

DAY ONE PROJECT

Ensuring Platform Transparency and
Accountability

Brandie M. Nonnecke

December 2020

The Day One Project offers a platform for ideas that represent a broad range of perspectives across S&T disciplines. The views and opinions expressed in this proposal are those of the author(s) and do not reflect the views and opinions of the Day One Project or its S&T Leadership Council.

Summary

Open-source investigations and public interest research using platform data (e.g., Facebook, YouTube) have enabled the collection of evidence of human rights atrocities, identified the role of foreign adversaries in manipulating public opinion before elections, and uncovered the prevalence and reach of terrorist radicalization and recruitment tactics.¹

However, it is becoming increasingly difficult for independent investigators and researchers to access platform data. Nascent data privacy legislation such as the EU General Data Protection Regulation (EU GDPR) and the California Consumer Privacy Act (CCPA) have placed increased pressure on platforms to restrict third party access to data. In an overly cautious interpretation of these laws, platforms are increasingly restricting third-party access to the data they collect.² In doing so, platforms shield themselves from public scrutiny and accountability.

To support transparency and accountability of platforms, the next administration should work with Congress to ensure that any new data privacy legislation proposed at the federal level does not inadvertently block the ability of third parties to gain access to platform data for open-source investigations and public interest research. Additionally, industry-wide principles and operational guides for the appropriate sharing and use of platform data for open-source investigations and public interest research are needed. The White House Office of Science and Technology Policy should take the lead by convening a workshop among key actors to make progress on these goals. Out of the workshop, a federal working group should be formed to develop principles and operational guides to support ethical third-party access to platform data, including the formation of technical standards to ensure data privacy and security.

Challenge and Opportunity

Platforms' restrictions on third-party data access not only diminishes the possibility of holding those platforms accountable for their effects on society through public interest research, but it also restricts the collection of evidence in open-source investigations.

¹ Aliaume, Leroy and Ben Strick. 2018. "Cameroon: Anatomy of a Killing." BBC, Sept. 23, 2018. https://www.youtube.com/watch?v=XbnLkc6r3yc&feature=emb_logo; Johnson, Carolyn Y. 2018. "Russian Trolls and Twitter Bots Exploit Vaccine Controversy," The Washington Post, Aug. 23, 2018. <https://www.washingtonpost.com/science/2018/08/23/russian-trolls-twitter-bots-exploit-vaccine-controversy/>; Timberg, Craig and Isaac Stanley-Becker. 2020. "Violent Memes and Messages Surging on Far-Left Social Media, A New Report Finds," The Washington Post, Sept. 17, 2020. <https://www.washingtonpost.com/technology/2020/09/14/violent-antipolice-memes-surge/>.

² Human Rights Watch. 2020. "'Video Unavailable' Social Media Platforms Remove Evidence of War Crimes," Oct. 4, 2020, <https://www.hrw.org/report/2020/09/10/video-unavailable/social-media-platforms-remove-evidence-war-crimes>; Silverman, Craig. 2019. "Exclusive: Funders Have Given Facebook a Deadline to Share Data with Researchers or They're Pulling Out." BuzzFeedNews, Aug. 27, 2019. <https://www.buzzfeednews.com/article/craigsilverman/funders-are-ready-to-pull-out-of-facebooks-academic-data>.

Public interest research conducted on platform data has enabled greater understanding of the role and effects of platforms on society. For example, research has uncovered the spread of misinformation and disinformation and their influence on public health and safety, the prevalence and effects of exposure to radicalizing content, and the effects of targeted political advertisements on public opinion and voting behavior.³ These research activities are facing greater opposition from platforms, with some researchers receiving cease-and-desist letters to discontinue collecting data.⁴

Open-source investigations, which collect publicly available data often hosted by platforms, have been used widely in human rights investigations and have enabled the documentation of human rights atrocities in Syria, Cameroon, and Sudan.⁵ Unfortunately, platforms are increasingly removing such evidence from public distribution and may be deleting it, undermining the ability of investigators to collect evidence and hold perpetrators accountable.⁶

Under pressure to ensure the privacy and security of data, platforms will likely continue to put in place more restrictive data-sharing policies.⁷ While protecting the privacy and security of data is important, so too is the need to access the data that platforms hold for public interest research and human rights investigations.

The California Privacy Rights Act (CPRA), which was passed in November 2020, has tried to strike this balance. The CPRA includes language allowing third-party access to private sector-held data for research purposes. To protect consumers' data, the CPRA includes requirements for consumer approval of data collection and use for research, implementation of de-identification techniques, technical safeguards to restrict re-identification, and security controls that limit who can access the data.⁸ Any newly introduced federal data privacy laws should incorporate the same standards.

³ Johnson, Carolyn Y. 2018. "Russian Trolls and Twitter Bots Exploit Vaccine Controversy," *The Washington Post*, Aug. 23, 2018. <https://www.washingtonpost.com/science/2018/08/23/russian-trolls-twitter-bots-exploit-vaccine-controversy/>; Timberg, Craig and Isaac Stanley-Becker. 2020. "Violent Memes and Messages Surging on Far-Left Social Media, A New Report Finds," *The Washington Post*, Sept. 17, 2020, <https://www.washingtonpost.com/technology/2020/09/14/violent-antipolice-memes-surge/>; NYU Online Political Transparency Project. n.d. "NYU Ad Observatory." Accessed Oct. 26, 2020. <https://adobservatory.org/>.

⁴ Horwitz, Jeff. 2020. "Facebook seeks shutdown of NYU Research Project into political ad targeting," *The Wall Street Journal*, Oct. 23, 2020. <https://www.wsj.com/articles/facebook-seeks-shutdown-of-nyu-research-project-into-political-ad-targeting-11603488533>.

⁵ O'Flaherty, Kate. 2018. "YouTube keeps deleting evidence of Syrian chemical weapon attacks." *Wired*, June 26, 2018, <https://www.wired.co.uk/article/chemical-weapons-in-syria-youtube-algorithm-delete-video>;

Aliaume, Leroy and Ben Strick. 2018. "Cameroon: Anatomy of a Killing." *BBC*, Sept. 23, 2018.

https://www.youtube.com/watch?v=XbnLkc6r3yc&feature=emb_logo; BBC. 2019. "Sudan's Livestream Massacre." July 11, 2019, https://www.youtube.com/watch?v=dR56qxM4kHA&feature=emb_logo.

⁶ Human Rights Watch. 2020. "'Video Unavailable' Social Media Platforms Remove Evidence of War Crimes," Oct. 4, 2020, <https://www.hrw.org/report/2020/09/10/video-unavailable/social-media-platforms-remove-evidence-war-crimes>.

⁷ Perez, Sarah. 2018. "Facebook rolls out more API restrictions and shutdown." *Tech Crunch*, July 2, 2018.

<https://techcrunch.com/2018/07/02/facebook-rolls-out-more-api-restrictions-and-shutdowns/>.

⁸ California Privacy Rights Act of 2020. Accessed Dec. 15, 2020. https://oag.ca.gov/system/files/initiatives/pdfs/19-0021A1%20%28Consumer%20Privacy%20-%20Version%203%29_1.pdf.

Plan of Action

The next administration should work with Congress to ensure that any newly proposed data privacy legislation targeting platforms does not overly restrict third-party data access for public interest research and open-source investigations.

To achieve this goal, proposed federal data privacy legislation should include:

- Language to enable third parties to access platform data for public interest research and open-source investigations, while guaranteeing data privacy and security.

To ensure data privacy and security, standards should be established to guide the:

- De-identification of data before granting third-party access;
- Implementation of technical safeguards to restrict re-identification of data; and
- Development of security controls that, in cases where sensitive data are used, limit and verify who can access the data.

The development of appropriate policies and data privacy and security standards will require collaboration between industry, government, civil society, academia, and other stakeholders. The next administration should start by convening a multi-stakeholder workshop and working group through the White House Office of Science and Technology Policy. The workshop will identify opportunities for collective action, such as establishing shared commitments among stakeholders to develop governance and technical processes that support ethical third-party data access. A working group can then be established to develop principles and technical guidance for the appropriate collection, storage, and sharing of platform data such that data privacy and security are ensured while supporting public interest research and open-source investigations.

Conclusion

Access to platform data has been critical to supporting open-source investigations documenting human rights abuses and public interest research on the effects of platforms on society. The next administration should ensure that any new data privacy legislation proposed at the federal level does not inadvertently block this impactful work. Additionally, the White House Office of Science and Technology Policy should convene a multi-stakeholder workshop and working group to establish industry-wide principles and operational guides for the appropriate sharing and use of platform data in a manner that ensures data privacy and security.

Frequently Asked Questions

How do we ensure researchers don't use platform data for nefarious purposes (e.g., the Cambridge Analytica scandal)?

Data made available through application programming interfaces (APIs) will fall under platforms' terms of service, and data made available through contractual arrangements will also have legal limitations. These terms of service and contractual agreements will specify prohibited practices, such as selling, re-identifying, or using data to discriminate.

How can platforms support public interest research and open-source investigations?

Platforms should continue to make anonymized data available through their public application programming interfaces (APIs). The data should be shared in a manner that preserves privacy and incorporates technical safeguards to restrict re-identification. In addition to APIs, platforms should create user-friendly dashboards that facilitate public engagement and access to data for public interest research and open-source investigations.

Why would industry participate?

Platforms have a history of making data available as a means to support complementary services and research. Their participation in the workshop and working group will enable them to tap into research expertise to evaluate and inform their services, guide their development of technical safeguards for data sharing, and support public trust by actively engaging in the development of procedures that better ensure transparency and accountability in their products and services.

DAY ONE PROJECT



About the Author

Brandie Nonnecke, PhD is Founding Director of the CITRIS Policy Lab, headquartered at UC Berkeley. Brandie has expertise in tech policy and internet governance. She is the founder of the Public Interest Research Alliance (PIRA), a multi-stakeholder, non-binding coalition committed to the establishment of shared principles and operational guides for the appropriate collection, storage, and use of platform data for public interest research. She is a Technology and Human Rights Fellow at the Carr Center for Human Rights Policy at the Harvard Kennedy School and recently served as a fellow at the Aspen Institute's Tech Policy Hub and at the World Economic Forum on the Council on the Future of the Digital Economy and Society. Her research publications, op-eds, and presentations are available at nonnecke.com.



About the Day One Project

The **Day One Project** is dedicated to democratizing the policymaking process by working with new and expert voices across the science and technology community, helping to develop actionable policies that can improve the lives of all Americans, and readying them for Day One of the next presidential term. For more about the Day One Project, visit dayoneproject.org.