



Generative AI guidance for taught students

Generative AI guidance for taught students	1
Can I use Generative Artificial Intelligence at Leeds?	1
Introduction to Generative AI for Taught Students	1
How does Generative AI work?	4
Key Strengths and Weaknesses of Generative AI	5
Key Strengths	5
Key Weaknesses	6
Use of Generative AI in assessments	8
Acknowledging the use of and referencing Generative AI	12
Generative AI Quick Checklist	13

Generative AI guidance for taught students

This guidance is for all taught students at the University of Leeds. Guidance for Postgraduate Research students is being developed separately and will be published soon.

Can I use Generative Artificial Intelligence at Leeds?

Welcome to the University of Leeds' latest information and guidance on using Artificial Intelligence (AI) in your studies. The guidance is intended to ensure that our use of Generative AI across Leeds is effective, ethical and transparent.

This guidance has been co-produced by University of Leeds academic staff, professional services staff and students from different disciplines. It is in line with the advice given at comparable universities, particularly UCL (from whom we have adopted this guidance with thanks).

This guidance will be reviewed every six months while AI discourse and capabilities develop. We will notify you of any major changes to the guidance.

Please check the [Generative AI website](#) regularly to see the latest advice and get in touch at AI@leeds.ac.uk if you have feedback or ideas for future advice we should include.

Introduction to Generative AI for Taught Students

This section introduces Generative AI (Gen AI) and highlights the basics of how it works, along with its strengths and weaknesses.

Generative AI has the potential to transform education and learning and to develop new forms of creativity. At the same time, it is full of biases, inaccuracies, and other issues. We need to hedge our moments to explore its full potential while recognising its flaws and issues.

Professor Jeff Grabill, Deputy Vice-Chancellor: Student Education

What is Generative AI (Gen AI)?

Gen AI is an Artificial Intelligence (AI) technology that automatically generates content in response to written prompts. The generated content includes texts, software code, images, videos and music.

Gen AI is usually trained using data from webpages, social media conversations, survey data, private datasets and other online content. It generates its outputs by identifying and repeating common patterns (for example, which words typically follow which words, or which pixels of images should come after which pixels). This is usually done by ingesting large datasets and statistically analysing the distribution of words or pixels or other elements in the data.

There are many other types of AI applications, that do not involve Gen AI, that are having an impact on teaching and learning. These will be explored in more detail as the guidance further develops.

Remember:

- While Gen AI is often accurate, even in complex conversations, it is not reliable and can make simple mistakes.
- Gen AI can simulate intelligent conversation, but it doesn't possess true consciousness or human-like thinking.
- Though Gen AI responses might seem insightful, it doesn't genuinely understand or comprehend topics as humans do.

This means that while many of us will be using Gen AI in creative ways in our work, we always need to verify that its outputs are accurate and appropriate for our needs.

Text Gen AI

In response to a human-written prompt, text Gen AI generates text that usually appears as if a human has written it.

Yet, just like human-written texts, text Gen AI outputs can be superficial, inaccurate, untrustworthy and full of errors.

Despite appearances, text Gen AI does not understand either the prompt written by the human or the text that it generates. You may have encountered this if you have ever tried to have a conversation with a text Gen AI tool and seen it contradicting itself. You may also have seen a text Gen AI tool choosing words that are not quite right for the context or intended meaning of the output. It is just providing the words that would statistically be most likely to appear in a response to the prompt you provided.

Every time that we use a text Gen AI tool, we need to consider its output from a critical perspective.

Examples of text Gen AI tools

There are many tools that provide text Gen AI functionality. The following list provide examples of some of the most popular ones. **Please note that the University of Leeds is not endorsing any of the tools in this list.**

- [Bard](#) (Google)
- [ChatGPT](#) (OpenAI)
- [Claude](#) (Anthropic)
- [Hugging Chat](#) (HuggingFace)
- [Llama](#) (Meta)

Examples of other Gen AI tools built on top of Gen AI tools

There are many tools that provide additional Gen AI functionality. The following list provides examples of some of the most popular ones. **Please note that the University of Leeds is not endorsing any of the tools in this list.**

- [ChatPDF](#) (summarises and answers questions about submitted PDF documents)
- [Elicit](#) (aims to automate parts of researchers' workflows, identifying relevant papers and summarising key information)
- [WebChatGPT](#) (Google Chrome extension that gives ChatGPT Internet access, to enable more accurate and up-to-date conversations)
- Microsoft has incorporated AI Chat into its Bing search engine and is implementing AI Chat across its Office portfolio with Copilot. You may have already seen this if you have used the [Edge browser](#) on a University PC.

Image, video, music Gen AI

Like Text Gen AI, image, video or music Gen AI can generate outputs based on human-written prompts. Some can also respond to visual or audio prompts.

Again, the appearance of image, video and music Gen AI outputs might appear novel. However, usually, they are only complex combinations of the millions of images, videos or music that they have ingested during their training.

Gen AI uses earlier works in its outputs and this creates substantial challenges in copyright, intellectual property and the ownership of content. You may have heard of the 2023 Writers Guild of America (WGA) strike in Hollywood on this topic, or the issues with a song created by AI mimicking the work of Drake and The Weeknd, which highlight deep issues with ethics and ownership when it comes to AI generated works.

Examples of image, video and music Gen AI tools

There are many tools that provide image, video or music Gen AI functionality. The following list provide examples of some of the most popular ones. **Please note that the University of Leeds is not endorsing any of the tools in this list.**

- [DALL•E](#) (OpenAI's image Gen AI tool)
- [DreamStudio](#) (Stable Diffusion's image Gen AI tool)
- [Midjourney](#) (Image Gen AI tool)
- [Runway](#) (Video Gen AI tool)
- [Boomy](#) (Music Gen AI tool)
- [Voicemod](#) (Music Gen AI tool).

How does Generative AI work?

Both text and image Gen AI are based on a set of AI techniques that have been available to researchers for several years and have been built one on top of another.

Text Gen AI

While it isn't necessary to understand exactly what each of the terms below means and exactly how the technologies work, it is helpful to have a basic understanding of their definitions and how they form a hierarchy of specificity. The following flow chart should provide a quick understanding of ChatGPT (and other text Gen AI) and that they are types of the following:

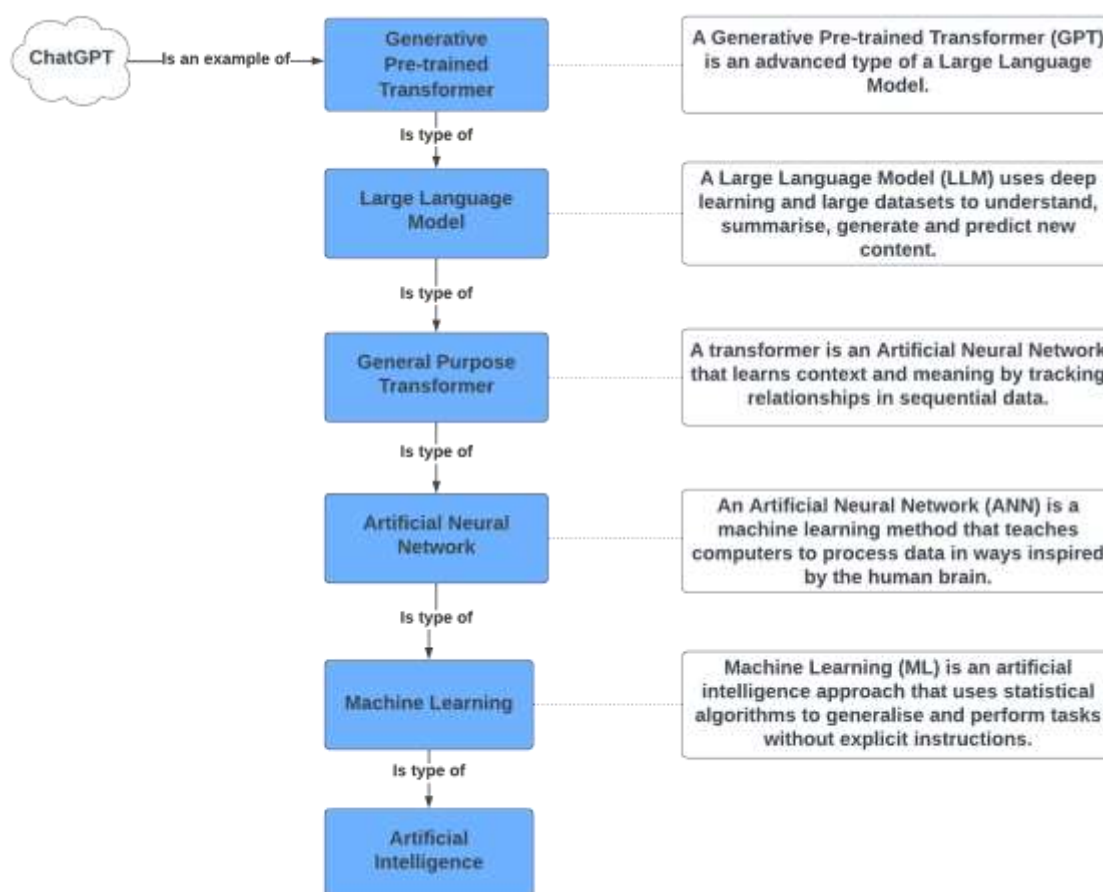


Image 1 – a flow chart highlighting the hierarchy of technologies used for a tool like ChatGPT.

Text description of image: ChatGPT is an example of a Generative Pre-trained Transformer (GPT), which is an advanced type of a large language model. A large language model uses deep learning and large datasets to understand, summarise, generate and predict new content. Large language models are a type of general purpose transformer. A transformer is an artificial neural network that learns context and meaning by tracking relationships in sequential data. An artificial neural network is a machine learning method that teaches computers to process data in ways inspired by the human brain. Machine learning is a type of artificial intelligence that uses statistical algorithms to generalise and perform tasks without explicit instructions

Key Strengths and Weaknesses of Generative AI

To understand the strengths and weaknesses of Gen AI, let's have a look at what's involved in training it. The training process involves the tool being provided with and processing huge amounts of data scraped from the internet and elsewhere.

Processing this data involves identifying patterns, such as which words typically go together (eg "Happy" is often followed by "Birthday"). This leads to several strengths and weaknesses, with the key ones described below.

Key Strengths

Diverse outputs

Gen AI can produce diverse and *seemingly* original outputs. It can create content and outputs through capturing nuances in language, based on the patterns which may not have otherwise been seen before in the data they were trained on. This helps open different perspectives and can give ideas on how to explore topics from a varied set of viewpoints.

Levelling the playing field

Gen AI can process and interpret human language in a conversational style, allowing it to generate contextually relevant responses to user prompts. It can also reformulate text to simplify or summarise it, which can help people start to understand more complex ideas. In addition, Gen AI can process and generate text in multiple languages. If used appropriately, Gen AI can be a great leveller for those who do not speak English as their first language or may not have the same literacy or language skills as others.

Organisational productivity

Gen AI can be fine-tuned for various tasks and domains, making them widely applicable (eg, chatbots, content generation and language translation). This can help boost organisational productivity, eg through answering questions in a human-like style, in reducing effort on tedious and monotonous tasks, in providing accessible summaries of complex topics, in doing automatic translations and transcriptions, etc.

Personalisation

Gen AI models can remember previous interactions, which results in more coherent and relevant conversation experiences for users. You can also ask some models to remember the style in which you write, in which you want your data to be presented and even ask it to test your knowledge against any piece of content. Gen AI can generate responses quickly, allowing for rapid interactions and real-time applications.

Industry applications

It is anticipated that most industry and workplaces will be using a form of Generative AI in the future to enhance and optimise their work. As Gen AI integrates into our

daily learning and work tools (such as Copilot and Microsoft Office, or the AI content generator in Grammarly, etc), it is important to develop your skills in using Gen AI effectively and ethically.

Key Weaknesses

Lack of trust and authenticity

Gen AI can generate information that appears factual but is often inaccurate. This is often called AI hallucinations. It is important to note that while Gen AI models appear to understand the content that they use and generate, they do not understand it. The data that they use for training have lots of inaccuracies and biases in them already. In addition, Gen AI can be used to create fake news, misinformation and 'deep fakes' easily. This could lead users to have misplaced trust in the Gen AI output.

Copyright and ownership

Gen AI output imitates or summarises existing content – mostly without the permission of the original content owners – but can give the appearance of creativity and originality. This generates challenges and issues of copyright, ownership, intellectual property and lack of authoritative legislation in this rapidly evolving area. It is important to keep this in mind when using Gen AI tools. You should not copy and paste any copyrighted text, or other sensitive or personal data, into an AI tool for it to use, as this data could be added into the AI tool's training dataset and could then be used illegally or unethically.

Carbon footprint

Training Gen AI requires huge amounts of power and indirectly generates huge amounts of carbon, with important consequences for climate change. For example, it is estimated that the training of GPT4 consumed between 51,772 and 62,318 megawatt hours of electricity and generated between 1,035 and 14,994 metric tons of carbon dioxide emissions, depending on the global location of the training. As an example, a 3000-mile roundtrip via air (such as London to Boston) emits 1 metric ton of carbon dioxide.

Feedback loop

The output of Gen AI is flooding the internet. This poses an interesting risk for future GPT models. These models themselves will be trained on online content that earlier GPT models have created (including all its biases and errors). This self-referential loop compounds the mistakes in data, might contaminate the training data and lead to the concept of model collapse. A model collapse occurs where the models quickly forget most of the original data that they learnt from.

Ethical, Social and Human costs

After the Gen AI model is trained, it is often checked and refined in a process known as Reinforcement Learning from Human Feedback (RLHF). In RLHF, Gen AI

responses are reviewed and validated by human reviewers. These human reviewers ensure that the Gen AI responses are appropriate, accurate and align with the intended purpose.

In the development of ChatGPT, the RLHF reviewers mostly were workers in global south countries such as Kenya. Workers were paid less than \$3 per hour to review the outputs of ChatGPT and identify any objectionable or toxic materials. This work has had a massive negative impact on many of those who were involved, including issues of exploitation and trauma.

Gen AI also tends to output standard answers that replicate the values and biases of the creators of the data used to train the models. This may constrain the development of plural opinions and further marginalise already marginalised voices.

It is important to reflect that Generative AI can help us achieve great things, but that it also comes at an ethical, social, environmental and human cost. Thus, it is important to consider these factors in our approach and support ethical frameworks around use of generative and other forms of AI. It is also very important to always check if the content produced is accurate and not an AI hallucination.

Use of Generative AI in assessments

The University of Leeds use a three-tiered traffic light system of Gen AI use in assessments. Staff and students should be aware of the categories and discuss in class whether an assignment falls wholly or partially within each.

As a student, you should always make sure that the [Academic Integrity principles](#) are followed in any work that you do, and when in doubt, always ask your tutor about appropriate use of Gen AI.

Introduction

As a general principle, you can use Gen AI to help you learn but cannot use AI to generate or falsify work. You may use Gen AI in ways that support your learning, enhance your ability to achieve your Programme Learning Outcomes and prepare you to succeed in your future careers. Using Gen AI to falsify work or breach guidelines for an assessment will undermine all these benefits and damage your learning. This is explained further in the [Academic Misconduct](#) procedure.

Categories of use of Gen AI in Assessment

There is a three-tier traffic light categorisation for using Gen AI in assessments: red, amber and green. This categorisation may be used by teaching staff to design and set assessments and for you to complete assessments in ways that will optimise – rather than damage – your learning. The three categories are not defined rigidly; rather they are a tool for ensuring that for each piece of assessment, staff and students have a shared understanding of whether generative AI tools can be used and, if they can, how, how much and where in the assessment process.

Regardless of the category, you should always take a critical approach to the use of any output from a generative AI tool, as these tools can generate superficial, inaccurate and unhelpful outputs. Gen AI tools can create content that is not peer reviewed or academically rigorous, often without any academic sources underpinning that content. Therefore, you should always analyse and verify the information Gen AI tools provide, rather than accepting it at face value. As a reminder, please review the Strengths and Weaknesses section above to understand more on why this is vital for your success.

RED category: AI tools cannot be used

RED CATEGORY

Under this category, you must not use Gen AI tools. The purpose and format of the assessments makes it inappropriate or impractical for AI tools to be used.

As an example, assessments under these categories may include demonstrating foundation level skills such as remembering, understanding, independently developing critical thinking skills and applying knowledge or demonstrating fundamental skills that will be required throughout the programme.

Such assessments are likely to be designed to support the development of knowledge and skills that you will require to be able to study successfully and effectively, including with the use of Gen AI tools in other contexts and in future assessments. There will be strong rationale for this category (for example, pedagogy, employability, etc). Your lecturer or module materials should explain this, but you can also ask for clarification on it.

Students believed to have ignored the categorisation will undergo the standard [Academic Misconduct procedure](#).

In an assessment in this category, you would still be permitted to use dictionaries, thesauri and spelling – and grammar-checking software, some of which are powered by AI, to help identify and correct spelling mistakes and grammatical errors. However, you should not use any software to rewrite sentences or make substantive changes to your original text, as this would be against the rules of this category.

We are currently working on what it means to provide reasonable adjustments while maintaining competence standards. If the meantime, if you have any questions or comments, please get in touch with us using AI@leeds.ac.uk

Red category examples

Examples of assessments where Gen AI is not allowed could include:

- In-person unseen examinations
- Class tests
- Some online tests
- Vivas
- Some laboratories and practicals
- Discussion-based assessments
- Where spoken and written language skills need to be assessed.

AMBER category: AI tools can be used in an assistive role

AMBER CATEGORY

Under this category, you are permitted to use AI tools in an assistive role as specified by the module tutor and required by the assessment.

Gen AI tools can be utilised to enhance and support the development of specific skills in specific ways, **as specified by the tutor and required by the assessment**. For instance, you might use Gen AI for tasks such as data analysis, pattern recognition, or generating insights.

It is anticipated that your tutor will be able to provide support and guidance to you in the use of Gen AI to ensure equity of experience. In these situations, the use of Gen AI is not in itself a learning outcome. There may still be parts of your assessment where use of Gen AI is inappropriate and this will be made clear to you in discussion with your tutor.

You should never provide any Gen AI tool sensitive or personal data in any prompt or content. For more information, please read the key information on [IT Security Considerations on the use of ChatGPT and AI LLM Engines](#).

Amber category examples

Examples of where Gen AI might be used in an assistive category include:

- drafting and structuring content
- as a support tutor
- supporting a particular process such as testing and debugging code or translating content
- providing ideas or inspiration to help you overcome a creative block.

It is also important to note that your tutors are also learning how to best design assessments that allow the use Gen AI to enhance your learning and development. Thus, an open discussion on this topic is encouraged at an early stage. Support guidance and case studies are also being created to support your tutors.

GREEN category: AI has an integral role and should be used as part of the assessment

GREEN CATEGORY

Under this category, you can use Gen AI as a primary tool and it should be used as part of the assessment.

In these assessments, you will demonstrate your ability to use Gen AI tools effectively and critically to tackle complex problems, make informed judgments and generate creative solutions. Here, the use of Gen AI and critical appraisal of its output will explicitly form part of the task and learning outcomes. The assessment will provide an opportunity to demonstrate effective and responsible use of Gen AI. The tutor should support and guide you in the use of Gen AI to ensure equity of experience.

Green category examples

Examples of where Gen AI tools could be used as an integral part of the assessment include:

- drafting and structuring content
- generating ideas
- comparing content (AI generated and human generated)
- creating content in particular styles
- producing summaries
- analysing content
- reframing content
- researching and seeking answers
- creating artwork (images, audio and videos)
- playing a supportive role and engaging in a conversational discussion
- developing code
- translating content
- generating initial content to be critiqued by students.

Note that the green category only applies to assessments where using Gen AI is explicitly part of the task. Assessments that do not specify this are not in the green category. Please check with your tutor if you are at all unsure. You can assume that unless the assessment is specified as one where Gen AI can explicitly be used, it will be in the red or amber category.

When using Gen AI tools to support your work, you should always save copies of the outputs that the Gen AI tool has produced for you. You may be asked to provide this information as part of your assessment or in any academic misconduct process.

Acknowledging the use of and referencing Generative AI

Gen AI is evolving rapidly and there is not yet consensus on how to acknowledge and reference it. In all cases, you are required to acknowledge truthfully what elements of any assignment are your own work or ideas and what has come from other sources – including Gen AI. You should make clear in your assignment where you have used Gen AI (eg for support with writing style but not for content).

The use of Gen AI **must** be acknowledged in an ‘Acknowledgements’ section of any piece of academic work where it has been used as a functional tool to assist in the process of creating academic work. If it is suspected that you have used a Gen AI tool to produce part of your work, but you have not acknowledged this use, this could be investigated under the Academic Misconduct procedure.

The minimum requirement to include in acknowledgement:

- Name and version of the generative AI system used eg ChatGPT-4.0
- Publisher (company that made the AI system) eg OpenAI
- URL of the AI system
- Brief description (single sentence) of context in which the tool was used.

For example:

“I acknowledge the use of ChatGPT-3.5 (Open AI, <https://chat.openai.com/>) to summarise my initial notes and to proofread my final draft.”

Further requirements may be stipulated by a department, academic programme, or individual teaching staff, or for a particular assignment, and will be made clear when an assignment is set. Details of such requirements will usually be provided in the assessment brief.

We are currently working on developing the referencing guidance further in collaboration with the sector and encourage you to keep checking the [Library's Referencing Pages](#) and the official guide for your School's referencing style for more information.

Generative AI Quick Checklist

If you are deciding on whether and how to use Gen AI tools, you may find the quick checklist below helpful.

- I have checked the assessment brief or have spoken to my tutor and can confirm that use of Gen AI is not prohibited for my assessment type.
- I have reviewed the Academic Integrity principles to help me avoid unintentional plagiarism.
- I have understood the risks and limitations of using Gen AI, including a recognition of issues of bias, sensitivity, accuracy, appropriate content and ethical issues.
- I have checked and critically reviewed any quotations, citations, or outputs that the Gen AI tool has generated.
- I have not submitted any Personally Identifiable Information (PII) to a Gen AI tool.
- I have ensured that no part of my assessment copies or paraphrases Gen AI outputs without acknowledgement.
- I have ensured that my assignment / research remains my own work.
- I have appropriately referenced the use of Gen AI tools in my work.
- I have saved copies of Gen AI outputs used in preparing my assessment. These copies may be asked to be provided as an appendix to my assessment or as part of any misconduct process.
- I have sought help from my tutors or the Library's Learning Development team when I was unsure about use of Gen AI or needed further guidance.

Please note that this checklist is not exhaustive and provides a framework for ensuring good academic practice is followed when using Gen AI tools.