FUTURE OF PRIVACY FORUM

1350 Eye Street NW, Suite 350, Washington, DC 20005 | 202-768-8950 | fpf.org

June 5, 2023

Via Electronic Mail

Comment Intake, Request for Information Regarding Data Brokers
Consumer Financial Protection Bureau
c/o Legal Division Docket Manager
1700 G Street NW
Washington, DC 20552

Dear Director Chopra and Consumer Financial Protection Bureau staff:

On behalf of the Future of Privacy Forum (FPF), we are pleased to submit comments in response to the Consumer Financial Protection Bureau (CFPB)'s "Request for Information (RFI) Regarding Data Brokers and Other Business Practices Involving the Collection and Sale of Consumer Information."[1] FPF is a global non-profit organization dedicated to advancing privacy leadership, scholarship, and principled data practices in support of emerging technologies.[2]

In 2021, FPF explored the landscape of the current data broker industry in testimony presented to the Senate Finance Subcommittee on Fiscal Responsibility and Economic Growth (Attached, Appendix A). Since then, emerging and evolving data practices have continued to create potential risks for individuals and to raise novel questions about the scope of the Fair Credit Reporting Act (FCRA). These practices include the growing use of generative Artificial Intelligence (AI) and Large Language Model (LLM) systems,[3] the increased sophistication of algorithmic profiling,[4] and the expanding significance of data aggregators within the financial system.[5]

---

[1] 88 Fed. Reg. 16951 (Mar. 21, 2023) https://www.federalregister.gov/documents/2023/03/21/2023-05670/request-for-information-regarding-data-brokers-and-other-business-practices-involving-the-collection.

[2] The opinions expressed herein do not necessarily reflect the views of FPF's supporters or Advisory Board.

[3] *See* Andrew Smith, "Using Artificial Intelligence and Algorithms, Federal Trade Commission," Federal Trade Commission (Apr. 8, 2020), https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms.

[4] *See* Federal Trade Commission, "Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues," (Jan. 2016), https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf.

[5] *See* "Finicity Promotes FCRA Accountability for Financial Data Aggregators," BusinessWire (Feb. 11, 2021), https://www.fintechfutures.com/techwire/finicity-promotes-fcra-accountability-for-financial-data-aggregators/.

Meanwhile, the exclusion of FCRA-covered activities from the state-level comprehensive privacy laws passed in recent years reinforces the critical need for federal leadership to establish jurisdictional clarity and address privacy risks.[6]

We recommend that the CFPB analyze the broad range of business activities that can be considered "data brokerage," and use the Bureau's regulatory instruments to address specific risks posed by emerging and evolving technologies and business practices, including:

- the use of LLMs and generative AI systems as research tools;[7]
- the role of "alternative risk scoring" and other forms of algorithmic profiling, particularly in the employment and financial contexts;[8]
- the increasingly important role of data aggregators within the financial system.[9]

Through Policy Statements, rulemaking, enforcement, and other tools at the CFPB's disposal, the agency should clearly communicate the Bureau's expectations to organizations and individuals, reinforce the ongoing application of existing law to new technologies, and address specific novel compliance issues raised by these activities.

1. THE TERM "DATA BROKER" ENCOMPASSES A WIDE RANGE OF BOTH FCRA AND NON-FCRA COVERED ACTIVITY

With this RFI, the CFPB "endeavors to gain insight into the full scope of the data broker industry."[10] The data broker industry is challenging to fully comprehend, at least in part because the term "data broker" can apply to a wide range of companies and business activities, and there is substantial disagreement regarding whether some activities and entities should be classified as "data brokers."[11] FCRA typically governs data brokers when they act as credit reporting agencies

---

[6] *See ex.,* The California Consumer Privacy Act (as modified by the California Privacy Rights Act), Cal. Civ. Code §1798.145(d); The Colorado Privacy Act, Colo. Rev. Stat. §6-1-1304(2)(i)(II); The Connecticut Data Privacy Act, Public Act No. 22-15 §3(b)(11); The Virginia Consumer Data Protection Act, § 59.1-576 (10).

[7] *See, ex.,* Andrew Smith, "Using Artificial Intelligence and Algorithms, Federal Trade Commission," Federal Trade Commission (Apr. 8, 2020), https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms.

[8] *See, ex.,* Mikella Hurley & Julius Adebayo, Credit Scoring in the Era of Big Data, 18 YALE J.L. & TECH. 148 (2016); Sahiba Chopra, Current Regulatory Challenges in Consumer Credit Scoring Using Alternative Data-Driven Methodologies, 23 Vanderbilt Journal of Entertainment and Technology Law 625 (2021).

[9] *See* "Finicity Promotes FCRA Accountability for Financial Data Aggregators," BusinessWire (Feb. 11, 2021), https://www.fintechfutures.com/techwire/finicity-promotes-fcra-accountability-for-financial-data-aggregators/.

[10] 88 Fed. Reg. 16951 (Mar. 21, 2023) https://www.federalregister.gov/documents/2023/03/21/2023-05670/request-for-information-regarding-data-brokers-and-other-business-practices-involving-the-collection.

[11] *See* US Government Accountability Office, "Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace" (Sept. 2013) at 4 (noting "Characterizing the precise size and nature of the industry can be difficult because definitions for resellers vary and data on resellers often are limited or not comparable."), https://www.gao.gov/assets/gao-13-663.pdf

(CRAs)[12] and prepare consumer reports[13] for the purpose of establishing an individual's eligibility for credit, employment, insurance, government benefits, and select other statutory purposes.[14] However, as Stacey Gray's 2021 testimony before the Senate Finance Subcommittee (Appendix A) explains, a broad swath of data broker activity falls outside of FCRA, including data aggregated and sold for the purposes of marketing and advertising,[15] appending and matching services,[16] some people search services,[17] fraud detection,[18] identity verification,[19] some alternative risk scoring,[20] and socially beneficial research initiatives.[21]

Adding to the confusion is the fact that many data brokers have both FCRA-covered and non-FCRA covered books of business, acting as CRAs for the purposes of some, but not all, of their business activities.[22] Non-financial regulatory frameworks, including state-level privacy and data broker registry laws, govern data broker activity in many of these non-FCRA contexts, as well as brokers without FCRA-covered business. While this comment does not attempt to provide a complete taxonomy of the data broker market, we do suggest that any new regulatory actions that the CFPB takes to regulate data brokers should reflect the industries' broad scope. Likewise, the CFPB's statutory authority supports the agency in targeted regulation of specific segments of the data broker industry, e.g. brokers that participate in the financial industry or brokers that operate as CRAs under FCRA.[23]

---

[12] The Fair Credit Reporting Act (hereinafter "FCRA"), 15 U.S. Code § 1681(f).
[13] FCRA, 15 U.S.C. § 1681(d).
[14] FCRA, 15 U.S.C. § 1681(a).
[15] *See* Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability" (May 2014) at 23, https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.
[16] *See* 2020 NAI Code of Conduct (Network Advertising Initiative), page 8-B, "audience matched advertising," https://thenai.org/wp-content/uploads/2021/07/nai_code2020-1.pdf.
[17] *See* Adi Robertson, "The Long, Weird History of Companies that Put Your Life Online," Wired (Mar. 21, 2017), https://www.theverge.com/2017/3/21/14945884/people-search-sites-history-privacy-regulation.
[18] *See* Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability" (May 2014) at 67, https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-%20report-federal-trade-commission-may-2014/140527databrokerreport.pdf.
[19] *See ex.,* Margaret Oates, "Identity verification: flows we've seen in CCPA data requests (2 of 2)" (July 7, 2022) at "Identity questionnaires," https://innovation.consumerreports.org/identity-verification-flows-weve-seen-in-ccpa-data-requests-2-of-2/.
[20] Hurley & Adebayo at 187 (noting that "many collection and reporting activities [conducted by alternative risk scoring companies] may fall outside of FCRA's bounds.").
[21] *See ex.,* Future of Privacy Forum, "Stanford Medicine & Empatica, Google And Its Academic Partners Receive FPF Award For Research Data Stewardship," (June 28, 2021), https://fpf.org/blog/stanford-medicine-empatica-google-and-its-academic-partners-receive-fpf-award-for-research-data-stewardship/.
[22] Hurley & Adebayo at 187.
[23] Dodd-Frank §1021 (codified as amended at 12 U.S.C. §5511).

2. STATE LAW EXCLUSIONS FOR FCRA-COVERED DATA AND ENTITIES UNDERSCORE THE NEED FOR LEGAL CLARITY FOR BOTH CONSUMERS AND COVERED ENTITIES

Many state-level privacy laws–including the two state-level comprehensive privacy laws currently in effect[24] and the two that will take effect July 1, 2023[25]–exempt FCRA-covered data, as well as GLBA-covered data or entities. Two main implications stem from this exemption framework: (1) because state privacy frameworks often do not cover FCRA-covered data, inadequate enforcement of FCRA where it applies may leave certain personal data unprotected and (2) clarity about where and how FCRA does apply is crucial for both consumers seeking to understand their rights in respect to certain data uses and regulated entities endeavoring to understand their compliance obligations when engaging in different activities. The CFPB is well-positioned to and should provide federal leadership by clarifying the scope of FCRA as applied to emerging data uses or in places where uncertainty persists.

3. THE CFPB SHOULD ADDRESS SPECIFIC RISKS OF EMERGING AND EVOLVING TECHNOLOGIES

We recommend that the CFPB continue to investigate and understand the broad range of business activities that can be considered "data brokerage" and that fall within the scope of FCRA, and use its range of regulatory tools to address the specific benefits and risks of emerging technologies and business practices. Emerging data practices that create particular risks for consumers and may in some instances be subject to novel considerations under FCRA include the use of generative AI, alternative risk scoring and algorithmic profiling; and financial data aggregation.

- Generative AI systems

Generative AI "is a type of AI that has been trained on data so that it can produce or generate content similar to what it has been trained on."[26] LLMs, meanwhile, are a subset of generative AI systems, which generate new text based on training data.[27] The training of generative AI systems and LLMs requires large quantities of data, which companies in some instances may acquire from

[24] *See ex.,* The California Consumer Privacy Act (as modified by the California Privacy Rights Act), Cal. Civ. Code §1798.145(d); The Virginia Consumer Data Protection Act, § 59.1-576 (10).
[25] The Colorado Privacy Act, Colo. Rev. Stat. §6-1-1304(2)(i)(II); The Connecticut Data Privacy Act, Public Act No. 22-15 §3(b)(11);
[26] Stephanie Wong, "Let's Look at LLMs: Understanding Data Flows and Risks in the Workplace," Future of Privacy Forum (Mar. 30, 2023), https://fpf.org/blog/lets-look-at-llms-understanding-data-flows-and-risks-in-the-workplace/.
[27] Id.

data brokers.[28] Such brokers act as "furnishers" under FCRA when they supply data to CRAs to be used for FCRA-covered purposes–including the production of analyses about a consumer's creditworthiness, eligibility for insurance, or fitness for employment–whether CRAs use AI or more traditional methods to compile that data in consumer reports.[29] Likewise, entities that use AI systems to make decisions based on such reports are acting as "users of consumer reports" under FCRA.[30]

In both of these instances, this activity would fall within the scope of and be subject to all of FCRA's requirements. For CRAs, this includes the requirement to use "reasonable procedures to ensure maximum accuracy" of the information included in consumer reports[31] and to investigate the accuracy of information contained in such reports when disputed by a consumer.[32] Such requirements will likely require novel interpretation as applied to AI systems, especially when the outputs of such systems are not readily explainable.[33] For example, if it is not clear exactly how a generative AI tool produced a certain output, are there ways for companies to demonstrate that outputs are maximally-accurate? Likewise, if a consumer identifies an inaccurate piece of data as having been included within a generative AI training data set, does correction of that inaccuracy require complete retraining of that generative AI model? To enforce FCRA in the generative AI context, the CFPB will have to wrestle with a number of complex and original questions such as these.

- Alternative risk scoring and other forms of algorithmic profiling[34]

Alternative risk scoring (also often called "alternative credit scoring") is an umbrella term for the process of using so-called "alternative data," including "payments data for non-loan products such as phone payments; rent, insurance, and utility bill payments; checking account transaction-level data; data related to a consumer's educational and occupational history;

---

[28] *See, ex.* Alexander Alben, "When Artificial Intelligence and Big Data Collide—How Data Aggregation and Predictive Machines Threaten our Privacy and Autonomy," AI Ethics Journal (2020), https://www.researchgate.net/profile/Alex-Alben/publication/346745060_When_Artificial_Intelligence_and_Big_Data_Collide-How_Data_Aggregation_and_Predictive_Machines_Threaten_our_Privacy_and_Autonomy/links/61f2bcfa8d338833e39bb096/When-Artificial-Intelligence-and-Big-Data-Collide-How-Data-Aggregation-and-Predictive-Machines-Threaten-our-Privacy-and-Autonomy.pdf; "Devin Coldewey, "AI is more data-hungry than ever, and DefinedCrowd raises $50M B round to feed it," TechCrunch (May 26, 2020), https://techcrunch.com/2020/05/26/ai-is-more-data-hungry-than-ever-and-definedcrowd-raises-50m-b-round-to-feed-it/.

[29] Andrew Smith, "Using Artificial Intelligence and Algorithms, Federal Trade Commission," Federal Trade Commission (Apr. 8, 2020), https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms.

[30] Id.

[31] FCRA, 15 U.S.C. § 1681e(b).

[32] FCRA, 15 U.S.C. § 1681i(a)(1).

[33] *See, e.g.,* Violet Turri, "What is Explainable AI?", Carnegie Mellon Software Engineering Institute Blog (Jan. 17, 2022), https://insights.sei.cmu.edu/blog/what-is-explainable-ai/.

[34] Paul Ohm and Scott Peppet, "What if Everything Reveals Everything?" in *Big Data is Not a Monolith* (Cassidy R. Sugimoto, Hamid R. Ekbia, and Michael Mattioli ed., 2016).

consumer behavioral data; and data derived from a consumer's social media network" to assess a consumer's creditworthiness.[35]

While alternative risk scoring can help provide consumers without traditional credit scores with access to credit, the creation and use of such scores also creates unique risks.[36] In particular, the "[d]ata points collected and used" for such alternative scoring "are increasingly vast,"[37] creating an environment in which it may be difficult if not impossible for companies to honor key consumer rights under FCRA, such as the right to correct inaccuracies[38] or the requirement that CRAs remove certain negative data points from consumer reports after a set amount of time has passed.[39]

- Consumer profiles generated by financial data aggregators

Finally, an emerging trend in the financial system is the increasingly active role of data aggregators, some of whom bundle and sell consumer financial information to third party businesses.[40] Such aggregators source bundled data from many consumers and different account types, such as savings, credit cards, and mortgages.[41] Third party business buyers can obtain insights into consumer behaviors from this bundled data, including information about real-time trends in people's spending, that may inform what products and services these parties offer to consumers.[42] These practices, even when they inform crucial decisions about the financial products or advice offered to consumers, are not always covered by FCRA.[43]

4. CONCLUSION

The use of generative AI in credit scoring-like processes, the rise of alternative risk scoring, and role of financial data aggregators within the open banking system all pose novel risks for consumers and raise thorny questions about the scope and application of FCRA. The CFPB is well-positioned to clearly reinforce the ongoing application of existing law to new technologies and address specific novel compliance issues raised by these activities through Policy Statements, rulemaking, enforcement, and the other regulatory tools at its disposal.

---

[35] Chopra at 628.
[36] Id. at 627 (noting that, "there are approximately forty-five million people, primarily from Black and Hispanic backgrounds, who are considered "unscorable" because credit-scoring firms are unable to provide an assessment of their credit risk using traditional scoring tools.").
[37] Hurley & Adebayo at 188.
[38] FCRA, 15 U.S.C. § 1681i.
[39] FCRA, 15 U.S.C. § 1681c
[40] *See, ex.* Julian Alcazar & Fumiko Hayashi, "Data Aggregators: The Connective Tissue for Open Banking," Federal Reserve Bank of Kansas City Payments System Research Paper (Aug. 24, 2022).
[41] Id.
[42] "Finicity Promotes FCRA Accountability for Financial Data Aggregators," BusinessWire (Feb. 11, 2021), https://www.fintechfutures.com/techwire/finicity-promotes-fcra-accountability-for-financial-data-aggregators/.
[43] Id.

The data broker industry is both diverse and complex, and is regulated by a range of frameworks, including state laws. As emerging data uses change the way that numerous entities, including data brokers, interact with, aggregate, and analyze consumer data, we commend the agency's attention to understanding the contemporary information ecosystem and FCRA's intersection with a range of data uses.

We look forward to answering any questions and to working with the CFPB on these important issues. If you have any questions regarding these comments please contact Felicity Slater at fslater@fpf.org (cc:info@fpf.org).

Sincerely,

Felicity Slater
Policy Fellow
Future of Privacy Forum

Written Testimony of Stacey Gray

Senior Counsel, Future of Privacy Forum

Before the US Senate Finance Subcommittee on Fiscal Responsibility and Economic Growth

"Promoting Competition, Growth, and Privacy Protection in the Technology Sector"

Tuesday, December 7, 9:30AM

Chair Warren, Ranking Member Cassidy, and Members of the Subcommittee, thank you for the opportunity to testify today on the important issue of consumer privacy in the technology sector. Specifically, I've been asked to discuss the subject of data brokers and consumer privacy, an important and highly relevant topic as Congress continues to work towards enacting a federal comprehensive data privacy law.

As a Senior Counsel at the Future of Privacy Forum,[1] I work on public policy related to the intersection of emerging technologies, business practices, and U.S. consumer privacy regulation. The Future of Privacy Forum is a 501c3 non-profit organization, based in Washington, DC, specializing in consumer privacy and dedicated to helping policymakers, privacy professionals, academics, and advocates around the world find consensus around responsible business practices for emerging technology.

Let me begin by observing that attention to this topic is not new. Privacy advocates, the Federal Trade Commission, and members of the Finance Committee and other Senate Committees[2] have long called for greater transparency, accountability, and regulation of the data broker industry. This includes reports from the Government Accountability Office (GAO) in 2013,[3] the Federal Trade Commission (FTC) in 2014,[4] and the research and

---

[1] https://www.fpf.org. The views expressed in this testimony are my own, and do not represent the views of FPF's supporters or Advisory Board. *See* Future of Privacy Forum, Advisory Board, https://fpf.org/about/advisory-board/; Supporters, https://fpf.org/about/supporters/.

[2] Majority Staff Report for Chairman Rockerfeller, "A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes," Committee on Commerce, Science, and Transportation, Office of Oversight and Investigations (Dec. 18, 2013), *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

[3] Government Accountability Office, "Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace," GAO-13-663 (Sept. 2013), https://www.gao.gov/assets/gao-13-663.pdf.

[4] Federal Trade Commission, "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.

advocacy of academic scholars and leaders, including Pam Dixon of the World Privacy Forum,[5] and my fellow witnesses here today.

Since many of these reports were published almost a decade ago, much has changed. There have been significant advances in machine learning, the ability of systems to learn, adapt, and generate inferences from large datasets, with varying accuracy. Adoption of consumer technology has also become nearly universal, with 97% of US adults now owning a smartphone,[6] and most adults owning several additional devices -- a fact which has led to fragmentation in marketing industries, and incentives for many businesses to collect even more data to attribute and measure behavior across devices.[7]

The legislative landscape is also evolving. Since 2018, California and two other states have passed non-sectoral consumer privacy legislation,[8] and three states have established limited data broker-specific regulation -- California,[9] Nevada,[10] and Vermont.[11] Some state efforts have focused on transparency, through the establishment of Data Broker Registries, while others, such as the California Privacy Rights Act, codify consumer rights to opt-out of the sale of data and limit the use of sensitive information. Much more work remains to be done.

In the context of this evolving landscape, I'd like to make two substantive points regarding the data broker industry, and then provide three recommendations.

1. First: Defining the term **"data broker"** is a challenge for many regulations, because it encompasses a broad spectrum of divergent companies and business activities. The GAO has used the phrase "information resellers,"[12] and the leading definition from current state law includes any commercial entity that "collects and sells [or licenses] the

[5] World Privacy Forum, https://www.worldprivacyforum.org/.
[6] Pew Research Center, "Mobile Fact Sheet" (Apr. 7, 2021), https://www.pewresearch.org/internet/fact-sheet/mobile.
[7] Jules Polonetsky and Stacey Gray, Future of Privacy Forum, *Cross-Device: Understanding the State of State Management* (2015), https://fpf.org/wp-content/uploads/2015/11/FPF_FTC_CrossDevice_F_20pg-3.pdf.
[8] *See* California Consumer Privacy Act, Cal. Civ. Code § 1798.100 (2018); California Privacy Rights Act, Cal. Civ. Code § 1798.100 (2020); Virginia Consumer Data Protection Act, Va. Code Ann. § 59.1-571 (2021); Colorado Privacy Act, Colo. Rev. Stat. § 6-1-1301 (2021).
[9] Data Broker Registration, Cal. Civ. Code §§ 1798.99.80-88 (2020).
[10] In 2021, Nevada updated its existing law governing operators of online services, providing consumer rights specific to qualifying data brokers. See Heather Sussman & David Curtis, Orrick, "Nevada Expands Online Privacy Law; Goes for Brokers" (July 1, 2021), https://www.orrick.com/en/insights/2021/07/Nevada-Expands-Online-Privacy-Law-Goes-for-Brokers.
[11] Vermont Data Broker Regulation, 9 V.S.A. § 2430 (2018).
[12] Government Accountability Office, "Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace," GAO-13-663 (Sept. 2013), https://www.gao.gov/assets/gao-13-663.pdf.

personal information of a consumer with whom the business does not have a direct relationship."[13]

Businesses that fall under this definition, including the 170 businesses registered and currently "active" in Vermont's Data Broker Registry,[14] or the 490 businesses currently registered in California,[15] use data for a wide range of purposes. Some of the information these businesses collect and sell is quite sensitive and closely linked to individuals, while other information is less sensitive or de-identified to some degree. Both registries, and most current definitions of data broker, exclude business activities that are regulated by the Fair Credit Reporting Act (FCRA)[16] (i.e., consumer reporting agencies and the use of credit reports for eligibility decisions in employment, insurance, and housing) or the Gramm-Leach Bliley Act (GLBA)[17] (i.e., financial institutions).

Commercial purposes that can fall outside of FCRA and GLBA include, but are not limited to:

- **Marketing and advertising -** Likely the largest category of typical "data broker" activities by revenue is for marketing and advertising,[18] including direct mail, online, and mobile advertising. Advertisers have long had the ability to purchase and curate lists of audiences (such as by demographics, zip code, or inferred interests).[19] Increasingly, data brokers and other large tech companies are interested in using web, mobile, and offline data to generate detailed predictions related to consumer purchasing intent, future behavior, psychological profiles,[20] lifestyle,[21] or sensitive information such as political affiliation or health

[13] Under the Vermont Data Broker Regulation, a Data Broker is "a business, or unit or units of a business, separately or together, that knowingly collects and sells or licenses to third parties the brokered personal information of a consumer with whom the business does not have a direct relationship." 9 V.S.A. § 2430(4)(A).
[14] Vermont Secretary of State, Corporations Division, "Data Broker Search" (last visited Dec. 3, 2021), https://bizfilings.vermont.gov/online/DatabrokerInquire/DataBrokerSearch.
[15] State of California Department of Justice, Office of the Attorney General, "Data Broker Registry" (last visited Dec. 3, 2021), https://oag.ca.gov/data-brokers.
[16] Fair Credit Reporting Act, 15 U.S.C. § 1681.
[17] Gramm-Leach-Bliley Act, 15 U.S.C. § 6801.
[18] *See* Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability" (May 2014) at 23, https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.
[19] In many cases, risks related to data depend on its use. For example, an audience list associated with "Interest in Motorcycles" could be used to send direct mail discounts from a local motorcycle repair shop, but could also be used by an insurance company to infer that individuals or households engage in risky behavior. *Id.* at vi.
[20] *See, e.g.,* AnalyticsIQ, "What We Do: Consumer Data" (last visited Dec. 3, 2021). https://analytics-iq.com/what-we-do.
[21] *See*, e.g. Experian's Mosaic ® USA (Dec. 2018) (last visited Dec. 3, 2021), https://www.experian.com/assets/marketing-services/product-sheets/mosaic-usa.pdf.

conditions.[22] Many advertising technology (ad tech) providers also use data to offer measurement for ad attribution, conversion, and related metrics.

- **Appending and matching services -** Many businesses provide matching services that allow companies to link, or append additional information, to their existing lists of customers.[23] In some cases, businesses offer specialized, isolated matching services, or "clean rooms," that allow for external partners to link datasets without sharing underlying data, often for reasons of data ownership or protecting privacy. For example, a healthcare institution might use a matching service to send information about clinical trials to patients with specific health conditions, without disclosing patient information to researchers.
- **People Search Databases -** People search databases are online search tools that provide free or paid access to information that can be found in public records, such as a person's home address, previous addresses, names of family members, DMV information, court records, and criminal records.[24]
- **Fraud detection -** Many companies offer commercial fraud detection services to institutions such as banks, healthcare institutions, and online retailers, to protect consumers and businesses against fraudulent activities.[25] Such services typically rely on a wide variety of data from public and private records, such as purchasing behavior, online behavior, or real-time behavioral data from devices.[26]
- **Identity verification -** The ability to accurately verify identity, or that an individual is who they say they are, is a key component of digital services across many

---

[22] Justin Sherman, "Data Brokers and Sensitive Data on U.S. Individuals" Duke Sanford Cyber Policy Program (Aug. 2021), https://sites.sanford.duke.edu/techpolicy/wp-content/uploads/sites/17/2021/08/Data-Brokers-and-Sensitive-Data-on-US-Individuals-Sherman-2021.pdf.

[23] *See* 2020 NAI Code of Conduct (Network Advertising Initiative), page 8-B, "audience matched advertising," https://thenai.org/wp-content/uploads/2021/07/nai_code2020-1.pdf.

[24] Examples of people search companies include Whitepages (whitepages.com); Truthfinder (truthfinder.com), BeenVerified (https://www.beenverified.com/), and Spokeo (https://www.spokeo.com/). *See also*, Adi Robertson, "The Long, Weird History of Companies that Put Your Life Online," Wired (Mar. 21, 2017), https://www.theverge.com/2017/3/21/14945884/people-search-sites-history-privacy-regulation, and Yael Grauer, "How to Delete Your Information From People-Search Sites" Consumer Reports (Aug. 20, 2020), https://www.consumerreports.org/personal-information/how-to-delete-your-information-from-people-search-sites-a6926856917.

[25] According to data released by the Federal Trade Commission, more than 2.1 million fraud reports were filed by consumers in 2020. Consumers reported losing more than $3.3 billion to fraud in 2020, up from $1.8 billion in 2019. Nearly $1.2 billion of losses reported last year were due to imposter scams, while online shopping accounted for about $246 million in reported losses from consumers. Federal Trade Commission, "New Data Shows FTC Received 2.2 Million Fraud Reports from Consumers in 2020" (Feb. 4, 2021), https://www.ftc.gov/news-events/press-releases/2021/02/new-data-shows-ftc-received-2-2-million-fraud-reports-consumers.

[26] *See, e.g.*, Tax N. et al. (2021) *Machine Learning for Fraud Detection in E-Commerce: A Research Agenda*. In: Wang G., Ciptadi A., Ahmadzadeh A. (eds) Deployable Machine Learning for Security Defense. MLHat 2021. Communications in Computer and Information Science, vol 1482. Springer, Cham. https://doi.org/10.1007/978-3-030-87839-9_2.

sectors.[27] Including for the estimated 1 billion people globally who do not have proof of identity and are thus prevented from accessing government services or excluded from basic financial services, individual "digital footprints" can offer opportunities for alternative approaches to digital identity verification.[28]

- **Alternative risk scoring -** Historically, credit scores provided by consumer reporting agencies (CRAs) include predictions of creditworthiness based on past loan repayment history and related information. A growing number of fintech and data broker companies have begun using data from other sources, such as rental history or payment of utility bills, to make similar predictions about risk.[29] Sometimes known as "alternative risk scoring," this can be used to extend lines of credit to consumers that are "thin-file," or have little to no formal credit history. However, such risk scoring has raised concerns about privacy, fairness, bias, and accuracy, when it involves predictions from data such as web browsing, search history, or social media. Alternative risk scoring is governed by FCRA when used for individual eligibility decisions, such as firm offers of credit, but in some cases may fall outside of the protections of FCRA, for example when involving household data or lead generation.[30]
- **Socially Beneficial Research Initiatives** - Commercial data contributes to a growing number of research initiatives that seek to harness data in support of socially beneficial goals, such as public health tracking, humanitarian efforts, disaster relief, and medical research. In 2020, FPF established an annual Award for Research Data Stewardship, recognizing collaborations between companies

---

[27] See Noah Katz and Brenda Leong, Future of Privacy Forum, "Now, On The Internet, Everyone Knows You're a Dog: An Introduction to Digital Identity" (Aug. 3, 2021) https://fpf.org/blog/now-on-the-internet-everyone-knows-youre-a-dog/. Notably, identity verification can also be an important responsibility for businesses in responding to consumer requests to access, delete, and control data under emerging consumer privacy laws. *See, e.g.*, Jennifer Ellan & Steven Stransky, "*The new CCPA draft regulations: Identity verification*," International Association of Privacy Professionals (June 30, 2020), https://iapp.org/news/a/the-new-ccpa-draft-regulations-identity-verification.

[28] Vyjayanti T. Desai, Anna Diofasi, and Jing Lu, *The global identification challenge: Who are the 1 billion people without proof of identity?*, World Bank (Apr. 25, 2018), https://blogs.worldbank.org/voices/global-identification-challenge-who-are-1-billion-people-without-proof-identity.

[29] *See generally*, Consumer Financial Protection Bureau, *CFPB Explores Impact of Alternative Data on Credit Access for Consumers Who Are Credit Invisible* (Feb. 16, 2017), https://www.consumerfinance.gov/about-us/newsroom/cfpb-explores-impact-alternative-data-credit-access-consumers-who-are-credit-invisible/.

[30] For an exploration of the boundaries of the Fair Credit Reporting Act, *see generally*, Testimony of Pam Dixon Before the US Senate Committee on Banking, Housing, and Urban Affairs: Data Brokers, Privacy, and the Fair Credit Reporting Act (June 11, 2019), https://www.banking.senate.gov/imo/media/doc/Dixon%20Testimony%206-11-19.pdf; and Sahiba Chopra, *Current Regulatory Challenges in Consumer Credit Scoring Using Alternative Data-Driven Methodologies*, 23 Vanderbilt Journal of Entertainment and Technology Law 625 (2021), https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1044&context=jetlaw.

and academic researchers that allow researchers to access commercial data with privacy and ethical safeguards.[31]

Some data broker activities provide clear benefits to consumers, such as the use of data for public health, or to protect financial accounts against fraudulent activity. Others primarily benefit the purchasers or users of the data, such as advertisers, with little or no accompanying benefit (or perceived benefit) to individuals. A key to effective regulation will be to draw nuanced distinctions based on sources of data, purposes of processing, limitations on sharing and sale, data sensitivity, and the potential for risk and harm to individuals and groups.

2. Second, the lack of a direct relationship with consumers that characterizes most "data brokers" is both at the heart of concerns around privacy, fairness, and accountability, while also presenting the greatest challenge for data privacy regulation.

Any business with a direct-to-consumer relationship, big or small, such as a retailer, restaurant, hotel, or social media network, can collect personal information about US consumers directly, indirectly, or through purchasing and appending it. In some cases, those "first party" companies can exercise enormous influence and market power.[32] However, there is still a degree of public accountability to users who are aware of who such companies are and can delete accounts or raise alarms when practices go too far. In addition, first party companies can directly present users with controls and tools to manage their data in an app, on a web site, through direct email communications, or other means.[33]

In contrast, a business lacking a direct relationship with consumers does not always have the same reputational interests, business incentives, or in some cases legal requirements, to limit the collection of consumer data, process it fairly, and protect it against exfiltration. In states such as California, where privacy law codifies the right to access, delete, or opt-out of the sale or sharing of data, consumers typically are not aware of what companies within the "data broker" category may process their information, how to reach them, or how to manage the hundreds of opt-out requests that would be necessary to control the disclosure of their information.[34]

---

[31] *See* Future of Privacy Forum Blog, FPF Issues Award for Research Data Stewardship to Stanford Medicine & Empatica, Google & Its Academic Partners (June 28, 2021), https://fpf.org/press-releases/fpf-issues-2021-award-for-research-data-stewardship/.
[32] Charlotte Slaiman, "Data Protection is About Power, Not Just Privacy," Public Knowledge (Mar. 3, 2020), https://www.publicknowledge.org/blog/data-protection-is-about-power-not-just-privacy.
[33] In some cases, the ability of advertisers to purchase data from data brokers can undermine the efforts of first party platforms to create greater transparency and control for users. *See, e.g.*, Privacy Risks with Facebook's PII-based Targeting: Auditing a data broker's advertising interface (FTC PrivacyCon), https://www.ftc.gov/system/files/documents/public_events/1223263/panel05_privacy_risks_fb_pii.pdf.
[34] *See* Maureen Mahoney, "California Consumer Privacy Act: Are Consumers' Digital Rights Protected?" Consumer Reports (Oct. 1, 2020),

At the same time, a lack of a consumer relationship means that businesses engaged in legitimate or socially beneficial data processing often cannot rely on traditional mechanisms of notice and consent. Affirmative consent, or "opt-in," may be impossible or impractical for a business to obtain, while "opting out" after the fact tends to be impractical for consumers to navigate. For this reason, consumer advocates and academics have long observed the problems of legal regimes that rely solely on consent: consumers can become overwhelmed with choices, and may lack the knowledge to assess future risks, complex technological practices (such as predictive analytics, machine learning, or AI), or future secondary uses.[35] These risks are especially acute in the data broker industry.

What does this mean? In some cases, consumer choice remains an appropriate component of consumer privacy frameworks; a lack of consent should prevent data processing in many circumstances. But choice cannot be the sole safeguard in consumer privacy rules. In other cases, data processing should not occur even *with* a person's consent, for example if the processing is inherently high-risk or harmful.[36]

In some circumstances, we should recognize there are socially beneficial uses of large datasets that cannot, for reasons of practicality or accuracy, hinge on consumer choice. For example, commercial research in the public interest may include allowing independent researchers to evaluate the effect of large platforms on mental health; understanding the effect of COVID-19 and public health efforts; enabling disaster relief, and mitigating bias and discrimination in AI.[37]

---

https://advocacy.consumerreports.org/wp-content/uploads/2020/09/CR_CCPA-Are-Consumers-Digital-Rights-Protected_092020_vf.pdf.

[35] *See e.g.*, Neil Richards and Woodrow Hartzog, "The Pathologies of Digital Consent," 96 Wash. U. L. Rev. 1461 (2019), *available at* https://openscholarship.wustl.edu/law_lawreview/vol96/iss6/11.

[36] Many proposals for federal privacy frameworks advanced by both industry and consumer advocacy groups have included categories of "prohibited" data practices that organizations processing personal information would be barred from engaging in, even with individual consent. *See e.g.*, Center for Democracy and Technology, CDT's Federal Baseline Privacy Legislation Discussion Draft (Dec. 13, 2018) (last visited Dec. 3, 2021), https://cdt.org/insights/cdts-federal-baseline-privacy-legislation-discussion-draft/ (proposing that federal law prohibit *per se* "unfair data processing practices," such as certain forms of biometric information tracking, precise geospatial information tracking, and probabilistic cross-device tracking); Compare to, *e.g.*, Privacy For America, "Principles for Privacy Legislation" (last visited Dec. 3, 2021), https://www.privacyforamerica.com/overview/principles-for-privacy-legislation/ (an industry-led proposal containing prohibitions on data misuse that would include (1) banning the use of data to make certain eligibility decisions outside outside existing sectoral laws, (2) banning the use of data to charge higher prices for goods or services based on certain personal traits, and (3) outlawing the use of personal information for stalking or other forms of substantial harassment).

[37] *See* Future of Privacy Forum & Anti-Defamation League, "Big Data: A Tool for Fighting Discrimination and Empowering Groups" (July, 2014), https://fpf.org/wp-content/uploads/2014/09/Big-Data-A-Tool-for-Fighting-Discrimination-and-Empowering-Groups-FINAL1.pdf.

In these cases, privacy law can offer other tools for protecting consumers, including: limits on collection of data; transparency; accountability; risk assessment and auditing; limitations on the use of sensitive data; and limitations on high-risk automated processing for making important decisions regarding individuals' life choices.

3. Recommendations:

First and foremost, Congress should pass baseline comprehensive privacy legislation that establishes clear limitations and rules for both data brokers and first-party companies that process individuals' personal information. Its primary purpose should be to address the gaps in the current U.S. sectoral approach to consumer privacy, which has resulted in incomplete legal protections. Currently, personal information collected within certain sectors, such as credit reporting, finance, and healthcare, are subject to longstanding federal safeguards, while commercial data outside of these sectors remains largely unregulated even when the data may be equally sensitive or high-risk.[38]

In the absence of comprehensive legislation, there are a number of steps Congress can take to address risks related to consumer privacy and data brokers. Legal protections specific to the industry (alone or as part of a comprehensive law) could play a useful role, for example, through a national registry or opt-out system that would build on, or standardize the work of California and Vermont. In practice, however, a comprehensive law that is not specific to particular technologies or business models will be most effective, fair, and interoperable with global frameworks such as the General Data Protection Regulation.

Other legal approaches include: 1) limiting the ability of law enforcement agencies to purchase information from data brokers, including information purchased as a workaround to evade the constitutional limitations on those agencies when seeking information directly; 2) enacting sectoral legislation for uniquely high risk technologies, such as facial recognition; or 3) updating existing laws, such as the Fair Credit Reporting Act, to more effectively cover emerging uses of data, for example in alternative consumer risk scoring.

Second, Congress should empower the Federal Trade Commission to continue using its longstanding authority to enforce against unfair and deceptive trade practices, through funding of enforcement, research, and consumer education; greater numbers of staff and

[38] For example, medical records held by hospitals and covered by the Health Insurance Portability and Accountability Act (HIPAA) are subject to federal privacy and security rules. However, equally sensitive commercial information or inferences about health conditions is largely unregulated when processed by app developers, search engines, or marketing and advertising firms, outside of the Federal Trade Commission's longstanding Section 5 authority.

the establishment of a Privacy Bureau, and civil fining authority to effectively police businesses.

And finally, legislators should ensure that, within reasonable limits, privacy regulation does not prevent the use of data for socially beneficial purposes that are in the public interest, such as identifying bias and discrimination, contributing to a fair and competitive marketplace, holding large platforms accountable through independent research, and contributing to generalizable scientific, historical, and statistical research and knowledge.

Thank you for this opportunity, and I look forward to your questions.