Use of generative Artificial Intelligence (AI) Policy

1. Purpose

This is the first iteration of the Commission's policy on the use of generative AI in the Commission. It should be read in conjunction with the Commission's Acceptable Use of Digital Technologies Resources.

The Commission is keen to enable and encourage the use of emerging technology, including generative AI, to enhance and support our work. Such use must however comply with relevant legislation and policy, including in relation to security, privacy, and ethical and responsible use.

2. Context

The policy has been informed by the Interim guidance - government use of generative AI tools and Adoption of Artificial Intelligence in the Public Sector issued by the Digital Transformation Agency (DTA) in collaboration with the Department of Industry, Science and Resources (DISR).

It will be reviewed and updated as necessary to ensure it remains consistent with relevant APS wide policy, guidance, and legislation.

3. Principles of use

The DTA's interim guidance provides guidance to assist APS agencies and staff adhere to Australia's AI Ethics Principles when using generative AI tools. The guidance sets out five principles (summarised at Appendix A) designed to support the responsible, safe, and ethical use of generative AI in the APS:

1	**Accountability**
2	**Transparency and explainability**
3	**Privacy protection and security**
4	**Fairness & human-centred values**
5	**Human, societal, and environmental wellbeing**

Any use of AI in the Commission must be consistent with these principles.

Further direction on the use of AI in the Commission is provided in the 'In Practice' section.

4. Governance

Generative AI may present opportunities for improvement and innovation in the ways we do our work. There are however challenges and risks around the use of AI tools which need to be understood and managed.

Oversight of the use of AI in the Commission will be incorporated into existing governance and accountability structures relating to outputs, risk, and compliance elements across the Commission's entire work program, not solely into ICT governance arrangements.

5. About generative AI tools

As noted above, generative AI presents opportunities and also poses risks. It is a rapidly evolving technology, and we are all learning about its capabilities and limitations. Generative AI is not infallible. Some of the risks and limitations to be considered include:

- Inaccuracies - outputs may contain information that is incorrect, unverified, out of date, or include fictional details.
- Outputs may use the intellectual property of others or replicate copyright protected work.
- Bias - AI tools are trained on large data sets which may be incomplete, incorrect, unrepresentative or reflect societal prejudices. As a result, outputs from these tools may reproduce biases present in the training data.

Publicly available AI tools

ChatGPT (OpenAI), Microsoft Bing, Google Bard, Dall-E are all examples of publicly available generative AI tools. They are not authorised for use in the Commission.

6. Review

Version	Comments	Author	Effective Date
1.0	Initial draft		February 2024
2.0	Added procedural notes including form/ process requesting use of any generative AI tools not authorised for general use - refer In Practice Notes.		March 2024

7. Approval

Consultation	Approved by	Name	Date
Management Committee	Chair	Danielle Wood	21/02/24

In Practice - Using generative AI tools in the Commission

These notes are designed to provide direction on the practical application of the policy.

Commission authorisation of AI tools/systems

- Consistent with the Commission's Acceptable Use of Digital Teachnolgies Resources policy, only authorised AI tools may be introduced and used in the Commission's ICT environment.

- The Commission will maintain a register of generative AI tools approved for use in the Commission.

Currently Authorised:

Microsoft Copilot (previously called Bing Chat Enterprise)	The Commission has the appropriate licences to use Microsoft's Copilot - an AI chat for the web, with the same functionality as tools like ChatGPT. Unlike ChatGPT and other similar tools however, Microsoft advise that data is protected, chat data is not saved or used to train the underlying large language models.
<u>**Trial**</u> **of Copilot for Microsoft 365** The Commission has agreed to participate in a trial of this AI tool which is being conducted by the Australian Government in partnership with Microsoft, through the Digital Transformation Agency (DTA).	**Copilot for Microsoft 365** is a different product to Microsoft Copilot. It is installed within the Microsoft Office applications and can assist users in various tasks and workflows within the Microsoft 365 suite. Data is kept secure within our tenancy.

Not Authorised:

Publicly available AI tools* Including: ChatGPT (OpenAI), GitHub Copilot, Microsoft Bing, Google Bard, Dall-E.	In consideration of the DTA recommendations, and the risks associated with the use of these AI tools, including in relation to privacy and information security, these are not authorised for use in the Commission.

Authorised via approval process:

- Staff may seek approval for use of a particular generative AI tool (publicly available or otherwise) by submitting a request to the Assistant Commissioner, Corporate Group.
- Depending on the circumstances, approval for use, if given, may be limited to a certain number of staff in the Commission.

If you wish to submit a request for use of an AI tool/platform (other than Microsoft Copilot) please complete this form: Request for use of AI tool.

*Note: If staff decide to use publicly available AI tools for personal use (i.e. for non-work-related purposes, using their own devices and systems) they must:
- not include any Commission or other government information or references
- not use Commission email addresses and/or passwords
- not introduce any content generated by such use to Commission ICT systems.

8. Using generative AI tools

1 Apply the DTA's **two golden rules when using generative AI** in your work:

1. You should be able to explain, justify and take ownership of your advice and decisions.

2. Assume any information you put into **any** generative AI tool could become public. Don't input anything into any tool that could reveal classified, personal, confidential, or otherwise sensitive information*.
   * This includes, for example, meeting notes from stakeholder engagement meetings and visits. Under no circumstances are these to be input/uploaded to **any** AI platform.

2 **Be open and transparent** about **how you are using AI in your work**. This includes sharing this information with your manager. This will among other things, support the Commission in meeting its responsibility to be clear when AI is being used to inform its activities.

When submitting work for review and/or approval, including material which may be distributed or published internally or externally, please provide the acknowledgement below:

> I acknowledge that this material was prepared using the assistance of outputs from *[insert name of generative AI tool(s) used (e.g Microsoft Copilot)]*. I have checked the material used for accuracy and quality.

3 Ensure you have approval to use any generative AI tool or system in the Commission ICT environment.

5 Implementation and use of any AI tool or system must support Commission business outcomes rather than business outcomes being driven by the adoption and use of AI.

5 Always carefully review outputs for accuracy, quality, and any bias before any use. AI outputs should be treated as tools to complement human judgement and expertise – not replace them. Remain critical and aware of the limitations and risks associated with the technology.

5 Consider intellectual property rights of third parties and copyright issues when using AI tools. (It is understood too, that at least at present, content generated by AI tools is not protected by copyright under Australian law).

5 When using any AI tool, ensure you:

- be very cautious about any links provided or generated by AI tools
- treat with care any files generated which have the potential to contain malicious code, such as macros in Microsoft Office files – these must not be used or distributed until vetted by the Digital Technologies team.
- do not input large government datasets.

5 Ensure your use of AI aligns with Australia's AI Ethics Principles.

AI User Group

An AI user group will be formed to support the practical application of the principles of AI use, share information about use cases in the Commission, and facilitate learning and knowledge sharing.

Membership of the group should include at least one representative from each work group. The group will be coordinated by the Director, Digital Technologies, and overseen by the Assistant Commissioner, Corporate Group.

9. Some definitions

Artificial Intelligence (AI)	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.
Generative AI	Takes its name from its **capacity to generate novel content**, as varied as text, image, music and computing code, in response to a user prompt.
	For example, conventional AI can be used to analyse features of a legal contract, such as whether the contract deals with intellectual property or privacy.
	By contrast, generative AI can be used to generate (i.e., draft) a new legal contract to cover those issues.
Large language models (LLMs) and **Multimodal foundation models (MFMs)**	LLMs and MFMs use machine learning algorithms to predict an output – such as an image or word – based on an input, such as a sequence of words.
	• **LLMs** specialise in generating human-like text by training on vast quantities of text.
	• **MFMs** are more complex as they use a wider range of information, including images, speech, numerical inputs and code and they are trained on the relationship between the various inputs.
	Note: What all these models do is recognise patterns in data and produce sophisticated answers based on those patterns.
	The models are not intelligent or able to necessarily determine fact from fiction in their inputs or training data.
Use Case	How, for what purpose, and in what context AI is used - or has the potential to be used. (This term is used widely, including in DTA and DISR advice and guidance material).
Publicly available generative AI tools	Widely available third-party AI platforms, tools or software that takes user input and uses generative AI to create output. Examples include: ChatGPT (OpenAI), MidJourney, , Microsoft Bing, Google Bard.

Principles - Summary from DTA Guidance

Appendix A

1. Accountability

- Accountability is one of Australia's AI Ethics Principles and one of the five APS Values . To remain open and accountable to the Australian community, users should be able to explain, justify and take ownership of their advice and decisions.
- AI tools can produce a convincing but inaccurate response to a query. A human with the appropriate skills, knowledge or experience should review any AI output before it is used .

2. Transparency and explainability

- It should be clear when generative AI tools are being used by government to inform activity.
- Users should consider including markings on briefings and official communications indicating if AI was used to generate any of the information.
- Users need to critically examine outputs to ensure their advice and decisions reflect consideration of all relevant information and do not incorporate irrelevant or inaccurate information.

3. Privacy protection and security

- Inputs into AI tools must not include or reveal classified information, or personal information.
- All activities need to align with legislation and policies relating to information and data (for example the *Privacy Act* and Protective Security Policy Framework).
- Government information must only be entered into these tools if it has already been made public or would be acceptable if made public. Employees determining if the information is suitable to be entered must have the appropriate delegation to do so.

4. Fairness and human centred values

- Generative AI tools may reproduce biases present in their training data. Before using AI-generated outputs, users should consider whether there is a process to ensure that outcomes are fair and meet community expectations.
- Again, users should be able to explain, justify, and take ownership of their advice and decisions.

5. Human, societal and environmental wellbeing

- Users should engage with generative AI tools to better understand the potential benefits and risks of the technology.
- Weigh the benefits and risks of use - traditional tools may be cheaper, safer, or better suited to the task.
- Consider intellectual property rights of third parties as well as broader copyright issues.
- Consider whether any use of Indigenous data and generative AI outputs is consistent with the expectations of First Nations peoples, particularly around Indigenous data sovereignty and governance, and with the forthcoming APS-wide Framework for Indigenous Data and Governance .