FUTURE OF PRIVACY FORUM

1400 Eye Street NW, Suite 450, Washington, DC 20005 | 202-768-8950 | fpf.org

November 8, 2021

California Privacy Protection Agency
Attn: Debra Castanon
915 Capitol Mall, Suite 350A
Sacramento, CA 95814
VIA EMAIL: regulations@cppa.ca.gov

RE: Future of Privacy Forum Comments, PRO 01-21

Dear Ms. Castanon and Members of the California Privacy Protection Agency,

The Future of Privacy Forum (FPF) welcomes this opportunity to weigh in on initial rulemaking under the California Privacy Rights Act. FPF is a 501(c)(3) non-profit organization that serves as a catalyst for privacy leadership and scholarship, advancing principled data practices in support of emerging technologies. Our primary office is in Washington, DC, and we work closely with our colleagues in Brussels, Singapore, Tel Aviv, and around the world. FPF seeks to support balanced, informed public policy and equip regulators with the resources and tools needed to craft effective regulation.[1]

In response to the Agency's invitation for comments, and with regard for the particular categories of information requested,[2] we offer resources and recommendations below regarding: automated decisionmaking, sensitive personal information, global opt-out signals, and de-identification.

Regulations under the California Privacy Rights Act should:

1. Establish guidelines for automated decisionmaking (ADM) that produces "legal or similarly significant effects."
2. Provide that information about "automated decisionmaking" follow NIST interpretability guidelines, and be meaningful and reasonably understandable to the average consumer.
3. Clarify a range of potential use cases for health and wellness data, by providing a principled, exemplar list of categories that are in or out of scope. In many cases, such distinctions will be based on context and reasonable use.
4. Ensure opportunities for socially beneficial commercial research using sensitive personal information.
5. Clarify the role of global opt-out signals in the context of today's labyrinth of existing permission frameworks, including in authenticated and non-authenticated platforms.
6. Establish an open process for authoritative approval of new global opt-out signals that meet the technical specifications of the Agency over time.

[1] The views herein do not necessarily reflect the views of our supporters or Advisory Board.
[2] California Privacy Protection Agency, Invitation for Preliminary Comments on Proposed Rulemaking Under the California Privacy Rights Act of 2020 (Sept. 22, 2021), https://cppa.ca.gov/regulations/pdf/invitation_for_comments.pdf.

7. Seek further input from de-identification experts and researchers to clarify key implementation issues for "deidentified data," including the role of technical, legal, and administrative controls, and Privacy Enhancing Technologies (PETs).

Part A. Automated Decisionmaking - 1798.185(a)(16)

1798.185(a)(16) requires the Agency to issue regulations governing access and opt-out rights with respect to the use of automated decisionmaking (ADM) technologies, including profiling. Although the CPRA does not specifically regulate automated decisionmaking (ADM), the concepts are useful for the purposes of understanding CPRA consumer rights (access, deletion, opt-out of sale and sharing, and limiting the use of sensitive personal information).

In general, we recommend that the Agency craft user controls to: (1) address the potential for harms caused by automated decisionmaking when it leads to "legal or similarly significant effects" on consumers, including clarifying when the use of sensitive personal information may be "necessary to perform the service or provide the goods reasonably expected by an average consumer" to identify and address bias and discrimination in high-risk decisions. Regulations should also provide (2) that information about "automated decisionmaking" follow NIST interpretability guidelines, and be meaningful and understandable to the average consumer.

1. Regulations should establish guidelines for automated decisionmaking (ADM) that produces "legal or similarly significant effects."

Strictly interpreted, "automated decisionmaking" encompasses almost every form of modern technology. This includes many routine, low-risk practices, such as loading a website, email filtering, or providing content recommendations.[3] However, some commercial automated decisions present serious risks to individual rights and autonomy, particularly in areas such as hiring, tenant screening, insurance, and other risk scoring.[4] Many of the most serious use cases fall outside the scope of CPRA (e.g., AI used in criminal sentencing, or by HIPAA-covered entities to make diagnosis decisions).

In order to distinguish higher risk automated decisionmaking from the broader world of all technology that involves "automation" (that is, all technology), a helpful guidepost would be to

[3] For a relevant comparison, the federal government maintains a list of automated processes in its Robotic Process Automation Use Case Inventory, providing detailed information on over 300 RPA Use Cases across the federal government. *See* U.S. General Services Administration, Federal Robotic Process Automation (RPA) Community of Practice, RPA Use Case Inventory, https://digital.gov/pdf/federal-rpa-use-case-inventory-compliant.pdf.

[4] *See, e.g.*, Miranda Bogen and Aaron Rieke, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias* (Dec. 2018), Upturn, https://www.upturn.org/static/reports/2018/hiring-algorithms/files/Upturn%20--%20Help%20Wanted%20-%20An%20Exploration%20of%20Hiring%20Algorithms,%20Equity%20and%20Bias.pdf.

align the CPRA regulations with Article 22 of the GDPR by applying heightened protections to automated decisions that lead to "legal or similarly significant effects."[5] The standard "legal or similarly significant effects" has the benefit of capturing high-risk use cases, while encouraging interoperability with global frameworks, for which a growing amount of legal guidance is becoming available.

According to leading guidance in the European Union,[6] decisions with "legal effects" include decisions that affect a person's legal rights, such as those that result in the cancellation of a contract, or entitlement to or denial of a benefit granted by law. "Similarly significant effects" includes decisions that do not necessarily alter a legal right, but have a similarly substantial impact on individuals, including in their circumstances and life opportunities. Commonly cited examples include: automatic refusal of an online credit application; decisions made by online job recruitment platforms; and decisions that affect other financial, credit, employment, health, or education opportunities.[7]

While the GDPR directly limits such processing (by prohibiting most "solely" automated decisionmaking that leads to legal or similarly significant effects), the statutory text of the CPRA likely does not offer such tools. Nonetheless, within the parameters of the law, California regulations can still create meaningful, workable safeguards for individuals. For example, regulations can clarify that:

- Automated decisionmaking (ADM) that leads to legal or similarly significant effects on consumers, can be subject to data protection impact assessments[8] to identify benefits and mitigate risks to consumers, per 1798.185(a)(15)(B)); and
- Businesses engaged in automated decisionmaking that leads to legal or similarly significant effects should have systems in place for identifying and addressing bias and discrimination, even in cases where such analysis may conflict with other rights. For example, the consumer right to "Limit the Use of Sensitive PI" applies generally to the use of any sensitive personal information under the CPRA, whether or not it involves automated decisions. Businesses seeking to address bias in automated decisionmaking may make inferences about race, ethnicity, or other sensitive information. In such cases,

---

[5] Art. 22 GDPR.
[6] European Data Protection Board, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (Oct 2017), https://ec.europa.eu/newsroom/article29/items/612053.
[7] In some cases, the Article 29 Working Party has noted that online advertising may be considered to have similarly significant effects under the GDPR, for example if it is particularly intrusive, targets vulnerable populations or uses knowledge of the vulnerabilities of individuals. This could include, for example, targeting "someone known or likely to be in financial difficulties . . . with adverts for high interest loans."
[8] This requirement exists in the GDPR. Art. 22 GDPR. Models for risk assessments include, for example: UK Information Commissioner's Office, Sample DPIA Template (Feb. 2018), https://ico.org.uk/media/2258461/dpia-template-v04-post-comms-review-20180308.pdf; FPF, Mobility Data Sharing Assessment (Aug. 2021), https://fpf.org/blog/fpf-and-mobility-data-collaborative-release-resources-to-help-organizations-assess-the-privacy-risks-of-sharing-of-mobility-data/.

regulations can encourage the development of accountability by clarifying that such uses are necessary to perform the service when high-risk decisions are involved.

2. **Regulations should provide that information about "automated decisionmaking" follow NIST interpretability guidelines, and be meaningful and reasonably understandable to the average consumer.**

1798.185(a)(16) also requires the Agency to establish rules for how businesses should comply with consumer access rights, when they involve automated decisionmaking. Access to information about the logic or functioning of an automated decision is most typically sought for decisionmaking that involves so-called "black box" algorithms in realms with high impact on consumers, such as in loan approval, hiring, or insurance. In developing regulations on this topic, California should follow NIST interpretability guidelines,[9] and require that responses to access requests be meaningful and understandable to average consumers.

Both the challenges and the need for providing meaningful information about AI-driven decisionmaking are not new. The Equal Credit Opportunity Act (ECOA)[10] and the Fair Credit Reporting Act (FCRA)[11] mandate customer-level explanations known as "adverse action notices" for automated decisions in the consumer finance space. Similarly, Article 22 of the GDPR requires businesses to "provide meaningful information" about the logic involved in automated decisionmaking about individuals that leads to legal or similarly significant effects.

In practice, however, it can be a challenge to provide truly meaningful, explainable, or interpretable AI for average consumers. Instead, what most consumers want to understand are the factors that led to a high-impact decision, and the main reasons for it. For example, in the case of an algorithmic decision tree for approval or denial of a loan: it is not enough to provide only "input data" (factors such as credit score and income) and "output" (in this case, approval or denial). In order for that information to be meaningful, a business would likely also need to share information about the relative salience (weight) of each factor. More complicated AI systems, such as neural networks, present an even greater challenge in situations where they are used to impact significant consumer decisions.

We recommend that California follow best practices and guidance from NIST's "Four Principles of Explainable Artificial Intelligence" (2020),[12] which articulates principles for explainable AI systems: "that the system produce an explanation, that the explanation be meaningful to humans, that the

[9] P. Jonathon Phillips et. al, *Four Principles of Explainable Artificial Intelligence*, U.S. Department of Commerce, National Institute of Standards and Technology (August 2020), https://www.nist.gov/system/files/documents/2020/08/17/NIST%20Explainable%20AI%20Draft%20NISTIR8312%20%281%29.pdf.
[10] Equal Credit Opportunity Act, 12 C.F.R. § 1002.9(a)(2).
[11] Fair Credit Reporting Act, 15 U.S.C. § 1681m.
[12] *See*, supra note 9.

explanation reflects the system's processes accurately, and that the system expresses its knowledge limits."

Further Resources:

- P. Jonathon Phillips et. al, *Four Principles of Explainable Artificial Intelligence*, U.S. Department of Commerce, National Institute of Standards and Technology (August 2020), https://www.nist.gov/system/files/documents/2020/08/17/NIST%20Explainable%20AI%20Draft%20NISTIR8312%20%281%29.pdf.
- European Data Protection Board, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (Oct. 3, 2017), https://ec.europa.eu/newsroom/article29/items/612053.
- *Explainable AI*, IBM, https://www.ibm.com/watson/explainable-ai.
- Aaina Agarwal, Patrick Hall, Sara Jordan, and Brenda Leong, *Five Things Lawyers Need to Know About AI* (October 2021), https://fpf.org/blog/five-things-lawyers-need-to-know-about-ai/.
- FPF, *The Privacy Expert's Guide to Artificial Intelligence and Machine Learning* (October 2018), https://fpf.org/wp-content/uploads/2018/10/FPF_Artificial-Intelligence_Digital.pdf.

Part B. Sensitive Personal Information - 1798.185(a)(19)(C)

1798.185(a)(19)(C) requires the Agency to issue regulations to "govern the use or disclosure of a consumer's sensitive personal information, notwithstanding the consumer's direction to limit the use or disclosure of [such information]." The CPRA enables consumers to direct businesses to limit the use of sensitive personal information to that use which is "necessary to perform the services or provide the goods reasonably expected by an average consumer who requests such goods or services."

3. **Regulations should help clarify a range of potential use cases for health and wellness data, by providing a principled, exemplar list of categories that are in or out of scope. In many cases, such distinctions will be based on context and reasonable use.**

The CPRA's inclusion of heightened protections for "sensitive personal information" aligns with trends in the European Union, and extends current protections in the United States for medical and health condition information.[13]

Regulations should help clarify the scope of this new category with respect to a range of potential use cases for commercial health and wellness data:

[13] The Federal Trade Commission has so far provided the strongest legal protections for commercial non-HIPAA health information in the United States. *See, e.g.*, Federal Trade Commission, *Flo Health, Inc.*, https://www.ftc.gov/enforcement/cases-proceedings/192-3133/flo-health-inc (involving improper disclosure of user data from fertility tracking apps).

- ***Diagnoses and medical conditions.*** Direct information about a consumer's medical diagnosis or health condition, such as an illness or injury, should be considered clearly within scope.
- ***Commercial data used to provide health-related products and services.*** Commercial data used to infer characteristics about a person's health should likely also be considered in-scope. For example, electronic medical record data, information traditionally subject to the FTC Health Breach Notification Rule,[14] or data collected through the use of consumer products that allow individuals to monitor vital signs (smart thermometer, glucose monitors, pulse oximeters, EKG app, etc.) would all be in scope. Similarly, health websites and apps that collect information with the intent to contact individuals about medications, or that request user reported health information (e.g. blood sugar, eating habits, or sleep patterns) and subsequently provide advice on health conditions or possible diagnoses would be considered sensitive.
- ***Fitness and wellness data.*** Regulations should clarify whether commercial wellness data, unrelated to a particular health condition or diagnosis (and not used for that purpose), is or is not in scope. For example, many health and fitness apps track information such as steps, workouts, meditation sessions, diet, or lifestyle information. Absent analysis of this data and generation of inferences regarding health conditions or diagnoses, this information is most appropriately categorized as non-sensitive.
- ***Inferences, educated guesses, and proxies.*** Finally, regulations should clarify that deliberate, sensitive inferences based on information that would otherwise be out of scope should be included as "sensitive personal information," based on its use for that purpose, regardless of accuracy. For example, social media, web search, browsing, or music/video streaming data, should likely all be considered out of scope generally. However, it should be clear that a business would be brought back in scope for limiting the use of "sensitive personal information" if it were to use such data to generate a health-related inference, such as to provide a targeting category of "likely to have X condition." In some cases, sensitivity will depend on context. For example, body characteristics such as height and weight may or may not "concern health," depending on use (e.g., to generate a BMI score, or, for example, to adjust a vehicle's safety settings).

Broadly speaking, "health" includes a wide range of potential information, and clarification here will be valuable. The CPRA already makes a useful distinction with respect to sensitive personal information based on context and use. Specifically, sensitive PI that is "collected or processed without the purpose of inferring characteristics about a consumer" is exempted from the opt-out requirement in Section 1798.121. We recommend that this distinction be applied broadly to the extent that sensitive categories of PI are treated differently in other areas of the law, such as the right to access (Section 1798.110), and restrictions on incompatible secondary use (Section 1798.100). In addition to the businesses' "purpose," the regulations should also consider reasonable context and risks to consumers.

[14] FTC Health Breach Notification Rule, 16 C.F.R. § 318.3 (2021).

4. **Regulations should ensure opportunities for socially beneficial commercial research using sensitive personal information.**

Finally, regulations should clearly encourage socially beneficial commercial research, including where it must be balanced against the consumer's right to access and delete information. For example, it may be beneficial for large platforms to conduct research on the effect of their services on consumers' mental health or time spent using online services. Similarly, businesses that provide direct-to-consumer health and wellness services may be continuously generating new health inferences, within ranges of potential accuracy. In some cases, it could be concerning, or even unethical, to inform individuals about low-confidence or ongoing health inferences, even while the research itself proves useful and could lead to new health breakthroughs in the future.

Notwithstanding the above, exempting data from access and deletion requirements for purposes of research should not allow for businesses to retain sensitive information for non-research purposes, for example if they are selling or disclosing data to third parties for marketing or other non-research purposes.

Part C. Opt-Out Signals - 1798.185(a)(19)

1798.185(a)(19) requires the Agency to issue regulations to define the requirements and technical specifications for opt-out preference signals sent by a platform, technology, or other mechanism. We support this and recommend that it serve as an opportunity to establish strong standards that will clarify and streamline options for consumers, and shape the adoption of similar tools in other jurisdictions. California should (1) clarify the role of opt-out signals in the context of today's labyrinth of existing permission frameworks; and (2) establish an open process for approval of new global opt-out signals that meet the specifications of the Agency over time.

5. **Regulations should clarify the role of global opt-out signals (i.e., not just one signal) in the context of today's labyrinth of existing permission frameworks, including in authenticated and non-authenticated platforms.**

In a fragmented data ecosystem, the adoption of universally accepted signals, including through user agents such as browsers or plug-ins, has become a practical necessity for individuals to control data collection. Without such signals, or other limits on data collection, opt-outs create an unworkable burden on individuals to identify, and individually contact, hundreds of commercial entities that might or might not process their data. It has been well documented that this is

unnavigable for average, and even very sophisticated, consumers.[15] Against this backdrop, the adoption of global opt-out signals in California, whether mandated[16] or voluntary and incentivized,[17] is a significant step forward.

However, the same fragmentation that compels adoption of global opt-out signals, has led to a mass of confusion for individuals and businesses who attempt to navigate sometimes conflicting opt-outs, settings, and signals. Today, consumers have a complicated web of options to express preferences and exercise some form of control over the sale, sharing, or use of personal information in mobile, web, and offline environments. The current volume of choices is largely unnavigable, yet relied upon by many businesses across sectors for legal, policy, and technical reasons. The confusion reflects decades of platform, business, and self-regulatory efforts to address privacy concerns without the underpinning of a single, comprehensive privacy law.

As an illustration, the options below exist today to control: browser and device-specific data; data within authenticated platforms; and offline or "physical world" information. Each of these cases raises unique issues for the adoption of global opt-out signals, which California should address through rulemaking to simplify and streamline existing systems for consumers and businesses.

***Browser and device-specific data*:**

Existing browser and device-specific controls include: privacy settings (to block cookies, block third-party cookies, "prevent cross-site tracking"); browser plug-ins; global signals such as Do Not Track or the Global Privacy Control; and self-regulatory mechanisms such as NAI Consumer Opt Out and DAA YourAdChoices. Similarly, control over mobile apps can be exercised through device settings (iOS and Android), including "Limit Ad Tracking," and app-specific permissions.

- ***Issue***: A business with browser or device-specific data (such as a cookie ID, or IDFA) may or may not be readily able to link that data to data from the same consumer on a different browser or device, or to traditional personal information processed separately.
- ***Recommendation***: Regulations should clarify that a global opt-out request associated with less readily available data (such as a cookie ID, browser information, or IDFA) should apply to the sale of all data with which the opt-out signal can be reasonably linked.

***Authenticated platforms*:**

Large and small platforms and businesses that have direct relationships with consumers increasingly offer their own "privacy dashboards" with settings for various uses of data occurring

---

[15] Maureen Mahoney, *California Consumer Privacy Act: Are Consumers' Digital Rights Protected?*, Consumer Reports & Digital Lab (Oct. 1, 2020), https://advocacy.consumerreports.org/wp-content/uploads/2020/09/CR_CCPA-Are-Consumers-Digital-Rights-Protected_092020_vf.pdf; Aleecia M. McDonald and Lorrie Faith Cranor, *The Cost of Reading Privacy Policies,* I/S: A Journal of Law and Policy for the Information Society, Vol. 4, No. 3 (2008), *543-568*, https://kb.osu.edu/bitstream/handle/1811/72839/ISJLP_V4N3_543.pdf.

[16] Cal. Code Regs. tit. 11 § 999.315 (Requests to Opt Out).

[17] CPRA Section 1798.135.

on and off-platform. These include settings for: social media; retail; financial institutions; educational services; and consumer technology (e.g., Smart TVs, vehicles, or children's toys).

- ***Issue**:* Any business with a direct consumer relationship must navigate possible conflicts between the known privacy settings of their users, and global opt-out signals received from devices that may or may not belong to their users.
- ***Recommendation***: Regulations should clarify what steps a platform should take when it receives a device or browser-specific global opt-out signal from (1) a consumer who has an account with the business and is authenticated (logged in); (2) a consumer who has an account with the business, but is not authenticated (not logged in); and (3) a consumer who does not have an account with the business.

***"Offline" or "physical world" information*:**

As the average consumer's number of devices grows, there is a growing industry for analyzing and using passive information sent by networked devices, or inferred from external behavior and appearance. Opt-outs sometimes include modifying device settings at the point of collection (such as MAC address randomization[18]), but more often are self-regulatory and limited in scope. Many opt-out mechanisms do not currently exist, or have yet to be developed (e.g., for video analytics, facial recognition, and augmented reality).

- ***Issue***: Global opt-out signals for offline data collection are largely limited in scope or do not yet exist.
- ***Recommendation***: Regulations should anticipate the development of such signals, and provide guidance for how to shape them. For example, California could encourage the widespread adoption of an SSID-based signal such as "_nomap" for the broader location industry that relies on network information within the control of individuals.[19] California could also encourage the development of user agents to control offline and Internet of Things (IoT) data.[20]

California should establish practical guidelines for businesses to navigate complex permissions systems, in a way that will simplify and streamline the current confusion for consumers.

6. **California should establish an open process for authoritative approval of new global opt-out signals that meet the technical specifications of the Agency over time.**

In a fragmented world of web, mobile, screenless IoT, and emerging technologies, there can rarely or never be "one opt-out signal to rule them all," at least without a corresponding trade-off in anonymity and privacy. New tools will continue to be developed. Each of them, like the controls

[18] *See, e.g.*, Apple Support, "Use private Wi-Fi addresses on iPhone, iPad, iPod touch, and Apple Watch," https://support.apple.com/en-us/HT211227 (last visited Nov. 8, 2021).
[19] *See, e.g.*, Google Maps Help, "Control access point inclusion in Google's Location services" (last visited Nov. 8, 2021), https://support.google.com/maps/answer/1725632?hl=en.
[20] *See, e.g.*, The Personalized Privacy Assistant Project, https://privacyassistant.org/.

and signals above, will necessarily function as a *partial* opt-out: applying to a certain kind of data, within a certain realm of processing.

As technology and business practices continue to evolve, under heightened pressure from platform rules and privacy regulation, it is likely that many more opt-out and consent tools will emerge. California should establish an open, authoritative process for approval of global opt-out signals that will be deemed to adhere to California law. We recommend that the Agency do so in consultation with technical, legal, and policy experts, as well as leaders in other jurisdictions that are developing similar tools, such as Colorado. In addition to granting businesses the benefit of clarity, consumers deserve to know which tools they choose will have legal effect.

In addition to a principles-based approach to establishing criteria for global opt-out signals (we agree, for example, that signals and opt-out tools should be easy-to-understand, and not contain defaults that misalign with the law[21]), the process should create procedural opportunity, for example by allowing civil society organizations or members of the public to propose new opt-out signals (their own or otherwise), allow stakeholders to weigh in, and involve a deliberative process that evaluates the many technical and policy factors in alignment with the Agency's criteria, such as: scale of adoption; alignment with criteria in other jurisdictions; and the extent to which the signal is possible to localize solely to consumers in California.

Part D. De-Identified and Pseudonymous Data - 1798.185(a)(2)

1798.185(a)(2) requires the Agency to issue regulations to update, as needed, the definitions of "deidentified" and "unique identifier" to address changes in technology, data collection, obstacles to implementation, and privacy concerns.

The CPRA defines "deidentified" as:

> "information that cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer, provided that the business . . . (A) takes reasonable measures to ensure that the information cannot be associated . . . (B) publicly commits to maintain and use the information in deidentified form . . . and (C) contractually obligates any recipients of the information to comply with all provisions of this subdivision." 1798.140(m).

While this language aligns with the longstanding approach of the Federal Trade Commission,[22] there remains little guidance or enforcement activity to help organizations understand how

[21] CPRA 1798.185(a)(19).
[22] U.S. Federal Trade Commission, *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers* (Mar. 2012).

"deidentified" will be interpreted by regulatory authorities in California. There are several key issues related to the interpretation and implementation of de-identification tools within the U.S., including California, that would benefit from greater certainty.

7. **Regulations should seek further input from de-identification experts and researchers to clarify key implementation issues for "deidentified data," including the role of technical, legal, and administrative controls, and Privacy Enhancing Technologies.**

We recommend that the Agency convene further specialized input, including through meetings and workshops, from leading de-identification experts, as well as researchers with experience using de-identified data safely. Public input, regulations, sector-specific guidance, and enforcement actions can serve to clarify key global issues related to de-identification, specifically:

- what constitutes "reasonable" measures to ensure that the information cannot be associated with a consumer or household;
- the status of certain types of protected, pseudonymized information, in which personal information is no longer attributable to a specific consumer without the use of additional information, provided that the additional information is kept separately and is subject to technical and organizational measures[23]; and
- technical measures needed to make de-identified data legally public under the CPRA.[24]

Evaluating privacy risk in de-identified data requires specialized, case-by-case assessments that consider a range of technical and contextual factors, which are often sector-specific. However, businesses often lack the internal expertise and capacity to deploy PETs in effective ways.[25] Although there are a growing number of vendors and practitioners offering such services, the availability of qualified PETs experts is extremely limited nationwide. By providing guidance on these topics, the Agency has an important opportunity to incentivize businesses' use of PETs to support the utility of data while mitigating risks to consumers.

Further Resources:

- NIST, *NISTIR 8053: De-Identification of Personal Data* (Oct. 2015), https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.

---

[23] See, e.g., Patrick Breyer v Bundesrepublik Deutschland, Judgment of the Court (Second Chamber) of 19 October 2016, https://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=EN.

[24] CPRA requires businesses processing deidentified data to "contractually [obligate] any recipients of the information to comply with all provisions of this subdivision." 1798.140(m). Greater clarity would be helpful with respect to how this provision applies to de-identified information released to the general public.

[25] For example, differentially private methods can be effective in providing mathematically sound guarantees of privacy, reflected by an epsilon value that indicates re-identification risk, sometimes called a "privacy budget." However, limited guidance exists to determine what these values should be for different contexts or types of data. See, e.g., Alexandra Wood, et al, *Differential Privacy: A Primer for a Non-Technical Audience*, 21 Vanderbilt J. Ent. & Tech. L. 209, 260 (2018), https://dash.harvard.edu/bitstream/handle/1/38323292/4_Wood_Final.pdf?sequence=1.

- Miranda Mourby et. al, *Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK*, Computer Law & Security Review, Volume 34, Issue 2 (Apr. 2018), 222-233, https://www.sciencedirect.com/science/article/pii/S0267364918300153.
- Khaled El Emam, Eloise Gratton, Jules Polonetsky, and Luk Arbuckle, *The Seven States of Data: When is Pseudonymous Data Not Personal Information?*, https://fpf.org/wp-content/uploads/2016/11/El-Emam_States-of-Data-Main-Article-short-v6.pdf.
- FPF, *A Visual Guide to Practical Data De-Identification* (June 2017), https://fpf.org/wp-content/uploads/2017/06/FPF_Visual-Guide-to-Practical-Data-DeID.pdf.
- Jules Polonetsky, Omer Tene, and Kelsey Finch, *Shades of Gray: Seeing the Full Spectrum of Practical Data De-identification*, 56 Santa Clara L. Rev. 593 (2016), https://digitalcommons.law.scu.edu/lawreview/vol56/iss3/3/.
- Claire McKay Bowen, *Personal Privacy and the Public Good: Balancing Data Privacy and Data Utility*, Urban Institute (August 2021), https://www.urban.org/sites/default/files/publication/104694/privacy-and-the-public-good_0_0.pdf.

Thank you for this opportunity to provide input on initial rulemaking under the California Privacy Rights Act. We welcome any further opportunities to provide resources or information to assist in this important effort.

Sincerely,

Stacey Gray, *Senior Counsel*

Jules Polonetsky, *CEO*

Future of Privacy Forum
1400 Eye St. NW Ste. 450
Washington, DC, 20005
info@fpf.org