February 15, 2022

The Honorable Zoe Lofgren
Chair
Committee on House Administration
U.S. House of Representatives
Washington, D.C. 20515

The Honorable Rodney Davis
Ranking Member
Committee on House Administration
U.S. House of Representatives
Washington, D.C. 20515

Dear Chairperson Lofgren, Ranking Member Davis, and members of the committee,

My name is Matthew Feeney. I am the director of the Cato Institute's project on emerging technologies. Tomorrow's hearing, titled "Big Data: Privacy Risks and Needed Reforms in the Public and Private Sectors," is relevant to my research, which focuses on how new and emerging technologies affect our civil liberties. The private and public use of "Big Data" is a critical issue at a time when there is bipartisan concern about the role of prominent technology companies and their influence on our civic and family lives.

Today's most popular online services collect vast troves of information about their users. As former Google CEO Eric Schmidt once put it in 2010, "With your permission, you give us more information about you, about your friends, and we can improve the quality of our searches. We don't need you to type at all. We know where you are. We know where you've been. We can more or less know what you're thinking about."[1] While services such as Google and Facebook remain popular among Americans, their collection of personal information prompts an understandable sense of unease, not least because law enforcement agencies at the state, local, and federal level analyze "Big Data" while conducting surveillance. Unfortunately, the state of 4th Amendment jurisprudence leaves much to be desired.

Congress can make reforms to address "Big Data" concerns, but these reforms should be carefully tailored to ensure that Americans are protected from unreasonable searches and seizures by the government without hampering growth and innovation in America's technology sector.

[1] Nick Saint, "Google CEO: 'We Know Where You Are. We Know Where You've Been. We Can More Or Less Know What You're Thinking About.'" *Business Insider*, October 4, 2010, https://www.businessinsider.com/eric-schmidt-we-know-where-you-are-we-know-where-youve-been-we-can-more-or-less-know-what-youre-thinking-about-2010-10.

GOVERNMENT USE OF BIG DATA

In recent years, the most prominent examples of government use of Big Data for surveillance were those revealed by former National Security Agency contractor Edward Snowden.[2] Bookshelves worth of books and articles have been written about the revelations, and an extensive accounting of them would require many more pages. For the purposes of this statement, it is enough to note that the Snowden revelations unveiled the vast scale of government access to personal information, which had previously been unknown to the public.

More recently, reporting suggests that the Central Intelligence Agency (CIA) has been conducting mass surveillance on Americans, a sad reflection on the state of reform many lawmakers and civil libertarians hoped for in the wake of the Snowden revelations.[3]

Government use of Big Data is not reserved to intelligence agencies that cite national security concerns as justification for surveillance. Domestic agencies have also engaged in surveillance fueled by Big Data. The Department of Homeland Security (DHS) has exploited the "border exception" to the Fourth Amendment, for example, allowing officers to conduct warrantless searches of electronic devices at international border crossings. In 2014, Palantir Technologies secured a contract to build a vast intelligence system capable of collecting data on a subject's "schooling, family relationships, employment information, phone records, immigration history, foreign exchange program status, personal connections, biometric traits, criminal records, and home and work addresses."[4]

State, local, and federal law enforcement agencies use surveillance tools to monitor Americans. Such tools include social media snooping software, cell site simulators, and facial recognition technology.[5]

Facial recognition is one of the most widely discussed surveillance technologies cited by civil libertarians. More than 2,200 law enforcement agencies and private companies in

---

[2] Lawfare, "Snowden Revelations," accessed February 14, 2022, https://www.lawfareblog.com/snowden-revelations.

[3] Matthew Guariglia and Andrew Crocker, "We Need Answers About the CIA's Mass Surveillance," *Electronic Frontier Foundation*, February 11, 2022, https://www.eff.org/deeplinks/2022/02/we-need-answers-about-cias-mass-surveillance.

[4] Spencer Woodman, "Palantir Provides the Engine for Donald Trump's Deportation Machine," *The Intercept*, March 2, 2017, https://theintercept.com/2017/03/02/palantir-provides-the-engine-for-donald-trumps-deportation-machine/.

[5] Brennan Center for Justice, "Third-Party Vendors of Social Media Monitoring Tools for Law Enforcement Agencies," November 17, 2021, https://www.brennancenter.org/our-work/research-reports/third-party-vendors-social-media-monitoring-tools-law-enforcement; American Civil Liberties Union, "Stingray Tracking Devices: Who's Got Them?" November 2018, https://www.aclu.org/issues/privacy-technology/surveillance-technologies/stingray-tracking-devices-whos-got-them.

twenty-seven countries have used Clearview AI, a controversial facial image search engine.[6] Clearview AI scrapes billions of images from social media platforms such as Instagram, Facebook, and Twitter. Investigators can search this database, uploading photos of suspects in the hopes of confirming their identity.[7]

Government use of Big Data is not reserved to law enforcement and intelligence agencies. Local governments are increasingly looking to "Smart City" technology to improve infrastructure.[8] Smarter infrastructure could result in fewer car crashes, less congested traffic, and more efficient public transport. Although not ostensibly related to mass surveillance, these initiatives have the potential to unveil intimate details. One of the best examples of how collection of personal transport data can reveal intimate details of private lives is a blog post (which has since been deleted) published by Uber detailing one-night sexual encounters among users.[9]

THE PRIVACY PARADOX

Any attempts at Big Data regulation or legislation will have to tackle the privacy paradox at the heart of debates surrounding data and privacy. On the one hand, Americans claim to care about their privacy, yet their behavior suggests otherwise. Amid complaints about "surveillance capitalism" we should keep in mind that Americans are keen users of services that collect information about them. Fortunately for those Americans more privacy-focused there are many services available that allow them to communicate and browse the web without revealing personal information.[10]

The Institute for Progress' Alec Stapp summarized the Privacy Paradox well, writing:

"There is a paradox at the heart of how people treat privacy. They say they value privacy while their actions imply they don't. In other words, their stated preferences contradict their revealed preferences. Is this just another example of cheap talk? Research from behavioral economics and related fields has shown that privacy valuations are highly context-dependent and subject to social-desirability bias and endowment effects."[11]

---

[6] Ryan Mac, Caroline Haskins, and Logan McDonald, "Clearview's Facial Recognition App Has Been Used By The Justice Department, ICE, Macy's, Walmart, And The NBA," *BuzzFeed News*, February 27, 2020, https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement.
[7] Kashmir Hill, "The Secretive Company That Might End Privacy as We Know It," *The New York Times*, November 2, 2021, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html.
[8] Jonathan Woetzel et al., "Smart cities: Digital solutions for a more livable future," *McKinsey & Company*, June 5, 2018, https://www.mckinsey.com/business-functions/operations/our-insights/smart-cities-digital-solutions-for-a-more-livable-future.
[9] Bradley Voytek, "Rides of Glory," *#UberData* (blog), Uber, https://web.archive.org/web/20140827195715/http://blog.uber.com/ridesofglory.
[10] End-to-end encrypted chat service Signal and the web browser TOR being among perhaps the most notable.
[11] Alec Stapp and Ryan Hagemann, "Regulatory Comment: Hearings on Competition and Consumer Protection in the 21st Century, The Intersection Between Privacy, Big Data, and Competition," *Niskanen*

Stapp went on to provide examples from a number of surveys showing that although many people claim to value their privacy, they often value their personal information at a low cost.

That "privacy" is not an absolute value must be considered by those concerned about "Big Data." Privacy is a condition people enjoy when they control information about themselves.[12] The age of social media has shown us the wide variety of ways in which people control such information. Some people like to post photos of their meals, themselves, or their romantic partners. Others would never post such content, choosing instead to volunteer personal information in political forums or e-commerce platforms. Others enjoy Amazon and YouTube recommendations based on viewing and purchasing histories. There is no one-size-fits-all approach to privacy. Fortunately, the market currently provides Americans of all privacy views a variety of social media platforms and communications tools.

Lawmakers considering regulation or legislation related to Big Data ought to be reassured by the fact that most Americans are aware that popular online services analyze their personal data and that alternatives are available to Americans concerned about their privacy. Contrary to what many lawmakers have been saying over the last few years, popular online communication and web browsing services are not monopolies.

REFORM

Those seeking to address concerns associated with Big Data should consider the following principles guiding reform:

1) Update legislation.

   The Electronic Communications Privacy Act (ECPA), passed in 1986, has yet to undergo a comprehensive revision accounting for more than thirty years of technological changes. The law allows law enforcement to access a wide range of intimate content such as emails, social media chats, and texts more than six months old without a warrant.[13] Congress should address the loopholes in ECPA that currently allow law enforcement to access troves of personal communications without a warrant.

Center, August 20, 2018, https://www.ftc.gov/system/files/documents/public_comments/2018/08/ftc-2018-0051-d-0027-152798.pdf.

[12] Jan Holvast, "History of Privacy," (conference paper, IFIP Advances in Information and Communication Technology, 2008), https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1037.5642&rep=rep1&type=pdf.

[13] Electronic Privacy Information Center, "Electronic Communications Privacy Act (ECPA)," accessed February 15, 2022, https://epic.org/ecpa/.

Congress should also improve upon Supreme Court precedent and impose a warrant requirement for searches of electronic devices at the border. Lawmakers from both parties have supported such legislation.[14]

2) Control, don't ban, facial recognition

More than half of American adults have images of their faces stored in databases that law enforcement officials can search while conducting facial recognition queries.[15] This is in large part thanks to the fact that many state departments of motor vehicles allow law enforcement to search their databases of driver's license photos.[16] The result is what Georgetown researchers have called the "Perpetual Lineup."[17] Concerns about facial recognition have promoted calls for a ban on the technology.[18] This is not the best approach.

Although facial recognition does pose a significant threat to civil liberties it is not without benefits. Facial recognition has the potential to ease queues at cinemas, grocery stores, and train stations. In addition, it can be used to enhance security and could be used to aid in searches for missing children or adults with dementia. Jettisoning these benefits is fortunately not necessary in order to protect civil liberties. As I have outlined before, there are policies that could ensure facial recognition is used by law enforcement in a narrow and transparent way that would not infringe on civil liberties. [19] These policies include a ban on real-time identification, removing data related to law-abiding Americans from databases, and open source, accuracy, and transparency requirements.[20]

3) Embrace transparency

A regrettable feature of modern American law enforcement is its lack of transparency. Too often, Americans learn about what surveillance devices their governments use from journalists rather than elected officials. Federal agencies should be required to regularly release information about 1) how often they ask technology firms for users' personal data, 2) what new surveillance tools its officers are using, and 3) how and when data related to closed cases are purged.

---

[14] Protecting Data at the Border Act, S. 1606, 116th Cong. (2019).
[15] Clare Garvie, Alvaro Bedoya, and Jonathan Frankle, "The Perpetual Line-up: Unregulated Police Face Recognition in America," *Georgetown Law Center on Privacy and Technology*, October 18, 2016, https://www.perpetuallineup.org.
[16] Ibid.
[17] Ibid.
[18] Fight for the Future, "Ban Facial Recognition," accessed February 14, 2021, https://www.banfacialrecognition.com.
[19] Matthew Feeney, "Should Police Facial Recognition Be Banned?" *Cato at Liberty* (blog), Cato Institute, May 13, 2019, https://www.cato.org/blog/should-police-facial-recognition-be-banned.
[20] Ibid.

4) Resist calls for "data ownership"

It is not uncommon in discussion about "Big Data" to hear "data ownership" proposals that would grant a property right in personal data. Although perhaps at first glance attractive, such proposals are conceptually confused and are not necessary to protect privacy.

Data is very different from traditional property. A property right entails a "bundle" of other rights, such as the right to exclude and alter property. If you have a property right in a house you can exclude others from it and alter it. It is also clear who owns a home and other pieces of property. Data does not have these qualities. Data is not excludable in the same way as a house. My enjoyment of data is not affected if you use the same data or alter the data for yourself. Similarly, it is not clear how such a property right would be allocated. Does my birthdate contain data about me, my mother, or my father? Should property rights in marriage data be assigned to one spouse or both?

Most importantly, allowing for a property right in data is not necessary for privacy, as Center for Growth and Opportunity scholar Will Rinehart has explained:

"[A] property right in data isn't needed to establish consumer privacy rights. For evidence of this fact, one only needs only to consult the current laws in the United States. The Gramm-Leach-Bliley Act (GLBA), the Fair Credit Reporting Act (FCRA), the Children's Online Privacy Protection Act (COPPA), and the California Consumer Privacy Act (CCPA), just to name a few, all protect privacy without creating property rights. As Stanford Law Professor Lothar Determann has said quite bluntly, 'no one owns data' because data are already 'subject to a complex landscape of access rights and restrictions.' Privacy regulation already defines certain kinds of entitlements to control and contract upon data. Adding a superordinate property right on top of these existing restrictions would make the entire enterprise all that more complicated and undermine current efforts to grant consumers control. If data property rights were implemented, for example, would an individual be able to limit critical information from being shared with credit rating agencies?"[21]

As lawmakers consider regulations and legislation they should be wary of the risks of regulatory capture. Market incumbent firms such as Microsoft, Google, and Facebook are well positioned to comply with whatever new regulations and legislation emerge

[21] *Hearing on Data Ownership: Exploring Implications for Data Privacy Rights and Data Valuation, Before the Committee on Banking, Housing, and Urban Affairs*, 116th Cong. (2019) (testimony of Will Rinehart, Director of Technology and Innovation Policy at the American Action Forum), https://www.banking.senate.gov/imo/media/doc/Rinehart%20Testimony10-24-19.pdf.

from Congress. History teaches us that these firms will invest in preparing for new regulation, and competition may suffer if regulations and legislation results in costs for market entrants. One of the best examples of this dynamic can be found in Google's preparation for the European Union's General Data Protection Regulation (GDPR), its flagship set of privacy regulations. Google reportedly spent hundreds of years worth of human time preparing for GDPR, a set of regulations that has benefitted the Silicon Valley giant:

"Google, for example, has seen some unique market benefits following the activation of GDPR in May 2018. From April to July of last year, Google's market share increased by 0.93% while most other adtech vendors in North America and Europe lost ground. The lowest rank of adtech companies — the top 150 to 100 — lost the most market share, a decline of 31.86% on average. Google was of course prepared for privacy regulation ahead of time, having spent "hundreds of years of human time" and, ostensibly, billions of dollars to shore up its defenses."[22]

Lawmakers on this side of the Atlantic should be wary of repeating Europe's experience.

Thank you for giving your attention to these important issues. I would welcome the opportunity to discuss my thoughts further with you or your staff.

Yours,

Matthew Feeney
Director, Project on Emerging Technologies
Cato Institute

[22] Alex Moazed, "How GDPR is Helping Big Tech and Hurting Competition," *Applico*, accessed February 14, 2022, https://www.applicoinc.com/blog/how-gdpr-is-helping-big-tech-and-hurting-the-competition/.