Imperial College London

Assessment Guidance

Working Group on Artificial Intelligence Tools in Teaching and Assessment

Aug 2023

Table of Contents

1. Introduction

The advent of widely available Artificial Intelligence (AI) content generation tools (*e.g.,* Fig. 1, below) makes it timely that we evolve the education we provide and review our teaching and assessment practices.
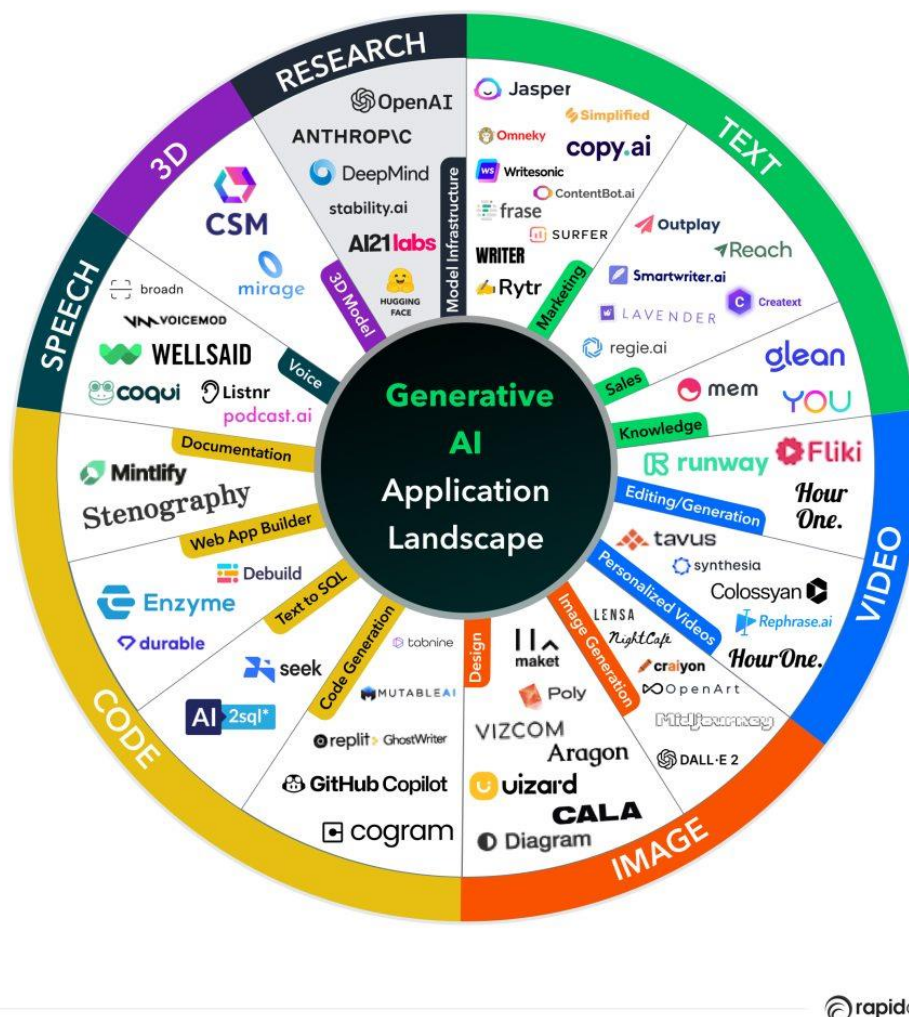

Figure 1. Examples of AI content generation tools (from: https://www.rapidops.com/blog/generative-ai-tools/)

Imperial College, other universities (*e.g.,* LSE, UCL), HE-representing bodies (*e.g.*, DfE, QAA, EUA, Russell Group), and tech-focussed bodies (*e.g.,* JISC, JCQ) *etc.* have published position/briefing/guidance/statement papers on the use of AI content generation tools in Education. The five 'principles' published by the Russell Group (with contributions from Imperial staff) encapsulate the overall approach of UK research intensive Universities:

1. Universities will support students and staff to become AI-literate.
2. Staff should be equipped to support students to use generative AI tools effectively and appropriately in their learning experience.
3. Universities will adapt teaching and assessment to incorporate the ethical use of generative AI and support equal access.
4. Universities will ensure academic rigour and integrity is upheld.
5. Universities will work collaboratively to share best practice as the technology and its application in education evolves.

The requirement to maintain academic integrity in our student assessments (principle 4, above) is the focus of this guidance document.

UK universities are required by the OfS to ensure that student assessments are valid, reliable and 'designed in a way that minimises the opportunities for academic misconduct and facilitates the detection of such misconduct where it does occur' ([OfS Nov 2022](#)). So, we need to ensure that our assessments are robust and that they provide opportunities for students to demonstrate their achievements against the stated learning outcomes through their original work. This is imperative for us to maintain the academic standards of our degrees and qualifications.

Of course, we should embrace the opportunities that AI offers to rethink and extend our teaching, learning and assessment practices. These tools are increasingly used in work and life, and we need to consider how we design authentic assessments and help our students learn how to use AI content generation tools productively, responsibly, and ethically. A 'Blue Skies thinking' sub-group of the College's Working Group on AI Tools in Teaching and Assessment has been established to help with this.

The guidance provided here has been provided by the 'Assessment Guidance' sub-group of College's Working Group on AI Tools in Teaching and Assessment. It is designed to help staff in the short term, and particularly in the run up to and during the 2023-24 Academic Year, to critique and adapt their existing assessment practices and embrace the opportunities afforded by AI content generation tools while minimising the potential for academic misconduct.

2. General Guidance for all Assessment Types

Departmental Guidance

Imperial takes challenges to academic integrity and to the value of its degrees with the utmost seriousness. Consequently, the College has detailed regulations and processes for ensuring academic integrity in formative and summative work. This Policy has been updated (May 2023) to provide 'an updated definition of academic misconduct in response to the proliferation of AI language models'. Specifically, the unacknowledged use of AI language models or other processing tools may now be considered plagiarism and/or dishonest practice as they confer an unfair advantage.

UG and PGT Departmental Teaching Committees (or equivalent committee) should define, in consultation with students where appropriate, what constitutes authorised use of AI content generation tools for students taking modules in their department. These local regulations must be formally endorsed and collated by the appropriate Faculty Education Committee (or equivalent committee) who should look to ensure a consistent approach between departments and programmes - excepting disciplinary nuances/imperatives.[1] Where use of AI content generation tools is authorised, Departments must clearly communicate this to colleagues and to students. This dissemination must include an accessible statement/document hosted on the appropriate Departmental (*e.g.,* examinations) web pages. College regulations will apply in cases where course leaders suspect that AI content generation tools have been used without authorisation. Given that there are no fully effective tools to detect the use of generative AI to date, this position underlines the importance of academic integrity and the ethical academic standards of behaviour that are expected of all members of our community, of which our students are an integral part.

To ensure academic standards are maintained, a department may choose to conduct 'authenticity interviews' on their submitted assessments. This means asking students to attend an oral examination on their submitted work to ensure its authenticity, by asking them about the subject or how they approached their assignment.[2] Students must be selected at random, and it must be explained to students that an invitation does not mean there is any specific concern that they have submitted work that is not their own. Such interviews can be conducted relating to a specific programme, module or element of assessment but must be organised by the appropriate UG and PGT Departmental Teaching Committees (or equivalent committee). As the prospect of the application of this new element of scrutiny to students' assessed work is primarily envisioned to be as a deterrent, if a department does choose to use authenticity interviews, this must be communicated to students at the start of session so that they are aware that they may be requested to take part. Such interviews have the potential to elicit significant anxiety with students and so comms and protocols associated with these should look to assure students that these are for the purpose of ensuring the academic integrity of the assessment.

Programme and Module Leader Guidance

The following specific suggestions may be helpful to staff when formulating assessments for the upcoming Academic Year:

1. Look to work within your approved programme and module specification parameters where possible. Module specifications set out the type of assessment and assessment weighting, but do not specify the questions or assessment criteria and so give latitude for adjusting assessments without the

---

[1] It is anticipated that review of these regulations at the Faculty level and then by QAEC will allow the development of a 1st iteration of College guidance/principles by the end of the Autumn Term AY 2023-24. Given the fast-paced advent of new AI content generation tools, these will likely need to be iterated thereafter on a regular basis.

[2] Detailed guidance for the conduct of authenticity interviews has yet to be developed in terms of what the outcomes of these interviews can be for the students and what constitutes sufficient evidence to assume misconduct. It is planned that QAEC develop this during Autumn Term AY 2023-24.

requirement to seek approval via the provisions in the [Modifications Policy] or report minor modifications to the Registry.

2. Familiarise yourself with the main AI content generation tools available (*cf. e.g.,* Fig. 1, above) and particularly those suited to the specific academic discipline. Assess their strengths and weaknesses, the types of assessment they can solve/aid answering, and whether the assessments you set are susceptible. However, be aware that if an upcoming assessment is submitted into an AI content generation tool, the tool will usually retain this data (although some offer opt-outs). This has the potential to compromise the integrity of the assessment as the data will eventually become 'public domain', although there is a latency associated with Reinforcement Learning from Human Feedback (RLHF) cycles and hence incorporation of inputted data into the Large Language Model (LLM). It is therefore strongly recommended that you only input *past* assessments into AI content generation tools to further your understanding of what these tools can do and how students may use them.[3]

3. Familiarise yourself with the types of uses that students report that they are using AI content generation tools from. These include the following (from [Lancaster University SU survey]):
   a. To assist with debugging code.
   b. To find out how to get started with a programming assignment.
   c. To explain module content and topics in plain language or summarise into key points.
   d. To generate ideas – help with getting started on an assessment.
   e. To revise – generating questions and model answers, explaining questions, explanations around how to approach specific types of questions/problems, to check their own answers to questions.
   f. To generate written coursework – ranging from generating essay outlines, paragraphs that are then paraphrased, to entire essays.
   g. To support neurodivergent student needs around grammar/punctuation/sentence structure and lengthy content summarisation (both also helpful for students who might not class themselves as neurodivergent).
   h. To reword text.
   i. For locating references (*e.g.,* using the specialist AI tool – [CiteAI]).
   j. To generate titles for coursework submissions (including dissertations).
   k. To write notes.
   l. For initial research and revision.
   m. To create mnemonics and flash cards to aid learning.
   n. For answers to online exams (both text and mathematical answers).

4. Provide clear guidance to students on what will be authorised/deemed appropriate and what will be unauthorised/deemed inappropriate use of AI content generation tools for every assessment.

5. Use of AI content generation tools might be authorised/deemed acceptable for certain purposes, for example:
   a. To improve the grammar of self-written work.
   b. To prompt new idea generation.
   c. To generate simplified explanations of complex topics to support learning.
   d. To provide formative feedback.
   e. To provide stylistic suggestions
   f. For preparatory topic area orientation searches

---

[3] The Business School IDEA Lab have introduced a '[Generative AI Stress Test]' for their taught module assessments. This comprises a four-phase process (Analyse, Review, Collaborate, Adapt) designed to evaluate and mitigate potential vulnerabilities in their assessments due to potential use of generative AI tools by students and to provide insights for staff into how students use AI tools. The focal point of the stress test comprises a meeting between a faculty member and staff from the IDEA lab in which past indicative assessments are tested using *Chat GPT4, Bard* and *Claude 2* against 6 criteria: Accuracy, Clarity, relevance, compliance, referencing, and ease of use. A variant of this approach was used in the College's workshop on *Transforming Education: Adapting Teaching and Assessment with Artificial Intelligence* (21/08/2023) facilitated by the EDU. Further such workshops may be provided by the EDU if there is demand.

6. Use of AI content generation tools might also be embraced in an 'educative' approach to their use in some assessments, to give the assessment greater authenticity (given that AI content generation tools are becoming commonplace in our lives and careers). This might involve assessment re-design to make the use of AI content generation tools integral to their completion. For example:
    a. Allow students to use AI content generation tools for *X* number of questions, clearly describing how they used the tool and citing the source, potentially with a permalink where available, while also commenting/reflecting on the outcome.
    b. Allow students to use generative AI to answer specific/complex questions and assess the prompt students use to "chat" with the generative AI. If questions are specific/complex, students need to use their knowledge in the subject to create the prompt.
7. In all cases where students are permitted to use AI content generation tools, it is advisable to:
    a. Stipulate that students reference/state/acknowledge which AI content generation tools they used and for what purpose, for example in the form of a template/coversheet. The Library provide [Guidance on Referencing](#).
    b. Strengthen academic integrity and plagiarism awareness training for students and review the uptake of the training. The Library provide [Guidance on Plagiarism Awareness](#) and on [Use of Generative AI Tools](#).
    c. Introduce training for students to inform of the key features of AI content generation tools and in particular their limitations (*e.g.,* hallucinatory referencing, fabricating 'facts', lack of post 2021 'knowledge' *etc*.).
    d. In line with College's expectation that departments provide access to all software packages required by students to succeed in their studies, ensure that all students have access to the requisite AI content generation tools.[4] Consider the advantage that paid subscriptions to tools such as ChatGPT can give to students who can afford it.
8. Where use of AI content generation tools is not authorised, take measures to prevent their unauthorised use. For example:
    a. Prevent students access to electronic devices during examinations – *e.g.,* conventional invigilated examination conditions or 'open book' exams allowing access only to hard copy notes.
    b. For examinations which are sat remotely or allow students to 'bring their own devices', apply appropriate 'lock-down' technologies as supported by ICT/Faculty EdTech teams/Exam Support Teams [*e.g.,* Lexis (Linux), Safe Exam Browser (Windows)].
    c. Do not rely on AI detection software. This is unproven technology that does not provide a robust basis for an academic misconduct investigation. In line with most UK HE providers and based on research suggesting that such tools are unreliable, College has currently opted not to support Turnitin's AI detection module.
9. Where use of AI content generation tools is not authorised, ensure that assessments and assessment criteria are developed to assess learning outcomes that test and give credit for higher-order skills that cannot yet be well replicated by AI content generation tools [*e.g.,* critical thinking and synthesis of new ideas, providing nuance, creativity and context, abductive reasoning (*e.g.*, see below)], and/or for which use of AI content generation tools is not advantageous/useful (*i.e.,* it would take more time for a student to explain to the AI content generation tool what they want, than to actually write up the answer themselves). For example:
    a. Devise questions whose answer requires discussing individual experiences/reflections on practical/hands-on applications that took place during term time, thus making them too context-specific for an AI content generation tool to properly address.
    b. Prepare questions whose answer requires making decisions based on a diverse set of data. For example: "***Solely*** *based on all the [examinable material] of this module so far and nothing else, what would X Y Z?*". The goal here is to limit the answer to a specific subset of knowledge too varied for AI to consume and/or isolate, such as specific/in-house PDF/PowerPoint

---

[4] Currently, OpenAI do not provide corporate or institutional licences for their subscription ChatGPT4 product.

slides/tutorials/lecture notes, Panopto videos, EdStem posts, *etc.*, so that, if an answer contains something outside of these constraints, it will raise alarms.

c. Construct an abductive-deductive-inductive process such as the following – with assessment potentially at any stage:

    i. *Abduction phase*: students are given a problem to solve that has some level of ambiguity or complexity e.g., why are so many obese in the UK? There are lots of inter-related causal factors involved in obesity, meaning that there isn't a singular root cause. Students are given, or asked to find, data relevant to the problem and to develop some conjectures (hypotheses) that could explain the problem (obesity). They would have to explain their thinking at this stage and an assessment could test the quality of the rationales for the competing conjectures.

    ii. *Deduction phase*: students start testing their hypotheses, which would require them to work out which tests were appropriate to each hypothesis. Again, they'd need to explain and justify their choices.

    iii. *Induction phase*: here the students would try to explain the problem (obesity) using the data they've used and hypotheses they've tested. They would need to explain how these, or combinations of them, contribute to understanding the problem.

d. The FoM have developed some guidance for developing authentic assessments for which current AI content generation tools are poorly suited to answering – see here.

e. Consider group projects and peer assessment.

f. Contemplate process-oriented, staged and/or programmatic approaches to assessment.

g. Ensure that the specific questions posed are original and refreshed regularly, if appropriate.

3. Case Studies for Different Assessment Types

It is planned to provide specific case studies of assessment approaches that exemplify the general guidance provided in Section 2. These will take time to assemble and will be curated as part of the Anatomy of Assessment project led by Prof. Martyn Kingsbury, Director of the Educational Development Unit (EDU) and of College's Centre for Higher Education Research and Scholarship (CHERS). It is hoped to provide case studies for selected types of assessment from *e.g.,* the following non-exhaustive list:

A. Coursework:

      i.   Essays
      ii.   Theses, dissertations, reports (*e.g.,* postgraduate milestone reports for ESA and LSR)
      iii.   Diary Entries
      iv.   Quizzes
      v.   Reflective writing
      vi.   Problem Sheets
      vii.   Marketing strategies

B. Examination:

      i.   Timed-assessments
      ii.   Written In-class Tests (*e.g.,* MCQ)
      iii.   Programming assignments

C. Practical:

      i.   Lab reports
      ii.   Poster Presentations
      iii.   Oral Presentations
      iv.   Placements
      v.   Clinical Work
      vi.   Oral Examinations
      vii.   Field Trips
      viii.   Group projects
      ix.   Coding-related practical

This section will therefore be updated as case studies are developed.

4. Additional Resources/References

The following sources have not been directly referenced but provided useful background/insights.

Student Facing URLs:

- [Edinburgh Napier](#) - student guidance.
- [SOAS](#) – guidance for students based on use cases, including a table.
- [University of Kent](#) - student guidance.
- [University of Edinburgh](#) – student guidance.
- [University of Exeter](#) – student guidance for referencing AI content generation tools.
- [University of Nottingham](#) – student guidance.
- [University of York](#) – student guidance.

Staff facing URLs:

- [University of Greenwich](#) - responsible use of generative AI in teaching and learning.
- [University of Greenwich](#) - risk assessment for AI in learning and teaching.
- [Queen Mary University](#) - staff guidance.
- [Monash University](#) – assessment guidance.
- [Harvard Graduate School of Education](#) – 'Navigating a World of Generative AI: Suggestions for Educators' (pdf).
- [Hong Kong University of Science & Technology](#) – guide to generative AI tool use in HE.
- [UCL](#) – generative AI Hub.
- [KCL](#) – generative AI guidance for teaching, assessment and feedback