GEN AI Guidance

What can GEN AI tools and apps do?

Generative artificial intelligence (GenAI) can create realistic, human-like text, images, code and art based on huge amounts of (usually public) data it has been trained on. It

- Can produce a range of useful outputs, like text, audio, images, and code
- Responds to natural language questions, so any employee can use it
- Is very good at understanding different types of data - useful given councils have large amounts of unstructured data in a large variety of formats.

What are examples of Gen AI?

Examples of GEN AI are ChatGPT, Bard, DALL-E, Otter.AI, GPT-4, Alpha Code, Cohere Generate and Google DeepMind but there are many more..

What must we do?

1. Do **not** use any GEN AI tools without first undertaking a Data Protection Impact Assessment (DPIA). Form available here.
2. Consider yourself accountable for everything the AI creates. Officers and councils can't blame the AI, so only use it when you can easily check and correct the AI's creations.
3. **Never** upload personal data about people. This data may be shared externally and be a data breach. Some Apps will have ways to prevent this e.g. If you are authorised after completing a DPIA to use ChatGPT, you can turn off 'Chat history and training' in settings.

What are the risks?

We need to ensure that any GEN AI tools we use are

- **Safe & Fair** - Evaluate models for their success, as well as potential biases, both during development and in ongoing ways post-deployment.
- **Transparent** - Share details of what data and AI models are used.
- **Accountable** - Humans are 'in-the-loop' where needed, and we have straightforward appeal processes for any automated or AI-informed decisions.
- **Private & Secure** - Ensure you are following data protection law, plus equalities legislation.

Example of a risk: A Teams meeting with an external partner attending who had the Otter.AI transcription app installed on their laptop. It then captured personal data discussed at the meeting. All attendees should be asked to disable AI apps in future

Dos and Don'ts Checklist

Do…

- **Undertake a DPIA before using GEN AI tools**
- **Do** maintain human oversight and responsibility for making final decisions on output produced
- **Do** use responsibly and ethically
- **Do** use in accordance with relevant organisation policy
- **Do** comply with relevant laws and regulations
- **Do** specify the definitions and scope of your prompts with care
- **Do** use to create draft briefings, reports, presentations, customer responses, etc.
- **Do** use to improve and refine existing content
- **Do** use to analyse publicly-available data
- **Do** fact check material generated by Generative AI LLMs
- **Do** be aware of the potential for disinformation and scams being generated
- **Do** take care to avoid use of output that may breach copyright or intellectual property rights
- **Do** be aware of risks including accuracy, bias, discrimination, confidentiality and security

Don't…

- **Don't** use to record and process confidential data and information
- **Don't** use to store or release non-public records
- **Don't** use for private individual records
- **Don't** let go of moral and ethical responsibility for output
- **Don't** use if you are in doubt about the security of data or information being input
- **Don't** assume that all of the output generated is factually correct

For more advice and guidance on GEN AI contact Information Governance & Risk on info.requests@leicester.gov.uk