

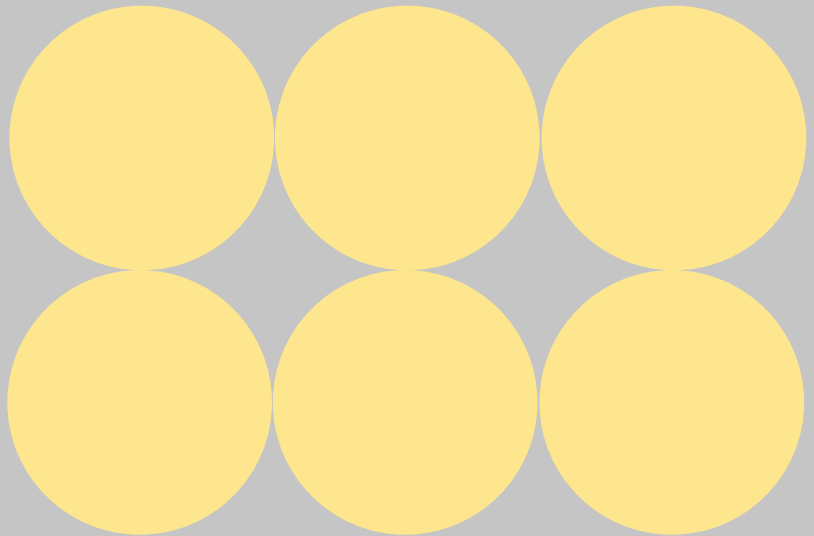


Legal Data
Intelligence™

The LDI Guide to Data Protection in AI Governance

By: Ryan Costello, Mike Kearney, and Briordy Meyers

September 2026



1. Introduction.....	2
2. LDI's Three "I"s and AI Governance.....	3
3. The Five AI Governance Dimensions	3
3.1 Data Governance and Privacy	4
3.2 People, Accountability, and Oversight	5
3.3 Third-party and Ecosystem Risk	6
3.4 Responsible AI Lifecycle	6
3.5 Technology and Security Assurance	7
4. Conclusion.....	8
5. Appendix: Cross-Reference of Major AI Governance Frameworks, Laws, Standards, and Selected Resources	9

1. Introduction

As organizations continue to rapidly adopt and integrate AI systems, corporate legal and privacy teams face the monumental task of transitioning from static AI governance checklists to dynamic, operational governance structures. AI governance is an enterprise risk discipline that requires clear accountability, strong information governance, secure technology, continuous assurance, and effective oversight of third-party providers.

The Legal Data Intelligence (LDI) Guide to Data Protection in AI Governance (Guide) is designed to provide corporate legal departments, Data Protection Officers (DPOs), and risk management professionals with an actionable, operational blueprint for embedding those capabilities across the AI lifecycle. The Guide merges the LDI Model's Three "I"s — **Initiate**, **Investigate**, and **Implement** — with five AI Governance Dimensions (see Figure 1). The result is a continuous, lifecycle-based compliance workflow that translates abstract ethical principles into enforceable, technical, and contractual controls rooted in data protection principles.

This Guide is inspired by, but not dependent on existing AI frameworks, laws, and regulations, including those identified in the Appendix. It uses established AI guidelines as important substantive guidance for legal data intelligence practitioners in developing an LDI-specific AI governance approach grounded in data protection principles; it is not intended to suggest one-to-one compliance with any particular framework, law, or regulation.

2. LDI's Three "I"s and AI Governance

This Guide is designed to operationalize governance efforts and data protection compliance across business operations. It prioritizes usability by organizing compliance actions within each Dimension into the Three "I"s of Legal Data Intelligence, each designed to ensure that governance efforts are practical and repeatable:

- **Initiate:** Focuses on foundational preparation — establishing clear organizational policies, aligning stakeholder roles (Chief Privacy Officer (CPO), Data Protection Officer (DPO), Chief Legal Officer (CLO), and Chief Information or Technology Officer (CIO/CTO)), verifying legal and regulatory requirements specific to the industry, and defining risk tolerance thresholds.
- **Investigate:** Focuses on active risk assessment and evaluation—conducting Data Protection Impact Assessments (DPIAs), Algorithmic Impact Assessments (AIAs), red teaming, threat modeling, and validating data and model characteristics prior to deployment.
- **Implement:** Focuses on operationalization and enforcement — deploying Privacy Enhancing Technologies (PETs), enforcing runtime guardrails, maintaining automated system logs, establishing incident response containment paths, and conducting continuous post-deployment monitoring.



Figure 1: The Interaction of AI Governance Dimensions with Data Protection and Privacy Principles

3. The Five AI Governance Dimensions

While LDI's Three "I"s provide usability, consistency, and progression across the model, the Dimensions provide the substantive AI governance elements. The Three "I"s organize the

considerations within each Dimension, translating a data protection compliance approach into practice. Although Data Governance and Privacy is defined as one Dimension of the AI governance elements, its considerations do not operate in isolation; they apply throughout all Dimensions shown here. The Guide also introduces considerations relevant to integrating AI governance into an overall risk program. Many considerations must be evaluated through a risk-based lens that accounts for the legal, regulatory, business, functional, and cost-related risks associated with implementing and operating AI systems. Organizations should assess the varying risk levels presented by individual AI systems and tailor their approach to their risk appetite, technology, industry, and regulatory profile.

3.1 Data Governance and Privacy

Trustworthy AI is fundamentally built upon lawfully processed data. The Data Governance and Privacy Dimension unifies data quality, retention, and privacy obligations under a single discipline: ensuring that the data ingested by models is legally obtained, accurate, and fit for purpose; that it is retained no longer than necessary and preserved where legally required; and that privacy by design and default is embedded as a non-negotiable prerequisite for lawful AI operations, protecting individual autonomy, identity, and dignity. Each of the considerations below should be addressed following a risk-based approach, calibrated to the risk tier of the AI system and the degree of human oversight required:

Initiate:

- Research your organization's data protection and privacy compliance obligations based on your business model and relevant jurisdictions. Align your data set approach for AI ingestion with your strategic risk appetite.
- Formulate a comprehensive, risk-based data management plan that defines minimum data quality attributes (e.g., accuracy, completeness, consistency, relevance, and validity) and verify a documented lawful basis (e.g., explicit consent, contract performance, legitimate interest) for all personal data before ingestion.
- Establish a resourced AI System Inventory logging active and decommissioned models, metadata, training datasets, first use dates, and curation pipelines, with formal retention policies for training data sets, synthetic media, and operation logs.
- Embed privacy controls (e.g., purpose limitation, data minimization, and transborder data flow compliance) directly into system design specifications.
- Establish Records and Information Management (RIM) accountability for identifying and classifying AI record types, defining retention and disposition requirements, and coordinating preservation obligations across AI systems and related data stores.

Investigate:

- Conduct any mandatory Data Protection Impact Assessments (DPIAs), Fundamental Rights Impact Assessments (FRIAs), and Algorithmic Impact Assessments (AIAs) to map complete data provenance, curation methodologies, and transformations, and inspect data sets for hidden biases, errors, or over-representation that could produce discriminatory outputs.
- Audit retention schedules against data minimization standards, confirming that training files and backups containing personal information are not held indefinitely under the guise of model fine-tuning. Analyze models for severe privacy vulnerabilities (e.g., training data memorization, de-anonymization of synthetic media, and PII leakage under adversarial queries) and evaluate impacts on data subject rights, including automated profiling opt-outs.

Implement:

- Deploy standardized data statements, model cards, and data labels for transparent, machine-readable provenance, supported by preprocessing transformations and continuous data lineage tracking to maintain a single, legally defensible source of truth.
- Retain technical documentation and auto-generated event logs for legally mandated periods (e.g., at least 10 years after a high-risk AI system is placed on the market) and implement secure, automated deletion and anonymization protocols upon expiration.
- Ensure that automated deletion, anonymization, overwriting, model retirement, and other disposition actions can be suspended for information subject to a legal hold, investigation, audit, regulatory request, or other authorized preservation requirement.
- Deploy state-of-the-art Privacy Enhancing Technologies (PETs) where needed (e.g., differential privacy, federated learning, data masking, cryptographic verification) alongside real-time AI interaction disclosures and functional channels for data subjects to access, correct, or delete their data.

3.2 People, Accountability, and Oversight

Humans are both the primary oversight mechanism and a significant source of operational risk in AI systems. The People, Accountability, and Oversight Dimension establishes clear accountability structures and mandatory AI literacy to mitigate cognitive biases, protect personal data, and satisfy emerging regulatory mandates, ensuring human oversight is meaningful and active rather than nominal. Each of the considerations below should be addressed following a risk-based approach, calibrated to the risk tier of the specific AI system and the degree of human oversight required:

Initiate:

- Define specific roles and responsibilities for all AI actors, including developers, operators, compliance auditors, and Records and Information Management personnel. Use role definitions applicable to the relevant jurisdiction — for example, the EU AI Act in the European Union and the The Information Commissioner's Office (ICO)'s AI and Data Protection Guidance in the United Kingdom — and codify organizational proficiency standards and baseline literacy metrics tailored to the risk tier of each AI system.

Investigate:

- Evaluate staff competencies and assess cognitive human-AI interaction risks.
- Specifically test for automation bias (the tendency of operators to blindly trust algorithmic outputs) and selective adherence to verify that oversight functions as intended.

Implement:

- Deliver ongoing, role-specific training on interpreting complex AI outputs, detecting algorithmic bias, and applying data protection rules.
- Empower human operators with clear authorization and procedural pathways to challenge, override, or decommission non-compliant outputs.

3.3 Third-party and Ecosystem Risk

The Third-party and Ecosystem Risk Dimension recognizes that most enterprises procure access to some, if not all, of the AI models they use. As a result, managing third-party relationships and data supply chains through legally binding agreements is essential to mitigating legal, privacy, and security risks in AI governance. Each consideration below should be addressed using a risk-based approach calibrated to the risk tier of the AI system and the degree of human oversight required:

Initiate:

- Engage legal, procurement, and DPO teams early.
- Formulate standardized contractual templates, Data Processing Agreements (DPAs), and Service Level Agreements (SLAs).
- Ensure those agreements explicitly address data ownership, vendor lock-in protections, liability allocation for model drift, and contemporaneous documentation of vendor performance, platform usage, and related governance evidence.
- Prohibit the use of proprietary company data for public model training without explicit consent, subject to applicable contractual terms, General Data Protection Regulation (GDPR) or UK GDPR data-processing obligations, and vendor-specific acceptable-use policies.

Investigate:

- Conduct comprehensive third-party due diligence, verifying vendor compliance with international data protection and privacy laws and technical standards.
- Review vendor transparency artifacts — technical documentation, conformity assessments, system cards, and model cards — for alignment with enterprise risk tolerance.
- Where specific contractual provisions related to privacy and AI governance cannot be negotiated, determine the need for other mitigating controls to ensure an acceptable level of risk associated with the use of third-party models.

Implement:

- Establish continuous vendor monitoring covering performance stability, security patches, and bias drift.
- Enforce robust contractual audit and verification rights that enable independent third-party inspection of vendor AI data pipelines and security controls and address the need for additional mitigating controls based on the results of those audits.

3.4 Responsible AI Lifecycle

Governing the full lifecycle — from design through deployment, change, and decommissioning — ensures that data protection by design and default is systematically embedded rather than treated as an afterthought. AI systems carry unique risks (e.g., new model releases, model drift, emergent behaviors, and non-deterministic outputs) that demand continuous oversight and structured change management rather than one-time checks or traditional patch management alone. Each consideration constituting the Responsible AI Lifecycle Dimension should be addressed using a risk-based approach calibrated to the risk tier of the AI system and the degree of human oversight required:

Initiate:

- Mandate early involvement of the DPO, legal counsel, and risk managers during the planning and design stage, documenting the system's intended purpose, data flows, legal context, and constraints before model building begins.
- Draft robust change management policies requiring formal compliance review and sign-off for substantial modifications to model weights, APIs, or database schemas, and plan for infrastructure interoperability. Ensure policy obligations are right sized for respective AI system risks.
- Create AI model and system end-stage criteria, with consistent retirement triggers such as persistent performance or bias drift, unsupported dependencies, material regulatory change, unacceptable security exposure, inability to maintain required human oversight, or disproportionate remediation cost. Align retirement decisions with data management, legal hold, and records retention policies. The Singapore InfoComm Media Development Authority (IMDA) Agentic AI Framework (2026) provides practical guidance on model retirement triggers and decommissioning criteria.

Investigate:

- Establish a structured Test, Evaluation, Verification, and Validation (TEVV) protocol, conducting pre-deployment evaluations with independent reviewers to stress-test capabilities and verify operation within defined privacy limits.
- Test regularly post-deployment for concept drift and data drift (e.g., performance degradation as operational environments diverge from the training baseline) and audit release schedules, compatibility layers, and hotfixes across integrated third-party systems.

Implement:

- Deploy in a controlled production environment with automated, tamper-evident logging of all data processing and inference events for auditability and set deterministic runtime thresholds to disengage or bypass the AI if outputs deviate from safety, accuracy, or data protection standards.
- Link continuous monitoring plans to automated alerting, and maintain meticulous version histories, model metadata, and instant rollback capabilities to restore a compliant model state when critical drift is detected.
- Implement an AI model and system retirement plan if and when an individual model meets the established retirement criteria.

3.5 Technology and Security Assurance

AI security extends traditional IT security to protect a unique and highly complex attack surface. Safeguarding model weights, code, data pipelines, and the underlying computational infrastructure is critical to preventing catastrophic data breaches and operational failures. Each of the considerations of the Technology and Security Assurance Dimension should be addressed following a risk-based approach, calibrated to the risk tier of the AI system and the degree of human oversight required:

Initiate:

- Identify AI-specific security threats (e.g., prompt injection, model inversion, data poisoning, model extraction, and evasion attacks) alongside traditional IT risks.
- Define the secure computational resources required to maintain system safety and formulate a robust incident response plan with explicit containment, escalation, and breach notification workflows for AI-related incidents. Identify and operationalize the notification windows applicable across relevant jurisdictions and sectors, rather than relying on a single statutory period; examples include 72 hours under EU GDPR Article 33 and UK GDPR Article 33, varying deadlines under US state laws, and sector-specific requirements in financial services and healthcare.

Investigate:

- Conduct regular AI red teaming, stress testing, and chaos testing using independent, qualified security assessors.
- Subject models to adversarial prompt injection and poisoning scenarios to proactively discover bypass risks and secure endpoints.

Implement:

- Implement robust security controls — input throttling, output filters, endpoint authentication, and watermarking.
- Continuously monitor the runtime environment and share threat intelligence, incident patterns, and security near-misses with AI-specific platforms such as MITRE Adversarial Threat Landscape for AI Systems (ATLAS) and the AI Incident Database, and with relevant vulnerability databases including Common Vulnerabilities and Exposures (CVE) and National Vulnerability Database (NVD).

4. Conclusion

The LDI Guide to Data Protection in AI Governance provides a practical, risk-based approach for translating AI governance principles into accountable and repeatable action. By combining LDI's Three "I"s of Initiate, Investigate, and Implement with the five AI Governance Dimensions, the Guide helps organizations move beyond static compliance checklists toward lifecycle governance rooted in data protection, legal accountability, human oversight, vendor assurance, secure technology, and continuous monitoring.

As agentic AI systems become more prevalent across legal and compliance functions, the Guide's phases will require iterative adaptation to address autonomous decision-chaining, multi-system interactions, and the heightened oversight demands these architectures present.

Effective AI governance depends on embedding privacy, security, transparency, and accountability into every stage of AI system design, deployment, operation, and retirement. Organizations that apply this process consistently can better align AI innovation with regulatory obligations, enterprise risk tolerance, and the broader expectation that AI systems be lawful, trustworthy, and responsibly managed.

5. Appendix: Cross-Reference of Major AI Governance Frameworks, Laws, Standards, and Selected Resources

This Appendix identifies major international and national AI governance frameworks, laws, standards, and selected research resources that provide important substantive guidance for developing LDI's Guide to Data Protection in AI Governance. Consistent with the Introduction, these authorities inform the Guide's data protection-based approach but do not make it dependent on, or suggest one-to-one compliance with, any framework, law, regulation, standard, or resource.

Resource	Issuing Body	Core Methodology and Risk Classification	Key Compliance Focus and Applicability
AI, Data Governance, and Privacy: Synergies and Areas of International Co-operation	Organisation for Economic Co-operation and Development (OECD)	Policy analysis of interactions among AI governance, data governance, and privacy.	Supports coordination of privacy, data governance, and AI policy across jurisdictions.
Assessment List for Trustworthy AI (ALTAI)	European Commission High-Level Expert Group	Self-evaluation tool structured around seven trustworthy-AI requirements.	Voluntary public- and private-sector assessment aligned with European norms.
Canada Directive on Automated Decision-Making & AIA Tool	Treasury Board of Canada Secretariat	Mandatory questionnaire calculating impact across four tiers.	Mandatory for Canadian federal public-sector automated decision systems.
Datasheets for Datasets	Gebru, Morgenstern, Vecchione, et al.	Dataset documentation framework covering motivation, composition, collection, preprocessing, uses, distribution, and maintenance.	Promotes transparency, accountability, and informed use of datasets.

Resource	Issuing Body	Core Methodology and Risk Classification	Key Compliance Focus and Applicability
EU Artificial Intelligence Act (Regulation (EU) 2024/1689)	European Parliament & Council of the European Union	Binding tiered risk framework covering prohibited, high-risk, transparency, and minimal-risk AI.	Applies to organizations placing AI on the EU market or putting it into service in the EU.
Extracting Training Data from Large Language Models	Carlini et al.	Empirical security research demonstrating extraction of memorized training data from language models.	Relevant to privacy testing, memorization risk, and data-leakage controls.
Human-AI Interactions in Public Sector Decision Making	Alon-Barkat & Busuioc	Research on automation bias and selective adherence to algorithmic advice.	Supports meaningful human oversight and testing of decision-maker reliance on AI outputs.
Human Rights, Democracy, and Rule of Law Impact Assessment (HUDERIA)	Council of Europe	Context-based risk and impact assessment for human rights, democracy, and rule-of-law effects.	Supports implementation of the Council of Europe AI Framework Convention.
IAPP-EY Professionalizing Organizational AI Governance Report	International Association of Privacy Professionals & Ernst & Young	Organizational research on maturing AI governance functions, roles, and operating models.	Useful for governance design, accountability, staffing, and program benchmarking.
IAPP-FTI Privacy and AI Governance Report	International Association of Privacy Professionals & FTI Consulting	Research on integrating privacy functions with enterprise AI governance.	Useful for privacy-led governance structures, responsibilities, and controls.

Resource	Issuing Body	Core Methodology and Risk Classification	Key Compliance Focus and Applicability
ICO Guidance on AI and Data Protection	UK Information Commissioner's Office (ICO)	Lifecycle guidance applying UK GDPR principles to AI, including accountability, fairness, transparency, minimization, security, and individual rights.	Applies to organizations processing UK personal data or deploying AI systems in the United Kingdom.
ISO/IEC 22989:2022	ISO / IEC	Foundational concepts and standardized terminology for artificial intelligence.	Provides a common vocabulary for governance, risk, technical, and contractual work.
ISO/IEC 42001:2023 (AIMS Standard)	ISO / IEC	Certifiable Plan-Do-Check-Act management system for governing AI risks and supply chains.	Global management system standard that can integrate with ISO 27001 and ISO 27701 programs.
Model Cards for Model Reporting	Mitchell, Wu, Zaldivar, et al.	Standardized documentation of model details, intended uses, evaluation factors, metrics, and limitations.	Supports transparent model reporting, review, and responsible deployment decisions.
NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0)	National Institute of Standards and Technology	Lifecycle operating model organized around GOVERN, MAP, MEASURE, and MANAGE.	Voluntary, technology-agnostic framework widely used by US agencies and enterprises.

Resource	Issuing Body	Core Methodology and Risk Classification	Key Compliance Focus and Applicability
OECD AI Principles / Revised Recommendation of the Council on Artificial Intelligence	OECD	Trustworthy AI principles covering inclusive growth, rights, transparency, robustness, security, safety, and accountability.	Voluntary international baseline referenced by governments and organizations.
Red Teaming for Generative AI: Silver Bullet or Security Theater?	Feffer et al.	Research evaluating the capabilities and limits of generative-AI red teaming.	Informs assurance design, test scope, assessor independence, and interpretation of findings.
Singapore Model AI Governance Framework (Traditional, Generative, and Agentic)	Infocomm Media Development Authority (IMDA) and Personal Data Protection Commission (PDPC) Singapore	Practical, voluntary guidance spanning governance, operations, stakeholder interaction, and runtime safeguards.	Includes self-assessment and evaluation resources for traditional, generative, and agentic AI.
Should It Stay or Should It Go? The Legal, Policy and Technical Landscape Around Data Deletion	Center for Democracy and Technology	Legal, policy, and technical analysis of data-deletion requirements and implementation challenges.	Relevant to retention, deletion, anonymization, backups, and defensible disposition.
The False Comfort of Human Oversight as an Antidote to A.I. Harm	Green & Kak / AI Now Institute	Critical analysis of limitations in relying on nominal human oversight to mitigate AI harm.	Supports stronger oversight design, authority, training, and escalation controls.

Resource	Issuing Body	Core Methodology and Risk Classification	Key Compliance Focus and Applicability
UK AI Governance: A Pro-Innovation Approach to AI Regulation	UK Department for Science, Innovation and Technology	Sector-led, non-statutory approach applying cross-sector principles through existing regulators.	Guides UK public- and private-sector AI governance alongside regulator-specific requirements.
US OMB Memoranda M-25-21 and M-25-22	White House Office of Management and Budget	Federal guidance on agency use and procurement of artificial intelligence technology.	Applies to US federal agencies and relevant contractors, not as general private-sector authority.