PRINCETON UNIVERSITY

CENTER FOR INFORMATION TECHNOLOGY POLICY
PRINCETON UNIVERSITY

February 25, 2020

Lisa B. Kim, Privacy Regulations Coordinator
California Office of the Attorney General
300 South Spring Street, First Floor
Los Angeles, CA 90013
Email: PrivacyRegulations@doj.ca.gov

Comments on Revised Proposed Regulations
Implementing the California Consumer Privacy Act

Thank you for the opportunity to provide comments to the California Department of Justice on the February 10, 2020 revised proposed regulations implementing the California Consumer Privacy Act.

We are academic researchers associated with the Center for Information Technology Policy (CITP) at Princeton University, with expertise in computer science, law, and policy.[1] We write to offer three specific recommendations that advance the Department's goal of protecting consumer privacy. We look forward to further opportunities to engage with the Department to provide additional analysis as the CCPA regulations evolve.

1. **Consent notices should avoid using dark patterns that burden consumer decision making.**

Dark patterns are user interface design choices that benefit an online service by coercing, steering, or deceiving users into making unintended and potentially harmful decisions. Mathur et al, Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites, Proc. ACM Hum. Comput. Interact. 3, CSCW, Article 81 (Nov. 2019) (attached). We have studied these user interface designs extensively. Recently, we published a study based on a crawl of over 11,000 shopping websites using automated techniques that detected a variety of dark patterns on over 10% of those sites that could

[1] In keeping with Princeton's tradition of service, CITP's Technology Policy Clinic provides nonpartisan research, analysis, and commentary to policy makers, industry participants, journalists, and the public. These comments are a product of that Clinic and reflect the independent views of the undersigned scholars.

mislead or confuse consumers. *Id.* Various academic studies have also examined the use of dark patterns around obtaining consumer consent to information collection. *See* Nouwens et al, Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence, CHI '20 CHI Conference on Human Factors in Computing Systems; Jamie Luguri and Lior Strahilevitz, Shining a Light on Dark Patterns, U of Chicago, Public Law Working Paper No. 719. A recent academic study reported on the use of dark patterns in obfuscating the consent notices required by the European Union's General Data Protection Regulation (GDPR). Utz et al, (Un)informed Consent: Studying GDPR Consent Notices in the Field, 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19); *see also* Norwegian Consumer Council, Deceived by Design (2018).[2]

The Department's design of a standardized opt-out button in §999.306(f) helps bring consistency across different providers and improves the ability of consumers to make informed choices. But the proposed design has a flaw that risks impairing a consumer's decision making because the button presents consumers with a pre-selected double negative choice by using a red cross next to the phrase "do not sell." As a result, consumers might be confused about whether or not the site has the ability to sell their information. We suggest that the Department adopt the design recommended in the study by Cranor et al., Design and Evaluation of a Usable Icon and Tagline to Signal an Opt-Out of the Sale of Personal Information as Required by CCPA (Feb. 4, 2020), which includes a check and a cross in the design and presents the choices in a neutral blue color. (*Id.* at p.32, shown below.)

The best icon to pair with current CCPA taglines to convey a "do not sell" opt-out is a *toggle* icon. This combination effectively communicates the presence of a choice, particularly one related to the sale of personal information.

Do Not Sell My Personal Information

More generally, the Department could assess whether providers make it equally easy for users to select among the choice to opt in or opt out of information sharing. For example, Facebook's GDPR consent flow opt in takes 3 clicks, while the opt out takes 11 clicks. *See* Deceived by Design. This suggests that consumers may not be presented with a fair choice.

[2] https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf

A key finding from the research literature is that service providers use a variety of design elements, including color, placement, size and language to obscure choices that consumers are likely to select if fairly presented. As a result, we support the new language in §999.315(c) that responds to such concerns by prohibiting user interface designs that have "the purpose or substantial effect of subverting or impairing a consumer's decision to opt-out."

The Department might consider developing a process to monitor how providers are presenting information choices after the CCPA regulations come into effect and provide additional guidance, as necessary, to prevent tactics that subvert or impair a consumer's decision making process. The Department could also provide more explicit guidance that explains how it will not simply evaluate business practices in a vacuum, but will examine how certain choices that enhance consumer privacy are presented relative to other options that may benefit the business.

2. The Department should clarify how the definitions of "personal information" and "sell" apply to common practices.

The Department's decision to provide additional guidance about how the CCPA applies to common practices helps clarify how the law will be interpreted. But we urge the Department to reconsider its analysis of Internet Protocol (IP) addresses and to offer guidance on cookies and similar tracking technologies.

a. Internet Protocol addresses are "personal information."

In the revised proposed regulations, the Department offers the following guidance on IP addresses: "For example, if a business collects the IP addresses of visitors to its website but does not link the IP address to any particular consumer or household, and could not reasonably link the IP address with a particular consumer or household, then the IP address would not be 'personal information.'" This guidance is problematic for several reasons.

First, IP addresses are used for identification. The purpose of an IP address is to route data to a particular user device or household. IP addresses can be—and often are—used as identifiers for linking individual-level or household-level information over time and across online services. Indeed, with the latest version of IP (IPv6) there may be additional information embedded in the address such as a device (MAC) address. Thus, IP addresses enable user or household tracking and singling out a user or device for contact, and may inherently contain some identifiable information.

Furthermore, associating an IP address with other forms of "personal information" is often technically trivial. Information that matches an IP address with an individual or a precise location is often publicly available on the internet, and commercial services offer precise IP address geolocation. Moreover, there are a number of businesses that possess a reliable mapping between individual user identities and IP addresses, including services that users log into, many third-party tracking and analytics services, and internet service providers. And even in a circumstance where an association between a user or household and an IP address is not already readily available, it is technically trivial to create that association by just sending an email to the user that includes invisible tracking content or induces the user to click a link. *See* Steven Englehardt, Jeffrey Han, and Arvind Narayanan, I never signed up for this! Privacy implications of email tracking, Proceedings on Privacy Enhancing Technologies; 2018 (1):109–126.

Second, the CCPA's statutory language recognizes that IP addresses are an example of "personal information." § 1798.140(o)(1) begins with setting forth the criteria for what constitutes personal information. The next subsection (o)(1)(A) identifies specific examples of identifiers that unambiguously constitute personal information, including "real name," "social security number," and "internet protocol address." That definition concludes with a catchall to capture "other similar identifiers" that satisfy the same criteria. Thus, there is no reason for treating an IP address any differently from identifiers such as a person's name or social security number.

Third, other regulatory agencies have concluded that IP addresses are indeed personal information. For example, the Federal Communications Commission concluded in a 2016 rulemaking that IP addresses were "personally identifiable information." The FCC explained:

> We disagree with commenters that argue that we should not consider MAC addresses, IP addresses, or device identifiers to be [personally identifiable information (PII)]. First, as discussed above, a customer's IP address and MAC address each identify a discrete customer and/or customer device by routing communications to a specific endpoint linked to the customer. Information does not need to reveal an individual's name to be linked or reasonably linkable to that person. A unique number designating a discrete individual—such as a Social Security number or persistent identifier—is at least as specific as a name. Second, MAC addresses, IP addresses, and other examples of PII do not need to be able

> to identify an individual in a vacuum to be linked or reasonably linkable. [Broadband internet access service (BIAS)] providers can combine this information with other information to identify an individual (e.g., the BIAS provider's records of which IP addresses were assigned to which customers, or traffic statistics linking MAC addresses with other data). As the Supreme Court has observed, "[w]hat may seem trivial to the uninformed, may appear of great moment to one who has a broad view of the scene and may put the questioned item of information in its proper context."[3]

The FCC offered this guidance when elaborating on a "reasonably linkable" standard, nearly identical to the standard in the CCPA. We see no reason for the Department to reach a different technical conclusion about networking technology than that reached by the federal telecommunications regulatory agency.

Regulators in the European Union have similarly concluded that IP addresses should be treated as personal information because they are reasonably linkable to individuals or households. In *Breyer v Bundesrepublik Deutschland* (2016), the Court of Justice of the European Union ("CJEU") explained that "a dynamic IP address registered by an online media services provider . . . constitutes personal data within the meaning of that provision . . . where the latter **has the legal means** which enable it to identify the data subject with additional data which the internet service provider has about that person." (Emphasis added.)

Fourth, the proposed guidance about the circumstances when IP address data can be linked to particular consumers or households could be read to only consider the data collected and maintained by a business. But the text of the CCPA does not contain either of these limitations; it provides an *objective* linkability standard, alternately phrased as "reasonably capable of being associated with" and "reasonably linked." We urge the Department to redraft the guidance to clarify that the linkability analysis is not simply confined to a business's own practices and data holdings and that information from third parties that could be obtained to identify consumers or households is relevant to the analysis.

Fifth, there is a practical concern that if the Department offers ambiguous guidance about when and how IP addresses are "personal information," that will detract from a predictable and uniform application of the law. Businesses of course

[3] https://www.fcc.gov/document/fcc-releases-rules-protect-broadband-consumer-privacy.

have significant commercial incentives to take the position that IP addresses are not subject to CCPA's privacy protections. Offering clarity on IP addresses now avoids foreseeable policy disputes in future about the circumstances when IP addresses are treated as personal information.

b. Third-party consumer tracking using cookies and similar technologies constitutes a "sale" of "personal information."

We recommend that the Department offer guidance on how CCPA applies to third-party consumer tracking using cookies and similar technologies (e.g., "supercookies" and "fingerprinting"), a pervasive business practice on the web and in mobile applications. *See e.g.*, Jonathan R. Mayer and John C. Mitchell, Third-Party Web Tracking: Policy and Technology;[4] Steven Englehardt and Arvind Narayanan, Online Tracking: A 1-million-site Measurement and Analysis, ACM CCS 2016.[5]

Like with the analysis of IP addresses, a tracking technology like cookies involves "personal information" because the data "is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household."

Tracking technologies that operate across online services also constitute a "sale" of personal information because such technologies are placed on sites in exchange for value. For example, when a third-party service collects consumer tracking information, it typically does so via content embedded in another business's site and offers an incentive for that business to host the tracking content. In other words, third-party tracking inherently involves personal information "[made] available . . . by [a] business to another business or a third party for monetary or other valuable consideration."

We recommend the Department offer guidance on the use of such technologies in the next round of proposed rulemaking.

3. If a consumer maintains a password-protected account with a business, logging into the account should be necessary and presumptively sufficient for verifying a consumer request.

[4] Available at https://jonathanmayer.org/publications/trackingsurvey12.pdf
[5] Available at https://www.cs.princeton.edu/~arvindn/publications/OpenWPM_1_million_site_tracking_measurement.pdf

Recent research has highlighted security risks associated with GDPR data request processes, because businesses are implementing new processes for customer authentication rather than using existing processes that have been vetted extensively. *See* Martino et al, Personal Information Leakage by Abusing the GDPR Right of Access.[6] These studies raise a concern for how businesses will respond to the access rights under the CCPA. The problem is that the new authentication methods add a whole class of newly recognized security risks, where attackers can circumvent established authentication protections by using weaker GDPR request processes.

We recommend that the Department specify that, if a consumer maintains a password-protected account with a business, logging into the account is a necessary step for verifying a consumer request. This is a technically simple precaution for businesses to implement, including in coordination with a third-party identity verification service. This step is also trivial for consumers—just one simple login to an existing account. Adding this step avoids creating new and often insecure authentication methods. It also reduces the risk of data leaks in which businesses respond to requests with extraneous data that does not pertain to the consumer making the request. *See* James Pavur and Casey Knerr. GDPArrrrr: Using Privacy Laws to Steal Identities. Black Hat USA 2019.

We also recommend that the Department specify that logging into a password-protected account is presumptively sufficient for verifying a consumer request. In many contexts, a user already has full access to and control over their data after logging into an account and there is no need to add unnecessary friction for consumers seeking to exercise their CCPA rights.

We acknowledge that there are circumstances where additional authentication beyond a login is appropriate, especially when the CCPA gives the end user access to more data than they would have in the ordinary course. We recommend setting a presumption that businesses can overcome in appropriate contexts (e.g., considering the factors that the Department proposes to articulate in § 999.323).

* * *

We appreciate the opportunity to participate in the rulemaking process and remain available to answer any questions the staff may have.

Respectfully submitted,

[6] https://www.usenix.org/conference/soups2019/presentation/dimartino

Marshini Chetty
Assistant Professor, Department of Computer Science, University of Chicago

Shaanan Cohney
Postdoctoral Research Associate, Center for Information Technology Policy, Princeton University

Mihir Kshirsagar*
Technology Policy Clinic Lead, Center for Information Technology Policy, Princeton University

Arunesh Mathur
Graduate Student, Department of Computer Science, Princeton University

Jonathan Mayer*
Assistant Professor of Computer Science and Public Affairs, Princeton University

Arvind Narayanan
Associate Professor of Computer Science, Princeton University

Ross Teixeira
Graduate Student, Department of Computer Science, Princeton University

Ari Ezra Waldman
Microsoft Visiting Professor of Information Technology Policy, Princeton University

* denotes principal comment authors.

Contact:

Website: https://citp.princeton.edu
Phone: 609-258-5306
Email: mihir@princeton.edu