# **OPERATIONAL POLICIES**

# **Generative Artificial Intelligence**



Policy Number: 012

Relevant Delegation: Not Applicable
Adoption Details: 20 June 2024

Last Review Details: N/A

## **POLICY OBJECTIVE:**

The purpose of this Policy is to provide a framework for the use of generative AI for the Shire of York and protocols to reduce unintended risk to business operations and reputation.

#### **POLICY SCOPE:**

This Policy is designed to provide a framework for the Shire to use generative AI to provide efficiencies across the organisation.

This Policy applies to all Shire Employees.

The requirements of this Policy apply to all uses of generative AI applications, which are applied to any Shire of York related purpose and regardless of whether the applications are hosted corporately or not.

These protocols cover the use of AI systems only.

#### **POLICY STATEMENT:**

#### INTRODUCTION

Generative AI refers to a set of technologies that has the capacity to generate novel content such as text, images, music, video and computing code, after being prompted by a user.

They are a subset of the broad field of artificial intelligence technologies. As this is a rapidly developing area, the scope of generative AI may change over time and therefore regular reviews of this policy may be required to allow for the use of future AI technology.

#### **PRINCIPLES**

Before the use of any generative AI software or platform on behalf the Shire of York, users should be familiar with the eight agreed Ethics Principles for responsible generative AI use created by the Australian Department of Industry, Science and Resources.

These principles are designed to ensure the use and deployment of AI is safe, secure, and reliable.

- 1. **Human, societal and environmental wellbeing**: Al systems should benefit individuals, society and the environment.
- 2. **Human-centred values**: Al systems should respect human rights, diversity, and the autonomy of individuals.
- 3. **Fairness**: Al systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.
- 4. **Privacy protection and security**: All systems should respect and uphold privacy rights and data protection, and ensure the security of data.
- 5. **Reliability and safety**: Al systems should reliably operate in accordance with their intended purpose.

- 6. **Transparency and explainability**: There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI, and can find out when an AI system is engaging with them.
- 7. **Contestability**: When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.
- 8. **Accountability**: People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

#### **PROVISIONS**

#### USING GENERATIVE AI AS A REPRESENTATIVE OF THE SHIRE

Before utilising generative AI as a representative of the Shire, Employees must become authorised by gaining approval from the Chief Executive Officer (CEO).

Employees must confirm with the IT Department which generative AI platforms are available and appropriate for use at the Shire of York based on agreed use cases.

Employees must ensure they are using the correct level of access for the agreed platform.

Only generative AI software or platform access purchased in accordance with this policy are to be used to create work or collateral on behalf of the Shire.

Prior to the use of any generative AI software or platform, Employees must receive instructions from the IT Department on the relevant licensing agreements, including any restrictions on use.

All Employees must receive training for all new software. This will be the responsibility of the IT Department.

Once you have been authorised, you will be referred to as an 'Authorised User' (see 'Definitions' below).

In the case of Authorised Users, where there is any question that the use of generative AI software or platform might not meet one of the Australian principles for responsible use, it should be reviewed by a line manager and the CEO before use of the system is allowed.

All users must ensure that all content generated by the use of generative AI software or platform is accurate and not misleading and complies with all relevant Shire of York Policies, including the Shire's Media Policy, Code of Conduct and Record Keeping Policy.

All users must adhere to the Terms of Use of the relevant generative AI platform, as well as copyright, privacy, defamation, contempt of court, discrimination, harassment and other applicable laws.

#### **USE BY THE CEO**

The CEO is not required to obtain authorisation to use generative AI software or platforms — though must follow the Shire's Code of Conduct and Rules of Engagement.

Disclosure of use is encouraged.

When using AI software or platforms to develop public collateral referring to the Shire the CEO must be aware of and check for the risk of 'hallucination' or the inadvertent creation of inaccurate, incorrect or misleading information by the AI platform.

## **RISK MITIGATION**

To reduce potential exposure to the Shire of York, all users must be aware of the following risks and employ the mitigation strategies:

- **Privacy**: strictly avoid the use of personal information in interactions with generative AI tools. Personal information includes but is not limited to names, addresses, contact details, and any identifiable data.
- **Confidentiality**: never use internal, sensitive, financial, or corporate information in generative AI tools to maintain confidentiality.
- *Hallucinations* (Al-generated inaccuracies): be cautious of relying on facts or information generated by Al. Verify the source of Al-generated information and ensure its accuracy through human oversight.
- **Bias**: actively work to prevent the introduction or reproduction of bias in Al outputs, particularly in protected categories such as race, age, ethnicity, sexuality, and gender.
- **Explicability**: avoid using generative AI for decision-making unless the AI's chain of reasoning is clear, apparent, replicable, and subject to human oversight.

#### **RELEVANT LEGISLATION:**

Not applicable.

#### **PENALTIES:**

Depending on the circumstances, non-compliance with this Policy may constitute a breach of employment or contractual obligations.

Any employee who fails to comply with this Policy may face disciplinary action and, in serious cases, termination of their employment.

Council Members who fail to comply with this Policy will be removed from the Shire's generative Al subscription database and may also face action under the Shire's Code of Conduct.

# **KEY TERMS/DEFINITIONS:**

**Generative AI** is a set of technologies that has the capacity to generate content such as text, images, music, video and computing code, after being prompted by a user. They are a subset of the broad field of artificial intelligence technologies.

Large language models (LLMs) are one of the technologies Generative AI are often built on. LLMs use algorithms trained on vast amounts of data, primarily text, which helps the model predict and approximate relationships between data. A prompt from a user can be 'read' by these models and used to generate often complex responses.

**Multimodal foundation models (MFMs**) are another technology used to build generative AI. Similar, but more complex than LLM, MFMs use a wider range of information, including images, speech, numerical inputs and code, and can generate responses based on this rich source of information.

**Authorised User** is a person who has been granted permission to use generative AI to speed up and scale different forms of work on behalf of the Shire of York.

Responsible Officer: Chief Executive Officer

Contact Officer: Administration and Governance Coordinator

Relevant Legislation: Not applicable

**Review History:** 

Date Review Adopted:	
Adopted	ELT - 20 June 2024

