Written Submission
AI Insight Forum: Privacy & Liability

Mark Surman
President of the Mozilla Foundation

November 8, 2023

Leader Schumer, Sen. Rounds, Sen. Heinrich, Sen. Young, and members of the Senate,

I'm grateful for the opportunity to speak to you today about two critical questions when it comes to artificial intelligence: how can we protect people's privacy in the AI era? And what does an appropriate framework for AI liability look like?

At Mozilla, we have a unique vantage point in finding answers to these questions: that of a non-profit foundation *and* a tech company. As a foundation, we've spent the past five years exploring what it takes to make AI trustworthy and, along with nine other philanthropic foundations, have joined Vice President Harris in announcing a $200 million investment in the trustworthy AI ecosystem. As a tech company, we're investing heavily in leveraging AI in our products and have set up our own AI R&D lab, Mozilla.ai.

Regulation is necessary

Before diving into these issues, I want to make clear that regulation is needed. We need safeguards to protect people and our institutions — and there is ample evidence that we need these sooner rather than later. In developing these safeguards, we must move swiftly but also with care. Regulation must not shut out competition as a byproduct by creating regulatory moats around the biggest players. We need regulation that is targeted in addressing the root causes of AI-enabled societal harms without entrenching the position of a handful of players dominating the AI industry. We're already seeing stakeholders pushing for regulation that would either exempt them or would lock in their current market position. For example, a restrictive licensing regime for AI would heavily tilt the odds against SMEs, AI start-ups, and open source developers by turning a small number of dominant companies into AI's gatekeepers.

However, good regulation will open up markets for everyone to share in the benefits of AI and at the same time protect people from AI-driven harms. So we should learn from the past and ensure that a handful of dominant players don't, again, get to create a playing field that first and foremost benefits themselves. This does not mean that SMEs or open source developers should not be subject to regulatory oversight or not have binding rules apply to their creations. However, the path to responsible release practices should be paved with diverse input, with an aim to foster and not forestall the technology ecosystem that has led to these innovations in the first place.

Privacy is the backbone of sound AI policy

Privacy is not merely a feature to be added to AI systems post-development; it is a foundational principle that must be woven into the fabric of AI development and governance. This applies to the entire lifecycle of AI — from preventing unwanted data collection for training models to preventing leakage of private information by an AI system after it has been deployed.

As the race to build large AI models accelerates, so does the race to accumulate vast swaths of data. In this context, proprietary and even personal data in particular become a key competitive advantage, with incentives to collect more and more of it only increasing. For example, social media companies could collect even more user data to train AI models and data brokers could be emboldened to scoop up more data to sell to companies with ambitions to develop their own AI models. Throughout the two and a half decades of Mozilla's history, we have argued that individuals' security and privacy on the internet must not be treated as optional and we have championed lean data practices. It should be no different when it comes to AI.

What we need to prevent is a race to the bottom when it comes to privacy — so passing a federal privacy law, setting effective rules of the road, is paramount. U.S. leadership in AI should not come on the back of Americans' privacy. That's why Mozilla endorsed the American Data Privacy and Protection Act last year, because we believe it will provide these vital protections, and we stand ready to work with you on a strong privacy proposal going forward. We also support the FTC moving to write privacy rules of the road in the interim. We can observe this in the EU, too. The EU Artificial Intelligence Act would be incomplete if it weren't built on the foundation of the robust privacy protections afforded by the GDPR. The EU's landmark privacy law already tackles the input side in AI development, requiring a legal basis (such as consent) to obtain personal data that can be used to train AI — something that's still missing in most of the US.

But this is only part of the challenge: What we also need is investment to develop more and better privacy-enhancing technologies (PETs), and more private ways of building AI. As we've said before, informed consent and user agency must remain the core pillar of data collection. But PETs can enable innovation by de-risking processes throughout the AI lifecycle, for example by enabling training and processing on-device or in secure environments, by leveraging encryption techniques, or by obfuscating data to prevent re-identification.

This takes significant resources — but it will be worth the investment and one that government agencies should support and encourage. At Mozilla, we are doing exactly that through our own AI research lab, Mozilla.ai. Open source can play an important role here to advance and diffuse privacy-preserving techniques in this context, for example by enabling people to run AI models on their private devices with private data, and without sending data to a third party.

All this requires finding answers to tough questions, some technical and some policy. What recourse should people have if their data is scraped and used to train AI models against their will? How can they object when a model has already and irreversibly been trained on their data? How can malicious actors be prevented from extracting people's data from AI models? There is no silver bullet solution here, but we know that comprehensive privacy legislation, empowering the FTC to undertake effective enforcement, and strong privacy and data governance requirements tailored to AI systems are the first steps.

Openness and transparency are key enablers in allocating responsibility

Assigning liability for harms caused by AI is incredibly complex — there's no straight line from harm to where it originated. For example, take a chatbot deployed to help people struggling with mental health issues by a company providing mental health services. The hypothetical company uses a pre-trained AI model from a leading AI lab and fine-tuned it for this specific context. Now say the chatbot starts providing advice that's actually detrimental to a user's mental health — who should be liable, the large AI lab or the mental health company?

Our work on the EU AI Act in the past years has shown the difficulty of identifying who's at fault and placing responsibility along the AI value chain. From training datasets to foundation models to applications using that same model, risks can emerge at different points and layers throughout development and deployment. At the same time, it's not only about where harm originates, but also about who can best mitigate it. After all, different actors along the AI value chain may have varying levels of expertise, resources, or insights into a given AI model.

Against this backdrop, it becomes clear that liability cannot just be focused on one particular layer of the AI 'stack'. Claims that AI should only be regulated at the application layer (i.e. those who deploy them in the final stages) should thus be met with skepticism. After all, some risks are baked into an AI model from the original development stage, such as harmful prejudice and bias or a tendency to 'hallucinate', i.e. for a model to produce false information. If a downstream actor were to integrate such a model into their own product, they may be able to mitigate these risks to some degree, but the upstream developer would in most cases be much better placed to do so. At the same time, not all risks emerge upstream. Some are very much specific to the field and context of application and subsequently need to be addressed downstream. Further, these examples paint a simplistic picture of the AI value chain — often, additional actors and components will be involved.

Any framework for imposing liability needs to take this complexity into account. What is needed is a clear process to navigate it. Regulation should thus aid the discovery and notification of harm (regardless of the stage at which it is likely to surface), the identification of where its root causes lie (which will require technical advancements when it comes to transformer models), and a mechanism to hold those responsible accountable for fixing or not fixing the underlying causes for these developments. To meet this standard, regulators will need to be well-resourced and equipped with the right expertise, which clearly points to the need for greater investment in public research, ability to hire leading technical talent within government agencies, and ensure that regulatory frameworks remain dynamic enough to allow for flexibility on the part of regulators.

Openness and transparency are key enablers in this regard. These principles can help scrutinize AI and allocate responsibility. Openness and open source — that is, in simplified terms, making different AI components (from models to code to data) openly accessible — can help make AI and its value chain more scrutable and enable independent study and evaluation of AI. And not only are open source models far easier to scrutinize; their existence also improves the auditability of 'closed source' models. For example, using open source AI, researchers developed a method to test large language models' safety filters, including those of popular proprietary chatbots, thus revealing a potential vector of attack.

Further, there are additional benefits of open source AI: especially when paired with strong competition rules that prevent anti-competitive practices, it spurs innovation and accelerates competition in the marketplace by providing common resources for the ecosystem at large. Competition spurs investment. And investment creates new jobs. Open source can provide building blocks to American start-ups building new products and it can prevent companies from being locked in by dominant vendors. Open source creates alternatives in the market, and without alternatives there is no choice for companies and consumers. Mozilla's own history is a case in point: Without the 25-year-old open source project that is our Firefox web browser, the browser market may never have seen real competition.

When it comes to liability, the role of open source warrants special consideration. Open source developers, too, should act responsibly and work to mitigate risks linked to the AI components they release. But liability frameworks should ensure that the open source AI ecosystem — and especially public-interest and community-driven work — isn't paralyzed by fear of liability claims. To harness both open source contributions to innovation as well as important safety and security research enabled through open source research and development, any liability framework needs to take a proportionate approach to open source. At the same time, we're seeing many actors deploy narratives of openness as a facade in order to advance their own interest. We should be wary of allowing open source to be co-opted by powerful commercial interests and regulation should take into account that not everything that is labeled as "open" actually is. When it comes to liability, it therefore is also important to ask *who* we are talking about — for example, a small non-profit, open source research organization is in a different position to comply with regulation than a large tech company releasing open source AI models as part of its broader business strategy.

Finally, we must look beyond a binary notion of open versus closed AI. First, openness is a spectrum rather than a binary state. Second, open and closed AI ecosystems can co-exist. In fact, they can even benefit from each other. We've already seen this in both AI — where many recent advances rely on open source techniques and frameworks — and in traditional software, where open source code is part of virtually every codebase. The challenge is not to determine which ecosystem should persist and which one should fade. It is about creating a balanced environment where both paradigms can flourish to fuel innovation, ensure competitiveness, and protect people's rights and safety.

Concluding thoughts

As progress in AI as well as its commercialization accelerates, it is critical that we take action to ensure that the benefits of AI are shared widely across society and to protect people from harm. Binding rules should be a part of this course of action, and privacy, openness, and transparency should be core principles underlying any regulatory framework. In contemplating new rules for AI, we therefore ask you to consider the following recommendations:

1. **Incentivize openness and transparency:** Open AI ecosystems facilitate scrutiny and help foster an environment where responsibility for AI-driven outcomes can be appropriately attributed. Moreover, openness in AI stimulates innovation by providing the building blocks with which the market can build competitive products. Ensuring that all projects, open source or not, meet minimum criteria of responsible release is different from effectively banning open source approaches due to hypothetical future harms. Openness is not a problem but a core part of the solution that will help a broad group of actors engage core questions in this space, including privacy or liability.

2. **Distribute liability equitably:** The complexity of AI systems necessitates a nuanced approach to liability that considers the entire value chain, from data collection to model deployment. Liability should not be concentrated but rather distributed in a manner that reflects how AI is developed and brought to market. Rather than just looking at the deployers of these models, who often might not be in a position to mitigate the underlying causes for potential harms, a more holistic approach would regulate practices and processes across the development 'stack'.

3. **Champion privacy by default:** Privacy legislation must be at the forefront of the AI regulatory framework. The American Data Privacy and Protection Act, endorsed by Mozilla, would represent a significant step towards providing the necessary privacy guarantees that underpin responsible AI. Until Congress passes a federal law, the FTC should push forward its critical Commercial Surveillance and Data Security rulemaking, and existing rules protecting consumers and competition need to be enforced.

4. **Invest in privacy-enhancing technologies:** Investment in privacy-enhancing technologies, with government funding at its heart, is crucial for the development of AI that protects individual privacy — beginning with data collection. Such investment not only aligns with ethical standards but also drives innovation in creating more responsible and trustworthy methodologies for AI development.

At Mozilla, we will continue to fight for and invest in more trustworthy AI. We hope you join this effort and stand ready to engage.