

Gradient descent in training graduate students

Recently, I walked out into the bright summer sun from the darkness of the graduate studies building at my university in a daze of mild horror. I realized that soon three years will have gone by since chatGPT and that we have squandered this time as instructors, trainers and academic leaders. At least I have. I have failed to discuss it in any organized way with trainees in my lab. I failed to develop my own clear opinions and guidelines on it and so, I realized, I have failed the trainees in my lab for some years now.

What precipitated all this was chairing my first PhD defense where the thesis was written in part with generative AI. How did I know this? Because it was flagged by the external examiner's report of this thesis. And thus I found myself reading a paragraph in the thesis' preface that I would otherwise have missed, describing how this 'original work' was 'improved' in its clarity, grammar and structure by several proprietary generative AI models. As chair for a PhD defense at my university, I am tasked with making sure the external examiner's concerns are addressed and that my university's exam policies are followed. I found myself reading and re-reading the policies on generative AI in doctoral theses, especially the FAQ on the use of 'Gen AI tools to assist with the writing of [a] thesis or dissertation.'

Students could 'normally assume that the use of AI tools for editorial (i.e. grammar, flow) and transcription purposes are broadly acceptable' but that the use of generative AI 'to get started with drafting' needed approval from the supervisor and supervisory committee. Literal outputs could not be included in the thesis unless they were cited as output of generative AI, otherwise it would constitute plagiarism. And it would be plagiarism if you did it at any stage—not just the final draft, but also when sharing drafts with your supervisory committee.

Unquoted ChatGPT as plagiarism made perfect sense to me. Yet it was also something I realized that I had never seen mentioned before. As academics we're trained against plagiarism. We're afraid of students using someone else's words, but this time it's not someone else's voice, it's something else's. An amalgamation of all the voices—our own much included through our academic writing—combined with an algorithm searching for a way to match the training dataset and minimize loss. A colleague's comments from earlier the spring echoed in my head, 'it sounds like me.' I suspect it does—ask it a few questions in an area of plant biology this colleague has revolutionized and the line between a summary and a straight-up rip-off of unquoted text provided by chatGPT is so thin I am not sure it exists.

For three years students could have been plagiarizing from chatGPT for their theses, dissertations and other academic writing and we've been mute. I had never heard a colleague compare chatGPT to plagiarism, but it is, and I don't think any of us ever told our students. At least no one I know has. I know because I have been asking around and the replies have astounded me—comments that it would be outrageous to quote chatGPT ('you'd look like an idiot, you should always change it just enough so you don't have to use quotation marks') and, in response to a blogpost I wrote about this, complaints about how arbitrary rules about plagiarism seem and a general side-stepping of the issue that effectively suggested we really should not consider generative AI as plagiarism. There was also a repeated undertone of mild horror from some colleagues that I have not figured out this is the pathway to less work on teaching students to write. A colleague whispered to me

that it was ‘a relief to not have to edit my students’ writing’ as though she was explaining a new drug everyone but me was taking.

Is everyone on the new drug of work-free writing and editing via the magic of generative AI? Maybe, but I think they’re missing out on the fun of science some. As Andrew Gelman wrote, ‘it is by writing that I explore and organize my thoughts. Writing logical natural-language text is not so different from deductive mathematical reasoning: the point is to state my assumptions, work out their implications, and confront the places where these implications do not comport with reality or with my current understanding of the world.’ If we don’t train our students in this, what are we doing? My sense is that we have fretted over undergraduates using generative AI, while offering endless seminars and grant opportunities on how to ‘integrate AI’ in our training, teaching, and beyond. We never told the future educators and researchers we’re training why they might not want to use generative AI for their writing, what they will lose, and also that it’s academic misconduct. At least it is to me now.

After charing the defense and discussing with colleagues and trainees in my lab, I have been looking for ways to minimize loss for the training I offer. I have decided on guidelines for those in my lab in how they use generative AI for their writing. I like writing and I like teaching writing, no matter how painful it is sometimes. I want to focus the time I have—for now at least—in training others in writing natural language so I don’t want trainees in my lab using generative AI for their text.

I realize my choice can feel limiting for those who are not native speakers of English and I am trying to leave room to not further disadvantage them. I also think we should have a broader conversation of how we want generative AI to help level the English-language playing field in academics, instead of hiding behind the benefit of generative AI for non-native English speakers as a reason to suggest we must all embrace it. I personally would rather have non-native speakers write a finished first draft in their native language then use generative AI to translate it. They can review it, tweak the translation to make sure it has their meaning, and submit the full workflow of this to a journal, which then publishes the original and translated versions. This is a conversation I want to have.

A conversation I don’t want to have is one where a student tells me they don’t understand why they would have to tell their supervisory committee that all the chapter drafts they were reading included chatGPT text. The student looked at me like I was crazy—why would they need to disclose this? My university has decided it could be academic misconduct. But no one—until me at the student’s PhD defense—told the student. And that’s the gradient descent we all seem to have agreed to.