Office of the Victorian Information Commissioner

Use of personal information with publicly available Generative AI tools in the Victorian public sector

This guidance outlines the Office of the Victorian Information Commissioner's (**OVIC**) position on the use of publicly available Generative Artificial Intelligence (**Generative AI**) tools within Victorian public sector organisations (**VPS organisations**).[1]

Generative AI is a type of Artificial Intelligence that creates content, such as text, images, music, audio and video. Generative AI tools do this by predicting the best response to the information entered, known as a prompt. Predictions, also known as outputs, are made based on the content of the prompt and the data on which the Generative AI tool has been trained.[2]

This guidance relates to VPS organisations' use of any Generative AI tool that is publicly available.[3] Publicly available Generative AI tools are platforms or software that can be accessed via a web browser or application. Generally, publicly available tools have minimal controls for how the information entered is used or protected. At the time of issuing this guidance, examples include free versions of ChatGPT, Google Gemini, Grammarly, Claude, Perplexity, Otter, QuillBot and Llama.[4]

The risks involved in using publicly available Generative AI tools are different to those involved in using tools that are purchased at an enterprise level. For guidance on using enterprise Generative AI tools in a secure environment, please refer to OVIC's guidance on the use of enterprise Generative AI tools in the Victorian public sector.

Personal information and publicly available Generative AI tools

VPS organisations should ensure their staff, contracted service providers and other personnel do not enter personal information into publicly available Generative AI tools.[5] Doing so will likely be a contravention of the Information Privacy Principles (**IPPs**) in the *Privacy and Data Protection Act 2014*

[1] For the purpose of this guidance, VPS organisations are organisations subject to Part 3 of the *Privacy and Data Protection Act 2014*. See section 3 of that Act for a list of organisations.

[2] For further information on Generative AI concepts, see the resource developed by the Centre of Excellence for Automated Decision-Making and Society, and OVIC: https://www.admscentre.org.au/genai-concepts/.

[3] This guidance replaces the guidance OVIC issued in its public statement on the use of personal information with ChatGPT in February 2024.

[4] The Australian Government and the Victorian Government have both issued a ban on DeepSeek.

[5] Personal information is defined in section 3 of the *Privacy and Data Protection Act 2014* as 'information or an opinion… that is recorded in any form and whether true or not, about an individual whose identity is apparent, or can reasonably be ascertained, from the information or opinion'. The definition of personal information excludes health information, which is covered by the *Health Records Act 2001*.

(**PDP Act**), and may cause significant harm to individuals whose information is used in publicly available Generative AI tools.[6]

Generative AI outputs can also reveal or create personal information. Personal information includes opinions about individuals. VPS organisations must treat outputs that contain personal information as a collection of personal information, and handle the information in accordance with the IPPs.[7]

Public sector information and publicly available Generative AI tools

VPS organisations should ensure their staff, contracted service providers and other personnel limit the public sector information that is entered into publicly available Generative AI tools.[8] VPS organisations should only enter public sector information that is already publicly known or approved for public release.[9]

The risk of using publicly available Generative AI tools with public sector information is context specific, and will differ depending on the Generative AI tool and its intended use.

Risks of using publicly available Generative AI tools

Using personal information with a publicly available Generative AI tool raises significant information privacy and other concerns. Some examples are outlined below.

- **Disclosure of personal information**: using personal information with a publicly available Generative AI tool may mean disclosing that information to the company that owns the tool. The information may be subsequently used, disclosed or accessed for unauthorised purposes unrelated to the original prompt. This may contravene IPPs 2.1 and 4.1, relating to use and disclosure of personal information, and information security.

  The information entered into a publicly available Generative AI tool is likely to be stored outside Victoria, and in many cases, in jurisdictions that do not have information privacy and information security laws equivalent to the PDP Act. This may contravene IPP 9, relating to transborder data flows.

[6] In November 2024, OVIC published a report on its investigation into the use of ChatGPT by a Child Protection worker. The report highlights a range of information privacy issues that stemmed from personal information being entered into a publicly available Generative AI tool: https://ovic.vic.gov.au/regulatory-action/investigation-into-the-use-of-chatgpt-by-a-child-protection-worker/.

[7] Comprehensive Guidelines to the Information Privacy Principles are available on OVIC's website: https://ovic.vic.gov.au/privacy/resources-for-organisations/guidelines-to-the-information-privacy-principles/.

[8] Public sector data, also known as public sector information, is defined in section 3 of the PDP Act as 'information (including personal information) obtained, received or held by an agency or body to which Part 4 applies'. Not all VPS organisations will have obligations under Part 4 of the PDP Act. However, OVIC recommends all VPS organisations limit the information they enter into publicly available Generative AI tools.

[9] This is consistent with the position taken in the Victorian Government's Administrative guideline on the safe and responsible use of Generative Artificial Intelligence in the Victorian public sector: https://www.vic.gov.au/administrative-guideline-safe-responsible-use-gen-ai-vps.

- Collection of personal information: generating personal information with a Generative AI tool constitutes a new collection of personal information. VPS organisations should be mindful of the necessity, lawfulness and fairness of collecting new personal information in this way. This may result in inaccurate information or opinions being generated and subsequently used or disclosed, which may contravene IPPs 1.1, 1.2, 3.1 and 10, relating to the collection of personal information, data quality, and sensitive information.

  IPP 1.3, which requires VPS organisations to provide a notice of collection to individuals when their personal information is collected, will also be relevant. It will be difficult for VPS organisations to provide adequate notice, where at the time of collection it is not known that personal information will be used with publicly available Generative AI tools, resulting in its disclosure and further use. Further, providing notice of collection where personal information is created by a publicly available Generative AI tool and therefore collected by the VPS organisation, may be difficult.

- Accuracy of information: it can be difficult to ensure the accuracy of personal information, for both that which is entered into a Generative AI tool and generated by it. Generative AI tools are not a reliable source of accurate information. They simply make predictions based on the prompt they receive and the training they've had. Relying on outputs that contain personal information may contravene IPP 3 in relation to data quality. Information accuracy issues also apply to non-personal information that is generated by a Generative AI tool.

- Information security: publicly available Generative AI tools raise a number of information security issues. For example, unauthorised disclosure of and access to personal information and other public sector information by third parties, which could lead to a contravention of IPP 4.1 and the Victorian Protective Data Security Standards.[10] This could lead to harm to individuals and VPS organisations, if the information is used for malicious intent.

- Retention of personal information: information entered into a publicly available Generative AI tool may be retained indefinitely by the company that owns the tool. Once information is entered into a publicly available Generative AI tool, it may not be possible to extract that information, leaving the VPS organisation without the ability to adhere to IPP 4.2.

  Generative AI tools are trained on vast amounts of data, including in many cases, information entered by a user.[11] Retaining this information is of benefit to the companies that own these tools, but where personal information is involved, may contravene IPP 4.2 relating to data

[10] Bodies that are subject to Part 4 of the *Privacy and Data Protection Act 2014* must comply with the Victorian Protective Data Security Standards: https://ovic.vic.gov.au/information-security/standards/.

[11] Some Generative AI tools provide users with the ability to disable the information they enter from being recorded, and opt-out of the information they enter being used to train the tool. Even if these features are disabled, privacy concerns around the disclosure of personal information and unauthorised access remain.

destruction. VPS organisations also have obligations under the *Public Records Act 1973* that must be adhered to.[12]

- Fairness: Generative AI tools must not be used to formulate decisions, undertake assessments, or for other administrative functions that may have consequences for individuals. The use of these tools with personal information may cause significant harm to those individuals whose information has been used. It risks unfair decisions being made about them based on information created by Generative AI, that may be inaccurate or of a diminished quality.

OVIC's minimum expectations

For VPS organisations using, or considering using, publicly available Generative AI tools, OVIC expects at a minimum they will:

- conduct security risk assessments before use, in accordance with their risk management framework to ensure that any risks are identified, assessed, treated and monitored accordingly[13]
- implement a formal process for the ongoing monitoring and review of risks and controls
- review the publicly available Generative AI tools' terms and conditions prior to use, and where possible, configure security settings to protect public sector information
- only permit public sector information that is already publicly known or approved for public release to be used with publicly available Generative AI tools
- require their staff, contracted service providers and other personnel to check the accuracy of generated content before using it
- develop organisational policies on the use of Generative AI tools, and communicate about the policies to staff, contracted service providers and other personnel
- train staff on how to safely and responsibly use Generative AI tools, including for what purposes publicly available Generative AI tools may be used, which tools are permitted, and what information is appropriate to enter into a publicly available Generative AI tool

[12] For further information on AI and recordkeeping, refer to the Public Record Office Victoria's guidance and policy: https://prov.vic.gov.au/recordkeeping-government/a-z-topics/AI.

[13] OVIC recommends that VPS organisations also conduct privacy impact assessments (**PIAs**) where a program or initiative involves the handling of personal information. As this guidance states that no personal information should be entered into publicly available Generative AI tools, PIAs will not be necessary if no personal information is handled. For guidance on PIAs, refer to OVIC's website: https://ovic.vic.gov.au/privacy/resources-for-organisations/privacy-impact-assessment/.

- notify OVIC if personal information or public sector information that is not already publicly known or approved for public release has been used with a publicly available Generative AI tool.[14]

VPS organisations are welcome to contact OVIC for guidance on managing the information privacy and information security impacts of publicly available Generative AI tools. OVIC can be contacted at enquiries@ovic.vic.gov.au.

Scenarios

The following scenarios have been prepared to illustrate the risks involved with using personal information in publicly available Generative AI tools.

Scenario 1 – Entering and generating content

A hiring manager is using a publicly available Generative AI tool to generate content for a recruitment selection report. They enter the CVs and interview notes captured by the panel, and the reference reports of five candidates into the Generative AI tool. They prompt the tool to generate six paragraphs recommending Candidate A for the role.

Trainers at the overseas company that owns the Generative AI tool subsequently review the content and prompt used with the tool, which is used to improve the model. This results in a disclosure of the personal information of five individuals, in contravention of IPP 2.1 and 4.1. As the content is stored outside of Victoria and retained by the company indefinitely, it is also a contravention of IPP 9 and IPP 4.2.

Scenario 2 – Using a publicly available Generative AI tool for decision making

A Victorian public sector employee is writing a report evaluating whether a prisoner should be granted parole, and uses a publicly available Generative AI tool to generate the content of the report, including the evaluation of risks. In doing so, they enter the prisoner's personal information. The company that owns the Generative AI tool stores its data outside of Australia.

The Generative AI tool generates an output in response to the prompt. The employee does not carry out their own independent evaluation and relies solely on the information contained in the output. The employee then submits the recommendation to their supervisor that the prisoner's parole application be rejected.

[14] VPS organisations should report information privacy and information security incidents to OVIC via the incident notification form: https://ovic.vic.gov.au/privacy/resources-for-organisations/information-security-and-privacy-incident-notification-form/.

The employee's supervisor reviews the report and notes that the reasoning for the opinion rejecting the parole application is flawed, and that the wrong recommendation had been reached. The supervisor rewrites the report themselves, undertaking analysis of the information to ensure its accuracy, and resubmits the report.

If it had not been for the supervisor's intervention, the use of the Generative AI tool could have had severe negative consequences for the prisoner. As the employee used a publicly available Generative AI tool, the information they entered into the tool resulted in an unauthorised disclosure of information. The use of the personal and sensitive information of the prisoner would contravene IPPs 2.1, 3.1, 4.1 and 9 at a minimum.

Scenario 3 – Using a publicly available Generative AI tool in a public school

A teacher uses a publicly available Generative AI tool to generate an email to a student's parents, expressing their concern about the student's behaviour and academic performance. In doing so, the teacher enters the notes they have taken about the student across the term to generate the output. The company that owns the Generative AI tool stores its data outside of Australia.

The teacher isn't happy with the tone of the output, so generates it again. In doing so, the Generative AI tool generates a new output with a different tone and different opinion about the student. The teacher then sends the content in an email to the student's parents.

The parents receive the email and raise concerns with the teacher about the opinions formed about their child, and question who wrote the letter. The teacher explains they used a Generative AI tool to generate the content. The parents are not satisfied the generation of the content about their child was necessary, and consider the use of the tool to be an unreasonably intrusive method of collecting information about their child.

As generating personal information using a Generative AI tool is considered a new 'collection' of personal information, there has been a contravention of IPPs 1.1, 1.2 and 1.3. Given the student's personal information was entered into a publicly available Generative AI tool, where data is stored outside Victoria, a contravention of IPPs 4.1 and 9 has also occurred.