Generative Artificial Intelligence in Higher Education: Why the 'Banning Approach' to Student use is Sometimes Morally Justified

Downloaded from: https://research.chalmers.se, 2026-10-17 03:07 UTC

Citation for the original published paper (version of record):
de Fine Licht, K. (2024). Generative Artificial Intelligence in Higher Education: Why the 'Banning Approach' to Student
use is Sometimes Morally Justified. Philosophy & Technology, 37(3).
http://dx.doi.org/10.1007/s13347-024-00799-9

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Philosophy & Technology (2024) 37:113
https://doi.org/10.1007/s13347-024-00799-9

RESEARCH ARTICLE

Generative Artificial Intelligence in Higher Education: Why the 'Banning Approach' to Student use is Sometimes Morally Justified

Karl de Fine Licht[1]

Received: 18 April 2024 / Accepted: 19 August 2024
© The Author(s) 2024
Keywords Generative artificial intelligence · Higher education · Prohibition · Ethics

1 Introduction

There has been significant discussion among academics and policymakers about managing the use of generative artificial intelligence tools, such as ChatGPT, Gemini, and GitHub Copilot, in higher education, particularly regarding student usage (Eke, 2023, Malmström et al., 2023, Yeadon et al. 2023). Many universities have adopted a 'nuanced approach,' which encourages responsible use of these tools to achieve high-quality outcomes while adhering to ethical principles and regulations (McDonald et al., 2024). This middle-ground stance is recommended by many involved in the policy debate (Gimpel et al., 2023; Rudolph et al., 2023; Slimi & Carballido, 2023). The belief is that these tools are largely beneficial, that their negative effects can be managed, and that they will become ubiquitous, making resistance futile. Not all institutions align with this approach—some are more welcoming of these technologies, even mandating their use, while others ban or strongly discourage their use (McDonald et al., 2024).[1]

However, there remains a lack of systematic analysis of differing stances within a unified framework that accommodates both depth and breadth. Additionally, a structured argument supporting any particular position is conspicuously absent. While there is considerable understanding of the primary potential issues associated with generative AI in education, such as concerns related to integrity, bias, costs, the digital divide, and overreliance (Rudolph et al., 2023; Qadir, 2023, Williams 2024), the

[1] The regulatory framework for generative AI use by students varies widely, from prohibition to mandatory adoption. A study by McDonald et al. (2024) examined guidance from 116 US universities with high research activity (R1). The study found that 63% of these institutions encourage generative AI use, with 41% offering detailed classroom integration guidance. Over half provided sample syllabi (56%), and 50% shared sample curricula and activities. This positive stance highlights AI's perceived benefits for education. In contrast, 27% of the institutions explicitly discourage or limit AI use, a stance that will be defended here.

✉ Karl de Fine Licht
karl.definelicht@chalmers.se

[1] Chalmers University of Technology, Vera Sandbergs Allé 8, 411 33 Gothenburg, Sweden

Published online: 31 August 2024

potential benefits are also well-acknowledged, including enhancements in productivity, a deeper comprehension of AI technology and the subject matter, and increased inclusivity (Mollick & Mollick, 2022, 2023; Rudolph et al., 2023). Nevertheless, these insights are often presented generally, without directly advocating for or against any specific approach. Recommendations typically emerge in discussion sections as secondary considerations following descriptive content. When they do argue for a position, it is often about implementation under reasonable resource constraints (Cassinadri, 2024).

In this paper, I will argue that universities are justified under certain conditions in banning student use of generative AI. The conditions are that (a) the faculty, students, and administration collectively support the ban after taking part in a reasonable process, and (b) the university is not well-resourced. If the university lacks sufficient resources, this justifies not utilizing these tools due to potential issues with student privacy and learning outcomes. Even if a university is adequately resourced, there are strong moral considerations against using these tools in higher education. These concerns include the substantial energy consumption required to operate and train large language models, which poses environmental impacts, specific risks to student privacy, and exploration. The purported benefits of generative AI for student learning are, at best, uncertain and, at worst, minimal, and do not outweigh the potential adverse effects identified by stakeholders. Therefore, in contexts where faculty, students, and administration advocate for a university-wide prohibition of generative AI for student use, substantial reasons support such a stance, making it a permissible or even morally justified decision.

This study employs a bottom-up approach to philosophical inquiry (e.g., Wolff, 2019). It emphasizes the importance of practicality and feasibility, considering real-world constraints and stakeholder perspectives when proposing solutions. While grounded in specific cases, this approach uses a form of reflective equilibrium, moving back and forth between particular judgments and more general principles. These general principles can be philosophical in nature, such as consequentialist or deontological theories, or normative considerations that we often follow and consider to be justified in many other relevantly similar situations. This allows for a nuanced understanding that avoids the oversimplification of complex issues. Importantly, we maintain an openness to revision based on new evidence or arguments, recognizing the dynamic nature of real-world problems. This ensures that the paper contributes to both practical and theoretical discussions, rather than focusing exclusively on one or the other.

This paper is structured as follows: Sect. 2 lays out the goals of universities and the means of achieving these goals. Section 3 argues for limiting student access to generative AI. Section 4 counters by discussing the benefits and the inevitable integration of AI in future professional environments, while also highlighting the risks of premature adoption. The final section summarizes the main arguments, reflects on the broader ethical implications, and suggests future research and policy directions.

2 Values and Goals

The key goals of a university are to teach students both specific and general knowledge and skills. The specific knowledge and skills are related to their chosen fields and careers, while the general ones are relevant across different domains. For example, Stanford University emphasizes promoting public welfare by teaching "the blessings of liberty regulated by law, and inculcating love and reverence for the great principles of government as derived from the inalienable rights of humankind to life, liberty, and the pursuit of happiness." Similarly, Harvard University commits to educating citizens and leaders through "the transformative power of a liberal arts and sciences education." There is, of course, criticism aimed at the current system; for example, Schwartz (2020) argues that we should largely instill intellectual virtues in our students. However, overall, providing students with the knowledge and skills needed for their careers and lives is an uncontroversial goal according to most scholars and universities.

Universities also often aim for students to enjoy their learning process as much as possible. This is why universities invest significant efforts and resources into student health and wellness (see, e.g., Leshner & Scherer, 2021) and why so many researchers focus on it (see, e.g., Worsley et al. 2022). There are pragmatic reasons for this, such as attracting a large pool of applicants, which allows them to select high-quality students. Most moral theories also support promoting student well-being from the university's perspective, assuming this does not adversely affect other important values. For example, some degree programs, such as Physics, may be challenging and occasionally unpleasant, but this is viewed as an unavoidable side effect due to the positive value the knowledge in physics brings to consenting adults. Nevertheless, universities should and often do strive to ensure that students have positive experiences.

Many universities also state that they should not engage in morally unacceptable actions or participate in such processes. (For a list of examples of universities, see Appendix A.) In their mission statements, these universities emphasize the importance of contributing positively to the community, using terms like justice, fairness, equality, and sustainability. These universities aim to avoid actions that violate human rights and generally have policies aimed at minimizing their contribution to climate change. Institutions such as MIT and the University of Edinburgh are actively engaged in addressing climate change, with MIT developing practical solutions and leading by example in their operations, and the University of Edinburgh contributing to the global sustainability agenda through research, teaching, and operational practices. Additionally, Yale University is dedicated to "improving the world today and for future generations through outstanding research and scholarship, education, preservation, and practice." Universities like UC Berkeley and Cambridge also commit to improving the world through education and research, serving society by transmitting and discovering advanced knowledge. Cornell University, for instance, emphasizes values like equality and sustainability. These examples illustrate the broad commitment to these principles.

There are also good reasons for universities to adhere to principles that avoid performing morally unacceptable actions or participating in morally adverse processes. According to most normative theories, agents have at least a *pro tanto* reason to act in accordance with what is morally right and avoid what is morally wrong (see e.g., Kant, 1785; Ross, 1930). Thus, universities at least have a reason to avoid performing such actions, all else being equal, and this could, in turn, justify having a mission statement that includes this principle, as mission statements are general guides for university actions. Additionally, universities often need to take on the role of being responsible institutions because they are thought to shape and mold the future generation. This contributes to why Ruben (2023) and other educational scholars emphasize ethical leadership when discussing how to improve our institutions, departments, and programs, which is compatible with giving weight to doing the right thing from a moral point of view.

2.1 The Case for the Banning Approach

In this section, I will argue that universities, especially those that are not well-resourced and have decided to ban generative AI after a reasonable and inclusive process, have strong moral reasons to prohibit and discourage student use of generative AI. This is because a prohibition will likely yield the best outcome when it comes to student learning, skills, and well-being. A prohibition will also result in the university having less involvement in outcomes and processes that are morally objectionable.

Thus, we could set up the argument in the following way:

P1: Universities have strong moral reasons not to unnecessarily compromise student learning and well-being.
P2: Universities have strong moral reasons to avoid participating in or directly contributing to morally adverse processes or outcomes.
P3: Allowing student use of generative AI without restrictions leads to unnecessarily compromised student learning and well-being.
P4: Allowing student use of generative AI results in universities contributing to morally adverse processes and outcomes.
Conclusion: Therefore, universities have strong moral reasons not to allow student use of generative AI.

We will discuss the premises in order, beginning with student well-being. Studies have found that students are stressed because they do not know to what extent they can use generative AI (see, e.g., Malmström et al., 2023). If the university's stance is that students are prohibited from using these tools, then this worry should be relieved. Banning or stopping the use of generative AI will also save money for students, who usually have financial difficulties and may suffer from paying for premium AI tools. Having even less money than they currently do could plausibly have a negative impact on their well-being since this is a general effect for people on the margin. Although this argument might not be so weighty that it alone decides the

issue, it is still important, especially given that universities generally value student well-being.

There are also good reasons to believe that the knowledge and skills students acquire might be compromised if universities that are not well-resourced adopt a prohibitive and discouraging stance towards generative AI. To begin with, if teachers are not knowledgeable about generative AI, they will not be able to assist students with these technologies effectively (see, e.g., Hattie, 2008). At a well-resourced university, it may be possible to (i) develop a course for faculty on how to use these tools effectively in their research and teaching, (ii) provide faculty with time off to attend this course instead of engaging in their regular teaching or administrative duties, or (iii) allow faculty to use the time they already have for course development and similar tasks. In this scenario, both students and teachers would receive support on how to consider the use of these tools in various contexts, thus to some extent alleviating concerns about using the tools inappropriately or in ways that could hinder their learning.

Nevertheless, there are a wide range of problems with this line of reasoning. First, most universities do not have any courses in generative AI in research and higher education; hence, (i) is not fulfilled (Lievens, 2023, Miller, 2024). While there are many rules and regulations in place, and very optimistic writings on how to use the tools in scenarios such as a flipped classroom, actual training on using these tools is sparse. Consequently, for faculty to become proficient in using these tools, they frequently have to independently seek out educational resources, which is not always an easy task. Second, many universities today struggle with having too many responsibilities and too few resources to manage them all (Mitchell et al., 2017), which explains why (ii) and (iii) are seldom fulfilled (Lievens, 2023, Miller 2024). Faculty members are already overwhelmed with responsibilities related to teaching, supervising PhD students, participating in faculty committees, and communicating their research to other parts of society. Hence, they do not have any spare time to learn about the tools, which violates (iii). Neither are they given time off from their ordinary duties, thus violating (ii). This situation leads to many faculty members not learning about the tools because it is simply too daunting a task.

Against this, one could argue that faculty do not need extensive training in these tools since they are so easy to use. Even though tools such as ChatGPT and Claude are easy to understand on the surface, there is much more to using them than one might initially realize. This is why many people argue for the 'ten-hour rule' for learning each individual tool (see, e.g., Mollick, 2024). Nonetheless, according to many experts, ten hours is too little to fully grasp these tools. The ten-hour rule is more about getting a feel for how the tools work and beginning to understand their benefits and limitations. This rule is separate from teaching students prompt engineering (even if NLP researchers believe this will soon be a thing of the past) and integrating these tools effectively and responsibly into their work processes. Thus, much more than ten hours will be needed to learn enough about the tools to educate others on them.

Assuming that (i)—(iii) are fulfilled, it is still not clear that allowing the use of generative AI tools would be beneficial for student learning. Empirical studies suggest mixed outcomes when using classic AI to acquire knowledge and skills (see,

e.g., Sawyer 2014, Luckin 2018, and Zawacki-Richter et al. 2019). For example, AI tools have long been cautioned against due to their potential to lead to superficial engagement with learning materials, prioritizing efficiency over depth of understanding. This is particularly concerning in subjects requiring rigorous analytical skills and a deep grasp of underlying principles, which are what we often try to teach students at the university. Today's generative AI models are often good at summarizing customized versions of text for specific purposes. Many of them can process, clean, and perform necessary statistical analyses on lab data without explicit instructions, given just the data. These capabilities, and others, can be problematic since extensive use by students may lead to superficial engagement with the material, preventing them from learning what they are supposed to be learning.

The introduction of generative AI might also undermine the relational aspects important for developing knowledge and skills needed after graduation (see, e.g., Garrison & Vaughan, 2008). With the introduction of AI in communities such as coding, where Stack Overflow has been a popular online community for programmers to ask and answer questions about software development, we have seen a significant drop in queries from countries where generative AI tools are available. In contrast, activity remains the same in countries like Russia and China, where these tools are not yet accessible. Something similar might, of course, happen in the university context, where students might opt for working on their own with AI tools instead of collaborating with their student colleagues.

One could argue that we should teach students to use the tools well to bypass these and other problems. However, since the use of generative AI with the capabilities of tools such as ChatGPT 4, Claude 3, and beyond is relatively new, we lack evidence-based learning strategies for its proper use. We simply do not yet have enough data to determine the best ways to teach students to enhance their learning using these tools (see, e.g., Adiguzel et al., 2023; Malmström et al., 2023). This makes it difficult even for well-resourced universities to teach these tools in a high-quality manner. That said, it is not enough to ban the use of tools for universities to ensure that students meet the learning objectives in their curriculum. They will also have to adjust their examination practices to make cheating more difficult. This can be done in part by making examinations focus on processes instead of outcomes, and by having oral and on-site proctored exams instead of take-home exams, just to name a few common examples from the literature (see, e.g., Rudolph et al., 2023). If the students have been involved in making the decision to ban student use of generative AI, it is more likely that they will perceive the decision as legitimate and adhere to it (de Fine Licht 2023).[2]

In support of P4—that unrestricted use of generative AI results in universities participating in morally adverse processes and outcomes—student privacy emerges as a critical concern. According to a standard notion of privacy, this pertains to the

[2] Many of the arguments presented above can be interpreted as arguments of *justice*, hence *supporting P4*. This is true for maximizing theories such as Utilitarianism or Prioritarianism, egalitarian theories such as those of Dworkin or Anderson.

student's control over how their data is collected, used, and shared. The moral implications of this can be understood as ranging from an absolute right—that no one ever can, e.g., procure data about you without your consent is always morally wrong—to prima facie rights, where we have a right to privacy, but it can be trumped by other rights, or as having some final or instrumental value, where our reasons to respect someone's privacy come in degrees. Irrespective of how we understand the underlying moral notion of privacy, it is *prima facie* problematic when someone collects your data without your informed and explicit consent, especially when it is done with purposes other than your own well-being in mind.

There are many arguments in favor of believing that students will share their data with these tools without explicit informed consent. First of all, in order to share your data with informed consent, you need to actually understand what it might imply when sharing your data. This means you not only need to know that you are sharing your data, but you also need to know what the consequences might be of sharing your data. Today, many students and the general population have only a vague sense of what private and public entities can do with seemingly insignificant pieces of data (see, e.g., Veliz 2021). With powerful classificatory or predictive AI systems, even seemingly insignificant data chunks can lead to new ways for the AI system to predict your actions, steer you in a desired direction, or discover other things about you, such as hidden diseases you may not know about. Thus, even if students are aware they are sharing their data, they do not do so on an informed basis. The fact that the generative AI systems are also 'black boxes' with no insight into how they work or operate does not make it easier to achieve informed consent.

Second, it is reasonable to expect that companies will continue to collect data, given the nature of generative AI, which relies on extensive datasets to train its algorithms, and the market logic in which these tools exist. The frontline generative AI tools available today are brought about by commercial entities. This is true even for foundations like OpenAI, which operate on a market logic, trying to maximize their own advantage. The commercial interests of AI providers can sometimes conflict with the privacy interests of users when companies prioritize innovation and market expansion (Véliz 2021). To produce high-quality tools such today's frontline models, a large amount of high-quality data is needed. The data quality depends on a wide range of factors, but today it needs to be created by humans and not generated by LLMs. Preferably, it should be correct or at least not offensive in a way that extremism or pornography is to many people. Since today's tools have been trained on much of the high-quality data that exists on the internet, new sources are becoming increasingly scarce and in higher demand. This gives companies an extra incentive to use student-generated data, as it is often of high quality in all the above-mentioned senses.

Third, the data that students provide when utilizing generative AI tools will likely be data they otherwise would not share with system providers. This not only gives companies additional incentives to collect student data but also puts students at greater risk. The standard recommendations for how students should interact with generative AI models, such as ChatGPT, include using them as a study buddy, interlocutor, or coach for future job interviews (see, e.g., Holmes & Miao, 2023). This encourages students to share a lot of personal data that they otherwise would not

have shared—including their hopes and dreams, practical knowledge, strengths and weaknesses, and so on—creating a significant risk of oversharing. This data is varied in nature and can be used for different practical purposes, making it attractive for companies to collect. However, the more data procured from a person, the easier it becomes to categorize that person. This is something students might have good reasons to want to avoid.

Thus, the data input into these systems must be meticulously secured and managed to prevent misuse. Without strict guidelines and robust data protection measures, the handling of this data can lead to significant privacy violations, putting students at risk of unauthorized data sharing (see, e.g., Bhutoria, 2022). However, many universities today don't even provide licensed tools to students where data privacy is guaranteed. If students do not receive licenses, for instance, it is easy to foresee that they will use 'free versions' where they trade data for product use, as seen with GitHub Copilot, which offers students access to the premium version of the coding tool in exchange for submitting their student ID and sharing data with the company. This scenario not only compromises student privacy but also places the institution at risk of violating data protection laws and ethical standards, thus engaging in morally objectionable activities. Moreover, this burden falls predominantly on the worst-off students who wish to save money for other expenses, and since these groups should be prioritized according to most normative theories, we have even stronger reasons to avoid this situation.

Another, more general support of P4, irrespective of the resources available to the university, is that universities should care about the environment and that generative AI, such as LLMs, has a significant and negative environmental impact. Many universities have ambitious aims when it comes to their work on environmental sustainability (see, e.g., Appendix A and Sect. 3 above). Additionally, when it comes to questions about environmental sustainability, many philosophers and political theorists, among others, have argued that we have a duty to consider the long-term impacts of technology on the environment and future generations (see, e.g., Jonas, 1984, Broome 2012, Steel, 2015; for an overview, see Caney 2021), which further strengthens the view that universities should care about the climate and their climate impact.

The training and usage of these technologies also consume vast amounts of energy (see, e.g., Bender and Gebru 2021). This high energy demand arises from the complex computational processes required to train and operate these models, which often involve handling and analyzing massive datasets across distributed computing networks. For instance, the energy required to conduct a single search query using a large language model is substantially greater than performing the same search using a traditional search engine like Google. The environmental costs of using generative AI are not limited to direct energy consumption. The infrastructure needed to support these AI systems, including data centers and network systems, also requires significant amounts of energy for cooling and continuous operation. This exacerbates the environmental footprint of deploying these technologies in educational settings.

This increased energy consumption contributes to higher carbon emissions, which is especially concerning given the urgent global need to reduce greenhouse gas emissions to mitigate climate change. Universities that encourage the widespread use of

these AI tools without considering their environmental impact are inadvertently contributing to environmentally unsustainable practices. Therefore, from an environmental perspective, the unrestricted use of generative AI in higher education aligns universities with practices that may be considered morally objectionable due to their contribution to environmental degradation. By adopting a more restrictive approach to the use of generative AI, universities can mitigate these adverse environmental impacts, aligning their operations with broader ecological and sustainability goals. This, in turn, supports the claim that P4 is valid when considering the environmental implications of generative AI usage. This approach parallels other policies universities have implemented around the world concerning traveling, garbage disposal, and other sustainability measures.

Of course, one might argue that we could curtail the use of generative AI by instructing students to only use the tools for specific purposes and not for tasks like searching for information, which they are not particularly effective at anyway. This would allow for limited use instead of banning and discouraging use altogether. However, the problem with this approach is that it seems unlikely that students, once familiar with these tools and their ease of use, would restrict their usage to only sanctioned activities. Ethical use of technology highlights the difficulty in ensuring that convenient tools are used responsibly (Vallor, 2016). These tools, while convenient, might yield sufficiently good results in many scenarios, potentially in a more curated fashion than traditional methods like Google. This ease of use could make students more likely to rely on these tools rather than older, less environmentally taxing methods. Thus, from an environmental perspective, there is reason to be restrictive and consider banning rather than merely advising students to be cautious with the tools for environmental reasons.

Another common argument against allowing the use of AI tools in higher education, often presented in support of P4, is that tools like ChatGPT have been developed under exploitative conditions, citing the training phase in Nigeria where individuals had to interact with earlier, more offensive and hallucinatory versions of the model for minimal compensation[3] This issue aligns with the university's stance on justice and fairness (see Appendix A). There is, of course, much to say about what the alternatives were for the workers, to what extent they were forced to work under these conditions, and so on, but it is clear that paying them such a low salary, and not compensating them for the suffering they endured are, according to many normative theories—egalitarian, utilitarian, and otherwise—highly problematic, and there are good *prima facie* reasons to avoid it.

Of course, comparing these issues with the broader ethical landscape of technology production reveals that the issue of exploitation is not unique to the production of generative AI. For instance, the components used in our computers and smartphones often involve mining and manufacturing processes that are not only exploitative but also environmentally destructive. Furthermore, major corporations like Microsoft and Google, which are either directly involved in or supportive of

[3] https://www.theguardian.com/technology/2023/aug/02/ai-chatbot-training-human-toll-content-moderator-meta-openai.

generative AI development, are also key players in industries reliant on these problematic supply chains. Thus, if one were to avoid products based on their ethical production challenges, consistency would demand reassessing the use of a broad range of modern technologies. While it seems to me that the arguments related to privacy and environmental sustainability may be more distinctive to generative AI, the exploitation argument is perhaps less so. However, I will not explore this further here, nor the implications of what universities should do in light of other potential moral wrongs they are involved in. But it should be noted that there is no easy way to convincingly argue that just because e.g., it is hard to do the right thing in one area, it's acceptable to do wrong in another.

The last argument we will discuss here is the digital divide (see, e.g., Rudolph et al., 2023). Due to insufficient resources, universities may not be able to provide students with licenses for premium versions of the relevant tools. This could result in some students lacking access to these tools, thereby facing challenges in learning how to use them effectively, which would be an argument in favor of P3. However, this argument does not seem as convincing as those mentioned previously. Currently, the variety of genuinely useful tools at the university is limited, and if students receive guidance on which ones to use, the financial burden should not be significant. Typically, students are expected to pay out of pocket for textbooks, computers, and other essentials, so it is not unreasonable or unfair to expect them to also cover the cost of a few software licenses. The students likely won't need access to all tools at all times, so the additional financial burden may not be substantial. Of course, if empirical investigation proves otherwise, then we would have another argument in favor of the banning approach. So, even though we might have reasons related to student well-being that suggest favoring P3, this does not necessarily translate into the conclusion that students cannot pay for the licenses in a way that is relevant for justice or their actual learning.

2.2 The Case Against the Banning Approach

There are many arguments against restricting the use of generative AI tools for students at universities. One of the more prevalent ones concerns the belief that these tools will become all-encompassing in the future (see, e.g., Mollick & Mollick, 2022, 2023; Rudolph et al., 2023; Gimpel et al., 2023; Williams 2024). The basic idea is that students will need to learn how to use these tools because they will be required in their future jobs. This, in turn, leads to the notion that we need to teach them how to use these tools at the university. Consequently, it supports making the use of the tools obligatory and creating new learning objectives where students not only learn to use the tools but also how to integrate them into their workflow.

However, if these tools are expected to become significantly important in the future but are not as critical now, it seems that we do not need to make these concessions immediately but rather later. After all, many technologies have appeared promising but have not panned out as imagined. Self-driving cars, for instance, have been touted as a viable option for many years now but have only been introduced

in a very limited way into traffic.[4] Similarly, other technologies that were perceived to have great potential have either not been used at all, or their applications have diverged widely from what was initially thought (Bauer, 2014; Douthwaite et al., 2001). Therefore, it seems somewhat rash to make changes in our pedagogical set-ups now when we have good reasons not to, potentially developing skills that we do not yet know the students will need.

Another similar argument is that since we cannot prove that students have used generative AI tools—this being due to the absence of reliable classifiers for AI-generated text, code, calculations, etc.—*we should not prohibit* the use of these tools (Weber-Wulff et al. 2023; Perkins et al., 2024; Elkhatat et al. 2023). The argument contends that if a prohibition cannot be enforced in a legally secure manner, it should not be implemented. However, even if there is some merit to this, it may not be as persuasive as it initially appears. For one, there are many prohibitions that we cannot enforce. For example, students are not allowed to have their essays ghostwritten by others, yet this is undetectable if done well, or even if done less well since it is hard to determine that the essay was not written by the student, even if it is, for example, much better than expected. It could simply be that the student made an honest effort, which yielded an unexpectedly good result. Since we have not previously had a general prohibition against take-home assignments and thesis work that is mostly done un-proctored, our usual approaches to dealing with issues of prohibitions and enforcement are more relaxed than the position against the prohibition of AI tools suggests. Additionally, there are cases where it is possible to determine that a student has cheated with these tools. For instance, the references may be fake, or when attempting to locate them, they do not exist, or the text is written in such a way that it raises suspicion, or when questioned, the student confesses to using a tool. Thus, there are at least some instances where unlawful use can be detected and punished.

Again, this argument can be slightly modified to suggest that the issue is not so much our inability to detect students' use of generative AI tools, but rather that undetected student use will result in unequal quality and quantity of work. Specifically, those who use the tools despite regulations may produce more and better output, thus outperforming their peers who abide by the rules, potentially leading to worse grades and fewer job opportunities for the latter. Currently, however, the results are mixed when it comes to output quality and quantity in highly skilled work. For example, it has been found that programmers can code about 55% faster utilizing GitHub Copilot (Kalliamvakou, 2022), with those who are less skilled at coding benefiting more. However, it has also been found that the code produced has about 50% lower quality (Harding & Kloster, 2024), which means that it is not necessarily beneficial to use these tools when doing thesis work. Consequently, those not relying on AI might not have much to worry about in terms of being outperformed. There is also research showing that generative AI tools of today can indeed help even experts in some areas while not in others (Dell'Acqua et al., 2023). But here again, the tools

[4] This in very limited settings in Arizona and San Fransico.

vary in their output quality, making it less likely that everyone using these tools will benefit so much that the bar will be raised and leave all others behind.

Of course, in the future, if generative AI becomes ubiquitous and integrated into everything we do, banning it would be as impossible as banning computers, the internet, or the use of Google. Moreover, if students are using generative AI all the time, a prohibition would be unfair to them, as adhering to such rules would be extremely challenging. If these tools are constantly at your fingertips, the temptation to use them will be significant. However, there may eventually be tools similar to today's 'Freedom app' for the internet, which blocks access to distracting sites. Yet, it would be difficult for students, already under considerable pressure, to voluntarily restrict their access, and it would be unreasonable to demand that they do so. But we are not at that point yet, and we might never reach it due to the challenges of establishing a viable economic model for such technologies. Therefore, we are not currently in this situation.

It has also been argued that prohibiting the use of generative AI tools in higher education settings poses a significant risk of diminishing the learning potential and outcomes for students compared to scenarios where these technologies are embraced (see, e.g., Mollick & Mollick, 2022, 2023). Generative AI, by design, facilitates access to a vast array of information, augments the learning experience with interactive and personalized content, and fosters a deeper understanding of both the subject matter and the technology itself. By restricting access to such tools, educational institutions risk denying students exposure to innovative learning methodologies that prepare them for a future where AI plays a central role in many professions. Furthermore, as argued in the same literature, the use of generative AI in education encourages critical thinking and digital literacy, skills that are essential in navigating and evaluating the accuracy of information in the digital age. In an environment where these tools are banned, students may find themselves at a disadvantage, lacking essential competencies and the ability to creatively leverage technology to solve complex problems. Embracing generative AI responsibly, with appropriate ethical guidelines and oversight, could therefore enhance educational outcomes significantly, ensuring students are not only consumers of information but also skilled navigators and innovators in a technology-driven world.

In addition, another argument related to the former is the potential of generative AI tools to significantly boost student productivity (Dell'Acqua et al., 2023). These advanced technologies streamline the research process, automate mundane tasks, and provide quick access to information, allowing students to allocate more time and energy to critical thinking and complex problem-solving. Generative AI can offer personalized study aids, summarize extensive texts, and generate drafts or outlines, thus accelerating the learning cycle and enabling students to cover more material in less time. The efficiency afforded by these tools can transform the educational experience, making it more engaging and less burdensome, which in turn can lead to higher-quality work and a deeper understanding of the subject matter. By prohibiting the use of generative AI, institutions may inadvertently hinder students' ability to work efficiently and effectively, putting them at a disadvantage in an increasingly competitive and fast-paced academic and professional landscape. Therefore, embracing these technologies with appropriate ethical considerations could not only

enhance learning outcomes but also prepare students to thrive in a future where leveraging AI for productivity gains is the norm.

Even though I agree with this to some extent, these arguments promoting the unrestricted use of generative AI in educational settings often gloss over significant concerns that necessitate a more cautious approach. First, the assertion that generative AI enhances learning outcomes by providing interactive and personalized content may be overly optimistic. Such claims could underestimate the potential for AI to foster dependency, where students may rely excessively on technology for answers and insights, thereby eroding their ability to think independently and critically. Over-reliance on AI tools can lead to a degradation of essential academic skills, such as critical analysis, logical reasoning, and the capacity for deep reading and comprehension, as mentioned above. This diminishes, rather than enhances, educational quality by producing graduates who are adept at manipulating tools but lacking in foundational knowledge and intellectual rigor.

Moreover, the suggestion that generative AI can unambiguously boost student productivity overlooks the complexity of learning processes. While AI might streamline certain tasks, this efficiency could detract from the learning experience by encouraging a transactional approach to education, where the focus shifts from understanding content to completing tasks. By facilitating quicker completion of assignments and other academic tasks, AI might indeed save time, but at the cost of students' full engagement with the learning material. This superficial engagement can lead to a shallow understanding of complex concepts, preparing students inadequately for real-world challenges where deep knowledge and the ability to navigate complex problems are paramount (see, e.g., Selwyn, 2011, 2016; Carr, 2020). Thus, even though there might be benefits to be had, there is a great risk of these not materializing under conditions where universities are not well-resourced. Additionally, if students, faculty, and others have reached the decision to say no to the use of the tools, the known risks should weigh heavier than the imagined benefits.

The final argument, which I believe is the strongest, pertains to the prioritization of faculty time. Prohibiting the use of generative AI could *reduce the incentive* for teachers and examiners to familiarize themselves with these tools, making it harder for them to design courses that prevent cheating. However, despite this argument's validity, it is not decisive against the position I have argued for in this paper. As previously discussed, there are straightforward methods to mitigate cheating, and if students support the decision, cheating should decrease further. Therefore, while the argument has some merit, it ultimately fails.

3 Conclusion

This paper has critically examined the widespread adoption and potential ban of generative AI tools in higher education. While there are arguments for integrating these technologies based on their potential to enhance learning and productivity, these arguments often overlook or downplay significant risks associated with unrestricted use, particularly in contexts with limited resources and skeptical stakeholders. In these cases, the potential for dependency on AI, erosion of critical thinking

Table 1 Ethical Policies, Codes of Conduct, and Mission Statements at Selected Universities. This table provides a list of universities along with direct links to their respective ethical policies, codes of conduct, and mission statements, emphasizing their commitment to responsible and ethical practices in academia

No	University	Link
1	Cornell University	Our Values & Impact
2	Delft University of Technology	Code of Conduct
3	ETH Zurich	The ETH Zurich
4	Harvard T.H. Chan School of Public Health	Standards of Conduct in the Harvard Community
5	Imperial College London	Mission
6	KTH Royal Institute of Technology	Ethical Policy
7	London School of Economics and Political Science (LSE)	Ethics Code
8	Massachusetts Institute of Technology (MIT)	Values
9	Penn Medicine	Core Values and Commitment to Diversity, Equity, and Inclusion
10	Politecnico di Milano	Our Values
11	RWTH Aachen University	Werte-Kodex
12	Stanford University	Our Vision
13	Technische Universität Berlin	University Mission
14	University of Cambridge	Mission and Core Values
15	University of Melbourne	Educational Values and Vision
16	University of Pennsylvania	Sustainability
17	University of Tokyo	Mission
18	University of Toronto	Mission, Vision & Values

skills, and perpetuation of educational inequalities present substantial challenges. Consequently, this paper has argued for a 'banning approach' in such educational settings, aligning with broader ethical considerations and the responsibility of educational institutions to safeguard both student learning and well-being.

Looking forward, it is essential for future research to continue exploring these themes, providing robust data, and developing policies that consider both the technological advancements and the ethical implications of AI in education. On the practical side, we also need to better equip our teachers and examiners so that they can teach these tools effectively when (or if) they become so prevalent that banning them is no longer justifiable. There are quite a few excellent educational resources available on how to use these tools. The problem lies in finding the time and securing funding for licenses so that faculty can begin to learn these tools in earnest. In the meantime, we also need to investigate more thoroughly how we should handle supportive stakeholders, such as students, in relation to the broader and narrower arguments that have been suggested here. It's not clear that we should permit the use of generative AI even if some of the stakeholders are favorable, but this needs to be scrutinized more closely since this is probably the case at many locations.

Currently, faculty around the world are trying to keep up on their own, often spending their own money on licenses for a wide range of generative AI tools and spending their spare time trying to figure out how to use them. This is, of course, one way to do it, but it is a slow and cumbersome method, which makes it even

more important to restrict student use, if possible. The decisions we make today will shape the ethical landscape of technology usage in educational settings, emphasizing the need for a thoughtful approach that balances innovation with responsibility.

4 Appendix

Table 1

Acknowledgements None

Author's Contributions I'm the sole author and I've done everything myself

Funding Open access funding provided by Chalmers University of Technology.

Data Availability Not applicable.

Declarations Ethics Approval and Consent to Participate.

Not applicable.

Consent for Publication Not applicable.

Competing Interests None.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Bauer, R. (2014). Failed Innovations—Five Decades of Failure? *Icon, 20*(1), 33–40.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610–623).

Bhutoria, A. (2022). Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a human-in-the-loop model. *Computers and Education: Artificial Intelligence, 3*, 100068.

Broome, J. (2012). *Climate matters: Ethics in a warming world (Norton global ethics series)*. WW Norton & Company.

Caney, S. (2021). Climate justice. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition). Retrieved from https://plato.stanford.edu/archives/win2021/entries/justice-climate/

Carr, N. (2020). *The shallows: What the Internet is doing to our brains*. WW Norton & Company.

Cassinadri, G. (2024). ChatGPT and the Technology-Education Tension: Applying Contextual Virtue Epistemology to a Cognitive Artifact. *Philosophy & Technology, 37*(1), 1–28.

de Fine Licht, K (2023). Integrating Large Language Models into Higher Education: Guidelines for Effective Implementation. *Computer Sciences & Mathematics Forum, 8*(1), 65. https://doi.org/10.3390/cmsf2023008065

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (24–013).

Douthwaite, B., Keatinge, J. D. H., & Park, J. R. (2001). Why promising technologies fail: The neglected role of user innovation during adoption. *Research Policy, 30*(5), 819–836.

Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology, 13*, 100060. https://doi.org/10.1016/j.jrt.2023.100060

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity, 19*(1), 17.

Garrison, R. D., & Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines*. Jossey-Bass.

Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mädche, A., & Vandrik, S. (2023). *Unlocking the power of generative AI models and systems such as GPT-4 and ChatGPT for higher education: A guide for students and lecturers* (No. 02–2023). Hohenheim Discussion Papers in Business, Economics and Social Sciences. https://doi.org/10.13140/RG.2.2.20710.09287/2

Harding, W., Kloster, M. (2024) Coding on Copilot -2023 Shows downward pressure on code quality. GitClear White Paper. Available here: https://www.gitclear.com/coding_on_copilot_data_shows_ais_downward_pressure_on_code_quality

Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.Véliz, C. (2021). *Privacy is power*. Brooklyn: Melville House.

Humby, C. (2006). "Data is the new oil." Presentation at the Association of National Advertisers' Marketing Maestros conference in 2006.

Jonas, H. (1984). *The imperative of responsibility: In search of an ethics for the technological age*. University of Chicago press.

Kalliamvakou, E. (2022) Research: quantifying GitHub Copilot's impact on developer productivity and happiness. Available here: https://github.blog/2022-09-07-research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/

Kant, I. (1785). *Groundwork of the Metaphysics of Morals*. Cambridge University Press.

Leshner, A. I., Sherer, L. A. (2021) Mental Health, Substance Use, and Wellbeing in Higher Education: Supporting the Whole Student. Consensus study report. Available here: https://www.ncbi.nlm.nih.gov/books/NBK566721/pdf/Bookshelf_NBK566721.pdf

Lievens, J. (2023). Artificial intelligence (ai) in higher education: tool or trickery? *Education and New Developments 2023 – Volume 2*.

Luckin, R. (2018). Machine Learning and Human Intelligence: The future of education for the 21st century. UCL IOE Press. UCL Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL.

Malmström, H., Stöhr, C., & Ou, A. W. (2023). Chatbots and other AI for learning: A survey of use and views among university students in Sweden. *Chalmers Studies in Communication and Learning in Higher Education*, *1*(10.17196) https://doi.org/10.17196/cls.csclhe/2023/01

McDonald, N., Johri, A., Ali, A., & Hingle, A. (2024). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *arXiv preprint arXiv:2402.01659*

Miller, R. E. (2024). Pandora's Can of Worms: A Year of Generative AI in Higher Education. *portal: Libraries and the Academy, 24*(1), 21–34.

Mitchell, M., Leachman, M., & Masterson, K. (2017). A lost decade in higher education funding state cuts have driven up tuition and reduced quality.

Mollick, E. R. (2024). *Co-Intelligence*. UK: Random House.

Mollick, E.R.; Mollick, L. New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments, 2022. SSRN. Available online: https://ssrn.com/abstract=4300783 (accessed on 30 July 2023).

Mollick, E.R.; Mollick, L. Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts, 2023. SSRN. Available online: https://ssrn.com/abstract=4391243 (accessed on 30 July 2023).

Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J., & Khuat, H. Q. (2024). GenAI Detection Tools, Adversarial Techniques and Implications for Inclusivity in Higher Education. *arXiv preprint arXiv:2403.19148.*

Qadir, J. (2023). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1–9). IEEE.

Ross, W. D. (1930). *The Right and the Good*. Oxford University Press.

Ruben, B. D. (2023). *Excellence in Higher Education Guide: A Framework for the Design, Assessment, and Continuing Improvement of Institutions, Departments, and Programs*. eBook: Routledge. https://doi.org/10.4324/9781003444664

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342–363.

Sawyer, R. K. (2014). The future of learning: Grounding educational innovation in the learning sciences. *The Cambridge handbook of the learning sciences*, 726–746.

Schwartz, B. (2020). Science, Scholarship, and intellectual value: A guide to what higher education should be like. *Journal of Moral Education 51*:1. https://doi.org/10.1080/03057240.2020.1772211

Selwyn, N. (2011). *Education and Technology: Key Issues and Debates*. Bloomsbury Publishing.

Selwyn, N. (2016). *Is Technology Good for Education?* Polity Press.

Slimi, Z., & Carballido, B. V. (2023). Navigating the Ethical Challenges of Artificial Intelligence in Higher Education: An Analysis of Seven Global AI Ethics Policies. *TEM Journal*, *12*(2). Véliz, C. (2021). *Privacy is power*. Brooklyn: Melville House.

Steel, D. (2015). *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy*. Cambridge University Press.

Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.

Véliz, C. (2021). Privacy is power: Why and how you should take back control of your data. Melville House.

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*(1), 26.

Williams, R. T. (2024, January). The ethical implications of using generative chatbots in higher education. In *Frontiers in Education* (Vol. 8, p. 1331607). Frontiers Media SA.

Wolff, J. (2019). *Ethics and public policy: A philosophical inquiry*. Routledge.

Worsley, J. D., Pennington, A., & Corcoran, R. (2022). Supporting mental health and wellbeing of university and college students: A systematic review of review-level evidence of interventions. *PLoS one, 17*(7), e0266725.

Yeadon, W., Inyang, O. O., Mizouri, A., Peach, A., & Testrow, C. P. (2023). The death of the short-form physics essay in the coming AI revolution. *Physics Education, 58*(3), 035027.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.