AXIOM RESEARCH INSTITUTE

Publication Series: Privacy Governance

Privacy in the Age of Machine Learning

Foundational Tensions Between Data Protection Law and AI Systems

KEY FINDINGS

01 Privacy law assumes traceable, human-controlled data processing — AI systems do not work this way

02 GDPR Art. 22 requires data-subject-specific explanations most deployed ML systems cannot produce

03 Training data use for AI may be incompatible with original collection purpose under Art. 5(1)(b)

04 Deleting training data does not delete model parameters that encode information derived from it

05 EU AI Act high-risk system obligations apply from August 2026 — assessment must happen now

Document ARI-2026-002 | Published May 2026 | axiomresearch.institute

About the Axiom Research Institute

The Axiom Research Institute is the independent research and publications division of Axiom Sovereign LLC, a fractional vCISO and GRC consulting firm. The Institute publishes primary research on post-quantum cryptography governance, AI privacy compliance, space cybersecurity, and emerging technology risk management. All publications are grounded in primary regulatory and standards sources and are designed to support practitioners, legal counsel, and executive leadership in navigating complex governance challenges.

About Axiom Sovereign

Axiom Sovereign LLC provides fractional vCISO services, GRC program design and assessment, vendor privacy risk analysis, and AI governance advisory for mid-market organizations in regulated sectors. Principal: Cory Missimore, CISSP, CISM, CISA, CIPP/E, CIPP/US, CIPM, AIGP.

How to Use This Paper

This is a foundational publication. It establishes the Institute's analytical framework and primary source grounding for this topic. Subsequent publications in the series address specific implementation domains and can be used independently or in conjunction with this paper. Implementation guides, assessment instruments, and templates derived from this paper are published separately and reference this document.

Document Number	ARI-2026-002
Publication Series	Privacy Governance
Published	May 2026
Version	1.0 — Initial Release
Suggested Citation	Axiom Research Institute. (2026). Privacy in the Age of Machine Learning (ARI-2026-002). Axiom Sovereign LLC.
Distribution	Public — Unrestricted
License	CC BY-NC 4.0 — Non-commercial with attribution

Disclaimer

EXECUTIVE SUMMARY

01

Privacy law and AI architecture are structurally incompatible — not incidentally.

GDPR, CPRA, and HIPAA were designed for traceable, human-controlled data processing. Machine learning systems optimize for statistical outcomes, encode training data as parameters rather than records, and resist the individual-level explanations that data subject rights require. These tensions are architectural, not resolvable by policy statement alone.

02

GDPR Article 32 makes quantum risk a present-tense privacy compliance obligation.

Article 32 requires "state of the art" security measures for personal data. NIST published post-quantum standards in August 2024. NSA has documented active Harvest Now, Decrypt Later collection operations. Personal data with long retention horizons — healthcare, biometric, financial — is both the primary HNDL target and the category most directly subject to Article 32. The argument that RSA-2048 satisfies "state of the art" is weakening with each passing month.

03

GDPR Article 22 creates an explainability requirement most organizations have not tested.

When an AI system makes or significantly contributes to a decision with legal or similarly significant effects, the data subject is entitled to a meaningful, specific explanation of the decision logic — not a description of how the system works in general. Most deployed machine learning systems cannot produce this. Most organizations have not assessed the gap.

04

Legal and technical AI governance programs operate separately. Neither is complete.

Privacy programs are not designed to evaluate model architecture against explainability requirements. AI governance programs are not designed to evaluate training data use against data protection law. The compliance gap sits between them, owned by neither, accruing exposure in both.

05

EU AI Act Annex III obligations apply from August 2026. The assessment window is closing.

Employment decisions, creditworthiness, insurance underwriting, clinical decision support, and essential services access — all within scope if AI contributes to the decision. The assessment, gap analysis, architectural remediation, and documentation development required for compliance take longer than most organizations allocate. For executives: the question to answer now is whether your organization has inventoried which deployed AI systems fall under Annex III. If that inventory does not exist, it needs to.

SECTION 1

The Foundational Mismatch

1.1 What Privacy Law Assumes About Data Processing

Modern data protection frameworks share a core architectural assumption. Data processing is a specific activity performed by an identifiable controller, for a stated purpose, using identified data, for a defined duration. GDPR Article 5 operationalizes this through six principles, three of which are directly relevant here: purpose limitation (data collected for specified, explicit, and legitimate purposes and not processed in ways incompatible with those purposes), data minimization (processing limited to what is necessary), and storage limitation (data retained only as long as the purpose requires, then deleted).

These principles assume that the relationship between data and purpose is knowable and verifiable, and that deletion of data terminates its influence on the organization's systems. Both assumptions break down in machine learning contexts.

1.2 What Machine Learning Systems Actually Do

A supervised machine learning model is trained on a dataset. The training process adjusts internal parameters — weights — to minimize prediction error against a ground truth label. When training concludes, those parameters encode a compressed statistical representation of the training data. The training data is not retained in the model. The patterns derived from it are.

> The purpose of training data use is not equivalent to the purpose of data collection. A health system collects patient records for clinical care. Using those records to train a clinical decision support model changes the processing purpose. GDPR Article 5(1)(b) requires that further processing be compatible with the original collection purpose. Training a commercial AI product on clinical data collected for direct patient care is not clearly compatible. The compatibility assessment has to happen before training begins. Most organizations have not done it.

Data minimization is structurally in tension with model accuracy. More training data produces more accurate models. "Minimum necessary" in machine learning contexts is not a fixed quantity — it is a function of the target accuracy level, which is itself a design choice. An organization cannot honestly claim principled data minimization unless it has documented why a specific dataset size was selected against a defined accuracy threshold.

Deleting training data does not delete the model's learned representation of it. The technical literature has established through membership inference attacks, model inversion attacks, and training data extraction research that trained models can, under specific conditions, reveal information about their training data. A

model trained on personal data is itself a system that processes information derived from personal data. Deleting the training dataset does not eliminate that system from the organization's information environment — and does not fully satisfy a GDPR Article 17 erasure request where the model's parameters encode derived information about the individual.

1.3 The Article 32 Problem — HNDL and AI as a Present-Tense Privacy Obligation

GDPR Article 32 requires controllers to implement "appropriate technical and organizational measures" to ensure a level of security appropriate to the risk — explicitly including "state of the art" protections. This provision has direct implications for both quantum risk and AI governance that most DPIAs have not analyzed.

For quantum risk: data stored or transmitted under RSA or ECC encryption is subject to Harvest Now, Decrypt Later collection today. If that data is personal data subject to GDPR, and adversaries are actively collecting it with the expectation of future decryption, then the question is not whether the organization will face quantum risk eventually — it is whether current encryption constitutes "state of the art" when NIST has published post-quantum alternatives and agencies have documented active collection operations. Organizations cannot dismiss this as a future compliance question. The argument that RSA-2048 satisfied Article 32 as of the date NIST published FIPS 203 is defensible. The argument that it satisfies Article 32 in 2026 is progressively less so.

> The intersection of HNDL and GDPR Article 32 is underanalyzed in most organizations' DPIAs and legal risk assessments. Personal data with long retention horizons — healthcare records, biometric data, financial history — is both the primary HNDL target and the category most directly subject to Article 32's "state of the art" requirement. This is not a future compliance question. The harvest is present-tense.

1.4 The Accountability Problem

GDPR Article 5(2) requires that the controller demonstrate compliance with the processing principles. The ability to demonstrate compliance requires explaining, for any processing activity, what data was used, for what purpose, under what legal basis, for what duration, and with what outcome.

Machine learning systems make this demonstration difficult. A readmission prediction model trained on 50,000 patient records does not process any particular patient's record at inference time in a way that produces a traceable audit record. The model's output is a function of all training data, its influence distributed across learned parameters in ways that cannot be attributed to specific records without specialized techniques. This creates a direct conflict with GDPR Article 22.

SECTION 2

The Regulatory Obligations in Specific Terms

2.1 GDPR Article 22

Article 22 covers decisions based solely on automated processing that produce legal effects or similarly significant effects on the data subject. Three exceptions exist — contractual necessity, Union or Member State law authorization, and explicit consent — each with conditions.

The contractual necessity exception requires the controller to implement measures that safeguard data subject rights, specifically including the right to obtain human intervention, express a point of view, and contest the decision. Nominal human review — a person who approves model outputs without substantive evaluation — does not satisfy this requirement. The EDPB's Guidelines 02/2022 on Article 22 are explicit on this point.

> The compliance question is operational: for every AI system that makes or significantly contributes to decisions with legal or similarly significant effects, can the organization produce — on demand, for a specific individual — a meaningful explanation of the decision logic for their particular case? For most organizations deploying ML in hiring, credit, insurance, clinical recommendation, or benefits determination contexts, the answer is no, and the gap has not been assessed.

2.2 CPRA and Automated Decision-Making Technology

California Civil Code Section 1798.185(a)(16) requires the California Privacy Protection Agency to establish rules governing automated decision-making technology. The CPPA's draft rulemaking (2023, ongoing revision) is broader than Article 22 in a specific way: it applies to significant decisions regardless of whether processing is solely automated. A human who reviews model outputs without substantive evaluation of the underlying logic does not convert automated processing into human decision-making under the draft rules.

2.3 The EU AI Act

Regulation (EU) 2024/1689 creates a risk-based framework that in specific respects exceeds GDPR obligations for high-risk AI deployments. Annex III high-risk categories include: safety components in critical infrastructure, biometric categorization, employment and workforce management decisions (recruitment, performance evaluation, promotion, termination), access to essential services (creditworthiness, insurance risk), law enforcement, migration and border control, and administration of justice.

Article 9 requires a risk management system throughout the system lifecycle. Article 10 requires data governance for training, validation, and testing datasets — directly engaging data minimization tensions.

Article 11 requires technical documentation including a description of the logic of the system, the same information GDPR Article 22 requires the controller to provide to data subjects. Article 13 requires that deployers receive sufficient information about purpose, accuracy conditions, limitations, and input data specifications.

Deployers of purchased high-risk AI systems carry their own obligations: fundamental rights impact assessments, human oversight implementation, and operational monitoring. For organizations purchasing AI tools from third-party vendors in Annex III contexts, Article 11 documentation is a procurement requirement. Vendors who cannot supply it are not compliant providers of high-risk AI systems.

2.4 HIPAA Applied to AI

OCR has not issued guidance specifically addressing AI training on ePHI as of this publication. The existing framework applies nonetheless. The minimum necessary standard (45 CFR 164.502(b)) requires limiting ePHI use to the minimum necessary for the intended purpose — not to the maximum available from the organization's records systems. The Privacy Rule's use and disclosure provisions require that ePHI use for purposes other than treatment, payment, or healthcare operations be either permitted or authorized. AI model training for commercial purposes requires analysis of whether the use falls within healthcare operations (45 CFR 164.501) or requires individual authorization.

> OCR enforcement in analogous contexts — health systems sharing patient data with technology vendors for non-clinical AI development without proper BAAs — indicates OCR will apply the existing framework when violations are surfaced. The absence of AI-specific guidance does not create a compliance safe harbor.

SECTION 3

The Organizational Governance Gap

3.1 Two Programs That Do Not Speak to Each Other

The dominant failure in organizational AI privacy governance is structural. Privacy and legal teams run privacy programs. Technical teams run AI governance programs. These programs operate largely in parallel, each addressing part of the compliance picture, neither addressing it completely.

The privacy program manages DPAs, privacy policies, data subject request handling, and DPIA completion. It is not typically staffed by people who can evaluate whether a specific model architecture can produce an Article 22-compliant explanation, or whether a machine unlearning implementation is sufficient to fulfill an erasure request. The AI governance program manages model development processes, model cards, bias evaluations, and deployment approvals. It is not typically staffed by people who can assess whether training data use is compatible with the original collection purpose under Article 5(1)(b), or whether an EU data subject's data used in US-based model training creates a Chapter V transfer issue.

> The compliance gap sits between these programs. Neither owns it. The legal exposure concentrates in the organization regardless.

3.2 DPIAs That Miss the Point

GDPR Article 35 requires a DPIA when processing is likely to result in high risk, specifically including systematic and extensive evaluation based on automated processing including profiling. AI deployments that make or contribute to significant individual decisions are within scope on their face. The question is not whether DPIAs are being completed. In many organizations, they are. The question is whether they are adequate.

An adequate DPIA for an AI system must address: whether the system can produce a data-subject-specific explanation satisfying Article 22; whether training data use is compatible with the original collection purpose under Article 5(1)(b) and what the legal basis is if not; what the minimum dataset necessary to achieve the specified performance requirement is and how that minimum was determined; how erasure requests for individuals in the training set will be fulfilled; and what the mechanism for portability fulfillment is. A DPIA that does not address these questions is a completed form, not a completed assessment.

3.3 Cross-Border Training

Organizations training AI models on EU personal data using US-based infrastructure are conducting a transfer of personal data to a third country under GDPR Chapter V. The transfer mechanism requirements

apply regardless of whether the organization considers the processing to be "internal."

The EU-US Data Privacy Framework (July 2023) restored adequacy as a transfer mechanism for DPF-certified US organizations. The framework is subject to a legal challenge from noyb filed with the Irish Data Protection Commission. The structural objection — that FISA Section 702 surveillance authority is incompatible with GDPR's requirement for essentially equivalent protection — was the objection that invalidated Privacy Shield in Schrems II. The DPF addresses this through executive order commitments and a redress mechanism, not through legislative change to FISA. Whether that is sufficient is a question the CJEU has not answered in this specific context.

> Organizations whose AI training infrastructure depends on DPF adequacy as its sole transfer mechanism are carrying concentration risk. The appropriate response is parallel mechanisms: maintain updated SCCs and completed Transfer Impact Assessments alongside DPF reliance, so that a change in DPF status does not create an immediate compliance gap.

SECTION 4

What Adequate Governance Looks Like

4.1 Integrating the Two Programs

Privacy governance and AI governance need shared review gates with defined integration points. This is not about merging teams. It is about establishing the specific decision points where both perspectives must be present before a deployment proceeds.

The DPIA template must be extended with mandatory AI addendum questions triggered by any deployment involving machine learning in a decision-adjacent context. Pre-deployment review must include concurrent legal and technical participation — not sequential sign-off — before any system in scope for Article 22, CPRA automated decision-making rules, or EU AI Act Annex III reaches production. Vendor AI due diligence requires the same questions applied outward: Article 11 documentation review, Article 22 explanation capability assessment, and training data legal basis verification for AI systems purchased from third parties in Annex III contexts.

4.2 Explainability in Practice

The legal standard is "meaningful information about the logic involved" — not a mathematically complete causal attribution. That standard is achievable for many deployment contexts, but it requires the question to be asked at design time, not at enforcement time.

For many high-risk applications, architectures that support interpretable outputs — logistic regression, decision trees, rule-based systems, constrained gradient boosting — can satisfy Article 22 without requiring post-hoc explanation techniques. The choice to deploy a less interpretable architecture is a design decision with legal implications that should be documented. Where complex architectures are necessary, model-agnostic techniques — LIME, SHAP, counterfactual explanations — can produce data-subject-specific outputs satisfying the legal standard in most contexts. The governance requirement is that the explanation mechanism is validated against the applicable legal standard before deployment.

4.3 Data Subject Rights for Training Data

An individual whose data was used in model training has rights under GDPR Article 17 (erasure) and Article 20 (portability). Deleting the training dataset does not satisfy the erasure obligation if the model's parameters encode information derived from that individual's data. The technical field of machine unlearning addresses this, but techniques are computationally expensive and not uniformly available across model architectures. Organizations using personal data for AI training should assess, before beginning training, how they would fulfill an erasure request against both the training dataset and the trained model, and document that assessment as an input to architecture selection.

SECTION 5

The Regulatory Timeline

5.1 EU AI Act Enforcement Schedule

Date	Obligation
February 2025	Article 5 prohibited practices effective. Includes subliminal manipulation, exploitation of group vulnerabilities, real-time
August 2025	GPAI model provider obligations (Articles 51–55). Technical documentation, training data summaries, cybersecurity ar
August 2026	High-risk AI system obligations (Annex III). Employment decisions, creditworthiness, insurance, essential services, clir

> Organizations that have not conducted an inventory of their AI deployments against Annex III categories should not assume they have no systems in scope. The categories are broad and the applications are common. Hiring tools that score resumes, credit models that inform underwriting, clinical decision support that informs treatment recommendations — each requires analysis before August 2026.

5.2 US State Law Trajectory

Colorado's Artificial Intelligence Act (SB 205), effective February 1, 2026, creates obligations for developers and deployers of high-risk AI systems: risk management practices, impact assessments, and transparency obligations. Colorado's framework is modeled on the EU AI Act and is the most comprehensive state-level AI governance law in the United States at time of publication. Texas, Illinois, and Virginia have advanced AI legislation at various stages. The direction is toward more requirements, not stabilization. An organization fully compliant with EU AI Act obligations will generally satisfy US state AI governance requirements derived from or modeled on the EU framework.

CONCLUSION

The governance frameworks most organizations apply to AI privacy compliance are insufficient because they were designed before AI was a material compliance domain and have not been updated to address the specific ways machine learning systems create tension with data protection obligations.

The tensions are not irresolvable. They require specific analysis at specific decision points: training data compatibility assessment before training begins, explainability mechanism validation before deployment, data subject rights procedures documented before models enter production, and cross-border transfer mechanism verification before EU data moves to non-EU training infrastructure. These are assessable,

documentable, and defensible — but only if the questions are asked at design time, not in response to a regulatory inquiry.

> The organizations that will demonstrate compliant AI governance when regulators examine them are not the ones with the most sophisticated AI systems. They are the ones that assessed specific legal obligations against specific system architectures before deployment, not after an inquiry.

For legal counsel and executives reading this paper: the immediate question is whether the organization has conducted an inventory of AI deployments against GDPR Article 22, CPRA automated decision-making provisions, and EU AI Act Annex III — and whether the DPIA process has been extended to ask the AI-specific questions this paper identifies. If neither has happened, that is the starting point.

Axiom Research Institute's subsequent publications in this series address DPIA methodology for AI deployments, training data compliance frameworks, and cross-border AI governance. Organizations seeking practitioner support for AI privacy governance may request an initial consultation at axiomsovereign.com.

REFERENCES

European Parliament and Council. (2016). General Data Protection Regulation. Regulation (EU) 2016/679. EUR-Lex.

European Parliament and Council. (2024). Artificial Intelligence Act. Regulation (EU) 2024/1689. EUR-Lex.

European Data Protection Board. (2022). Guidelines 02/2022 on the application of Article 22 of the GDPR in the context of automated individual decision-making.

European Data Protection Board. (2017). Guidelines on Data Protection Impact Assessment (WP248 rev.01).

State of California. California Privacy Rights Act. California Civil Code Sections 1798.100–1798.199.4.

California Privacy Protection Agency. (2023). Automated Decision-Making Technology Regulations (Draft Rulemaking).

State of Colorado. (2024). Colorado Artificial Intelligence Act. SB 205. Effective February 1, 2026.

U.S. Department of Health and Human Services. (2013). HIPAA Privacy Rule. 45 CFR Parts 160 and 164.

Court of Justice of the European Union. (2020). Data Protection Commissioner v. Facebook Ireland Limited and Maximillian Schrems. Case C-311/18.

European Commission. (2023). Adequacy Decision for the EU-US Data Privacy Framework. Commission Implementing Decision 2023/1795.

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. IEEE Symposium on Security and Privacy.

Carlini, N., et al. (2021). Extracting training data from large language models. USENIX Security Symposium.