January 15, 2025

National Telecommunications and Information Administration
U.S. Department of Commerce
1401 Constitution Avenue NW
Washington, D.C. 20230

Comment submitted electronically via https://www.regulations.gov

RE: NTIA Ethical Guidelines for Research Using Pervasive Data Request for Comment, Docket NTIA-2024-0004-0001

Dear NTIA Officials:

The undersigned civil society organizations are writing in response to the National Telecommunications and Information Administration's (NTIA) request for public comments[1] on the possibility of the NTIA issuing ethical guidelines for the use of "pervasive data."[2] As experts on privacy, technology, civil liberties, and human rights issues, we wish to offer our perspectives on the special risks and threats posed by a key source of data about people: commercial data brokers. Given that this profit-driven industry's practices and data products undermine or conflict with core ethics principles of human subjects research—including "respect for persons, beneficence, and justice"[3]—we urge the proposed NTIA ethical guidelines to caution the academic research community against reliance on commercial data brokers.

As essential background, the United States lacks a robust, comprehensive federal consumer privacy and data security law to regulate the commercial collection and sale of personal data. Instead, Americans must rely on a scattershot array of federal and state laws that apply to only certain industries or certain types of data in limited contexts, such as patient records held by medical care providers,[4] data collected by consumer reporting agencies like

[1] Nat'l Telecomms. & Info. Admin., U.S. Dep't. Commerce, Ethical Guidelines for Research Using Pervasive Data, Request for Comments, Docket No. 241204-0309 (Dec. 11, 2024), https://www.federalregister.gov/documents/2024/12/11/2024-29064/ethical-guidelines-for-research-using-pervasive-data.
[2] The NTIA's Request for Comments broadly defines pervasive data as "data about people gathered through online services," which we adopt for the purposes of this comment letter. As the NTIA further explains, "Pervasive data may include text, images, videos, biometric information, information about a data subject's behavior (purchases, financial standing, media consumption, search history, medical conditions, location, etc.), and other information that makes up a person's digital footprint."
[3] Chad Boutin, Nat'l Inst. of Standards and Tech., U.S. Dep't. Commerce, NIST Researchers Suggest Historical Precedent for Ethical AI Research (Feb. 15, 2024), https://www.nist.gov/news-events/news/2024/02/nist-researchers-suggest-historical-precedent-ethical-ai-research.
[4] Health Insurance Portability and Accountability Act, Pub. L. 104-191, 110 Stat. 1936 (1996).

credit bureaus,[5] and personal data sought by certain foreign entities.[6] Certain states have enacted their own consumer privacy laws,[7] including some that directly address data brokerage,[8] but each state law comes with its own idiosyncrasies and limitations while those outside the state remain uncovered.

Under this patchwork legal regime, commercial data brokers are stockpiling and selling all sorts of data about people at vast scales—with little oversight and little respect for their privacy. The digital trails that Americans generate with virtually every action they take online today are especially appealing targets for data brokers. Sourced from third parties such as internet service providers, corporate websites, social media platforms, messaging boards, and smartphone app developers, data about people gathered through online services can reveal sensitive details about us and our lives—from name, age, addresses, email addresses, and phone numbers to income, education, professional affiliations, relationships, religion, precise location and movement patterns, medical conditions, and political preferences.[9] This data may even include information about us that cannot be changed after a privacy breach incident, such as facial images, fingerprints, and other biometric data.

Data brokers aggregate this data into exhaustive dossiers on millions of Americans that often also include inferences about other attributes, such as shopping habits and religious beliefs, and predictions about behavior like susceptibility to different kinds of advertising. These inferences are often wrong.[10] Thus data brokers are often trafficking people's profiles that can subject them—for example, gun owners[11] or visitors to an abortion clinic[12]—to potentially severe consequences, including stalking, targeting for financial scams, false accusations, or

---

[5] Fair Credit Reporting Act, 16 U.S.C. § 1681f and 16 U.S.C. § 1681b.
[6] See EPIC, DOJ Finalizes Mixed Bag Data Broker Regulation (Jan. 8, 2025), https://epic.org/doj-finalizes-mixed-bag-data-broker-regulation/.
[7] International Association of Privacy Professionals, US State Privacy Legislation Tracker (last updated Jan. 6, 2025), https://iapp.org/resources/article/us-state-privacy-legislation-tracker/.
[8] For example, the California Consumer Privacy Act, codified in Cal. Civ. Code § 1798.100, and "Daniel's Law," P.L. 2021, c.371 (A6171 2R CC).
[9] EPIC, Data Brokers (accessed Jan. 10, 2025), https://epic.org/issues/consumer-privacy/data-brokers/; Emile Ayoub and Elizabeth Goitein, Closing the Data Broker Loophole, Brennan Center for Justice (Feb. 13, 2024), https://www.brennancenter.org/our-work/research-reports/closing-data-broker-loophole; Steven Melendez and Alex Pasternack, Here are the data brokers quietly buying and selling your personal information, Fast Company (Mar. 2, 2019), https://www.fastcompany.com/90310803/here-are-the-data-brokersquietly-buying-and-selling-your-personal-information.
[10] See Suzanne Smalley, 'Junk Inferences' by Data Brokers Are a Problem for Consumers and the Industry Itself, Record (June 12, 2024), https://therecord.media/junk-inferences-data-brokers; see also Nico Neumann et al., "Data Deserts and Black Boxes: The Impact of Socio-Economic Status on Consumer Profiling," Management Science 70, no. 11 (Jan. 2024): 8003, https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2023.4979?j (consumers "with higher incomes or living in affluent areas" are more likely to be profiled accurately by data brokers).
[11] See Corey G. Johnson, Without Knowledge or Consent, ProPublica (Oct. 24, 2024), https://www.propublica.org/article/gunmakers-owners-sensitive-personal-information-glock-remington-nssf.
[12] Joseph Cox, Inside the U.S. Government-Bought Tool That Can Track Phones at Abortion Clinics, 404 Media (Oct. 23, 2024), https://www.404media.co/inside-the-u-s-government-bought-tool-that-can-track-phones-at-abortion-clinics/.

criminal prosecution. And when government agencies are the customers, data brokers are enabling the circumvention of people's constitutional and statutory rights when they sell data to the government that would otherwise require legal process like a warrant if the government sought to collect the same data directly from a communications company or the people targeted.[13]

Data brokers are obtaining and commodifying this data about people without their awareness, let alone their express informed consent.[14] Companies face enormous incentives to amass consumers' personal data however possible,[15] including through trickery[16]—data which is then aggregated and repackaged by data brokers for sale. This widespread corporate surveillance is fueled by ineffective, antiquated "notice-and-choice" or "notice-and-consent" models for obtaining consumer consent,[17] which assume that people who click on "yes" on a website to hurry past a company's privacy disclaimer have given meaningful consent to their data being harvested and exploited. As the NTIA has correctly noted, notice-and-choice "mandates have resulted primarily in long, legal, regulator-focused privacy policies and check boxes, which only help a very small number of users who choose to read these policies and make binary choices."[18]

Thus, commercial data brokers do not collect or provide data based on opt-in consent—a business model which fundamentally conflicts with researchers' ethical and legal obligations to get human subjects' informed consent and to respect participant privacy. As researchers know, they must adhere to numerous legal and institutional requirements for research that involves human beings, such as the Common Rule[19] requiring investigators to get informed consent from their research participants. These safeguards both protect the rights

---

[13] Comment Letter to the Office of Management and Budget Regarding Commercially Available Information, Brennan Center for Justice, Demand Progress Education Fund (Demand Progress), the Electronic Privacy Information Center (EPIC), the Surveillance Technology Oversight Project (S.T.O.P.), and 12 other civil society organizations (Jan. 9, 2025), https://www.brennancenter.org/our-work/research-reports/comment-submitted-office-management-and-budget-regarding-executive-branch.

[14] Gennie Gebhart, Electronic Frontier Foundation, Bad Data "For Good": How Data Brokers Try to Hide Behind Academic Research (Aug. 15, 2022), https://www.eff.org/deeplinks/2022/08/bad-data-good-how-data-brokers-try-hide-academic-research.

[15] Fed. Trade Comm'n, Data Brokers: A Call for Transparency and Accountability (May 2014), https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

[16] Fed. Trade Comm'n, FTC Report Shows Rise in Sophisticated Dark Patterns Designed to Trick and Trap Consumers (Sept. 15, 2022), https://www.ftc.gov/news-events/news/press-releases/2022/09/ftc-report-shows-rise-sophisticated-dark-patterns-designed-trick-trap-consumers.

[17] Claire Park, New America Foundation, How "Notice and Consent" Fails to Protect Our Privacy (Mar. 23, 2020), https://www.newamerica.org/oti/blog/how-notice-and-consent-fails-to-protect-our-privacy/.

[18] Nat'l Telecomms. & Info. Admin., U.S. Dep't. Commerce, Developing the Administration's Approach to Consumer Privacy, Request for Comments, Docket No. 180821780-8780-01 (Oct. 11, 2018), https://www.federalregister.gov/documents/2018/09/26/2018-20941/developing-the-administrationsapproach-to-consumer-privacy.

[19] Office for Human Research Protections, U.S. Dep't Health and Human Servs., What Regulations Protect Research Participants? The Common Rule (last reviewed Oct. 17, 2023), https://www.hhs.gov/ohrp/education-and-outreach/about-research-participation/protecting-research-volunteers/principal-regulations/index.html.

and welfare of the participants and promote the public's trust in the scientists and institutions conducting the research. However, as discussed herein, data brokers do not meet this ethics standard; to the contrary, these businesses are financially incentivized to *disregard* people's privacy.

Given the myths of data anonymization, researchers should disbelieve data brokers' claims to provide "anonymous" data to sidestep the privacy and informed consent concerns. Even when personal data is supposedly "anonymized" and delinked from people's unique identities, it can be pieced together with other data fragments to reidentify individuals.[20] In fact, different studies have shown that only three pieces of information—zip code, birthday, and gender—are required to pinpoint the unique identity of 87 percent of Americans[21]; commercial data brokers already collect and sell this and more data about us. As another example, our whereabouts and daily movement patterns give away intimate details such as where we sleep at night, what modes of transportation we use, and whom we visit regularly. One study found that only two randomly chosen time and location data points were needed to uniquely characterize 50 percent of people.[22] Also, due to the possibility of future technological advancements in artificial intelligence, computing power, forensic science, data tracking, and data analysis techniques, current methods such as scrubbing, scrambling, or aggregating data to obscure people's unique identities cannot be guaranteed to be forever "future-proof."

The data products sold by data brokers are also often junk, which means their use conflicts with researchers' ethical commitments to scientific integrity and beneficence towards human subjects. The data products peddled by data brokers are routinely riddled with errors, systemic biases,[23] junk inferences,[24] and other flaws; for instance, past research indicated that at least 40 percent of the attributes that data brokers had on people were found to be

---

[20] See Justin Sherman, Big Data May Not Know Your Name. But It Knows Everything Else, Wired (Dec. 19, 2021), https://www.wired.com/story/big-data-may-not-know-your-name-but-it-knows-everything-else/; Jennifer Valentino-DeVries et al., Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret, New York Times (Dec. 10, 2018), https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacyapps.html.

[21] Paige Collings, Electronic Frontier Foundation, Debunking the Myth of "Anonymous" Data (Nov. 10, 2023), https://www.eff.org/deeplinks/2023/11/debunking-myth-anonymous-data.

[22] Gennie Gebhart, Electronic Frontier Foundation, Bad Data "For Good": How Data Brokers Try to Hide Behind Academic Research (Aug. 15, 2022), https://www.eff.org/deeplinks/2022/08/bad-data-good-how-data-brokers-try-hide-academic-research.

[23] See Rashida Richardson et al., Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, And Justice, NYU Law Review 192 (Feb. 2019), https://www.nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson-Schultz-Crawford.pdf.

[24] See Suzanne Smalley, 'Junk Inferences' by Data Brokers Are a Problem for Consumers and the Industry Itself, Record (June 12, 2024), https://therecord.media/junk-inferences-data-brokers; see also Nico Neumann et al., Data Deserts and Black Boxes: The Impact of Socio-Economic Status on Consumer Profiling, Management Science 70, no. 11 (Jan. 2024): 8003, https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2023.4979?j (consumers "with higher incomes or living in affluent areas" are more likely to be profiled accurately by data brokers).

inaccurate or no longer accurate.[25] These personal data errors can result in devastating outcomes for people, such as when landlords, lenders, or employers wrongfully deny applications. Because the original human subjects ultimately profiled by data brokers are not aware of these data inaccuracies, or even that the data collection and monetization are taking place, they are also unable to correct the mistakes. By relying on commercial data brokers, academic researchers are thus taking major risks of perpetuating uncorrectable flaws in data about real life people, adding conclusions based on tainted data to the public domain, and undermining the integrity and credibility of their own work.

Reliance on commercial data brokers undermines the research community's broader ethical and legal obligations, including transparency and deterrence of academic fraud. That is because the privacy violations inherent in commercial data brokers' data products risk further ethical breaches if the studies based on them are published, as doing so would likely require providing public access to the underlying data. Public access might not even be allowed at all; some data brokers require non-disclosure agreements that prohibit their customers from publicly releasing the data or that give the data broker control over how its involvement as the supplier is disclosed.[26]

These pitfalls of data brokers must be understood in the context of other regulations, practices, and norms that run in the opposite direction to ensure public access, encourage data sharing and collaboration with colleagues, promote research integrity and reproducibility, and guard against fabrication of data and other academic fraud. In particular, the White House Office of Science and Technology Policy has directed federal funding agencies, "no later than December 31, 2025, to make publications and their supporting data resulting from federally funded research publicly accessible without an embargo on their free and public release."[27] Scientific journals also often require submitting authors to deposit the data underlying their publications in publicly accessible repositories.[28] Meanwhile, the ability of peer reviewers, outside investigators, watchdogs, and journalists to uncover acts of research data falsification or manipulation depends on their access to the studies' underlying data sets.[29] That

---

[25] Levi Kaplan, Alan Mislove, and Piotr Sapieżyński, Measuring Biases in a Data Broker's Coverage, PrivacyCon 2017 Conference (July 2017), https://www.ftc.gov/system/files/ftc_gov/pdf/PrivacyCon-2022-Kaplan-Mislove-Sapiezynski-Measuring-Biases-in-a-Data-Brokers-Coverage.pdf

[26] Gennie Gebhart, Electronic Frontier Foundation, Bad Data "For Good": How Data Brokers Try to Hide Behind Academic Research (Aug. 15, 2022), https://www.eff.org/deeplinks/2022/08/bad-data-good-how-data-brokers-try-hide-academic-research.

[27] A. Nelson, W.H. Office of Science and Tech. Policy, OSTP Memorandum on Ensuring Free, Immediate, and Equitable Access to Federally Funded Research (Aug. 25, 2022), https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf.

[28] See, e.g., Nature, Scientific Data Journal, Data Repository Guidance, available at https://www.nature.com/sdata/policies/repositories; PLOS Recommended Repositories, available at https://journals.plos.org/plosone/s/recommended-repositories.

[29] See, e.g., Stephen Dubner, Freakonomics Radio, Can Academic Fraud Be Stopped? (Jan. 1, 2025), https://freakonomics.com/podcast/can-academic-fraud-be-stopped-update/.

is, the success of the overall scientific research endeavor hinges on broader disclosure of the data in question—which risks becoming foreclosed by the use of commercial data brokers.

In summary, we offer this comment letter as policy and legal experts on the commercial data broker industry and its extensive history of shady business practices and ethical, legal, and privacy lapses. **Should the NTIA decide to draft and issue ethics guidelines for research using pervasive data, we hope that you will incorporate these concerns about data brokers and counsel against their use by researchers given the perils for the human subjects and the researchers alike.**

Thank you for considering our views, and if you have any questions, please do not hesitate to contact us.

Sincerely,

Brennan Center for Justice
Demand Progress Education Fund
Electronic Privacy Information Center (EPIC)
Surveillance Technology Oversight Project