July 14, 2023

The Honorable Rohit Chopra

Director

Consumer Financial Protection Bureau

1700 G St NW

Washington, DC 20552

Dear Director Chopra:

As the Consumer Financial Protection Bureau (CFPB) works to better understand the information ecosystem and the important role “data brokers” play in collecting and disseminating data, we write to caution the Bureau against regulations that would hamstring societally productive uses of data or infringe on First Amendment protected speech.

The Software & Information Industry Association (SIIA) is the principal trade association for those in the business of information. SIIA represents over 450 companies reflecting the broad and diverse landscape of digital content providers and users in academic publishing, education technology, and financial information, along with creators of software and platforms used by millions worldwide, and companies specializing in data analytics and information services. Our mission is to protect the three prongs of a healthy information environment essential to that business: creation, dissemination and productive use.

Of course, marketplaces have limits, and privacy standards are essential to the business of information. SIIA supports a comprehensive, uniform federal privacy standard that provides meaningful privacy protections to consumers, including the selling and sharing of personal data.

At the outset, this RFI suggests the definition of a “data broker”as *“an umbrella term to describe firms that collect, aggregate, sell, resell, license, or otherwise share consumers’ personal information with other parties.”*[[1]](#footnote-1) SIIA members publish information ranging from business-to-business news, to curricula about the Civil War, to securities pricing, to databases of case law and other public records, to scientific, technical and medical articles. These publications are used for purposes ranging from academic research to corporate due diligence. Our members routinely help both businesses and government deter, prevent, and track down fraudulent criminal activity. These products and services enable commerce, prevent crime, and provide the building blocks of ideas: the invisible backbone of functioning markets and a functioning democracy.

As we explain in more detail below, that ecosystem is not an accident, but a direct result of constitutional design. The line between a First Amendment-protected publisher and a company that violates privacy is a fine one that is simply not recognized in the RFI’s definition. The Bureau’s definition invites a category error in which constitutionally protected publishing activity becomes conflated with real privacy or other harms that the Bureau is empowered to prevent. Those errors will, if enshrined in regulation, chill a number of the beneficial (and First Amendment-protected) activities in which SIIA members engage. We are particularly concerned that regulating such a broad swath of entities as “data brokers” will stymie valuable activities based solely on publicly-available data that protect Americans’ security and safety.

We do not believe this is the Bureau’s intended result. We offer the following comments with the hope that the CFPB will work constructively towards regulation that targets the *misuse* of data rather than its publication.

1. **“Data brokers” engage in societally productive enterprises that benefit consumers by enabling services that are currently taken for granted.**

Individuals, government agencies, and commercial enterprises depend on timely access to data, and its predictable transmission supports billions of dollars in commerce and crucial decisions in people’s everyday lives. SIIA member companies enable data uses and activities to which Americans are accustomed and which protect their security and safety.

These include:

* **Law enforcement.** Federal and local law enforcement agencies rely on “data brokers” to locate suspects, victims, and witnesses to crimes. Even agencies such as the Federal Bureau of Investigation rely on products provided by “data brokers” because they enable “FBI investigative personnel to perform searches from computer workstations and eliminates the need to perform more time-consuming manual searches of federal, state, and local records systems, libraries, and other information sources.”[[2]](#footnote-2)These private-sector tools can also organize and analyze publicly-available data, infer patterns, and efficiently sort relevant from irrelevant information. The government routinely uses these tools in investigations of large-scale, complex, and sometimes international criminal schemes that use shell companies to avoid detection. They have been used to investigate, for example, human trafficking and fentanyl distribution networks.
* **Combating money laundering, corruption, and terrorism.** Financial institutions and other businesses rely on both non-public and publicly available data sources to help them meet “know your customer,” anti-money laundering, anti-terrorism, and anti-human trafficking obligations, and to comply with other financial laws, regulations, and industry practices. For example, banks have specific obligations when opening accounts for Politically Exposed Persons (“PEP”) who are close relatives of senior government officials. PEP lists are compiled from publicly available media combined with non-public data. These services enable the implementation of best practices in line with international objectives for corporate governance and efforts to combat bribery and corruption around the world.
* **Child support enforcement.** State and local agencies use data sets provided by brokers to locate individuals who are delinquent in paying their child support obligations. The Association for Children for Enforcement of Support reports that public record information provided through commercial vendors helped locate over 75 percent of the “deadbeat parents” they sought.[[3]](#footnote-3)
* **Insurance.** Insurers of all shapes and sizes access such data each day to underwrite policies and pay claims.
* **Product safety.** Companies use this information to provide consumers and auto dealers with a vehicle's accident history, alerting consumers to whether they are potentially buying a “lemon,” and to put both dealers and consumers on notice that the vehicle is subject to a safety recall.
* **Tax Compliance and prevention of waste and abuse.** Governments use databases of real estate records to detect tax avoidance. Moreover, they use our members’ tools to investigate potential abuse of public benefits programs, such as pandemic payment fraud, by using inferences from a combination of public real estate records, social media posts, and other public sources.
* **Fraud prevention.** Identity theft and other forms of fraud are a constant threat to consumers and businesses alike. Companies often leverage public record data to authenticate consumers in order to prevent identity theft and fraud. This can include an insurance company obtaining data from the DMV—or a vendor reselling this data—in order to authenticate a consumer’s true identity and risks. It can also include asking “out of wallet” questions, which are those a fraudster would be unlikely to know, such as “Which of the following five addresses is a past home address of yours?” or “Which of the following cars did you once own?” The answers to these questions can be found in state real property records or publicly-available Uniform Commercial Code filings.
* **News-gathering and publishing.** Newspaper companies regularly obtain the information used in brokered data products to report on crimes, detect possible corruption or conflicts of interest, and publish stories involving the operation or safety of motor vehicles.

These valuable activities are enabled or performed by “data brokers,” and have emerged as a result of an ecosystem that has favored the free exchange of public domain information. It would be both a policy and, as discussed in more detail below, a constitutional mistake to regulate the publication of public domain information in ways that interfere with well-intentioned and ethical SIIA members that are enhancing consumer security and convenience.

1. **There is a legislative and regulatory consensus that “data brokering” is First Amendment-protected activity.**

An undifferentiated hostility to publishing activity will cause a number of undesirable policy results. Interests in privacy must be balanced against other societal interests, among them the freedom of speech. The Supreme Court has made clear that “the creation and dissemination of information is speech for First Amendment purposes.”[[4]](#footnote-4) The State may not infringe these rights to protect a generalized interest in consumer privacy.[[5]](#footnote-5)

California expressly recognized this concern during the drafting of the California Privacy Rights Act (CPRA), which was designed to clarify and remedy a series of problems within the California Consumer Privacy Act (CCPA). The CCPA contained an exemption for publicly-available information, but attempted to narrow the definition only to data used for a purpose compatible with that for which it was publicly maintained or made available in government records.[[6]](#footnote-6)

The constitutionally protected public domain consists not only of information released by the government, but also that which is widely available in private hands. For the reasons described in the attached memorandum, the CCPA’s restrictions on the dissemination of publicly available information violated the First Amendment rights of both the businesses whose speech they burdened and the users of the information who are entitled to receive it.[[7]](#footnote-7) This problem may be worsened by discriminating among speakers, or alternatively on the basis of speech content, which separately—and likely facially—violates the First Amendment.[[8]](#footnote-8)

In response to these speech-based concerns, California amended the CCPA to thoroughly protect the public domain. [[9]](#footnote-9)  California’s data broker registry does not cover public domain information. [[10]](#footnote-10) Since then, every state to enact a privacy bill has placed some variation of this language into their privacy laws.

1. **Even if publicly-available information is exempted, it remains a violation of the First Amendment to restrict the transfer of derivatives of this data, as well as to render these protections dependent on within which dataset the public information is found.**

The public domain that the First Amendment protects is, in most cases, a one-way ratchet. It extends beyond the publication of public domain information to the creation of inferences from such information. And, while public domain information can be combined with other, private information that can be regulated, personal information retains its first-amendment protection once removed from the context that created the consumer’s reasonable expectation of privacy in that information. Thus, although publicly-available information may be combined with non-public data subject to regulatory restrictions, this combination does not alter the status of nor reduce the first-amendment protection for the public data. In addition to failing judicial scrutiny, such attempts to regulate public data would produce absurd results.

The Colorado Privacy Act (CPA), for example, includes language restricting the use of “personal data obtained or inferred.”[[11]](#footnote-11) According to rulemaking interpreting the CPA, “sensitive data inferences” are covered information subject to consumer rights of deletion and the like.[[12]](#footnote-12) The problems with this construct become readily available when considered in a mundane fact pattern. Suppose, for example, an investigative reporter attempts to find out whether a fugitive named John Smith is living in Wisconsin, and cross-references court filings, real estate records, and newspaper articles to infer that Mr. Smith is living there. If the reporter infers that he is in fact living there, that is opinion. If the reporter confirms it, it becomes fact. In either case, the publisher of that information has a First Amendment right to disseminate it. An undifferentiated fear of the collection of public information will not sustain a regulation against the heightened First Amendment scrutiny that courts would apply.[[13]](#footnote-13)

Such an interpretation would also produce unintended and unworkable outcomes, enabling bad actors to veto their inclusion in databases and publications that businesses and non-profits provide to customers and others who use it for legitimate purposes such as public safety, fraud detection, and public health.

This position comes subject to two important caveats. First Amendment protection is not absolute, and certainly a narrowly tailored statute that advanced a compelling interest would pass muster, even with public domain information. But drafting such a speech regulation must be done with great precision, which appears to be lacking in the way that the notice discusses “data brokers.”

Second, once publicly available information is removed from the sweep of regulation, the Bureau has more latitude to regulate that which is truly “private.”[[14]](#footnote-14) However, once publicly-available information is excised from private information, it remains public: “John Smith” may not request deletion of his real estate records because they were once used in a credit report. Any other approach would also nonsensically shrink the pool of public information over time that is available for a raft of positive public endeavors.

**\* \* \***

Protection of privacy is a legitimate legislative priority, and SIIA supports efforts to provide meaningful protections to consumers while enabling critical, societally-beneficial uses of information and respecting constitutionally guaranteed free speech interests. We look forward to continued engagement with the Bureau on these important issues.

Thank you for considering our views.

Respectfully submitted,

Chris Mohr, President

Paul Lekas, Senior Vice President for Global Public Policy and Government Affairs

Software & Information Industry Association

1. FR Doc. 2023-05670 Filed 3-20-23; 8:45 am (emphasis added). The definition goes on to state: “Data brokers encompass actors such as first-party data brokers that interact with consumers directly, as well as third-party data brokers with whom the consumer does not have a direct relationship. Data brokers include firms that specialize in preparing employment background screening reports and credit reports. Data brokers collect information from public and private sources for purposes including marketing and advertising, building and refining proprietary algorithms, credit and insurance underwriting, consumer-authorized data porting, fraud detection, criminal background checks, identity verification, and people search databases.” [↑](#footnote-ref-1)
2. *Departments of Commerce, Justice, and State, the Judiciary, and Related Agencies Appropriations for Fiscal Year 2000: Hearings on H.R. 2670/S.1217 Before Subcomm. for the Depts of Commerce, Justice, and State, the Judiciary, and Related Agencies of the S. Comm. on Appropriations,* 106th Cong. 280 (1999). [↑](#footnote-ref-2)
3. Comments of Gail H. Littlejohn, Vice President, Gov't Affairs, & Steven M. Emmert, Dir., Gov't Affairs, Reed Elsevier Inc., LEXIS-NEXIS Group (Mar. 31, 2000), *available at* http:// www.sec.gov/rules/proposed/s70600/littlej1.htm; *see also Financial Information Privacy Act: Hearings on H.R. 4321 Before the H. Comm. on Banking and Financial Services*, 105th Cong. 100 (1998) (statement of Robert Glass). [↑](#footnote-ref-3)
4. *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 570 (2011). [↑](#footnote-ref-4)
5. See generally *E. Volokh, Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People from Speaking About You*, 52 Stan. L. Rev. 1049, 1081 (2000). [↑](#footnote-ref-5)
6. *See* Cal. Civ. Code § 1798.140(o)(2). California law also has activity-based exemptions that permit valuable activities. [↑](#footnote-ref-6)
7. *Stanley v. Georgia*, 394 U.S. 557, 564 (1969). [↑](#footnote-ref-7)
8. *See Rosenberger v. Rector & Visitors of Univ. of Va.*, 515 U.S. 819, 828 (1995) (“In the realm of private speech or expression, government regulation may not favor one speaker over another.”). [↑](#footnote-ref-8)
9. “‘Personal information’ does not include publicly available information or lawfully obtained, truthful information that is a matter of public concern. For purposes of this paragraph, “publicly available” means: information that is lawfully made available from federal, state, or local government records, or information that a business has a reasonable basis to believe is lawfully made available to the general public by the consumer or from widely distributed media, or by the consumer; or information made available by a person to whom the consumer has disclosed the information if the consumer has not restricted the information to a specific audience.” Cal. Civ. Code § 1798.140(v)(2). [↑](#footnote-ref-9)
10. *See* Cal. Civ. Code § 1798.99.80. [↑](#footnote-ref-10)
11. Colo. Rev. Stat. § 6-1-1303(25)(a). [↑](#footnote-ref-11)
12. 4 Colo. Code Regs. § 904-3. [↑](#footnote-ref-12)
13. *See Cox Broad. Corp. v. Cohn*, 420 U.S. 469, 494-495 (1975) (“even the prevailing law of invasion of privacy generally recognizes that the interests in privacy fade when the information involved already appears on the public record”). [↑](#footnote-ref-13)
14. *Compare* 12 CFR § 1016 (Regulation P). [↑](#footnote-ref-14)