

INNOVATIONS

A professor accused his class of using ChatGPT, putting diplomas in jeopardy

AI-generated writing is almost impossible to detect and tensions erupting at a Texas university expose the difficulties facing educators



By [Pranshu Verma](#)

Updated May 18, 2023 at 2:19 p.m. EDT | Published May 18, 2023 at 6:00 a.m. EDT



Human read



Listen 9 min

Students at Texas A&M University at Commerce were in celebration mode this past weekend, as parents filed into the university's Field House to watch students donned in cap and gown walk the graduation stage.

But for pupils in Jared Mumm's animal science class, the fun was cut short when they received a heated email Monday afternoon saying that students were in danger of failing the class for using ChatGPT to cheat. "The final grade for the course is due today at 5 p.m.," the instructor warned, according to a copy of the note obtained by The Washington Post. "I will be giving everyone in this course an 'X,'" indicating incomplete.

Mumm, an instructor at the university's agricultural college, said he'd copied the student essays into ChatGPT and asked the software to detect if the artificial intelligence-backed chatbot had written the assignments. Students flagged as cheating "received a o."

He accompanied the email with personal notes in an online portal hosting grades. "I will not grade chat Gpt s***," he wrote on one student's assignment, according to a screenshot obtained by The Post. "I have to gauge what you are learning not a computer."

The email caused a panic in the class, with some students fearful their diplomas were at risk. One senior, who had graduated over the weekend, said the accusation sent her into a frenzy. She gathered evidence to prove her innocence — she'd written her essays in Google Docs, which records timestamps — and presented it to Mumm at a meeting.

The student, who spoke to The Post under the condition of anonymity to discuss matters without fear of academic retribution, said she felt betrayed.

“We’ve been through a lot to get these degrees,” she said in an interview with The Post. “The thought of my hard work not being acknowledged, and my character being questioned. ... It just really frustrated me.” (Mumm did not return a request for comment.)

The rise of generative artificial intelligence, which underlies software that creates words, texts and images, is sparking a pivotal moment in education. Chatbots can craft essays, poems, computer code and songs that can seem human-made, making it difficult to ascertain who is behind any piece of content.

While ChatGPT cannot be used to detect AI-generated writing, a rush of technology companies are selling software they claim can analyze essays to detect such text. But accurate detection is very difficult, according to educational technology experts, forcing American educators into a pickle: adapt to the technology or make futile attempts to limit the ways it’s used.

The responses range the gamut. The New York City Department of Education has banned ChatGPT in its schools, as has the University of Sciences Po, in Paris, citing concerns it may foster rampant plagiarism and undermine learning. Other professors openly encourage use of chatbots, comparing them to educational tools like a calculator, and argue teachers should adapt curriculums to the software.

Yet educational experts say the tensions erupting at Texas A&M lay bare a troubling reality: protocols on how and when to use chatbots in classwork are vague and unenforceable, with any effort to regulate use risking false accusations.

“Do you want to go to war with your students over AI tools?” said Ian Linkletter, who serves as emerging technology and open-education librarian at the British Columbia Institute of Technology. “Or do you want to give them clear guidance on what is and isn’t okay, and teach them how to use the tools in an ethical manner?”

Michael Johnson, a spokesman for Texas A&M University at Commerce, said in a statement that no students failed Mumm’s class or were barred from graduating. He added that “several students have been exonerated and their grades have been issued, while one student has come forward admitting his use of [ChatGPT] in the course.”

He added that university officials are “developing policies to address the use or misuse of AI technology in the classroom.”

In response to concerns in the classroom, a fleet of companies have released products claiming they can flag AI generated text. Plagiarism detection company Turnitin unveiled an AI-writing detector in April to subscribers. A Post examination showed it can wrongly flag human generated text as written by AI. In January, ChatGPT-maker OpenAI said it created a tool that can distinguish between human and AI-generated text, but noted that it “is not fully reliable” and incorrectly labels such text 9 percent of the time.

Detecting AI generated text is hard. The software searches lines of text and looks for sentences that are “too consistently average,” Eric Wang, Turnitin’s vice president of AI, told The Post in April.

Educational technology experts said use of this software may harm students — particularly nonnative English speakers or basic writers, whose writing style may more closely match what an AI generated tool might generate. Chatbots are trained on troves of text, working like an advanced version of auto-complete to predict the next word in a sentence — a practice often resulting in writing that is by definition eerily average.

But as ChatGPT use spreads, it’s imperative that teachers begin to tackle the problem of false positives, said Linkletter.

He says AI detection will have a hard time keeping pace with the advances in large language models. For instance, Turnitin.com can flag AI text written by GPT-3.5, but not its successor model, GPT-4, he said. “Error detection is not a problem that can be solved,” Linkletter added. “It’s a challenge that will only grow increasingly more difficult.”

But he noted that even if detection software gets better at detecting AI generated text, it still causes mental and emotional strain when a student is wrongly accused. “False positives carry real harm,” he said. “At the scale of a course, or at the scale of the university, even a one or 2% rate of false positives will negatively impact dozens or hundreds of innocent students.”

At Texas A&M, there is still confusion. Mumm offered students to a chance to submit a new assignment by 5 p.m. Friday to not receive an incomplete for the class. “Several” students chose to do that, Johnson said, noting that their diplomas “are on hold until the assignment is complete.”

Bruce Schneier, a public interest technologist and lecturer at Harvard University’s Kennedy School of Government, said any attempts to crackdown on the use of AI chatbots in classrooms is misguided, and history proves that educators must adapt to technology. Schneier doesn’t discourage the use of ChatGPT in his own classrooms.

“There are lots of years when the pocket calculator was used for all math ever, and you walked into a classroom and you weren’t allowed to use it,” he said. “It took probably a generational switch for us to realize that’s unrealistic.”

Educators must grapple with the concept of “what does it mean to test knowledge.” In this new age, he said, it will be hard to get students to stop using AI to write first drafts of essays, and professors must tailor curriculums in favor of other assignments, such as projects or interactive work.

“Pedagogy is going to be different,” he said. “And fighting [AI], I think it’s a losing battle.”