

FHML Examiner Guidelines for the use of Artificial Intelligence (AI) tools like ChatGPT to generate texts

What is ChatGPT?

ChatGPT (Generative Pre-trained Transformer) is free software that generates answers to questions based on a vast knowledge database with relatively high accuracy and readability, hardly distinguishable from human texts. It can produce summaries, answers to essay questions, papers, programming codes, etc.

ChatGPT cannot reason or produce original knowledge. However, it can combine information from different sources in such a confident and convincing way that it gives the semblance of understanding and reasoning.

Why does chatGPT impact education and assessment?

ChatGPT can generate texts that students could use for their assignments without assessors being able to tell the difference between student- and software-generated submissions. This is not necessarily a problem. It becomes a problem for education and assessment when students use ChatGPT to demonstrate knowledge and skills required to meet the learning outcomes of a course, without having acquired such knowledge or skills.

AI tools like ChatGPT bring challenges and opportunities. The main challenge is to ensure that students still obtain the knowledge and skills that we want them to obtain, regardless of whether or not they use ChatGPT. The main opportunity is to use it in education and to approach our assessments differently, in creative ways tailored to our changing world.

ChatGPT is now a part of our world. This means that we have to tolerate or even facilitate its use as long as we ensure the quality of our education and assessment. Below, we provide suggestions for doing so.

Note that you are not required to implement all suggestions and that some suggestions merely emphasise what education and assessment should already look like within a PBL system that is based on CCCS and UM's vision on assessment. Also, note that the Board of Examiners may instigate specific guidelines and requirements for assessments and exams in the remainder of the current academic year and for the next academic year.

How could you adapt your course to AI tools like ChatGPT to ensure the quality of education and assessment?

When designing and conducting teaching activities:

- Educate students on the use of AI tools like ChatGPT: what is it?, what can it produce?, and what are pitfalls and disadvantages?
- Have students critically relate different sources instead of merely summarising them.
- Have students provide key quotes from their readings and explain why they capture the gist of what they read.
- When using AI tools like ChatGPT in education: Have ChatGPT answer learning goals, then have students discuss these answers based on their readings.

When designing assessments:

- Add non-written components (e.g. presentation, poster, debate, oral exam or inquiry) to written assignments produced at home.
- Consider use of oral exams.
- Have students integrate content from the course into their written assignments (with specific references to the literature used in your course).
- Have students provide reasons for why they (dis)agree with something, referring to learning materials used in the course.

When using AI tools like ChatGPT in assessment:

- Have students use ChatGPT to formulate an answer to a question, and then have them critically assess this answer with arguments based on the learning materials used in the course (with specific references).

- Ask students to mark text that is generated by ChatGPT or another AI tool in their assignments or exams.
- Ensure that the necessary intended learning outcomes are still assessed by checking the work produced by the students and not the AI tool. For example, by asking students to critically reflect on the answers generated by the AI tool or by comparing the AI generated answer with their own answer (or specific sources).

Important:

To prevent any misunderstandings, include in the course syllabus and in the exam instructions that students are not allowed to use ChatGPT or other similar text or image-producing AI tools in assignments, presentations or exams:

"Unless otherwise mentioned, AI generated text used to answer exam questions/assignments can be seen as commissioned work that represents plagiarism and fraud and will be sanctioned under the Rules and Regulations."

When assessing assignments:

- Look for unusual language or formatting: texts (partly) produced by ChatGPT may be atypical of student work, with unusual or repeated phrases or differences in style, syntax, spelling and punctuation between different sections of the same text; pay extra attention to the content and reasoning of such sections/texts.
- Currently, there are a number of AI text detection tools (see appendix). These currently available software tools are still beta versions, which provide probabilistic results. It is not yet known how reliable their checks are. Therefore, we advise you to use them cautiously and only as a means to confirm several other findings (e.g. writing style as mentioned above, lacking claims, no real argumentation, meagre referencing, fake sources).
- Always check for plagiarism in written assignments: texts produced by ChatGPT may contain excerpts from other sources without proper attribution, which plagiarism detection software such as Ouriginal can detect.
- Always forward your suspicion to the Board of Examiners that will investigate it and -if confirmed- sanction the student.
- The Board of Examiners will sanction exam work that was AI generated by considering it plagiarism (commissioned work) and/or fraud (an action that makes it impossible to evaluate the student's knowledge).

Important information about assessment methods in your course

- The Boards of Examiners (BoEs) will request the programme coordinators / programme management to make a risk analysis of current and foreseen assessments and exams, for period 5-6 of the current academic year and for the upcoming academic year (be mindful that assessment plans of period 1 and course-overarching trajectories need to be ready before the end of this academic year). What are the risks that AI tools are used for the various assessment formats and in the longitudinal trajectories, and what are the consequences? What needs to change given the potential use of AI tools?.
- Diversity of assessments within each programme is key. The BoEs require that, as of academic year 2023-2024, at least 30% of the (written) assignments in a programme are full (student) identity proof. Examples of methods that fulfil these requirements are oral exams, presentations, and supervised MECC Exams (Testvision or on paper).
- Monitoring different stages in the production of longer papers or a thesis is advisable. E.g. integrate in-between presentations / oral inquiries and use these in the assessment, and/or as part of the internship assessment.
- Add to assessment plans that when there is suspicion of use of an AI tool in a written assignment/assessment, the coordinator may perform an additional inquiry to make sure that ILOs are met, e.g. via an oral exam, written assignment under surveillance etc.

Disclaimer

Open AI text generators, such as ChatGPT are constantly evolving, which means that a fix today may not be a fix tomorrow. Therefore, our list of suggestions will be periodically updated.

Appendix: currently available AI text detection tools

So far, at least five AI text detection tools are available on the internet:

<https://gptzero.me/>

<https://openai-openai-detector.hf.space/>

<https://gltr.io/dist/index.html>

<https://writer.com/ai-content-detector/>

<https://platform.openai.com/ai-text-classifier>

At the moment there is limited experience, and as such no preferred tool can be indicated yet. We mention these here, so you are aware that in case of suspicion, these are tools that could be used for further investigation. We will keep you updated on the status of these and other tools that might become available.

NB: For GPTzero, the text should be at least 10 sentences long with a maximum of 5000 characters. You also have the option of uploading a paper via the upload button.

For Open AI, you can enter max. 510 tokens (characters, spaces); we recommend around 200 tokens (characters, spaces) per sample. The result is considered reliable after around 50 tokens but the more text is included, the more reliable the outcome is.

The AI Text Classifier (platform.openai) requires a minimum of 1,000 characters (approx. 150-250 words).