Te Papa Al Guidelines

A	pprov	ed Gen-Al tools for Kaimahi to use	2
Pı	rincip	les related to enabling and empowering kaimahi	4
	1. skills	Emphasising talent and creativity to develop internal and external Al knowledge and s, whilst safeguarding the knowledge, expertise and creativity of kaimahi	4
		Prioritising the productivity and job satisfaction of kaimahi and the organisation by aging AI to simplify tasks, boost efficiency, and reduce frustration, and create new ortunities	5
	3.	Enabling environments for safe experimentation, testing, and evaluation of new AI ideas echnologies, fostering innovation without risk to operational systems	3
Ρı	rincip	les related to maintaining trust and confidence	8
	4. Tiriti,	Recognising Māori rights and interests including the application of Mana Taonga, Te Māori Data Governance in decision making	8
	5. Ahur	Considering security, privacy, policy and legislation by incorporating Te Rautaki - Hāpai ea, Tūrangawaewae, Papatūānuku, Te Papa policies and legislation in decision making. 12	
		Ensuring human oversight, transparency and accountability so that content created or ified by AI must be reviewed by a qualified human to ensure the robustness of the output s reliable, fair, and consistent with organizational standards.	
	discı	Informed use and serving the public AI so that use and processes should align with the aki and our policies ensuring that they do not unjustly harm, exclude, disempower, or iminate against individuals or particular groups, while actively reducing bias and porting data ethics	7
	8. on Pa	Considering environmental impacts when using AI so as not to further impact negatively	
	Appe	endix A - Definitions	2

This document is to be used in conjunction with the Te Papa Artificial Intelligence Policy. The following descriptions and examples should help apply the policy to your work. Because the AI space is changing so rapidly, not every situation is likely to be covered by this guideline, and we will continue to update it as new information, questions, and use cases arise.

Al Acknowledgement: The creation of the content in this document was created with the support of GenAl tools, including ChatGPT and Copilot, but the final output reviewed by a multiple Te Papa staff.

Approved Gen-Al tools for Kaimahi to use

These free AI tools are available to Te Papa kaimahi, highlighting their usage limits, data handling practices, and login options. These tools have been selected based on their reliability and compliance with our AI Policy and Guidelines.

Last updated 15 August 2025

Al Product	Daily Usage Limit	Uses Your Data for Training	Caches Your Data	Log on
CoPilot Use this image in the Edge browser to open CoPilot	Unlimited	No	Yes Your cached data in the free version of Copilot is protected, not discoverable by others.	Already logged on as part of your Te Papa logon and password
ChatGPT current version (as at 15 August 2025) https://chat gpt.com/	Limits on some models within a 5 hours window https://help.openai.com/en/articles/9275245-chatgpt-freetier-faq	By default, Yes – OpenAI may use chats unless you disable chat history in settings to opt out https://help.openai. com/en/articles/773 0893-data-controls- faq	By default, Yes – OpenAI will cache prompts and answers unless you disable chat history in settings to opt out https://help.openai. com/en/articles/859 0148-memory-faq	Free Need to create an account. For work purposes you must use your Te Papa email address
Claude https://clau de.ai/	Limited Short conversations allow around 45 messages every 5 hours. Longer conversations or large attachments reduce this to about 15 messages, as Claude re-reads the entire conversation each time you send a message.	No	Yes Cached prompts have a minimum lifetime of 5 minutes, refreshed with each use	Free Need to create an account. For work purposes you must use your Te Papa email address

Te Papa AI Guidelines Version: 1.1 Version date: 15 August 2025 PM#14310623

Principles related to enabling and empowering kaimahi

 Emphasising talent and creativity to develop internal and external AI knowledge and skills, whilst safeguarding the knowledge, expertise and creativity of kaimahi

Explanation

We want to encourage innovation and skill development in AI, and kaimahi should be confident using one or more of the approved AI tools to use. (see table above) We also recognise that this requires the need to maintain and protect the collective and individual intellectual and creative contributions and context of kaimahi and those that we work with. This means always being aware of ethical considerations, fair recognition, or cultural respect (see also principles 3 and 8).

Kaimahi know who to ask if they have a question, e.g around safeguarding knowledge.

Examples

- Kaimahi who work regularly with imaging or video, have access and confidence in the tool that suits those processes, such as the built-in AI in Adobe Photoshop
- Kaimahi who work with lots of data should be confident in using data processing AI such as Google BigQuery, Google Sheets with Gemini, ChatGPT with Data Analysis & Report AI.
- Al workflow is defined as a way of saving time, reducing manual work, or improving how something gets done by bringing AI in at the right points in the process.

Q&A

Q: How are we safeguard people's creativity?

A: This is a challenge that many industries are grappling with as AI becomes more capable. Here are some key ways to protect workers' knowledge and creativity:

- Human-Al Collaboration: Al should enhance creativity, not replace it. Workers can focus on innovation, judgment, and originality while Al handles repetitive tasks.
- Intellectual Property Rights: Upholding copyright and trademark laws help ensure that human-created work isn't misused or undervalued by AI-generated content
- Transparency & Ethics: AI-generated content should be clearly labelled so human-created work gets proper recognition.

 Fair Recognition Models: Consider how to fairly compensate creatives when Algenerated work is derived from human knowledge

2. Prioritising the productivity and job satisfaction of kaimahi and the organisation by leveraging AI to simplify tasks, boost efficiency, and reduce frustration, and create new opportunities

Explanation

We will focus on enhancing both kaimahi well-being and organisational effectiveness by strategically implementing AI as a tool to augment our work. AI has the potential to significantly enhance job satisfaction by automating routine and repetitive tasks, freeing kaimahi to concentrate on more engaging, complex, and creative mahi. This shift not only improves efficiency and productivity but also contributes to a reduction in frustration and stress among kaimahi, fostering a more positive and engaged workforce that is better equipped to effectively serve its audience.

It's important to emphasise that the strategic use of AI is not designed to replace kaimahi, but rather to be empowering by optimising their work and enabling focus on critical thinking, collaborative and creative activities.

Examples

- Al tools can automate tasks such as data entry, report generation, and basic communication responses, allowing kaimahi to dedicate more time to strategic planning, relationship building, and problem-solving. For example, you are writing a report and have some key data that you would like to add. Using a secure Al platform, you enter that data with some prompts on the output. you use this as the basis to start your report.
- Al can assist in summarising large documents, extracting key information, and identifying trends in data, enabling kaimahi to make more informed decisions and reducing the cognitive load associated with information overload. You want to understand some key trends in visitation data so with clear prompts, you have the ability to upload documents to analyse and summarise the trends, but you need to ensure the data in the document is not sensitive and/or contains identifiable personal information.
- All can be used as a tool to brainstorm ideas, generate initial drafts of creative content, and provide feedback on written work, supporting kaimahi in their creative endeavours and potentially sparking new avenues for innovation.

You are starting a new project and procrastinating about where to start. Using an AI tool, you upload the project details and timelines with clear prompts to help you get started.

 Al can facilitate personalised learning and development opportunities, tailoring training materials to individual kaimahi needs and skill gaps, thereby enhancing their professional growth and job satisfaction. For example, you are learning a new skill or want to strengthen an area of your work, Al can provide learning resources or develop tutorials to help you.

Q&A

Q: How can Al make my job more satisfying?

A: All can take care of repetitive tasks, giving you more time to focus on work that matters. Tools like All proofreading or task management can free up your day and make things flow better. Just remember to check All outputs for accuracy.

Q: Will AI take my job?

A: No – Al's here to support, not replace. It can handle routine bits so you can focus on the human side of your work: thinking, creating, and connecting with others.

Q: How do I learn to use AI well?

A: You'll get support. There's training on prompt writing, responsible use, and plenty of examples. Members of the AI Exploration are around for chats, and regular drop-ins are a chance to learn from each other.

Q: What makes a good Al prompt?

A: Be clear and specific. Instead of "summarise this," try "Summarise this 10-page report on visitor feedback into ten bullet points, focusing on accessibility and key action points."

Q: Can AI help with creative work?

A: Definitely. It's great for brainstorming, first drafts, or fresh ideas. You'll still need to tweak things, but it's a good way to get started.

Q: How can AI simplify tasks and boost efficiency?

A: All can speed things up and reduce admin. For example, Outlook's Suggested Replies help write emails faster, and All can automate data entry to free up time for other work.

Q: Is the use of Al mandatory - can I opt-out of using it?

A: While AI tools are available to assist and enhance your work, participation is optional. You can choose whether to use some AI based on your needs and comfort level. It's best to discuss any concerns with your manager.

3. Enabling environments for safe experimentation, testing, and evaluation of new AI ideas and technologies, fostering innovation without risk to operational systems.

Explanation

We must balance risk with productivity and creativity, by providing guidelines to enable safe experimentation and use. Such guardrails include providing support in choosing and using AI models, and training in best practice. People Managers can support AI experimentation with AI within the content of Kaimahi roles, while also reinforcing relevant policy principles and other safeguards. We will also, at a high level, monitor the impacts and uses of AI across the organisation. Existing policies such as the Privacy Policy, ICT Policy, and others also provide information on acceptable use. See also **Principle 5** Considering security, privacy, policy and legislation

The AI Exploration Group (see appendix for names) and your manager is available to help you navigate both the opportunities and challenges while you experiment. They can discuss key issues such as ethics, environmental impact, security risks, and data sovereignty, ensuring you make informed decisions while exploring AI's potential.

Example

You are undertaking some research across a broad range of business activities within Te Papa and want to ensure that you are using AI for efficiency, whilst not releasing confidential information externally without realising it. You consult the Te Papa approved AI tools to see what can be used, follow Te Papa AI policy and discuss potential methodology with identified kaimahi AI champions. You may then consider contracting specialist advice and sharing this work with relevant teams or developing a templated approach that may then be re-used by other kaimahi.

Q&A

Q: What frameworks guide our use of Al?

A: Start by checking Te Papa's list of approved AI tools (see table above). All staff also complete mandatory training on <u>Privacy Awareness induction</u> and <u>Information</u> <u>Technology Induction</u> / <u>Annual Refresher course</u>.

Q: How is Te Papa monitoring Al use across the organisation?

A: We've run a staff-wide survey on current and planned AI use and will repeat it regularly to track changes and impact. Some projects may involve deeper analysis where AI is a key focus.

Q: How do I get a subscription for an AI tool?

A: First, check if the tool is approved under Te Papa's AI guidelines. All licences and subscriptions are managed by IT. You can <u>contact IT</u> to discuss your needs and any associated costs.

Q: What if the AI tool I want isn't approved yet?

A: You can request for it to be reviewed by submitting a request through the <u>ICT Request</u> Link.

Q: What training is available to help me use AI responsibly?

A: Core training on Privacy and ICT Security is part of induction. Introductory AI training, including how to choose and use tools, is available on the Kupenga page.

Q: Will digital literacy be part of job descriptions?

A: Yes, where relevant. As digital tools become more central to our work, confidence using them is increasingly important. Support and training will be available.

Principles related to maintaining trust and confidence

4. Recognising Māori rights and interests including the application of Mana Taonga, Te Tiriti, Māori Data Governance in decision making

Explanation

Māori data includes any digital or digitised information that relates to Māori people, culture, or knowledge—including *mātauranga Māori* (Māori knowledge). This data is considered a *taonga* (treasure), meaning it holds cultural and historical significance and should be treated with care and respect. It also means we need to consider Article Two of Te Tiriti o Waitangi.

Before using AI with Māori data, it's important to think about Māori rights and interests. Here are some key questions to consider:

- Mana Taonga principle How have we respected and upheld the Mana Taonga principle?
- Governance and authority Do Māori have control over how their data is managed, used, and interpreted? How can we ensure this happens?

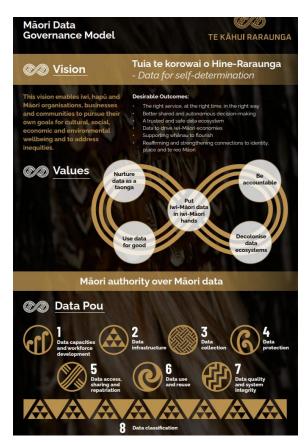
- Culturally safe practices What responsibilities do we have to handle this data in a way that respects mātauranga Māori, Māori perspectives and tikanga?
- Potential risks (BADDR¹ practices) Could AI use Māori data in ways that disconnect it from its proper cultural context? For example, does it risk losing its connection to mātauranga (knowledge), whānau (family), hapū (sub-tribes), and iwi (tribes)? How can we prevent this?

In simple terms, the goal is to ensure that Māori data remains connected to its rightful owners and is handled with cultural integrity and fairness.

Examples

- Using AI you will need to consider how to consider privacy, and for Māori this
 includes individual and collective privacy (see Privacy Commissions talk), data
 security (for sensitive information such as financial, collection location, valuation or
 provenance, commercially sensitive data or data about, by, for Māori (see Pou 5 and
 6 of the Māori Data Governance Model below).
- Misrepresentation or Loss of Māori Context. Al works by grouping and analysing lots
 of data, but it might accidentally misrepresent Māori identities or remove important
 cultural context. For example, if Al looks at Māori taonga as individual instead of
 seeing them as part of whānau (family), hapū (sub-tribes), or iwi (tribes) knowledges,
 it might overlook key aspects of Māori identity and ways of living.
- Using AI without considering bias (see Principle 7 too). AI learns from existing
 information, but sometimes that information is biased. For example, AI might have
 been trained on documents where Māori knowledge was taken without permission
 (appropriated) or where outsiders described Māori culture in ways that are incorrect
 or harmful. If AI repeats these mistakes, it can spread false narratives that don't
 reflect true Māori perspectives.

¹ Blame Māori by directly or indirectly situating the dominant group as the ideal group, and/or Māori as being culpable for their poorer outcomes. • Aggregate data in ways that misrepresent or miss key aspects of Māori identities and lived experience. • Decontextualise data, by focusing on Māori individuals and families outside of their social and/or cultural context, and is • Deficit-based, implying that Māori are inherently deficient. • Restrict access to Māori data under the control of statistical agencies and official institutions.



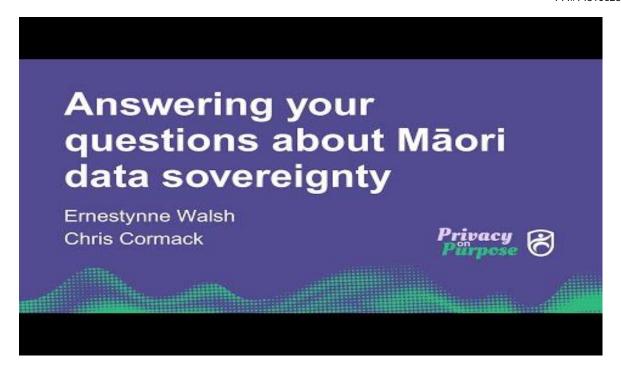
Te Kāhui Raraunga Māori Data Governance Model

Te Kāhui Raraunga Charitable Trust (Te Kāhui Raraunga or TKR) is an independent body established in 2019 to lead the action required to realise the advocacy of the Data Iwi Leaders Group (Data ILG).

Q&A

Q: Where is Te Papa getting its direction on Māori data governance and sovereignty from?

A: Te Papa is using the resources from <u>Te Mana Raraunga</u>. You may also want consider the <u>Principles of Māori Data Sovereignty</u>, Te Kāhui Rarauna <u>Māori Data Governance</u> <u>Model</u>, and the <u>Care Project for Indigenous Data Governance</u>. There are also some good talks on Māori data sovereignty and governance from the 2025 Privacy Commission <u>Do it. Yourself.</u> 3 min video <u>Māori Data Governance</u> <u>| Te Kāhui Raraunga</u>



Q. What are some of the concerns and worries of embracing AI from a Māori perspective?

A. One big worry is that AI might take elements of Māori culture and use them without proper understanding or respect. This is called cultural appropriation, and it can remove important meaning and context.

For Te Papa, a major risk is sharing Māori information publicly without approval from whānau (family), hapū (sub-tribes), or iwi (tribes). This would go against the Mana Taonga Principle, which ensures Māori communities have a say in how their cultural treasures (taonga) are treated. Another concern is decontextualisation, where Al presents Māori knowledge (mātauranga) in a way that separates it from its original stories, communities, and meanings. If Al removes these connections, Māori culture could be misrepresented or misunderstood.

The key takeaway is that Māori data and knowledge should be handled with respect, understanding, and proper consultation.

Q. How will we work with iwi and hapu to ensure their information and knowledge is safe?

A. Transparency and honest with information sharing with guidance by data sovereignty experts and Te Papa values and mana taonga principle. Work with the Iwi Partnerships team early on.

5. Considering security, privacy, policy and legislation by incorporating Te Rautaki - Hāpai Ahurea, Tūrangawaewae, Papatūānuku, Te Papa policies and legislation in decision making.

Explanation

New Zealand does not currently have specific laws that regulate the use or deployment of AI. Instead, AI is regulated through existing technology-neutral laws covering areas like privacy, consumer protection, intellectual property, and human rights. Using public models (such as free versions of OpenAI or other online models) you will need to consider how to handle privacy (data that includes information that could identify people), data security (for sensitive information such as financial, collection location, valuation or provenance, commercially sensitive data or data about, by, for Māori (see also Principle 4)).

Examples

- You're writing a Proposal Document for ELT, and you really want to get it right. You
 copy the budget table from your document and ask Te Papa's Microsoft 365 Copilot
 (because the data is protected in the Te Papa MS system) to check the maths and
 see if you've missed any key costs. Then you ask Copilot to review the document for
 spelling, grammar and overall accuracy and completeness, and to write an executive
 summary for you, which you then adapt.
- You want to use tools such as Google's Cloud Vision API or Microsoft's Azure Computer Vision to automatically label and categorise objects with images on collection items. These tools can identify objects, artworks, and even specific features within the items AI to assist the curatorial practice in labelling collection items, identifying patterns, and uncovering new insights from data, and in turn, making them more accessible. You've ensured you're only using images that are out of copyright or have appropriately licenced, and do not include taonga or other sensitive cultural material. (See the blog / for an example)

Q&A

Q: Is it safe to upload documents containing sensitive information, such as financial data, personal details, or content that is not yet approved or published?

A: No. If you're using the free version of ChatGPT or another public AI tool, avoid entering confidential museum information. Inputs may be used to improve the AI, even

if they're not made public. Paid versions like ChatGPT Enterprise or Team don't use your data for training. Always check the platform's privacy terms.

Q: How do I know if I am considering security, privacy, policy, and legislation?

A:

- Read and understand your responsibilities by reviewing the Information Technology Policy and Privacy and Personal Information Policy.
- Complete the mandatory training: Privacy Awareness induction and Information Technology Induction / Annual Refresher course.
- Refer to the Kupenga privacy page for guidance on the Privacy Act and Te Papa's obligations.
- Visit the Kupenga Information Security page for additional help and support.

Q: How can I learn about Te Papa's relevant security, privacy, policy, and legislation?

A:

- Read the Public Service Al Framework.
- Explore AI-related guidance from the Privacy Commissioner.
- Review broader public service advice from the <u>Centre for data ethics and</u> innovations.

Q: How can I explore the impacts of AI from a Tūrangawaewae perspective?

A:

To explore the impacts of AI from a Tūrangawaewae perspective, consider the following resources:

- "Tūrangawaewae: A place to stand. A Discourse of Cultural Awareness or Hope" by Mere Kēpa This paper critiques the development and advancement of indigenous knowledge while taking care of taonga within contested boundaries and spaces. It provides insights into how cultural awareness and governance can be integrated with modern technologies.
- "Tūrangawaewae, Tikanga and Talking to Each Other" by Deidre Brown This article discusses the principle of tūrangawaewae and its importance in Māori culture, highlighting how modern architecture and design projects incorporate tikanga and kaitiakitanga values. It provides a broader context for understanding how AI and other technologies can align with these principles.

Q: What is the rights and licensing status of Al-generated content?

A: New Zealand is still adapting its intellectual property laws to AI advancements, and discussions are ongoing about how to regulate AI-generated content fairly. Talk to our Rights team with any questions.

6. Ensuring human oversight, transparency and accountability so that content created or modified by AI must be reviewed by a qualified human to ensure the robustness of the output and is reliable, fair, and consistent with organizational standards.

Explanation

While AI has made considerable progress in recent years, it is far from infallible and cannot be used without oversight, this includes considering how others we may be working with may be using AI. It also requires clearly communicating when AI is being used, how it is being used, and what tools are being used (including tool versions if applicable) for any use of AI for either internal or external purposes.

Examples

- You run a Generative AI prompt to create first-pass descriptions of newly digitised collection items. The draft stays in "pending" until you (or another human) confirm factual accuracy, tone, and cultural sensitivity. The public record notes: "This text was created using Generative Artificial Intelligence and has been verified by Te Papa."
- A conversational AI chatbot answers website queries about opening hours, ticket prices, and collection highlights. A Te Papa team reviews weekly chat logs and corrects any errors. The chat window header states: "This assistant uses AI technology and is regularly audited by Te Papa staff—feedback welcome."
- You run a text-analysis AI model to group thousands of feedback comments by theme. A human then manually audits a 10 % sample for accuracy and screens out personal data before reporting. Reports note the use of AI and human verification.

Q&A

Q: Is there a standard format for acknowledging the use of AI?

A: The Library team has put together guidelines for acknowledging the use of AI in writing.

Also below are some general suggestions for acknowledging the use of AI in out outputs.

For outputs where genAl was queried for information or provided insights for a person producing the output:

- "informed by [tool*]"
- "some information sourced through [tool]"

For outputs where genAl provided a starting point that was improved upon by a per:

- "written based on a first draft generated by [tool]"
- "developed with the support of [tool]" "edited by [tool]"
- "redacted by [tool] with human review"

For outputs that were entirely generated by genAI with human review:

- "generated by [tool]"
- "summarised by [tool]"
- "processed by [tool]"
- "transcribed by [tool]"
- "image generated by [tool] with the prompt [prompt]"

For complex research, audience facing, of collection based use of AI, you should follow best practice including keeping a record of your prompts and outputs, and also tracking cost and possibly carbon use. More advice to come. Contact the AI group for more advice.

Q: Can I add Al-generated content into Emu?

A: Yes, AI-generated content can be added to EMu, but only with the explicit permission of the Manager, Digital & Collections Information Systems. Before inclusion, the content must be reviewed and verified by a qualified human. The Digital Collections and Access Team with work with you depending on your use case.

Al-generated content must also be clearly attributed. This includes:

- Tagging the AI platform as the creator using its associated Parties record (e.g., OpenAI).
- Adding a clear statement in the relevant text field.

^{*}When referencing [tool], state the name and version of the tool.

For example, if you use the GPT-40-mini model to generate collection item descriptions:

- Attach the OpenAl Parties record to the 'Described by' field.
- Include a statement in the Description field such as:

"This text was created using Generative Artificial Intelligence and has been verified by Te Papa. It was generated using the GPT-4o-mini model from OpenAI."

Q: What makes someone a qualified human reviewer?

A: They know the subject well (e.g. curatorial or language expertise) and understand Te Papa's style, ethics, and cultural protocols.

Q: Who's responsible for Al-generated content?

A: The kaimahi or team who reviews and signs it off. Al is just a tool—people remain accountable.

Q: Can Al be used without human input?

A: No. A person must start the task, monitor it, and approve or audit the output.

Q: How often should AI outputs be audited?

A: Regularly for routine tasks like chatbot replies. Do a spot-check after any major AI update or policy change.

Q: Do I need to note AI use in internal documents?

A: Yes. Mark them "Draft generated with Al—human-reviewed" to keep things clear for others.

Q: Should I keep records of AI prompts and outputs?

A: Yes. It helps with accountability, reproducibility, and future audits.

Q: How do I make sure external partners follow our AI approach?

A: Let them know, review their terms, and check in with IT if unsure.

Q: When to acknowledge I have used AI

A: Generating content yes, spell check no.

7. Informed use and serving the public AI so that use and processes should align with the Rautaki and our policies ensuring that they do not unjustly harm, exclude, disempower, or discriminate against individuals or particular groups, while actively reducing bias and supporting data ethics.

Explanation

Harm can come from reinforcing AI bias, consciously or unconsciously, which it is why it is important to be very considered and deliberate in thinking about potential impact. Most public LLMs are trained on the data available online and in digitally available forms, which is already biased to English language, Western concepts and information, and includes various forms of prejudice, stereotypes, and imbalances in representation. By using these potentially biased models, they can impart that bias onto our work, therefore further propagating that original bias.

To avoid this, in addition to following Principle 6: Ensuring human oversight, transparency and accountability, there are other mitigations that can assist avoiding bias.

Examples

Below are examples of possible bias

- Using AI to translate into te reo Māori or Pacific languages. The accuracy of these language translations within mainstream AI is still inconsistent, and needs human expertise and verification from relevant Te Papa kaimahi.
- If using AI to come up with ideas for exhibitions etc, AI is more likely to suggest Western ideas that are better represented in existing data sets, or seen as more "popular", lessening the likelihood of other ideas, and in turn reinforcing the bias
- Using AI to do rostering may not consider context like childcare needs, balance of skills across a shift
- Using AI to generate images of people

Here are other types of bias: selection bias, confirmation bias, measurement bias, stereotyping bias, out-group homogeneity bias.

Example scenario of fairness, equity and GenAl

(from Bias, discrimination, fairness, equity and GenAl | NZ Digital government

You've been using GenAI to source information that will be used to make decisions about who to prioritise for support.

You should ask:

- an expert in the field and/or
- members of the community that the information is about.

This helps you test that it's correct, factual and unbiased.

- You should also cross-check with official and authoritative sources about the potential for harm if the information is at risk of being incorrect or biased.
- After completing your checks, you conclude that the sources are authoritative, accurate and balanced.

Confident that your sources are not unintentionally biasing the output, you proceed to use the GenAI provided sources in your work and test the outputs.

Q&A

Q: How can I tell if something is biased?

A: Here are a few articles that give further help in navigating AI bias

- Bias, discrimination, fairness, equity and GenAl | NZ Digital government
- Navigating the Al Bias: Exploring Tools and Techniques | by Aruna Pattam |
 Medium
- Bias in Al | Chapman University

Q: What if AI produces content that is biased or culturally insensitive?

A: Pause, redact or discard the output; failure to do so can erode public trust. You may also choose to escalate to the provider for example, ChatGPT has a form to report bias.

8. Considering environmental impacts when using AI so as not to further impact negatively on Papatūānuku

Explanation

Like other uses of technology at scale (the internet, smart phones, cars, etc), the creation of AI models, and use of the resulting tools, can be energy intensive, and therefore carbon emission intensive, and make heavy use of water for cooling. As a

result, when using AI tools, the smallest tool able to do the task to the appropriate quality should be used.

Kaimahi can understand how AI sits within Te Papa's broader approach to sustainability as part of the rautaki and its commitment to Papatuanuku.

Examples

- You are working on mahi that is directly associated with environmental impact, and as part of the kaupapa, the team is actively looking to minimise carbon footprint as a core objective for the project. Team members ensure that they understand their project needs and requirements, discuss how AI may be used as part of the project, including its environmental impact. Then incorporate training to upskill team members and identify an approach or method for AI use to ensure they are using the most efficient or least 'compute' effort.
- You are undertaking an externally funded project where sustainability approach and measures are incorporated into the criteria for eligibility, how can you ensure that your use of AI as part of the project aligns with the criteria.
- A team is looking to automate a laborious image editing process and think that using AI could help. Team members discuss the use of AI and ask, "Could we do this without AI?" they decide to talk to a member of the technology team to check if there is a more efficient way to automate the task.

Q&A

Q: How can I understand the impacts of AI from a Papatūānuku perspective?

A: You can read about the environmental impacts of AI from a New Zealand perspective in the report titled "AI for the Environment in Aotearoa New Zealand" by the AI Forum.

Q How can I assess the impact of my actions on the environment?

A Pinpointing the exact environmental cost of every AI action is complex. AI systems require considerable energy for processing and cooling servers. While Te Papa is working to choose efficient tools and understand their impact, you can help by being thoughtful. Consider: 'Is this AI task necessary?' Simpler tasks often have lower impacts. Your mindful use contributes to our collective effort to respect Papatūānuku.

Understand the main causes of environmental impact:

 Most of an AI model's energy and resource consumption comes from its initial creation (training) Typically, this is not something that Te Papa would

- do. This environmental impact has already occurred when kaimahi use Al models.
- Use of Image generation and very large text generation models (inference)
 can be resource intensive, although less so than training new ones.
- o Water impact of up to a 1L of water used per prompt in data centre cooling.
- Impact of power use can cause emissions if a data centre is powered by fossil fuels, consider the grid intensity where the data centre is located.

You can read more about the environmental impacts of AI from a New Zealand perspective in the report titled "AI for the Environment in Aotearoa New Zealand" by the AI Forum.

Q What could be some ways to minimise my environmental impact whilst using AI at work?

A: Choose lightweight alternatives

 Models often have variants called mini, light or turbo which are less energy intensive to use.

Seek expert advice

o If you are using AI to help you with something you aren't familiar with, first ask someone who you think might be an expert in this area, they might have a non-AI and therefore more environmentally friendly solution to your problem.

Reduce use

- The more queries/tasks you make, the more energy will be consumed overall, though individual queries typically consume small amounts of energy.
- Intensive models such as image and audio generation or transcription services use the most energy. Disable automatic AI services such as meeting notetakers if not needed.

Q How do I measure the environmental impact of specific AI systems.

A It's often very difficult to assess the impact of a specific solution. There are many considerations from the size of the prompt type of hardware, location, and backend efficiency.

Q Can I opt out of using AI if I prefer not to use it due to environmental impact or will everyone at Te Papa be using it, so I don't have a choice?

A The use of AI is not mandatory. While AI tools are available to assist and enhance your work, participation is optional. You can choose whether or not to use AI based on your needs and comfort level, however it's useful to remember that we already have a proportion of AI built into the systems that we use such as MS Word and Co-Pilot.

Appendix A - Definitions

The table below outlines the various terms, definitions and concepts that are referenced in this policy, and frequently used in Al. These definitions come from the <u>Glossary of Al terms | NZ Digital government</u>.

Term	Definition			
AI- Generative (GenAI)	Generative AI is a type of AI system that can create or generate new content such as text, images, video and music based off models and patterns detected in existing datasets. (OECD)			
Al system	An Al system is a machine-based system 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 Al systems vary in their levels of autonomy and adaptiveness after deployment.			
Data ethics	Data ethics refers to the study and practice of ethical issues related to data, including its generation, recording, processing and distribution, and use. It encompasses principles and standards that guide the responsible and fair handling of data to ensure the rights and privacy of individuals are protected. Key principles of data ethics: Privacy: Ensuring that personal information is collected, stored, and used in ways that protect individuals' privacy and comply with legal requirements. Transparency: Being open about how the data is collected, used, and shared. Consent: Obtaining informed consent from individuals before collecting their data. Security: Protecting data from unauthorised access, breaches, or cyberattacks to maintain its confidentiality, integrity, and availability. Fairness: Ensuring data practices do not result in discrimination or bias and that data will be used in ways that are both equitable and just. Accountability: Holding organisations and individuals accountable for their data practices and ensuring there are mechanism in place to address any issues if they arise.			
Human oversight	Human in the loop refers to the involvement of human oversight and decision-making in the processes that involve AI and automated systems. This approach allows for critical decisions, especially those impacting individuals, to be reviewed, verified, and influenced by human judgement and expertise			
Māori Data Governance	 Māori data governance refers to the principles and practices that ensure Māori data is collected, managed, and used in a way that respects Māori values, rights, and interests. Key aspects of Māori data governance include: Data Sovereignty – ensuring Māori data is subject to Māori governance and control. Ethical use, promoting the ethical use of data to enhance wellbeing of Māori people, language, and culture. Advocating for Māori involvement in the governance of data repositories and decision-making processes. Safeguarding the quality and integrity of Māori data. 			