CHAPTER 25.

COMMERCIAL GENERATIVE AI IS A GOOD IDEA FOR TEACHING WRITING! ✨ *SCHOOLS SHOULD INSIST THAT TEACHERS LEAD THE DEVELOPMENT OF AI EDUCATIONAL TECHNOLOGIES*

Marit J. MacArthur
University of California, Davis

The choice to adopt a particular commercial application in education should depend on whether it supports our purpose and our ethics. Technology can increase efficiency and broaden access to knowledge—or facilitate oppression and deepen the digital divide. Because corporate America needs to protect client privacy and corporate intellectual property, many companies are taking a slow, thoughtful approach to generative artificial intelligence (GenAI), building in-house tools (Lohr, 2023) within so-called walled gardens (meaning that the tools run on internal or otherwise highly secure cloud servers, and do not share any data with third parties). As consumers, educators need to follow this example, reject through GenAI marketing hype, assess commercial platforms in relation to our purpose of teaching and learning, and resist those that do not fit our needs. In this volume, Leah Heilig and Josh Chase (2025) helpfully suggest that instructors and students apply slow design principles to make reflective choices about writing with or without GenAI platforms, "to resist the culture of expediency surrounding GenAI adoption." Expediency also encourages educators to assume that commercial GenAI is the only game in town. This is rather ironic for two reasons: first, many colleges and universities have faculty and students in computer science, design and related fields who are capable of building basic open-source, non-commercial AI tools; and second, the advent of GenAI has made it easier to create open-source specialized bots and programs.

To put it simply, commercial GenAI platforms are a bad idea for teaching writing. They are overpriced. They may exploit student writing as training data and fail to protect student privacy. They may interfere with student learning

DOI: https://doi.org/1010.37514/PER-B.2026.2777.2.25

because they were not designed for education, where their impact was untested when OpenAI, Google, Anthropic and others forced a global AI pedagogy experiment, with significant "moral hazards" (Vee, 2024). Commercial platforms are not the only game in town. And yet they tempt school administrators, who often put too much trust in Big Tech (even as faculty across disciplines critique the industry), and who face frequent pressure to reduce instructional costs.

However, affordable GenAI applications, developed by educators for writing support, with strong data protections, *do* have a place in education. Playlab.ai and Hugging Face provide resources and opportunities for educators to collaborate and develop open-source (non-proprietary) tools. And educator-led platforms such as PapyrusAI and MyEssayFeedback, though they currently rely on ChatGPT APIs (through OpenAI and Microsoft, respectively), offer greater data and privacy protections than commercial platforms. Though such applications need iterative testing and development, they have real potential to improve equity in writing support, writing outcomes, workforce preparation and AI literacy—and to provide opportunities for students to get involved in applied, interdisciplinary research on GenAI and writing.

COSTS, TRAINING DATA, STUDENT LEARNING

AI companies clearly see big profit opportunities in education, inspired by schools' dependence on commercial learning management systems (e.g., Canvas, Google Classroom). For instance, the California State University—the largest public university system in the U.S.—signed a $16.9 million contract with OpenAI for ChatGPT Edu in February 2025, without consulting faculty and amid "a $2.3 billion budget gap" (Barajas, 2025). And just as there are "hidden costs" to GenAI, as Whitney Lew James points out in this volume, there are hidden profits for these companies when they build their chatbots into existing ed tech. The UC system, like many other colleges and universities across the United States, makes Microsoft Copilot (which relies on the GPT 4 series, through Microsoft's partnership with OpenAI) available to all students, and individual campuses are also making Gemini available. Costs for Copilot and Gemini, part of campus Microsoft Office and Google Workspace licenses, are unclear.

Student use of these increasingly ubiquitous platforms invites conflict, of course, with academic integrity policies. The policy at UC Davis targets "any work … generated by software or artificial intelligence" ("Code of Academic Conduct," 2025). When students prompt Microsoft Copilot, ChatGPT and ChatGPT Edu to complete assignments for them, they not only risk accusations of cheating, of course, but fail to learn. And these platforms are not designed to help students use them appropriately. OpenAI overtly markets ChatGPT to students for writing

support, offering advice that many students would struggle to apply (e.g., "Delegate citation grunt work," "Develop your ideas through Socratic dialogue" ["A student's guide to writing with ChatGPT," 2024]), in a sign that in their rush to profit, developers also display the "curse of expertise": experts often forget how they developed high-level skills, and thus struggle to teach novices (Hinds, 1999). Making these platforms available to students places a burden on overwhelmed instructors to teach students appropriate use of inappropriate tools. These platforms can also interfere with the development of students' writing skills, with student-teacher relationships—so crucial to motivation and success—and with teacher training, because future educators learn by responding to student writing and assessing individual students' needs in discipline-specific contexts.

While paying high fees to Big Tech for such tools, schools may also be giving away training data that has great monetary value: even when writing by students and faculty is ostensibly not harvested as training data, platform usage patterns are. This is another hidden profit of commercial platforms. Ed tech has a history of profiting from such data. The flawed plagiarism detector Turnitin.com was sold in 2019 for $1.75 billion to Advance Publications, after being trained on huge amounts of uncompensated student writing, without student consent (Newton, 2019). Initially, OpenAI attracted users with the seemingly free service of ChatGPT, and in exchange, naïve users often unwittingly shared their own writing and usage patterns, under the deceptive, default "Data controls" setting: "Improve the model for everyone" ("Data controls FAQ," 2024). While ChatGPT Edu claims to offer "[r]obust security, data privacy, and administrative controls," OpenAI is the subject of numerous lawsuits for its use of copyrighted material and continues to "cut corners to harvest [training] data" ("Introduction to ChatGPT Edu," 2024; Baker Hostetler, 2024; Metz et al., 2024;). Interestingly, after turning down a licensing deal with OpenAI in 2023 and suing the company for copyright infringement, *The New York Times* recently made a deal to share some of its editorial content with Amazon for AI training, for an undisclosed amount (Grynbaum & Metz, 2025). Like major media outlets, colleges and universities should respect and protect writing as intellectual property—writing by students and faculty both—neither giving it away nor selling it to the highest bidder. As James puts in this volume, "the creative work of generations should be priceless." And as Kathryn Conrad (2024) further argues in "A Blueprint for an AI Bill of Rights for Education," students "should be able to opt out of assignments that may put [their] own creative work at risk for data surveillance and use without compensation or that might put [their] privacy at risk."

In this light, it is unfair and unethical, and shortsighted in terms of workforce needs, for schools to pay big bucks for high-cost commercial GenAI. This is so precisely because of GenAI's vampiric reliance on "training data"—a euphemism,

in this context, for "an enormous amount of human cultural production" (Bryan, in this volume). GenAI platforms' efficacy dramatically degenerates when trained on their own output unless that output is edited and otherwise curated by humans (Shumailov et al., 2024). Access to a pipeline of college graduates who are strong writers, readers, and editors matters to AI companies, which signal the increased monetary value of human writing and their dependence on higher education in two ways: 1) they compensate media companies for new content, by expert professional writers and journalists, as training data, via expensive licensing agreements, and 2) they pay college graduates $20–$40/hour to write for and train GenAI and "annotate data," in other words comment on and correct AI output (David, 2024; Lu, 2024). The further development of AI deeply relies on the ability of human experts to *write*—in Big Tech's terms, to generate new training data—and *to read and critically assess and edit* AI-generated text (MacArthur, 2023, 2025), in order to collaborate effectively with GenAI in the workplace. Clearly, the economy needs schools to develop future workers with expertise in their fields and strong reading, writing and editing skills, including the ability to analyze the immediate rhetorical situation (audience, purpose, genre, and real-world, local context) as only humans can.

Prompt "engineering" is also a misleading term here, because it is not engineering—it's writing with awareness of the rhetorical situation. To "engineer" a prompt means this: to write instructions for GenAI, detailing the immediate rhetorical situation, which a chatbot cannot guess because it is not embedded in the physical, social world (MacArthur, 2025). Students must practice *a lot*, with some guidance and response from teachers, to learn to read and write effectively enough to collaborate with an AI ghostwriter, because writing is human communication and "experience cannot be automated" (Tenen, 2024). If ChatGPT-like tools interfere with the goals of writing instruction, what would an appropriate AI platform look like?

EDUCATOR-LED ALTERNATIVES TO COMMERCIAL GENAI TOOLS

To benefit from GenAI in developing their writing skills, students need AI tutoring platforms designed and overseen by writing teachers. Some well-intentioned proponents of AI tutoring platforms overlook the centrality of human instructors, fondly citing a well-known study (Bloom, 1984) showing that one-on-one and one-on-two tutoring, compared to "conventional" classroom teaching, can lift a student's course grade by two standard deviations (2 sigma) to the 98th percentile, exceeding performance gaps commonly attributed to income/race/ethnicity. The tutorial method was first documented at Oxford University in the

fifteenth century (Moore, 1968). As noted in a promotional blog for the nonprofit Khan Academy's Khanmigo, a tutoring platform based on OpenAI's GPT, "it has always been too costly to offer this level of attention to everyone" (Khan Academy, 2023). But they "believe that AI can bridge that gap," offering "personalized" services to "learners" at $4/month. While Khanmigo has lofty goals, it underestimates educators' crucial role. Effective human tutors and teachers develop meaningful relationships with individual students that build their confidence and their motivation to learn and collaborate.

Indeed, the fantasy that deploying such a GenAI tutor can cheaply realize the immense benefits of one-on-one tutoring or small-group teaching overlooks a key fact: the tutors in Bloom's study were human. They offered not only personalized feedback, but supportive human attention and rhetorical awareness. Tutored students in the study had more positive "attitudes [toward] and interests" in learning, compared to students in "conventional [classroom] instruction" and spent much more "time on task" (90+% vs. 65%), (Bloom, 1984, p. 4). No matter how personalized GenAI gets in its probabilistically determined advice, it cannot offer a human relationship, write a letter of reference based on that relationship, or apprehend an assignment's real-world audience and purpose without prompting. Because a student does not know what she does not know, she may struggle to imagine what questions to ask GenAI, in the process of trying to develop domain-specific knowledge, including writing skills. A teacher can write appropriate prompts for GenAI, to guide students with a particular assignment, and a teacher can judge the accuracy and relevance of GenAI output—including AI-generated writing advice—while a student may struggle to do both.

One early example of ChatGPT's interference with learning, and an educator-led solution that still deploys GenAI, came from UC Davis in the summer of 2023. Ty Feng was a graduate student TA for a programming languages course which, at first, allowed students to use ChatGPT for homework but not exams. When an unusual number of students performed poorly on exams, instructors faulted overreliance on ChatGPT. Feng, advised by Professor Hao-Chuan Wang, developed CourseAssistAI, which is trained on instructors' own course materials to create specialized tutoring chatbots, overseen by instructors. It uses ChatGPT's API but does not share data with OpenAI, and charges $5/student/month, with lower costs for a campus license.

Educators are also developing affordable, human-supervised, appropriate platforms for writing instruction, such as PapyrusAI and MyEssayFeedback (MEF). Full disclosure: I am co-leading a large California Education Learning Lab grant project applying and testing MEF. But rather than promoting these platforms in particular, I share their design and functionalities as examples of alternatives to commercial GenAI.

PapyrusAI was developed at the University of California Irvine's Digital Learning Lab, led by Professor of Education Mark Warschauer and tested in writing courses by a team of experts in educational research, writing instruction in STEM, and computer science. With funding from the National Science Foundation and UCI Beall Applied Innovation, Papyrus AI is built on ChatGPT's API, but student writing is not retained by OpenAI for training purposes. Students can use the platform's large prompt library, which instructors can modify or add to, to seek feedback during the writing process in an iterative conversation. Thus far, it is available on a limited basis within the UC system and at select colleges and K-12 schools in California, at a cost of about $1/user/month. Further expansion is planned.

MyEssayFeedback (MEF) is a publicly available, not-for-profit tool developed by Eric Kean, who taught college mathematics for 20 years and also developed LibreTexts' OER homework platform ADAPT, with extensive volunteer input from Anna Mills, a writing instructor at the College of Marin and member of the MLA-CCCC Joint Task Force on AI and Writing. It uses the ChatGPT API through Microsoft Azure, with strong privacy protections typical for business applications. MEF applies an OER library of instructor-authored prompts, which teachers can revise and adapt, to generate AI feedback on students' drafts. Used at 23 colleges and universities in 10 states and five countries as of 2024, MEF can be embedded in Canvas. Director of the University of Mississippi's AI Institute for Teachers Marc Watkins reflected that "faculty need to be in the driver's seat" and many students "appreciated [MEF]'s rapid response time." One student wrote: "the biggest takeaway for me … was that A.I. [*sic*] is not something to [fear] … [I] was not … heavily relying on its feedback to completely change my paper … I know how to ask questions of an A.I. tool that go deeper than just giving me answers, because they can help me lead myself to the answer of what better writing would look like" (MyEssayFeedback, 2024). Costs for MEF vary as usage grows, but for the above-mentioned grant project in California, Microsoft token credits will limit costs to $2–$3/student/course. Kean has also offered to license the tool to the entire California Community College, CSU and UC systems for a single fee (Kean, 2024). Though schools would cover processing of the essays (largely token costs), the potential cost savings, compared to commercial platforms, are enormous. Many more educator-led tools, some of them open-source, are being developed with resources through Playlab.ai and Hugging Face. There are far too many to list here.

CONCLUSION

Instead of diverting limited funds from writing instruction toward commercial GenAI, schools can insist that educators guide the development and integration of more affordable, appropriate tools, and support educators to do so. We must

network through professional associations, listservs and social media and collaborate to learn about non-commercial GenAI platforms for writing support, develop and test new ones, and advise school administrators to make the right choices. We should negotiate with Big Tech to lower API costs for instructor-led applications (through token credits, for instance), and, in collaboration with students and faculty in computer science, design and related disciplines, explore open-source options through platforms like HuggingFace and the nonprofit Playlab AI. Federal, state and private funding sources can support instructors, to give them the time and resources for this important work. Ideally, the outcomes will be increased AI literacy for writing instructors and students, improved equity in writing support and writing outcomes, and a future workforce well-prepared with the necessary reading, writing and editing skills to train and collaborate with GenAI.

REFERENCES

Baker Hostetler. (2024). *Case tracker: Artificial intelligence, copyrights and class actions*. https://tinyurl.com/yn27a5p8

Barajas, J. (2025, August 19). *Inside Cal State's big $17 million bet on ChatGPT for all. LAist.* https://laist.com/news/education/chatgpt-california-state-university-csu-ai-deal

Bloom, B. S. (1984) The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher 13*(6), 4-16.

Bryan, M. D. (2026). AI is completely unlike any other writing software: AI is strange and rhetorical just like other writing software. In C. Basgier, A. Mills, M. Olejnik, M. Rodak, & S. Sharma (Eds.), *Bad ideas about AI and writing: Generative practices for teaching, learning, and communication*. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/PER-B.2026.2777.2.05

CourseAssistAI. (2024). *Transform learning with your AI-powered course assistant.* https://courseassistai.com

Conrad, K. (2024). A blueprint for an AI bill of rights for education. *Critical AI, 2*(1). https://doi.org/10.1215/2834703X-11205245

David, E. (2024, January 4). OpenAI's news publisher deals reportedly top out at $5 million a year. *The Verge.* https://tinyurl.com/3canx55d

Grynbaum, M., & Metz, C. (2025, May 29). The Times and Amazon announce an A.I. licensing deal. *The New York Times*. https://tinyurl.com/yc265bx6

Heilig, L., & Chase, J. (2026). We must act fast to address AI in writing classrooms: Slow design can build more sustaining classroom practice with (or without) AI. In C. Basgier, A. Mills, M. Olejnik, M. Rodak, & S. Sharma (Eds.), *Bad ideas about AI and writing: Generative practices for teaching, learning, and communication*. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/PER-B.2026.2777.2.38

Hinds, P. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied, 5*(2) 205-221.

James, W. L. (2026). GenAI can be accessed for free: AI poses significant hidden costs to individuals, society, and the planet. In C. Basgier, A. Mills, M. Olejnik, M. Rodak, & S. Sharma (Eds.), *Bad ideas about AI and writing: Generative practices for teaching, learning, and communication*. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/PER-B.2026.2777.2.22

Kean, E. (2024, August 1). Personal Communication.

Khan Academy. (2023). Khan Academy's 7-step approach to prompt engineering. https://blog.khanacademy.org/khan-academys-7-step-approach-to-prompt-engineering-for-khanmigo/

Lohr, S. (2023, August 29). The A.I. revolution is coming. But not as fast as some people think. *The New York Times.* https://www.nytimes.com/2023/08/29/technology/ai-revolution-time.html

Lu, Y. (2024, April 4). Now hiring: sophisticated (but part-time) chatbot tutors. *The New York Times.* https://www.nytimes.com/2024/04/10/technology/ai-chatbot-training-chatgpt.html

MacArthur, M. (2023, September 23). AI, expertise and the convergence of writing and coding. *Inside Higher Ed.* https://www.insidehighered.com/opinion/views/2023/09/28/ai-and-convergence-writing-and-coding-opinion

MacArthur, M. (2025, June 25). Large language models and the problem of rhetorical debt. *AI & Society*. https://link.springer.com/article/10.1007/s00146-025-02403-w

Metz, C. (2024, April 8). How tech giants cut corners to harvest data for A.I. *The New York Times*. https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html

Moore, W. G. (1968). *The tutorial system and its future*. Pergamon Press.

MyEssayFeedback. (2024). *Fast, AI-powered formative feedback.* https://myessayfeedback.ai

Newton, D. (2019, March 11). A new weapon in the $1.7 billion war on academic cheating. *Forbes.* https://www.forbes.com/sites/dereknewton/2019/03/11/a-new-weapon-in-the-1-7-billion-war-on-academic-cheating/

OpenAI. (2024). Data controls FAQ. https://help.openai.com/en/articles/7730893-data-controls-faq

OpenAI. (2024). Introducing ChatGPT Edu. https://openai.com/index/introducing-chatgpt-edu/

OpenAI. (2024). A student's guide to writing with ChatGPT. https://openai.com/chatgpt/use-cases/student-writing-guide/

PapyrusAI. (2024). https://www.papyrusai.org

Shumailov, I., Shumaylov, Z., Zhao, Y. et al. (2024). AI models collapse when trained on recursively generated data. *Nature, 631*, 755-759. https://doi.org/10.1038/s41586-024-07566-y

Tenen, D. Y. (2024). *Literary theory for robots. How computers learned to write*. Norton.

University of California, Davis. (2025). *Code of academic conduct.* https://ossja.ucdavis.edu/code-academic-conduct

Vee, A. (2024). The moral hazards of technical debt in large language models: Why moving fast and breaking things is bad. *Critical AI, 2*(1). https://doi.org/10.1215/2834703X-11205182