Testimony of Professor Nathaniel Persily

James B. McClatchy Professor of Law

Co-Director of the Stanford Cyber Policy Center

Stanford Law School[1]

Before the United States Senate Committee on the Judiciary

Subcommittee on Privacy, Technology, and the Law

"Platform Transparency: Understanding the Impact of Social Media"

Submitted May 2, 2022

Thank you, Mr. Chairman and Members of the Committee, for inviting me today to testify on the need for greater transparency from the internet platforms. My name is Nate Persily. I am the James B. McClatchy Professor of Law at Stanford Law School and Co-Director of the Stanford Cyber Policy Center. Perhaps most notably for purposes of this hearing, I was also the cofounder of Social Science One, an effort to get internet platforms, such as Facebook, to share privacy-protected data with outside researchers.

I want to use my remarks today to explain the purposes greater transparency serves, the types of transparency that are necessary, and the reason that only

[1] Affiliation for identification purposes only; appearing in personal capacity.

governmental regulation – not voluntary efforts by the companies – can ensure the benefits of transparency while protecting user privacy. **We cannot live in a world where Facebook and Google know nearly everything about us, and we know next to nothing about them. These platforms have lost their right to secrecy, and it is well past the time that someone other than the firms' own data scientists be granted access to the data that reveal the impact of these platforms on the information ecosystem**.[2]

It has become fashionable recently to describe these platforms as the global public square. I happen to disagree with that characterization, because these platforms, by design, organize and prioritize some kinds of communication over others and structure their environments to create a specific user experience. If anything, the internet itself is the public square, whereas the platforms are privately controlled spaces with specific rules for interaction. For those to whom the metaphor appeals, however, it usually is in service of an argument that these platforms should be more open and less censorious in their content moderation. That debate is important to have, and we should hope that different platforms adopt different algorithms and content moderation rules so that the

[2] For an expanded version of the arguments presented here, see Social Media Platforms and the Amplification of Domestic Extremism and Other Harmful Content: Hearing before the United States Senate Committee on Homeland Security and Governmental Affairs, Oct. 28, 2021 (testimony of Nathaniel Persily); Nathaniel Persily, "A Proposal for Researcher Access to Platform Data: The Platform Transparency and Accountability Act," 1 *Journal of Online Trust and Safety* (2021), https://doi.org/10.54501/jots.v1i1.22; Nathaniel Persily & Joshua A. Tucker, *How to fix social media? Start with independent research*, Brookings Inst., Dec. 1, 2021, https://www.brookings.edu/research/how-to-fix-social-media-start-with-independent-research/ ; Nathaniel Persily, Opening a Window Into Tech: The Challenge and Opportunity for Data Transparency, Stanford Cyber Policy Center (2020), available at https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/cpc-open_windows_np_v3.pdf; Nathaniel Persily & Joshua A. Tucker, "Conclusion: The Challenges and Opportunities for Social Media Research," in *Social Media and Democracy: The State of the Field and Prospects for Reform* (N. Persily & J. Tucker eds., 2020).

online marketplace of ideas can be a true market where users can "vote with their clicks" to opt into different sets of rules.

Whatever one might say of the metaphor, both the enforcement of the rules that govern the public square and the behavior of speakers within it should be transparent. In this respect, the large platforms fall short. The public does not have a window into the dynamics of speech regulation in these critically important spaces or how content spreads within and across them. The role of law in this area, at the very least, is to make the public square genuinely public: to allow for outsiders to have the same understanding as insiders with respect to how these platforms operate.

I. The Goals of Transparency

In discussions of potential platform regulations, transparency proposals are often viewed as less significant than other kinds of interventions, such as reforms of Section 230 of the Communications Decency Act. That assessment, however, understates the foundational role that transparency can play in a larger regulatory regime and how regulated industries adapt based on the information they are forced to disclose. Transparency proposals, such as the Platform Accountability and Transparency Act, will lead to fundamental change at the platforms, while at the same time educating policymakers considering other forms of tech regulation.

First and most important, greater transparency will affect the behavior of the firms themselves. When the platforms know that they are being watched and that outsiders will have access to the same data as insiders, they will inevitably change their policies and enforcement of them. Forcing disclosure means that content moderation or algorithmic manipulation cannot happen in the dark without consequences. Under a regime such as that proposed in PATA, the public will routinely be made aware of changes in policies and practices related to advertising, content moderation, and algorithmic design. If a platform's policies or affordances are responsible either for propagation of dangerous content or for undue censorship, the platform will not be able to hide it from public view.

Second, transparency is critical for the creation of good public policy. At present there are fundamental disagreements between outsiders and insiders as to the scale of online harms and the responsibility of the platforms for causing or amplifying them. Critics focus on the number of posts (or related likes or reposts) of problematic content (disinformation, hate speech, etc.), whereas platform defenders tend to argue that such content is a tiny share of the average user's feed. Moreover, it is now part of conventional wisdom that the algorithms and recommendation engines (let alone advertising) of the major platforms routinely amplify hate speech, disinformation, and other problematic content. However, most of the academic investigation of these topics suggests it is more common for people to seek out this problematic content rather than have it force-fed to them. If Congress or any legislature around the world is to make policy with respect to online content, it is critical that we have answers to these questions. We can only get answers if outsiders are given access to the same data from which the platform advocates draw their conclusions.

Third, the public has a right to know how these platforms operate and the scale and character of online harms. Much, if not most, of human interaction is now occurring through digital communication. It has become fashionable to describe these platforms as information monopolies, usually in service of an antitrust argument about the dangers of concentrated power over critical speech forums. But these platforms are information monopolies in a different sense: they, and they alone, hold the keys to the data that is most informative as to contemporary society. Not only are outsiders largely in the dark about the role that platforms play in exacerbating social problems, but the principal information about society itself is now in a few private hands. If we want to understand pretty much any contemporary problem – e.g., the dynamics of the pandemic, the causes and scale of teenage depression, the challenges to journalism, the depths of political polarization, foreign election interference – we need to have access to data that are locked up in these private companies. In short, transparency is about more than looking over the shoulders of the platforms; it is a necessary predicate to understanding most contemporary policy challenges "in the real world" as well.

II. Transparency for Whom and About What?

Transparency, like all values related to regulation of the information ecosystem, comes with tradeoffs. Unfettered transparency could lead to violations of user privacy, as well as gamesmanship by bad actors seeking to outflank platform countermeasures designed to protect users. The goal of platform transparency legislation should be to make public all that can be made public without compromising user privacy and without providing the tools for adversaries to game the platforms' rules and algorithms. As with any regulatory challenge, transparency can be done well or poorly, and it is the role of policy to confront these tradeoffs honestly with the goal of maximizing the benefits of transparency (outlined above) while minimizing its costs.

The proposed Platform Accountability and Transparency Act (PATA) seeks to do just that. Like any draft legislation, it is not perfect and needs refinement.[3] But it goes farther and provides more detail than any previous efforts. Most critically, it focuses on what I consider to be three legs to the transparency stool: (1) broad obligations for public disclosures; (2) protection for researchers analyzing publicly available data; and (3) supervised access for vetted researchers to the data accessible to the platforms own data scientists.

PATA requires large platforms to issue public reports or representative data on several critical topics. It gives the Federal Trade Commission power to require reports on widely shared content, including metrics for engagement and reach, as well as disclosure of the types of datasets in the platform's possession related to users or content. The proposal also requires the disclosure of advertising data, including the identity of advertisers, the ad content itself, and information as to targeting and reach. Finally, it requires disclosures related to algorithms and metrics: in particular, a summary of the inputs to the algorithms and the "optimization objective of such models."

Second, PATA immunizes researchers from civil and criminal liability for scraping of non-private data from platforms, if they meet a number of specified requirements. This provision responds to the controversy sparked by the NYU Ad

[3] In her submitted testimony, my colleague, Daphne Keller, highlights some areas of PATA in need of refinement. I agree with most of her arguments. However, all of these objections can be satisfied with modifications that do not undermine the critical requirements of the largest platforms to provide broad public disclosures, access of vetted researchers to platform data, and civil and criminal immunity for researchers doing public interested work on public platform data. One point of hers that I would like to stress is the importance of limiting these transparency obligations to the largest platforms – namely, Facebook/Meta, Google/YouTube, TikTok, and Twitter. We do not want the transparency obligations to inhibit the rise of potential competitors to these platforms or to place excessive burdens on small platforms that might be forced to shut down were they forced to comply.

Observatory's efforts to study Facebook advertising.[4] Researchers there provided a browser plugin to volunteers with Facebook accounts to capture information about the ads that the platform served to them. Citing concerns about privacy arising from its settlement with the FTC following the Cambridge Analytica scandal, Facebook disabled the accounts of the NYU researchers. Afterwards, the FTC took the extraordinary step of clarifying that the settlement did not prevent Facebook from allowing such researcher access. The sordid saga concerning the Ad Observatory only further reinforces the need for law to give direction as to what kinds of research are permitted or forbidden on these large platforms. Despite some recent court decisions in their favor, researchers continue to avoid pursuing certain projects for fear that they might be prosecuted or held liable under the Computer Fraud and Abuse Act.

Third, PATA provides a secure pathway for vetted researchers to gain access to data already in the firm's possession, which otherwise might only be analyzed by the firm's own data scientists. Some data, even in aggregated and anonymized form, might be too sensitive to be made public. A secure pathway for vetted researchers, with safeguards to protect the privacy of that information, is absolutely critical to answering the most pressing questions concerning the impact of platform policies on the information ecosystem.

Although the public reporting requirements will be quite useful to understanding certain phenomena, only secure access will allow research on "who" has seen/engaged with "what," "when," and "why." These are the critical inquiries undergirding any

[4] Laura Edelson & Damon McCoy, "We Research Misinformation on Facebook. It Just Disabled Our Accounts," *New York Times*, Aug. 10, 2021.

analysis of social media phenomena. Researchers do not need to analyze individual accounts – and indeed, they should be criminally punished if they try to do so. However, they need to understand which large demographic groups (age, race, gender, region, etc.) have engaged with which types of content (organic content, media, advertising etc.). They also need to understand when they have done so (e.g., in the period just before an election, after a major event, or in the normal pace of the platform) and "why" (for example, was the content referred to them by a recommendation or algorithmic system or did they search for it?).

If we had access to platform information along the three paths described above, what might we learn? Renée DiResta, Laura Edelson, Brendan Nyhan, and Ethan Zuckerman recently published some examples of the questions that could be asked and answered in an article in *Scientific American* titled "It's Time to Open the Black Box of Social Media" (Apr. 28, 2022):

> - Research suggests that misinformation is often more engaging than other types of content. Why is this the case? What features of misinformation are most associated with heightened user engagement and virality? Researchers have proposed that novelty and emotionality are key factors, but we need more research to know if this is the case. A better understanding of why misinformation is so engaging will help platforms improve their algorithms and recommend misinformation less often.
>
> - Research shows that the delivery optimization techniques that social media companies use to maximize revenue and even ad delivery algorithms themselves can be discriminatory. Are some groups of users significantly more likely than others to see potentially harmful ads, such as consumer scams? Are others less likely to see useful ads, such as job postings? How can ad networks improve their delivery and optimization to be less discriminatory?
>
> - Social media companies attempt to combat misinformation by labeling content of questionable provenance, hoping to push users

towards more accurate information. Results from survey experiments show that the effects of labels on beliefs and behavior are mixed. We need to learn more about whether labels are effective when individuals encounter them on platforms. Do labels reduce the spread of misinformation or attract attention to posts that users might otherwise ignore? Do people start to ignore labels as they become more familiar?

- Internal studies at Twitter show that Twitter's algorithms amplify right-leaning politicians and political news sources more than left-leaning accounts in six of seven countries studied. Do other algorithms used by other social media platforms show systemic political bias as well?

- Because of the central role they now play in public discourse, platforms have a great deal of power over who can speak. Minority groups sometimes feel their views are silenced online as a consequence of platform moderation decisions. Do decisions about what content is allowed on a platform affect some groups disproportionately? Are platforms allowing some users to silence others through the misuse of moderation tools or through systemic harassment designed to silence certain viewpoints?

Different people prioritize different questions when it comes to the platforms. Some worry about the scale and character of online harms, such as hate speech, disinformation, incitement and child endangerment. Others worry that the platforms are engaging in undue censorship, perhaps tilting the speech marketplace toward a particular ideology or party. Transparency and outsider access to platform data respond to both of those concerns. Indeed, while the political parties may diverge with respect to the other forms of tech regulation they support, greater transparency to reveal the nature of online harms and the ways that platforms are responding to them represents a bipartisan first step on the road to sound policy supported by either side.

III. Why Government Regulation Is Necessary

In an ideal world, the platforms, on their own, would make the data described above available for outside analysis. To be fair, some significant exceptions exist to the general rule of platform data hoarding. Twitter has established relationships with a range of academics to provide access to real-time and historical Twitter data[5] and have provided curated datasets on influence operations. Facebook worked with Social Science One (described below), acquired and promoted CrowdTangle (a platform that allows for analysis of public posts), and has recently teamed up with a large group of independent social media researchers to analyze the impact of the platform on the 2020 election.[6] To my knowledge, neither Google nor TikTok have embarked on similar efforts, although some individual academics have gotten favored access to all of these platforms, sometimes with a requirement of pre-publication review of research.[7]

We cannot and should not depend on the generosity of platforms, however, for analysis of data in the public interest. Indeed, we tend to forget that the data they hold is, in the end, about us, the users. Analysis of our behavior and communication – especially when it has profound civic consequences – should not be the sole province of the firm's

[5] See Twitter API, Academic Research Access, at https://developer.twitter.com/en/products/twitter-api/academic-research#:~:text=New%20and%20existing%20Twitter%20developers,an%20academic%20institution%20or%20university.

[6] See Nick Clegg & Chaya Nayak, New Facebook and Instagram Research Initiative to Look at US 2020 Presidential Election, Aug. 31, 2020, https://about.fb.com/news/2020/08/research-impact-of-facebook-and-instagram-on-us-election/; Talia Stroud, Joshua A. Tucker, Annie Franco, & Chad P. Kiewiet de Jonge, A Proposal for Understanding Social Media's Impact on Elections: Rigorous, Peer-Reviewed Scientific Research, Aug. 31, 2020, https://medium.com/@2020_election_research_project/a-proposal-for-understanding-social-medias-impact-on-elections-4ca5b7aae10.

[7] For a survey of existing and proposed approaches to platform data sharing, see Elizabeth Hansen Shapiro, Michael Sugarman, Fernando Bermejo, & Ethan Zuckerman, *New Approaches to Platform Data Research* (Feb. 2021); Jacob N. Shapiro, Natalie Thompson, & Alicia Wanless, *Research Collaboration on Influence Operations Between Industry and Academia: A Way Forward*, Carnegie Endowment for International Peace, Dec. 2020, https://carnegieendowment.org/files/Shapiro_Thompson_Wanless_Instantiating_Models_final.pdf.

employees, who are inexorably tied to the profit maximizing mission of the firm. Platform data scientists are restricted in the questions they can ask of the data and usually prevented from releasing their findings to the outside world. Although every effort must be undertaken to protect user privacy, some independent actors must have access to the data under the firm's control.

This was the motivation for the creation of Social Science One, an effort to provide independent researchers with access to platform data. Social Science One was run by academics for academics, and was independently funded by a set of foundations. Researchers and their projects were vetted by the Social Science Research Council; Facebook played no role in approving either researchers or research projects.

Facebook's principal role was, at the inception and to this day, to provide datasets for analysis by these vetted academics. To Facebook's credit, the project did produce one of the largest social media datasets in existence, which has been analyzed by close to one hundred researchers and is updated every few months. That dataset includes URLs (shared publicly at least 100 times) with certain categories of engagement data attached, to enable researchers to get a rough sense of which URLs were seen or engaged with by which large categories of users (by age, gender, and region). We had hoped to provide a much richer dataset akin to those required by PATA with the URLs dataset constituting a step in that direction.

Social Science One faced two sets of problems almost from its inception. The first and most significant was concern over protecting privacy. Within weeks of the announcement of Social Science One, the Electronic Privacy Information Center (EPIC) filed a "cease and desist" letter with the Federal Trade Commission and the European Union Data Protection Board. EPIC alleged that the research effort, by its very nature, violated user privacy and therefore violated both the General Data Protection Regulation of the European Union and the 2011 consent decree Facebook entered into with the FTC.

EPIC's letter raised the Cambridge Analytica scandal as a reason that academics should not be granted access to platform data. The cloud cast by the Cambridge Analytica scandal over Social Science One (and parallel research efforts) cannot be overstated. That scandal involved an academic's misuse of social graph data and its concomitant transfer to the international political consulting firm, Cambridge Analytica, which allegedly used it and other data to develop profiles for voter targeting. Later the FTC settled with Facebook, which paid $5 billion dollars to end the dispute, roughly a year into the existence of Social Science One. The scandal revealed how much data is in the possession of these platform, but it also now serves as a clarion call for the establishment of a legally sanctioned and regulated process that will simultaneously grant researcher access while ensuring government oversight to protect user privacy.

These pressures regarding user privacy constrained Social Science One as it tried to realize independent researchers' hopes of broad data access. Statistical noise, following the principles of differential privacy, was added to the URLs dataset, described above. In other words, even though the data were aggregated at the URL level (that is, no individual level data appeared in the dataset) and only URLs that were shared 100 times were included, nevertheless Facebook felt it necessary to remove any risk that someone might somehow figure a way to break down the aggregated data to surmise what a given individual may have seen on the platform. I continue to disagree that such a strategy was legally necessary to protect privacy, but these were sincerely held concerns that stymied many of our efforts to establish a system for independent analysis of platform data.

The second challenge that Social Science One confronted concerned the reliability of the data that Facebook delivered. Although it came to light after I resigned as co-chair of Social Science One, Facebook data scientists discovered that the URLs dataset neglected to include about a third of the United States population in the available data. In particular, the dataset did not include users for whom Facebook had not identified a political affinity, thereby likely leaving out of the URL interaction data many political moderates and others whose political views were not easily classified. Once the error was detected, Facebook updated the data within a month, but some papers that had used the data had already been published.

The experience of Social Science One is instructive as we consider legislative options to require platforms to share data with outside researchers. Until there is a countervailing legal requirement that pushes firms to share data, they will always perceive the costs of granting access as exceeding the benefits. Moreover, the life-cycle of Social Science One highlights how critical it is to do as much as possible to protect and safeguard user privacy in the design of these new data sharing initiatives. Only national legislation that creates a system of federal oversight that compels access to data under strict privacy protections can serve as a pathway both to providing the necessary information the public deserves and to respecting the public's legitimate expectations of privacy.

Conclusion

Greater transparency for the large internet platforms is that rare form of tech regulation that should appeal to both Republicans and Democrats. Whether one is most concerned about the scale of disinformation and extremism on the platforms or the platforms' potential bias in moderating content, greater access to platform data represents a prerequisite to gathering evidence and crafting public policy in this domain. It might take years for the parties to come together on regulation related to content moderation, antitrust, privacy, and online advertising. In the meantime, it is critical – now – to set up a data sharing regime that can allow policymakers and the public alike to understand how new digital technologies are affecting society.

If the United States does not act, Europe will. The new Digital Services Act provides for a series of public disclosures and data-sharing efforts. A working group from the European Digital Media Observatory (EDMO)[8] will soon release its code of conduct on researcher access to platform data. Greater transparency seems inevitable at this point. The question is which government will take the lead and how comprehensive legally compelled transparency will be.

In the span of less than a decade, what had been a widely shared utopian view of the potential of digital technologies for democracy has turned decidedly apocalyptic. Previously heralded as giving voice to the voiceless, the platforms are now blamed for worsening public health crises, teen depression and suicide, election-related conspiracies,

[8] See European Digital Media Observatory, Launch of the EDMO Working Group on Access to Platform Data, Aug. 31, 2021, https://edmo.eu/2021/08/30/launch-of-the-edmo-working-group-on-access-to-platform-data/.

ideological censorship, political polarization, and even genocide. All of this might be true. Until the platforms disclose more information to the public and subject themselves to outside, independent oversight, we will not be able to understand the prevalence of these problems or the responsibility of the platforms for exacerbating them.