Statement of Julia Stoyanovich

Associate Professor of Computer Science & Engineering and of Data Science,
Director of the Center for Responsible AI at New York University

U.S. Senate AI Insight Forum: High Impact AI

November 1, 2023

Leader Schumer, Senators Rounds, Heinrich, and Young, thank you for inviting me to participate in this important forum! I am an associate professor of Computer Science & Engineering and of Data Science, and the founding Director of the Center for Responsible AI at New York University. My academic research is focused on AI and data engineering systems, and on how to incorporate legal requirements and ethical norms into the way these systems are designed, developed and used.[1] I teach responsible AI to students[2], practitioners in industry and government[3], and members of the public[4]. And I have been deeply involved in AI governance and regulation in New York City[5], New York State[6], and elsewhere, since 2017.

My main point today is that most AI systems are high-impact.

The impacts of AI systems can be positive or negative. As a positive example, there is substantial promise in combining human expertise with AI in radiology to improve diagnostic accuracy. Much additional work is needed to realize this potential: technical research and validation, upskilling of radiologists to use AI while exercising agency and control, and engagement with patients and their families to further validate the systems and build trust. This work is challenging, but it's worth it, and it would be irresponsible not to try!

On the negative side, we worry about several high-impact scenarios. I highlight some of them:

- The most straightforward is when <u>mistakes made by AI systems lead to catastrophic irreversible harms,</u> even to the loss of human life. Examples include autonomous vehicles that have been known to cause fatal traffic accidents, and lethal autonomous weapons that may confuse a civilian in a wheelchair with a motorized combatant[7].

---

[1] This research is supported in part by the National Science Foundation through Awards 1741047 1926250, 1934464, 1922658, 2326193, and 2312930.

[2] https://dataresponsibly.github.io/courses/

[3] https://dataresponsibly.github.io/algorithmic-transparency-playbook/

[4] See free public education course "We are AI: Taking Control of Technology" at https://dataresponsibly.github.io/we-are-ai/ , along with an accompanying comic book series https://dataresponsibly.github.io/we-are-ai/comics

[5] https://www.nyc.gov/site/adstaskforce/index.page

[6] https://www.osc.state.ny.us/state-agencies/audits/2023/02/16/artificial-intelligence-governance

[7] https://www.ohchr.org/en/calls-for-input/2021/report-special-rapporteur-rights-persons-disabilities-artificial-intelligence

- In other cases, individual mistakes made by - or with the help of AI - may be less severe, but, taken together, they lead to cumulative harms that impact a large number of people and may be difficult to reverse. An example is systematically denying jobs to older individuals or to individuals with disabilities.
- There are also cases where, rather than predicting and assisting us, AI systems deliver a self-fulfilling prophecy by steering us towards products, services and opinions. For example, misinformation and disinformation may rapidly spread online, leading to political polarization and threatening our democratic institutions.
- And then, there are cases where AI systems simply don't work. They generate arbitrary answers, or make arbitrary predictions. For example, an AI chatbot may confidently give an incorrect answer to your question, and even fabricate supporting evidence. And an AI hiring assistant may consider you more or less fit for a job if your resume is provided in PDF format vs. in raw text.[8]

Furthermore, irrespective of whether AI predictions are beneficial, harmful or inconsequential, the use of this technology has substantial environmental impacts.[9]

Congress can and should play a central role in ensuring that we realize the benefits of AI while controlling the risks – to individuals, organizations, society at large, and the environment. In my statement, I make three recommendations towards this goal.

Recommendation 1: Congress should mandate public disclosure about the use of AI, starting with the use of this technology by the federal government.

It is crucial to involve members of the public – the totality of individuals who are impacted by AI systems in their everyday lives – in deliberations about the foreseeable risks of AI and the appropriate legal standards for oversight, both before AI systems are put in place and continuously throughout their use.

A public accountability mechanism that is both powerful and simple to enact is to notify people when they are interacting with AI, or when they are being subjected to AI-assisted decisions. Such simple disclosure must be in place especially when AI is used in high-impact domains, including hiring and employment, education, housing, credit and lending, and healthcare. And it must be in place whenever AI is used in decision-making by a government entity. Disclosing AI use is important not only because it can help us watch out for and collectively control the risks, but also because it helps uphold human dignity and autonomy.

---

[8] https://dl.acm.org/doi/10.1145/3514094.3534189
[9] https://themarkup.org/hello-world/2023/07/08/ai-environmental-equity-its-not-easy-being-green

Earlier this Fall, we offered a course on algorithmic transparency to professionals in the news, media and journalism industry. In a follow-up interview, one of the participants noted: *"If you are questioning whether or not you need to tell people [about AI], you need to tell people."*

Disclosing AI use may seem trivial, yet it's anything but. New York City's government has been signaling its intention to improve public accountability in its use of automated decision systems (ADS) – that include AI – since 2017, starting with the appointment of the ADS Task Force, on which I served. Unfortunately, results have been less than satisfactory so far, despite substantial public pressure. The NYC government has issued several roadmaps and action plans over the years, each promising to start working on an approach to AI governance and public disclosure. And yet, the use of a disappointingly small number of systems has been disclosed by the NYC government so far.[10] Part of the reason for this is that there is no independent oversight of NYC government's transparency efforts, and so, ironically, the government remains unaccountable in its own commitment to accountability. Congress should learn from this experience and ensure that public disclosure mandates – in both government and industry – are enforced by an independent oversight entity.

Beyond disclosing AI use, public disclosure should clearly and concretely explain what goals the AI system pursues and how it was validated against these goals. Is the AI helping you apply for as many potentially relevant job openings as possible? Or is it helping you selectively apply for just a few openings that are a perfect match for your qualifications? Does the system work? And how do we know that it works? Notably, these basic questions – and much of the methodology to test for them – are the same for simple AI systems like those used in scoring and ranking, and for complex systems that use generative AI.

Let's take hiring and employment as a concrete high-impact example. New York City Local Law 144 (LL 144) is the first effort in the U.S. to place guardrails around the use of automated decision systems in hiring and employment. This law requires that job seekers be notified before they apply that they will be screened by an AI. And it also requires (a limited form) of bias audit for these tools. LL 144 is a good first step, but it falls short of weeding out the bad actors among the tool vendors: The law does not include any provisions to explain to job seekers why they were screened out by the tool, or, more generally, whether and how the tools were validated. This omission is conspicuous, and it risks eroding public trust.[11]

---

[10] See the report by the NYC Mayor's Office of Operations' Algorithms Management and Policy Officer in 2020, https://www.nyc.gov/assets/oti/downloads/pdf/reports/ampo-agency-compliance-cy-2020.pdf   See also results of the AI Governance audit, conducted by the Office of New York State Comptroller, published in February 2023, https://www.osc.state.ny.us/files/state-agencies/audits/pdf/sga-2023-21n10.pdf

[11] We conducted public engagement activities as part of deliberations on LL 144, see https://dataresponsibly.github.io/documents/Bill1894Showreel.pdf

Indeed, there is evidence to suggest that recommendations of many employee selection tools are inconsistent and arbitrary[12]. Tools that don't work hurt job seekers and employees, subjecting them to arbitrary decision-making with no recourse. Tools that don't work also hurt employers, they waste money paying for software that doesn't work, and miss out on many well-qualified candidates based on a self-fulfilling prophecy delivered by a tool.

Finally, <u>disclosure should be meaningful rather than misleading</u>. I will give another example from New York City, this time focusing on the high-stakes process used by the NYC Department of Education (DOE) to match 5th graders with middle schools. The process is based on a lottery: every child has a random number assigned to them by the DOE, which determines their priority for admission to a school of their choice. After substantial public pressure, the DOE started disclosing lottery numbers to the parents. However, these numbers are released as long alpha-numeric sequences that represent hexadecimal numbers. (For example, my son's lottery number is a53c44d7-37c6-41a7-95d7-01c579814016.) This is problematic, and is rightfully perceived as deliberately confusing and obstructive. Few parents are familiar with hexadecimal numbers, only the first two symbols are important ("a5" in my son's case). A more meaningful way to disclose this information is to convert the hexadecimal number to a percentile that parents will understand (approx. 65th percentile, in my case).

In summary, Congress should help establish public disclosure standards, and mandate their use in government and in industry. It should use financial incentives – or apply financial pressure – to ensure that these standards are also met by state and local governments.

Recommendation 2: Congress should substantially increase federal investment in responsible AI research, education and training.

Responsible AI is a new interdisciplinary field that combines methodologies from computer science, social science, psychology, policy, and law. I have been active in this field for about a decade, and can attest to tremendous progress. However, much additional research is needed to develop new and commoditize existing responsible AI tools and methodologies, making them readily available to those developing, using, and overseeing AI. We must substantially increase federal investment in responsible AI research to meet these demands.

Further, <u>we cannot understand the impact – and especially the risks – of AI systems without active and thoughtful participation of everyone in society</u>, either directly or through their trusted representatives. To think otherwise is to go against our democratic values. To enable broad participation, we must substantially increase federal investment in responsible AI education

---

[12] "Resume Format, LinkedIn URLs and Other Unexpected Influences on AI Personality Prediction in Hiring: Results of an Audit," Rhea et al., AAAI/ACM AIES 2022, available at https://dl.acm.org/doi/10.1145/3514094.3534189

and training for everyone - children and college students; working professionals who build and use AI; auditors; policymakers and regulators; and those impacted by AI – the public at large.

We cannot in good faith rely on for-profit entities to do responsible AI research and education – any more than we can rely on oil companies to be the arbiters of research and education on climate change. Instead, we must direct federal funding towards US-based academic institutions – through the National Science Foundation, the National Institutes of Health, and other funding agencies – to frame, lead and conduct this work, and to bring in industry participation only as appropriate.

That being said, commercial entities have been reaping outsized financial benefits from AI, while leaving the rest of society to bear the brunt of controlling these tools – an impossible task without sufficient access to information and to other resources. Some companies have been making voluntary commitments to responsible AI and AI safety. For example, Anthropic, Microsoft, Google, and OpenAI announced "a new AI Safety Fund with over $10 million to accelerate academic research on frontier model safety."[13] This amount is negligible compared to the combined valuation of each of these companies (in the trillions for Microsoft and Google, and in the billions for Open AI and Anthropic). And it's grossly insufficient to meet the demands of responsible AI and AI safety research.

Congress should ensure that commercial entities that capitalize on AI contribute a substantial portion of their profits to responsible AI research and education. A promising option is to impose a **high-impact AI tax**, similar to carbon tax, and to use the collected funds to support responsible AI research, education and regulation efforts.

Recommendation 3: Congress should mandate a comprehensive inquiry into the environmental impact of AI. And it should financially disincentivize unnecessary or unproductive use of AI to limit these impacts.

The use of AI has substantial environmental impacts, with costs incurred through hardware manufacturing and operation, data collection and storage, model training and maintenance, and interactions with the already trained models. Environmental impact must be factored into the deliberation about the benefits and the harms of AI. We do not currently have sufficient information – or even reliable ways to get information – to quantify this impact.

Congress should mandate a comprehensive inquiry into the environmental impact of AI, and it should support open academic research on ways in which this impact can be reduced. Once more information is available, Congress will be in a position to financially disincentivize unnecessary or unproductive use of AI by quantifying its end-to-end environmental footprint and factoring that information into the taxation system.

---

[13] https://www.frontiermodelforum.org/