SIRO NOTICE
Senior Information Risk Owner

2024/002
ChatGPT and Generative Artificial Intelligence Tools

SIRO NOTICES

SIRO Notices are Welsh Government notices of organisation wide changes to security related procedures that must be adopted by all Divisions. Deputy Directors are required to confirm, through their Directors, that the changes have been implemented in their areas via the annual Internal Control Questionnaire.

Use of Microsoft 365 Copilot, ChatGPT, Google Bard and other Generative Artificial Intelligence (AI) Tools

This is an updated version of the SIRO notice (2023-001) that was issued in April 2023.

Microsoft 365 Copilot and Copilot for Bing
Microsoft 365 Copilot and Copilot for Bing operate within our network boundary. This means that they are suitable for use with OFFICIAL-SENSITIVE data as the queries do not leave our network and they are not used to train any of the large language models.

All staff have access to Copilot for Bing, but only staff with licenses have access to Microsoft 365 Copilot.

Other AI Tools e.g. ChatGPT
We continue with the formal Welsh Government position on the use of other Generative AI tools such as ChatGPT, as detailed below. As new tools evolve, we will be developing a list of AI tools which have been approved for use across the organisation by the Solution Design Authority.

Remember:

- OFFICIAL-SENSITIVE or personal information should not be used in any queries or prompts.

- Queries or prompts that could lead to issues should they be made public must also be avoided when using these tools e.g. a query from a CEO asking “what’s the law on making people redundant” has obvious significance when the CEO is identifiable. Consequently, be very cautious when entering any query or prompt about Welsh Government, even if you are using the tool via a personal account.

- The output of any searches or prompts should only be used as a basis for further work and due diligence - it should never be fully relied upon without manual review.

- You can use your gov.wales email address to register for AI tools such as ChatGPT - ***but you must register with a unique password*** i.e. one that is different to that used to access any Welsh Government systems.

The scope of this notice also includes the paid-for ChatGPT+ and Open AI API. For clarity, you must not buy licenses for any specific generative AI tools – including, but not limited, to OpenAI API – nor implement any AI functionality within existing applications without approval from the Solution Design Authority. If you are in any doubt about the use of AI tools beyond ChatGPT, please get in touch with colleagues in the Departmental Security Unit.

If you have worked with the Data Science Unit to develop bespoke AI tools, you should work with the Unit on privacy impact and ethics assessments to understand what information can be used and the limitations of those models.

See Annex 1 for further information about ChatGPT and using the tools safely.

Remember to…	**Please do…**
Handle sensitive or personal information securely and not share publicly via ChatGPT or other AI tools. We must always comply with UK GDPR. Only enter information that is already publicly available into these Generative AI tools Microsoft 365 Copilot and Copilot for Bing are the only AI tools suitable for OFFICIAL-SENSITIVE information.	Explore and understand more about Generative AI tools. Learn about their shortcomings and think about how it might help in your area of work, and share your ideas with colleagues
Double-check the output – be it text, code or numbers – and be mindful of the currency (or sometimes lack of) of the information used by the tool. Generative AI has limits and faults, so we should always treat outputs with caution.	Discuss and enquire about the application of Generative AI in the organisation, for instance through reaching out to your Digital Champion to find out more.
	Experiment with Generative AI, testing out a variety of asks, but being mindful of its risks and limits as outlined in this document.

Tim Moss
Welsh Government Senior Information Risk Owner (SIRO)

Issue date: 10 October 2024

Annex 1 – ChatGPT and using Generative AI safely

What is ChatGPT and how does it work?

Chat Generative Pretrained Transformer (to use its full name) is one of a growing number of 'generative AI' tools to be developed. Generative AI has the ability to create human-like responses across an almost limitless range of subjects and almost any language. This makes tools such as Chat immensely powerful and attractive – it achieved more than 1 million plus subscribed users in just 5 days in late 2022. For comparison, it took Netflix almost 3.5 years to reach the same milestone.

Are there any drawbacks?

In short, yes. With any highly complex technology we can always expect some risks that need to be carefully considered. AI tools such as Chat have the potential to revolutionise the way we live and work, but we must also be mindful of some of the things we know, as well as those we don't.

So, what are the risks?

- The data that allows Chat to function is predominantly internet-based content. This means that its accuracy **cannot be guaranteed**.

- It stores and re-uses **every question** it's asked, to re-train and fine tune its data model. The global electronics giant Samsung suffered a high-profile data breach resulting from employees entering company-sensitive data in to ChatGPT.

- On occasions, it has been found to respond with offensive language.

- It also generates a what are known as 'AI hallucinations' – incorrect or misleading information but presented in a way that makes it highly believable.

Surely there are benefits we can gain from new A.I technologies?

We know there is interest in how we could use Generative AI and other new technology to help us with our work - from simplifying processes to analysing large scale consultations or writing early drafts of correspondence. In using any new technology, we need to balance the potential benefits with the risks outlined above.

As a national government, as well as managing the clear security risks we also need to be mindful of the ethical and reputational implications of automating our work, as well as ensuring Trade Union colleagues are fully engaged in line with our principles of social partnership and fair work.

What now?

The power and potential of AI technologies is growing at an exponential rate. We are already working as an organisation to better understand how AI can be used to support some of our key ambitions, such as improving public service delivery and growing our economy here in Wales. As part of WG2025 we will continue to actively consider how this could be part of our improvement programme. We are also continually monitoring guidance issued from other

parts of government, including the Central Digital and Data Office and the National Cyber Security Centre, to learn from the approaches they are taking. However, we are not yet in a position where DDaT and IT Services are able to offer any support around the use of ChatGPT or any other AI tools.

For now, the key message is one of caution – as an organisation, but also as individuals, we must ensure we take the utmost care of the data we are responsible for. We'll need to strike a careful balance of embracing these sorts of technologies, whilst not putting our data at risk.

As you'd expect, we'll be keeping a very close eye on these technologies as they rapidly develop.

Remember to…	**Please do…**
Handle sensitive or personal information securely and not share publicly via ChatGPT or other AI tools. We must always comply with UK GDPR. Only enter information that is already publicly available into these Generative AI tools Microsoft 365 Copilot and Copilot for Bing are the only AI tools suitable for OFFICIAL-SENSITIVE information.	Explore and understand more about Generative AI tools. Learn about their shortcomings and think about how it might help in your area of work**,** and share your ideas with colleagues
Double-check the output – be it text, code or numbers – and be mindful of the currency (or lack of) of the information used by the tool. Generative AI has limits and faults, so we should always treat outputs with caution.	Discuss and enquire about the application of Generative AI in the organisation, for instance through reaching out to your Digital Champion to find out more.
	Experiment with Generative AI, testing out a variety of asks, but being mindful of its risks and limits as outlined in this document.