Guidance for the safe and responsible use of Generative AI in the Victorian Public Sector

# Overview

The Victorian Government is committed to digital transformation, innovation and community safety. Generative artificial intelligence (Generative AI) tools offer opportunities to enhance public sector productivity and the way the VPS delivers policies and services.

This document supports the [*Administrative Guideline for safe and responsible use of Generative AI in the Victorian Public Sector*](https://vic.gov.au/administrative-guideline-safe-responsible-use-gen-ai-vps)(the Generative AI Guideline). This documentprovides guidance on the use of Generative AI tools in Victorian Government, and applies to the use of publicly-available and agency-approved Generative AI tools for official work purposes.

Given the pace at which Generative AI is evolving and its potential uses, this guidance provides general advice only and should not be regarded as legal advice. You should always consult your organisation’s legal area where required.

### About Generative AI

AI is a technology characterised by its ability to perform complex tasks that historically only a human could do, such as reasoning, learning, problem ‑solving, and decision-making. Generative AI is subset of AI that refers to any tool that generates new content and information including text, images, code and audio in response to user-prompts.

Generative AI works by learning patterns and relationships in information it is trained on, then using that knowledge to predict plausible material in response to user prompts. The accuracy and usability of materials created using Generative AI relies on the quality of prompts and the tools’ algorithm and training material.

### Opportunities for the use of Generative AI

When managed safely and responsibly Generative AI provides for the possibility of substantial productivity gains and innovative solutions in the VPS. Uses for Generative AI include:

* Content generation: creating first drafts of content such as proposed topics for a speech, marketing materials and articles
* Creative thinking: by suggesting ideas for design and concepts, such as infographics and presentations
* Accessibility of information: such as language translation and improving readability by simplifying text
* Personalised and accessible services: resolving questions more efficiently and consistently by integrating chat features with existing systems.

### Who is this guidance for?

This guidance is for employees, contractors, consultants and volunteers engaged directly or indirectly by public sector bodies and public entities as defined in the *Public Administration Act 2004* with access to public sector information (public sector personnel).

The *Administrative Guideline for the safe and responsible use of Generative Artificial Intelligence in the Victorian Public Sector* sets out the implementation approach for organisations in applying the guideline.

### Applying Australia’s AI Ethics Principles

Victoria agreed to the [National framework for the assurance of artificial intelligence in government](https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government) (the National AI Assurance Framework) in June 2024. The framework provides a nationally consistent approach to assuring AI in government and assists applying [Australia’s AI Ethics Principles](https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles.).

The Department of Government Services is developing a tool to apply the National AI Assurance Framework in the Victorian Public Sector.

Prior to the release of the Victorian Government’s AI Assurance Framework, public sector personnel are encouraged to make use of the National AI Framework.

### Definitions

Refer to the definitions provided in the Generative AI Guideline.

# Why safe and responsible use of Generative AI is important

While presenting new opportunities, Generative AI also presents limitations and increases risks that require careful management. In particular, content and information generated by Generative AI may not be accurate, fair or transparent.

Generative AI does not understand context, content or meaning beyond the prompt provided by the users. Generative AI models learn from their training data, and if training datasets contain biased or inaccurate information, generated content can inherit these biases and inaccuracies. Additionally, there is limited, or no, way to explain how Generative AI content is produced.

Increased risks also apply to the information you use when generating content. When using Generative AI tools personnel should be mindful that information used does not exceed the protective marking determined by their agency as appropriate for that tool. Misuse of data can cause an [interference with privacy](https://ovic.vic.gov.au/privacy/for-the-public/privacy-complaints/) which is defined in the *Privacy and Data Protection Act 2014*.

Personnel need to be aware of Generative AI limitations and have checks and assurances in place to ensure safe and responsible use.

# Considerations for public sector personnel in applying the Generative AI Guideline

The section outlines considerations you should take before using either publicly-available or agency-approved Generative AI tools for official work purposes. This list may be complemented by specific guidance from your employer in relation to organisation- or sector-specific policies or requirements.

* Review and comply with the most up to date Generative AI Guideline.
* Review and comply with any relevant policies and guidelines from your organisation. These can generally be found on your organisation’s intranet.
* Complete any relevant Generative AI training offered by your organisation.
* Review any materials you intend to use with the Generative AI tool to ensure that they do not include information that is unsuitable for sharing with the tool. This applies to both any files you intend to upload or information you include directly in your prompt.
* Carefully structure your prompt.
* Review and have consideration to the practical guidance when using Generative AI tools.

# Updates to this guidance

This guidance will be reviewed and updated by the Department of Government Services concurrently with the Generative AI Guideline.

Material changes to this guidance will be approved by the Whole of Victorian Government Artificial Intelligence Interdepartmental Committee.

Suggestions for changes to the guidance can be provided to the Department of Government Services via email to [vicgov.ciso@dpc.vic.gov.au](mailto:vicgov.ciso@dpc.vic.gov.au) and will be considered in the context of the next review.

# Where to go for help

For queries on this guidance, please contact the Cyber Security, Data and Digital Resilience Division, Department of Government Services, at [vicgov.ciso@dpc.vic.gov.au](mailto:vicgov.ciso@dpc.vic.gov.au).

The following bodies can assist with guidance regarding VPS use of Generative AI:

* The [Victorian Public Sector Commission](https://vpsc.vic.gov.au/) can help with issues relating to the [VPS Code of Conduct](https://vpsc.vic.gov.au/ethics-behaviours-culture/codes-of-conduct/code-of-conduct-for-victorian-public-sector-employees/) and the [Code of Conduct for Directors of Victorian Public Entities](https://vpsc.vic.gov.au/ethics-behaviours-culture/codes-of-conduct/code-of-conduct-for-directors-of-victorian-public-entities/).
* The [Office of the Victorian Information Commissioner](https://ovic.vic.gov.au/) can assist with data protection and personal information matters, including [understanding privacy obligations when using AI](https://ovic.vic.gov.au/privacy/resources-for-organisations/artificial-intelligence-understanding-privacy-obligations/).
* The [Health Complaints Commissioner](http://www.hcc.vic.gov.au) can help with the handling of health information.
* The [Victorian Equal Opportunity and Human Rights Commission](https://www.humanrights.vic.gov.au/) can assist with matters relating to the Victoria’s Charter of Human Rights.
* [Public Record Office Victoria](https://prov.vic.gov.au/) can provide advice on [AI and recordkeeping](https://prov.vic.gov.au/recordkeeping-government/a-z-topics/AI), including its [recordkeeping policy](https://prov.vic.gov.au/sites/default/files/files/documents/ai_tech_and_recordkeeping_policy_v1_2024.pdf).

# When using Generative AI tools

Public sector personnel are to have regard to the following practical guidance when using Generative AI.

|  |  |
| --- | --- |
| Maintain accountability, transparency and explainability | |
| Generative AI tools do not understand real-world context and often produce compelling materials containing incorrect or incomplete information.  Government decisions are open to scrutiny including through formal merit and judicial reviews processes. Generative AI tools provide no, or limited, explanation on how materials are created. Relying on materials created by these tools without being able to explain their conclusions may be unfair and may lead to decisions not being upheld.  **When using Generative AI tools:** | |
| **Do** | **Don’t** |
| Remain responsible and accountable for content you create, share or use while performing your official work.  Fact-check information against verifiable sources.  Attribute any content to the appropriate Generative AI tool.  Be familiar with the Public Record Office of Victoria advice [Artificial Intelligence (AI) – Capturing and managing records generated by or using AI technologies.](https://prov.vic.gov.au/sites/default/files/files/documents/ai_tech_and_recordkeeping_policy_v1_2024.pdf)  [Maintain a record of the information y](https://prov.vic.gov.au/sites/default/files/files/documents/ai_tech_and_recordkeeping_policy_v1_2024.pdf)ou use to generate Generative AI tool outputs and the Generative AI generated content where they are incorporated into official documents. | Copy and paste sections of Generative AI content into your work without consideration of attribution and intellectual property obligations.  Ask these tools to answer a question you cannot independently validate the answer to.  Use these tools to make decisions, undertake assessments or use them for other administrative actions that may have consequences for individuals, groups or organisations. |

|  |  |
| --- | --- |
| Protect privacy and safeguard public sector information. | |
| You have an obligation to safeguard classified, confidential and personal information you use in your work. Entering such information into a Generative AI tool could cause an [interference with privacy](https://ovic.vic.gov.au/privacy/for-the-public/privacy-complaints/) or a data breach  Many existing publicly-available Generative AI tools are operated by private, foreign owned companies. Information provided to these tools may be stored or used overseas and outside the jurisdiction of Australian privacy and data security laws. It may also not be possible for public sector bodies to retrieve any information provided to these tools.  When using Generative AI tools: | |
| Do | Don’t |
| Only use publicly available information with publicly available Generative AI tools.  Be aware of the permitted purpose, privacy and security considerations of agency-approved Generative AI tools, and only input information in accordance with what is authorised by that agency.  Use privacy-enhancing options in tools where available (for example some Generative AI tools allow users to opt out of their data being used to train the vendor’s model).  Read and comply with the [public statements and guidance on AI](https://ovic.vic.gov.au/privacy/resources-for-organisations/) made by the Office of the Victorian Information Commissioner.  You should carefully review the information provided to a Generative AI tool as any information you provide may be re‑used or divulged to a third-party without a legitimate need to know. | Input, upload or share any public sector information not available in a publicly available publication into publicly-available Generative AI tools.  Input, upload or share any public sector information into agency-approved Generative AI tools that exceeds the protective marking determined as appropriate for that tool.  Disclose or input any information into an agency-approved Generative AI tool that does not comply with its permitted purpose and privacy and security assessments. |

|  |  |
| --- | --- |
| Be aware of fairness and human rights | |
| You have an obligation to respect and promote the human rights set out in the [Charter of Human Rights and Responsibilities](https://www.humanrights.vic.gov.au/legal-and-policy/victorias-human-rights-laws/the-charter/) by:   * making decisions and providing advice consistent with human rights; and * actively implementing, promoting and supporting human rights.   The accuracy and usability of material created by Generative AI tools relies on the tools’ algorithms and training material. Incomplete or inaccurate training data will lead to incorrect and potentially harmful AI-generated outputs that can be discriminatory, unfair, unjust or incompatible with human rights.  When using Generative AI tools: | |
| Do | Don’t |
| Fact-check information against verifiable sources.  Ensure that any use of Generative AI is consistent with your responsibilities under the Victorian *Charter of Human Rights and Responsibilities Act 2006*.  Be mindful of community expectations around Generative AI use, and consider any reputational risks from the Victorian Government being seen to use such tools in its work. | Assume or trust that outputs generated are unbiased, accurate or understand local context or nuance.  Rely on Generative AI as the only input to your work (it is just one input). It should not replace your own research, analysis and content development.  Allow Generative AI to make decisions for the Victorian Government. |

|  |  |
| --- | --- |
| Consider reliability and safety | |
| A Generative AI tool’s output depends on the quality of the prompts used as well as the tool’s algorithm training data. How prompts are phrased and the sequencing of questions and instructions can generate different outputs. Even identical prompts will often produce different results.  Generative AI models can also be used to automate and generate believable outputs that harm, deceive and damage (including at scale), for example generating content such as fake images, compelling lies and sophisticated phishing schemes.  When using Generative AI tools: | |
| Do | Don’t |
| Pay attention to the wording and sequencing of prompts to ensure the desired results.  Use agency-approved Generative AI tools ahead of publicly‑available Generative AI tools.  Use a trusted Generative AI user interface tool from a reputable vendor.  Review the terms of service for the tool and consult your legal team where appropriate.  Consult with your legal team on whether you can comply with a tool’s terms or opt out of where possible especially if they change from time to time. | Attempt to bypass a Generative AI’s built-in safety measures and ethical restrictions (jailbreaking) to produce an inappropriate response. |

|  |  |
| --- | --- |
| Consider legal obligations | |
| The Generative AI Guideline requires that you meet all legislative, regulatory and administrative obligations when using Generative AI tools for official work purposes. In addition to privacy legislation, you should also consider your obligations under the Victorian *Charter of Human Rights and Responsibilities Act 2006, Public Records Act 1973 and the Public Administration Act 2004*.  You should be aware that legal precedents around ownership and use of AI generated content and its relationship with copyright law are unclear. There are risks of breaching intellectual property laws, contractual obligations and legal professional privilege when using Generated AI tool in your work.  When using Generative AI tools: | |
| Do | Don’t |
| Meet all your existing employment obligations – including the VPS or Director Code of Conduct, privacy, security, human rights, anti-discrimination, administrative and other laws or policies.  When generating images, videos, or voice content, ensure that the solution provider provides legal coverage, including fair treatment of artists and respect for intellectual property rights.  Seek advice from your legal team about Generative AI tools where required. | Enter any information that may breach another person’s intellectual property rights or legal professional privilege.  Enter information that may attract claims of public interest immunity or executive privilege.  Copy and paste sections of Generative AI content into your work without consideration of what your obligations are in relation to attribution and intellectual property. |

# Example scenarios

This section provides example use cases for Generative AI tool. This is not an intended to be an exhaustive list of applications and should not be taken as an endorsement of specific tools or uses. The appropriateness of use case is based on individual circumstances and the nature of the content and information being used.

### Examples of safe and responsible use

| Helping to find a venue for an event | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| You have booked a venue for community consultation workshops which members of the public have nominated to attend, but with just one week to go the venue has cancelled your booking.  You need to quickly book an alternative venue for your workshop. | You ask the publicly-available Generative AI tool to provide you with a list of venues including contact details, capacity, average price and customer reviews.  You call the top three venues recommended by the Generative AI tool to confirm cost and availability before making a booking. | Yes. You used Generative AI to find and summarise publicly-available information.  You called the venue to confirm the information was correct before acting on the Generative AI tool’s recommendations.  Additional consideration  If the event being held was confidential or had specific security requirements, event details should not be entered into a publicly available Generative AI tool. |

| Help with a simple data analysis. | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| You are putting together a spreadsheet and would like to analyse the data.  The data in the spreadsheet relates to a project you are working on. The data does not include personal, sensitive or health information.  Your organisation has an agency-approved Generative AI tool that can be used for official work purposes where no personal, sensitive or health information is present. | You ask the Generative AI tool to analyse the data and provide any relevant insights.  You review the analysis provided by the tool to ensure it is working as intended. | Yes. Your use complies with the tool’s permitted purpose and does not contain any content or information that cannot be used in the tool.  You reviewed and validated the output of the Generative AI tool before using it.  Don’t forget to label the analysis as being created by Generative AI. |

| Summarising a report | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| A reputable thinktank has just officially released a 500-page report on a topic you are presenting to colleagues next week.  You have an incredibly busy week full of other deadlines and will not have the opportunity to read the full 500-page document and update your presentation. | You use a Generative AI tool to summarise key findings of the report and provide you with the references to the pages on which they appear.  You review sections of the report that the Generative AI tool has referenced to confirm the summaries provided are accurate.  When you give the presentation, you disclose that the summary was generated by Generative AI.  You advise your audience to review the report and to not rely on the summary provided by Generative AI. | Yes. The information in the report is public information. You are an expert on the topic so can identify incongruent information if it exists.  You also validated the summary back against relevant sections of the report to ensure it accurately reflects the content of the report.  You labelled the summary as being created by Generative AI. |

| Taking notes for a meeting | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| You are hosting an online meeting with your team to discuss a work project. The online meeting software you are using has been approved by your organisation for official work purposes.  You need to take minutes for the meeting but there is no one in your team available to help. | You use a feature in the online meeting software to record the meeting and take notes.  The software says that the feature is powered by Generative AI, so you ask everyone’s permission to turn on the feature.  You remind everyone in the meeting not to discuss any sensitive topics. | Yes. The Generative AI tool has been approved by your organisation for use, and your asked everyone’s permission first before using it.  Don’t forget to disclose to your team that you are recording the meeting.  Don’t forget to turn off the Generative AI tool if someone no longer consents to the recording.  Don’t forget to review the Generative AI tool’s meeting notes to ensure they are accurate. If you’re not sure, review the recording of the meeting.  Don’t forget to review access to the meeting to ensure that only those who require them have access. |

### Examples of inappropriate use

| Drafting an email | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| You receive an email from a difficult stakeholder with which you have been discussing a sensitive matter.  You wish to respond in a firm yet polite way, but you are not sure how to do it. | You copy and paste the email chain into a Generative AI tool and ask it to draft a response for you. | No. The information you provided to the Generative AI tool is sensitive and not publicly available. The tool may inappropriately share this information with other users. |

| Taking a group photo | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| You have organised the department’s executive board strategic planning retreat. At the end of the last day the participants gather for a group photo to commemorate the occasion.  Several executives had to leave earlier in the day, but you have access to their profile photos. | You upload the group photo and the absent executive’s individual photos into a Generative tool and ask it to produce a group photo that includes all participants.  You use the photo for an all-staff email about the event. | No. The photo is a record of the event, and it is not consistent with the VPS Code of Conduct to knowingly produce an incorrect record.  It may also be a breach of privacy to use someone’s photo without their consent. |

| Analysing responses to a request for tender | | |
| --- | --- | --- |
| The situation | Gen AI use case | Appropriate use? |
| You are on a procurement panel and been provided with several lengthy responses from vendors that detail their proposals, pricing and associated service catalogues.  You have been asked to analyse the proposals and rank the proposal to determine the best offer. | You upload all the documents to a publicly available Generative AI tool and ask for them to be compared and a report generated with a summary and rank for each proposal. | No. Not only do the proposal documents contain confidential information, you also have no way of explaining how the offers were compared and how the Generative AI tool arrived at the ranking for each vendor.  It is also unlikely that existing Victorian Government procurement and policies would have been adequately applied. |