

ChatGPT: Opportunity or death knell for assessment as we know it?

In the last few weeks, a lot of column inches in the media have been devoted to a new piece of software, ChatGPT. Developed by OpenAI and launched in November 2020, it uses Artificial Intelligence (AI) approaches in a programme that can formulate answers in a conversational manner to questions and follow-up questions, from the most basic to the highly sophisticated.

According to the OpenAI website, ChatGPT has been made freely available “...to get users’ feedback and learn about its strengths and weaknesses. During the research preview, usage of ChatGPT is free” (<https://openai.com/blog/chatgpt/>). A further iteration of the software, ChatGPT Plus, developed from the feedback on the initial package, was launched on the 1st of February, but this time there is a subscription fee of \$20 (just over £16) per month. For now, this package is available to customers in the United States and the freely available Chat GTP will remain available, but it is only a matter of time before the more advanced package becomes universally available along with the offerings of other providers in this field.

So, what is ChatGPT, and why has it caused such interest?

ChatGPT is what is described as a language learning model, and an example of a developing market. For example, Microsoft announced an extended partnership with OpenAI for a “...multi-year, multi-billion dollar investment to accelerate AI” (Microsoft, 2023). These language learning models can draw on literally millions of pieces of information from the internet to formulate a response to a question, although much of the content available to ChatGPT is based on 2021 data, though that will undoubtedly change to more up-to-date material, from which it builds a model, based on statistical probability, of the words and sentences that tend to follow whatever text came before. And of course, it uses the information according to the rules written by human programmers. In some ways, an advanced version of the predictive text facility we are all familiar with on our mobile phones.

But it is more sophisticated than that in its use of ‘reinforcement leaning’ a technique “...which uses human feedback in the training loop to minimize harmful, untruthful and/or biased outputs” (Ramponi, 2022)

Since its launch, many commentators have noted its ability to produce answers to questions in fluent language in the form of “convincing essays – even if some facts it uses aren’t strictly true” (Hern, 2022) leading some to suggest AI poses a serious threat to authentic and credible assessment across the age range, but particularly for GCSE, A Level, degree and vocational qualifications where students routinely produce unsupervised written work.

The published aims of the AI developers “...to accelerate AI breakthroughs to ensure these benefits are broadly shared with the world” (Microsoft, op. cit.) appear laudable, and the capability to generate lesson plans for example looks inviting. But one of the questions some people are beginning to ask is familiar in the world of educational assessment is will this software fundamentally undermine the integrity of coursework, assignments and essays?

Spotting plagiarised or inauthentic work is nothing new – schools, colleges and universities have invested considerable time, effort and money in recent years to keep up with the rise in software programmes: software such as Turnitin, Grammarly and Quetext are examples. There has also been a rise in what are described as ‘essay factories’ that can, for a fee, produce answers for students across the whole range of levels and qualifications. And despite moves to make essay generation services illegal (Gov.UK, 2022), they are still widely available.

Malpractice in UK examinations has been adjudged by the Independent Commission on Examination Malpractice (2019) as being at "...a very low level" with the report stating that "...heads of centre should build and maintain an ethical culture in which malpractice by students and staff does not take place" (p.3).

The question with ChatGPT however, is whether it marks a significant shift which will make it harder if not impossible for assessors to judge the authenticity of students work, be that in internally set homework, or as part of an assessment submitted for a high-stakes qualification such as a GCSE, university degree or license to practice. Put bluntly, the fear is that AI could be used by students to pass off a chatbot's answer as their own (Parker, 2022). That's the downside.

But like many technological initiatives past and present, it also has huge potential, much of which we are only starting to understand at this early stage. ChatGPT and other similar programmes also have the potential to have a positive impact on teaching, learning and, yes, assessment, across all age ranges and settings. Such programmes offer tremendous scope for independent learning, enabling users to access amazing resources that can introduce students to knowledge previously unavailable or unimaginable. In this day and age, students of all ages need to be able to effectively use information and ideas sourced from the internet. But in doing so they also need the insight and ability to judge the validity, accuracy and trustworthiness of what they see and hear – exactly the skills and expertise that teachers and lecturers will help them develop through the establishment of an ethical culture that promotes "...a proper approach to examinations, the need to avoid malpractice of any sort and the sanctions that follow from proven malpractice" (Independent Commission on Examination Malpractice, op. cit. p.33).

The upshot is that we, at the CIEA, like all those professionally engaged in assessing students' and judging their work, need to be aware of what ChatGPT and other such programmes do, and what they mean for our work. One thing is for sure, we are going to be hearing a lot more in the coming months about how AI is going to transform or undermine current assessments. Determining the exact nature of the opportunities and threats and the best way to respond to them is something only humans can decide. No AI programme can do that – at least not yet!

Keeping an open mind about the future of AI's impact on educational assessment is important and it is equally important that we need to reflect on and contribute to an informed discussion on the opportunities and challenges ahead. The CIEA would therefore welcome feedback from members on their views on this development and its impact - positive or negative - on their assessment work.

Please share your views by emailing administratorciea@herts.ac.uk

Gov.UK. 2022. Essay mills are now illegal – Skills Minister calls on internet service platforms to crack down on advertising. The Education Hub, Gov.Uk. Available online from: <https://educationhub.blog.gov.uk/2022/04/28/essay-mills-are-now-illegal-skills-minister-calls-on-internet-service-providers-to-crack-down-on-advertising/>

Hern, A. 2022. AI-assisted plagiarism? ChatGPT bot says it has the answer for that. The Guardian. Accessed online. Available from: <https://www.theguardian.com/technology/2022/dec/31/ai-assisted-plagiarism-chatgpt-bot-says-it-has-an-answer-for-that>

Independent Malpractice Commission. 2019. Report of the Independent Malpractice Commission. Joint Council for Qualifications (JCQ). Available from: <https://www.jcq.org.uk/examination-system/imc-home/>

Microsoft. 2023. Microsoft and OpenAI extend partnership. Accessed online. Available from: <https://blogs.microsoft.com/blog/2023/01/23/microsoftandopenaiextendpartnership/>

Parker, K. 2022. Could an AI bot be writing students' homework? tes Magazine. Accessed online. Available from: <https://www.tes.com/magazine/teaching-learning/secondary/chatgpt-could-ai-bot-be-writing-students-homework>

Ramponi, M. 2022. How ChatGPT actually works. Accessed online. Available from: <https://www.assemblyai.com/blog/how-chatgpt-actually-works/>