Comments of the Software and Information Industry Association

Notice of Proposed Rulemaking on Executive Order 14117: "Preventing Access to Americans' Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern"

Department of Justice, National Security Division
Submitted November 29, 2024

The Software & Information Industry Association (SIIA) appreciates the opportunity to provide comments on the Department of Justice's (Department) Notice of Proposed Rulemaking (NPRM) to implement Executive Order 14117 (the Order) on "Preventing Access to Americans' Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern".

SIIA is the principal trade association for those in the business of information, including its aggregation, dissemination, and productive use. Our members include roughly 375 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.

We recognize the laudable goals of the Order and, as articulated in our response to the Department's ANPRM, appreciate the thoughtful approach taken by the Department to avoid unintended consequences of overly prescriptive requirements.[1] In this vein, we hope to point out specific concerns we have around potential inconsistencies and impracticalities in the Order, including the need for definitional clarity and regulatory alignments to prevent widespread confusion and unnecessary compliance costs.

[1] Comments of the Software & Information Industry Association. Advanced Notice of Proposed Rulemaking on Sensitive Personal Data. Department of Justice, National Security Division. Submitted on April 19, 2024. https://www.siia.net/wp-content/uploads/2024/04/SIIA-DOJ-ANPRM-Comment-4.19.24.pdf.

1. **The definition of "bulk sensitive data" should be revised to clarify that it does not apply to all data, and specifically to include a definition of "sensitive personal data."**

The NPRM's proposed definition of "bulk U.S. sensitive personal data" does not actually include a reference to "sensitive personal data."[2] Without this reference, the scope of the "bulk U.S. sensitive personal data" definition could be construed to cover all data – not just bulk "sensitive" data. This, in turn, risks that the Rule could be interpreted to apply to every "data brokerage" transaction.

We do not believe this is the Department's intent. Rather, the NPRM suggests that the data covered by the Rule is designed to fall into either the category of "bulk U.S. sensitive personal data," or, alternatively, "government-related data." To clarify this, we recommend supplementing the definition of "bulk sensitive data" with a clarification that the term means "a collection or set of bulk *sensitive* personal data relating to U.S. persons…" (emphasis added).

2. **The definition of "bulk sensitive data" should not include de-identified, anonymized, pseudonymized, or encrypted data.**

The proposed definition of "bulk U.S. sensitive personal data" specifically includes data that is "anonymized, pseudonymized, de-identified, or encrypted."[3] At a high level, we do not believe this makes sense or aligns with the intent of the EO to protect Americans' sensitive data, since such data is tied neither to an individual nor to an individual's device.

Although we appreciate the studies and reports around reidentification cited in the NPRM, the works cited do not definitively link the definition of covered data in the NPRM to the types of data that would, or even could, be used to reidentify individuals. Instead, the NPRM uses anecdotal evidence of reidentification to justify its conclusion that anonymous and deidentified data could be reidentified or present national security risks. As just one example, the NPRM's cited evidence relies heavily on the use of precise geolocation data as a source of risk.[4] Yet it fails to address the more nuanced question of how coarse geolocation data reduces this risk, and how it could reduce or eliminate reidentification.

---

[2] "Provisions Pertaining to Preventing Access to U.S. Sensitive Personal Data and Government-Related Data by Countries of Concern or Covered Persons," 89 Fed. Reg. 86116, 86205 (Oct. 29, 2024), *available at* https://www.federalregister.gov/documents/2024/10/29/2024-24582/provisions-pertaining-to-preventing-access-to-us-sensitive-personal-data-and-government-related-data#sectno-reference-202.206.

[3] *Id.*

[4] *See Id.* at 86127.

Deidentification, anonymization, pseudonymization, and encryption are fundamentally privacy-protective practices, and should be encouraged. As such, we believe the Department should exempt from the scope of "sensitive personal data" and "government-related data" personal data that has been deidentified, anonymized, pseudonymized, or encrypted in such a way as to render it no longer "personal data" as defined under U.S. state comprehensive privacy laws. This ensures that understandings of personal data meet existing requirements and guidance under U.S. privacy law. It would also prevent adding to the complex federal and state patchwork of regulations and confusing consumers and compliance efforts by redefining the term "personal data" to cover these types of data uniquely within the context of this EO.

Because the states employ different definitions for various anonymization practices, we would also encourage the Department to be clear about how it is using these terms. The terms "anonymized," "pseudonymized," and "de-identified," in particular, can be confusing because there is little agreement—among either technical and policy experts, legal practitioners, regulatory authorities, or state laws themselves—about how they should be used, and indeed which constitutes the highest level of anonymization. The Department should take care to determine how these categories should be defined according to their technicalities and process, or alternatively the level of risk they present.

3. The definition of "data brokerage" should be both clarified and cabined to the scope of the NPRM.

The Rule defines "data brokerage" as "the sale of data, licensing of access to data, or *similar commercial transactions* involving the transfer of data from any person ("the provider") to any other person ("the recipient"), where the recipient did not collect or process the data directly from the individuals linked or linkable to the collected or processed data" (emphasis added).[5] At the same time, the Rule defines "transaction" as "any acquisition, holding, use, transfer, transportation, exportation of, or dealing in any property in which a foreign country or national thereof has an interest."[6]

As such, "data brokerage" includes any scenario in which a covered entity, as part of a commercial transaction, provides a different covered entity with access to data that the recipient did not collect or process. We are concerned this definition is unintentionally broad because it technically extends beyond the Rule's definition of "bulk sensitive data." To make this

[5] *Id.* at 86207.
[6] *Id.* at 86205-06.

definition consistent with the remainder of the NPRM, we strongly suggest clarifying that "data brokerage" applies only to "bulk sensitive data."

Furthermore, the definition's inclusion of "similar commercial transactions," in addition to sales and licensure of data, is quite vague and leaves the scope of its application open to interpretation, especially when the underlying "data brokerage" definition includes both sales and licensing. It is thus unclear what "similar commercial transactions" means. For example, as subsequently discussed, it leaves open the question of whether digital advertising platforms that contract with third party advertisers to target ads within their own "walled gardens" are covered by this definition. We suggest that this part of the definition should be removed or narrowed to cover the specific types of commercial transactions the rule intends to cover.

Lastly, we appreciate the approach towards "data brokerage" rather than "data brokers," which eliminates problems associated with defining entire entities according to lines of business. However, we suggest that it will ease compliance efforts and reduce confusion for those entities engaging in "data brokerage" to align this definition's activities closer to emerging norms in U.S. state laws around data broker activities, such as the language of Vermont's or California's data broker registry laws.[7]

4. The definition of "data brokerage" should specifically exempt data sharing platforms.

The Rule's definition of data brokerage could also create significant confusion specifically around whether data sharing platforms are included. This is because the definition could be read to include companies that provide a platform for other entities to share data or advertise to potential customers. However, such data sharing platforms do not determine what data is shared or review the data before it is shared. In addition, such platforms do not typically insert themselves into the contractual relationship between data providers and their end-customers. As such, we suggest the Rule should clarify that it explicitly exempts such platforms from coverage. Without this, platforms may be put in the position where they need to review all data exchanges *and* the underlying data sets, creating significant new privacy and data security risks. Data sharing platforms would also be forced to intrude on contractual relationships between data providers and their end-customers which these platforms would not normally be a part of. Any such requirements should fall on the data provider, not the platform.

[7] 9 V.S.A. § 2430(4)(A); Cal. Civ. Code § 1798.99.80..

5. **The definition of "covered person" could lead to significant confusion and unintended consequences in specific situations.**

The definition of "covered person" under the proposed rule is "(1) is 50 percent or more owned, directly or indirectly, by a country of concern; (2) is organized or chartered under the laws of a country of concern; or (3) has its principal place of business in a country of concern."[8] Although thoughtful in its application to subsidiaries based in countries of concern, there are certain specific scenarios where this definition would still benefit from additional clarity.

We are particularly concerned that these definitions could create a chain of accountability scenario as data is used downstream. We suggest considering a scenario where a U.S. entity sells a data set to a third party. This third party has an office with employees in a country of concern, which has both American employees and employees with ties to the foreign government. A U.S. company sells to this third party's U.S. office and that third party permits that data to be accessed by the foreign office – and therefore the employees and government of this country of concern. In this case, it is unclear whether the U.S. company would need to confirm who each of their purchasers' employees are in each of their foreign offices to avoid liability. It is also unclear whether this U.S. business would be required to implement initial safeguards necessary to block these employees' access.

A related, but separate, concern is also the list of entities whose association with a country of concern restricts them from receiving data from U.S. companies. This is especially salient for entities on the list that are owned by a country of concern or an entity located in those countries. While there may be ways to determine how not to do business with a company located in a country of concern, the question of ownership is often nearly impossible for companies to determine from a compliance perspective.

6. **Uncertainty regarding diligence requirements will impose an extraordinary compliance burden on U.S. companies attempting to comply in good faith.**

The cost of compliance for businesses to conduct KYC is already significant. A Thomson Reuters study revealed that banks are already spending an average of $60 million annually on KYC compliance. In one finding, banks were spending up to 48 days to onboard a new business client, while in another, business clients reported that they were contacted an average of eight times during the onboarding process.[9]

[8] 89 Fed. Reg. at 86206.

[9] KYC: A Sound Principle but Complex Reality (2016), *available at* http://info.risk.thomsonreuters.com/COB-FI-Survey-Gated.

KYC diligence requirements where the method of execution is unclear is likely to only add to this cost and frustration. For example, to comply with the Order, it will be critical to be able to assess the degree of connection between a vendor or customer and either a country of concern or a covered individual. Some of this information, such as residence or country of incorporation, may be on hand. However, the extent to which specific sets of information are indicia of control or influence by a country of concern or covered individual may not be readily apparent. These indicia may also be substantially different than metrics used in either KYC or FCPA analysis.

We understand the Department's interest in creating flexibility for companies to create a risk-based compliance approach similar to that undertaken by the Office of Foreign Assets Control (OFAC). We nevertheless encourage the Department, during the course of its implementation, to weigh heavily the benefits of clarity that is provided by both advisory opinions and licenses as described under the Rule.

7. **The Rule would benefit from clarifying its scope regarding the employment of covered individuals by non-US affiliated companies.**

Although the Rule includes a discussion about the employment of covered individuals by U.S. companies, its scope is less clear regarding the employment of covered individuals by non-U.S. affiliated companies. Specifically, it is not clear the extent to which a foreign entity that includes a number of U.S. affiliates is covered. It is important for these entities to have clarity to avoid compliance ambiguity regarding the status of their worldwide employees, particularly those employees of those non-U.S. entities operating within designated countries of concern.

* * *

Thank you for considering our feedback to the proposed Rule. We are happy to discuss any of these comments in further detail, if helpful.