AI Governance Framework Playbook

Templates and checklists for establishing an AI governance programme

LAWSEL ADVISORY

www.thelawsel.com

Table of Contents

CHAPTER 1

Establishing an AI Governance Committee

Committee Charter Template

Purpose: To provide oversight and strategic direction for the organisation's development, deployment, and use of artificial intelligence systems.

Scope: All AI systems developed internally, procured from third parties, or used in the organisation's products and services.

Composition:

- Executive Sponsor (C-suite) — accountable for AI governance
- Chief Technology Officer or VP Engineering — technical oversight
- General Counsel or Head of Legal — regulatory compliance
- Head of Product — commercial context and prioritisation
- Data Protection Officer — privacy implications
- Ethics Advisor (internal or external) — responsible AI considerations
- Business Unit Representatives — as needed for specific decisions

Meeting Cadence:

- Quarterly standing meetings (minimum)
- Ad hoc meetings triggered by: new high-risk AI deployments, significant incidents, material regulatory changes

Decision Authority:

- Approve deployment of high-risk AI systems
- Set AI risk appetite and tolerance levels
- Approve the AI governance policy and framework
- Escalate to the Board on matters exceeding delegated authority

Reporting:

- Quarterly report to the Board covering AI inventory, risk status, incidents, and regulatory developments
- Annual AI governance maturity assessment

CHAPTER 2

AI Risk Assessment Framework

Risk Categories

Assess each AI system against the following risk categories:

1. Harm to Individuals
   - Could the system's output cause physical, psychological, financial, or reputational harm to individuals?
   - Does the system make or influence decisions about individuals' access to services, employment, education, or rights?
2. Bias and Fairness
   - Could the system produce discriminatory outcomes based on protected characteristics?
   - Is the training data representative of the affected population?
   - Have bias testing and fairness metrics been defined?
3. Privacy
   - Does the system process personal data?
   - Does it involve profiling, tracking, or surveillance?
   - Is a DPIA required?
4. Security and Robustness
   - Is the system vulnerable to adversarial attacks or manipulation?
   - What happens if the system fails or produces incorrect outputs?
   - Are there adequate fallback mechanisms?
5. Transparency and Explainability
   - Can the system's decisions be explained to affected individuals?
   - Are users aware they are interacting with AI?
   - Is there adequate documentation of how the system works?
6. Accountability
   - Is there clear ownership and responsibility for the system?
   - Are there human oversight mechanisms?
   - Is there an audit trail?

Risk Scoring

For each risk category, score:

Likelihood: 1 (Rare) | 2 (Unlikely) | 3 (Possible) | 4 (Likely) | 5 (Almost Certain)
Impact: 1 (Negligible) | 2 (Minor) | 3 (Moderate) | 4 (Major) | 5 (Severe)

Risk Score = Likelihood × Impact

- 1–5: Low Risk — monitor, no specific controls required beyond standard practices

- 6–12: Medium Risk — implement targeted controls, review quarterly
- 13–19: High Risk — require AI Governance Committee approval, implement comprehensive controls
- 20–25: Critical Risk — escalate to Board, consider whether to proceed

CHAPTER 3

AI Model Documentation Template

Model Card

Complete for every AI model in production:

Model Overview:

- Model name and version
- Model owner (team and individual)
- Date of last update
- Intended purpose and use cases
- Out-of-scope uses (what the model should NOT be used for)

Training Data:

- Data sources and provenance
- Data collection methodology
- Data size and time period
- Known biases or limitations in the data
- Data preprocessing steps
- Licence terms and consent basis for training data

Performance:

- Evaluation metrics and results
- Performance across demographic groups (fairness metrics)
- Known failure modes
- Confidence thresholds

Limitations:

- Known limitations and biases
- Scenarios where the model performs poorly
- Environmental or contextual dependencies

Monitoring:

- Key performance indicators tracked in production
- Drift detection methodology
- Retraining triggers and schedule
- Alerting thresholds

Regulatory:

- Risk classification (EU AI Act category)
- Applicable regulations
- Compliance status

CHAPTER 4

Ethical AI Review Process

Review Triggers

An ethical AI review is required when:

- A new AI system is proposed that will affect individuals
- An existing system is being applied to a new use case
- A system's performance data reveals potential bias
- A complaint or incident raises ethical concerns
- The AI Governance Committee requests a review

Assessment Criteria

The review should assess:

Fairness:

- Does the system treat all demographic groups equitably?
- Have disparate impact metrics been measured?
- Are there mechanisms to detect and correct bias?

Transparency:

- Can affected individuals understand why a decision was made?
- Is there meaningful human oversight?
- Are limitations clearly communicated?

Accountability:

- Is there a clear chain of responsibility?
- Can decisions be audited and challenged?
- Is there a complaints and redress mechanism?

Safety:

- What happens when the system fails?
- Are there adequate safeguards?
- Have edge cases and adversarial scenarios been tested?

Privacy:

- Is data use proportionate and necessary?
- Are privacy rights protected?
- Is there meaningful consent where required?

Approval Workflow

- Requestor submits AI use case proposal and risk assessment
- Ethics lead conducts initial screening (within 5 business days)
- If low risk: ethics lead may approve with conditions
- If medium/high risk: full review by Ethics Review Panel
- Ethics Review Panel meets within 15 business days
- Panel issues: Approved / Approved with Conditions / Referred to AI Governance Committee / Rejected
- Conditions are documented and tracked to completion
- Approved systems are added to the AI inventory with review date

CHAPTER 5

AI Incident Response Plan

Incident Classification

Severity 1 — Critical:

- AI system causes or contributes to physical, financial, or significant reputational harm
- Systemic bias discovered affecting a large population
- Regulatory enforcement action

Severity 2 — Major:

- AI system produces incorrect outputs affecting business decisions
- Bias detected in a specific use case
- Customer complaint involving AI fairness or transparency

Severity 3 — Minor:

- Model performance degradation detected by monitoring
- Near-miss identified (potential harm avoided by safeguards)
- Internal feedback about AI concerns

Response Steps

Immediate (within 4 hours of detection):

- [ ] Alert the AI Governance lead and incident response team
- [ ] Assess severity and classify the incident
- [ ] If Severity 1: consider disabling or restricting the AI system immediately
- [ ] Preserve logs and evidence

Short-Term (within 48 hours):

- [ ] Conduct root cause analysis
- [ ] Determine scope and affected individuals
- [ ] Assess legal and regulatory notification obligations
- [ ] Implement interim controls
- [ ] Prepare internal and external communications

Medium-Term (within 2 weeks):

- [ ] Implement permanent fix
- [ ] Validate fix through testing
- [ ] Complete regulatory notifications if required
- [ ] Notify affected individuals if required
- [ ] Update AI risk assessment

Post-Incident:

- [ ] Conduct post-incident review within 30 days
- [ ] Document lessons learned

- Update governance framework and controls
- Report to AI Governance Committee and Board (for Severity 1–2)

CHAPTER 6

AI Vendor Due Diligence

Vendor Assessment Checklist

When procuring AI systems or services from third parties:

Model Transparency:

- Can the vendor explain how the AI system works at a level sufficient for your compliance needs?
- Will the vendor disclose training data sources and methodology?
- Does the vendor provide model cards or equivalent documentation?
- Can the vendor demonstrate bias testing results?

Compliance:

- Has the vendor classified the system under the EU AI Act (if applicable)?
- Can the vendor provide conformity documentation?
- How does the vendor handle regulatory changes?
- Does the vendor maintain an AI governance programme?

Data Practices:

- How does the vendor use your data? (training, improvement, analytics)
- Is your data segregated from other customers' data?
- What happens to your data if the contract ends?
- Where is data stored and processed?

Security:

- What security certifications does the vendor hold?
- How does the vendor protect against adversarial attacks on the AI system?
- What is the vendor's incident response process for AI-specific issues?

Performance:

- What SLAs does the vendor offer for AI performance (accuracy, latency)?
- How does the vendor monitor for model drift?
- What is the retraining and update schedule?
- How are model updates communicated and validated?

Ready to discuss your compliance needs?

Book a free 30-minute consultation to explore how LawSel Advisory can support your business.

BOOK A CONSULTATION

calendly.com/rini-thelawsel/30min

GET IN TOUCH

rini@thelawsel.com
www.thelawsel.com