

AI Explainability Guide

Outliers Professionals — Data & AI Centre™ · Guide · Professional

Resource ID: res_ai_explainability_guide | Audience: Data/AI & business leaders

Purpose

This guide — AI Explainability Guide — establishes the standards, controls and operating expectations required to deliver trusted, governed and value-creating data and AI outcomes across the enterprise. It is part of the Outliers Professionals Data & AI Centre™ catalogue and aligns with NDPR/NDPA, ISO/IEC 27001, ISO/IEC 23894 (AI risk), ISO/IEC 42001 (AI management), NIST AI RMF, EU AI Act tiering principles, DAMA-DMBOK and CDMC.

Who Should Use It

Audience: Data/AI & business leaders. Typical users include the Chief Data Officer, Chief AI Officer, Head of Analytics, Data Governance Council, Data Stewards, Model Risk Officers, CISO, DPO, Internal Audit, Risk & Compliance, and accountable business owners of data domains and AI use cases.

When To Use It

Use at programme inception, during annual policy refresh, when on-boarding a new data domain, when proposing a new AI/ML use case, at every model lifecycle gate (intake → build → validation → deployment → monitoring → retirement), during regulatory or audit reviews, and when reporting to the Board Data & AI / Risk Committee.

Step-by-Step Usage Guide

1. Confirm scope and accountable owner.
2. Inventory existing artefacts, data domains, models, prompts and decisions in-scope.
3. Map applicable obligations (NDPA, sectoral regulators, internal policy).
4. Complete each working section in this document with enterprise-specific evidence.
5. Calibrate risk tiers, controls and approval gates using the Outliers Data & AI tiering ladder (Tier 1 informational → Tier 4 high-risk autonomous).
6. Route the draft through Data Governance Council → Model Risk → Legal/DPO → Internal Audit.
7. Obtain executive approval and version-control the final artefact in the Data & AI evidence room.
8. Schedule the next review (default: annual; quarterly for Tier 3/4 models).

Governance & Control Considerations

- Three-lines accountability: business owner (1LoD), Data Governance & Model Risk (2LoD), Internal Audit (3LoD).

- Data quality dimensions: accuracy, completeness, consistency, timeliness, validity, uniqueness — each rule logged in the Data Quality Rules Library.
- Model controls: documented purpose, training data lineage, validation, bias/fairness testing, explainability, human-in-the-loop, monitoring and drift thresholds.
- Privacy: NDPA lawful basis, DPIA where required, minimisation, retention, cross-border transfer assessment.
- Security: classification (Public / Internal / Confidential / Restricted), encryption in transit and at rest, access reviews, secrets management.
- Prompt & GenAI controls: prompt logging, sensitive-data redaction, jailbreak monitoring, vendor due-diligence, output review thresholds.

Review & Approval Workflow

Draft → Data Governance Council review → Model Risk validation (if AI) → DPO/Legal sign-off → Executive Data & AI Committee approval → Board Data & AI Committee noting. Sign-offs captured in the approval register with date, version and rationale. Material changes trigger re-approval; minor edits logged as point releases.

Implementation Notes

- Embed obligations into the SDLC / MLOps pipeline as policy-as-code checks where feasible.
- Pair every Tier 3/4 AI use case with a Model Card and a live Drift Log.
- Wire dashboards to the canonical KPI library: data quality index, model performance, drift incidents, privacy incidents, AI value-tracking.
- Provide role-specific enablement (executive, manager, practitioner, end-user) before go-live.
- Run an annual Data & AI maturity assessment to recalibrate priorities.

Sample Working Template

Model ID	Use Case	Risk Tier	Owner	Last Validation	Drift Status	Action
MDL-001	Credit scoring	Tier 4 — High	Risk Analytics	2026-04-12	Within bounds	Quarterly review
MDL-014	Customer churn	Tier 3	Marketing	2026-03-30	Yellow — PSI 0.18	Re-train
MDL-027	Fraud triage	Tier 4 — High	Fraud Ops	2026-05-08	Green	Continue

© *Outliers Professionals*. For internal use under the *Data & AI Centre™* licence. Customise to your enterprise before adoption.