Written Testimony of

Udbhav Tiwari

Director of Global Product Policy, Mozilla

Before the

United States Senate Committee on Commerce, Science, and Transportation

on

"The Need to Protect Americans' Privacy and the AI Accelerant"

July 11, 2024

—

Chair Cantwell, Ranking Member Cruz, and esteemed members of the Committee,

Thank you for the opportunity to testify on the critical issue of protecting Americans' privacy in the age of artificial intelligence (AI). My name is Udbhav Tiwari, and I am the Director of Global Product Policy at Mozilla. Today, I will discuss the urgent need for comprehensive privacy legislation, the importance of data minimization, and the role of privacy-enhancing technologies in fostering responsible AI development. America can continue to be a leader in AI by putting in place clear rules of the road on privacy that will help to spur beneficial competition and create a level playing field for players of all sizes instead of a race to the bottom.

About Mozilla

At Mozilla, we approach tech policy issues from a unique vantage point as a non-profit foundation, open-source community, *and* a tech company. We build the open-source Firefox web browser, Mozilla VPN, and products like Solo, an AI-powered website builder for micro-SMBs and solopreneurs. These products are used by hundreds of millions of individuals around the world. Mozilla's mission is to ensure the internet is a global public resource, open and accessible to all. To fulfill this mission, we are constantly investing in the security of our products, the privacy of our users and in advancing the movement to build a healthier internet. This includes supporting research in areas like competition and the impact of harmful social media recommendations that have helped to inform policymakers around the world. Mozilla has influenced major companies to adopt better privacy practices and empowered people directly with tools to better understand and protect their data online. For Mozilla, individuals' security and privacy on the internet are fundamental rights and must not be treated as optional.

With this goal, for the past five years Mozilla has been committed to advancing trustworthy AI. Mozilla recently published a paper, Accelerating Progress Toward Trustworthy AI, that outlines how Mozilla and our allies are advancing openness, competition, and accountability in AI. Mozilla is putting its resources behind these priorities as well: The Mozilla Foundation, which is the full owner of the Mozilla Corporation, has been dedicating 100% of its $30M a year budget to philanthropic activities, advocacy, and programmatic work on this topic. Mozilla is also investing another $30M in research and development on trustworthy AI, $35M in responsible tech startups — including startups with a focus on trustworthy AI — through Mozilla Ventures, and building the data infrastructure needed to create AI that works in multiple languages via Common Voice. We're seeking to help ensure that every person and every community can safely build, use, and assess AI.

The Imperative of Comprehensive Privacy Legislation

At Mozilla, we believe that comprehensive privacy legislation is foundational to any sound AI framework. Without such legislation, we risk a "race to the bottom" where companies compete by exploiting personal data rather than safeguarding it. Maintaining U.S. leadership in AI requires America to lead on privacy and user rights.

Privacy is a critical component of AI policy, not just because AI has the potential to accelerate privacy related harms, but because at its heart, AI is made possible by the utilization of tremendous amounts of data. How that data is collected today by AI companies varies, but there is little question that AI, especially generative AI, has created a dynamic that pushes companies to collect as much data as possible, creating a race to accumulate vast swaths of data.

In this context, any additional data a company can collect, including proprietary and personal data, could become a key competitive advantage as companies try to create “data moats,” to ward off would-be competitors. This means that without regulation, business incentives will drive companies to collect more and more data however they can. For example, big tech companies could collect even more user data to train AI models, while data brokers would be incentivized to scoop up additional data to sell to third parties with AI ambitions, as the value of data increases.

AI systems thrive on data, and the drive to develop advanced AI models has intensified the demand for vast amounts of personal information. This data collection, often done without adequate consent or deceptive choices, poses significant risks to individual privacy and security.

By championing policies that promote innovation, create clear rules of the road for companies, and protect fundamental user rights, we can create both a competitive and level playing field for the American AI industry and prepare domestic champions for global leadership. At the core of these policies should be data minimization.

Data Minimization: A Cornerstone of Privacy and AI Policy

Data minimization is a crucial principle that ensures only the necessary data is collected and used for specific purposes - minimizing risk for business and enhancing consumer trust. This approach reduces the risk of privacy breaches, limits the potential misuse of personal information, and mitigates the burden of consumers having to defend their own privacy. In the context of AI, data minimization can be achieved through several strategies including:

1. **Informed Consent**: Individuals must be meaningfully informed about how their data will be used and be asked explicit consent when their personal data is used to train AI models. This transparency builds trust and empowers users to make informed decisions about their personal information while providing clarity to businesses. Mozilla has campaigned on behalf of consumers to push the industry to do better when it comes to transparency.
2. **Privacy by Design**: Integrating privacy considerations into the design and development of AI systems ensures that privacy is not an afterthought but a fundamental component. This holistic approach of privacy by design - encompassing the entire AI lifecycle, from data collection to deployment - is a core element of Mozilla’s recent “Privacy for All” campaign.

Investment in Privacy-Enhancing Technologies

While legislation is essential, technical advances must work hand-in-hand with them to create a more safe and private future. The ecosystem needs significant investment in privacy-enhancing technologies (PETs) to develop AI systems that respect and protect individual privacy. PETs enable innovative solutions that reduce risk in the AI lifecycle while ensuring privacy is maintained without stifling technological progress. For example, PETs can enable training and processing on-device or obfuscate machine learning outputs via differential privacy to mitigate the risks of re-identification.

At Mozilla, we are committed to advancing privacy-preserving approaches in both the browser and for AI in general. We are focused on developing tools that prioritize user privacy by running on user devices (via projects such as Llamafile) and by developing technologies such as Interoperable Private Attribution (IPA) that minimize the need for data collection via browsers by leveraging multi-party computation. **Local or on-device processing** enables AI models to run on local devices rather than centralized cloud servers, significantly reducing the amount of data transmitted and stored by service providers. Mozilla's AI-enabled translation feature in Firefox, for example, performs translations locally using machine learning - ensuring that user data (such as page URLs and content) is not sent to either Mozilla or third-party servers who can then use data collected for their own purposes.

The Role of Openness in Privacy-Preserving AI

Openness is an essential ingredient for improving verifiable and meaningful privacy in AI technologies. While there are different degrees in the spectrum of openness, making AI components and systems more openly available to the wider community, improves transparency. Transparency, in turn, enables scrutiny - without which we have little ability to govern and hold the current large models to any privacy standards we might expect of other online businesses.

Without knowing what data is being collected and how it is being leveraged by AI systems, we will have little ability to govern the technology effectively in the interests of consumers. Without incentives that encourage openness across the AI stack, a handful of dominant companies will win the race to the bottom, while citizens lose any effective means of control over their privacy.

Open approaches also play a vital role in promoting innovation, preventing the concentration of power in the hands of a few companies. They enable the economic benefit of AI to be more widely shared, amongst businesses of different sizes and capabilities - leading to increased investment and job creation. We also have clear evidence from open source development practices, that openness allows for diverse input and collaboration, fostering the development of privacy-preserving techniques that can benefit everyone rather than relying on security through obscurity.

Mozilla's commitment to openness is exemplified by our efforts to enable AI models to run on private devices with private data, via projects such as Llamafile. This reduces the need for data to be sent to centralized servers, mitigating privacy risks and promoting user control.

AI Can Amplify Privacy Violations

Online manipulation, targeted scams, and online surveillance are not new risks in our digital lives. Bad actors often seize on any opportunity they can, whether through spam emails or sophisticated deep fakes, to harm the average American. However, AI technologies can supercharge such harms by enabling personalization and scale with much lower barriers to access such capabilities than previously possible. AI technologies introduce unique privacy challenges that must be addressed proactively, where some of the most pressing concerns include:

- **Profiling and Manipulation**: AI can infer sensitive attributes about individuals, leading to potential privacy violations if used for targeted content or discrimination. This is especially true for advertising, a field where AI and machine learning have already been leveraged for years to predict the wants and desires of unsuspecting consumers. In addition, the growth of generative AI has led to advertisers creating highly customized campaigns, from text to images to videos, raising the likelihood of hyper-targeted manipulation at low costs.
- **Consent and Data Rights**: Using personal data to train AI models raises questions about consent, especially when individuals are unaware their data is being used for training. In the case of sensitive categories, such as kids or health data, existing protections need to be updated for newer use cases that AI enables. These updates can include stricter anonymization standards, inclusion of data inferred by AI systems (such as behavioral profiles and

predictive analytics), and improved disclosure requirements for when such sensitive data is leveraged for training AI models.

- **Bias and Discrimination**: AI systems trained on biased data can perpetuate and amplify these biases, resulting in discriminatory outcomes. We've already seen prominent companies be taken to court for such practices by the US government and AI will only create more avenues for such algorithmic discrimination.
- **Data Exploitation**: Generative AI systems trained on vast datasets may inadvertently reveal private information, posing risks to the privacy and security of the average American or for businesses whose employees use such systems. For example - employees using a cloud-based generative AI platform to create meeting notes by feeding in potentially sensitive or confidential information could inadvertently lead to that platform leaking that data to other parties in the future, something we've already seen occur in the recent past.
- **Deepfakes and Identity Misuse**: AI-generated content can convincingly depict individuals doing things they never did, threatening privacy and reputation.

To mitigate these risks, we need comprehensive federal privacy legislation, strong regulatory oversight, and continued investment in PETs. We must also ensure that AI systems are transparent and accountable, with mechanisms in place to address privacy violations and provide recourse for affected individuals, underpinned by disclosure.

Protecting Civil Liberties in the Age of AI

AI's potential to impact civil liberties cannot be understated. The same technologies that drive innovation can also be used to infringe upon fundamental rights and used by big tech companies to trample individuals' privacy. Therefore, it is imperative that AI development and deployment are guided by principles that protect civil liberties. This includes safeguarding freedom of expression, preventing unlawful surveillance, and ensuring that AI systems do not perpetuate discrimination or bias.

At Mozilla, we have long championed the protection of civil liberties. We believe that privacy is not just a feature but a fundamental right that must be upheld in all technological advancements. Our commitment to privacy preserving data practices and AI reflects our dedication to protecting users' civil liberties in the digital age.

The Path Forward

As we navigate the complexities of AI and privacy, it is crucial to strike a balance between innovation and protection. Regulation must be designed to address the root causes of AI-enabled societal harms without entrenching the position of a few dominant players. We should avoid restrictive licensing regimes that could stifle competition and innovation, particularly for small and medium-sized enterprises (SMEs) and open-source developers.

Instead, we need a regulatory framework that promotes responsible AI development and deployment, safeguards individual privacy, and fosters a diverse and competitive AI ecosystem. This includes urgently creating clear rules and enforcement mechanisms to stop bad actors from exploiting the innate privacy rights of Americans' online. America can continue to be a leader in AI by putting in place clear rules of the road on privacy and data protection that will help to spur beneficial competition and create a level playing field for players of all sizes instead of a race to the bottom. The improved consumer trust engendered by better privacy practices leads to increased brand loyalty, enhancing the competitive edge that American technology companies currently enjoy globally. Robust privacy practices also attract international partnerships and investments, positioning businesses to compete more effectively in the global marketplace - where privacy is increasingly a competitive differentiator.

Conclusion

In conclusion, protecting Americans' privacy in the age of AI is a critical challenge that requires comprehensive legislation, policies that support openness, investment in privacy-enhancing technologies, and a commitment to data minimization. At Mozilla, we are dedicated to advancing privacy-preserving AI and advocating for policies that promote innovation while safeguarding individual rights.

We urge Congress to pass binding federal privacy legislation and enforce strong privacy regulations. By doing so, we can ensure that AI development proceeds in a manner that respects privacy, promotes trust, and benefits all Americans.

Thank you for the opportunity to testify today. I look forward to your questions and to working with you to protect Americans' privacy in the AI era.