve
 - Daniel Ho Approve
 - Ayanna Howard Not present
 - Jon Kleinberg Approve

- Ramayya Krishnan Approve
- Ashley Llorens Approve
- Haniyeh Mahmoudian Approve
- Christina Montgomery Approve
- Liz O'Sullivan Reject
- Fred Oswald Approve
- Frank Pasquale Not present
- Trooper Sanders Approve
- Navrina Singh Approve
- Swami Sivasubramanian Approve
- Keith Strier Approve
- Reggie Townsend Approve

Approve: 20 | Reject: 1 | Not present/no response: 4

Vote 2 (to supersede Vote 1)

- O'Sullivan introduced a proposal to supersede the prior statement with a revised statement: "Arguments about existential risk should not detract from the necessity of addressing existing risks of AI."
- O'Sullivan explained that the Safety, Trust, and Rights WG unanimously approved only this first half of the statement. O'Sullivan asserted that ongoing debates about the appropriate prioritization of AI risks focus solely on *risk*; discussing AI benefits in this context is superfluous.
- Vogel motioned a vote to supersede the prior statement with the revised statement. In the presence of quorum, each Member of the Committee was polled and decided by majority vote to reject the proposal, letting the result of the first vote state.

- Miriam Vogel Reject
- James Manyika Reject
- Amanda Ballantyne Approve
- Sayan Chakraborty Not present
- Jack Clark Reject
- David Danks Reject
- Victoria Espinel Reject
- Paula Goldman Reject
- Susan Gonzales Reject
- Janet Haven Approve
- Daniel Ho Reject
- Ayanna Howard Not present
- Jon Kleinberg Approve

- Ramayya Krishnan Reject
- Ashley Llorens Reject
- Haniyeh Mahmoudian Reject
- Christina Montgomery Reject
- Liz O'Sullivan Approve
- Fred Oswald Reject
- Frank Pasquale Not present
- Trooper Sanders Reject
- Navrina Singh Approve
- Swami Sivasubramanian Reject
- Keith Strier Reject
- Reggie Townsend Reject

Approve: 5 | Reject: 17 | Not present/no response: 3

Closing Remarks

- Vogel thanked Committee members for their constructive discussion and applauded the progress made during the meeting.
- Members of the public are encouraged to share comments and questions with NAIAC and NAIAC-LE by emailing <u>NAIAC@nist.gov</u>. They can visit <u>ai.gov/naiac</u> to subscribe for Committee updates. A summary of the meeting will also be posted on <u>ai.gov/naiac</u>.
- Chambers adjourned the meeting at 1:43 PM Eastern Time.