The AI Insight Forum on Innovation
Oct. 24, 2023

Remarks delivered by Suresh Venkatasubramanian

Leader Schumer and Senators Heinrich, Rounds, and Young: I thank you for inviting me to testify at this important forum on artificial intelligence (AI) and innovation. I'm a professor of computer science and the director of the Center for Technological Responsibility at Brown University. I recently completed a stint as a White House tech policy advisor in the Biden administration, and included in my portfolio was developing the recently released Blueprint for an AI Bill of Rights.[1] I have spent the last decade studying and researching the impact of automated systems (and AI) on people's rights, opportunities and ability to access services. I've also spent time working with civil society groups like the American Civil Liberties Union and Data & Society and have advised state and local governments on sound approaches to governing the use of technology that impacts people's lives.

Whom are we innovating for?

The topic of this forum is innovation: how we reimagine what's possible. When we imagine what's possible, we don't often ask, "for whom?" The push for equity in technology starts by centering the needs of those who have traditionally been marginalized — because innovation must be for all and not just for the wealthiest and most privileged among us. Innovation that ignores the needs of those already exploited, harmed and marginalized by technology will only defer the problem to the future, when it will be harder to effect change. And without addressing issues of equity, we won't be able to truly innovate beyond where we currently are.

Unfortunately, those investors and CEOs profiting off of extractive technologies that harm much of the population and benefit a few may tell you that equity and innovation are at odds with each other. I'm here to tell you, based on my 30 years of experience as a technologist and researcher, that this is a false choice.

Equity in technology is really a different *kind* of innovation. It centers people rather than the technology. It seeks to mitigate the real and persistent harms accruing from insensible deployment of technology, including badly designed and ineffective systems, systems that are discriminatory and use data recklessly, and systems that are inscrutable and unaccountable. *Equitable innovation* builds trust between people, communities and institutions and the algorithms that bind them all together. By focusing on people and our needs, equitable innovation creates value that goes above and beyond any specific use case, as I now outline.

Consider the years of work that have gone into identifying how AI decision-making systems might exhibit discriminatory behaviors.[2] We now have many ways to measure the degree of disparate outcomes and identify the sources of bias. We have numerous tools to mitigate such

---

[1] https://www.whitehouse.gov/ostp/ai-bill-of-rights/
[2] These harms have been extensively documented in the Blueprint for an AI Bill of Rights.

biases — one of the methods I developed is now included in a system that IBM distributes widely. What these tools enable us to do now goes well beyond simply addressing discriminatory behaviors. We can build systems that are sensitive to the needs of individual groups of people, even those not defined by race or gender categories. We can test and evaluate systems before deploying them to make sure that their performance is consistent across different groups of people. And in a fascinating twist, we can use any exhibited bias in algorithms as a lens on the data used to develop it so that we can identify sources of bias that come from human decision-making.[3] Equitable innovation builds trust and higher-quality systems and helps us learn more about how people make decisions.

We want explanations for AI systems so that we can tell whether a system has come to a conclusion erroneously and so that an impacted person can appeal a decision they feel has been wrongly rendered upon them. But scientists have repurposed the same technology that allows us to explain the predictions of an AI system so that they can use AI to identify patterns in data and use the explanatory technology to understand why the system found those patterns. Those insights have provided the seeds for new scientific discoveries in areas ranging from cosmology[4] to chemistry,[5] and more are likely to emerge with time.

Data is the fuel for machine learning. On the surface, concerns around data privacy would only hinder our ability to collect the data needed to build complex predictive models. However, innovation in *privacy-enhancing technologies,*[6] including privacy-preserving machine learning and data minimization, has given us the ability to have our cake and eat it too. The most striking example of this is in the way data from the U.S. census is collected and aggregated for analysis. We can protect individual privacy as well as make inferences on data in a way that is secure from data leakage and that builds trust.

It is not a coincidence that the three examples above showcase innovations that started as ways to correct concerns laid out in the Blueprint: algorithmic discrimination, data governance and explanations. When we innovate equitably, everyone benefits.

Who innovates?

While Big Tech has adopted the mantle of innovators over the last two decades, I see a different set of innovators who have truly reimagined what's possible in an AI-driven world. Researchers, mainly in academia, have discovered new ways to evaluate the behavior and performance of AI tools and have radically changed our understanding of how to build and evaluate the decision-making tools we use in the world. Community advocates[7] and partners in civil society[8] have reimagined what we want — and do not want — from the tech tools that

---

[3] BBC. *Amazon scrapped 'sexist AI' tool*. 10 Oct 2018. https://www.bbc.com/news/technology-45809919

[4] Hughes et al. *The GALAH Survey: A New Sample of Extremely Metal-poor Stars Using a Machine-learning Classification Algorithm.* May 2022. https://ui.adsabs.harvard.edu/abs/2022ApJ...930...47H/abstract

[5] Raccuglia et al. *Machine-learning-assisted materials discovery using failed experiments.* *Nature,* volume 533, pages73–76 (2016). https://www.nature.com/articles/nature17439

[6] White House Office of Science and Technology Policy. *Advancing a vision of privacy-enhancing technologies.* 28 Jun 2022. https://www.whitehouse.gov/ostp/news-updates/2022/06/28/advancing-a-vision-for-privacy-enhancing-technologies/

[7] The Algorithmic Justice League. https://www.ajl.org/

[8] Data & Society. https://datasociety.net/

increasingly take up space in our lives, and they have shown the way to building better assistants that help us live on our own terms. Journalists[9] have documented the impact of automated decision-making in society and, in doing so, have helped us reimagine how to govern and monitor the algorithms we increasingly use.

The tools that researchers, civil society and journalists have built are everywhere. They have made it possible for the National Institute for Standards and Technology to develop a risk-management framework[10] that is practical and realistic. They have made it possible for the EU to put out legislation[11] that will govern the development and deployment of AI. And they inspired the Blueprint for an AI Bill of Rights[12] — itself a blueprint for a reimagined, possible and achievable future.

And all of this happened while those charged with innovating — those many corporations that occupy the bulk of forums like these — sat on their hands. They fired those within their organizations who sought to innovate (and fired those who attempted to protect those whom innovation too often ignored). They merely said NO to a different imagining of the future; to tech designs that could benefit all and, above all, to anything except business as usual.

Inaction is not innovation. And as a computer scientist who helped bring about the revolution in responsible AI innovation through my own research as well as through building a community of researchers[13] who are now training their own students, I'm disappointed and disheartened at this lack of imagination and the lack of faith in American ingenuity.

How will we get equitable innovation?

I have painted a picture of a vibrant ecosystem of innovation. All of this is happening within universities and civil society and yet within (very) few companies. The problem is one of incentives. I have spoken with C-suite executives, heads of engineering, compliance professionals and data scientists in the trenches at many companies, and they all point to incentives that mitigate against supporting equitable innovation.

First, there is no competitive advantage in embracing this innovation if in the very short term a company loses out to its competitors who don't. I have repeatedly heard that the lack of a level playing field with rules that govern all players, as well as the uncertainty around what rules are likely to emerge from Washington, are reasons to wait and see rather than start deploying the practices that equitable innovation calls for.

Second, the pace of development in AI creates a sense of urgency that companies feel compelled to respond to, even without a clear strategy. Companies are pouring billions of dollars into strategic repositioning around AI but are not yet aware of the pitfalls associated with it — pitfalls that only come with experience; consideration; and, above all, time.

---

9 The Markup. https://themarkup.org/
10 National Institute of Standards and Technology. *AI Risk Management Framework. https://www.nist.gov/itl/ai-risk-management-framework*
11 https://artificialintelligenceact.eu/wp-content/uploads/2023/08/AI-Mandates-20-June-2023.pdf
12 https://www.whitehouse.gov/ostp/ai-bill-of-rights/
13 https://facctconference.org/

We must act now. We must allow innovators to experiment; but to do that, we need to prioritize the people impacted by technology and ensure that the field of play is level, the rules are clear, and the game is not causing harm. **It is time for Congress to step in and establish this level playing field by passing legislation that prioritizes equitable innovation.**

Making sure we all benefit from technology.

Guardrails — the literal ones we build on curving mountain roads like in Utah where I have lived for many years — make us safer. I can drive through a mountain pass, and as long as my car has headlights, I can see the guardrails and make sure I stay off the edge. I don't need to buy a more expensive car with built-in radar, LIDAR or other sensors to detect that I might be nearing a cliff. The guardrail — a simple piece of technology — benefits everyone.

Equitable innovation benefits all of us — not just a few. As guardians and representatives of the public trust, **Congress should ensure that technologies put out into the world are *by design* engineered to benefit all of us, the American people.** This means that AI systems are tested for safety and effectiveness; that they don't discriminate; they use our data prudently and minimally; they are transparent and explainable; and that technology doesn't make it harder for any of us to access its benefits.

Establishing ground rules to play.

I'll go back to my guardrails example. I grew up outside the United States in a town where lanes on roads were merely suggestions, if they were even present, and where cars, pedestrians, animal carts and motorcycles all shared one road in a vibrant and yet chaotic mess. The lack of lanes did not make the traffic go faster. Rather, traffic moved much, much slower.

Establishing rules of the road for technology creates a playing field with clear expectations of how systems behave. These expectations lay a foundation of trust; hold all participants to the same standards (no exception for expensive cars!); and allow for innovation to grow—most cars now have lane drift detectors because they rely on the lane markers for guidance.

A critical piece of establishing ground rules is accountability. It's not enough for entities building products to "self-evaluate." They need to show their work to an independent authority, or an entity who is accountable to the people.

When tech companies claim they can't show their work for risk of losing proprietary information, they are trying to rewrite the rules of the game to suit their own purpose. They are also ignoring the inconvenient fact that we regularly demand that companies that make cars, planes, drugs or really any kind of food show their work to ensure that their products are safe, and this has not compromised the protection of proprietary information.

Further, transparency and reporting are not just accountability mechanisms. They are ways to encourage learning in the process of innovation. I'm a professor, and when I ask students to show their work, they don't complain that I'm hindering their innovative thinking. They do as

asked, and it makes them learn better and solve problems faster in the future. Similarly, transparency and reporting are an accountable mechanism to learn how complex AI systems work on the ground[14] and adapt their design accordingly.

Ensuring access to innovation.

There's a growing debate about access to the most powerful AI tools out there and whether there should be limits on who can build them and deploy them. Access to innovation must be available to all of us, and that access is crucial for equitable innovation for several reasons.

First, we need to harness the creativity that exists in all sections of our population, and we want to make everyone's path to innovate as easy as possible. Doing so will bring a plethora of new voices and perspectives to a space that has traditionally been dominated by a few, and innovative new ideas will emerge from this chorus. Our diversity is our strength.

Second, we want competition. Competition makes technology better — but only if there's real competition. Right now, the resources required to innovate exclude all but a few players. **Congress should pass the CREATE AI act and establish the National AI Research Resource to make the playing field that much more level.**

Third, we want to make sure the systems we build are reliable and safe from vulnerabilities. Decades of experience in building secure systems has taught us that the best way to make systems reliable and safe is to have them be open for all to stress-test. The White House-initiated open red-teaming exercise for generative AI at DEFCON this year[15] is a good illustration of the power and value of expanding access to AI systems.

Conclusion

Where I grew up, it was virtually impossible to get cars that had built-in seatbelts. When I visited one time with my family, my children were young, and we spent days finding a rental car that had seatbelts on all seats. The driver, however, wasn't wearing his, and my 4-year-old son whose only experience with cars and seatbelts was in the United States, remonstrated with the driver, ultimately convincing him to wear his seatbelt. Today, seatbelts are far more common in cars there, and most people do wear them. They've seen the value and benefit from it.

I think of this when I think about American leadership and equitable innovation. We can foster a new wave of innovation that leverages the vast amount of scientific discovery and innovation already coming from our academies and civil society. The U.S. can be the example that other countries — already looking to see how the U.S. will think about technology governance — will follow. We will be a model for inclusive, forward-thinking, and enlightened technology design that serves us rather than controls us.

---

[14] Data and Society. *AI on the ground.* https://datasociety.net/research/ai-on-the-ground/

[15] White House Office of Science and Technology Policy. *Red-Teaming Large Language Models to Identify Novel AI Risks.* 29 Aug 2023. https://www.whitehouse.gov/ostp/news-updates/2023/08/29/red-teaming-large-language-models-to-identify-novel-ai-risks/