Before the
National Telecommunications and Information Administration
Washington, DC 20230

Privacy, Equity, and Civil Rights Request for Comment	))	Comments from Public Knowledge[1]
	)	
Docket Number: NTIA-2023-0001	)	
	)	

[1] Authored by Sara Collins, Senior Policy Counsel, based on Public Knowledge's comments to the Federal Trade Commission in their Advance Notice of Proposed Rulemaking: Trade Regulation Rule on Commercial Surveillance and Data Security, 87 FR 51273. *See* comments at https://publicknowledge.org/policy/comments-in-ftc-filing-on-commercial-surveillance-and-poor-data-practices/.

Unlimited Data Collection, Use, and Sharing Amplifies Existing Harms Felt by Marginalized Communities

The National Telecommunications and Information Administration's (NTIA) call for comments on the intersection of privacy, equity, and civil rights is an important step in continuing the Biden Administration's priority of promoting equity and increasing support for marginalized communities.[2] Building off the listening sessions hosted by NTIA in December of 2021, Public Knowledge offers its own comments to illustrate the impact of data collection and processing on marginalized groups. These adverse impacts are not relegated to one data type or specific use case; rather the impacts can be felt broadly across the entire ecosystem.

Since the inception of the internet more than two decades ago, the exponential increase in the volume of consumer data collected through the internet and the subsequent exploitation of that consumer data to target advertisements, aid in discrimination, and put consumers at great risk is astonishing. This is because, as then FTC Commissioner Chopra testified at a hearing in 2019, data has unique features that are unlike any other asset in the economic marketplace.[3] Those features are:

> *First, it is not a finite resource, like precious metals or minerals.*
>
> *Second, data is not "consumed" in the traditional sense… it can be copied and shared.*
>
> *Third, data gets more valuable as you collect more of it.*[4]

[2] Advancing Racial Equity and Support for Underserved Communities Through the Federal Government, Exec. Order No. 13985, 86 FR 7009 (Jan. 20, 2021), https://www.govinfo.gov/content/ pkg/FR-2021-01-25/pdf/2021-01753.pdf.

[3] Online Platforms and Market Power, Part 3: The Role of Data and Privacy in Competition: Hearing Before the Subcomm. on Antitrust, Commercial, and Administrative Law, 116 Cong. 2 (2019) (Statement of Rohit Chopra). https://www.ftc.gov/system/files/documents/public_statements/1549812/chopra_-_testimony_at_hearing_on_online_platforms_and_market_power_part_3_10-18-19.pdf

[4] Id.

These three features, plus the fact that data has become so easy to collect and store, mean companies have little incentive to limit their data collection and consumers cannot be expected to effectively fight back against the push to collect more data.[5] The primary reason for all this opaque data sharing is targeted advertising.[6] Targeted advertising is fueled by personal data.[7] The theory of targeted advertising is that by having as much information as possible about your potential customers, you can more effectively spend your advertising dollars. This is why we have seen privacy scandals in almost every sector — from makeup and beauty to weather forecasting, to radio apps.[8] However, the harms of increased data collection and processing are not felt by consumers equally. Data use can oftentimes replicate bias and discrimination that exist in the physical world.

Unfortunately, just putting stricter controls on racial, sexual orientation, gender and religious identity data cannot effectively combat the discriminatory effects of data processing. The context in which the data is collected, as well as the assumptions made about individuals, all can be used to discriminate.

For example, in 2021 a Catholic Substack, The Pillar, was able to out a gay priest using location data collected by the dating app Grindr.[9] While the location data was not identified as

[5] *See* Harold Feld, *The Market for Privacy Lemons. Why "The Market" Can't Solve The Privacy Problem Without Regulation,* Wetmachine (Feb. 15, 2019) https://wetmachine.com/tales-of-the-sausage-factory/the-market-for-privacy-lemons-why-the-market-cant-solve-the-privacy-problem-without-regulation/ (applying Akerlof economic analysis to demonstrate that rational actors cannot believe any promise of protecting private information, thus preventing any functioning market from emerging).

[6] *But See* Ed Markey, United States Senator for Massachusetts, *Markey Report Reveals Automobile Security And Privacy Vulnerabilities* (Feb. 9, 2015) https://www.markey.senate.gov/news/press-releases/markey-report-reveals-automobile-security-and-privacy-vulnerabilities (Data is also collected even without a monetization model in place, because it may become monetizable in the future).

[7] Anna Dorothea Ker, *The Big Business of Ad Tech*, The Privacy Issue (Jan. 22, 2020) https://theprivacyissue.com/data-tracking/big-business-ad-tech.

[8] Thomas Brewster, *A Load of Apple iPhone Apps Are 'Covertly' Selling Your Location*, Forbes (Sept. 7, 2018), https://www.forbes.com/sites/thomasbrewster/2018/09/07/a-load-of-apple-iphone-apps-are-covertly-selling-your-location/?sh=5a7e7cc85832.

[9] Joseph Cox, *The Inevitable Weaponization of App Data Is Here*, Vice (Jul. 21, 2021), https://www.vice.com/en/article/pkbxp8/grindr-location-data-priest-weaponization-app.

being the outed priest, using publicly available information and other data purchased from data brokers, The Pillar was able to identify the priest with a high degree of confidence. This public outing forced the priest to resign his position in the Catholic Church. The data that was used for this public outing was not sexual orientation information, rather it relied on tracking the movements and app usage of one person in order to infer their belonging to a marginalized community. And the weaponization of this kind of personal information, oftentimes referred to as doxxing, disproportionately affects the most marginalized among us.[10]

Since the *Dobbs* decision, there has been heightened scrutiny of how period tracking apps collect, use, and retain data.[11] While an app like Flo is unlikely to have data about a person's abortion, or even their pregnancy, it does track whether a user has missed a period. A missing period can be a useful proxy to determine that a woman is pregnant. Flo had promised that data would only be shared to "provide services in connection with the app,"[12] but that was not the case. Flo was sharing health related information, like a user's pregnancy status, with a variety of third-party marketing and analytics firms. To add insult to injury, the health information being shared with third parties was not being securely transferred. An investigation by the Wall Street Journal made it clear that outside attackers could see the unencrypted health information being transferred from Flo to its third-party marketers, like Facebook.[13] This free flow of information from Flo meant that law enforcement had easy access, through either sale or surveillance, of

[10] Access Now, *What is doxxing, how it endangers women, and what we can do about it,* (Mar. 7, 2022) https://www.youtube.com/watch?v=63GJ00AxZmA.

[11] Rina Torchinsky, *How period tracking apps and data privacy fit into a post-Roe v. Wade climate*, NPR (Jun. 24, 2022), https://www.npr.org/2022/05/10/1097482967/roe-v-wade-supreme-court-abortion-period-apps).

[12] In the Matter of Flo Health, Inc., Docket C-4747, Complaint (Jan. 13, 2021), https://www.ftc.gov/system/files/documents/cases/flo_health_complaint.pdf.

[13] Dam Schechner, *You Give Apps Sensitive Personal Information. They Tell Facebook.*, Wall Street Journal (Feb. 22, 2019), https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636.

what women were pregnant and if their pregnancies ended earlier than expected. As states continue to criminalize abortions,[14] there will be significant incentive to surveil women in order to enable arrests or warrants seeking more information. And that initial surveillance could be predicated on as little information as a missed period noted in an app like Flo. While the collection of this type of data by a period tracking app is useful for women, the downstream effects of having an unregulated data market mean that weaponization of their personal information is likely to occur.

Algorithmic Decision-Making and Artificial Intelligence Systems Codify Pre-Existing Bias and Inequality

Another way data can be used to discriminate is through automated decision-making systems, usually referred to as algorithms. Oftentimes algorithms are marketed as being a more scientific or rigorous way to make important decisions; however, this assertion does not match up with reality.[15] Algorithmic decision-making systems can, and frequently do, suffer from a wide array of deficiencies and biases. There are a variety of reasons that AI systems can “fail.” A system may be assigned a task that is conceptually or practically impossible; it may be the product of an imprecise translation of business objectives;[16] it may interact with the real world in unexpected ways, with too few risk-mitigation measures; or it may suffer from overstated or misrepresented capabilities.[17] Additionally, the data on which an algorithm relies may preclude proper performance, if the data is inaccurate, incorrectly coded, fragmented, or incomplete.[18]

[14] *See* Center for Reproductive Rights, *After Roe Fell: Abortion Laws by State,* https://reproductiverights.org/maps/abortion-laws-by-state/ (last accessed Mar. 6, 2023).

[15] Sara Collins, *21st Century Snake Oil: The Consequences of Unregulated, Unproven AI*, Tech Policy Press (October 13, 2021).

[16] Samir Passi & Solon Barocas, *Problem Formulation and Fairness*, Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (2019).

[17] Inioluwa Deborah Raji, I. Elizabeth Kumar, Aaron Horowitz & Andrew D. Selbst, *The Fallacy of AI Functionality*, 2022 ACM Conference on Fairness, Accountability, and Transparency (2022).

[18] Sharona Hoffman & Andy Podgurski, *Big Bad Data: Law, Public Health, and Biomedical Databases*, 41 J. Law Med. Ethics 56 (2013); Milena A. Gianfrancesco, Suzanne Tamang, Jinoos Yazdany & Gabriela Schmajuk,

While "prediction" is used loosely in the world of computer science, it carries a specific connotation in conversations among consumers and policymakers: the forecasting of future events. While advances in AI have led to breakthrough performances in a variety of tasks, there are significant limits on the ability of algorithms to predict—in the colloquial sense—future outcomes.[19] For instance, algorithmic models have performed poorly, and no better than two-variable regression models, in predicting recidivism.[20]

Failures of predictive systems in the justice system are particularly perilous. COMPAS, a recidivism risk prediction tool, is used across the nation in pretrial, parole, and sentencing decisions. Researchers at Dartmouth, however, showed that the capabilities of the tool were grossly overstated, and that the performance of the tool was indistinguishable from predictions made by individuals with little to no criminal justice expertise.[21] Most troublingly, analysis by ProPublica "found that black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than black defendants to be incorrectly flagged as low risk."[22] Put simply, countless individuals are required to stay in jail for no other reason than the color of their skin.

Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data, 178 JAMA Intern. Med. 1544 (2018). For one example of the harms that can be caused by such issues, see the "unsafe and incorrect" recommendations for cancer patients generated by IBM's artificial intelligence unit, "Watson." Eliza Strickland, *IBM Watson Heal Thyself: How IBM Watson Overpromised and Underdelivered on AI Health Care*, IEEE Spectrum, https://spectrum.ieee.org/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care; Casey Ross, Ike Swetlitz, Rachel Cohrs, Ian Dillingham, Nicholas Florko & Maddie Bender, *IBM's Watson Supercomputer Recommended 'Unsafe and Incorrect' Cancer Treatments, Internal Documents Show*, STAT News (July 25, 2018), https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/?utm_source=STAT+Newsletters&utm_campaign=beb06f048d-MR_COPY_08&utm_medium=email&utm_term=0_8cab1d7961-beb06f048d-150085821.

[19] Arvind Narayanan & Matt Salganik, "Limits to prediction: pre-read," Princeton University (Sept. 1, 2020);

[20] Julia Dressel & Hany Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, 4 Science Advances (2018).

[21] *Id.*

[22] Jeff Larson, Surya Mattu, Lauren Kirchner & Julia Angwin, How We Analyzed the COMPAS Recidivism Algorithm, ProPublica (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

In another use of predictive systems by the justice system, a man in Chicago was targeted for surveillance by the police because an algorithm used by the police department predicted that he was likely to be involved in a shooting. The surveillance attracted attention from his neighbors, who decided he had cooperated with the Chicago police – resulting in him being shot multiple times in the leg.[23] Both of these are examples of racial profiling that has been masked by the latest technology.

These algorithmic systems are prone to perpetuating, and sometimes even amplifying, historical prejudices.[24] As one technologist put it: "Machine learning is like money laundering for bias. It's a clean, mathematical apparatus that gives the status quo the aura of logical inevitability."[25]

This is causing real-world harms. Algorithms used to allocate health care have been shown to make different decisions for groups of different races.[26] Algorithms and associated data mining techniques used for employment decision-making risk replicating the biased hiring practices that currently exist, while simultaneously obfuscating scrutiny of harm.[27] Facial recognition software performs differently on people of varying races and genders, and is especially likely to misidentify the gender of women with dark skin.[28] The federal government itself has performed studies that found facial recognition technologies to be less reliable for

[23] Matt Stroud, *Heat Listed*, The Verge (May 24, 2021), https://www.theverge.com/c/22444020/chicago-pd-predictive-policing-heat-list.

[24] A. Caliskan, J. J. Bryson, A. Narayanan, *Semantics derived automatically from language corpora contain human-like biases*. Science 356, 183–186 (2017).

[25] Maciej Ceglowski, *The Moral Economy of Tech*, Society for the Advancement of Socio-Economics (SASE) Conference (June 26, 2016).

[26] Heidi Ledford, *Millions of Black People Affected by Racial Bias in Health-Care Algorithms*, Nature (Oct. 26, 2019), https://www.nature.com/articles/d41586-019-03228-6.

[27] Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 Wm. & Mary L. Rev. 857 (2017), https://scholarship.law.wm.edu/wmlr/vol58/iss3/4.

[28] Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 Proceedings of Machine Learning Research 1 (2018).

Asian and African-American faces relative to Caucasian faces.[29] This is in part because of the effects of biased, inaccurate, or skewed data, which cause significant harm in many contexts and may be especially pernicious in so-called "predictive policing."[30]

Recommendations

While the first and best step to address these harms should be Congress passing a comprehensive privacy law, it is not the only action that would have beneficial effects. The executive branch and independent agencies could use their authority to start to move the needle. Some examples of this include the Federal Trade Commission's Advance Notice of Proposed Rulemaking on Commercial Surveillance and Data Security[31] (in which NTIA provided expert comments[32]) , Federal Communications Commission's Notice of Proposed Rulemaking updating and strengthening data breach of customer proprietary network information,[33] and the Equal Employment Opportunity Commission's guidance on the requirements for using algorithmic assessment for job applicants in accordance with the Americans with Disabilities Act.[34]

But that does not mean more cannot be done. The NTIA is well positioned to be that catalyst of action. NTIA should continue to participate in rulemaking activities by providing expert comment and technical assistance when warranted. NTIA could also encourage other

[29] Chad Boutin, *NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software*, National Institute of Standards and Technology, Department of Commerce (Dec. 19, 2019), https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software.

[30] Rashida Richardson, Jason M. Schultz & Kate Crawford, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U. L. Rev. 192 (2019).

[31] Trade Regulation Rule on Commercial Surveillance and Data Security, 87 FR 51273 (announced Aug. 22, 2022).

[32] National Telecommunications and Information Administration ANPR Comment (Nov. 21, 2022), https://www.ntia.doc.gov/files/ntia/publications/ ftc_commercial_surveillance_anpr_ntia_comment_ final.pdf.

[33] Data Breach Reporting Requirements, 88 FR 3953 (proposed Jan. 23, 2023).

[34] *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees,* EEOC-NVTA-2022-2 (May 12, 2022) https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence.

agencies[35] to start rulemaking processes of their own or issue guidance. The slow speed of Congress to address this problem should not hamper the administrative agencies charged with protecting the public. This report is an important first step in what should become a whole government push to protect people's privacy.

[35] *See* Congressional Research Service, *Data Protection Law: An Overview,* (Mar. 25, 2019) https://crsreports.congress.gov/product/pdf/R/R45631 (a comprehensive list of current federal privacy laws and the agencies who enforce them).