Office of the Victorian Information Commissioner

Use of personal information with ChatGPT

OVIC Public Statement

Anique Owen, Assistant Commissioner – Privacy Guidance and Dispute Resolution

What we'll cover

- An overview of ChatGPT, and how it works.
- The privacy risks when using ChatGPT.
- Key considerations when using ChatGPT.

What is ChatGPT, and how does it work?

- ChatGPT is a generative artificial intelligence platform developed by OpenAI.
- It uses natural language processing to respond to a prompt from a user and generate human-like text, known as an ‘output’.
- ChatGPT can detect patterns, context and meaning from prompts and generates an output most likely to be related, word by word, until the stop condition is met.
- ChatGPT’s Large-Language Model (**LLM**) is trained on publicly availably data from books, webpages, news articles and other sources to detect patterns, context and meaning.

Privacy risks when using ChatGPT

Inputting personal and sensitive information into ChatGPT raises significant privacy concerns, and will contravene several Information Privacy Principles, including but not limited to:

- IPP 1, Collection;
- IPP 2.1, Use and Disclosure;
- IPP 3.1, Data Quality;
- IPP 4.1, Data Security;
- IPP 4.2, Data Retention; and
- IPP 9, Transborder Data Flow.

Scenario

A VPS employee is writing a report evaluating whether a Prisoner should be granted parole and uses ChatGPT to generate the content of the report, including the evaluation of risks. In doing so, they input the Prisoners personal and sensitive information.

ChatGPT generates an output in response to the prompt. The employee then submits the recommendation that the Prisoner's parole application is rejected to their supervisor.

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.

Considerations when using ChatGPT

1. Is this ‘personal information’?

2. If it isn’t ‘personal information’, is the information already publicly known, or if disclosed would not cause harm to an individual, or damage to an organisation?

Core messages when using ChatGPT

- VPS organisations must ensure staff and contracted service providers do not use personal information with ChatGPT.

- ChatGPT must not be used to formulate decisions, undertake assessments, or used for other administrative actions that may have consequences for individuals, for example, evaluations, assessments, or reviews. Doing so is a contravention of the Information Privacy Principles (IPPs), and may cause significant harm to individuals whose information is used with ChatGPT.

- If an organisation becomes aware that personal information has been used with ChatGPT it should treat the occurrence as an information security incident and notify OVIC immediately.

For privacy guidance, please email:

privacy@ovic.vic.gov.au