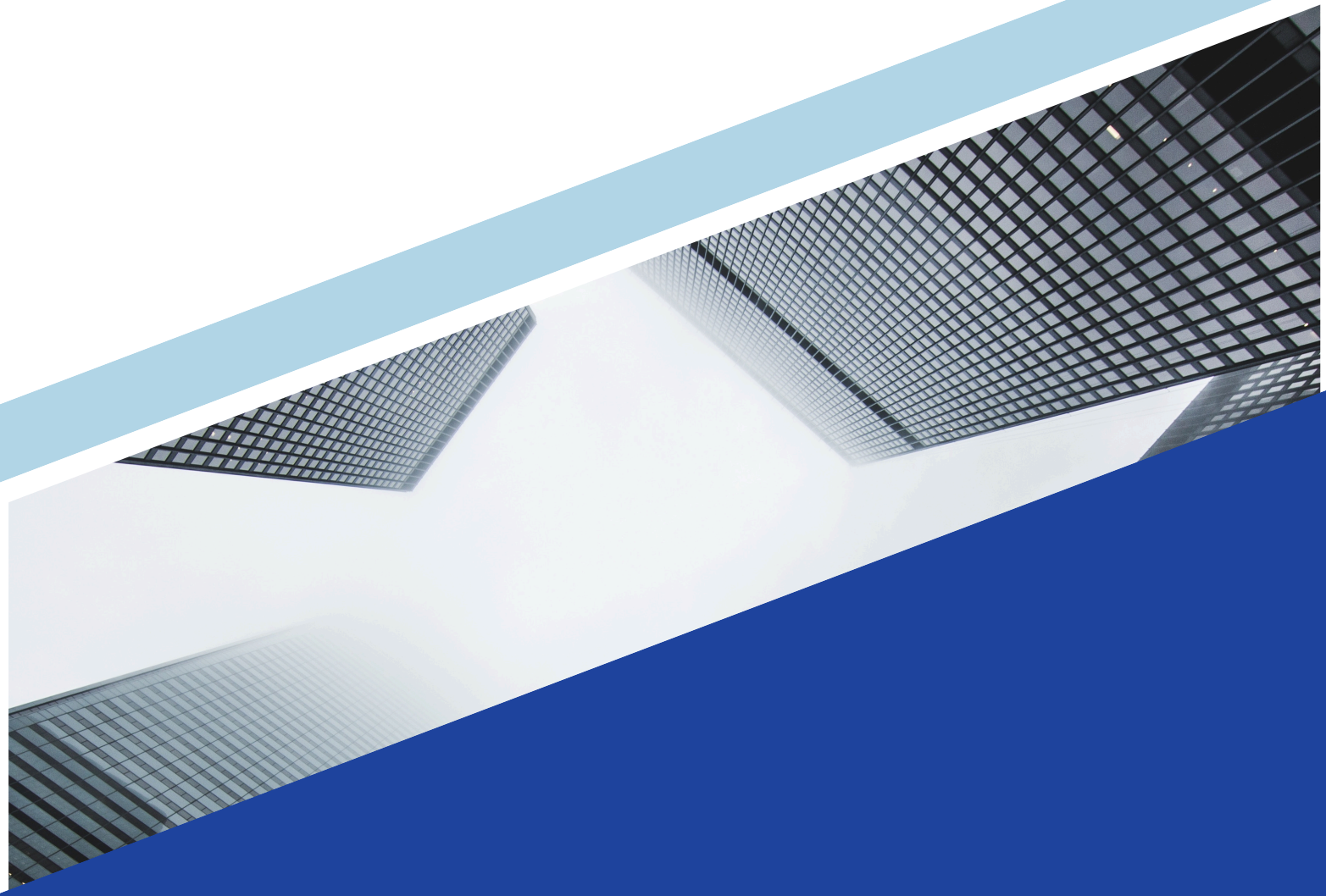
industry intelligence inc.®

AI Data Protection Framework



Effective Date: September 19, 2025
Version: 1.3

1990 S BUNDY DR,
LOS ANGELES, CA 90025
WWW.INDUSTRYINTEL.COM

Scope & Purpose:

Abstract: This AI Data Protection Framework serves as the operational companion to Industry Intelligence Inc.'s Data Security & Governance Policy, detailing the specific frameworks, guidelines, and technical controls we implement to ensure responsible AI usage and client data protection. While the Data Security & Governance Policy establishes our overarching security commitments and principles, this Framework provides the practical implementation standards and procedures that enforce those policies within our AI-enabled systems.

This document explains how Industry Intelligence Inc. ("IndustryIntel") integrates artificial intelligence into our products and services while maintaining the highest standards of data protection, privacy, and security for our clients.

Applies to: All current and future AI-enabled products/services and IndustryIntel Workforce Teammates (collectively, "Teammates")

Audience: Clients, partners, employees, and contractors

Standards basis: NIST AI RMF 1.0 and NIST AI 600-1

Data Governance & Protection Commitments

- No training on your data (default). Client content (prompts, files, chats, outputs, and metadata) is not used to train or fine-tune foundational models unless expressly authorized in a contract. We configure vendors to disable provider-side training where controls exist; if "no-training" cannot be contractually guaranteed, we do not use that vendor for client data.
- Purpose-limited processing & tenant isolation. Client data is processed strictly to deliver contracted services (e.g., summarization, tagging, retrieval-augmented Q&A) with logical and technical separation by customer/tenant.
- Data minimization & privacy-enhancement. We prefer retrieval over retention, avoid personal data where feasible, and apply privacy-enhancing techniques (redaction, tokenization, aggregation) consistent with Privacy-Enhanced design.
- Provenance & transparency. AI-assisted outputs are labeled. We preserve and display source links for generated summaries/answers and, where practical, propagate provenance signals (e.g., watermarks/signatures).
- Human oversight for consequential outputs. Analysts can review, challenge, and override system outputs; we expose limitations, confidence cues, and citations to support Explainable/Interpretable use.
- Secure by design. Encryption in transit/at rest, least privilege, MFA, secure SDLC, and monitoring/red-teaming against AI-specific threats (prompt injection, data poisoning, model misuse).
- Fairness and bias management. We measure and mitigate harmful bias and performance disparities and document context/limitations, consistent with Fair — with harmful bias managed.
- Incident handling & disclosure. We maintain AI incident response and disclosure processes (including after-action reviews and logging) appropriate to severity and context.

Framework foundation (NIST-aligned)

We operationalize NIST's AI RMF Core — GOVERN, MAP, MEASURE, MANAGE — and its trustworthiness characteristics (valid & reliable; safe; secure & resilient; accountable & transparent; explainable & interpretable; privacy-enhanced; fair with harmful bias managed).

2.1 GOVERN — organization-wide guardrails

- Policies & risk tolerance. Documented acceptable-use, data handling, and AI risk policies; inventory of AI systems; safe decommissioning; leadership ownership of risk.
- Roles & training. Clear accountability across product/ engineering/ analyst/ security/ legal; periodic training on AI risks and provenance.
- Third-party & IP governance. SLAs/DPAs cover "no training," content/IP ownership, incident duties, provenance expectations, audit rights; approved-provider list with ongoing review.
- Incident processes. Policies for AI incident logging, red-teaming, disclosure, and continuous improvement.

2.2. MAP — understand context, data, and impacts

- Use-case scoping. Intended purpose, users, context, assumptions, foreseeable misuse/abuse, and off-label use documented for each application and Teammate.
- GAI-specific risks. Assess confabulation, information integrity, data privacy/memorization, information security (incl. prompt injection), harmful bias & homogenization, value-chain risk.
- Data/rights mapping. Source lineage, rights/permissions, and provenance practices are documented for any corpora used for retrieval, tuning, or evaluation.

2.3 MEASURE — tests, metrics, and monitoring

- TEVV (Test, Evaluation, Verification, and Validation). Pre-deployment and in-production testing for factual grounding, bias/disparities, robustness to attacks, explainability quality, and provenance efficacy; track confabulation rates and safety refusals.
- Feedback loops. In-product flags and analyst review drive prompt/tooling updates, dataset refinements, and control hardening over time.

2.4 MANAGE — controls, go/no-go, and continuous improvement

- Release gates & rollbacks. Risk-based go/no-go criteria; kill switches and rollback plans for high-risk features; escalation paths to leadership.
- Incident response. Containment, client notice where appropriate, root-cause analysis, and preventive actions; periodic exercises; update playbooks per lessons learned.

Data protection controls (privacy & security in practice)

- Collection & minimization. Only what is necessary to deliver features. Prefer client-side redaction or server-side tokenization. No unrelated mining or secondary use.
- Isolation & access. Tenant-level segregation; strict least-privilege access, auditable logs. Context data (prompts/retrieved passages/outputs) retained only as needed for delivery, support, and safety.
- Encryption & retention. Encryption in transit/at rest; short, contract-controlled retention for logs and chats; secure destruction on decommission.
- Provenance & transparency. Show sources and timestamps for AI-assisted outputs; label generated content.
- Bias & fairness. Evaluate performance across segments (languages, dialects, geographies, user cohorts); document limitations; use diverse reviewers and counter-prompt tests.

- Security for GAI (Generative Artificial Intelligence). Threat modeling; defense-in-depth against direct/indirect prompt injection; validation/sanitization of external content prior to model use; adversarial testing; detection of PII leakage; value-chain hardening.
- Third-party governance. Contracts cover "no training," IP/content ownership, incident SLAs, provenance expectations, and audit rights; maintain approved-provider list and monitor changes.
- Client controls. On request: disable transcript storage; enforce stricter pre-call redaction; custom retention.

Service-specific addendum

Teammates are AI-assisted modules that analyze domain-specific corpora (e.g., market, regulatory, trade, patents, supply chain, economy, and other future domains) and provide retrieval-augmented answers, summaries, tagging, and impact analysis for client workflows.

Disclosures. Teammates are AI-assisted research/analysis tools—not legal, investment, or other professional advice. Users should verify with cited primary sources.

Cross-Teammate controls (mapped to NIST AI 600-1 risks):

- Confabulation & information integrity. Retrieval-augmented generation from curated sources; answers require citations; confidence/recency cues; block or flag unsupported claims for human review.
- Data privacy & memorization. Client prompts/uploads remain within the client tenant; excluded from provider/model training by default.
- Information security. Hardened against prompt injection and tool misuse; adversarial testing and sandboxing; fallbacks and rollbacks defined.
- Harmful bias & homogenization. Uniform templates for impact narratives; diverse reviewer QA; measure and remediate disparities across domains and languages.

- Business Intelligence value-chain & component integration. Supplier risk assessment for datasets, models, and tools; contracts mandate content provenance and incident duties; inventory upstream dependencies and maintain traceability.
- Incident disclosure. Material misrepresentations or provenance failures are logged as AI incidents; we correct outputs and notify affected users as appropriate.

Program Operations & Continuous Improvement

- GOVERN — Policies, inventory, roles/training, third-party controls, and incident governance (AI RMF Table 1).
- MAP — For each application/Teammate, document purpose, context, data/rights/provenance, foreseeable misuse, and GAI-specific risks (confabulation, information integrity, privacy, security, harmful bias/homogenization, value-chain).
- MEASURE — TEVV plans for factuality, bias, robustness, explainability, and provenance efficacy; structured feedback and monitoring (AI RMF Table 3; AI 600-1 actions).
- MANAGE — Risk-based go/no-go, kill-switches, rollbacks, incident response & disclosure, and continuous improvement (AI RMF Table 4; AI 600-1 incident guidance).

Assurances & Standards Alignment

- Assurance that your content is private and not used to train models by default.
- Traceability from AI outputs to sources and decision context.
- Accountability via governance, measurement, and incident processes.
- Fitness-for-purpose through human oversight for consequential insights.
- Alignment to NIST guidance (AI RMF 1.0; GAI Profile).

Notes on references

- The following documents are directly relevant to this policy, and are referenced within this document:
- NIST AI RMF 1.0 (NIST AI 100-1): trustworthiness characteristics and the Core functions (GOVERN/MAP/MEASURE/MANAGE). DOI: 10.6028/NIST.AI.100-1.
- NIST AI 600-1 (Generative AI Profile): risk set (confabulation; information integrity; data privacy; information security; harmful bias & homogenization; value-chain, etc.) and suggested actions across governance, testing, and incident disclosure. DOI: 10.6028/NIST.AI.600-1.

Contact Us

www.industryintel.com

sales@industryintel.com