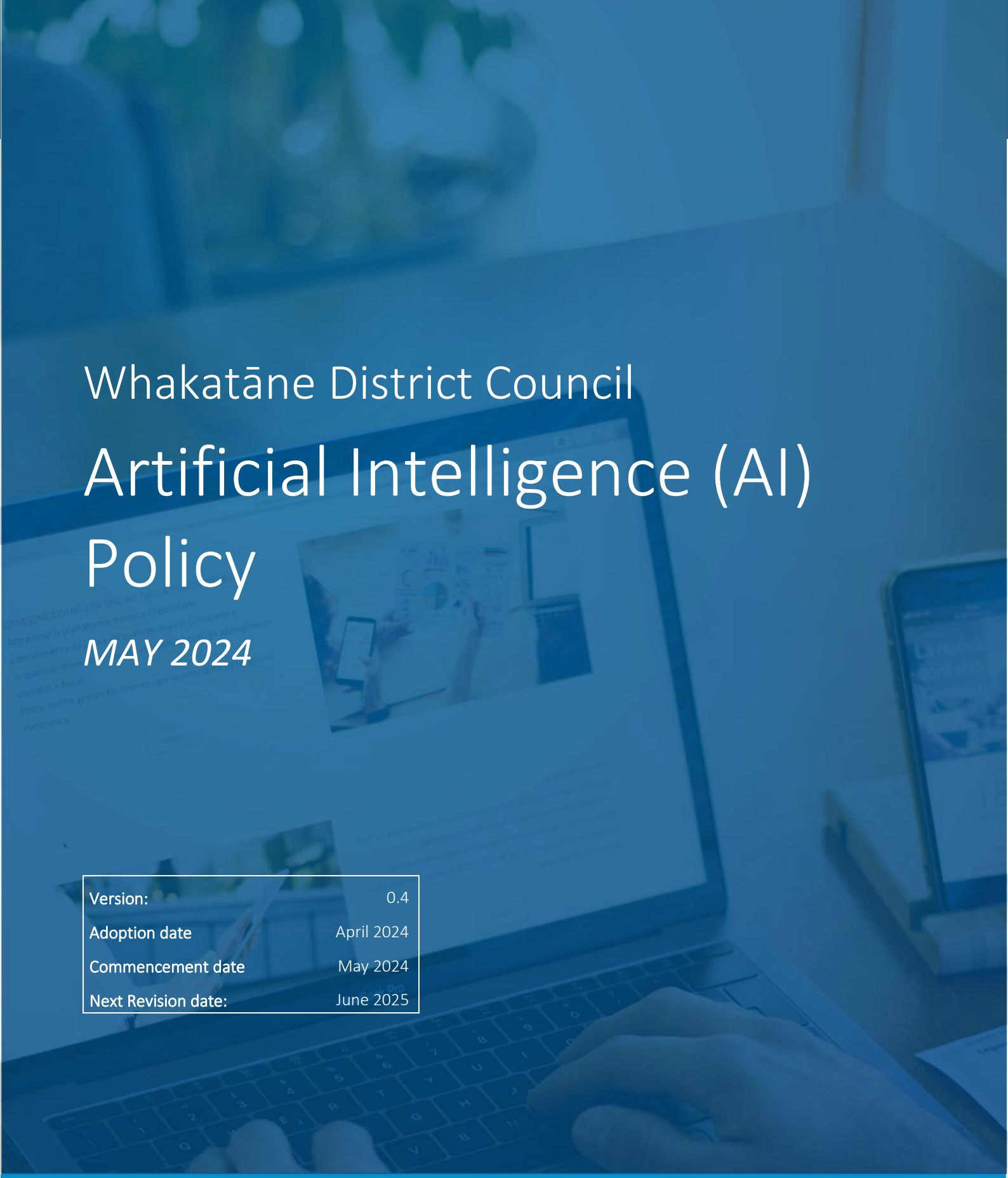

Whakatāne District Council

# Artificial Intelligence (AI) Policy

*MAY 2024*

| | |
|---|---|
| Version: | 0.4 |
| Adoption date | April 2024 |
| Commencement date | May 2024 |
| Next Revision date: | June 2025 |

whakatane.govt.nz

WHAKATĀNE District Council
*Kia Whakatāne au i ahau*

# Artificial Intelligence (AI) Policy

## 1. Purpose

The purpose of this policy is to establish a framework for the ethical use of a Generative Artificial Intelligence Large Language Models (AI) including Copilot, ChatGPT, Bard, Bing, or other similar tools. This policy applies to employees, contractors, temporary staff, or other third parties (all referred to as 'employee' in this document).

AI tools should only be used in a way that promotes fairness and avoids biases to prevent discrimination and promote equal treatment and contribute positively to the Council's goals and values.

This policy is designed to ensure that the use of AI is ethical, lawful, and complies with all applicable laws, regulations, and Council policies, and to complement Council's existing information and security policies.

Any use of AI technology for Council activities should be done with full acknowledgement of the policies, terms and conditions of the AI developer/vendor. Particular attention should be given to Privacy, Access, Data Usage and Data Deletion policies.

## 2. Scope

This policy applies to all employees, contractors, temporary staff, or third parties with access to AI, through Council-owned or BYOD (bring your own device) for Council activities.

## 3. Policy

### *3.1 Use of AI*

Employees are authorised to use AI for work-related purposes. This includes tasks like generating text or content for reports, emails, presentations, images and customer service communications.

Before accessing AI technology, employees should first notify their People Leader of their intent to use AI, the reason for using it and what type of information will be uploaded, and the generated output and expected distribution of content.

### *3.2 Copyright*

Employees should follow copyright laws when using AI. AI cannot be used to create content that impacts the intellectual property rights of others, including but not limited to, copyrighted material. If an employee is unsure if a particular use of AI constitutes copyright infringement, they should contact their People Leader before using AI.

*3.3 Accuracy*

Employees are responsible for reviewing any outputs, and are accountable for ensuring the accuracy of AI generated output before use/release. **If an employee has any doubt about the accuracy of information generated by AI, they should not use that information.**

*3.4 Confidentiality*

Confidential information mustn't be entered into any AI tool, because it could enter the public domain. Employees should follow all applicable data privacy laws and organisational policies when using AI. **If an employee has any doubt about the confidentiality of information, they should not use it.**

*3.5 Ethical use*

AI should be used ethically and comply with all applicable legislation, regulations and organisational policies. Employees shouldn't use AI to generate content that is discriminatory, offensive or inappropriate. If there are any doubts about the appropriateness of using AI in a particular situation, employees should check with their People Leader or the Manager Digital Services.

*3.6 Label*

Content produced via AI must be identified and disclosed as containing AI-generated information.

Footnote example: ***Note:*** *This document contains AI generated content. AI generated content has been reviewed by the author for accuracy and edited/revised where necessary. The author takes responsibility for this content.*

*3.7 Development and use of API and plugin tools*

API and plugin tools enable extra access to, and functionality for, AI services to improve automation and productivity outputs; however, they also represent additional risks. OpenAI's Safety Best Practices guidelines recommend the following concepts should be included when developing API and plugin tools for internal systems:

- Adversarial testing
- Human in the loop (HITL)
- Prompt engineering
- "Know your customer" (KYC)
- Constrain user input and limit output tokens
- Allow users to report issues
- Understand and communicate limitations
- End-user IDs

API and plugin tools must be rigorously tested for:

- Moderation – to ensure the model properly handles hate, discriminatory, threatening, etc.
- Factual responses – provide a ground of truth for the API and review responses accordingly

# 4. Risks

The use of AI has inherent risks that employees should be aware of. A comprehensive risk assessment should be conducted for any project or process where AI is proposed to be used. The risk assessment should consider the potential impact of potential risks regarding legal; accuracy of output; bias and discrimination; security (including technical protections and security certifications); and data sovereignty and data protection.

### *4.1 Legal*

Information entered into AI might enter the public domain. This can release non-public information and breach regulatory requirements, customer or vendor contracts, or compromise intellectual property.

Any release of private or personal information without the authorisation of the information's owner could result in a breach of the principles of the Privacy Act 2020, specifically:

- Principle 5 – Storage and security of information
- Principle 9 – Limits on retention of personal information
- Principle 10 – Use personal information
- Principle 11 – Disclosing personal information
- Principle 12 - Disclosure outside New Zealand

This may also include a breach of s19 of the Bill of Rights Act 1990, *Freedom from Discrimination*.

Any unauthorised release of public information and records may result in a breach of the principles of the *Information and Records Management Standard* issued under s27 of the Public Records Act 2005.

• Information and records must be protected from unauthorised or unlawful access, alteration, loss, deletion and/or destruction.

• Access to, use of and sharing of information and records must be managed appropriately in line with legal and business requirements.

Use of AI to compile content may also infringe on regulations for the protection of intellectual property rights including the Copyright Act 1994.

*4.2 Accuracy of output*

AI uses algorithms to generate content, and there is a risk that AI may generate inaccurate or unreliable information. Employees should exercise caution when relying on AI generated content and should always review and edit responses for accuracy before using the content. Employees should critically assess and question/challenge AI generated output before use. **If an employee has any doubt about the accuracy of information generated by AI, they should not use it.**

*4.3 Bias and discrimination*

AI may produce bias, discriminatory or offensive content. Employees should use AI responsibly and ethically, and comply with Council policies and applicable laws and regulations.

*4.4 Security*

AI might store sensitive data and information, which could be at risk of being breached or hacked. Council must assess AI technical protections and security certification before use. **If an employee has any doubt about the security of information input into AI, they should not use it.**

*4.5 Data sovereignty and data protection*

While an AI platform might be hosted internationally, information created or collected in New Zealand, under data sovereignty rules, is still under jurisdiction of New Zealand laws. The reverse also applies. If information is sourced from AI hosted overseas for use in New Zealand, the laws of the source country regarding its use and access may apply. AI service providers should be assessed for data sovereignty practice by any organisation wanting to use AI.

# 5. Compliance

Any violations of this policy should be reported to the Manager Digital Services.

# 6. Review

This policy will be reviewed periodically and updated as necessary to ensure continued compliance with all applicable legislation, regulations and organisational policies.

# 7. Acknowledgment

By using AI, employees acknowledge that they have read and understood this policy, including the risks associated with the use of AI. Employees also agree to comply with this policy and to report any violations or concerns to People and Capability.

# 8. Definitions

| Term | Definition |
|---|---|
| API | Application Programming Interface - a set of functions and procedures allowing the creation of applications that access the features or data of an operating system, application, or other service |
| Author | The person that writes the document and would be considered the owner |
| Generative AI | Generative artificial intelligence is artificial intelligence capable of generating text, images, videos, or other data using generative models, often in response to prompts |
| OpenAI | A U.S. based artificial intelligence (AI) research organisation founded in December 2015, researching artificial intelligence with the goal of developing "safe and beneficial" artificial general intelligence |
| Plugin Tools | Computer software that adds new functions to a host program without altering the host program itself |
| Public Domain | Creative materials that are not protected by intellectual property laws such as copyright, trademark, or patent laws |

Original authors: **Jackie Lyons, CPA, NACD.DC, QFE, Cyber Risk Governance℠**; Mike Manson, Chief Executive, ALGIM