Responsible Artificial Intelligence (AI) Procedure

Section 1 - Context

(1) RMIT endorses a considered and ethical use of Artificial Intelligence (AI), acknowledging that it is increasingly part of the future of life and work and presents both opportunities to benefit and potential to harm individuals, communities, and the environment.

(2) RMIT's endorsed AI assets will not harm human beings nor pose unreasonable safety risks. Adoption of RMIT's endorsed AI assets should be appropriate and proportional to achieve given legitimate aims, balancing benefits and risks.

(3) This procedure establishes the ethical principles and risk-based framework which underpin the safe and responsible adoption of AI to support the functions and activities of the RMIT Group.

Section 2 - Authority

(4) Authority for this document is established by the Information Governance Policy.

Section 3 - Scope

(5) This procedure applies to all individuals who use Artificial Intelligence solutions (AI assets) to procure, create, manage, handle, use or process RMIT Group Information, including RMIT Group staff, casual employees, contractors, visitors and honorary appointees. This also includes third parties (suppliers) and agents of the organisation who are bound to RMIT policy where their contract of engagement with RMIT specifically provides for this.

(6) AI assets managed in compliance the Research Policy are outside the scope of this procedure.

(7) The RMIT Group, which includes RMIT University and its controlled entities, will be referred to as RMIT hereafter.

Section 4 - Procedure

Principles

(8) Human, societal, and environmental wellbeing: AI assets should benefit students, staff, researchers, alumni, industry, the environment, and the broader communities in which RMIT operates according to our mission.

(9) Human-centred values: An AI asset must operate in accordance with human rights and respect and preserve the dignity of individuals and groups.

(10) Fairness: The benefits of AI should be accessible to all students, staff, researchers and broader communities in which RMIT operates. AI assets should promote fairness, non-discrimination, and diversity.

This document may be varied, withdrawn or replaced at any time. Printed copies, or part thereof, are regarded as unauthorised and should not be relied upon as the current version. It is the responsibility of the reader to refer to RMIT's Policy Register for the latest version of this document.

(11) Transparency: RMIT, including all staff and students, are responsible for providing transparency and reasonable disclosure in the use of AI. It should be clear when and how AI is used in learning, teaching, research, business and operational activities.

(12) Accountability: RMIT staff responsible for the different phases of the AI asset lifecycle should be identifiable and accountable for the outcomes of the AI assets, and human oversight and control of AI assets should be maintained.

(13) Contestability: RMIT will ensure timely and just processes exist to challenge the use, consequences and outcomes of an AI asset when it significantly impacts a person, community, group or the environment.

(14) Privacy and Security: RMIT must ensure that any AI asset used by the institution upholds the privacy rights and data protection of AI Consumers and ensure the security of data in keeping with policy, legislation of the region and requirements of a public institution.

(15) Reliability: AI assets endorsed by RMIT that are built or under licence agreements should be tested to operate reliably in accordance with their intended purpose.

(16) Education and Literacy: AI assets endorsed by RMIT should be accompanied by relevant education on their intended purpose, operations, outcomes and risks to facilitate discussion between educators, researchers, students, and professional staff.

Existing Governance Controls

(17) The AI Governance Framework will be operationalized through existing governance controls, and where needed, further assessment may be undertaken for the context of AI.

(18) Existing risks assessments, indicated below, must be completed before the AI Governance Framework is initiated. These assessments will also determine if the initiative is introducing an AI component(s):

a. Privacy Impact Assessment(PIA)
b. Security Risk Assessment (SRA)
c. Third-Party Risk Assessment (TPRA).

AI Governance Framework

Enterprise Catalogue

(19) Once the above governance assessments are completed and the initiative has been determined to introduce AI will be assessed for entry into the RMIT Data and AI Catalogue.

(20) The RMIT Data and AI Catalogue is used to catalogue AI assets, for visibility, oversight and accountability across the lifecycle of AI assets.

(21) An initiative which introduces AI will qualify for entry into the RMIT Data and AI Catalogue if RMIT can modify its AI parameters.

(22) The modification of parameters includes activities that RMIT may perform to influence the behaviour and outputs of the AI/Machine Learning (ML) model, including model training, hyperparameter tuning, Large Language Model (LLM) fine tuning, modifying LLM system prompts and LLM selection.

(23) Initiatives which introduce AI which RMIT is unable to modify the parameters of will be assessed and managed according to the existing Risk Management Policy and related procedures to determine if the risk is within one or more Domain Risk Appetite Statements.

(24) Initiatives which introduce AI which are assessed as being able to have their parameters modified by RMIT, will be determined to be AI Assets and added into the RMIT Data and AI Catalogue.

Transparency Monitoring

(25) The Risk Exposure Tool will be used to assess the level of risk AI assets within the enterprise catalogue present to RMIT.

(26) If an AI asset poses a risk, it is categorised as a tier 2 risk within the risk profile of the relevant college, portfolio, or controlled entity as per Risk Management Policy. Some AI assets may be subject to transparency reporting.

(27) The purpose of a transparency report is to:

a. help the AI Consumers and stakeholders understand how an AI asset behaves, what it does, its intended application, and precautionary actions required as part of its adoption
b. enable compliance with RMIT policies, including the Privacy Policy.

Risk-based Oversight and Reporting

(28) Risks associated with AI assets will be reported to governance committees (i.e. Council, Audit and Risk Management Committee, Academic Board etc.) in accordance with the Risk Management Policy.

Roles and Responsibilities

(29) An individual or group assumes the role of AI Sponsor when they initiate, fund, and/or oversee the implementation and adoption of AI assets. AI Sponsors are responsible for:

a. the realisation of value and benefits of AI assets, in alignment to RMIT strategy
b. ownership and management of risks of AI assets and ensuring that, if necessary, the AI asset is catalogued.
c. independently overseeing and ensuring compliance to RMIT policies, including the Risk Management Policy, Information Technology and Security Policy, and Information Governance Policy.

(30) An individual, group or third party assumes the role of AI Developer when they create, design, build, test, and/or deploy an AI asset. AI Developers are responsible:

a. for selecting the appropriate architecture, data, algorithms and ensuring the quality, performance, and robustness of the AI asset.
b. creating entries in the catalogue at the direction of the AI Sponsor and AI Assurer.

(31) An individual, group or third party assumes the role of AI Operator when they monitor and evaluate or verify an AI asset and its outcomes. AI Operators are responsible for:

a. monitoring, maintenance, and ongoing compliance of AI assets
b. creating entries in the catalogue at the direction of the AI Sponsor and AI Assurer.

(32) An individual assumes the role of an AI Consumer when they interact, use or benefit from the AI assets or AI generated outputs. AI Consumers are responsible for:

a. using the AI asset in a responsible and ethical manner
b. providing feedback and reporting any issues or concerns.

(33) An individual member of the Data and AI Risk Management Stewards Working Group (AISG) is assigned to act in

the role of AI Assurer and must be independent from that of any AI Sponsors, AI Developers and AI Operators. They are responsible for:

a. independently reviewing AI assets and AI risks to enable risk management processes
b. reviews the level of risk and informs the AI Sponsor
c. independently ensuring the roles and responsibilities defined in this procedure are followed.

Section 5 - Definitions

(Note: Commonly defined terms are in the RMIT Policy Glossary. Any defined terms below are specific to this policy).

Artificial Intelligence (AI)	The application of advanced computational methods such as machine learning, deep learning and neural networks to produce outputs including large language models which can perform complex, human-like tasks. Refer to the definition for AI.
AI asset	A technology, application, software or model that performs AI automated, machine learning, generative, or combined actions for which RMIT Group can modify the parameters for specific processes and functions to achieve a defined purpose.
Modifying the parameters of the AI	Include activities that RMIT may perform to influence the behaviour and outputs of the AI/Machine Learning (ML) model, including model training, hyperparameter tuning, Large Language Model (LLM) fine tuning, modifying LLM system prompts and LLM selection.

Status and Details

Status	Current
Effective Date	31st July 2024
Review Date	14th March 2028
Approval Authority	Chief Operating Officer
Approval Date	29th July 2024
Expiry Date	Not Applicable
Policy Owner	Fiona Notley Chief Operating Officer
Policy Author	Nicola Burgess Interim Executive Director, Technology
Enquiries Contact	Data Management and Governance +61 3 9925 8000

This document may be varied, withdrawn or replaced at any time. Printed copies, or part thereof, are regarded as unauthorised and should not be relied upon as the current version. It is the responsibility of the reader to refer to RMIT's Policy Register for the latest version of this document.