College and Digital Pathways Staff Guidance: Use of Generative Artificial Intelligence

Kaplan International Pathways is dedicated to supporting students and college and digital pathways staff in the responsible and ethical use of generative Artificial Intelligence.

Providing our students with the skills they need to be successful now and in the future is essential. With the continuous advancements in generative AI, we have committed to reviewing this guidance **monthly and providing updates as needed** during the 2023 – 24 academic year.

Feedback

Your feedback matters. Please let us know how this guidance can be improved. We will use these ideas to inform the monthly updates.

Updates

September 2023: Academic Integrity

October 2023: New and improved guidance on the VLE including:

For students: an updated and refreshed Academic integrity for students | Pathways (kaplaninternational.com) package which includes a section on generative AI in the Student Success Area | Pathways (kaplaninternational.com) on the VLE.

For staff: a new Artificial Intelligence | Pathways (kaplaninternational.com) section on the Pathways Learning, Teaching and CPD Space | Pathways (kaplaninternational.com) on the VLE.

November 2023: Reminder to critically evaluate generative AI in the 'Referencing the use of generative AI' section

December 2023: Student Declaration – position statement updated in line with discussions held by Generative AI Working Group.

January 2024: The 'Generative AI: defined' section has been updated to remain current. ChatGPT has now been trained on data up to January 2022. The QAA has more comprehensive guidance available: Generative Artificial Intelligence (qaa.ac.uk). Further reading has been relocated to the Artificial Intelligence | Pathways (kaplaninternational.com) section on the Pathways Learning, teaching and CPD site.

February 2024: Link to the Summary of the AI SWOT with recommendations (kaplaninternational.com) added.

Next update: March 2024

Contents

Generative AI: defined

TIP: **It is important that we learn more about artificial intelligence tools and have discussions to support students to understand the strengths and weaknesses, as well as the ethical concerns associated with the use of AI.**

Generative Artificial Intelligence (generative AI or Gen AI) is a type of artificial intelligence software that produces outputs such as text, images, computer code, audio or video. At a basic level, generative AI works by analysing vast sets of data, identifying patterns, and producing output based on what a human would be statistically likely to create (JISC, 2023). Typically, these Large Language Models (LLMs) such as ChatGPT or Bing produce unique responses every time (JISC, 2023) and are often interacted with through a 'chat-bot' interface using 'prompts' in a question-and-answer format.

While there has been a recent increase in the use of generative AI tools, artificial intelligence more broadly is prevalent in the technology we use day to day. Common applications of artificial intelligence include predictive text, search engines, recommendations from music or TV streaming services or digital voice assistants.

Though OpenAI's products such as ChatGPT are currently the most prominent, it is important to remember that there are other types of generative AI software available. Examples include:

Microsoft Copilot (microsoft.com) (Microsoft products using OpenAI's technology)

Google Bard - (bard.google.com) (Google's AI tool based on its own models)

Dall.E (openai.com) (can produce images)

Midjourney (midjourney.com) (can produce images)

Humata (humata.ai) (can find and summarise sources)

Elicit (elicit.org) (can create literature reviews)

It is likely that the references to specific AI software in this document will quickly become outdated as new and more powerful as AI systems are released. For a more detailed insight into each of these tools see A Generative AI Primer - National centre for AI (jiscinvolve.org) for an HE specific primer that has been updated this month, January 2024.

Nevertheless, although particular platforms may change, there are some principles that we can apply to student use of *any* powerful generative AI platform.

What are the strengths and weaknesses of generative AI?

TIP: **Critically evaluate the content generative AI produces. At its most basic, generative AI is a source of information. A useful checklist to support the critical evaluation of information produced by generative AI has been produced by the University of Newcastle:** Critical Evaluation of generative AI information.pdf (ncl.ac.uk)

As an example of why it is important to evaluate the output of generative AI, consider the example in Figure 1. ChatGPT was asked to complete a calculation and then provide an explanation. You'll notice that the first answer of 261 was different to the final answer of 277.

When generative AI gets things wrong, they are sometimes referred to as a 'hallucinations'. In The A-Z of AI: 30 terms you need to understand artificial intelligence (bbc.com) from the BBC, it explains "[s]ometimes if you ask an AI like ChatGPT, Bard or Bing a question, it will respond with great confidence – but the facts it spits out will be false" (Fisher, 2023).

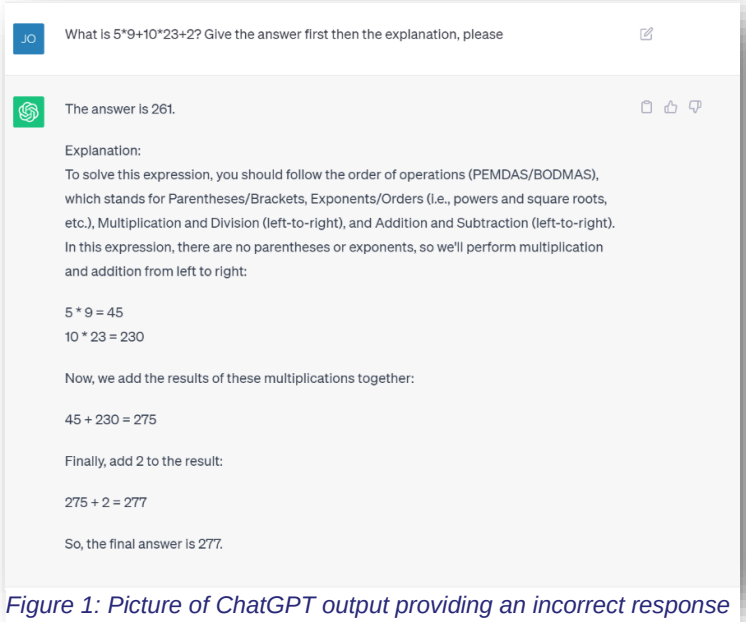

Figure 1: Picture of ChatGPT output providing an incorrect response to a maths question (OpenAI, 2023).

This confidence is particularly worrying considering the database ChatGPT learned from was the internet up to January 2022 (JISC, 2024), and consequently, there is a significant risk of bias in its responses. As Fisher (2023) points out, this "discrimination would be obscured by supposed algorithmic impartiality". Additionally, only the paid ChatGPT-4 has access to the internet. The free version is, therefore, not in real time.

Figure 2: Picture of ChatGPT output being out of date (OpenAI, 2024)

Using the internet as ChatGPT's training database not only introduces cultural and political bias (JISC, 2024) but also means it has limited knowledge post January 2022. For example, consider the output from ChatGPT in Figure 2.

Generative AI is very good at creating structured texts following common genres and format conventions, as well as applying the requested style, tone and voice (UNESCO, 2023). This means it is particularly useful for rapidly generating persuasive texts such as essays and creating highly formatted information such as computer code.

In analysing ChatGPT 3, which is now out of date by two versions, Sobieszek & Price (2022) argue that LLMs prioritise being "plausible instead of truthful". The way in which ChatGPT acknowledges sources when asked to produce a text following academic writing conventions highlights this. Often, a well-structured, very plausible looking essay produced by ChatGPT will be punctuated by entirely fabricated citations and references. It is well worth bearing in mind that these are the sources we can see. The unreferenced information in the output is far more difficult to evaluate and verify. This challenge is compounded by the plausible presentation of the text that can "hijack our intuitive capacity to evaluate the accuracy of its outputs" (Sobieszek & Price, 2022).

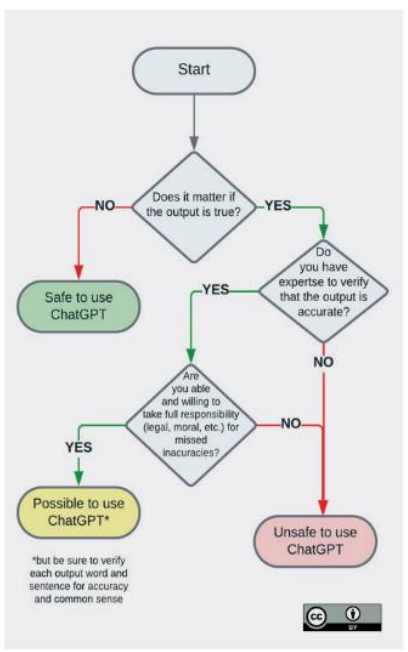

Figure 3: Flowchart devised by Aleksandr Tiulkanov, AI and Data Policy Lawyer, January 2023

TIP **Given the strengths and weaknesses outlined, consider following the simple decision path (Figure 3.) when considering using generative AI. Remember, you must follow the Kaplan GenAI Do's and Don'ts (PDF) when using generative AI.**

The following table from JISC (figure 4.), summarises the key strengths and weaknesses of generative AI and acts as a useful summary.

Capabilities	Limitations	Concerns
• It can write plausible sounding text on any topic. • It can generate answers to a range of questions, including coding, maths-type problems and multiple choice. • It is getting increasingly accurate and sophisticated with each release. • It generates unique text each time you use it. • It's great at other tasks like text summarisation.	• It can generate plausible but incorrect information. • ChatGPT is only trained on information up until January 2022 (but those with the paid ChatGPT Plus service have access to a version that can access the internet, and has a slightly later training date) • Limited ability to explain the sources of information for its responses (this varies between Chatbots)	• It can and does produce biased output (culturally, politically etc) • It can generate unacceptable output. • It has a high environmental impact, concerns around human impact and ownership of training material. • Security and privacy concerns around the way users' data is used to train the models. • There is a danger of digital inequity.

Figure 4: JISC. (2024) "A Generative AI Primer". A Generative AI Primer - National centre for AI (jiscinvolve.org)

Student use of generative AI

TIP **It is important that we discuss generative AI with students to help them understand more about:**

- **how reliable it is,**
- **what acceptable use looks like, and;**
- **how to reference and acknowledge its use.**

This section of the guidance should be read in conjunction with the Student Guidance | Pathways (kaplaninternational.com). The student facing guidance is more concise and tries to outline acceptable uses of generative AI while offering advice on how to avoid Academic Integrity issues when submitting summative assessments.

If students use generative AI for their summative assessments, they must acknowledge its use by following the referencing guidance. Colleges may also wish to ask students to complete a declaration before submitting their work, as outlined below.

Academic Integrity Policy

TIP **We should not assume that a strong piece of work is plagiarised or the result of generative AI. Suspicions of unacknowledged use of artificial intelligence tools must be accompanied by evidence.**

The development of Generative Artificial Intelligence continues to evolve rapidly. It is recognised that it is a useful tool that if used properly can support student learning. It is also a tool that could be used inappropriately to obtain an unfair advantage.

Submissions of summative assessments suspected to be generated either wholly or in part by an AI tool without appropriate referencing should be flagged as suspected **plagiarism** by the marker. Any suspected cases will be dealt with via the existing Academic Integrity process as outlined in Chapter 7. Academic Integrity | Pathways (kaplaninternational.com) in the Academic Standards and Quality Manual | Pathways (kaplaninternational.com).

TIP **The Quality Assurance Agency (QAA) has compiled a comprehensive collection of resources related to generative A: Generative Artificial Intelligence (qaa.ac.uk).**

Referencing the use of generative AI

The use of generative AI must be appropriately acknowledged within students work following the Harvard APA standard How to cite ChatGPT (apastyle.apa.org).

The student guidance has been updated to remind students to critically evaluate the information provided by generative AI tools and, where possible, use a second source.

Citation

(OpenAI, 2023)

Quotation

When given a follow-up prompt of “What is a more accurate representation?” the ChatGPT-generated text indicated that “different brain regions work together to support various cognitive processes” and “the functional specialization of different regions can change in response to experience and environmental factors” (OpenAI, 2023; see Appendix A for the full transcript).

Referencing

Open AI, (2023), ChatGPT 14 March [Large Language Model] https://chat.openai.com/chat

Images

Julius Caesar ‘Image’

Note. Image generated with the prompt “Julius Caesar” by OpenAI, ChatGPT, 2023 (https://chat.openai.com/chat).

Student Declaration

A recommendation by the Quality Assurance Agency (QAA) is that assessment submissions could include a student declaration (which could be via Turnitin or an additional cover sheet) that work produced is their own, and that any sources of information including assisted technology has been fully acknowledged (QAA, 2023). You may wish to consider adding a declaration to an assessment for a module in your college's portfolio of local modules i.e., those not taught at another college. Adding a declaration to a summative assessment is a local decision for each college.

The following is a suggested format from Monash University (monash.edu) in Australia that students can follow to declare the use of generative AI content in a summative assessment. You may wish to adapt the wording or style for your assessments, in consultation with your Academic Leader (or equivalent).

I acknowledge the use of [insert AI system(s) and link] to [specific use of generative AI]. The prompts used include [list of prompts]. The output from these prompts was used to [explain use].

TIP **This has not been included in the student guidance. If this is required on an assessment you must include it as part of the assessment instructions.**

Prevention and Detection

An important concern around the use of generative AI is how students use it in their summative assessments. We need to work with students to guide them in the responsible use of generative AI tools in both their learning and summative assessments.

AI detection software **should not be used** to identify or evidence suspected breaches of academic integrity. We must maintain the integrity we expect of our students in our approach and avoid bias and inconsistency. All submissions are subject to the same transparent standards. Software-driven AI detection tools have been shown to be unreliable (Weber-Wulff et al, 2023) as well as carrying significant privacy, equity and data security risks due to a lack of transparency.

Remember, processing student information through unauthorised external tools is a significant General Data Protection Regulation (GDPR) risk. Do not use external AI detection tools with any student assessment or information.

Similarly, there are difficulties for Academic Integrity Panels when comparing work suspected to have been generated by AI to previous student work. It is difficult to verify that the previous work is the students without using a sample taken from an invigilated exam, which would not be fair because the two texts will have been written under different expectations. Panels must take care to ensure they are comparing work of comparable format from the same time period. For more information, see Chapter 7. Academic Integrity | Pathways (kaplaninternational.com) in the Academic Standards and Quality Manual | Pathways (kaplaninternational.com).

In the recent update to the Academic Standards and Quality Manual | Pathways (kaplaninternational.com), specific reference was made to offering the student an opportunity to demonstrate that the work was their own during the Student Academic Integrity Meeting. This is a useful chance to discuss the students work in cases where there is evidence of unacknowledged generative AI use that has led to a suspicion being raised.

TIP **It is important that students are given an opportunity to demonstrate the work is their own. The student is not being reassessed on the Learning Outcomes of the assessment. The Academic Integrity process, and especially the Student Academic Integrity Meeting, is a learning experience that should improve the students understanding of Academic Integrity no matter the outcome.**

The following measures have been compiled from a range of sources to support colleges with the prevention and detection of unacknowledged use of generative AI tools.

- **Communicate with students** on the limitations and potential harm of such tools to them and their education (QAA, 2023) but also the positive applications and use cases.
- **Follow existing processes**, maintain fairness and demonstrate the integrity we ask of our students.
- Encourage **students to submit drafts**, not only will they benefit significantly from the feedback but they also have an early chance to avoid any academic integrity issues by showing the genesis of their work.
- **Identify high-risk assessments** such as written assignments that focus on the retrieval of facts or opinions.
- Check what is referenced, AI tools will often use **fabricated citations** (Cotton et al., 2023).
- **Discuss these assessments with colleagues** and agree common attributes you would expect to see in a typical submission such as linking a personal reflection to real events.
- **Familiarise yourself with the output generative AI tools** produce by entering past questions – do not use current questions as this could improve the accuracy of the model.

TIP **On the 30th August, Andy Rosen contacted staff with an invitation to experiment with generative AI tools by providing access to an online platform specifically designed to provide colleagues with a safe tool to experiment with. The toolkit can be accessed here | (createwithdaisy.com). Colleagues are encouraged to use this safe and controlled platform to experiment with their assessments while it is available.**

Colleagues will have two-months to access the platform. The sign-up instructions are as follows:

"*Skip the payment options, click on the sign in button on the upper right side of the screen, and enter your work email address (only Kaplan domains will be recognized) and you will be emailed a link to create a new password and login for the platform*."

AI tools tend to produce similar outputs when given the same prompts such as a specific coursework question (Cotton et al., 2023). Take considerable care to maintain fairness by reviewing all submissions with this in mind if a suspicion is raised. Early examples of formulaic submissions may be missed as a baseline is established so further evidence should be used to support any suspicion raised.

TIP **It is recommended that academic teams agree a list of the potential indicators of unacknowledged use of generative AI tools. This may be at the college or subject level depending on the context. These can be shared with students and applied equally and consistently to all marked work.**

- The following is an abridged list of potential indicators from the Joint Council for Qualifications:
- A default use of American spelling, currency, terms and other localisations
- A lack of direct quotations and/or use of references
- Inclusion of references which cannot be found or verified (some AI tools have provided false references to books or articles by real authors)
- A lack of reference to events occurring after a certain date (ChatGPT v3.5 is trained on data up to January 2022)
- Instances of incorrect/inconsistent use of first-person and third-person perspective
- A variation in the style of language throughout the work
- A lack of graphs/data tables/visual aids where these would normally be expected
- A lack of specific local, topical or personal knowledge
- Content being more generic in nature rather than relating to the student themself, or a specialised task or scenario, if this is required or expected
- The inadvertent inclusion by students of warnings or provisos produced by AI to highlight the limits of its ability, or the hypothetical nature of its output

None of the above indicators alone are evidence of students using generative AI tools without acknowledgement. When marking teams agree a list of potential indicators and a number of them are spotted in concert with other suspicions, they pay prove useful in supporting the academic integrity process. It is also recommended that the list is shared with students, so they know what is expected of their work.

Sharing practice

TIP **Review the summary and recommendations from the large-scale AI SWOT that was undertaken across all colleges here:** Summary of the AI SWOT with recommendations (kaplaninternational.com)

It is important to share practice to further consolidate our experience and deliver a consistent approach via our existing forums of networks; Academic Management Group, Curriculum and Assessment Teams (formerly Assessment Collectives), Academic Integrity Working Groups etc.

TIP **Your feedback matters. Please feedback on your thoughts and experiences using generative AI. This is a developing situation and we need to work together to ensure students have the best possible opportunity for success in their studies and future careers.**

References

TIP **All of the following references have hyperlinks at the end for ease of access.**

Cotton, D., Cotton, P., and Shipway, J. (2023). “Chatting and cheating. Ensuring academic integrity in the era of ChatGPT.” Chatting and cheating. Ensuring academic integrity in the era of ChatGPT (doi.org)

Fisher, R. (2023, July 20). The A-Z of AI: 30 terms you need to understand artificial intelligence. BBC Future. What you should know about artificial intelligence from A - Z (bbc.com)

JCQ. (2023) “AI Use in Assessments: Protecting the Integrity of Qualifications.” AI Use in Assessments: Protecting the Integrity of Qualifications (jcq.org.uk)

JISC. (2024) “Generative AI - a primer”. Jisc National Centre for AI A Generative AI Primer - National centre for AI (jiscinvolve.org)

OpenAI. (2023). ChatGPT (Sep 05 version) [Large language model] http://chat.openai.com/chat

QAA. (2023) “The rise of artificial intelligence software and potential risks for academic integrity: A QAA briefing paper”. The rise of artificial intelligence software and potential risks for academic integrity: A QAA briefing paper for higher education providers (qaa.ac.uk)

Sobieszek, A., Price, T. (2022). Playing Games with Ais: The Limits of GPT-3 and Similar Large Language Models. Minds & Machines 32, 341–364 Playing Games with Ais: The Limits of GPT-3 and Similar Large Language Models (doi.org)

UNESCO. (2023) “ChatGPT and Artificial Intelligence in higher education.” ChatGPT and Artificial Intelligence in higher education (unesco.org)

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J.G., Popoola, O., Sigut, P., & Waddington, L. (2023). Testing of Detection Tools for AI-Generated Text. ArXiv, abs/2306.15666. Testing of Detection Tools for AI-Generated Text (arxiv.org)

Further reading

Further reading on generative AI can be found on the dedicated Course: Artificial Intelligence | Pathways (kaplaninternational.com) page in the Pathways Learning, Teaching and CPD Space (Pathways Learning, Teaching and CPD Space | Pathways (kaplaninternational.com) on the VLE.