Date: January 15, 2025
To: National Telecommunications and Information Administration
Re: Ethical Guidelines for Research Using Pervasive Data

Building prosperity through innovation relies on building understanding through research, and ethical guidelines preserve the compact of trust between researchers and society as a whole. We thank the NTIA for investigating this issue through its notice and request for public comments on Ethical Guidelines for Research Using Pervasive Data of Dec. 11, 2024 ("the RFC") and for the opportunity to share our perspective.[1]

Public Knowledge promotes freedom of expression, an open internet, and access to affordable communications tools and creative works. We work to shape policy on behalf of the public interest, and fight for a creative and connected future for all. While Public Knowledge does not have specific expertise on research ethics or research practices, we write in support of developing voluntary research guidelines for the use of pervasive data for three key reasons:

1. We care deeply about the privacy rights of users, and the potential harms that can emerge from irresponsible or unethical use of pervasive data.

2. We are concerned about a growing erosion of trust in the open, public environments of the internet.

3. We support frameworks to govern ethical and responsible use of public data separate from overly restrictive, economically-motivated intellectual property (IP) regimes.

Protecting User Privacy

The NTIA's RFC rightfully underscores the critical need to safeguard user privacy when utilizing pervasive data for research purposes. Public Knowledge emphasizes that protecting user privacy is not merely a matter of individual rights but a cornerstone of maintaining societal trust in digital environments. Irresponsible or unethical uses of pervasive data can lead to significant harms, including exploitation of personal information, discriminatory practices, and chilling effects on online expression. These risks underscore the urgent need for ethical guidelines to ensure that privacy protections are meaningful and robust.

To this end, Public Knowledge supports the development of voluntary guidelines that place privacy at the center of research involving pervasive data. Such guidelines should prioritize transparency, minimize unnecessary data collection, and respect user consent and expectations.

---

[1] While we appreciate the notice and comment opportunity, the timing of the RFC has made contributing to this call challenging, and other resource-constrained public interest stakeholders may have missed this opportunity entirely. We hope the NTIA will, as it stated in the RFC, provide other opportunities for stakeholders to engage on this topic.

Additionally, they must explicitly address the potential for cumulative societal harm arising from unchecked data practices, such as eroding trust in the open internet and exacerbating inequities.

We further urge that these guidelines be developed in alignment with broader privacy principles that protect users from harm while preserving the open and collaborative nature of the internet. By taking these steps, the NTIA can help establish a framework that not only safeguards individual privacy but also strengthens the trust and integrity needed for ethical research in the digital age.

Preserving the Open Internet

The RFC highlights a specific harm that is of particular concern: the potential for "decreased willingness to post or access information online." While protecting privacy is fundamentally oriented around preventing harms to individuals, there are also collective risks to irresponsible and unethical data usage.

The creation of the Internet, World Wide Web, and the explosion of user-generated content on online platforms has given rise to an unprecedently vast, spectacular, shared body of human knowledge and creativity. Yet, the surveillance and targeted advertising business model of online platforms, the use of publicly accessible content for commercial generative AI training, online harassment and stalking, and forces like them are creating growing concerns about the future of the open internet. Without proper stewardship, we are at great risk of eroding the shared commons of the Internet.

Ethical guidelines for the use of pervasive data presents an opportunity to restore and renew some of the trust needed to continue in the great shared experiment that is the Internet. Research guidelines should take into consideration user expectations, be oriented towards public benefit, and be rooted in the values and principles of openness, sharing, and reciprocity that are foundational to our digital culture. As discussed further below, guidelines should also balance protections for users and the commons by properly taking into account the public and open nature of some forms of pervasive data; researchers should not be unduly restricted from taking advantage of publicly accessible data, lest we lose the open and public character that makes our digital landscape so valuable.

Perils of Intellectual Property Expansion

Though it was not specifically addressed in the RFC, we feel it is important to explicitly address the overlap between user rights, legal obligations, and intellectual property regimes. Copyright law has often been weaponized to protect privacy or control information, even where it may not

be properly applicable.[2] Copyright is a particularly effective bludgeon online thanks to the Digital Millennium Copyright Act (DMCA), and this creates an emphasis on "ownership" and other property-like attitudes towards personal data. This perspective is a dangerous one to import into research ethics because copyright and IP legal regimes must be understood in fundamental tension with the First Amendment, and are intended as economically-focused doctrines designed specifically to incentivize the creation of scientific and creative works.

When developing ethical research guidelines, it is appropriate to consider how user expectations ought to be factored into the use of user data, however the attitudes seeping from IP into the discourse about the use of data in research should be weighed very carefully. The desire for commodification of every aspect of data--whether that data is created intentionally or not, for economic purposes or not--is a symptom of steadily increasing copyright expansionism that is not founded in the law or the core purpose of copyright. The NTIA should be mindful of these perspectives, even when developing voluntary guidelines, so as not to inadvertently contribute to these misconceptions and harmful attitudes. Emphasis should be centered more on privacy rights, expectations of privacy and use, and user protection.

Conclusion

Public Knowledge appreciates the NTIA's thoughtful engagement on the important issue of ethical research using pervasive data and the opportunity to contribute to this dialogue. By prioritizing user privacy, fostering trust in digital spaces, and supporting open and collaborative research practices, the NTIA can help ensure that ethical guidelines reflect both individual rights and the collective value of the internet. We encourage the NTIA to continue its leadership in addressing these challenges and look forward to future opportunities to collaborate on advancing these critical principles.

NICHOLAS P. GARCIA
Senior Policy Counsel
Public Knowledge

---

[2] See, e.g.
https://www.techpolicy.press/new-research-highlights-xs-failures-in-removing-nonconsensual-intimate-media/;
https://cdt.org/wp-content/uploads/pdfs/copyright_takedowns.pdf; https://www.eff.org/takedowns