Why give researchers access to platform information?

When Frances Haugen brought the “Facebook files” to light, the documents revealed internal research on how certain users were vulnerable to harmful messaging, from eating disorders to misinformation about COVID-19. Given the pervasiveness of social media, and potential for damage to social institutions and individuals, we can't depend on whistleblowers to divulge information about platforms or rely on platforms to research themselves. Independent researchers provide an important check by conducting impartial analysis about platforms.

What is sensitive data, and why do researchers need access to it?

Much of the information researchers need to study harmful online misinformation can be made available safely to the public, through universal ad transparency and certain types of high engagement content. However, access to some types of data worthy of outside study requires additional layers of protection to ensure user privacy. Think of these as “sensitive data.” This is information that does not contain any personally identifiable information, that is, it does not include information about specific users, such as names or photos. But, it may contain elements that could potentially be used to identify users if combined with other data sources.

“Clean rooms” provide access to sensitive data while protecting users

“Clean rooms” are a way of providing vetted researchers access to sensitive data in a controlled environment. Often there is a mechanism for approval of certain researchers or research projects, who are given access to information but must conduct analysis in a closed environment—generally they are not permitted to download the data. Researchers are monitored closely and are not permitted to make data public without a privacy review. Some platforms voluntarily provide this type of access to their advertisers, and some companies, such as Meta, have begun providing limited information to approved researchers. This type of data sharing model is also widely used in other industries, such as finance and biomedical data.

Proposals containing clean room provisions

Researchers are calling for mandates for “clean room” access to platform data, rather than relying on platforms’ voluntary actions. Such provisions are included in transparency proposals before Congress:

- The Platform Accountability and Transparency Act would give authority to the National Science Foundation (NSF) to approve proposals submitted by university-affiliated researchers to study social media platforms.
- The Digital Services Oversight and Safety Act creates a new office within the Federal Trade Commission that would provide certification to researchers who apply and set rules for what information they would access and how.

Clean room data access, as defined in Platform Accountability and Transparency Act

The Platform Accountability and Transparency Act provides a model for clean room data access for vetted researchers. Under the proposal, researchers and projects would be subject to a number of conditions designed to protect user privacy and safety. These include:

1. “Qualified researchers” must be affiliated with universities and specifically identified in a research proposal that is approved by the National Science Foundation (NSF).
2. Research proposals must be approved by an institutional review board at a qualified researcher’s affiliated university. These review boards are charged with protecting human research subjects.
3. The NSF would be required to publish its research criteria for qualified research projects, and will also, as part of its review of proposals, consider cybersecurity and privacy risks. Approved projects will be provided data that meet safeguards for data protection.
4. Researchers must submit a pre-publication version of their research to the relevant platforms and to the FTC for review to confirm the analysis does not expose personal information, trade secrets, or confidential commercial information.

Source: Platform Accountability and Transparency Act

For more information about NYU Cybersecurity for Democracy:
info@cybersecurityfordemocracy.org | cybersecurityfordemocracy.org | Twitter: @cyber4democracy