GOVERNMENT OF
WESTERN AUSTRALIA

Large Language Models

WA Public Sector Guidance

```
#include <stdio.h>
int main() {

   int number1, number2, sum;

   printf("Enter two integers: ");
   scanf("%d %d", &number1, &numb

   // calculating sum
   sum = number1 + number2;

   printf("%d + %d = %d", number1,
   number2, sum);
   return 0;
}
```

What are Large Language Models[1]?

Generative AI models generate novel content such as text, images, audio and code in response to prompts.

A large language model (LLM) is a type of generative AI that specialises in the generation of human-like text.

Who is this guidance applicable to?

- All WA Government organisations and employees operating on behalf of the agencies.
- Non-government suppliers and personnel that access WA Government information and resources.

While this guidance is intended to be applied to 'open' generative LLMs, in which data may be stored by third parties, and prompts used to on-train the model, it is still relevant to private LLMs where data does not leave the tenant boundary.

What are some of the possible risks associated with using Large Language Models?

Some of the risks associated with LLMs relate to ethics and fairness, legal and privacy, data integrity and accuracy and protections of information privacy.

LLMs are trained on large amounts of data. Depending on the quality of this data, LLMs can produce inaccurate and biased or discriminatory results.

Inputs into models may be accessible to third parties or used to on-train the model which could result in data leakage.

Intellectual property and copyright of outputs should be considered, and agencies must ensure compliance with relevant policies and legislation.

When you are procuring or using an AI solution for the first time at your agency, or for a different purpose, you must undertake a self-assessment against the WA Government AI Assurance Framework.

[1] The definitions of 'generative AI' and 'a large language model' (LLM) are based on the definitions in Bell, G., Burgess, J., Thomas, J., and Sadiq, S. (2023, March 24). Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs).

DOs and DON'Ts for the safe, responsible and ethical use of LLMs in the WA Public Sector:

DO	Don't
Assume that inputs into online LLM AI tools will automatically be considered as available in the public domain.	Enter any personal information about any person.
Remain responsible and accountable for your work, including critically assessing all AI generated outputs and validating quality and accuracy with other sources.	Enter any health information about any person.
Thoroughly research and source reliable references for content where necessary.	Enter inputs or use outputs in a way that may breach another person's intellectual property rights.
Meet all your existing employment obligations- including in relation to handling official information, privacy, security, human rights, anti-discrimination, administrative and other laws	Use these tools for any query that is complex or sensitive, or where local context and nuance is critical.
Where required, attribute content that has resulted from the use of these tools	Ask these tools to answer a question you don't already know the answer to or cannot validate the answer to – you need the knowledge to decide whether it can be trusted or contains bias.
Regularly monitor and review uses of these tools to ensure they are being used responsibly ethically, safely, and according to law and government policy (including the WA Government AI Policy).	Use these tools to replace your own research, analysis and content development or embed them into your work in a way that means it cannot be done without them.
Opt out of the LLMs retaining and using your data for training purposes, if possible.	'Copy and paste' sections of AI-generated content into your work. If you do copy and paste any AI generated content, consider what your obligations are in relation to attribution and intellectual property.
Consider that third party services and vendors may store data and operate out of other legal jurisdictions which may have differing legislation and protections.	Engage and implement third-party AI services that have not undergone adequate security and risk assessments.
Familiarise yourself with terms and conditions specific to the AI application.	

Use Case Examples

Nick needs to develop a project plan and is wondering if he can use ChatGPT to create a baseline project plan that he can then improve upon.

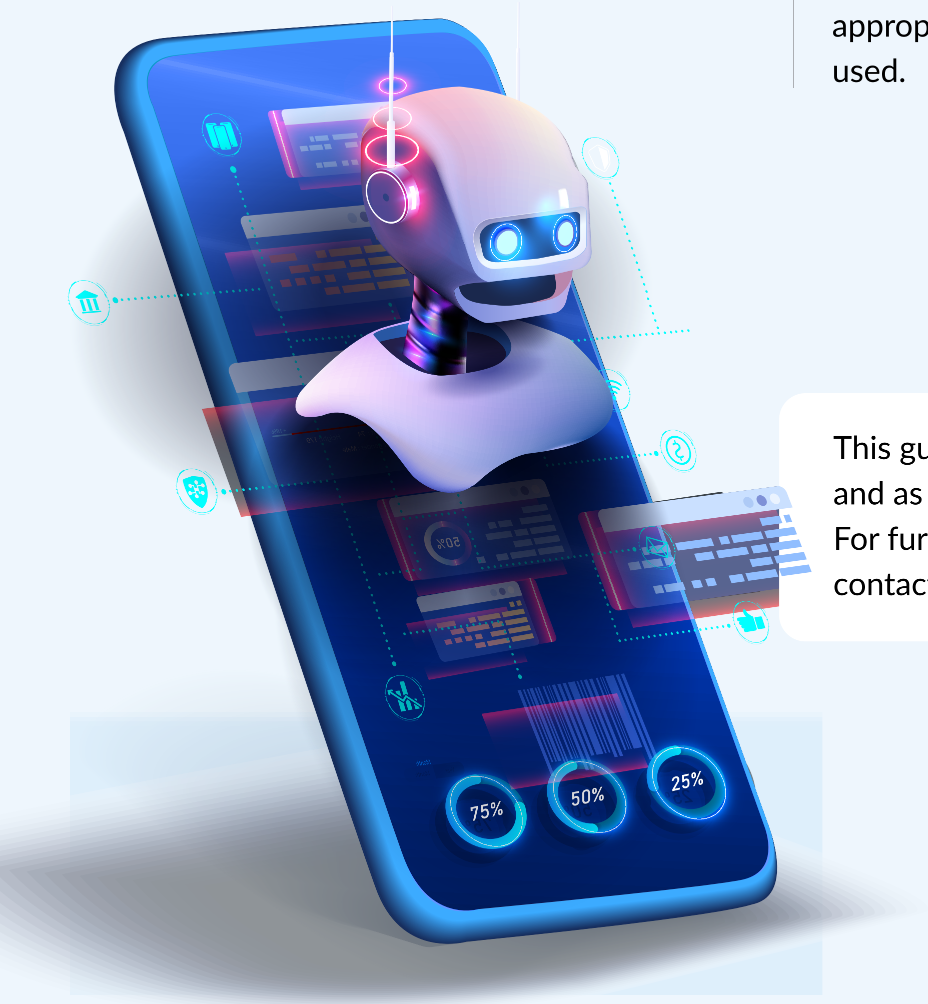
What should Nick do?

Nick can get the template of a project plan from ChatGPT. Nick must refrain from entering any details of the project, such as the project name, agency, names of systems/ software, high level requirements or staff members involved. These details could provide sensitive information about the project to ChatGPT.

Roque is writing technical requirements for a tender that needs to go out urgently. Roque wants to type his requirements into Bard AI to confirm some of the technical specifications around monitor resolutions.

What should Roque do?

Roque can use Bard AI to find information about technical aspects, however Roque should take care not to input any details of his agency's specific requirements or any organisational information. Roque should also take care not to mention the word tender or any market sensitive information that could, combined with his official email address, indicate a tender is being prepared by his agency. The information received should also be thoroughly validated by a human for appropriateness and accuracy before being used.



This guidance will be reviewed as AI evolves, and as risks are better understood.
For further inquiries or questions, please contact ai-dgov@dpc.wa.gov.au.