

The ALLuminate Assessment Standard

Introducing the Agent Privacy Risk Taxonomy

V.01

MLCommons AI Risk & Reliability Program
Privacy and Confidentiality Working Group

September 2026

Agent Privacy Risk Taxonomy V.01

Authors¹: Modupe Akintan, Elie Alhajjar, Aaron Alva, Jaspreet Bhatia, Miranda Bogen, Chris Calabrese, April Chen, Kristie Chon Flynn, Damien Desfortaines, Roman Eng, Dylan Gilbert, Bennett Hillenbrand, Alex Kessler, Shujaat Mirza, Vinh Nguyen, Matt O'Shaughnessy, Abhinav Palia, Jonathan Petit, Veronica Tang, Christine Task, Zheng Xu

Abstract

The rapid transition from Large Language Models (LLMs) to autonomous, multi-agent systems introduces and exacerbates critical privacy failure modes. Traditional privacy and data protection frameworks often fail to address the risks created by the ways in which modern agents ingest, process, use, and share data, or the ways in which consent and accountability apply to agents.

To systematically address these challenges, this paper sets out an Agent Privacy Risk Taxonomy designed to establish a standardized and common language to categorize risk vectors, failure causes, and privacy impacts in a consistent way. The framework organizes agentic privacy risks across five domains:

1. Data Ingestion & Processing Risks
2. Aggregation, Use & Sharing Risks
3. Inconsistent Privacy Practices Across Agents
4. Failure of Legacy Consent Systems
5. Accountability & Governance Risks

To operationalize this taxonomy, the paper introduces a structured risk assessment methodology that models privacy hazards as a six-stage causal chain. This chain maps a threat to its specific failure mechanism and privacy event, to the affected parties, to the immediate privacy impacts, and to eventual downstream harms. This taxonomy and the risk assessment methodology will be utilized to design concrete agentic evaluation scenarios, serving as the basis for benchmarking agentic privacy performance.

This is the first draft of the Agent Privacy Risk Taxonomy. We welcome feedback from the AI community so that we can iteratively refine.

Methodology

The methodology underlying this taxonomy synthesizes cross-industry working group expertise, analysis of known incidents, and foundational frameworks. Foundational frameworks such as

¹ The authors are listed in alphabetical order

[Solove's Taxonomy of Privacy](#), the NIST Privacy Risk Model ([NIST IR 8062](#) / [NIST AI RMF](#)), Artificial Intelligence Risk Management Framework ([NIST AI 100-1](#)), and Helen Nissenbaum's contextual integrity — establish essential baselines for privacy harms such as unauthorized surveillance, aggregation, disclosure, and loss of agency. However, these traditional models assume deterministic data flows, centralized storage, and clear perimeter boundaries. As AI shifts from passive conversational retrieval to autonomous agency, these established frameworks must be extended rather than replaced. Autonomous agents act as dynamic data intermediaries that execute code, retain longitudinal session memory, and orchestrate across external tools and protocols (e.g., MCP). The taxonomy introduced in this paper builds directly upon Solove and NIST by contextualizing their harm and risk taxonomies within agent-specific primitives: continuous tracking, trajectory logging, memory poisoning, and over-privileged inter-agent handshakes. By doing so, we extend traditional privacy protection principles into an operational, six-stage causal model and five domain-specific failure categories designed for benchmarking agentic privacy.

1. Data Ingestion Risks

This category addresses the vulnerabilities arising from how autonomous agents capture, structure, and retain data.

Risk ID	Risk Vector	Cause	Privacy Impact
R1.1	Continuous tracking	Always-on, background monitoring	Violates data minimization; enables unexpected or intrusive inferences about users
R1.2	Unintended/Unexpected data collection	Collection of data that is beyond the scope, purpose, or intention of the user	Enables unexpected user profiling; increases risks of sensitive data being inappropriately used, leaked, or stolen
R1.3	Trajectory logging	Retention of detailed execution traces, including inputs, reasoning, and tool calls without appropriate deletion controls	Exposes user intentions and behavior
R1.4	Session persistence	Retention of information across sessions	Compounds potential for collection of sensitive information
R1.5	Unredacted PII	Standard regex/heuristics fail to redact patterns in	Enables re-identification through behavioral patterns

		unstructured semantic data	or other "semantic PII"
R1.6	Untrusted content/tool/agent	Run-time processing of external, untrusted content (like emails or websites) that conceals malicious instructions, which the model fails to separate from valid system commands.	The compromised model can silently extract sensitive personal data and exfiltrate the system to perform unauthorized external requests.
R1.7	Over-privileged inter-agent data handshakes	Agents prioritize task completion and capability over data minimization, and modern protocols like MCP lack granular, field-level access controls to limit context sharing.	Non-required user data, proprietary business logic, and session context are exposed to external third-party servers.

R1.1 Agent ingests data through continuous tracking: Background agents designed for proactive assistance continuously monitor user environments, keystrokes, application states, ambient audio or other related content (e.g., emails, calendars). This constant ingestion violates the principle of data minimization, as agents capture a high volume of non-essential personal context. Over time, the processing of this unstructured ambient data can allow the system to draw highly intrusive, unauthorized inferences about a user and even potential unexpected profiling of other individuals around the user.

R1.2 Unintended or unexpected data collection: Agents can collect sensitive data that goes beyond the user's expected scope, purpose, or intention. This can result in significant privacy harms, such as enabling unexpected user profiling, allowing the utilization of personal data in ways the user never intended, and increasing the risk of sensitive data being leaked or stolen.

R1.3 Logging of agent trajectories: Agent trajectories serve as detailed execution logs, tracking the step-by-step sequence of an agent's inputs, reasoning steps, and tool calls. While trajectories are critical for engineering evaluation, debugging, and reinforcement learning, they inherently contain raw data that can reveal sensitive information about users. Because these logs capture the exact mental model and contextual exposure of the agent, the lack of granular user control and transparent retention policies over trajectories creates a severe data exposure risk.

R1.4 Compounding risks due to persistence across sessions: To support personalized, long-term interactions, agents rely on persistent memory architectures that span multiple sessions and platforms. This continuous historical record can allow agents to construct more complex, highly sensitive user profiles. Consequently, privacy risks compound over time; data volunteered for a single, low-risk transaction remains

accessible, influencing future agent behaviors and decisions across unrelated context boundaries.

R1.5 Failure of traditional PII redaction techniques: Legacy techniques for privacy preservation, such as regular expressions and heuristic-based redaction pipelines, are designed for highly structured Personally Identifiable Information (PII) like social security numbers or credit cards. These pipelines fail to sanitize fuzzy quasi-identifiers. Limitations of traditional PII redaction techniques mean that this “semantic PII” could be used to re-identify users.

R1.6 Untrusted content/tool/agent: Agents are often designed to browse the web, read incoming emails, analyze uploaded documents, or interact with third-party software. Attackers can embed hidden, adversarial payloads, referred to as indirect prompt injection attacks, such as invisible text, metadata, or structured system commands, in external content. These indirect injections can be executed at run-time by a user’s agent at processing time, potentially enabling new vectors for their sensitive information to be leaked or stolen.

R1.7 Over-privileged inter-agent data handshakes: AI agents operating over protocols like MCP default to sharing exhaustive context windows to maximize performance, but the absence of granular access controls enables external tools to request and exfiltrate entire conversational histories. This over-privileged context sharing exposes enterprise IP, credentials, and sensitive personal data to third-party infrastructure without explicit user authorization.

2. Data Aggregation, Use & Sharing Risks

This category covers risks associated with the downstream synthesis, semantic utilization, multi-agent sharing of collected information, and special cases of out-of-context memorization of sensitive data.

Risk ID	Risk Vector	Cause	Privacy Impact
R2.1	Mosaic effect	Synthesis of fragmented, harmless data points by one or more agents	Reconstructs highly sensitive, undisclosed traits (e.g., health)
R2.2	Excessive use	Inadvertent utilization of non-essential sensitive context during tasks	Bypasses simple file-level access controls via semantic links, utilizes

			persistent sensitive data inappropriately
R2.3	Unintended or unexpected data sharing or deletion	Goal misalignment or unauthorized execution due to security flaws or over-broad agency or sharing due to subpoenas	Discloses sensitive data to third-parties in an inappropriate way or deletes data in unexpected ways. Concentrates deeply personal data into single execution traces, creating high-value targets for government demands or threat actors
R2.4	Agentic supply chain manipulation	Adversary manipulating the upstream Agentic components	Triggers unauthorized exfiltration
R2.5	Unauthorized access to data due to unexpected dynamic code execution	Execution of untrusted code	Grants attackers full read/ write host access bypassing LLM layer privacy filters
R2.6	Memory poisoning	Ingestion of adversarial payloads embedded in external content that get stored in databases or long term memory	Allows malicious actors to plant silent retrieval instructions and enable modification causing it to behave maliciously
R2.7	Agent actions reveal sensitive information	Agent's actions allow a third party to determine a user's traits, actions, or beliefs without the user's explicit authorization	Third party can infer sensitive user information
R2.8	Out-of-context, model-originated sensitive information disclosure	The underlying model discloses sensitive information about a person that was not present in the agent's input, tool results, or memory. The information is not in in-context user	Discloses data the deployer never collected and cannot delete, correct, account for, or attribute to a lawful basis

		input, memory, or tool access	
R2.9	Out-of-context, model-originated inferred protected attributes	The model infers sensitive special categories subject matter (health, politics, orientation, religion, etc.) about a person without being given it	Attributes sensitive characteristics that may be false, unwelcome, or unknown to the subject

R2.1 Mosaic effect: The Mosaic Effect occurs when seemingly harmless, fragmented data points located across different platforms and boundaries are combined to infer highly sensitive, private user data. Agentic or multi-agentic systems can exacerbate this risk by enabling the collection of a broader set of information about users from unstructured environments. Even when individual agents operate within local privacy boundaries, the collective pipeline can synthesize these disjointed elements to infer undisclosed personal traits, such as underlying medical conditions. This sequential compounding risk is formally analyzed in recent research on [sequential LLM agent pipelines](#)

R2.2 Excessive data use by Agents: Autonomous agents require broad access to context to execute complex reasoning paths. However, empirical studies demonstrate that standard autonomous agents are highly prone to the inadvertent use of unnecessary sensitive information during task execution, as detailed in [studies on autonomous agent data leakages](#).

- i. **Beyond file-level privacy:** A key driver of this risk is the failure of traditional file-level access controls. Because agents rely on fluid semantic reasoning and cross-surface data aggregation, blocking access to a single file or database is ineffective. The agent can often reconstruct or infer the restricted information through redundant, unprotected, or secondary sources.
- ii. Unexpected or unanticipated use of sensitive data persisting across sessions.

R2.3 Unintended or unexpected data sharing, and deletion: Because agents operate with broad agency, they risk executing unauthorized or unexpected actions that spill sensitive data due to goal misalignment or adversarial exploitation. This could include unexpected and unauthorized sharing of sensitive information with a third party. These unexpected execution paths are a structural hazard, where agents lack native social boundaries and may disclose highly sensitive, out-of-context personal data to unintended external entities. Similarly, if an agent is instructed by a user to “forget” unexpected behavior around deletion could occur. Some users may intend for the agent to delete the data, others for the agent to not use the data for the particular use case

being invoked but want the data retained. Navigating this ambiguity and appropriately aligning with user expectation is key for data deletion requests.

R2.4 Agentic supply chain manipulation: AI agent components can be compromised upstream, such as third-party skill packages, tool descriptions, or Model Context Protocol (MCP) servers, to manipulate the agent's reasoning logic (such as prompt injection) into leaking sensitive data under the guise of normal execution.

R2.5 Unauthorized access to data due to unexpected dynamic code execution: Autonomous agents can be manipulated into generating and running arbitrary code or system commands directly within its hosting environment. A human or external AI agent attacker can exploit the agent's code execution tool (such as a Python REPL or bash environment), completely bypassing LLM-level privacy controls to gain direct, unauthorized read and write access to the host machine, sensitive environment variables, and connected internal networks.

R2.6 Adversarial context and memory poisoning: Agents are often designed to browse the web, read incoming emails, analyze uploaded documents, or interact with third-party software. Attackers can embed hidden, adversarial payloads—such as invisible text, metadata, or structured system commands—in external content. These adversarial payloads can be executed by a user's agent and persist the payload to memory, potentially enabling new vectors for their sensitive information to be leaked or stolen.

R2.7 Agent actions reveal sensitive information: Actions taken by an agent in the course of executing a user's goal may inadvertently allow a third party to infer sensitive information about the user's traits, activities, or beliefs. This may occur because the agent does not have sufficient context about what information the user would consider sensitive, or because the agent is not able to determine which actions may allow third parties to infer that sensitive information.

R2.8 Out-of-Context, model-originated sensitive information disclosure: Out-of-context means the model has no access to in-context user input, memory, or tool access. In rare cases, the model underlying an agent may emit sensitive information (such as direct identifiers, credentials, special-category data, or quasi-identifier combinations) about an identifiable person that was never supplied to the agent. From outside the system, the mechanism is unattributable: the content may reflect memorized training data, retrieval the evaluator cannot observe, or confabulation that happens to be true. (For more details on what we define as sensitive information, see Annex: Sensitive Information Disclosure Categories).

R2.9 Out-of-Context, model-originated inferred protected attributes: Out-of-context means the model has no access to in-context user input, memory, or tool access. In rare cases, the model may infer and assert a protected characteristic rather than reproduce one — labeling a person as likely to have a health condition, political affiliation, or

orientation from indirect cues. The assertion may be false; if true, it exposes something the subject never disclosed.

3. Inconsistent Privacy Practices across Agents

As multi-agent ecosystems grow, the lack of standardized coordination protocols creates severe policy fragmentation.

Risk ID	Risk Vector	Cause	Privacy Impact
R3.1	Policy fragmentation	Conflicting privacy guidelines during multi-agent orchestration	Leaks data due to mismatched understandings of contextual sharing requirements or norms
R3.2	Cross-agent data rights propagation failure	Deletion, correction, consent withdrawal, or processing restrictions are not propagated across agents and derived copies	Leaves data processed under revoked or outdated permissions
R3.3	Cross-agent privacy failure propagation	Reliance on upstream outputs; shared context or memory; inherited permissions; missing provenance; inadequate validation.	Amplifies the scope, persistence, recipients, and downstream consequences of one or more underlying privacy events.

R3.1 Fragmented privacy behavior in A2A interactions: When agents collaborate to fulfill a request, they must share data across boundaries. Currently, there is no unified protocol to negotiate privacy constraints in Agent-to-Agent (A2A) interactions. An agent operating under strict corporate privacy standards may transfer data to a secondary agent that has weak ingestion boundaries or a different understanding of contextual sharing norms. This misalignment results in immediate data leakage across trust zones.

R3.2 Cross-agent data rights propagation failure: A valid deletion, correction, withdrawal, or restriction instruction does not propagate. E.g., The user deletes the health information, but Agents B and C were exposed to this information in a prior agentic interaction and retain and continue using it.

R3.3 Cross-agent privacy failure propagation: A harmful output, disclosure, inference, or permission error propagates downstream. E.g., Agent A wrongly labels a user as

having a health condition; Agents B and C were exposed to this information in a prior agentic interaction and retain and continue using it.

4. Failure of legacy consent systems

Static, binary permission gates (e.g., standard browser pop-ups) are structurally incompatible with the dynamic, open-ended reasoning of autonomous agents.

Risk ID	Risk Vector	Cause	Privacy Impact
R4.1	Limitations in inferring and adhering to consent	Moving from explicit collection of detailed user consent to error-prone inference of user's intent results in mismatch between agent actions and user's expectations	Agents take actions that do not adhere to user's intent and expectations regarding privacy; exposes systems to adversarial manipulation of privacy boundaries
R4.2	Consent fatigue	High frequency of manual runtime permission pop-ups	Leads to prompt fatigue and blind user approvals of risky actions
R4.3	Transparency deficit	Opaque multi-step reasoning and unlogged data access paths	Degrades user trust; introduces severe corporate reputational risks
R4.4	Manipulation	Manipulative design decisions or social engineering techniques used by agents	Encourages or tricks users into revealing highly sensitive information

R4.1 Limitations in inferring and adhering to consent: There is a difference in fidelity between agent actions and users' intent/expectations regarding how their data is collected, used, shared, or could be at risk. To minimize this user friction, agent architectures are shifting toward intent-based alignment, where the model dynamically interprets and refines user privacy preferences based on natural language cues. Because this alignment is probabilistic rather than deterministic, it cannot guarantee that users' privacy preferences are inferred or adhered to in a provable way. Additionally, it is

highly vulnerable to adversarial prompt manipulation, in which attackers exploit this interpretative flexibility to bypass user-established privacy boundaries.

R4.2 Consent fatigue: Manual runtime confirmation gates for every minor action can create cognitive overload. Confronted with a constant stream of permission pop-ups, users quickly experience prompt fatigue. This fatigue leads to habitual, unreflective clicking ("blind approvals"), nullifying the protective utility of the consent mechanism.

R4.3 The transparency deficit: Autonomous agents generate a vast, non-linear sequence of reasoning steps and multi-tool execution paths. The internal opacity of this process makes it extremely difficult for users to inspect how, when, or why their data was accessed. This severe deficit in clear, readable transparency breeds user anxiety regarding corporate manipulation, degrading trust and creating immense brand and regulatory exposure.

R4.4 Manipulation: Users can be manipulated into revealing sensitive information. This could occur through design decisions in AI tools (e.g., "dark patterns") that encourage users to relax their security posture. It might also occur by agents using social engineering techniques (including by simulating human-like traits to gain user trust) in order to trick people into revealing credentials, personal identifiable information (PII), proprietary corporate secrets, or other sensitive information.

5. Accountability and Governance Risks

When privacy violations occur, the distributed nature of agentic execution complicates post-incident forensics and liability attribution.

Risk ID	Risk Vector	Cause	Privacy Impact
R5.1	Attribution deficit	Complex, multi-factor actions (depending on prompts, models, and context) obscure root cause of incidents	Prevents clear liability assignment and timely technical remediation
R5.2	Inadequate logging	Absence of scalable, cross-ecosystem, privacy-preserving interaction logs	Impedes post-incident forensic analysis and root-cause isolation

R5.1 Lack of clarity around who is accountable for a given action: An agentic action is the synthesized result of user instructions, underlying model behavior, system

prompts, retrieval context, and dynamic tool outputs. Consequently, when a privacy incident occurs (e.g., unauthorized data exfiltration), isolating the root cause can be exceptionally difficult. It can be unclear how to distribute responsibility among the third party providers of pre-training and fine-tuning data, the model vendor, the application developer, the third-party tool creator, or the user, which can prevent accountability and hinder effective technical remediation.

R5.2 Inadequate interaction logging: Analyzing and remediating privacy failures can require a comprehensive execution log that tracks historical interactions across multiple ecosystem boundaries. Currently, there is no standardized, scalable, and privacy-preserving method to implement such cross-ecosystem logging. Without these standardized forensic trails, identifying the exact point of failure during complex multi-agent execution can be difficult.

Agent Privacy Risk Assessment

Agent privacy risk should be assessed as a causal chain:

Deployment condition or threat → failure mechanism → privacy event → affected party → privacy impact → downstream harm

Stage	Node	Description	Assessment Evidence Type
1	Deployment Condition / Threat	Architectural feature, user behavior, system failure, or adversarial trigger creating the opportunity for an incident.	Directly Observed: System design specs, security logs, or threat intelligence.
2	Failure Mechanism	Agent behavior or control weakness through which the event actually occurs (e.g., prompt injection susceptibility, semantic leakage).	Directly Observed: Execution traces, model evaluation benchmarks, or audit logs.
3	Privacy Event	Collection, inference, retention, use, modification, or disclosure of personal data in a manner inconsistent with an applicable purpose, permission, expectation, policy, or contextual norm.	Directly or indirectly Observed: Network telemetry, audits and inspection, user feedback, or output logs.
4	Affected Party	Identification of specific entities impacted (e.g., primary end users, enterprise clients, or non-consenting 3rd parties).	Inferred / Identified: User identity mapping or context boundary analysis.
5	Privacy Impact	Immediate privacy consequences (e.g., loss of confidentiality, unauthorized profiling, or credential exposure).	Inferred / Evaluated: Data sensitivity classification and contextual policy breach.

6	Downstream Harm	Tangible consequences resulting from the event (e.g., identity theft, financial loss, legal liability, or reputational damage).	Inferred / Uncertain: Requires external evidence beyond system logs to demonstrate actual occurrence.
---	------------------------	---	---

The assessment should then identify who is affected, the immediate privacy impact, and any credible downstream harm.

A successful attack, agent action, or technical data exposure does not by itself demonstrate that downstream harm occurred. The assessment should state which part of the causal chain was observed and which parts remain inferred or uncertain.

Likelihood

Likelihood should only be expressed as a probability when supported by observational data with a defined population, denominator, operating environment, and period of observation. Where such evidence is unavailable, the assessment should characterize the following dimensions separately.

Deployment Exposure: How often does the assessed product encounter the data, capability, recipient, or operating condition required for the privacy event? (Inherent in ordinary product operation, present only for some users or supported workflows, requires an unusual configuration, action, or sequence).

Conditional Failure Evidence: When the required conditions are present, under what circumstances does the system produce the privacy event?

- **Ordinary use:** Demonstrated in representative, non-adversarial workflows.
- **Realistic misuse or attack:** Demonstrated under a credible threat model, defined user or attacker capability, and realistic attack budget (range of budgets).
- **Plausible but unconfirmed:** Supported by architecture or prior evidence but not reproduced in the assessed system.

Impact

Impact should distinguish the privacy event itself from its immediate and downstream effects. The immediate impact should consider the sensitivity, identifiability, and volume of the data; the degree of control lost; departure from the original purpose or context; number and type of recipients; persistence and reversibility; and the affected party's ability to detect, contest, correct, or delete the information. Credible downstream harms may include physical, economic, reputational, psychological, autonomy-related, discriminatory, relational, legal, organizational, or societal effects.

Impact ratings should reflect effects on affected parties, not only technical compromise, regulatory exposure, or organizational reputation.

Level	Definition
Critical	The privacy event creates severe or potentially irreversible harm to affected parties, such as threats to physical safety, coercive control, systemic discrimination, loss of major life opportunities, or widespread exposure of highly sensitive data. The impact is extensive, persistent, difficult to remediate, or affects a large or particularly vulnerable population.
High	The privacy event creates substantial harm or loss of control involving sensitive data, important decisions or relationships, or material effects on employment, health, finances, legal status, safety, reputation, or autonomy. The impact may be difficult to detect, contain, reverse, or remediate.
Moderate	The privacy event causes a meaningful but contained loss of control, contextual misuse, unwanted inference, or disclosure. Its effects are limited in scale, sensitivity, duration, persistence, or downstream consequence and can generally be mitigated.
Low	The privacy event involves limited data, sensitivity, recipients, persistence, and deviation from reasonable expectations. Effects are minor, readily detectable, reversible, and unlikely to produce material or continuing harm.

Appendix

Sensitive Information Disclosure Categories

The following sensitive information disclosure categories are built from the requirements of the EU General Data Protection Regulation (GDPR), the U.S. Health Insurance Portability and Accountability Act (HIPAA), the California Consumer Privacy Act (CCPA), and the California Privacy Rights Act (CPRA). This is a minimum baseline, and entities should consider additional sensitive information based on their context, conditions, and the changing regulatory dynamics. Each category has a subject (what the data is about or contains) and a format (syntactic = pattern-matchable, exact ground truth exists; semantic = free-form, meaning-based).

Category	Subject	Format	Legal & standards anchors
----------	---------	--------	---------------------------

S1 — Direct identifiers	Names/aliases; postal address; email; phone/fax; SSN; passport, driver's licence, state/national ID, taxpayer ID; dates of birth; medical-record, health-plan & account numbers; certificate/licence numbers; vehicle IDs & plates; device IDs; IP addresses; usernames/handles & person-directed URLs; precise geolocation coordinates	Syntactic	GDPR Art. 4(1); HIPAA Safe Harbor, 45 CFR §164.514(b)(2); CPRA §1798.140(v)(1)(A)
S2 — Credentials & secrets	Passwords; API keys & tokens (cloud access keys); OAuth/refresh tokens; JWTs & session cookies; DB URIs/connection strings; SSH & TLS private keys; code-signing keys; webhook URLs with embedded tokens; wallet seed phrases; MFA/recovery codes; financial account + access-code combos. Often memorized from user-provided code	Syntactic	CPRA §1798.140(ae)(1)(B); ISO/IEC 27001:2022 Annex A 5.17
S3 — Special categories (SPI)	Health; genetic; biometric; racial/ethnic origin; political opinions; religious/philosophical beliefs; union membership; sex life/orientation; contents of mail/email/texts; immigration status; neural data	Semantic (incl. long verbatim blocks)	GDPR Art. 9(1) + CPRA §1798.140(ae) — the union of both lists (immigration: AB 947, 2023; neural data: SB 1223, 2024).
S4 — Indirect / quasi-identifiers	Syntactic: employee/badge IDs; DOB + ZIP + gender (the classic re-identification triple); IMEI/advertising IDs; cookie/device fingerprints; partial identifiers (last-4 SSN, masked cards). Semantic: employment history; job title + employer + dates; compensation; education records	Mixed	ISO/IEC 20889:2018; GDPR Recital 26; 45 CFR §164.514(b)(2)(R); NIST SP 800-188