November 8, 2021

California Privacy Protection Agency
Attn: Debra Castanon, Chief Privacy Officer
915 Capitol Mall, Suite 350A
Sacramento, CA 95814

RE: Invitation for Preliminary Comments on Proposed Rulemaking Under the California Privacy Rights Act of 2020

Dear Ms. Castanon:

The Software & Information Industry Association (SIIA) appreciates the opportunity to submit comments about Proposition 24, The California Privacy Rights Act of 2020 (CPRA), which extends the California Consumer Privacy Act of 2018 (CCPA).

SIIA is the principal trade association for the software and digital information industries worldwide. Our members include nearly 450 companies, many based in California or primarily serving California residents. Our members include a range of broad and diverse digital content providers and users in specialized content industries, 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. On behalf of our members' wide interests and services, SIIA has long advocated for privacy protections.

Our members publish a variety of information projects including scientific, technical and medical journals, business to business publications, and databases of news articles and court decisions. They depend on the First Amendment-protected vibrant public domain consisting of both information released by the government and that which is widely available in private hands. The transmission of publicly available information is fully protected by the First Amendment, and we are gratified that CPRA fixed the CPPA's free speech defects. The CPRA revises the definition of personal information (and, separately, the definition of sensitive personal information), to exempt publicly available information from its definition.[1]

Our comments focus on honing the practical aspects of implementing CPRA, particularly as it concerns the wide range of members we serve. We also identify compliance-related challenges raised by several of the rulemaking topics.Our comments reinforce two specific recommendations on behalf of our members regarding amendments to CPRA: 1) revise and

[1] CA Civ Code §1798.140 (v) (2) - 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 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.

more narrowly define the term "sensitive personal information", to avoid first amendment conflicts; and 2) determine the focus, scope and impact of automated decision-making in order to determine how its output is governed and implemented under CPRA.

1. Clarify CPRA's Requirements with Respect to Limiting Use and Disclosure of Sensitive Personal Information

First, our members agree with the proposition that consumers should be able to opt-out of the sharing and sale of personal information when it violates reasonable expectations of privacy. But, as it concerns notifying consumers about businesses' use of their information, our members recommend a streamlined approach. CPRA requires companies to provide a clear link allowing consumers to limit the business' use of their sensitive personal information.[2] We recommend a simpler process of operationalizing this, by grouping this option with other consumer rights, rather than having to comply with this aspect of the rule as a standalone requirement.

Second, CPRA defines "sensitive personal information"[3] to include a wide range of personal information, which is inclusive of: highly identifiable information that imposes a high risk; personal information that may already be governed by existing privacy laws; and lower-risk information that appears to be closely tied with publicly available information. As a practical matter, businesses may not be able to fulfill a consumer's request to access, limit and delete sensitive personal information (and certain personal information) if the use is reasonably necessary to fulfill a business or service-related obligation, or in circumstances where security and integrity may be compromised. CPRA explicitly adopts exemptions for businesses to comply with consumer's rights, if the action would disrupt the business' ability to exercise or defend legal claims.[4] CPRA also includes business exemptions for deleting consumers' personal information, in the instances where security and integrity are at odds.[5] Therefore, it would be beneficial for the Agency to extend a similar protection with respect to the consumers' right to access their personal information, when security or integrity of the business are at question. The Agency could do so by clarifying that security or integrity are an example of such an exemption to defend a legal claim. Circumstances where businesses can be exempt from fulfilling such requests could include: conducting biometric screenings for authentication purposes, providing fraud prevention or anti-money laundering services, providing age-appropriate content to minors, and participating in similar security and compliance-related activities. These activities may involve third parties and service providers

[2] CA Civ Code § 1798.135 (2018). Amended by Proposition 24.

[3] CA Civ Code § 1798.140 (ae) (1-3): "Sensitive personal information means personal information that reveals: (A) A consumer's social security, driver's license, state identification card, or passport number; (B) A consumer's account log-in, financial account, debit card, or credit card number in combination with any required security or access code, password, or credentials allowing access to an account; (C) A consumer's precise geolocation; (D) A consumer's racial or ethnic origin, religious or philosophical beliefs, or union membership; (E) The contents of a consumer's mail, email, and text messages unless the business is the intended recipient of the communication; (F) A consumer's genetic data."

[4] CA Civ Code § 1798.145. Exemptions.

[5] CA Civ Code § 1798.105. Consumers' right to delete personal information.

working to fulfill these obligations and allowing customers the opportunity to opt-out would not be plausible.

A separate but related area for potential revision is to exempt inference-based data out of the definition of personal information. As currently written, personal information includes inferences drawn "from any of the information used to build a profile about a consumer, which include the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes[6]". From a practical standpoint, businesses can use a combination of some sensitive personal information (e.g., trade union membership) and publicly available data (e.g., public records) to build inferences, probabilities, correlations, or couple publicly available data with proxy data (e.g., zip codes) to compile such inferences. If inferences are built from publicly available data and a combination of other personal and sensitive personal information, they should be exempt. Otherwise, there is a risk of regulating "potentially" sensitive personal information, which has not been fully validated as such.

The UK Information Commissioner's Office (ICO) provides further clarification about the situations when inference-based data should not be considered a "special category" of information[7], a term which may be comparable to CPRA's enforcement of sensitive personal information. The ICO suggests that the determination of whether processing inference-based data would trigger GDPR Article 9[8] is dependent on two factors: the level of certainty of the inference and the intent behind the inference. For example, inferences that are educated guesses would not trigger Article 9, whereas inferences processed specifically to treat someone differently on the basis of that inference would do so. It is important for CPRA to capture these nuances when it comes to implementation. Using the ICO's guidance is a relevant and logical model for the Agency to provide guidance, after stakeholder input, regarding appropriate use of inference-based data.

Inference-based data is the backbone of many businesses providing tailored or niche services, including profiling (discussed in detail, next) that require a range of information to fulfill a business obligation. Further, by allowing consumers to opt-out of the sharing and sale of inference-based data, it would significantly limit consumer choice and perpetuate inequities if only some consumers limit the sale of their data -- data that is pivotal to businesses that use this data to provide tailored services.

A more focused definition of sensitive personal information can be found in Virginia's privacy law, the Virginia Consumer Data Protection Act (VCDPA)[9], which includes the most

---

[6] CA Civ Code § 1798.140 (v)(1)(K).

[7] What is special category data? UK Information Commissioner's Office. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/#scd7

[8] Art. 9 GDPR – Processing of special categories of personal data - General Data Protection Regulation (GDPR) (gdpr-info.eu). https://gdpr-info.eu/art-9-gdpr/.

[9] VCDPA defines sensitive information to include: "Personal data revealing racial or ethnic origin, religious beliefs, mental or physical health; diagnosis, sexual orientation, or citizenship or immigration status; the processing of genetic or biometric data for the purpose of uniquely identifying a natural person; the personal data collected from a known child; or precise geolocation data."

relevant and high-risk personal information to avoid overbreadth and overreach. The VCDPA takes a different approach than CPRA: it excludes lower risk information, such as trade union membership, and does not include inference-based data in the definition of personal information.

We also note that proposing a narrower definition of sensitive personal information in CPRA that excludes driver's license, passport, and financial information would avoid inherent under sampling challenges that stem from collecting data on historically disadvantaged communities, because this type of information may not exist, may be less likely to exist, or may be less likely to be accurate when collected on these populations[10]. Some of this data is already governed by other laws (HITECH Act, HIPAA, GLBA, and others) and therefore can be exempted. We should avoid making policy decisions based on data that is not representative or could be processed and used in unbalanced and inappropriate ways.

2. Establish a Principles-Based Approach to Automated Decision-Making Technology, Focused on Fully-Automated Decision-Making Affecting Legal Rights with a Tailored Consumer Opt-Out

We urge the Agency to exercise prudence in approaching regulations on "automated decision-making technology," a concept that has no predicate in California statute or regulation. Automated technologies are used to render billions of decisions each day. Yet most of these decisions are not sufficiently tied to legal rights of natural persons and, we submit, have no meaningful effect on consumers to out-balance the potentially significant consequences that expansive rulemaking in this area will have on California and its residents.

Regulation of automated decision-making should, to the maximum extent possible, be both risk-based and technology-neutral. Because not all automated decision-making creates the same privacy risk, regulations should be tailored to the harms created by that risk. A standard that presumes that the use of automated decision technology is undesirable would hamstring many beneficial uses of automated technologies. Instead, we respectfully suggest that regulation in this context focus on decisions made solely on an automated basis that produce legal or similarly significant effects on a consumer. This will help ensure that California's rules provide an ongoing privacy framework to withstand technological advances.

In applying this approach, we respectfully offer the following guiding principles.

A. Distinguish Between Automated Decision-making and Automated Decision-making Technology

First, the Agency should pay close attention to the distinction between *automated decision-making*, and automated decision-making *technology (*and respective engines) *that drive the decision-making process*. The Agency's phrasing of the questions within Topic 2 suggests awareness of this distinction. Respectfully, we submit that the Agency should focus its rulemaking on the *decisions* that are generated by automated processes rather than on the

[10] Big Data and discrimination: perils, promises and solutions. A systematic review. Jornal of Big Data. SpringerOpen. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0177-4

technology itself. Regulating technology more broadly will have significant, unforeseen (and negative) consequences for consumers and businesses alike.

Companies across a wide range of industries today use technology to generate or inform a broad set of decisions. Automation serves a range of purposes, from personalizing and customizing content for groups of people and specific purposes, authenticating mobile apps, and providing fraud-detection alerts and security alerts and features, which powers basic processes in banking, retail, security, tech, publishing, automobile, and other industries.

B. Decisions Should be Based on Impact on Natural Person's Rights

Second, the Agency's approach to automated decision-making should be guided by clear objectives. The CPRA already provides the outer limits of those objectives. Specifically, the statute ties rulemaking on automated decision-making technology to the definition of "profiling":

> "Profiling" means any form of automated processing of personal information, as further defined by regulations pursuant to paragraph (16) of subdivision (a) of Section 1798.185, to evaluate certain personal aspects relating to a natural person, and in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, Interests, reliability, behavior, location or movements."[11]

This definition indicates a concern with how automated processes may be used to profile natural persons in a manner that has a direct effect on that natural person. Yet the definition of "profiling" leaves open what sort of "decisions" should be subject to regulations.

We recommend that the "decisions" should be those that have a direct effect on the legal rights of the natural person subject to automated decision-making. This approach is informed by the approach taken by the European Union. Article 22 of the GDPR protects consumers from decisions "based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."[12] We recommend that any rulemaking on automated decision-making be closely tied to those situations in which ***profiling is done through automated processes that have a direct effect on the legal rights of a natural person.***

C. Tailor Determination of Impact to Risk-based Approach

Third, we recommend the Agency focus its rulemaking on fully automated decisions – those that do not involve any human support. Similar to the intent of GDPR Article 22, the Agency could clarify that consumers have an opportunity to opt-out of automated

[11] CA Civil Code § 1798.140(z).

[12] Art. 22 GDPR – Automated individual decision-making, including profiling - General Data Protection Regulation (GDPR) (gdpr-info.eu). https://gdpr-info.eu/art-22-gdpr/

decision-making, only in instances when the consumer faces a "legal or similar effect" that is based on the automated decision-making (and does not involve human input or intervention). This creates a narrowly tailored, situation-based approach to opt-out that would require a high level of necessity, so as not to bog down businesses with unnecessary consumer requests to opt-out of routine, automated transactions.

Taking this approach once step further, and to avoid the situational vagueness that is embedded in the GDPR, we recommend that additional input be gathered to aggregate a list of specific use cases for when opt-out would be necessary and considered to have a legal or similar effect. The agency could share this list with businesses and consumers to provide further clarity and examples of when the opt-out would apply. GDPR Article 22 also allows some exemptions with regard to this right, including using data to enter into contracts and processing that is authorized by law or in circumstances when explicit consent is granted by the consumer.[13]

We propose narrowing opt-out requests to scenarios based on the potential for significant impact to the consumer, and using certain criteria to determine the impact, such as: a) whether the decision-making is based on fully or partially automated technology; b) the level of risk and material harm the decision would impose on the consumer; c) whether human input is a part of the decision-making process; d) the benefits to the business and the public from the use of the technology; and e) the irreversibility of the decision. A risk-based approach could allow consumers to opt out of profiling in life-altering or particularly challenging situations, such as access to essential goods or services (for insurance, healthcare, criminal enforcement, or other related purposes and activities).

Using such an approach, consumers should also be granted the right to obtain human intervention and the right to challenge decisions with legal or similarly significant effects, aligned to GDPR Article 22(3). Therefore, we recommend inclusion of an appeal process to ensure appropriate recourse. The appeal process would allow for additional consumer support in understanding the information granted to them and addressing any changes they require with regard to opting out of the information or opting back into the automated process, as needed. Additional, business process-related concerns about authenticating and answering consumer requests – including the types of information to be provided by the business, how to make this information most useful and readable to the consumer in Plain English, and how to standardize this information – could be answered through an additional request for stakeholder input or a rulemaking process.

D. Overly Prescriptive Rules on Profiling May Limit Innovation

Fourth, while we appreciate the steps that the legislature took to carve out publicly available information, we remain concerned about the manner in which the statute's broad limitations on "profiling" may inadvertently chill the expression of protected expression. The statute defines profiling as "any form of automated processing of personal information... to evaluate certain personal aspects relating to a natural person, and in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation,

[13] Artificial Intelligence (AI) and the GDPR - Part one - Data Protection - PwC UK blogs. https://pwc.blogs.com/data_protection/2019/01/artificial-intelligence-ai-and-the-gdpr-part-one.html

health, personal preferences, Interests, reliability, behavior, location or movements."[14] It is not difficult to envision, for example, an investigative reporter who might consult one of our members' products to identify a potential source through a combination of covered and publicly available information. All kinds of research draw inferences from both types of information, and we do not believe the intent of the legislature was to chill it unintentionally. Once again, we believe that the Agency's efforts would benefit from a more detailed and specific administrative record, by attempting to define these circumstances in more detail.

Notwithstanding these principles, we believe there are unique considerations when analyzing data generated by automated decision-making technologies and recommend the Agency host additional stakeholder input and hearings specifically to discuss this issue. To both adequately protect privacy and allow for innovation in the use and development of artificial intelligence, we urge policy makers to engage in fact-finding to fully understand this developing but technologically essential ecosystem.

In sum, we believe that changes along the lines above will both make CPRA a national model and support increased interoperability with other state consumer privacy laws. We thank you for the opportunity to submit comments. Please do not hesitate to reach out with further questions on this or other consumer privacy-related matters.

Respectfully submitted,

Paul Lekas, Senior Vice President for Global Public Policy

Divya Sridhar, Senior Director for Data Policy

Software & Information Industry Association

[14] CA Civil Code, sec. 1798.140 (z)