Surya Mattu
Digital Witness Lab, Co-founder & Lead
Center for Information Technology Policy

Statement to the U.S. Senate AI Insight Forum on Innovation
Russell Senate Office Building
Washington, D.C.

___

November 1st, 2023

Thank you for giving me the opportunity to present my views on this important issue to the Forum. Here in the twenty-first century, we have witnessed the rapid adoption of digital technology. Computer algorithms have become the invisible infrastructure of our society. And the harms they propagate are not easy to spot.

In today's digital world, injustice lurks in the shadows of the Facebook post that's delivered to certain groups of people at the exclusion of others. Or the hidden algorithms used to profile candidates during job interviews. Risk-assessment algorithms decide who deserves prison time and are used to accuse families of welfare fraud. Algorithmic systems have been integrated into every aspect of society, and meanwhile, regulatory mechanisms have struggled to keep up.

Artificial intelligence is the next frontier in algorithmic injustice. As you decide how to regulate this emerging technology, I urge you to consider how independent scrutiny by journalists and researchers has played a crucial role in checking the power of technology. To date, we have invented our own ways to watchdog systems. Our work has been at the forefront of uncovering harmful practices perpetuated by technology companies. I urge you, as you move forward, to mandate meaningful forms of transparency and disclosures that allow individuals to examine how the technology they rely on every day, and that profits from their data, works. Built-in transparency also empowers journalists, researchers, and community advocates to study the impact of a given technology on society with precision.

I have spent the past decade working as a journalist, watching our world become more ruled by algorithms. Last year, I started the Digital Witness Lab at Princeton University to build independent, public, and open-source tools that scrutinize the role algorithms play in society. As journalists, our job is to audit power–to witness and document harms. And as power has moved into the realm of algorithms, we have had to adjust. We have had considerable success harnessing two seemingly banal tools that were not designed to provide transparency, but which watchdogs have appropriated for that purpose. These are tools that every company uses to track what's happening on its own webpage: browser add-ons and browser developer tools.

Browser add-ons are small programs that can be installed directly onto a web browser, allowing users to augment how they interact with a given website. They are commonly used for bookmarking sites such as Pinterest, as well as operating tools like password managers and ad-blockers. However, they are also incredibly useful for enabling people to collect their own data within a tech platform's walled garden.

Similarly, browser developer tools were made to allow web developers to test and debug their websites' user interfaces. As the internet evolved and websites became more complex, these tools evolved too. They added features like the ability to inspect and change source code, monitor network activity, and even detect when a website is accessing your location or microphone. These are powerful mechanisms for investigating how companies track, profile, and target their users.

I have put these tools to use as a data journalist. I was able to show how a marketing company logged users' personal data from the sites of clinical health trial providers and mortgage companies.[1] Companies were collecting a user's data even before they clicked "submit" on a form to sign up for the company's service. More recently, I documented how the Meta Pixel tool (formerly the Facebook Pixel tool) quietly scoops up people's data as they use hospital websites[2], fill out federal student loan[3] applications, and file their taxes online[4]. Our investigation into the data leaks from tax-filing tools was cited as one of the motivating factors for the "Attacks on Tax Privacy" Senate report[5] that was published in July of this year.

Perhaps the most troubling aspect of our investigation was that there are already existing laws in place that are supposed to protect private data.

The risk of leaking sensitive personal information from websites has been well documented for over a decade. Laws like HIPAA exist to specifically prevent such information from being shared with any third-parties.

Yet our investigation last year found a third of top hospitals' websites sent patient data to Facebook.

One of the most consistent lessons I have learned from my experience as a data journalist is that even when a harm is understood and addressed with regulation, harm often goes unnoticed in the absence of independent scrutiny and persistent monitoring.

Artificial intelligence presents new potential harms, even as we struggle to monitor the behavior of more established technologies. Specific regulations will continue to lag, because the reality is that a lot of the *harms that AI will cause can't be predicted*. The most harmful

---

[1] Before You Hit 'Submit,' This Company Has Already Logged Your Personal Data, https://gizmodo.com/before-you-hit-submit-this-company-has-already-logge-1795906081

[2] Facebook Is Receiving Sensitive Medical Information from Hospital Websites, https://themarkup.org/pixel-hunt/2022/06/16/facebook-is-receiving-sensitive-medical-information-from-hospital-websites

[3] Applied for Student Aid Online? Facebook Saw You, https://themarkup.org/pixel-hunt/2022/04/28/applied-for-student-aid-online-facebook-saw-you

[4] Tax Filing Websites Have Been Sending Users' Financial Information to Facebook, https://themarkup.org/pixel-hunt/2022/11/22/tax-filing-websites-have-been-sending-users-financial-information-to-facebook

[5] Report: Attacks on Tax Privacy: How the Tax Prep Industry Enabled Meta to Harvest Millions of Taxpayers' Sensitive Data, https://www.warren.senate.gov/oversight/reports/in-new-report-senators-warren-wyden-lawmakers-reveal-massive-likely-illegal-breach-of-taxpayer-privacy-by-tax-prep-companies-with-meta_call-for-agencies-to-investigate-prosecute

effects likely won't be caused by direct consumer interaction with the technology, but rather by the secondary and tertiary effects of integrating AI into existing products and services.

And while simple browser tools work for scrutinizing a website for privacy violations, we will need a far more robust transparency framework to study the downstream harms of integrating AI. We need a platform-independent transparency framework–something that I like to call an inspectability API. [6]

An application programming interface (API) is a way for companies to make their services or data available to other developers. For example, a developer building a mobile app may want to use the phone's camera for a specific feature. They would use the iOS or Android Camera API. Another common example is an accessibility API, which allows developers to make their applications accessible to people with disabilities by doing things like making the user interface legible to screen readers.

An inspectability API would allow individuals to export data from the products and services they use every day and share it with researchers, journalists, and advocates in their communities. Companies could be required to implement this API to adhere to transparency best practices, much as they are required to implement accessibility features to make their apps and websites usable for people with disabilities.

Determining how such a framework would work in the context of AI-powered systems requires further research. But at the very least it could involve a *standardized, machine-readable disclosure procedure* to inform users when they are being subject to AI-driven decision making, as well as inform them about what *data of theirs is being used by the company for training AI products*. For instance, requiring disclosure of training data sources for services like ChatGPT would allow content creators to know when their work is being monetized without their consent.

Such a framework would put the onus on tech companies to be more deliberate and explicit in their use of these technologies. Tech companies do not often set out to harm their users, but nor do they necessarily spend a lot of time thinking about the downstream consequences of their products. Forcing companies to disclose their actions could mitigate the unintended consequences of developing large-scale data driven technologies by simply forcing companies to think ahead.

HIPAA should have ensured hospital websites don't share sensitive patient information with Meta, but crafting regulations is not where the work ends. Putting regulations into place for AI will have no meaning if there is no way for researchers, journalists, and other community advocates to independently scrutinize the downstream effects these technologies have in the lives of regular people, especially those belonging to the most vulnerable communities.

---

[6] The Internet Is Turning Into a Data Black Box. An 'Inspectability API' Could Crack It Open, https://www.wired.com/story/inspectability-api-app-transparency/