Responsible AI Governance Checklist

Fairness · Explainability · Privacy · Security · Compliance

HOW TO USE THIS CHECKLIST

Complete this checklist before deploying any AI/ML model to production. Assign each item to an owner and record completion date. Items marked [BLOCK] must be resolved before go-live. Items marked [REVIEW] require documented justification if incomplete.

FAIRNESS & BIAS

- [ ] [BLOCK] Sensitive attributes (race, gender, age, etc.) identified and reviewed for proxy leakage.
- [ ] [BLOCK] Bias audit conducted across demographic subgroups using disaggregated metrics.
- [ ] [REVIEW] Fairness metric selected (demographic parity, equalized odds, calibration) and documented.
- [ ] [REVIEW] Human override mechanism exists for consequential model decisions.
- [ ] [REVIEW] Bias monitoring scheduled post-deployment with alerting thresholds.

EXPLAINABILITY & TRANSPARENCY

- [ ] [BLOCK] Model explainability method selected (SHAP, LIME, attention, rule extraction) per risk tier.
- [ ] [BLOCK] Model card / factsheet completed and stored in the model registry.
- [ ] [REVIEW] End-users informed that an AI system is influencing decisions affecting them.
- [ ] [REVIEW] Appeal or contestation mechanism provided where required by law or policy.
- [ ] [REVIEW] Training data provenance documented and accessible to audit.

PRIVACY & DATA PROTECTION

- [ ] [BLOCK] Data Privacy Impact Assessment (DPIA) completed for all PII/sensitive data.
- [ ] [BLOCK] PII is not present in model features unless legally justified and minimised.
- [ ] [BLOCK] Data retention and deletion schedule aligned with GDPR/CCPA/local law.
- [ ] [REVIEW] Consent or legitimate interest basis documented for all training data sources.
- [ ] [REVIEW] Right-to-be-forgotten process tested for models trained on personal data.

Responsible AI Governance Checklist

Security · Regulatory Compliance · Sign-off

SECURITY & ADVERSARIAL ROBUSTNESS

- [ ] [BLOCK] Model endpoints are authenticated and access-controlled (not open to public).
- [ ] [BLOCK] Adversarial input testing conducted (prompt injection, evasion attacks where applicable).
- [ ] [REVIEW] Model extraction / inversion attack surface assessed for sensitive models.
- [ ] [REVIEW] Rate limiting and anomaly detection applied to inference endpoints.
- [ ] [REVIEW] Third-party model/API vendor security assessment completed.

REGULATORY COMPLIANCE

- [ ] [BLOCK] Applicable regulations identified (EU AI Act, GDPR, CCPA, FCRA, HIPAA, FCA, etc.).
- [ ] [BLOCK] Model risk tier assigned (EU AI Act: unacceptable / high / limited / minimal).
- [ ] [REVIEW] Regulatory notification or pre-approval submitted where required.
- [ ] [REVIEW] Legal review of model outputs for defamation, IP, or liability exposure.
- [ ] [REVIEW] Compliance training completed by all team members working on the AI system.

OPERATIONAL GOVERNANCE

- [ ] [BLOCK] AI incident response playbook written and tested.
- [ ] [BLOCK] Kill switch / rollback procedure documented and tested in staging.
- [ ] [REVIEW] Post-deployment review scheduled (30/60/90 day checkpoints).
- [ ] [REVIEW] Model retirement criteria and sunset plan documented.
- [ ] [REVIEW] Governance checklist review cycle set (annual minimum).

SIGN-OFF

Role	Name	Date	Signature
Model Owner			
Data Engineering Lead			
Privacy / Legal			
Security			
Executive Sponsor			