CALIFORNIA PRIVACY PROTECTION AGENCY
400 R ST. SUITE 350
SACRAMENTO, CA 95811
cppa.ca.gov

Request for Input from the CalPrivacy Audits Division

Emerging Technology, Inference, and Identifiability: CalPrivacy Audits Technologist Request for Input

As the California Privacy Protection Agency (CalPrivacy) continues to operationalize the California Consumer Privacy Act (CCPA) within a rapidly changing technical landscape, the newly formed Audits Division believes that technical and interdisciplinary research can help to inform effective oversight. Accordingly, the Audits Division requests input to support its development of audit framework and compliance inquiry functionalities.

Emerging technology systems that rely on automated data processing raise questions about how persons subject to CCPA are practically applying core statutory concepts—particularly protections and requirements related to personal information, sensitive personal information (SPI), and deidentification.

This post is intended to be the first in a series of requests for technical input.

1. Inference

As many emerging technologies are capable of inferring sensitive attributes from data that is not itself categorized as sensitive, we seek a deeper understanding of inferences that may be probabilistic, latent, or emergent from model behavior rather than explicitly designed outputs.

Probabilistic Inference: A conclusion generated by assigning statistical likelihood values to possible outcomes based on observed input data and a trained model. Rather than producing a single deterministic answer, the system estimates the probability distribution over possible states — for example, estimating the probability that someone is pregnant based on purchase patterns or scroll behavior. The output reflects the model's confidence, not a verified fact about the individual.

Layperson translation: The system makes an educated guess. It sees patterns across thousands of people and says, "Based on what we know about people like you, there's a 73% chance you have this condition." It doesn't know — it's betting. But that bet can still be acted on as if it were true.

Latent Inference: A conclusion derived not from explicit, observable input variables but from latent (hidden) representations learned during model training. Neural networks and embedding models compress raw data into intermediate feature spaces; inferences drawn from these spaces may have no direct correspondence to any single input variable the subject provided. The model effectively "discovers" attributes that were never explicitly measured.

Layperson translation: The system figures out something about you that you never told it — and possibly don't even know yourself. It doesn't ask your race, political views, or mental health status, but it learns to decode signals buried in your shopping history, typing speed, or scroll patterns that it associates with a particular race, set of political views, or mental health status. Those inference come from data that seem unrelated.

Emergent Inference: An output or capability that arises from the interaction of a model's scale, architecture, and training data — not from any deliberate design decision to produce that output class. Emergent inferences are not predictable from examining the model's components individually; they are unintended capabilities that can appear at scale without a targeted training objective. They may include the ability to infer psychological states, political orientation, health conditions, or identity characteristics without any explicit training objective targeting those attributes.

Layperson translation: Nobody programs the system to do this — it just learns to do it on its own, as it gets bigger and smarter. The developers don't build a tool to detect depression or predict immigration status, but the model, having processed enough data, develops that ability as a side effect.

Key input

- What technical methods are or can be used to detect or evaluate the presence of sensitive attributes within latent representations (e.g. embeddings)?
- What evaluation techniques are or can be developed to assess a model's capacity to infer sensitive attributes, even where those attributes are not explicitly collected or labeled (i.e. methods that inspect a system's intermediate reasoning or logged outputs)?
- What meaningful thresholds are or can be used to distinguish between statistical correlation and actionable sensitive inference? What methodologies best assess whether an inference was "acted on" (i.e.counterfactual black box testing, architecture evidence, temporal tracing)?
- In what ways might businesses be integrating services from others that affect consumers' privacy, including consumers' ability to exercise their CCPA rights?

2. Reidentification

Advances in technology introduce new vectors for reidentification, including:

- model inversion and extraction techniques
- linkage across datasets and contexts
- outputs that may reproduce or encode elements of training data

Key input

- What are robust and defensible standards for deidentification, including when data is used to train or interact with machine learning models?
- How does the integration of multiple services into a single system affect reidentification risk?

- What methodologies can organizations use to measure and mitigate risks associated with:
    - membership inference
    - training data extraction
    - attribute inference through model interaction?
- How can auditors adapt existing technical evaluation methods for reidentification risk — developed for adversarial ML security research — into a standardized, audit-usable protocol?

Invitation for engagement

Ongoing dialogue between regulators, researchers, and practitioners can help the Agency to evaluate and mitigate risks to consumers' privacy in this rapidly changing technical landscape. CalPrivacy's Audits Division accordingly invites researchers and practitioners to engage with, and provide input on, the topics in this post.

Where to Submit Input

Electronic Submission: Input may be submitted electronically to AuditTeam@cppa.ca.gov. Please include "Technical Input: Emerging Tech" in the subject line.

Mail Submission: California Privacy Protection Agency
Attn: Audits Division – Request for Input
400 R St., Suite 350
Sacramento, CA 95811

Note:

This request does not implement, interpret, or make specific the law enforced or administered by the CPPA, establish substantive policy or rights, or constitute legal advice.

Materials and input submitted to CalPrivacy in response to this request may be subject to disclosure under the California Public Records Act. If you have unpublished research, draft manuscripts, or proprietary information you do not wish to be subject to such disclosure, please do not submit them.