Cyber Security NSW generative artificial intelligence (AI)

End-user guidance

Version 2

28 August 2023

Department of Customer Service

NSW GOVERNMENT

Cyber Security NSW generative artificial intelligence (AI)

Purpose

The purpose of this guidance is to advise end users in NSW Government on public generative AI tools and what you should and should not do to ensure responsible, safe and ethical use in the workplace. This guidance will be updated and evolve with consultation and new developments in AI.

Users should follow their entity's internal policies on generative AI in the first instance.

Definitions

There is no single agreed definition of AI. This guidance uses definitions based on the International Organisation for Standardisation (ISO).[1]

What is AI?

AI refers to 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. AI systems are designed to operate with varying levels of automation.

What is generative AI?

Generative AI is a subset of AI that generates novel content such as text, images, audio and code in response to prompts. Generative AI technologies use large language models (LLMs) that specialise in the generation of human-like text.

What is a public generative AI tool?[2]

Any tool that is available to the public, takes user input and uses generative AI to create output. Examples include:

- ✓ ChatGPT (OpenAI)
- ✓ Copilot X (Github)
- ✓ BingAI (Microsoft)
- ✓ BardAI (Google)
- ✓ DALL.E
- ✓ MidJourney.

[1] The definitions of 'Artificial Intelligence (AI)', 'machine learning', 'algorithm' are based on the respective ISO definitions (ISO/IEC 22989:2022);

[2] https://architecture.digital.gov.au/generative-ai

Opportunities and benefits of AI

AI is already playing an important role across the NSW Government in delivering more efficient and improved services to NSW communities. Examples include medical diagnosis systems that use patient symptoms and medical history, Spatial Digital Twin to assist in town planning, and the introduction of driverless trains.

Public generative AI tools present numerous opportunities that the NSW Government can leverage, such as:

✓ content generation

for articles, blog posts, marketing materials

✓ idea generation

for creative concepts, taglines, training icebreakers

✓ research assistance

for research on specific topics, summaries, topic snapshots or overviews

✓ technical help

for troubleshooting or technical walkthroughs on specific Excel or Word tasks

✓ language translation

to translate text between languages or simplify complex text.

Risk considerations

While generative AI technologies present opportunities for increased productivity across the NSW Government, they also introduce ethical, privacy and security risks that should be considered and addressed prior to use.

Risks: Ethics

NSW Government employees should not utilise the outputs of public generative AI if it could result in negative impacts for NSW Government entities or NSW communities.

The information provided by public generative AI tools is often unverified, may not be factual, or may be unacceptably biased. All outputs must be fact checked. Irrelevant or inaccurate information must not be used for any official purpose. Additional assessments must be undertaken to ensure that the outputs reflect consideration of all relevant information and are ethical and responsible.

When using public generative AI tools, users must be able to explain and justify their actions and decisions. Humans must remain the final decision maker.

More information on ethical use of AI can be found in the NSW AI Ethics Policy:

https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-ethics-policy

Do		Don't	
Use responsibly and ethically.	✓	Use generative AI outputs that are: • unethical • irresponsible • biased • inaccurate • discriminative.	✗
Fact check and verify outputs before using for any official purpose.	✓	Allow generative AI to make decisions for the NSW Government.	✗
Ensure outputs reflect consideration of all relevant information.	✓	Use outputs that infringe on copyright or violate intellectual property rights.	✗

Use case 1

Bob is conducting a security assessment of a potential vendor. The report will be circulated with management to decide if the vendor is suitable. Bob wants to use Bard AI to research the vendor and generate a security questionnaire.

What should Bob do?

✓ Bob can use Bard AI to find information on the vendor or ask recommendations on the type of questions he can ask in a security assessment.

! Bob should take care not to input any official non-public details related to a NSW Government entity.

✓ The generated responses should be thoroughly validated by Bob or another colleague for appropriateness and accuracy before being used.

✓ Bob should remember that generative AI lacks the context and human industry expertise to provide analysis or opinion.

Risks: Privacy

When using public generative AI, users must comply with all applicable legislative requirements and laws. Any collection, storage, use and disclosure of information must comply with the Privacy and Personal Information Protection Act 1998 and the Health Records and Information Privacy Act 2002.[3]

Personal or health information should not be inputted or disclosed to public generative AI tools.

- **Personal information** is any information that identifies an individual such as written records that may include an individual's name and address, photographs, images, video or audio footage.
- **Health information** is any personal information or opinion about an individual's physical or mental health; health services provided to an individual or to be provided in the future; information collected in connection with organ donation; or other personal information that is genetic information about an individual arising from a health service provided.

There are a number of additional acts and regulations that promote the protection of personal and health information in NSW. More details are available in the NSW AI Ethics Policy: https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-ethics-policy/key-considerations

Do

Comply with applicable legislative requirements and laws. ✓

Seek advice from your privacy or security officer if in doubt or where guidance can't be followed. ✓

Don't

Disclose or input personal or health information. ✗

Use case 2

Nairn is developing a spreadsheet to calculate customer enquiries by enquiry type for his team. The formula being used presents an error message. Nairn plans to type the formula into Bing AI to find out why and how to fix it.

What should Nairn do?

✓ Nairn can ask Bing AI to check whether the formula is correct and troubleshoot how to fix it.

! Nairn must take care not to use real data or mention any organisational information.

✓ Employees can leverage generative AI to troubleshoot basic problems and errors in day-to-day tasks.

[3] https://www.ipc.nsw.gov.au/privacy/nsw-privacy-lawssotres

Risks: Data and security

NSW communities must have confidence that their data and information are always used safely and securely. Any unauthorised disclosure or data breach may undermine trust in the NSW Government.

Most publicly available AI platforms store, process and host data outside of Australia. NSW Government employees must ensure any data or information shared complies with all applicable legislative requirements or laws, such as the State Records Act 1998 (NSW).

https://staterecords.nsw.gov.au/recordkeeping/using-cloud-computing-services-implications-information-and-records-management

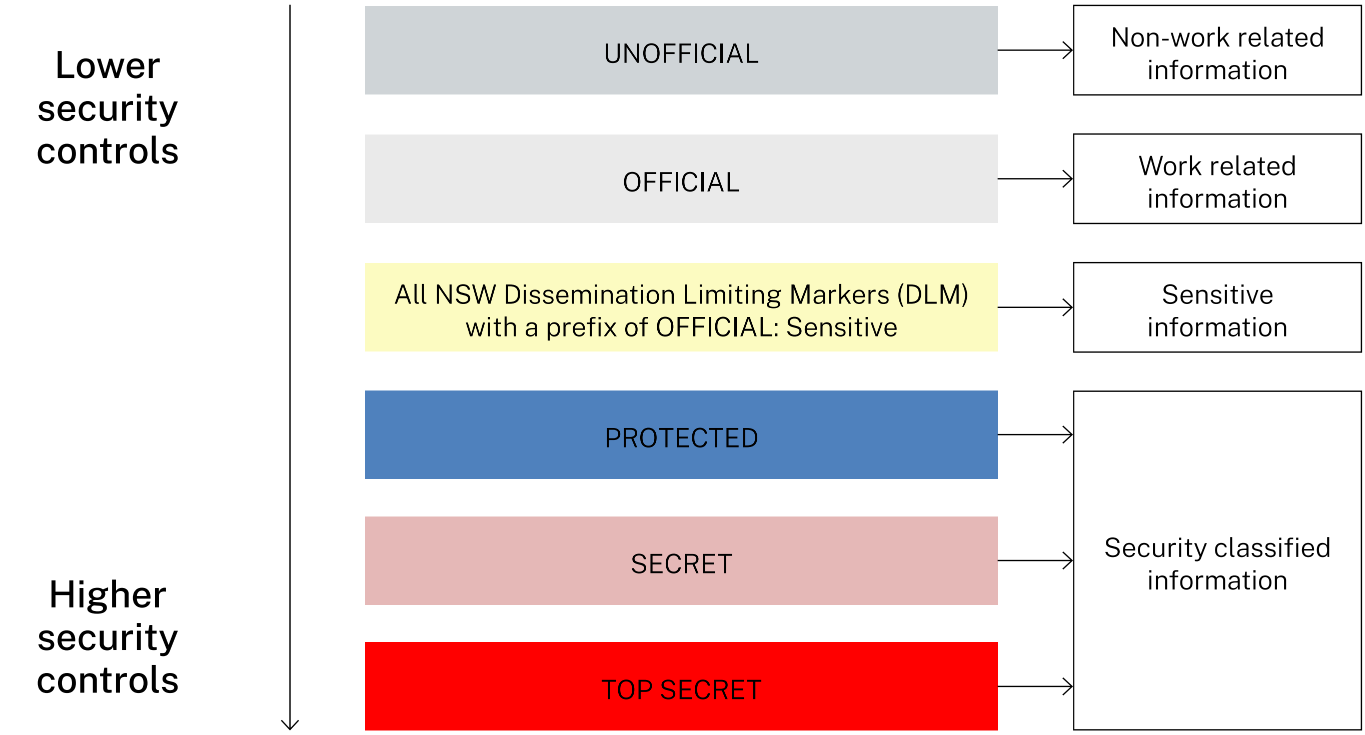
When using public generative AI tools, users must not input sensitive or classified information.

Official information may only be disclosed if it is already publicly available, or there is a reasonable expectation that the information is acceptable to be made publicly available. Employees determining whether the information in question is suitable for public release must have the appropriate delegation to do so.

NSW Government employees should not input information that would allow public generative AI tools to extrapolate classified or sensitive information based on the aggregation of content entered over time.

Unauthorised disclosure of NSW Government or Commonwealth classified information can lead to substantial fines and/or imprisonment.

NSW information labels



Lower security controls

Higher security controls

UNOFFICIAL → Non-work related information

OFFICIAL → Work related information

All NSW Dissemination Limiting Markers (DLM) with a prefix of OFFICIAL: Sensitive → Sensitive information

PROTECTED

SECRET

TOP SECRET

→ Security classified information

For more information on NSW information classification, labelling and handling guidelines:

https://data.nsw.gov.au/nsw-government-information-classification-labelling-and-handling-guidelines

Do		Don't	
Comply with applicable legislative requirements and laws.	✓	Input official, sensitive or classified information.	✗
If account creation is required, use a corporate email.	✓	Create an account unless registration is a requirement.	✗
Disable training and logging features.	✓	Open any AI-generated links.	✗
Disable chat history.	✓	Open AI-generated files.	✗
Enable multifactor authentication where available.	✓	Use unofficial generative AI websites, applications or plugins.	✗
Reference any AI-generated content.	✓	Use AI-generated code in government systems.	✗
Report to your Chief Information Security Officer (CISO) or security team if unsure, or where guidance can't be followed.	✓	Input or validate code from any government systems.	✗
		Input large government datasets.	✗
		Amend or downgrade NSW Information Labels.	✗

Use case 3

Alice is creating a report and PowerPoint presentation to present the results of an employee survey to the executive team. The results are labelled Official: Sensitive – NSW Government. She is short of time and plans to use ChatGPT to help create the content for the report and presentation.

What should Alice do?

! Alice must not upload any of the report or results, as it is labelled Official: Sensitive – NSW Government. Disclosure of this data into the public domain could be damaging to the NSW Government and its employees.

✓ Alice should take care when using ChatGPT to source content for official government documentation. Any quoted statistics should be properly verified and referenced. Alice should avoid clicking on any links generated by ChatGPT.

✓ A safer option would be for Alice to ask ChatGPT for presentation or report layout recommendations or ideas.

Users should follow their entity's internal policies on public generative AI tools in the first instance. If there is any doubt, contact your security team or CISO. For enquiries relating to this guidance, please contact: info@cyber.nsw.gov.au

Department of Customer Service

2/24 Rawson Place
Haymarket NSW 2000

Office hours:
Monday to Friday
9am - 5pm

E: info@cyber.nsw.gov.au
W: digital.nsw.gov.au/policy/cyber-security