



Department of
the Premier and Cabinet

Artificial Intelligence Policy

Western Australian Government

Produced and published by:
Department of the Premier and Cabinet
Office of Digital Government

Published March 2024

Contact:
Office of Digital Government
Dumas House,
2 Havelock Street West Perth WA 6005
Email: ai-dgov@dpc.wa.gov.au

This policy is based on content developed by:

- NSW Government. This content is © State of New South Wales.
For current information go to www.nsw.gov.au .
- Australian Government Department of Industry, Science and Resources.
For current information go to <https://www.industry.gov.au/>

This document, the Western Australian Government Artificial Intelligence Policy, Version 1 is licensed under a Creative Commons Attribution 4.0 International Licence. You are free to re-use the work under that licence, on the condition that you attribute the Government of Western Australia (Office of the Government Chief Information Officer) as author, indicate if changes were made, and comply with the other licence terms. The licence does not apply to any branding or images.

License: creativecommons.org/licenses/by/4.0/legalcode

Attribution: © Government of Western Australia (Office of Digital Government) 2024

Notice Identifying Other Material and/or Rights in this Publication: The Creative Commons licence does not apply to the Government of Western Australia Coat of Arms. Permission to reuse the Coat of Arms can be obtained from the Department of the Premier and Cabinet:
Email commonbadge@dpc.wa.gov.au

Contents

Background	4
Purpose	4
Scope	4
Objectives.....	4
Defining artificial intelligence	4
Policy requirements	5
Human, social and environmental wellbeing	6
Human-centred values.....	6
Fairness.....	6
Privacy protection and security	7
Reliability and safety	7
Transparency, explainability and contestability	7
Accountability.....	8
Accuracy	8
Relevant Legislation, Policies and Other Documents	8



Background

Artificial Intelligence (AI) is a rapidly evolving field with the potential to provide great benefit to WA Government organisations. To create an enabling environment in which benefits can be maximised, the risks associated with AI must be mitigated and considered in all AI usage.

Purpose

The WA Government AI Policy (the Policy) establishes general principles WA Government organisations must apply to develop and use Artificial Intelligence (AI) tools ethically, safely and responsibly. The Policy aims to support agencies to manage the benefits and risks associated with AI systems and applications. The Policy will be updated as the technologies evolve and as a better understanding of benefits and risks develops.

Scope

The Policy applies to WA Public Sector Agencies.

This policy is applicable to the development and usage of all AI solutions, including traditional AI systems and generative AI technologies.

Objectives

- Mitigate risks associated with AI: Provide WA Government employees with robust principles to identify and address potential risks and safeguard against possible harm.
- Enable safe usage of AI: Through considered risk mitigation, an environment that fosters AI enablement can be established.

Defining artificial intelligence

There are numerous different definitions of AI, encompassing a range of technologies as they evolve. This policy adopts the following broad definition of AI based on definitions from the International Organization for Standardization ISO/IEC 22989:2022:

“An engineered system that generates predictive outputs such as content, forecasts, recommendations, or decisions for a given set of human defined objectives or parameters without explicit programming. AI systems are designed to operate with varying levels of automation.”

Policy requirements

This policy uses a principles-based approach to set out essential considerations for all WA Government employees producing or using AI solutions in their work. In all cases, human judgement and intervention are vital in harnessing AI's potential to be implemented ethically, safely and responsibly.

These principles must be considered in all stages of the AI lifecycle including

- design, data and modelling (such as planning, data collection and model building)
- development and validation (such as training and testing)
- deployment (including the use of existing AI technologies)
- monitoring and refinement (including fixing any problems that occur).

Benefit to the community and agency must be considered in every instance. WA Government employees must consider whether an AI solution would assist in improving the productivity of the agency, assist stakeholders and/or improve agency capabilities. WA Government employees must also consider whether a non-AI solution could be used to achieve the same result. When AI is being employed, it must be because it is the best solution for the task.



Human, social and environmental wellbeing

Throughout their lifecycle, AI systems must benefit individuals, society and the environment.

This principle aims to clearly indicate from the outset that AI systems must be used for beneficial outcomes for individuals, society and the environment. AI system objectives must be clearly identified and justified. AI systems designed for legitimate internal business purposes, like increasing efficiency, can have broader impacts on individual, social and environmental wellbeing. Those impacts, both positive and negative, must be accounted for throughout the AI system's lifecycle, including impacts outside the organisation.

Human-centred values

Throughout their lifecycle, AI systems must respect human rights, diversity, and the autonomy of individuals.

This principle aims to ensure that AI systems are aligned with human values. AI systems must enable an equitable and democratic society by respecting, protecting and promoting human rights, enabling diversity, respecting human freedom and the autonomy of individuals, and protecting the environment. Human rights risks need to be carefully considered, as AI systems can equally enable and hamper such fundamental rights. AI systems must not undermine the democratic process, and must not undertake actions that threaten individual autonomy, like deception, unfair manipulation, unjustified surveillance, and failing to maintain alignment between a disclosed purpose and true actions and outcomes.

Fairness

Are the models trained and tested on relevant, accurate, and generalisable datasets and is the AI system deployed by users trained to implement them responsibly to manage and mitigate bias.

This principle aims to ensure that AI systems are fair and that they enable inclusion throughout their entire lifecycle. AI systems must be user-centric and designed in a way that allows all people interacting with them to access the related products or services. This includes both appropriate consultation with stakeholders, who may be affected by the AI system throughout its lifecycle, and ensuring people receive equitable access and treatment.

The best use of AI will depend on data quality and relevant data. It will also rely on careful data management to ensure potential data biases are identified and appropriately managed. Measures must be taken to ensure AI produced decisions are compliant with anti-discrimination laws. Non-discrimination is essential to ensure that AI technologies serve all individuals equally.



Privacy protection and security

Compliance with appropriate data policies and legislation, for example, the forthcoming WA Privacy & Responsible Information Sharing (PRIS) and the Commonwealth Privacy Act 1988.

Respect for privacy and data protection is paramount when using AI systems. This includes ensuring proper data governance, and management, for all data used and generated by the AI system throughout its lifecycle. For example, maintaining privacy through appropriate data anonymisation and informed consent principles. AI systems must incorporate privacy by design principles.

There must also be appropriate data and AI system security measures in place. This includes the identification of potential security vulnerabilities, and assurance of resilience to adversarial attacks.

Reliability and safety

Throughout their lifecycle, AI systems must reliably operate in accordance with their intended purpose.

This principle aims to ensure that AI systems reliably operate in accordance with their intended purpose throughout their lifecycle. This includes ensuring AI systems are reliable, accurate and reproducible as appropriate. AI systems must not pose unreasonable safety risks and must adopt safety measures that are proportionate to the magnitude of potential risks. AI systems must be monitored and tested to ensure they continue to meet their intended purpose, and any identified problems must be addressed with ongoing risk management as appropriate. Further, the connection between data, and inferences drawn from that data by AI systems, must be sound and assessed in an ongoing manner.

Responsibility must be clearly and appropriately identified for ensuring that an AI system is robust and safe.

Transparency, explainability and contestability

So affected stakeholders can know how the AI model reached its decision.

Data must be used safely and in accordance with relevant legislation. The people of WA must have access to an efficient and transparent review mechanism if there are questions about the use of data or AI-informed outcomes. This includes disclosing when an AI system is engaging with them, or when AI has been used to produce an outcome that impacts them.

There must be sufficient access to and understanding of the algorithm, and inferences drawn, to make contestability effective. In the case of decisions significantly affecting rights, there must be an effective system of oversight, which makes appropriate use of human judgment.

Accountability

Consider who is responsible for each element of the model's output and how the designers and implementers of AI systems will be held accountable.

Decision-making remains the responsibility of organisations and individuals. AI-based functions and decisions must always be subject to human review and intervention. Mechanisms must be put in place to ensure responsibility and accountability for AI systems and their outcomes at all stages of the AI lifecycle. The organisation and individual accountable for the decision must be identifiable as necessary and AI systems that have a significant impact on an individual's rights must be accountable to independent review. This includes providing timely, accurate, and complete information for the purposes of independent oversight bodies.

Accuracy

Have the outputs been checked to ensure that they are factual, reliable and unbiased.

AI has the potential to provide valuable insights but can also provide incorrect information and decisions. These inaccurate AI outputs can have far-reaching implications, especially when AI is being used in public services and can impact the wider community. The responsibility is with all WA Government employees involved in the development or use of AI technology to ensure the accuracy of inputs and outputs.

Relevant Legislation, Policies and Other Documents

Many existing policies, laws and regulations already have embedded components that are applicable in the usage of AI and so agencies must continue to be mindful of these. Agencies must continue to comply with Acts and legislation.

Some of the legislation and policies that are applicable include (but are not limited to):

- *Privacy and Responsible Information Sharing (PRIS) legislation - forthcoming*
- *Public Sector Management Act 1994*
- *State Records Act 2000*
- *Freedom of Information Act 1992*

