Written Testimony of Terah Lyons
Executive Director
The Partnership on Artificial Intelligence to Benefit People & Society

House of Representatives Oversight & Government Reform Committee
Subcommittee on Information Technology
Hearing on
Game Changers: Artificial Intelligence Part III – AI and Public Policy

April 18, 2018

Chairman Hurd, Ranking Member Kelly, and Members of the Subcommittee, thank you for the opportunity to discuss a unique and unprecedented multi-stakeholder body—the Partnership on Artificial Intelligence to Benefit People & Society (PAI, or the Partnership)—and to provide PAI's perspective on the important issues raised by artificial intelligence (AI). We appreciate the Subcommittee's interest in the potential challenges and opportunities presented by artificial intelligence, the possible barriers to effective development and deployment of artificial intelligence, and the possible solutions in response to these challenges and barriers. The Partnership tackles both substantive AI and AI governance challenges in its mission to resolve such questions.

Like you, we believe that AI is and will be beneficial to people and society—providing valuable services, improving efficiency for citizens, companies, and governments, helping people with a wide range of individual goals, and even saving lives. We also recognize that AI invites important questions, such as the potential impact on our workforce; the possibility of entrenching bias; and the ramifications for privacy. One thing is clear—if we allow a gulf to open up between the entities developing AI technologies and the people and societies that are impacted by them, we will both limit progress and cause harm in the process. Every organization involved in the Partnership on AI has accepted that premise as part of their membership, and each organization is committed to maximizing the benefits of AI while working to address its potential challenges.

In fact, our early work spans several of your priority issues. PAI's initial Working Groups—the main bodies through which PAI examines AI best practices, marshals resources toward high-priority AI research, and identifies aspirational pro-social AI projects—are focused on the labor impacts of automation; fairness, accountability, and transparency of AI systems; and the ethical and effective application of AI in safety-critical scenarios.

You are also interested in exploring how government and other agents can most effectively address those challenges. We believe that PAI's multi-stakeholder approach is vital in delivering positive outcomes in technology development and technology governance. We must

work together to find a path forward in understanding and delivering solutions to the complex challenges associated with AI development—and in maximizing the benefits of AI—by leveraging diverse interests, expertise, and perspectives. Congress' support for similarly diverse and nuanced conversations will be important to this work, now, and in the future.

We would like to be clear that PAI is not a lobbying organization. PAI does, however, intend to be a *resource* to policymakers—for instance, in conducting research that informs AI best practices and exploring the societal consequences of certain AI systems, as well as policies around the development and usage of those systems. Toward those ends, PAI is fortunate to have a diverse membership of experts which spans sectors and fields – AI researchers and engineers, companies, academics from a wide range of disciplines, advocates for interests related to human rights, civil liberties, and other leaders from civil society. It is a deep and growing bench, with unique perspectives and an emphasis on equity informed by its equal representation of for-profit and not-for-profit members on our Board of Directors. We believe that our diversity is vital to the success, sustainability, and durability of the organization and its mission, and critical to effective and inclusive multi-stakeholder engagement.

My testimony today will discuss:

- **The promise of AI**
- **The formation of PAI and the need it serves**
- **PAI's four goals**
- **PAI's early work**
- **What is next for PAI and AI governance**

The Promise of AI

Artificial intelligence technologies present a significant opportunity for the United States and for the world to address some of humanity's most pressing and large-scale challenges, to generate economic growth and prosperity, and to raise the quality of human life everywhere.

While we might be years or decades away from fully realizing that opportunity, it is important to recognize that AI technologies are already solving important challenges today. Researchers are building systems that show increasing abilities to conduct pattern recognition in healthcare, including augmenting human experts in interpreting radiologic studies and detecting cancers from images at human expert levels, with the increasing availability of large datasets from which algorithms can learn. AI advances also show promise at providing new efficiencies and optimizations, in applications ranging from the products we purchase and the transportation technology and infrastructure we use, to the ways in which we consume energy. AI technologies are also leading to enhanced understanding of global systems at scale through the collection and analysis of data about global health, the environment, and more.

There are many similar examples about how AI is being used in beneficial ways today with the promise of even more advancements in the future. To illustrate the trajectory of some of these innovations, AI has the potential to be a new economic engine to drive growth, opportunity, and prosperity. Over the last decade, a global explosion of smartphones have

become the infrastructure for a new app economy, catalyzing new businesses, both big and small. The next decade could see the creation of an AI economy, ushered in by a new wave of efficiency and optimization. Healthcare, nutrition, environmental conservation, economic inclusion, and accessibility and mobility, among other domains, all have the potential to see radical shifts for the positive as a result of AI.

In order to ensure a strong future where we harness the latent value of AI methods, we must begin thinking carefully, starting now, about the challenges of development and implementation posed by these technologies and how we can address them to ensure a future that benefits us all.

The Formation of PAI and the Need it Serves

With the promises and risks of AI in mind, PAI was established to study and formulate best practices on AI technologies, to advance the public's understanding of AI, and to serve as an open platform for discussion and engagement about AI and its influences on people and society.

We are at a critical juncture for AI's development and its applications. The development of AI technologies is now intersecting with the application of these systems and techniques. Advancements in algorithms, computing power, and rich data, are helping the field to solve technical challenges that will improve domains such as perception, natural language understanding, and robotics, bringing great value in the years ahead.

However, with technological advancement comes concerns and challenges associated with responsible development and the impacts of technologies on people's lives. These concerns include the safety of AI technologies, the fairness and transparency of systems, and the intentional as well as inadvertent influences of AI on people and society, including potential influences on privacy, democracy, criminal justice, and human rights. We must also be mindful of concerns associated with the potential unintended harms that "AI-for-good" applied in some circumstances and manners may impose, and must work to ensure that good intentions in technology application and deployment translate into positive outcomes for the people and communities they impact.

The Partnership's formation was first rooted in leading technologists collectively understanding the need for AI best practices, dialogue, and common understanding related to issues such as these. From the outset, the Partnership was designed as a multistakeholder entity, and from its founding has grown to include a community of technology companies, academic institutions, civil society organizations and other types of not-for-profit institutions, and increasingly global perspectives. Diversity of thought is critical to PAI's success, as is considering the impact of AI on diverse constituencies. An organization comprised of for-profit and not-for-profit members has more nuanced perspective to draw upon in identifying and effectively addressing these challenges. To that end, we aspire to an even more representative community than PAI currently reflects, with engagement from even more diverse interests than currently represented, specific application areas and sectors not currently involved in our Partner

community, and underrepresented interests that may give voice to constituencies needed but not often heard in matters of technology governance.

The Partnership's goal is to involve deep subject-matter experts, across a variety of disciplines, in our work. PAI also aims to focus both on identifying and developing issues and topic areas ripe for consensus, as well as grappling with more challenging and nuanced questions associated with the development and deployment of AI. Through this issue identification, and through enumerating, interrogating, and leveraging the varying perspectives and interests of our multi-stakeholder body, it is the Partnership's aspiration to incite critical reflection about the impacts of AI on people and society, and to constructively shape practices, norms, expectations, and behavior related to responsible AI development.

The Partnership's activities are deliberately determined by its Partners, but from the outset of the organization, the hope and intention has been to create a place for open critique and reflection. Crucially, the Partnership is an independent organization; though supported and shaped by our Partner community, the Partnership is ultimately more than the sum of its parts, and will make independent determinations to which its Partners will collectively contribute, but never individually dictate.

It is our intention that PAI will be collaborative, constructive, and work openly with all, both those who find themselves aligning with recommendations and guidance on best practices and those who differ in assessments—or who are skeptical of our work. Crucially, PAI has been explicitly designed to bring together varying perspectives in a structure that ensures balanced governance by diverse stakeholders. In the same manner that PAI welcomes and encourages thoughtful debate among its members, PAI also welcomes debate among society on the best ways to engage these topics.

PAI's Goals

The formation of PAI was based on a clear set of goals and overarching Tenets (see Appendix). Based on these goals and founding principles, substantive Thematic Pillars were developed to guide the organization's work, in pursuit of the following outcomes:

- ***Develop and Share Best Practices***

  Support research, discussions, identification, sharing, and recommendation of best practices in the research, development, testing, and fielding of AI technologies. Address such areas as fairness and inclusivity, explanation and transparency, security and privacy, values and ethics, collaboration between people and AI systems, interoperability of systems, and of the trustworthiness, reliability, containment, safety, and robustness of the technology.

- ***Provide an Open and Inclusive Platform for Discussion and Engagement***

  Create and support opportunities for AI researchers and key stakeholders, including people in technology, law, policy, government, civil liberties, and the greater public, to communicate directly and openly with each other about relevant issues to AI and its influences

on people and society. Ensure that key stakeholders have the knowledge, resources, and overall capacity to participate fully.

- ***Advance Public Understanding***

Advance public understanding and awareness of AI by multiple constituencies, including writing and other communications on core technologies, potential benefits, and costs. Act as a trusted and expert point of contact as questions, concerns, and aspirations arise from the public and others in the area of AI. Regularly update key constituents on the current state of AI progress.

- ***Identify and Foster Aspirational Efforts in AI for Socially Beneficial Purposes***

Seek out, support, celebrate, and highlight aspirational efforts in AI for socially benevolent applications. Identify areas of untapped opportunity, including promising technologies and applications not being explored by academia and industry R&D.

We believe that artificial intelligence technologies hold great promise for raising the quality of people's lives and can be leveraged to help humanity address important global challenges such as climate change, food, inequality, health, and education. Our mission is in service of unlocking this potential through intentional, inclusive, multidisciplinary analysis, design, and accountability.

PAI's Early Work

PAI's efforts and early-stage programs are currently organized around a set of Thematic Pillars reflecting key challenges and opportunities in AI:

- ***AI Labor, and the Economy***

AI advances will undoubtedly have multiple influences on the distribution of jobs and nature of work. While advances promise to inject great value into the economy, they can also be the source of disruptions as new kinds of work are created and other types of work become less needed due to automation.

Discussions are rising on the best approaches to minimizing potential disruptions, making sure that the fruits of AI advances are widely shared and competition and innovation is encouraged and not stifled.

- ***Safety-Critical AI***

Advances in AI have the potential to improve outcomes, enhance quality, and reduce costs in such safety-critical areas as healthcare and transportation. Effective and careful applications of pattern recognition, automated decision making, and robotic systems show promise for enhancing the quality of life and preventing thousands of needless deaths.

However, where AI tools are used to supplement or replace human decision-making, we must be sure that they are safe, trustworthy, and aligned with the ethics and preferences of people who are influenced by their actions.

- ***Fair, Transparent, and Accountable AI***

AI has the potential to provide societal value by recognizing patterns and drawing inferences from rich amounts of data. Data can be harnessed to develop useful diagnostic systems and recommendation engines, and to support people in making breakthroughs in such areas as biomedicine, public health, safety, criminal justice, education, and sustainability.

While such results promise to provide great value, we need to be sensitive to the hidden assumptions and biases in data as well as the biases reflected in decisions about system design. This can lead to actions and recommendations that replicate those biases, and suffer from serious blind spots. Such failings can lead to the unfair and systematic exclusion of groups of people from consequential resources and services, such as in financial services, job opportunities, and education.

Researchers, officials, and the public should be sensitive to these possibilities and we should seek to develop methods that detect and correct those errors and biases, not replicate them. We also need to work to develop systems that can explain the rationale for inferences.

- ***Collaborations Between Humans and AI Systems***

A promising area of AI is the design of systems that augment the perception, cognition, and problem-solving abilities of people. Examples include the use of AI technologies to help physicians make more timely and accurate diagnoses and assistance provided to drivers of cars to help them to avoid dangerous situations and crashes.

Opportunities for R&D and for the development of best practices on AI-human collaboration include methods that provide people with clarity about the understandings and confidence that AI systems have about situations, means for coordinating human and AI contributions to problem solving, and enabling AI systems to work with people to resolve uncertainties about human goals.

- ***Social and Societal Influences of AI***

AI advances will touch people and society in numerous ways, including potential influences on privacy, democracy, criminal justice, and human rights. For example, while technologies that personalize information and that assist people with recommendations can provide people with valuable assistance, they could also inadvertently or deliberately manipulate people and influence opinions.

We seek to promote thoughtful collaboration and open dialogue about the potential subtle and salient influences of AI on people and society.

- ***AI and Social Good***

AI offers great potential for promoting the public good, for example in the realms of education, housing, public health, and sustainability. We see great value in collaborating with public and private organizations, including academia, scientific societies, NGOs, social entrepreneurs, and interested private citizens to promote discussions and catalyze efforts to address society's most pressing challenges.

Some of these projects may address deep societal challenges and will be moonshots – ambitious big bets that could have far-reaching impacts. Others may be creative ideas that could quickly produce positive results by harnessing AI advances.
Beyond the specified thematic pillars, we also seek to convene and support projects that resonate with the tenets of our organization. We are particularly interested in supporting people and organizations that can benefit from the Partnership's diverse range of stakeholders, and are welcoming input from stakeholders regarding what work would be most valuable for an entity like PAI to undertake. We are open-minded about the forms that these efforts will take.

Working Groups

The organization's earliest programming work has established Working Groups associated with three of the above Thematic Pillar areas: AI, Labor, and the Economy; Safety-Critical AI; and Fair, Transparent, and Accountable AI. Through these structures, Partners opt into Working Groups addressing each Pillar. Co-Chairs—one each representing a for-profit and from a not-for-profit Partner—lead each Working Group. This diverse leadership helps to ensure that the Partnership empowers and draws upon the full breadth of its perspectives, and is designed to help ensure that all members have a seat at the table. Over time, we will launch additional Working Groups focused on other areas of work. And as the work of these sub-communities develops, we look forward to updating the public as to PAI's organizational priorities, projects, and work products.

Providing an Open and Inclusive Platform for Engagement

From the founding of the Partnership on AI, we have been optimistic about the high potential of our Working Groups to generate meaningful insights and to incite cross-sector, multidisciplinary reflection about opportunities and challenges in the AI field. As an organization, the Partnership believes that transparency, diversity, and inclusivity are essential in advancing this work. From the outset, we intend to create a culture where Working Group members listen closely to those with whom they *disagree*—understanding that a central value of the Partnership is in our capacity to learn from differing interests, perspectives, and areas of expertise. Together, we will work to leverage the power of our diversity to create an environment which promotes meaningful collective reflection, and produces impactful decisions and outputs. The project of maintaining the strength and integrity of our community will rely on commitment to these ideals from across our organization.

The Partnership is dedicated to creating and supporting opportunities for AI researchers and key stakeholders, including people working in technology, law, policy, government, advocacy—and the broader public—to communicate directly and openly with each other about relevant issues to AI and its influences on people and society. The Partnership is also dedicated to ensuring that its stakeholders have the knowledge, resources, and overall capacity to participate fully. To ensure that valued perspectives are never marginalized, the Partnership is in the midst of developing a full program of support and capacity-building for civil society organizations (CSOs), which face unique challenges in supporting the type of work, often voluntary, required of meaningful participation in a multistakeholder organization like the Partnership. Though the organization is in early stages of developing this program area, the

Partnership is committed to fostering a sustainable multistakeholder model which enables the full participation of interested Partners.

As citizens, communities, and businesses across the U.S. and around the world grapple with important challenges associated with the development of AI technologies, we would invite members of Congress and the broader public to engage with PAI to ensure we are focusing our attention and the attention of our Partners, accordingly.

What is Next for PAI and AI Governance

The work of the Partnership is motivated by the belief that because the societal implications of AI development and deployment are complex and multifaceted we cannot determine solutions alone—neither as singular entities, or as a singular discipline. As the AI field and multidisciplinary understandings of its impact evolve over time, the Partnership will be a place where stakeholders remain engaged on these topics, and work to update, refine, and advance perspectives and solutions to associated challenges.

The Partnership will be offering perspectives on topics like those related to our Thematic Pillars and current and future Working Groups. We are committed to supporting the following immediate areas of work:

- **Engagement of Experts**

The regular engagement of experts across multiple disciplines (including but not limited to psychology, philosophy, economics, finance, sociology, public policy, and law) to discuss and provide guidance on emerging issues related to the impact of AI on society.

- **Engagement of Other Stakeholders**

The engagement of AI users and developers, as well as representatives of industry sectors that may be impacted by AI (such as healthcare, financial services, transportation, commerce, manufacturing, telecommunications, and media) to support best practices in the research, development, and use of AI technology within specific domains.

- **Third-Party Support**

The design, execution, and financial support of objective third-party studies on best practices for the ethics, safety, fairness, inclusiveness, trust, and robustness of AI research, applications, and services. The identification and celebration of important work in these fields. The support of aspirational projects in AI that would greatly benefit people and society.

- **Informational Material Development**

The development of informational materials on the current and future likely trajectories of research and development in core AI and related disciplines.

Conclusion

We are excited about the future, and recognize that the Partnership is an ambitious project with high expectations both from its members and the greater public. With the launch of

our first Working Groups, we are getting to the pressing work at hand and look forward to making a valuable and lasting contribution to advancing trusted, ethical, inclusive, multidisciplinary, and beneficial AI.

Thank you for this opportunity to speak about the Partnership on AI, its goals, and its efforts. We look forward to being a resource for Congress, other global policymaking bodies, and the public, to help ensure that the benefits of AI are realized and the challenges are addressed as PAI pursues its mission.

ADDITIONAL MATERIALS

Tenets of the Partnership on AI

Members of the Partnership on AI share the following tenets:

1. We will seek to ensure that AI technologies benefit and empower as many people as possible.

2. We will educate and listen to the public and actively engage stakeholders to seek their feedback on our focus, inform them of our work, and address their questions.

3. We are committed to open research and dialogue on the ethical, social, economic, and legal implications of AI.

4. We believe that AI research and development efforts need to be actively engaged with and accountable to a broad range of stakeholders.

5. We will engage with and have representation from stakeholders in the business community to help ensure that domain-specific concerns and opportunities are understood and addressed.

6. We will work to maximize the benefits and address the potential challenges of AI technologies, by:

    a. Working to protect the privacy and security of individuals.

    b. Striving to understand and respect the interests of all parties that may be impacted by AI advances.

    c. Working to ensure that AI research and engineering communities remain socially responsible, sensitive, and engaged directly with the potential influences of AI technologies on wider society.

    d. Ensuring that AI research and technology is robust, reliable, trustworthy, and operates within secure constraints.

    e. Opposing development and use of AI technologies that would violate international conventions or human rights, and promoting safeguards and technologies that do no harm.

7. We believe that it is important for the operation of AI systems to be understandable and interpretable by people, for purposes of explaining the technology.

8. We strive to create a culture of cooperation, trust, and openness among AI scientists and engineers to help us all better achieve these goals.

Members of the Partnership on AI

Trust is critical to the adoption of AI and realization of its benefits. Developing a framework for AI best practices – for it to work, for public acceptance, to make sure we have done it right – must include relevant and impacted stakeholders. A paramount directive of PAI is that its membership be reflective, in an equitable way, of both profit and non-profit entities. We also believe that, given the global nature of technology, that our membership include organizations from around the world.

Since we came together to form PAI we have been uplifted by support from around the world, with enthusiasm about our mission, tenets, and goals coming from companies, nonprofits, and academics alike. We continue to work to build a diverse, multi-stakeholder organization for open and constructive dialogue on AI.

At present, PAI has more than 50 organizations as members, with more than 120 representatives of these organizations contributing to its first Working Groups. Membership comprises civil society organizations, advocacy organizations, academic research laboratories and institutes, and for-profit organizations, including six of some of the largest technology companies in the world. The Board represents similarly diverse perspectives, with members hailing from organizations including large companies, as well as representatives of the ACLU, the MacArthur Foundation, OpenAI, and The Association for the Advancement of Artificial Intelligence. We think there is strength in this early ideological diversity and look forward to continuing to build upon it with representatives from new industries, new geographies, and new perspectives not often heard in discussions related to technology governance.

A complete list of Partners and Board members can be found at:
https://www.partnershiponai.org/partners/
https://www.partnershiponai.org/board-of-directors/

Committee on Oversight and Government Reform
Witness Disclosure Requirement — "Truth in Testimony"

Pursuant to House Rule XI, clause 2(g)(5) and Committee Rule 16(a), non-governmental witnesses are required to provide the Committee with the information requested below in advance of testifying before the Committee. You may attach additional sheets if you need more space.

Name:

1. Please list any entity you are representing in your testimony before the Committee and briefly describe your relationship with each entity.

Name of Entity	Your relationship with the entity
Partnership on AI	Executive Director

2. Please list any federal grants or contracts (including subgrants or subcontracts) you or the entity or entities listed above have received since January 1, 2015, that are related to the subject of the hearing.

Recipient of the grant or contact (you or entity above)	Grant or Contract Name	Agency	Program	Source	Amount

3. Please list any payments or contracts (including subcontracts) you or the entity or entities listed above have received since January 1, 2015 from a foreign government, that are related to the subject of the hearing.

Recipient of the grant or contact (you or entity above)	Grant or Contract Name	Agency	Program	Source	Amount

I certify that the information above and attached is true and correct to the best of my knowledge.

Signature: Terah Lyons Date: 4/04/18

Terah Lyons' Bio:

Terah Lyons is the founding Executive Director of the Partnership on AI, a multistakeholder nonprofit initiative focused on advancing the benefits and addressing the challenges of machine intelligence founded by Amazon, Apple, DeepMind, Facebook, Google, IBM, and Microsoft. She is a former Policy Advisor to the U.S. Chief Technology Officer in the White House Office of Science and Technology Policy (OSTP) and a Mozilla Foundation Technology Policy Fellow. In her capacity at the White House, Terah led a policy portfolio in the Obama Administration focused on machine intelligence, including AI, robotics, and intelligent transportation systems. In her work at OSTP, she helped establish and direct the White House Future of Artificial Intelligence Initiative, oversaw robotics policy and regulatory matters, led the Administration's work from the White House on civil and commercial unmanned aircraft systems/drone integration into the U.S. airspace system, and advised on Federal automated vehicles policy. She also advised on issues related to diversity and inclusion in the technology industry and entrepreneurial ecosystem.

Prior to her work at the White House, Terah was a Fellow with the Harvard School of Engineering and Applied Sciences based in Cape Town, South Africa. She previously worked with David Gergen at the Harvard Kennedy School of Government Center for Public Leadership examining leadership and U.S. politics, and with Amy Kaslow working to chronicle domestic and international economic reconstruction, poverty, and contemporary conflict and genocide. She is a graduate of Harvard University.