# Responsible AI Adoption Policy

# Section 1 - Key Information

| | |
|---|---|
| **Policy Type and Approval Body** | Administrative – Vice-Chancellor |
| **Accountable Executive – Policy** | Vice-Chancellor |
| **Responsible Manager – Policy** | Pro Vice-Chancellor (AI) and Chief AI Officer |
| **Review Date** | 14 November 2027 |

# Section 2 - Purpose

(1) This Policy outlines the principles for using, developing, and managing Artificial Intelligence (AI) solutions at La Trobe University. The goal is to ensure that AI is used in alignment to our strategy and commitment to responsible AI implementation and adoption.

(2) AI is a powerful technology that can benefit the University and its community. However, it's important to use AI in a way that is safe, ethical, and complies with data privacy and protection laws. The University's approach to AI demonstrates our Cultural Qualities:

a. Accountable: we apply clear lines of responsibility and instill transparency in our AI use.
b. Connected: we engage across disciplines and functions, using cross-University collaboration to identify, manage and monitor issues and risks.
c. Innovative: we apply a risk-based approach, classifying AI according to its level of risk and applying corresponding levels of oversight and controls to mitigate that risk. We use stakeholder feedback to inform how AI may be enable the right outcomes and the risks it may pose.
d. Care: we actively consider environmental sustainability and the impact of unwanted bias, data security and potential benefits in the use of AI. Although they may be supported by AI applications, significant decisions will be made by humans to promote a safe and inclusive community.

# Section 3 - Scope

(3) This Policy applies to all University campuses, staff, students, contractors, honorary and adjunct staff, and all administrative activities under the control of the University.

(4) Whilst the Policy applies to the processes supporting research activities, it does not apply to the Research or its outcomes. Research operates under the University's Research Governance Policy, Research Governance Framework, specific standards, contractual requirements, and the Australian Code for the Responsible Conduct of Research (2018).

# Section 4 - Key Decisions

| Key Decisions | Role |
|---|---|
| Approve the implementation of medium and high risk AI use cases | Responsible AI Adoption Committee (RAAC) |
| Approve the implementation of low risk AI use cases | Chief Information Officer |

# Section 5 - Policy Statement

(5) The University is committed to using AI in a responsible manner that accords with Australia's AI Ethics Principles, the voluntary framework published by the Australian Government's Department of Industry, Science and Resources:

a. Human, social and environmental wellbeing: AI is only used where there are beneficial outcomes for university stakeholders, the environment and/or the Community. Assessment of these outcomes should include both positive and negative impacts, including those occurring externally to the University.
b. Human-centred values: La Trobe's AI aligns to human values and human rights risks are considered. The University will not support AI uses that breach Human Rights.
c. Transparency and explainability: La Trobe's AI operates with transparency and considers the expectations of data subjects and individuals affected by its use. The AI and its related tools and algorithms are explained in low or non-technical terms to promote accessibility and understanding. Where University systems involve humans interacting with AI, disclosure will be made to ensure user awareness.
d. Fairness: unfair bias and discrimination are proactively and regularly evaluated and eliminated in AI datasets, hypotheses, and algorithms. This is achieved through using safeguards, deliberate considerations and quality data.
e. Reliability and safety: AI operates within a system of robust data protection and cybersecurity controls, including a risk assessment and risk mitigation plan.
f. Privacy protection and security: Data used in AI development is managed in accordance with the University's Data Governance and Information Security Policies. Privacy risk is actively considered as part of the University's risk assessment process.
g. Contestability: Uses of AI that impact staff, students, or other individuals must be clearly identified and include a mechanism for review or inquiry to challenge the system outcome.
h. Accountability: fit for purpose accountability mechanisms are embedded in decision-making processes, including clear lines of responsibility, transparent decision logic and avenues for redress in the event of adverse outcomes. Where decisions are made based on AI, they must be subject to human intervention through nominated roles that ensure accountability for oversight. Such decisions must be able to be overturned by human experts, within a control framework.

(6) Staff, students, contractors, honorary and adjunct staff must not input, upload, transmit, or otherwise disclose any University information into publicly accessible or open-source artificial intelligence (AI) systems, University information includes personal, health and sensitive information pertaining to staff, students and other clients and stakeholders to the University, research data, commercially sensitive information including contracts, partnerships, funding arrangements, intellectual property, internal communications, legal advice, governance documents and operational data.

(7) An AI Adoption Risk Assessment must be completed for:

a. the procurement/acquisition of all new AI Systems/Tools; and
b. existing University approved AI Systems/Tools for which new use cases are proposed.

(8) The risk of a particular AI tool will be assessed having regard to its intended purpose and how it may be used in practice (regardless of the intended purpose).

(9) The University's AI governance will adopt a risk-based approach to ensure that obligations and oversight are proportionate to the risk posed. AI uses and systems will be classified into risk categories, based on their potential to violate individual's rights:

| Category | Description | General examples |
|---|---|---|
| **Low** | Presents a limited risk to individuals; AND<br>Does not involve sensitive, personal or confidential information | Spam filters, video games, using generative AI applications for general tasks involving no sensitive, personal or confidential University data |
| **Medium** | Applications with limited transparency, such that individuals may not realise they are engaging with AI; AND<br>Does not involve sensitive, personal or confidential information | Chat bots, generated images (picture, voice, video) |
| **High** | Could have a detrimental impact on health and safety or affect access to a fundamental right such as education, employment or justice; OR<br>Uses sensitive, personal or confidential information and the use complies with relevant University policies | AI assessment and shortlisting of job candidates, systems to evaluate learning outcomes, triage applications for health and welfare services |
| **Unacceptable – will not be pursued** | Techniques that may cause significant harm; OR<br>Uses sensitive, personal or confidential information in a way that does not comply with University policies | Untargeted scraping of facial images from CCTV footage, behavioural manipulation that causes harm |

(10) The University supports the proposed voluntary guardrails outlined in the Voluntary AI Safety Standard. We will ensure that our AI systems are developed and used responsibly, with a focus on accountability, risk management, data governance and human oversight.

(11) The University's Code of Conduct (the Code) outlines the ethical, professional and legal standards used as the basis of decisions and actions. Consistent with the Code, members of the University Community are individually accountable for their actions, including their own use of AI. Individuals must ensure they adhere to University Policies, including the Information Security Policy, Code of Conduct, Data Governance Policy, Privacy Policy and Records Management Policy and University issued directions regarding the use of AI.

(12) AI must only be used for its approved purpose. Any use of an AI system outside of its approved purpose must be assessed separately in accordance with this Policy.

(13) Ensuring trust is a foundational aspect of La Trobe's AI approach, recognising that:

a. Potential uses of AI may not be considered appropriate where there is a high inherent risk from its application. La Trobe recognises that a balance is required between opportunity and risk;
b. Algorithmic bias may result in erroneous or unjustified differential treatment which could have unintended or serious consequences for the environment, groups of individuals and/or for their human rights;
c. AI is a dynamic area of technology and is subject to increasing levels of regulation. As a result, this Policy and associated Procedure must be reviewed and updated to ensure it aligns to current and emerging regulatory standards and government advice;
d. As a public body, the University has responsibilities to the community that are outlined in its founding legislation. Continued trust in its performance underpins the ability to deliver its mandate.

(14) The University's adoption, ethical application and implementation is governed through its Responsible AI Adoption

Committee (RAAC). The RAAC is responsible for ensuring AI projects and initiatives align with University strategy, promote good practices and that risks are managed to deliver value. The RAAC reports into the Senior Executive Group (SEG) and endorsement from the RAAC is required for all new AI Use Cases. The RAAC delegates endorsement for low-risk use cases to the Chief Information Officer to support operational requirements.

(15) All uses of AI will have a nominated business owner. The RAAC will provide a report of approved AI systems and applications to the Senior Executive Group annually.

(16) Education, including building AI awareness, literacy and user capability through training, tools, and guidelines is essential to ensure principles and requirements are applied systemically. AI Business owners must ensure that roles responsible for implementing or managing AI have sufficient expertise or receive appropriate training to enable them to apply this Policy.

# Section 6 - Procedures

(17) All Institutional Data is managed within the University's Data Governance Policy and Framework (under development). Data and its use are subject to legislative requirements as outlined in the Framework. In addition to these requirements, the use of AI brings ethical considerations, particularly where it is to be used in making judgments or decisions that involve humans. The University outlines the expectations of its staff in the La Trobe Code of Conduct.

(18) All requests to use AI with existing datasets or to migrate additional data sets into data repositories should be initiated using the online enquiry function accessible through the intranet.

(19) Requests to implement technology that incorporates AI capabilities should be directed to the IS team through the Ask ICT function.

(20) Review through the RAAC forms an input to the University approval processes. The Responsible AI Adoption Ethics Committee (RAAEC) will review and provide recommendations to the RAAC for any tools rated medium or high level that require RAAC approval. The AI Business owner is responsible for submitting requests to use AI to the Responsible AI team.

(21) An AI Adoption Risk Assessment must be completed for all new uses of AI other than those classified as low risk. Depending on the data sets used and risk aspects of the usage, a Privacy Impact Assessment may also be required. This process also includes proposals to provide University data to a third-party for use in an AI application.

(22) A central record of AI applications, the RAAC assessment, and the ethics review for high-risk applications is to be maintained by the Information Services (IS) Division through their AI Accelerator function. Risk will be managed in accordance with the University's Risk Management Framework, including identifying appropriate controls. Potential controls for higher-risk applications include evaluation of algorithms and data sets for bias and accuracy, and auditing AI activities and outcomes. All high-risk applications will be reviewed annually for continued alignment with the principles.Uses that can be demonstrated as still required and compliant will be retained, with the remaining uses disassembled and terminated.

(23) Any complaints, concerns or risks relating to the use of AI at La Trobe should be reported to through the relevant channels operating for that activity, for example academic misconduct or research integrity. Data breaches should be reported using the University's data breach process.

# Section 7 - Definitions

(24) For the purpose of this policy and procedure:

a. Artificial Intelligence System / Application: a machine-based system or application that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment. (Source: OECD 2023) see the OECD Explanatory memorandum on the updated OECD definition of an AI system | OECD Artificial Intelligence Papers | OECD iLibrary (oecd-ilibrary.org) for explanatory material about the definition.
b. AI Business Owner: role accountable for a specified AI application. This role is responsible for ensuring appropriate implementation in accordance with RAAC approval, including appropriate training and support for users, monitoring system use and compliance with ethical requirements.
c. Responsible AI: an approach to developing and deploying artificial intelligence from both an ethical and legal standpoint (source: International Standards Organisation).

# Section 8 - Authority and Associated Information

(25) This Policy is made under the La Trobe University Act 2009.

(26) Associated information includes:

a. AI Adoption Risk Assessment
b. Responsible AI at La Trobe
c. Privacy by Design: Effective Privacy Management in the Victorian public sector
d. Privacy Policy
e. Data Governance Policy
f. Information Security Policy
g. Code of Conduct
h. Records Management Policy
i. About Data and Analytics - Intranet

## Status and Details

| | |
|---|---|
| **Status** | Current |
| **Effective Date** | 23rd December 2025 |
| **Review Date** | 14th November 2027 |
| **Approval Authority** | Vice-Chancellor |
| **Approval Date** | 23rd December 2025 |
| **Expiry Date** | Not Applicable |
| **Responsible Manager - Policy** | Phil Laufenberg<br>Pro Vice-Chancellor (AI) and Chief AI Officer |
| **Enquiries Contact** | Office of the Vice-Chancellor<br>+61 3 9479 2000 |

## Glossary Terms and Definitions

**"student"** - Student is defined in the La Trobe University Act 2009 as: (a) a person enrolled at the University in a course leading to a degree or other award; or (b) a person who is designated as a student or is of a class of persons designated as students by the Council.

**"staff"** - Staff means any person employed by the University as per the definition in the La Trobe University Act 2009 (Vic).