Federal Register Notice 87 FR 15274, https://www.federalregister.gov/documents/2022/03/17/2022-05683/request-for-information-on-federal-priorities-for-information-integrity-research-and-development, May 15, 2022

Request for Information on Federal Priorities for Information Integrity Research and Development

Mozilla

Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO) and National Science Foundation (NSF)

RESPONSE OF MOZILLA TO REQUEST FOR INFORMATION ON FEDERAL PRIORITIES FOR INFORMATION INTEGRITY RESEARCH & DEVELOPMENT

Table of Contents

1. About Mozilla

Mozilla is a unique public benefit organization and open source community formed as a non-profit foundation in the United States. We have a strong reputation for our commitment to ensuring that privacy and security are fundamental to the internet. This is one of our guiding principles that recognises, among other things, that the internet is integral to modern life; the internet must remain open and accessible; security and privacy are fundamental; and that a balance between commercial profit and public benefit is critical.[1] These principles, in addition to our Data Privacy Principles[2], provide the basis for the way we develop products, manage the consumer data we collect, how we select and interact with partners, and how we shape our public policy and advocacy work.

Our Public Mission & Incentives

Mozilla's story originated in 1997 with Netscape Navigator, the original consumer browser and a popular browser of the 1990s. In a historic move for competition, Netscape publicly released its new browser engine (called "Gecko") under an open source license to enable others to verify, improve, and reuse the source code in their own products. The company was later the subject of the failed acquisition strategy of a powerful digital gatekeeper, when AOL purchased it in 1999. Although Netscape did not last following its acquisition by AOL, its open source browser engine Gecko has continued to shape the internet.

The non-profit Mozilla Foundation was created in 2003 to continue work on open source browser technology and with a larger mission to preserve the open internet. Firefox v1.0 was released in 2004 using Gecko with volunteer open source code contributions from around the world, and it was one of the first major consumer facing products to be built in this way using open source methodology. Today localization developers continue to make Firefox available in local languages and with local customizations for their communities to access the internet. Other developers have forked the Firefox codebase and used the Gecko browser engine to create new browsers with different features. The most well known example is Tor, an anonymity browser frequently used by journalists and human rights activists. While it has officially been blocked in Russia,[3] reliance on Tor has increased recently as a means to gain access to the open internet.[4]

In 2005, the Mozilla Foundation created a wholly-owned taxable subsidiary, the Mozilla Corporation, to serve its public mission through open source technology and product

[1] Mozilla's 10 Principles, https://www.mozilla.org/about/manifesto/
[2] Mozilla's Data Privacy Principles, https://www.mozilla.org/en-US/privacy/principles/
[3] Maria Xynou, Arturo Filastò. Russia Started Blocking Tor. OONI, December 17, 2021. https://ooni.org/post/2021-russia-blocks-tor/
[4] Sam Schechner and Keach Hagey. Russia Rolls Down Internet Iron Curtain, but Gaps Remain. WSJ, March 12, 2022. https://www.wsj.com/articles/russia-rolls-down-internet-iron-curtain-but-gaps-remain-11647087321

development of Firefox. In addition to remaining the sole shareholder of the Corporation, the Foundation advocates for better privacy, trustworthy AI, and digital rights and runs philanthropic programs in support of a more inclusive internet. These programs currently include fellowships and awards that invest in community leaders who are developing technology, policy, education and norms that will ultimately protect and empower people online.

2. Greater Transparency into Hidden Harms

A great amount of harm, including but not limited to the effects of disinformation, happens on major tech platforms outside the view of regulators and the public. These platforms offer highly sophisticated targeting tools that allow peddlers of disinformation to narrowly segment their audience, tailor content accordingly, and reach people most susceptible to their messages. Each person has their own individualized, potentially misleading experience.

This highly personalized experience means that harm enabled by platforms through their targeting systems is not easily identified by regulators, watchdog groups, or researchers. Because the experience is so personalized, harm can only be shown anecdotally, rather than systematically. There is dangerously little insight into what people experience, what ads are presented to them and why, and what content is recommended and why. This creates an asymmetry of information between those who produce disinformation and those seeking to understand it.

To address this, we need greater access to platform data (subject to strong user privacy protections), greater research tooling, and greater protections for researchers. This is why Mozilla has invested in building tools for researchers and why we support legislative solutions to provide greater insight into online disinformation, discrimination, and deception currently hidden from the public and from regulators.

The remainder of this submission discusses these gaps in more detail and is responsive to specific questions in the Request for Information concerning key research challenges, barriers for conducting information integrity R&D, and support for technological advancement in the fields of measurement and research platform development.

3. Robust Research Platforms are Essential to Understanding the Information Ecosystem & Protecting Information Integrity

The ability to understand public life on the internet is largely concentrated in the hands of private actors. These for-profit companies have a vested interest in perceptions of public life online. Such interests raise questions from external researchers, among many others, about disinformation on platforms and whether such platforms fairly disclose information that ensures better accountability.

Unfortunately, the data and tools made available by major platforms to understand topics like threats to information integrity remain inadequate. Most of the voluntary public-facing measures by major platforms have failed. Researchers have found, for example, that ad transparency tools are often nearly unusable.[5] At the same time, major companies continue to threaten legal action against good faith security research into disinformation.

We need far more robust research tools along with a deep pool of subject matter experts capable of taking advantage of these tools. More specifically, the software projects and infrastructure necessary to understand online public life require significant domain expertise, technical expertise, and capital to build and maintain. Efforts like the Markup's CitizenBrowser project, or New York University's Ad Observer have noted their substantial startup costs, as well as ongoing operational costs. These barriers to entry consequently mean that the capability to study online public life remains accessible to few organizations.

The research produced by these teams has begun to shift the understanding and perceptions of legislators, regulators, and the public. New regulations across the globe, such as the General Data Protection Regulation (GDPR)[6], California Consumer Privacy Act (CCPA)[7], and the Digital Services Act (DSA)[8], are starting to move the broader data economy to center on informed user consent for data collection, and user control of the data collected from them.

Mozilla sees an opportunity to lower barriers for researchers to study and understand life online, and to provide users and the public with consent-driven approaches to exchanging their data.

Mozilla's Rally Project - What it is and How it Works

To explore these opportunities, Mozilla launched Rally, an opt-in platform for consumers to donate their data to researchers and causes they support. Mozilla envisions Rally as a platform through which users can affirmatively control their data and how it's used.

In its pilot phase, Rally has worked with journalists, academics, and non-profit researchers to develop and launch projects to understand the public's experiences online through user contributed browser and interaction data. The browser is a critical tool to allow people to navigate the online world. It can therefore provide significant insight into what people are experiencing online, what information they are consuming,

[5] Laura Edelson. Facebook's political ad spending numbers don't add up. Medium, October 12, 2020. https://medium.com/online-political-transparency-project/transparency-theater-facebooks-political-ad-spending-numbers-don-t-add-up-d7a85479a002

[6] The EU General Data Protection Regulation, https://gdpr.eu/

[7] The California Consumer Privacy Act, https://oag.ca.gov/privacy/ccpa

[8] The EU Digital Services Act, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

and how they are being manipulated, as long as that data can be collected in a privacy respecting manner. Rally aims to do just that, taking advantage of instrumentation tools already in the browser to potentially power research and studies by reputable parties outside Mozilla.

Current Rally studies[9] include one focused on "Political and COVID-19 News" across the web. This study can help us understand how web users are exposed to, consume, and share these types of information, which can inform efforts to distinguish trustworthy and untrustworthy content. Another study focuses on local news. The results will help build our understanding of how the modern news environment works, and which alternative funding models for local journalism may be feasible. These studies demonstrate the potential power of research platforms like Rally to contribute to information integrity R&D.

Users can sign up for Rally, select the projects they would like to contribute data to, and are prompted to install a browser extension in order to participate. Projects launched on Rally are reviewed by the Mozilla Rally team in order to vet the design and implementation of the studies so that users can have the confidence that Rally studies adhere to Mozilla's data practices.

The web extensions built for Rally projects rely on Rally's software development toolkits to ensure that data measured on a user's machine is encrypted and transmitted securely to the Mozilla data platform. From there, data is placed in each project database, where only a restricted list of researchers and Mozilla Rally staff are permitted access. Researchers then access and run analyses on row level user data on the Mozilla data platform.

4. Vetting Qualified Researchers & Proposals

For research platforms like Rally, granting access to the *right* parties is important to ensuring our tools are not abused and users are not placed at risk. Currently the burden is on the Rally team to determine what researchers to work with, what credentials researchers need in order to gain access to our platform, and what specific research proposals are in the public interest. This is a responsibility we take seriously. To make this determination, we review both the researchers and their proposed studies. Unfortunately, this level of diligence can be cumbersome, limiting our ability to offer Rally to a large number of researchers.

This diligence burden currently must be independently borne by each research platform seeking to support work on information integrity. Further, research vetting may present prohibitive barriers to other companies or other research platforms that lack Mozilla's

[9] Mozilla Rally current studies, https://rally.mozilla.org/current-studies/

technical expertise, resources, and mission focus, making them less willing to take on this burden, less likely to offer research tools, and more likely to make mistakes when they do so.

Unlocking the potential of solutions like Rally and allowing them to offer tooling at scale to a diversity of researchers requires relieving this burden on individual parties and companies. These tools are potentially a shared resource and the diligence burden ideally could be shared rather than duplicated. To address this need, there should be an independent or governmental body that can take on responsibility for this vetting. Current proposals in Congress[10] envision just such an approach, asking the National Science Foundation (NSF) to establish a process to solicit research applications and vet qualified researchers. We encourage NSF to explore creating such a program, even absent a legislative mandate. At a minimum, the Federal government through the Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO) and the NSF could work to create consensus standards that companies like Mozilla adopt and use.

Such an adjudicative body or standards need to address the following two areas:

- **Researcher Vetting.** Currently Mozilla's Rally platform is available to a very small number of researchers, mostly at marquee universities, allowing us to use university reputation as a proxy for researcher credibility. While this approach helps manage risk effectively, it has obvious shortcomings and limits the number of researchers that may benefit from Rally. We encourage the NSF and NITRD NCO to explore how to establish some type of standard or credentialing process, one not strictly limited by university affiliation, to vet researchers prior to their application to the research platform. Such a credential would be an incredibly valuable signal that Mozilla and others could use as criteria to determine who gains access.

- **Public Interest Value, Study Ethics & Methods**. Separate from researcher vetting, there must be standards established for research proposals themselves to ensure they are ethically designed and would contribute to the public interest. We recognize that, despite Mozilla's long track record of work to build a healthier Internet, we should not be the sole arbiter of what constitutes public interest research. Nor are we necessarily the right party to judge the ethics of the studies or the soundness of their scientific methods. This is a function better served by mechanisms like Institutional Review Boards (IRBs), which are today available to university affiliated researchers. There should be an independent evaluation

[10] Platform Accountability and Transparency Act (PATA), https://www.coons.senate.gov/imo/media/doc/text_pata_117.pdf

process that can assess both the soundness of the method and the public interest value of the research.

5. Policy Tools to Address Threats to Information Integrity

Robust research platforms are not, however, enough. Legislative or regulatory action is needed to complement and support independent research into disinformation. Recent history has shown that major tech platforms do not have sufficient incentive to provide the necessary level of transparency and access to researchers and that they must be required to do so.[11] To that end, Mozilla's policy advocacy focuses on two particular areas: researcher safe harbor and ad transparency. We strongly support provisions of the European Union's Digital Services Act[12] that would mandate disclosure of all ads on tech platforms. We are similarly encouraged by recent Congressional proposals that seek to provide accountability and transparency into the real world impact of disinformation, such as the Platform Accountability and Transparency Act (PATA)[13], which would provide important protections for researchers and require disclosure of ads.

Mandating a Safe Harbor to Protect Public Interest Researchers

Public interest research is key to shedding light on hidden harms and disinformation. Experts engaged in this research need to be protected. Accordingly, Mozilla has called for a safe harbor to allow researchers, journalists, and others to study disinformation and access relevant datasets, free from threat of legal action.

Mozilla often hears of researchers who are concerned that companies or governments may take legal action against them for their legitimate research – including civil or criminal penalties under laws such as the Computer Fraud and Abuse Act (CFAA), violations of Terms of Service, and more. Facebook, for example, has blocked research tools and threatened legal action against researchers seeking to investigate election integrity and misinformation online.[14]

These actions by platforms not only put researchers themselves at legal risk, but also stifle vital transparency into real world harm by deterring individuals and institutions doing critical work. Indeed, research tools and initiatives most vital to public interest -

[11] Marshall Erwin. Why Facebook's claims about the Ad Observer are wrong. Mozilla blog, August 2021. https://blog.mozilla.org/en/mozilla/news/why-facebooks-claims-about-the-ad-observer-are-wrong/

[12]Owen Bennett. Mozilla publishes position paper on EU Digital Services Act. Mozilla Open Policy & Advocacy blog, May 18, 2021. https://blog.mozilla.org/netpolicy/2021/05/18/mozilla-publishes-position-paper-on-eu-digital-services-act/

[13] PATA, https://www.coons.senate.gov/imo/media/doc/text_pata_117.pdf

[14] Jeff Horwitz. Facebook Seeks Shutdown of NYU Research Project Into Political Ad Targeting. WSJ, October 23, 2020. https://www.wsj.com/articles/facebook-seeks-shutdown-of-nyu-research-project-into-political-ad-targeting-11603488533

most capable of identifying patterns of harm or threats to information integrity on major tech platforms - may receive the greatest scrutiny and be subject to the greatest legal exposure. This is the pattern we have seen with New York University's Ad Observatory project, which has offered research tools effective at identifying harms on tech platforms, and as a result, has had to withstand sustained legal attack.

The risk of legal exposure, both for ourselves and our research partners, is something that Mozilla must be mindful of when offering a research platform like Rally. In Mozilla's case, we have the legal expertise and resources such that we cannot be intimidated by spurious legal threats from major platforms. This is unlikely to be the case, however, for many of our potential research partners.

A safe harbor would protect and promote research in the public interest as long as researchers handle data responsibly and adhere to professional and ethical standards, such as those developed to support the vetting process described above. There is enormous value this can provide to the public. Mozilla has one of the earliest Bug Bounty programs in software. We make clear that we will not threaten or bring any legal action against anyone who makes a good faith effort to comply with our vulnerability notification policy because this encourages security researchers to investigate and disclose security issues. Their research helps make the internet a safer place.

Ad Transparency to Combat Hidden Harm & Disinformation

Advertisements are a critical vehicle for dissemination of disinformation. Research has found ad targeting helps peddlers of disinformation find and build their initial audience in the early stages of a disinformation campaign. Once that audience has been created, organic growth then takes over.

The public and regulators need far greater access to data about ads and ad targeting. To address this gap, Mozilla supports Executive or Congressional action requiring social media advertisers and companies to publicly disclose the ads and targeting criteria appearing on social media platforms. Ad disclosure should include the following:

- Apply to all advertising, so as not to be constrained by arbitrary boundary definitions of 'political' or 'issue-based' advertising;
- Include disclosure obligations that concern advertisers' targeting parameters for protected classes as well as aggregate audience demographics, where this makes sense given privacy and other considerations;
- Provide Broad Access to Data. Data should be broadly available to regulators, researchers, journalists, and watchdog groups, rather than restricting access to privileged stakeholders.

Previous regulatory initiatives aiming at ad transparency to combat disinformation have generally focused on ‘political’ advertising. Focusing on purely ‘political’ advertising (e.g. advertising copy developed by political parties) is too narrow and is insufficient to capture the complex web of actors involved in politically-motivated disinformation online. Moreover, a broad ads disclosure framework could also drive transparency with respect to what is known as ‘issue’ advertising. Experience has shown how disclosure obligations that include this broader category of political ads put platforms in a challenging position, requiring them to decide what is 'political' in nature, which can vary depending on context, jurisdiction, and time. A focus on all ads would negate this line-drawing challenge.

Further, the inclusion of all ads allows for the identification and analysis of other forms of systemic harm that may be occurring in the current ad ecosystem. Indeed, other types of advertising that are not overtly political in nature may nonetheless be deceptive or may be targeted in a way that discriminates towards particular groups.

6. Support for Technological Advancement

Building projects to understand online public life requires significant domain expertise, technical expertise and start up and operational capital. This is not something that individuals with deep subject matter expertise on topics like information integrity can necessarily do. Our work at Mozilla seeks to reduce barriers to entry in this research space through projects such as Rally.

Growing the overall research ecosystem requires complementary efforts to support and catalyze researchers and domain experts as they invent novel ways of measuring and understanding online life.

This requires far more extensive support from organizations like the NITRD NCO and the NSF. In particular, in addition to the best practices and researcher vetting described above, we encourage you to consider how grantmaking, funding, and programs can help build and maintain measurement systems and platforms capable of supporting the researcher ecosystem. Government engagement on issues like disinformation and threats to information integrity can be fraught, raising challenging issues regarding free expression and speech. Funding support for research platforms however sidesteps these challenges and doesn’t require the government to weigh in on particular disinformation topics. This is a unique, important role that you can play to advance work on this topic.

7. Conclusion

Mozilla is encouraged that the agencies have undertaken the process of reviewing questions pertaining to Federal priorities for research and development efforts to

address misinformation and disinformation. It is essential to develop policies and tools to foster transparency and trust in the online ecosystem. The issues addressed in this paper reflect Mozilla's perspectives and recommendations in key areas. They are not intended to be exhaustive and we would be happy to provide additional detail or further information if helpful.