# Artificial Intelligence Risk Management Policy. V1.0

## 1. Purpose and Objective

The purpose of this AI Risk Management Policy ("Policy") is to establish a structured and consistent approach for identifying, assessing, mitigating, and monitoring risks associated with the development, deployment, and maintenance of Artificial Intelligence (AI) systems. This Policy ensures that the organisation's AI activities remain compliant with legal and regulatory requirements, uphold ethical standards, and align with the organisation's broader strategic objectives and defined risk appetite. *(RM-1)*

## 2. Scope

**2.1 Applicability.** This Policy applies to all AI initiatives, projects, products, and services developed or utilised by the organisation, including but not limited to:

- Proprietary AI models built in-house.
- Third-party AI tools integrated into operational processes.
- Pilot or proof-of-concept AI deployments within any business unit.

**2.2 Impact-Based Scaling of Requirements.** To ensure proportionality, the organisation applies an impact-based approach to AI risk management. The level of oversight, documentation, and control requirements shall scale with the potential impact of the AI system, ensuring that governance activities are commensurate with possible harm or consequence.

- **Low-Impact Systems**: May follow streamlined documentation and governance pathways, provided minimum baseline requirements are met (e.g., registration, basic monitoring, responsible use adherence).
- **Moderate-Impact Systems**: Must undergo formal risk assessment, maintain documented risk register entries, and implement standard controls. Review and validation must be completed prior to production deployment.
- **High-Impact and Critical-Impact Systems**: Require enhanced governance measures, including:
    - Formal approval by the Risk Management Committee (RMC) prior to deployment
    - Detailed mitigation and control plans for all material risks
    - Extended validation (e.g., fairness testing, robustness evaluation, stakeholder impact review)
    - Ongoing quarterly reviews post-deployment
    - Escalation of residual risks to the AI Governance Committee for acceptance

Impact classification must be assigned during the initial risk assessment according to guidance published by the RMC, and periodically re-evaluated to reflect changes in scale, use, exposure, or operating environment.

**2.2 Exclusions.** Certain low-risk or experimental AI activities may be granted limited exceptions if they do not pose significant material, legal, ethical, or operational risks. All requests for exceptions must follow the procedures outlined in Section 13 of this Policy. *(RM-1)*

**2.3 Related Policies.** This Policy complements and should be read in conjunction with the AI Governance Policy and related policies including:

- AI Model and Data Lifecycle Management Policy (LC-1)
- AI Incident Management and Response Policy (IM-1)

## 3. Definitions

For the purposes of this Policy, the following definitions apply:

- **Risk**: The combination of the likelihood of an event and its potential consequences—financial, operational, reputational, or otherwise—that could adversely affect the organisation's objectives. For avoidance of doubt, this includes the potential for harm to a customer, employee or other stakeholder as a direct result of the organisations development or use of an AI system.
- **Risk Appetite**: The level and type of risk the organisation is willing to accept in pursuit of its strategic objectives.
- **Risk Tolerance**: The acceptable variance or thresholds around the organisation's risk appetite for specific categories of AI risk (e.g., legal, operational, reputational).
- **Risk Register**: A centralised repository (document or system) in which identified risks, their assessments, and their respective controls are recorded and maintained.
- **Control Owner**: An individual or team assigned responsibility for implementing and maintaining specific risk controls to ensure ongoing risk mitigation.

Additional AI-related terminology, if not defined here, shall align with definitions provided in the organisation's AI Governance Policy or corporate glossary.

## 4. Policy Statement

The organisation is committed to systematically identifying, assessing, and mitigating risks arising from AI systems at every stage of their lifecycle. All AI initiatives shall be subject to rigorous risk evaluation processes, aligned with the organisation's risk appetite (Section 4.1) and tolerance levels (Section 4.2). This Policy establishes clear accountabilities and controls to ensure AI-related risks are managed proactively, transparently, and in compliance with applicable regulations and internal guidelines. *(RM-1)*

## 5. Risk Appetite and Risk Tolerance

**5.1 Risk Appetite Statement.** The organisation recognises the transformative potential of AI but also acknowledges the inherent risks associated with advanced data-driven technologies. To balance innovation with responsible deployment, the organisation maintains a **moderate to low** risk appetite for AI-related initiatives. This means the organisation is open to adopting AI solutions that offer substantial operational or strategic benefits, provided that associated risks—such as bias, privacy breaches, regulatory non-compliance, or reputational harm—are clearly identified, assessed, and effectively controlled. *(RS-5, RO-1)*

**5.2 Risk Tolerance Levels.** Risk tolerance levels provide clarity on how much risk is acceptable within specific categories. For AI projects, the following guidelines apply:

- **Regulatory and Legal Compliance**: **Minimal Tolerance.** The organisation will not engage in AI activities that carry a high risk of violating applicable laws or regulations. *(RO-1)*
- **Data Privacy and Security**: **Low Tolerance.** The organisation emphasizes strong data governance practices. Use of personal or sensitive data in AI systems must follow stringent privacy and security controls. *(PR-1)*
- **Ethical and Reputational Impact**: **Low to Moderate Tolerance.** AI systems that could result in significant ethical dilemmas or reputational damage must undergo thorough review and require approval from the Risk Management Committee. *(RS-5, CO-2)*
- **Operational Disruption**: **Moderate Tolerance.** The organisation allows a reasonable level of experimentation with AI solutions, but projects that present a high risk of severe operational disruption are subject to heightened controls and monitoring. *(OM-1)*

- **Unanticipated Costs: Moderate to High Tolerance.** The Organisation accepts a reasonable level of unforeseen or non-recoverable costs associated with AI experimentation, including pilot initiatives that may not result in immediate business value. This tolerance supports innovation, provided such costs are pre-approved within defined budgets and do not create material financial exposure.

**5.3 Ongoing Assessment of Appetite and Tolerance.** The Risk Management Committee (RMC) shall review the stated risk appetite and tolerance levels at least annually or in response to major organisational or regulatory changes. Revisions to these thresholds require formal endorsement by Executive Leadership.

## 6. Risk Management Framework

This Policy implements a structured and proactive framework for managing AI risk, covering the full lifecycle of AI risk activities—risk categorisation, identification, assessment, and control selection. The framework ensures that all AI-related risks are surfaced early, assessed consistently, treated appropriately, and monitored effectively throughout the AI System lifecycle. The core components of this framework are described below, supported by additional guidance published as required by the RMC.

**6.1 Risk Identification.** AI risks should be identified systematically, beginning as early as the ideation or proposal stage of a project and continuing through development, deployment, and operations. Teams are expected to inquire beyond generic checklists to uncover AI-specific risk types, including technical, ethical, operational, and strategic risks. *(RM-2, IM-1)*

Recommended methods for risk identification include:

- **Pre-mortem analysis**: Envisioning plausible future failures and working backward to identify contributing factors.
- **Incident pattern mining**: Learning from historical AI failures and external incident databases.
- **Horizon scanning**: Identifying risks across short, medium, and long-term timeframes.
- **Red teaming**: Simulating adversarial attacks or misuse scenarios.
- **Dependency chain analysis**: Mapping out upstream and downstream components that influence the AI system's stability or integrity.

At least one structured technique (e.g. a pre-mortem or red-teaming workshop) must be applied for each major AI project. Identified risks shall be recorded in the AI Risk Register along with their source, description, and associated context. Risks must be revisited periodically and updated following incidents or system changes.

**6.2 Risk Assessment.** Each identified risk must be formally assessed according to guidance published by the RMC. This includes four key elements:

- **Likelihood**: How probable it is that the risk will occur, scored on a five-point qualitative scale.
- **Impact severity**: The potential harm or consequence of the risk materialising, considering legal, operational, financial, or reputational damage.
- **Impact velocity**: How quickly the risk could unfold and affect the organisation, if triggered.
- **Dynamic feedback effects**: Whether the risk could self-amplify or create cascading failures, such as model feedback loops or emergent behaviours.

A 5x5 Risk Matrix is used to combine likelihood and impact. Risks with fast velocity or feedback potential must be prioritised even if their raw risk score appears moderate. These dynamic factors are particularly relevant to AI due to the self-adjusting nature of many systems and their potential for non-linear effects.

The overall risk classification (Very Low, Low, Medium, High, Critical) will determine the required treatment pathway and level of governance oversight. *(RM-2, RM-3)*

**6.3 Risk Control and Mitigation.** For all risks rated Medium or higher, appropriate control measures must be selected. Control options and selection criteria are described in guidance developed by the organisation's AI Risk Management Committee. Controls should be proportionate to the risk's assessed level and aligned with the organisation's stated risk appetite and tolerance thresholds.

Examples of standard control types include:

- **Technical controls**: Fairness testing, drift detection, model explainability enhancements, rate-limiting, or fail-safes.
- **Process controls**: Human-in-the-loop mechanisms, approval gates, governance reviews.
- **Organisational controls**: Role-based access, escalation procedures, and training.
- **Third-party risk controls**: SLAs and contractual safeguards where AI systems are sourced externally.

Each control must be assigned to a Control Owner who is responsible for implementation and for updating the risk's treatment status. Where mitigation is not feasible, risks may be escalated for acceptance or avoidance decisions in accordance with defined approval thresholds.

Residual risk—after controls have been applied—must be reassessed to confirm whether it falls within acceptable tolerance with the application of the selected controls. *(RM-3, RS-5)*

**6.4 Risk Monitoring.** AI systems must be monitored on an ongoing basis for any indicators that an identified risk is emerging or that existing controls are no longer effective. This includes monitoring:

- Input data characteristics (e.g. distribution shifts, data quality degradation)
- Model performance metrics (e.g. accuracy, fairness, precision/recall)
- System behaviour (e.g. anomalies, unexpected user outputs)
- User feedback, complaints, or error reports

Monitoring procedures must be documented for each AI system and aligned with the controls outlined in Section 6.3. Where possible, automated monitoring tools should be implemented. Thresholds and triggers must be established for key indicators, and any breaches must be escalated in accordance with the organisation's Incident Management procedures. *(RM-4, IM-1, IM-2)*

**6.5 Continuous Improvement.** The risk management framework must evolve with changing technologies, use cases, and regulatory expectations. The Risk Management Committee shall review emerging AI risk types and control approaches on a regular basis. *(RM-4, GL-1)* Lessons learned from internal incidents, external case studies, audits, and stakeholder feedback shall inform periodic revisions to:

- The risk taxonomy and identification methods
- The scoring and prioritisation model
- The catalogue of control strategies and treatment guidelines

## 7. Roles and Responsibilities

AI risk management is a shared organisational responsibility. While the Risk Management Committee (RMC) provides central oversight, all employees and teams involved in the development, deployment, oversight, or procurement of AI systems are responsible for identifying, escalating, and mitigating risks in accordance with this Policy. Each role below contributes to the implementation and assurance of effective AI risk management:

**7.1 Risk Management Committee (RMC).** The RMC, functioning under the AI Governance Policy (GL-1), is accountable for operationalising and maintaining the Organisation's AI risk management framework. The RMC shall:

- Conduct regular reviews of the AI Risk Register to ensure alignment with the Organisation's stated risk appetite and tolerance thresholds.
- Evaluate risks escalated due to severity, novelty, or systemic impact, and determine the appropriate treatment or acceptance pathway.
- Recommend updates to risk controls, taxonomies, and risk assessment methodologies.
- Report significant findings, trends, or control weaknesses to Executive Leadership, ensuring alignment with broader strategic and compliance objectives.

**7.2 AI Governance Committee.** The AI Governance Committee provides executive oversight for risk management in high-impact AI systems. The Committee shall:

- Approve or reject risk acceptance decisions for critical or novel AI use cases.
- Review critical incident reports and lessons learned to verify that systemic risks are appropriately remediated.
- Ensure alignment between organisational risk posture and AI deployment strategies.

**7.3 System-Level Risk Roles.** The following roles have responsibilities for the effective execution of this Policy:

- **Project Leaders**: Responsible for the initial identification of AI-related risks during planning and development. Ensure all relevant risks are documented, assessed, and updated throughout the project lifecycle. Coordinate closely with the RMC on material risk changes.
- **Control Owners**: Accountable for the design, implementation, and maintenance of assigned risk controls. Monitor control effectiveness, document any limitations, and report deviations or failures to the RMC.
- **AI Governance Lead**: Leads investigation and resolution of risks involving system behaviour, model drift, performance degradation, or unintended outcomes. Supports root cause analysis for risk-related incidents.
- **Legal and Compliance**: Monitors AI initiatives for legal and regulatory compliance. Advises on emerging legal risks, reviews third-party AI engagements, and recommends control enhancements to reduce exposure.
- **Data Protection Officer (DPO):** Evaluates risks related to the processing of personal or sensitive data by AI systems. Ensures compliance with applicable privacy laws and supports the design of privacy-preserving controls.
- **Internal Audit**: Independently assesses the effectiveness of AI risk management practices. Conducts audits of risk documentation, escalation processes, and control implementations. Reports findings to the AI Governance Committee and recommends improvements.

## 8. Implementation, Monitoring, and Continuous Improvement

**8.1 Policy Distribution and Adoption.** This Policy shall be distributed to all relevant stakeholders, including AI project teams, data scientists, product managers, compliance officers, and other personnel involved in AI initiatives. All AI systems must comply with this Policy within three (3) months of the effective date or within a timeframe determined by the Risk Management Committee (RMC).

**8.2 Ongoing Monitoring.** The RMC shall monitor the implementation and operational effectiveness of this Policy using defined performance indicators, including:

- Frequency and severity of risk escalations
- Resolution timeframes for identified risks
- Adherence to approved risk appetite and tolerance thresholds
- Identified trends, recurring issues, or control failures shall trigger further review and remediation actions.

**8.3 Audit and Compliance.** An annual audit of AI risk management practices shall be conducted by Internal Audit or an appointed external auditor. The audit shall assess compliance with this Policy; the adequacy and effectiveness of controls; and completeness and accuracy of the AI Risk Register. Findings shall be reported to the AI Governance Committee and used to inform updates to risk controls and guidance.

## 9. Training and Awareness

**9.1 Mandatory Training.** Employees, contractors, and relevant stakeholders involved in AI initiatives are required to complete training on this Policy and associated relevant guidance *(RO-2)*. The organisation's Learning and Development team, in coordination with the RMC, shall develop and deliver training materials focusing on:

- Understanding the organisation's risk appetite and tolerance.
- Recognising and reporting potential AI risks.
- Applying risk controls and adhering to escalation protocols.

**9.2 Recordkeeping and Training Compliance Monitoring**. Completion of training shall be recorded and monitored by the Learning and Development team. Non-completion of mandatory training may result in restricted access to AI development tools, datasets, or production environments until compliance is achieved.

## 10. Exceptions and Exemptions

All requests for exceptions to this Policy must be documented and submitted to the RMC. Each request must include a clear justification of the business need, an analysis of potential risks, alignment with the organisation's risk appetite, and any compensating controls to be implemented. The RMC may grant temporary or conditional exemptions at its discretion if the overall risk remains acceptable.

## 11. Policy Review

**11.1 Policy Maintenance and Review.** The RMC is responsible for maintaining and reviewing this Policy to ensure it remains aligned with legal, regulatory, technological, and organisational developments. This Policy shall be reviewed at least annually or earlier if prompted by significant internal or external changes.

**11.2 Substantive revisions**. Substantive revisions to this Policy, including changes to the risk appetite statement, risk classification thresholds, or role assignments, must be:

- Endorsed by the RMC
- Approved by Executive Leadership
- Communicated to all affected stakeholders

*Approved by:*

---

[Name], Chief Technology Officer

*Date:*
[Date]

---