TECHNOLOGY

The First Year of AI College Ends in Ruin

There's an arms race on campus, and professors are losing.

By Ian Bogost
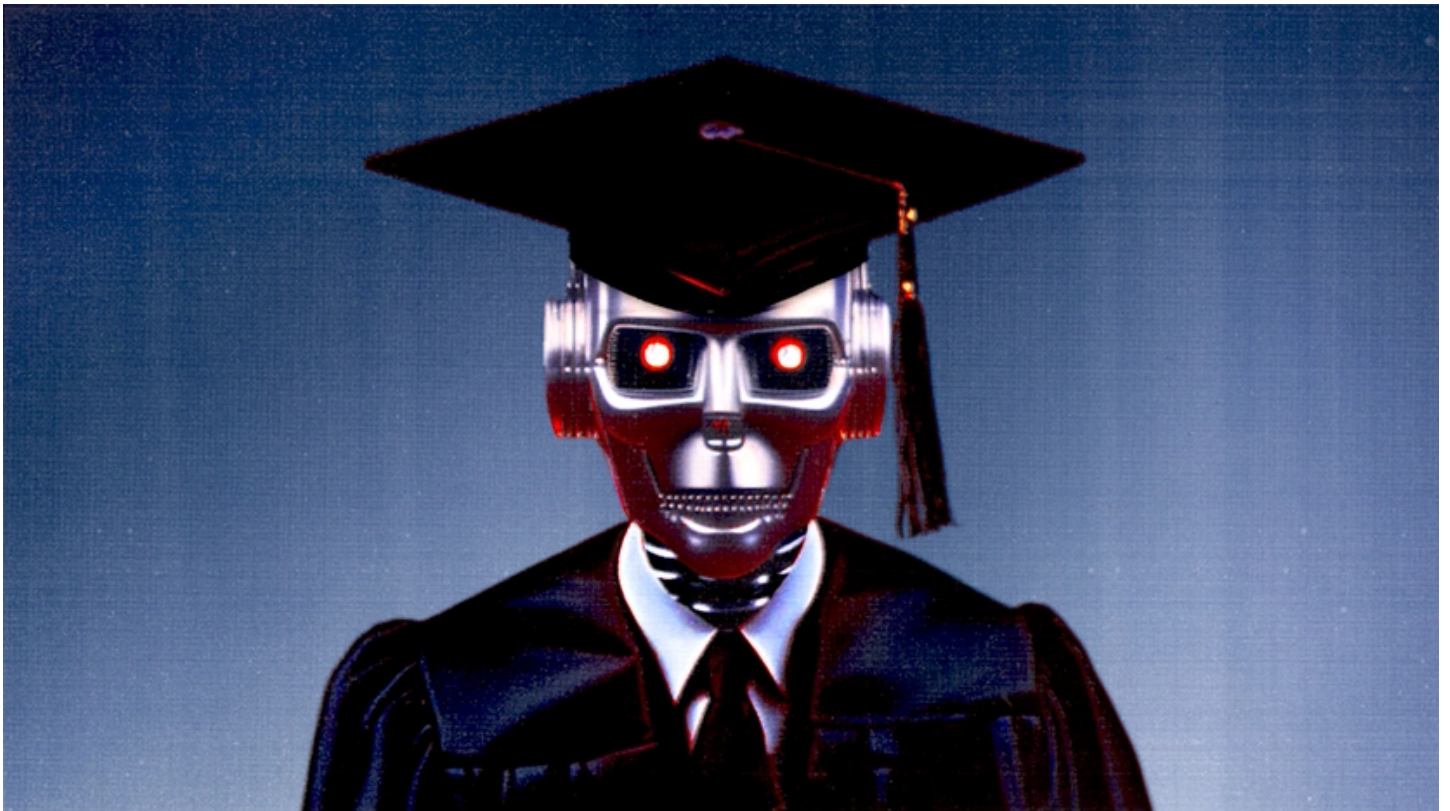

Illustration by Erik Carter / The Atlantic

MAY 16, 2023

SHARE ⌄   SAVED STORIES ↗   SAVE ▯

Listen to this article

⟲15 ▶ 15⟳   00:00   12:19

Listen to more stories on Hark

This article was featured in One Story to Read Today, a newsletter in which our editors recommend a single must-read from The Atlantic*, Monday through Friday. Sign up for it here.*

O NE HUNDRED PERCENT AI. That's what the software concluded about a student's paper. One of the professors in the academic program I direct had come across this finding and asked me what to do with it. Then another one saw the same result—*100 percent AI*—for a different paper by that student, and also wondered: *What does this mean?* I did not know. I still don't.

The problem breaks down into more problems: whether it's possible to know for certain that a student used AI, what it even means to "use" AI for writing papers, and when that use amounts to cheating. The software that had flagged our student's papers was also multilayered: Canvas, our courseware system, was running Turnitin, a popular plagiarism-detection service, which had recently installed a new AI-detection algorithm. The alleged evidence of cheating had emerged from a nesting doll of ed-tech black boxes.

This is college life at the close of ChatGPT's first academic year: a moil of incrimination and confusion. In the past few weeks, I've talked with dozens of educators and students who are now confronting, for the very first time, a spate of AI "cheating." Their stories left me reeling. Reports from on campus hint that legitimate uses of AI in education may be indistinguishable from unscrupulous ones, and that identifying cheaters—let alone holding them to account—is more or less impossible.

ONCE UPON A TIME, students shared exams or handed down papers to classmates. Then they started outsourcing their homework, aided by the internet. Online businesses such as EssayShark (which asserts that it sells term papers for "research and reference purposes only") have professionalized that process. Now it's possible for students to purchase answers for assignments from a "tutoring" service such as Chegg—a practice that the kids call "chegging." But when the AI chatbots were unleashed last fall, all these cheating methods of the past seemed obsolete. "We now believe [ChatGPT is] having an impact on our new-customer

growth rate," Chegg's CEO admitted on an earnings call this month. The company has since lost roughly $1 billion in market value.

Other companies could benefit from the same upheaval. By 2018, Turnitin was already taking more than $100 million in yearly revenue to help professors sniff out impropriety. Its software, embedded in the courseware that students use to turn in work, compares their submissions with a database of existing material (including other student papers that Turnitin has previously consumed), and flags material that might have been copied. The company, which has claimed to serve 15,000 educational institutions across the world, was acquired for $1.75 billion in 2019. Last month, it rolled out an AI-detection add-in (with no way for teachers to opt out). AI-chatbot countermeasures, like the chatbots themselves, are taking over.

RECOMMENDED READING

The Internet Is Rotting

JONATHAN ZITTRAIN
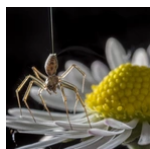
Spiders Can Fly Hundreds of Miles Using Electricity

ED YONG

Now, as the first chatbot spring comes to a close, Turnitin's new software is delivering a deluge of positive identifications: This paper was "18% AI"; that one, "100% AI." But what do any of those numbers really mean? Surprisingly—outrageously—it's very hard to say for sure. In each of the "100% AI" cases I heard about, students insisted that they had not let ChatGPT or any other AI tool do all of their work.

But according to the company, that designation does indeed suggest that 100 percent of an essay—as in, every one of its sentences—was computer generated, and, further, that this judgment has been made with 98 percent certainty. A Turnitin spokesperson acknowledged via email that "text created by another tool that uses algorithms or other computer-enabled systems," including grammar checkers and automated translators, could lead to a false positive, and that some "genuine" writing can be similar to AI-generated writing. "Some people simply write very predictably," she told me. Are all of these caveats accounted for in the company's claims of having 98 percent certainty in its analyses?

Perhaps it doesn't matter, because Turnitin disclaims drawing any conclusions about misconduct from its results. "This is only a number intended to help the educator determine if additional review or a discussion with the student is warranted," the spokesperson said. "Teaching is a human endeavor." The company has a guide for

humans who confront the software's "small" risk of generating <u>false positives</u>. Naturally, it recommends the use of still more Turnitin resources (an <u>AI-misuse rubric</u> and <u>AI-misuse checklist</u> are available) and doing <u>more work</u> than you ever would have done in the first place.

<u>Read: ChatGPT is about to dump more work on everyone</u>

In other words, the student in my program whose work was flagged for being "100% AI" might have used a little AI, or a lot of AI, or maybe something in between. As for any deeper questions—exactly *how* he used AI, and whether he was wrong to do so— teachers like me are, as ever, on our own.

SOME STUDENTS probably are using AI at 100 percent: to complete their work absent any effort of their own. But many use ChatGPT and other tools to generate ideas, help them when they're stuck, rephrase tricky paragraphs, or check their grammar.

Where one behavior turns into another isn't always clear. Matthew Boedy, an English professor at the University of North Georgia, told me about one student so disengaged, he sometimes attended class in his pajamas. When that student submitted an uncharacteristically adept essay this spring, Boedy figured a chatbot was involved, and <u>OpenAI's verification tool</u> confirmed as much. The student admitted that he

hadn't known how to begin, so he asked ChatGPT to write an introduction, and then to recommend sources. Absent a firm policy on AI cheating to lean on, Boedy talked through the material with the student in person and graded him based on that conversation.

A computer-science student at Washington University in St. Louis, where I teach, saw some irony in the sudden shift from giving fully open-book assignments earlier in the pandemic to this year's attitude of "you can use anything except AI." (I'm withholding the names of students so that they can be frank about their use of AI tools.) This student, who also works as a teaching assistant, knows firsthand that computers can help solve nearly every technical exercise that is assigned in CS courses, and some conceptual ones too. But taking advantage of the technology "feels less morally bankrupt," he said, "than paying for Chegg or something." A student who engages with a chatbot is doing some kind of work for themselves—and learning how to live in the future.

Another student I spoke with, who studies politics at Pomona College, uses AI as a way to pressure-test his ideas. Tasked with a research paper on colonialism in the Middle East, the student formulated a thesis and asked ChatGPT what it thought of the idea. "It told me it was bogus," he said. "I then proceeded to debate it—in doing so, ChatGPT brought up some serious counterarguments to my thesis that I went on to consider in my paper." The student also uses the bot to recommend sources. "I treat ChatGPT like a combination of a co-worker and an interested audience," he said.

The Pomona student's use of AI seems both clever and entirely aboveboard. But if he borrows a bit too much computer-generated language, Turnitin might still flag his work for being inauthentic. A professor can't really know whether students are using ChatGPT in nuanced ways or whether they've engaged in brazen cheating. No problem, you might say: Just develop a relationship of mutual trust with students and discuss the matter with them openly. A good idea at first blush, but AI risks splitting faculty and student interests. "AI is dangerous in that it's extremely tempting," Dennis Jerz, a professor at Seton Hill University, in Greensburg, Pennsylvania, told me. For students who are not invested in their classes, the results don't even have to be good— just good enough, and quick. "AI has made it much easier to churn out mediocre work."

Faculty already fret over getting students to see the long-term benefit of assignments. Their task is only getting harder. "It has been so completely demoralizing," an English teacher in Florida told me about AI cheating. "I have gone from *loving* my job in September of last year to deciding to completely leave it behind by April." (I am not printing this instructor's name or employer to protect him from job-related repercussions.) His assignments are typical of composition: thesis writing, bibliographies, outlines, and essays. But the teacher feels that AI has initiated an arms race of irrelevance between teachers and students. "With tools like ChatGPT, students think there's just no reason for them to care about developing those skills," he said. After students admitted to using ChatGPT to complete assignments in a previous term—for one student, *all* of the assignments—the teacher wondered why he was wasting his time grading automated work the students may not have even read. That feeling of pointlessness has infected his teaching process. "It's just about crushed me. I fell in love with teaching, and I have loved my time in the classroom, but with ChatGPT, everything feels pointless."

The loss that he describes is deeper and more existential than anything academic integrity can protect: a specific, if perhaps decaying, way of being among students and their teachers. "AI has already changed the classroom into something I no longer recognize," he told me. In this view, AI isn't a harbinger of the future but the last straw in a profession that was almost lost already, to funding collapse, gun violence, state overreach, economic decay, credentialism, and all the rest. New technology arrives on that grim shore, making schoolwork feel worthless, carried out to turn the crank of a machine rather than for teaching or learning.

What does this teacher plan to do after leaving education, I wonder, and then ask. But I should have known the answer, because what else is there: He's going to design software.

A COMMON LINE about education in the age of AI: It will <u>force</u> teachers to <u>adapt</u>. Athena Aktipis, a psychology professor at Arizona State University, has taken the opportunity to restructure her whole class, preferring discussions and student-defined projects to homework. "The students said that the class really made them feel human in a way that other classes didn't," she told me.

But for many students, college isn't just a place for writing papers, and cutting corners can provide a different way of feeling human. The student in my program whose papers raised Turnitin's "100% AI" flag told me that he'd run his text through grammar-checking software, and asked ChatGPT to improve certain lines. Efficiency seemed to matter more to him than quality. "Sometimes I want to play basketball. Sometimes I want to work out," he said when I asked if he wanted to share any impressions about AI for this story. That may sound outrageous: *College is for learning, and that means doing your assignments!* But a milkshake of stressors, costs, and other externalities has created a mental-health crisis on college campuses. AI, according to this student, is helping reduce that stress when little else has.

Similar pressures can apply to teachers too. Faculty are in some ways just as tempted as their students by the power of the chatbots, for easing work they find irritating or that distract from their professional goals. (As I pointed out last month, the traditional recommendation letter may be just as threatened by AI as the college essay.) Even so, faculty are worried the students are cheating themselves—and irritated that they've been caught in the middle. Julian Hanna, who teaches culture studies at Tilburg University, in the Netherlands, thinks the more sophisticated uses of AI will mostly benefit the students who were already set to succeed, putting disadvantaged students even further at risk. "I think the best students either don't need it or worry about being caught, or both." The others, he says, risk learning less than before. Another factor to consider: Students who speak English as a second language may be more reliant on grammar-checking software, or more inclined to have ChatGPT tune up their sentence-level phrasing. If that's the case, then they'll be singled out, disproportionately, as cheats.

One way or another, the arms race will continue. Students will be tempted to use AI too much, and universities will try to stop them. Professors can choose to accept some forms of AI-enabled work and outlaw others, but their choices will be shaped by the software that they're given. Technology itself will be more powerful than official policy or deep reflection.

Universities, too, will struggle to adapt. Most theories of academic integrity rely on crediting *people* for their work, not machines. That means old-fashioned honor codes will receive some modest updates, and the panels that investigate suspected cheaters will have to reckon with the mysteries of novel AI-detection "evidence." And then everything will change again. By the time each new system has been put in place, both technology and the customs for its use could well have shifted. ChatGPT has existed for only six months, remember.

Rethinking assignments in light of AI might be warranted, just like it was in light of online learning. But doing so will also be exhausting for both faculty and students. Nobody will be able to keep up, and yet everyone will have no choice but to do so. Somewhere in the cracks between all these tectonic shifts and their urgent responses, perhaps teachers will still find a way to teach, and students to learn.