Department of Finance
An Roinn Airgeadais
www.finance-ni.gov.uk

The use of Generative AI in the NICS

Guidance Paper

DOCUMENT CONTROL

The Department of Finance Departmental Chief Information Security Officer is the owner of this document and is responsible for ensuring it's distribution. The document will be reviewed annually.

DOCUMENT INFORMATION

Title	The use of Generative AI in the NICS – Guidance Paper	**Version**	2
Author	[redacted]	**Date**	30/05/2023
Reviewed By	[redacted]	**Date**	07/06/2023
Approved by	[redacted]	**Date**	07/06/2023
TRIM Reference	[redacted]	**Review Date**	07/06/2025

CHANGE HISTORY

Ver. No.	Ver. Date	Revised By	Description	Reviewer	Status
2	05/06/2024	[redacted]	Wording updated to provide consistency in terminology of reference to Generative AI.	[redacted]	Approved

DISTRIBUTION

Group	Date
[redacted]	11 June 2024
[redacted]	11 June 2024
[redacted]	11 June 2024
[redacted]	11 June 2024

Contents

Document Control 2
1 Introduction 5
1.1 Aim 5
2 Generative AI 5
3 How might the NICS use Generative AI Models 5
4 How can I safely provide Generative AI Models with information 6
5 When using Generative AI, consider the three How's 7
6 Guidenlines for responsible use of AI systems and generated content 7
6.1 Data privacy and security 7
6.2 Bias and discrimination 7
6.3 Ethical considerations 7
6.4 Transparency and Explainability 8
6.5 Accountability 8
6.6 Training 8
6.7 Continuous monitoring and improvement 8
7 Requirements for AI-generated content 8
7.1 Proofing 8
7.2 Editing 8
7.3 Fact-checking 9
7.4 Starting point, not the finished product 9

8 References 9
9 ICO Guidance 9
10 NICS Goverance and Approvals 9
11 Appendix 1 – Questions when processing Personal Data 11
12 Appendix 2 – Glossary 12
13 Appendix 3 – Useful Links 13

1 INTRODUCTION

1.1 AIM

The aim of this guidance document is to ensure a consistent approach on the use of Generative AI across the NICS.

2 GENERATIVE AI

Generative AI is a broad label used to describe any type of artificial intelligence (AI) that can be used to create new text, images, video, audio, or code. Large Language Models (LLMs) are part of this category of AI and produce text outputs.

ChatGPT is an example of a generative AI language model developed by OpenAI and is based on the GPT (Generative Pre-trained Transformer) architecture. ChatGPT is trained on a massive dataset of text from the internet, including books, articles, and websites, and it uses deep learning techniques to generate responses that are contextually relevant and grammatically correct. It can be used for a variety of natural language processing tasks, such as language translation, summarisation, and conversation generation.

With all AI technologies, there are concerns about the ethical implications of use, particularly in areas such as bias, privacy, and the potential for misuse. As such, responsible use of Generative AI, along with appropriate regulation and oversight is essential.

NICS recognises the potential benefits of using artificial intelligence (AI) to improve efficiency and productivity in the workplace. However, we also acknowledge the importance of using AI responsibly and ethically, particularly when it comes to generating content.

3 HOW MIGHT THE NICS USE GENERATIVE AI

- Generative AI: There are many ways NICS Departments could use Generative AI to improve operations and engagement with customers. Here are a few examples:

- CHATBOT: A business area could use Generative AI to develop a chatbot that can answer common customer inquiries and provide basic support. Generative AI can help the chatbot generate more natural-sounding responses and improve the overall customer experience.

- CONTENT CREATION: Generative AI can help a business area generate content for their website, social media channels, and marketing materials. This can save time and resources while still producing high-quality content that is engaging and informative.

- CUSTOMER ENGAGEMENT: A business area could use Generative AI to engage with customers on social media platforms, such as Twitter or Facebook, by generating responses to customer inquiries or comments. This can help improve customer satisfaction and loyalty.

- LANGUAGE TRANSLATION: If a business area serves customers who speak different languages, Generative AI can be used to translate content quickly and accurately.

- RESEARCH AND ANALYSIS: Generative AI can help a business area analyse customer feedback or reviews to identify common themes or issues. This can help the business make data-driven decisions and improve its products or services.

It's important to note that while Generative AI can be a valuable tool it is essential to use it responsibly and ethically. This includes ensuring that any AI-generated content is fact-checked and edited by humans, and that data privacy and security are prioritised.

4 HOW CAN I SAFELY PROVIDE GENERATIVE AI MODELS WITH INFORMATION

When evaluating the use of Generative AI Models to create content or automate certain business tasks which may involve providing information either through fine-tuning or prompt augmentation an appropriate solution must be selected.

When considering a Public Generative AI Models, users should be aware of the risk that data stored online may be hacked, leaked, or more likely accidentally made publicly accessible. This could potentially include user-identifiable information. A further risk is that the operator of the Generative AI Models is later acquired by an organisation with a different approach to privacy than was true when data was entered by users.
For **cloud provided** Generative AI Models, the terms of use and privacy policy become key (as they are for public Generative AI Models s). Users must understand:

- How the data they use for fine-tuning or prompt augmentation is managed.

- Is it available to the vendor's researchers or partners? If so, in what form?

- Is data shared in isolation or in aggregation with other organisations?

- Under what conditions can an employee at the provider view queries?

Self-hosted Generative AI Models are likely to be highly expensive. However, following a security assessment (which should include referring to the NCSC's principles for the security of machine learning), they may be appropriate for handling organisational data. In particular, organisations should refer to NCSC guidance on securing your infrastructure and data supply chains.

Sensitive information or personal data should never be put into these tools. Beyond existing data protection laws NICS has no oversight over how data, which is entered into web-based Generative AI model is then used. Therefore, you should not put information into Generative AI models that, if compromised or lost, could have damaging consequences for individuals, groups of individuals, a department or for NICS more generally.

5 WHEN USING GENERATIVE AI, CONSIDER THE THREE HOW'S

1. How your question will be used by the system. These systems learn based on the information you enter. Just as you wouldn't share work documents on social media sites, do not input such material into Generative AI tools.

2. How answers from Generative AI can mislead. These tools can produce credible looking output. They can also offer different responses to the same question if it is posed more than once, and they may derive their answers from sources you would not trust in other contexts. Therefore, be aware of the potential for misinformation from these systems. Always apply the high standards of rigour you would to anything your produce, and reference where you have sourced output from one of these tools.

3. How Generative AI operates. A Generative AI tool, such as a Large Language Model, will answer your question by probabilistically choosing words from a series of options it classes as plausible. These tools cannot understand context or bias. Always treat with caution the outputs these tools produce and challenge the outputs using your own judgement and knowledge.

6 GUIDENLINES FOR RESPONSIBLE USE OFGENERATIVE AI MODELS AND GENERATED CONTENT.

6.1 DATA PROTECTION AND SECURITY

Ensure that any data used to train Generative AI models is handled with appropriate care and that any AI-generated content is handled in compliance with data protection laws.

6.2 BIAS AND DISCRIMINATION

AI systems can reflect and amplify human biases and prejudices. Ensure that Generative AI models / AI generated content do not discriminate against individuals or groups based on protected characteristics such as race, gender, or age.

6.3 ETHICAL CONSIDERATIONS

Ethical concerns surrounding the use of Generative AI models in the workplace, such as the potential impact on employment and the ethical use of AI in decision-making must be considered.

6.4 TRANSPARENCY AND EXPLAINABILITY

Ensure the use of Generative AI models is transparent and explainable and that the user understand how AI content / decisions are made and why specific outcomes are generated. All AI-generated content must be clearly labelled as such, and the use of AI in generating content should be transparent to staff and customers.

6.5 ACCOUNTABILITY

Understand that Departments are accountable for their outcomes generated by Generative AI models and for any harm caused by their AI-generated content.

6.6 TRAINING

Staff who work with Generative AI models should be trained on how to use them effectively, ethically and that they understand the limitations of the technology.

6.7 CONTINUOUS MONITORING AND IMPROVEMENT

Continued use of the technology requires ongoing monitoring and modification of Generative AI models to ensure that they are functioning as intended and are not causing unintended consequences.

7 REQUIREMENTS FOR AI-GENERATED CONTENT

7.1 PROOFING

All AI-generated content must be proofread and checked for accuracy by a human before being published or shared. This includes checking for spelling errors, grammar mistakes, and factual inaccuracies.

7.2 EDITING

AI-generated content must be edited to ensure that it is well-written, coherent, and engaging. This includes ensuring that the content is structured in a logical manner and that it is appropriate for the intended audience.

7.3 FACT-CHECKING

AI-generated content must be fact-checked to ensure that all information is accurate and up-to-date. This includes verifying sources, checking statistics, and ensuring that any claims made in the content are supported by evidence.

7.4 STARTING POINT, NOT THE FINISHED PRODUCT

AI-generated content should be viewed as a starting point, not the finished product.

8 REFERENCES

Whether using the outputs from Generative AI either verbatim or with minor alterations, it is important to make clear to those reading that a Generative AI tool has been used. To do this the tool(s) should be cited in a footnote, with its URL and any sources used as inputs. If

9 DATA PROTECTION GUIDANCE

As with any other processing of personal data, Generative AI developments must take the approach of data protection by design and by default . This isn't optional – if you're processing personal data, it's the law.

Data protection Legislation still applies when the personal information that you're processing comes from publicly accessible sources. If you're developing or using generative AI that processes personal data you need to ask yourself the questions at Appendix 1.

To assist organisations considering this new technology, the ICO has updated its Guidance on AI and Data Protection; which should be used in conjunction with their risk toolkit, which helps organisations to identify and mitigate data protection risks. Advice is also available from the ICO's Regulatory Sandbox and new Innovation Advice service.

10 NICS GOVERANCE AND APPROVALS

As per the 'NICS Use of Generative AI –Policy'.

NICS staff wishing to utilise the technology must complete the embedded Proforma.

Aligning with the wider UK approach, but acknowledging the current absence of a fully developed Regulatory Framework, NICS staff wishing to utilise the technology must;

- Receive approval from Grade 6 (minimum) for the specific use case(s) in which they intend to use AI;
- Engage with the Departmental ITSO and provide the SIRO with written assurance that a security assessment has been carried out and the Generative AI Model aligns with the Principals detailed within the AI Cyber Security Code of Practice, NCSC guidance 'Principles for the security of machine learning', 'securing your infrastructure' and 'data supply chains' that there is no risk to NICS;
- Engage with the DPO to ensure a full understanding of the terms of use and privacy notice of the Generative AI Model they intend to use. Consider a Data Protection Impact Assessment to ensure compliance with UK Data Protection legislation and that there is low/no risk to personal data.
- Receive approval from the Departmental Senior Information Risk Owner for the specific use case(s) in which they intend to use AI;
- Adhere to the guidance on 'The use of Generative AI in the NICS – Guidance Paper'
- Provide assurance to the G6, DPO and ITSO and SIRO during and on completion of the Generative AI usage that all data, privacy, information, security and ethical regulations / standards have been adhered to.

11 APPENDIX 1 – QUESTIONS WHEN PROCESSING PERSONAL DATA

Questions when processing Personal Data

1. Have you considered all policies, procedures and guidance provided by your department to assist you in meeting your data protection obligations and engaged with your DPO if necessary?

2. What is your lawful basis for processing personal data? If you are processing personal data you must identify an appropriate lawful basis, such as consent or legitimate interests.

3. Are you a controller, joint controller or a processor? If you are developing generative AI using personal data, you have obligations as the data controller. If you are using or adapting models developed by others, you may be a controller, joint controller or a processor.

4. Have you prepared a Data Protection Impact Assessment (DPIA)? You must assess and mitigate any data protection risks via the DPIA process before you start processing personal data. Your DPIA should be kept up to date as the processing and its impacts evolve.

5. How will you ensure transparency? You must make information about the processing publicly accessible unless an exemption applies. If it does not take disproportionate effort, you must communicate this information directly to the individuals the data relates to.

6. How will you mitigate security risks? In addition to personal data leakage risks, you should consider and mitigate risks of model inversion and membership inference, data poisoning and other forms of adversarial attacks.

7. How will you limit unnecessary processing? You must collect only the data that is adequate to fulfil your stated purpose. The data should be relevant and limited to what is necessary.

8. How will you comply with individual rights requests? You must be able to respond to people's requests for access, rectification, erasure or other information rights.

9. Will you use generative AI to make solely automated decisions? If so – and these have legal or similarly significant effects (e.g. major healthcare diagnoses) – individuals have further rights under Article 22 of UK GDPR.

12 APPENDIX 2 – GLOSSARY

AI	Artificial Intelligence
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
GPT	Generative Pre-trained Transformer
ICO	Information Commissioners Office
ITSO	IT Security Officer
LLM	Large Language Models
NCSC	National Cyber Security Centre
NICS	Northern Ireland Civil Service
SIRO	Senior Information Risk Owner

13 APPENDIX 3 – USEFUL LINKS

AI and data protection risk toolkit | ICO

AI regulation: a pro-innovation approach - GOV.UK (www.gov.uk)

Artificial intelligence | ICO

National AI Strategy - GOV.UK (www.gov.uk)

Principles for the security of machine learning - NCSC.GOV.UK

Regulatory Sandbox | ICO

Secure your infrastructure - NCSC.GOV.UK

Supply chain security guidance - NCSC.GOV.UK