Higher Education’s Generative Artificial Intelligence Paradox: The Meaning of Chatbot Mania

Dr Jürgen Rudolph, Dr Mohamed Fadhil Bin Mohamed Ismail, Dr Stefan Popenici

Abstract

Higher education is currently under a significant transformation due to the emergence of generative artificial intelligence (GenAI) technologies, the hype surrounding GenAI and the increasing influence of educational technology business groups over tertiary education. This commentary, prepared for the Special Issue of the Journal of University Teaching & Learning Practice (JUTLP) on “Enhancing student engagement using Artificial Intelligence (AI) and chatbots,” delves into the complex landscape of opportunities and threats that AI chatbots, including ChatGPT, introduce to the realm of higher education. We argue that while GenAI offers promise in enhancing pedagogy, research, administration, and student support, concerns around academic integrity, labour displacement, embedded biases, environmental sustainability, increased commercialisation, and regulatory gaps necessitate a critical approach. Our commentary advocates for the development of critical AI literacy among educators and students, emphasising the necessity to foster an environment of responsible innovation and informed use of AI. We posit that the successful integration of AI in higher education must be grounded in the principles of ethics, equity, and the prioritisation of educational aims and human values. By offering a critical and nuanced exploration of these issues, our commentary aims to contribute to the ongoing discourse on how higher education institutions can navigate the rise of GenAI, ensuring that technological advancements benefit all stakeholders while upholding core academic values.

Citation

Introduction

The development of full artificial intelligence could spell the end of the human race... It would take off on its own and re-design itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn’t compete and would be superseded. (Hawking, 2014)

Stephen Hawking’s foreboding quote is a stark reminder of how the current AI revolution continues to expose the vulnerability of humankind. Despite this apocalyptic warning, it is also important to realise that AI holds promise for higher education institutions worldwide, from improving the teaching and learning experience to facilitating administrative tasks (Humble & Mozelius, 2022). However, this promise is enveloped with a myriad of challenges and dangers. Consequently, it is imperative to scrutinise GenAI’s opportunities and threats critically, including a serious exploration of the assumption that “full artificial intelligence” can be achieved.

The AI proponents’ narrative posits that recent advancements in machine learning, natural language processing, and data analytics have propelled AI to the forefront of global educational innovation (Doroudi, 2023; Luan et al., 2020). The current dominant narratives reflect that these advancements stand positioned to offer unprecedented opportunities for personalised learning experiences, innovative ways of assessing learners, streamlining administrative processes, leaner operational workflows, and data-driven decision-making (Alam & Mohanty, 2023). Supposedly, one of the most significant opportunities presented by AI in higher education is its potential to completely transform the learning experience. Adaptive learning platforms powered by AI algorithms can analyse student data in real time, adapting educational content to individual learning preferences and students’ pace in learning (Kabudi et al., 2021). Moreover, the AI proponents’ argument goes that AI-driven educational tools such as chatbots and virtual assistants can augment the role of educators, providing real-time feedback, generating insights from vast datasets, and automating routine tasks such as grading and administrative paperwork (Vashishth, 2024).

It is of critical importance to highlight that ‘artificial intelligence’ is a highly problematic and loaded concept despite its seemingly magical advantages. The term is a marketing label vulnerable to exploitation and exaggeration (Popenici, 2023a). Lindgren (2023a) notes that “AI is driven by myths that animate it as magic” (p. 94), while Crawford (2021) argues that “AI is neither artificial nor intelligent” (p. 8). Lindgren (2023b, p. 17) further posits that AI, being “the subject of evolving wars of definitions”, is an “empty signifier”, a “ubiquitous apparatus... entangled with human experience”, and “part of the technological unconscious”. AI is an ideology-driven, socio-political mythology, “constantly repeated and performed in marketing talk, hyped-up conferences, tech evangelism, business manifestos, and overblown media reporting” (Lindgren, 2023b, p. 17).

When it was created in the 1950s, AI was an overhyped phenomenon that was underwhelming in its delivery (Woolridge, 2020). In the 2010s, with voice assistance and self-driving cars, robotics, and automated healthcare, AI again rose to become the decade’s buzzword (Metz, 2022). For the general public, the magical concept of AI is experienced through dystopian manifestations in the form of Hollywood blockbusters such as The Terminator or The Matrix.
AI, therefore, remains a heady mix of real technological advances, unfounded hype, wild predictions, and legitimate concerns for the future (Rudolph et al., 2023b).

More recently, GenAI has fired the public's imagination. GenAI chatbots can be defined as conversational agents that harness deep learning models to produce human-like content in terms of text, images, and video in response to prompts (Michel-Villarreal et al., 2023; Hart, 2024). Caldarini et al. (2022) defined chatbots as intelligent computerised systems capable of mimicking human conversation in its natural form and processing natural language to engage in human conversations. In other words, it can be seen as an automated dialogue system that can possibly replace humans in a dialogue and potentially address the needs of many users all at once (Zhai & Wibowo, 2023). Thus, AI chatbots can potentially assist as virtual teaching assistants, automated assessors, research enablers, and administrative assistants (Labadze et al., 2023).

One of the foremost concerns is the ethical implications of AI algorithms, particularly in admissions, financial aid provision, and academic evaluations (Tsamados et al., 2021), also due to a well-documented record of AI bias and problematic algorithmic decision-making. Moreover, the rapid proliferation of AI in education raises concerns about data privacy, security, and ownership (Berendt et al., 2020). As educational institutions amass vast quantities of student data for AI-driven analytics and personalisation, questions arise regarding data governance, consent, and the potential for misuse, commercial exploitations that stand against students' interests or unauthorised access. Safeguarding sensitive information and upholding principles of privacy and confidentiality are paramount in the ethical deployment of AI technologies in higher education.

Figure 1.
Yin and yang: Threats and opportunities of GenAI for higher education.
With the current hype, it is more important than ever before to assess if we are at a historic, revolutionary moment in AI development that is beneficial for humanity or if we are heading towards a more apocalyptic scenario, à la Stephen Hawking. Between Chomsky et al.’s (2023) evaluation of GenAI as “high-tech plagiarism” and a “way of avoiding learning” and Bill Gates’s view of it being as important as the invention of the computer or the Internet (The Economist, 2023c) we may find many areas of extreme relevance for higher education. Similarly, higher education reactions to chatbots have been on a continuum between feeble attempts to ban them and proactively including them in the curricula (Rudolph et al., 2023b).

Confronted with the current chatbot mania within higher education, it is critical to explore the different facets of chatbots and their potential roles in higher education. Before concluding with implications for higher education, this commentary will examine ten opportunities and threats, each associated with GenAI.

### Opportunities

For the past decade, chatbots have been increasingly integrated into various facets of higher education, including assessments, admissions, consultations, and campus life (Luckin et al., 2016; Rudolph et al., 2023b). Using AI-driven virtual assistants may improve efficiency and create leaner workflows. By automating routine tasks such as course registration and student inquiries, staff can attend to more complex and pressing issues. Chatbots can potentially enhance student engagement and retention by providing personalised recommendations, proactive reminders, and timely interventions (Tian et al., 2021). GenAI features impressive qualities that allow it to create original and novel content, such as text, images, videos, and music (Anantrasirichai & Bull, 2022).

Although there are studies that have explored the advantages, disadvantages, opportunities, and threats of using ChatGPT in higher education (e.g. Firat, 2023; Gamage et al., 2023; Ifelebuegu et al., 2023; Limna et al., 2023; Rasul et al., 2023; Author, 2023a, 2023b; Van Wyk, 2024), it is necessary to delve deeper and explore the opportunities that GenAI has. This section briefly discusses ten opportunities that AI offers to higher education. These opportunities are broad and are presented below in ten key areas across the higher education landscape.

#### Teaching Assistant

A recent noteworthy finding is that GPT-4 scored exceptionally well on various creativity tests, exceeding 91% of humans on an Alternative Uses Test for creativity and 99% on the Torrance Tests of Creative Thinking (Haase & Hanel, 2023; Shimek, 2023). With such striking capabilities, higher education can likely benefit from AI chatbots. One notable advantage of AI chatbots is that they can be deployed as teaching assistants (Mollick & Mollick, 2023). Tools such as ChatPDF and Claude 3 can be beneficial for reviewing and summarising lengthy articles and texts (Rudolph et al., 2023d). For example, Claude 3 users can input 200,000 tokens (equivalent to 150,000 words or hundreds of pages) in a single prompt (Anthropic, 2024). Ideally, teachers would like to read through all the necessary texts and resources to prepare their lessons from scratch, but this may not always be possible with a hectic teaching schedule. Chatbots may assist teachers with such tasks within minutes, freeing them up to
attend to students and have more time for other critical tasks such as curriculum planning, assessment creation and grading.

A common claim is that educators in higher education can use GenAI to brainstorm and adapt materials for their lessons. Depending on the intent, prompts in GenAI can be refined and repeated, thus churning out multiple ideas, resources, and materials within minutes (Kohnke et al., 2023; Eager & Brunton, 2023). Integrating critical cognitive functions such as brainstorming, summarising, and drafting will likely lead to significant productivity gains (Mollick & Euchner, 2023) not only in administrative and research-related work but also in teaching-related processes such as the creation of learning activities, lesson plans and learning resources, and assessments (Al-Naqbi et al., 2024). Educators can consider using GenAI intentionally as a round-the-clock virtual tutor, which helps students practice their skills, for instance, when learning a new language (Ifelebuegu et al., 2023). However, educators must exercise this purpose with care and be aware that Gen AI might be useful not as their replacement but as their assistant.

GenAI can be used by educators to craft personalised learning plans, create effective discussions and interactions, and deliver appropriate and practical feedback (Liu et al., 2023). It has the potential to improve the manner in which pedagogical approaches are used in class according to the appropriate context. It suggests that introducing GenAI can shift the focus from transferring knowledge to processing knowledge and from disciplinary learning to interdisciplinary learning (Chiu, 2024). One other benefit of GenAI as a teaching assistant is that it can draft teaching scripts for new teachers. This is helpful in creating coherent texts for teachers to use when giving explanations and providing instructions in class. These texts and explanations can be improved with a good prompt (e.g., using a ‘persona prompt’ and telling GPT-4 that it is a ‘Professor of Higher Education’ when asking it to draft text). Crawford et al. (2023c) posit that GenAI’s very inherent limitations (leading to subpar quality, inaccuracies, biases, and a lack of originality) present fertile ground for ‘teachable moments’, thereby providing an avenue to enhance student competencies in navigating an AI-influenced landscape. Lodge et al. (2023) advocate for a comprehensive strategy that prioritises self-regulated learning and co-regulation of learning, emphasising the critical importance of autonomy, adaptability, and profound comprehension in mastering AI-enriched educational landscapes while highlighting the role of individual agency within a synergistic human-AI learning network to optimally prepare learners for an AI-integrated future.

Nonetheless, optimistic views on the potential benefits of using AI in education fall short of addressing the implications of inherent algorithmic bias within AI outputs and how this feature impacts educators by providing materials shaped by a concerning set of values. The history of technology reveals why these issues currently impact the development and applications of AI, and a deeper analysis helps us understand how white supremacy and other dangerous ideas about race, gender, and ability are embedded in today’s technology (Broussard, 2023).

**Feedback and Assessments**

It is also proposed that GenAI is helpful in facilitating the creation of assessments and the provision of feedback in the classroom. Besides being used as teaching assistants, GenAI can be instrumental in creating feedback to assist students. The challenge for educators is to secure unbiased feedback and identify AI ‘hallucinations’ as accurate and appropriate feedback is critical in students’ conceptual and cognitive development (Van der Kleij & Lipnevich, 2021). Grading students’ written responses and providing personalised feedback
can take a substantial amount of time. It is commonplace in many higher education institutions to prefer high-stakes final exams over continual formative assessments. However, these forms of exams typically incentivise rote learning and discourage risk-taking and innovation (Goita & Kone, 2023). In addition, the traditional manual paper grading limits individualised evaluations and feedback. Such exams also discourage collective learning and instead promote unnecessary competition among students (Naseer et al., 2024).

Educators can feed GenAI with the assignment requirements and the relevant scoring criteria, but students can also use AI to fill in the results of quizzes and complete their assessments with a simple prompt. A good prompt will increase the chance that GenAI provides students with the correct and helpful feedback to facilitate their learning. It is also possible that the AI system will suggest that students review the assignment guidelines and challenge them to work on the areas of improvement (Chen et al., 2020). When integrating such quizzes into the curriculum, educators must take into account the veracity of the learning process and observe if students are merely completing quizzes without learning the materials effectively. Educators must consider evaluating students’ learning through various forms of assessment to ensure rigour in the assessment process.

It is important to note that assessments go beyond just a grading process (Author, 2023a). Assessments involve the use of different forms of questions to demonstrate students’ learning and measure their understanding beyond the subject matter, providing real-time feedback, adapting to varying abilities, assessing the learning process, and documenting ongoing progress as opposed to only achieving a single score to indicate competency or mastery of a particular subject (Adom et al., 2020). It is possible that AI-enhanced formative assessments can be designed and adapted to suit different contexts, reflect varying levels of difficulty, facilitate the learning development of students through formative assessments, and support students’ learning through real-time feedback to encourage reflection and remedial action. However, it is necessary to understand that a techno-optimistic view of AI’s use in education and assessment can be counterproductive and, in fact, detrimental. This is because studies have demonstrated that creating fair and equitable assessments is one area that AI struggles with (Madaio et al., 2022; Yu, 2020; Fenu et al., 2022). It is also imperative that if AI is considered for developing assessments, careful considerations must be taken to ensure that its design is ethical, its implementation is careful, and its performance is reliable.

**Adaptive Curriculum Development**

Several studies have highlighted the benefits of using GenAI for adaptive curriculum development (Ruiz-Rojas et al., 2023; Shimizu et al., 2023; Singha & Singha, 2024). Unlike traditional AI models that operate based on predetermined patterns (Drisgas et al., 2009), GenAI applies advanced algorithms, such as Generative Adversarial Networks (GANs) and Recurrent Neural Networks (RNNs), to acquire different sets and patterns of information from vast datasets and produce authentic content that closely resembles the ones created by humans (Khosravi et al., 2022). As such, it is possible for GenAI to transform learning experiences and empower developers to create more personalised and engaging content for their students.

Videos can be used as resources for the curriculum. Using AI editing tools, generative text-to-video apps, and video productivity apps can be used to create content to complement the
curriculum (Rebelo, 2023). For instance, Runway can be used to experiment with GenAI and Visla to turn a script into a video (Rebelo, 2023). Moreover, GenAI is capable of creating immersive and interactive learning experiences through the creation of virtual learning environments (Bondarenko et al., 2020). Simulations and virtual labs create dynamic learning environments, allowing students to experiment and make mistakes without real-world consequences, encouraging critical thinking and problem-solving skills. This practical approach to learning can significantly boost student engagement and comprehension, making complex subjects more accessible and manageable. For example, Wolfram is useful for teaching and learning mathematics, and Video Insight is suitable for summarising long YouTube videos (Mollick & Mollick, 2023).

In addition, given its promising adaptive ability, GenAI can take the same curriculum and adapt it for different levels and varying learning abilities using similar materials and resources (Kadaruddin, 2023). Furthermore, it is possible for educators to use GenAI to keep the curricula updated with the latest information, statistics, and case studies (Chang & Kidman, 2023). Finally, given the dialogical capabilities of GenAI, such as GPT-4, it may also be a helpful partner for thinking dialogically through one’s ideas when developing the curriculum using GenAI as a critical partner.

Research Assistant

GenAI can also be useful in the field of research. It has the potential to help researchers analyse hypotheses in a given subject matter and use them to generate new ones (Aithal & Aithal, 2023). GenAI is also beneficial in conducting literature reviews, analysing large datasets, and summarising findings on a particular subject. It can establish patterns and correlations and provide insights that may not be immediately obvious to human researchers in a short time, thus making the research process more efficient (Ooi et al., 2023). It can also help combat writer’s block by bouncing ideas and drafting routine written correspondence such as interview and research collaboration invites (Gunser et al., 2021).

Trinka, for example, is an AI-powered writing assistant that is helpful in demonstrating the impact of GenAI in the research process. It has the necessary tools to help educators correct mistakes in their work (Sagre & Ahlawat, 2023). Interestingly, it also makes suggestions to improve the tonality and phrasing of the research article one is writing. In addition, Trinka is helpful in checking that the research paper we want to submit is aligned with a journal’s style guide. Tambo and Snyder (2024) highlighted how GenAI can ‘boil down’ scores of documents into a succinct, digestible summary. This brings to mind Scholarcy, another AI tool that helps create summaries of research articles and highlights the critical parts of the article for the researcher. One other useful AI-powered tool is Scite. It helps researchers and scholars assess the quality of articles and can indicate the context in which an article has been cited and how it was applied in that context (Lund & Shamsi, 2023).

Besides the core research tasks, GenAI can also be used for data visualisation – creating figures, diagrams, process flows, and graphs (Suryadevara, 2020). The different types of AI-driven tools, such as Appy Pie’s Free AI Graph Maker, Chartify.ai, and Graph Maker, allow users to create graphs quickly without the need for complex coding skills (Rudolph et al., 2023d). Another tool, VizGPT, also has a user-friendly chat interface to create data visualisations using natural language queries (Mikami, 2023). However, Kooli (2023) and
Wang et al. (2023) highlighted how using Gen AI in education raises ethical concerns, particularly about the authenticity and accuracy of the generated content. There is a need for safeguards to be put in place to ensure the credibility and reliability of AI-generated materials, and for most reputable journals, GenAI is not accepted as a co-author (Zhang, 2023; Crawford et al., 2023b).

**Automated Grading**

Automated assessment grading is another significant benefit that GenAI offers. Grading systems driven by GenAI can quickly evaluate assignments and assessments using predetermined rubrics and grading frameworks, providing immediate and constructive feedback (Luckin, 2017). The speed and expediency GenAI offers in the grading process can reduce the workload of educators and assessors but requires careful application. One pronounced ethical challenge in AI use in assessment is ensuring fairness and equity (Popenici, 2023b). As such, it is important to ensure that automated grading systems are observed for biases and inconsistencies in their application. Although using Natural Language Processing algorithms enables automated grading systems to interpret written feedback and revise it for accuracy and validity (Darvishi et al., 2022), it is equally critical for educators to carefully evaluate these interpretations for inaccuracies and biases.

A grading system powered by GenAI algorithms can be useful for reviewing students' answers for coherence, competence, and quality (Alasadi & Baiz, 2023). The system can provide educators with reports that can highlight gaps in students’ knowledge. In addition, a GenAI system can possibly identify patterns in students’ answers and recognise and highlight areas where students struggle or have misconceptions (Preiksaitis & Rose, 2023). It is possible for educators to use this information to help struggling students and revise it to address the gap in students’ learning. The consideration of using AI to facilitate these efforts, however, must be approached with caution.

**Collaborative Circles**

The Socratic method has been touted as a cornerstone of engaging classroom discussions and dialogues (Mahoney et al., 2023). Based on critical thought and active interactions among learners, dialogic learning has also been defined as a form of authentic learning (Levine & Rascoff, 2020). Although there is some scepticism about creating effective discussion platforms online (Donelan & Kear, 2023), if carefully thought through and appropriately implemented, engaging online discussion platforms can be helpful (Sharma & Yadav, 2023). GenAI can facilitate setting up such discussion platforms and foster creative collaboration among students and educators (Mollick & Euchner, 2023). It can be especially challenging to get students to come together in groups to discuss and collaborate in some contexts (Hew et al., 2023). To facilitate interdisciplinary discussion and get students to engage in critical conversations and think creatively, it is vital for them to work in groups. GenAI can help by organising discussion forums online and monitoring the content posted in these forums (Tomić et al., 2023). Through this monitoring, GenAI can decipher the effectiveness of peer learning through the discussion forums, the gaps in their knowledge, and the progress of their learning (Bozkurt et al., 2023).
Kim et al. (2022) described in their study how GenAI can present real-world problems for students to think about and discuss. GenAI applications can stimulate the discussion by offering scaffolding-driven, detailed, step-by-step instructional support (Kim et al., 2022). Further to this support, as discussed in the previous section, GenAI can measure the level of participation in the discussion forum and evaluate students’ engagement in the forum. Furthermore, Liu et al. (2023) suggested that GenAI, used for setting up collaborative circles, would allow students to be evaluated in real-time based on their own merit and possibly reduce incidences of plagiarism, ghosting, and collusion. AI-driven discussion platforms such as Packback have been used to reduce instructor and grading workload while maintaining the quality of student learning and ensuring effective student-teacher engagement (Rutner & Scott, 2022). If applied in the appropriate contexts and its strengths correctly utilised, AI applications can replicate face-to-face classroom discussions.

**Student Intervention**

One potential application of GenAI in higher education could be to facilitate student intervention. GenAI can use data from students’ academic performance to provide suggestions for educators to help them identify at-risk students. In addition, GenAI can provide suitable recommendations and optimise intervention programs to help educators assist these students (Jiao et al., 2022). One benefit of this proactive approach is that it provides opportunities for educators to initiate early intervention and the provision of appropriate support for struggling students (Celik et al., 2022). In addition, providing these interventions will likely facilitate the improvement of students’ performances and wellbeing.

Integrating AI into student support and intervention will create more personalised, process-oriented instruction and learning (Luckin & Cukurova, 2019). Educators, however, must realise the futility of complete reliance on AI and be determined that any use of AI analytics must be managed with care and caution. Educators play a central role in assessing and identifying students who require intervention. Using the educator’s assessment, Gen AI can be used to suggest changes to pedagogical approaches, the crafting of appropriate intervention strategies, or to improve the quality of learning resources (Ouyang et al., 2023). Providing reminders for at-risk students, personalised learning plans for students who need support, and tutoring and learner models (Sandoval et al., 2018) is useful for educators and students. Despite this opportunity to use AI for the students, it must be emphasised that educators must take on the lead role in managing the use of GenAI for initiating student intervention and assessing the veracity of students’ results.

**Administrative Support**

Administrative professionals in most higher education institutions perform critical but routine tasks such as managing schedules, documenting student enrolment and performance, handling standard communications, organising meetings, and providing support for the management and teaching staff (Adeshola & Adepoju, 2023). However, administrative staff are now confronted with a harsh reality in that the jobs that they perform can now be automated and driven by GenAI. Such systems can also help the organisation manage the security of the data and the data recovery process during a crisis or an emergency (Raimundo & Rosário, 2021).
Besides the advantages of its application in academia, which have been discussed in previous sections, GenAI can also be used in the administrative process in higher education institutions. For example, GenAI can suggest how to improve our written correspondence and craft templates for email replies (AlAfnan et al., 2023). Moreover, given its distinct ability to automate routine processes (Javaid et al., 2023), using GenAI to automate routine administrative tasks in higher education, such as consolidation of grades, creating performance summaries of students, crafting out evaluation reports, and maintaining student enrolment and attendance records (Meyer et al., 2023). Automating such routine tasks is particularly advantageous as administrative staff can focus on more critical processes such as student recruitment, decision-making, focusing on operational tasks, and supporting teachers in the classroom (Chiu, 2023).

Inclusive Education

Inclusive education is defined as education in which students with disabilities have access to the standard curriculum in a general education classroom (Bryant et al., 2019). Some of these students are challenged by poor vision, suffer from speech impairment or auditory ailments, or are limited by attention deficit hyperactivity disorder (ADHD) or other similar disorders (Spaeth & Pearson, 2023). It is critical, however, that these students have equitable access to education and receive resources that are adapted to their learning needs.

One distinct advantage of GenAI is how it has the potential to improve accessibility to education for vulnerable communities when the learning needs of these students are considered. GenAI can create accessible forms of educational materials for students with disabilities, such as descriptions for visual educational content. In addition, GenAI can create alternative formats for course materials or textbooks and provide text-to-speech support for visually impaired learners (Mina et al., 2023). Harnessing the features of GenAI to adapt to different learning needs, paces, and styles is helpful, particularly for neurodivergent learners (Jesse, 2024). Furthermore, neurodivergent learners may find the GenAI interface to be a non-threatening learning environment, and the interactive dialogues are helpful for giving them confidence in their learning (Roscoe et al., 2022). Although the introduction of GenAI is fraught with challenges, its capabilities may be significant in providing students with special needs access to educational resources and experiences.

Wellbeing Initiatives

The pace in many higher education institutions is fast and unrelenting (Treve, 2021). The wellbeing of students, educators and support staff has increasingly become critical. AI can be an excellent collaborative partner in supporting the wellbeing of both teachers and students. If used judiciously, GenAI may be able to offer valuable resources to support the stress and anxiety levels of students (Fitzpatrick et al., 2017) while reducing the workload for teachers (Hashem et al., 2024).

Besides supporting the wellbeing of students, GenAI can gain insights into the wellbeing of teachers by analysing the impact of the organisations’ wellbeing initiatives and providing appropriate changes and suggestions (Budhwar et al., 2023). This approach allows the constant adaptation of wellbeing initiatives to help teachers cope better. In addition, by monitoring operation workflows and teachers’ tasks, the organisation can utilise the data from these analyses to optimise teachers’ workload and revise their teaching schedules (Hashem
et al., 2024). This can significantly reduce work-related stress, allowing teachers to manage their time and workloads better and enjoy a healthier work-life balance (Pataranutaporn et al., 2021).

Mollick (2022) bluntly described GenAI, as a consummate bullshitter that holds significant potential. With this thought, if we seek to be moral agents in higher education, it becomes critical to not only explore opportunities presented by GenAI but also be pertinently aware of the threats it possesses.

**Threats**

*I lived in daily fear, lest the monster who I had created should perpetrate some new wickedness.* (Victor Frankenstein in Shelley, 2017, p. 73).

GenAI has sparked unprecedented hype and enthusiasm as well as alarmist predictions for universities, notably Robin Raskin’s forecast of a reduction by half in U.S. higher education institutions (Anderson & Rainie, 2023). The advent of GenAI has unveiled a plethora of threats to higher education and the broader societal fabric, encompassing concerns over authorship, copyright infringement, and the potential for misuse in activities such as spamming, disseminating fake news, creating malware, and hacking (Marcus, 2023; Rudolph et al., 2023a, 2023b). A notable moment in March 2023 saw Elon Musk, Steve Wozniak, and prominent AI researchers advocating unrealistically for a halt to GenAI development, citing significant risks to democracy, employment, and education (Future of Life Institute, 2023). This section delves into ten critical threats posed by AI to higher education. While varying in scope, these threats resonate across the higher education landscape, reflecting broader socio-economic implications. Our analysis acknowledges the interconnected nature of higher education challenges within a wider societal and economic context.

**Academic Integrity Challenges**

In a dystopian academic future scenario, assessments are generated, completed, and graded through GenAI, culminating in an educational void where learning by both students and teachers is conspicuously absent (Popenici et al., 2023). This encapsulates the dire implications of GenAI on academic integrity within the sphere of higher education amidst a disturbing trend of declining human intelligence (Popenici, 2023a).

The proliferation of GenAI technologies like ChatGPT poses significant challenges to traditional assessment methods, particularly essays and online examinations. The concern that students might utilise AI for completing assessments, thereby circumventing plagiarism detection, is not unfounded. Despite advancements in AI detection tools, the effectiveness of conventional anti-plagiarism software, such as iThenticate and Turnitin, remains questionable against AI-generated content (Perkins, 2023; Chaka, 2023, 2024; Hassoulas et al., 2023; Ifelebuegu, 2023; Mohammadkarimi et al., 2023; Sullivan et al., 2023). It is relevant to note that OpenAI, the creator of ChatGPT, provided in a public release an answer to educators to the important question “Do AI detectors work?” with a succinct “No, not in our experience.” It further noted that the advanced tool designed to detect AI-generated text “labeled human-written text like Shakespeare and the Declaration of Independence as AI-generated” (OpenAI, 2023). The indistinguishability of AI-generated responses from human work presents a formidable challenge in maintaining fairness in evaluation and upholding academic standards.
(Michel-Villareal et al., 2023). Such developments risk devaluing academic qualifications should academic misconduct like cheating go unchecked (Anft, 2023).

Critiques of written assignments as uninspiring and ineffective in assessing true student learning persist (McMurtrie, 2023). While the efficacy of assignments is debatable, the imperative of preserving academic integrity is undeniable. An emerging, albeit imperfect, solution advocates for students to formally recognise the use of GenAI, disclosing its application in their work (Rasul et al., 2023). Resistance to altering traditional assessment models – predominantly assignments and examinations – is evident. However, the advent of GenAI necessitates exploring alternative assessment forms, such as oral or pen-and-paper exams, in-class essays, oral presentations, group discussions, practical laboratories, and fieldwork (Author, 2023a).

The response of higher education institutions to the integration of GenAI into student work is notably sluggish. In his research of US university leaders’ responses to GenAI, Anft (2023) reports a lack of proactive measures among US university leadership to adapt assessment strategies and develop comprehensive policies to govern the use of GenAI. A mere fraction has initiated dialogue or established committees to deliberate its implications, revealing a significant gap in policy and practice (Anft, 2023).

Dilution or Automation of University Teacher Roles

The advent of AI in education threatens to relegate university teachers from subject specialists and mentors to mere facilitators of AI-curated content, undermining their expertise and value in the educational landscape. Universities worldwide face a unique dilemma: despite their unprecedented number and economic significance, they experience a profound crisis of confidence and identity (Collini, 2012). Since the pandemic, there has been an increase in the precarity and casualisation of academic labour, which is undervalued, overused, and stigmatised (Solomon & Du Plessis, 2023). The trends of the fall of the faculty (Ginsberg, 2011), the loss of faith in university education (Collini, 2012), the demise of homo academicus (Fleming, 2021), the denigration of the teaching profession, the devaluation of learning and the commodification of higher education (Popenici, 2023a) all precede the rise of GenAI in higher education.

Historical cycles of educational technology – from Edison’s predictions about motion pictures imminently replacing books (Terzian, 2019) to the envisioned role of radio as a ‘Master Teacher’ (Cook, 1938), television’s educational promises (Terzian, 2019), and the advent and challenges of computers in classrooms (Watters, 2021) – have often promised more than they delivered (Kefalaki et al., 2022). Longstanding predictions are that AI will replace teachers (Houser, 2017). Recent developments in GenAI technologies renew these debates, with financial pressures on universities fuelling the discourse (Chan & Tsi, 2023; Devlin, 2023). For instance, Haw (2019) worried that “swapping expensive lecturers for cheap, versatile machines that don’t go on strike, don’t need sleep, and respond to students within nanoseconds will be hard to resist”.

Selwyn (2019) identified the digital automation of teaching as a pivotal challenge for education in the coming decades. The role of teachers, while not supplanted by robots, is increasingly encroached upon by a myriad of digital entities such as software, applications, and various AI-driven tools tasked with pedagogical functions. This technological encroachment puts the esteemed professional standing of university lecturers at risk, signalling a potential shift in the educational paradigm (Selwyn, 2019). We answer the questions of whether (ro)bots (1) can, (2) will, or (3) should replace teachers with (1) ‘yes’, (2) ‘if we let them’ and (3) ‘no’. Whether teachers will be relegated to an AI assistant’s role or automated away entirely will be a societal, economic and political choice that additionally depends on whether we fall prey to a blind
technological solutionism (Morozov, 2013). Turkle (2011) has long cautioned against the diminished human interaction that comes with learning alongside robots. Turkle’s (2011) warnings about the illusion of companionship offered by robots – absent the complexity, richness, and demands of genuine human relationships – underscore the nuanced and potentially subtle harms that may come with integrating machines into learning environments.

Quality, Accuracy, and Ethical Concerns

GenAI, epitomised by models such as GPT-3 and its more expansive successor, GPT-4, presents many quality, accuracy and ethical challenges. GPT-3’s training on a vast corpus of 45 terabytes of text, sourced from a variety of internet-based entities like Common Crawl and Wikipedia, set a precedent for the scale of data utilised in AI development (Brown et al., 2020; Rudolph et al., 2023a). GPT-4 expanded this dataset further, processing an “internet scale” amount of data, as described by OpenAI’s president, Greg Brockman, indicating a dataset wide enough to encapsulate the linguistic diversity of English online (Metz, 2023).

However, the quality and accuracy of the input data for these models raise significant concerns. The adage ‘garbage in, garbage out’ aptly summarises the issues of biased, obsolete, or flawed data leading to potentially misleading or harmful AI outputs. Critics argue that despite AI-generated text’s sophisticated syntax and grammar, it lacks creativity, critical thinking, and wisdom, producing plausible text without genuine intellectual depth (Popenici, 2023b). The bots are but “stochastic parrots” (Bender et al., 2021).

Tests have revealed that AI chatbots, including GPT-4, routinely fail to reference academic literature accurately or rely on substandard sources, underscoring the models’ limitations in scholarly and intellectual rigour (Rudolph et al., 2023b). The reliance on easily accessible but not always reliable sources like Wikipedia or Khan Academy and the omission of open-source academic literature highlight a significant gap in AI’s reference generation capabilities. The unchecked input into chatbots opens the door to misinformation and the propagation of ‘junk science’, with ChatGPT being critiqued for its ‘mansplaining’ – its overconfidence in inaccurate answers (The Economist, 2023a). Furthermore, the advent of GenAI technologies, capable of generating ultrarealistic images, audio clips, and videos, exacerbates concerns over deepfakes and misinformation, raising ethical questions regarding deception and the erosion of digital communication trust (Suleyman & Bhaskar, 2023).

Cybercriminals can replicate family voices from brief audio clips, leading to deceptive emergencies. AI-generated deep-fake pornography recorded an exponential rise, increasingly affecting people across the world, including students, teachers, schools and institutions of higher education (Kassova, 2024; OECD AI, 2023). Digitally cloned celebrities like Tom Hanks and Taylor Swift are falsely promoting products, and the internet is rife with manipulated videos featuring faked politicians (The Economist, 2024). The increasing sophistication of AI technologies makes it challenging to distinguish between genuine and artificial content. In early 2023, an AI-generated fake image of the Pentagon in flames rattled for a short time the stock exchange in the United States, with S&P 500 briefly dropping 0.3% (Sytsma et al., 2024). As AI advances, creating near-perfect digital fabrications becomes more feasible, outpacing the capabilities of detection software. This scenario raises concerns about misuse:

Dystopian possibilities abound. It will be difficult, for example, to avoid a world in which any photograph of a person can be made pornographic by someone using an open-source model in their basement, then used for blackmail... Perhaps anyone will be able to produce
a video of a president or prime minister announcing a nuclear first strike, momentarily setting the world on edge (The Economist, 2024).

The integration of GPT technology into mainstream products, such as Microsoft’s suite, heralds a shift towards a partially ‘cyborgised’ existence, even without physical implants (Metz & Weise, 2023). This evolution brings to the fore the pressing need for a careful examination of GenAI’s societal implications, particularly concerning informational integrity and ethical considerations.

**Technological Colonialism, Mono-Culturalism, and the Concentration of Power**

The integration of external AI technologies in higher education posits the risk of institutions becoming excessively reliant on tech giants for educational content and pedagogy. This issue nests within the broader discourse on the concentration of power and wealth by Big Tech, including AI powerhouses like Alphabet, Amazon, Apple, Meta, Microsoft, Nvidia, and OpenAI in the US, Samsung in South Korea, and Alibaba, Huawei, and Tencent in China (cf. Abdalla & Abdalla, 2021; Jacobides et al., 2021). Intense corporate rivalries and the “war of the chatbots” are further complicated by the geopolitical AI arms race, notably between the US and China, with entities like the Beijing Academy of Artificial Intelligence at the forefront of research (Rudolph et al., 2023b; Metz, 2022; Lee, 2018).

Core infrastructure platforms (Van Dijck et al., 2018), used by all the other platforms to operate, have enabled Big Tech companies such as Alphabet, Meta and Microsoft to collect and combine diverse data flows, resulting in a “Big Technification” of every aspect of social life (Hendrikse et al., 2022). Big Tech’s centralisation of AI development and control and their strategies to dominate AI innovation mark a convergence around AI. Such oligopolisation not only accrues disproportionate economic benefits to these entities but also amplifies wealth inequality, impacting societal welfare and the public good (Rikap, 2023). The narrative of Big Tech’s influence extends to controlling and manipulating information dissemination, privacy norms, and cultural narratives, often prioritising profit over societal needs. This is exemplified by the recent turbulence within OpenAI’s governance (shortly after the board fired CEO Sam Altman, most of the board members were sacked and Altman reinstated) and the commercialisation of large language models (LLMs) by Google, Meta, Anthropic, Inflection and others, reflecting capitalism’s adaptability and pervasiveness (Klein, 2023). Developments of the most influential and widely used LLMs frequently demonstrate a disconnect and incongruence between corporate capital interests and educational aims.

Acemoglu and Johnson (2023) assert that technology’s potential to foster widespread prosperity is contingent upon economic, social, and political choices, highlighting the stark disparities in wealth and access. The disproportionate accumulation of wealth by a small elite, juxtaposed with the abject poverty affecting hundreds of millions, underscores the exacerbating inequalities (Rudolph et al., 2023b). The disparities in access to AI technologies accentuate existing educational inequalities, privileging those with superior resources. Digital inequity, stemming from uneven distribution of technology and internet access, risks widening the knowledge gap among students in the global North and South. It is imperative for higher education institutions to mediate these disparities by facilitating access to technology and making GenAI an inclusive tool (Adarkwah et al., 2023; Kouam & Muchowe, 2024).

**Erosion of Graduate Attributes**

The advent of GenAI raises concerns about the continued attainability of graduate attributes. As society gravitates towards an increased reliance on AI for tasks traditionally demanding human ingenuity, the threat of a diminished capacity in essential cognitive skills such as critical
thinking and problem-solving among students looms. The capability of AI to produce, or at least 'co-pilot', art, music, and literature prompts a critical reflection on the future of human creativity. By outsourcing creativity to, for instance, text-to-image AI such as DALL-E 3 and Midjourney, its depersonalisation poses a risk of homogenising culture and diminishing the value placed on human creativity and personal expression. In addition, the increasing encroachment of AI into areas requiring personal interaction threatens to undermine the development of interpersonal skills and emotional intelligence, further compromising the holistic development of graduates (Limna, 2023). This development challenges the nurturing of a suite of graduate skills identified as pivotal for future success, including but not limited to critical thinking, communication, collaboration, leadership, adaptability, digital literacy, and ethical professionalism (Rasul et al., 2023).

**Graduate Employment**

The ascendancy of artificial intelligence heralds significant disruptions across the labour market, with automation and AI tools precipitating worker displacement in myriad sectors. Amidst this technological upheaval, the creation of new job opportunities remains shrouded in uncertainty, both in terms of volume and accessibility to those displaced, posing a risk to economic stability and employment levels. Under the prevailing dynamics of informational capitalism, AI integration into the workforce paradoxically engenders a demand for low-status, precarious labour (Lindgren, 2023a). This phenomenon perpetuates the cycle of digital labour, wherein humans inadvertently support AI’s operational framework, often without remuneration, echoing historical patterns of labour exploitation and monotonous work conditions under the guise of technological advancement (Lindgren, 2023a; Berman, 1992).

Historically, societal attitudes towards labour have oscillated dramatically, with compulsory work enforced upon the impoverished to stave off idleness, as chronicled in medieval British laws and by Orwell (1933; Rudolph et al., 2023c). Conversely, the leisure class has been at times dissuaded from labour, advocating for reduced work hours to enrich personal and civilisational development (Russell, 2004). The contemporary dialogue on labour, amplified by the rise of automation and AI, reiterates the necessity to re-evaluate work-life paradigms, contemplating frameworks like Universal Basic Income to mitigate the fallout from potential mass job displacements (see Bastani, 2020; Susskind, 2021). Regardless of how this debate will conclude, we can anticipate a seismic shift caused by GenAI on the job market, significantly impacting career prospects for recent and future graduates.

The labour practices underpinning AI development, particularly the exploitation of low-wage workers in the Global South for tasks like data labelling, underscore the industry’s reliance on hidden human labour. OpenAI’s engagement with Sama, employing workers from Kenya, Uganda, and India, reveals the distressing psychological impacts and inadequate compensation tied to moderating hazardous content, highlighting the exploitative dimensions of AI’s supply chain (Perrigo, 2023; Yalalov, 2023).

Furthermore, the notion of skill mismatch and obsolescence emerges as AI-driven education struggles to align with the evolving demands of the job market, raising concerns about the future relevance of current educational curricula (Crawford, 2021). Crawford (2021) articulates the multifaceted layers of exploitation inherent in AI’s ecosystem, from the underpaid labour essential for AI’s construction and maintenance to the deceptive efficiencies of ‘fauxtomation’. This critique extends to the broader implications of automation, highlighting the widening chasm between the beneficiaries of technological advancements and those relegated to increasingly insecure, under-compensated roles.

**Privacy and Surveillance**
In the contemporary digital landscape, data has transcended its personal and intimate confines, evolving into a ubiquitous commodity – “everything that could be captured” (Crawford, 2021, p. 113). This transformation, facilitated by paradigms such as ‘data mining’ and the adage ‘data is the new oil’, marks a pivotal shift in the perception of data from a personal asset to an inert, exploitable resource. The rhetoric surrounding this shift underscores a significant departure from the notion of data as subject to individual ownership and control.

The discourse on AI, particularly from figures like Peter Thiel (2019), posits AI as a technology with profound military utility, emphasising its application in mundane yet critical tasks such as computer vision and data analysis. These capabilities play a crucial role in highlighting the dual-use nature of AI technologies, which are applicable in both civilian and military contexts (Crawford, 2021). It would be ill-advised to ignore the fact that the inception of AI in the late 1950s and its later developments remain closely associated with military interests and funding.

The advent of GenAI, with its unparalleled data processing and generation abilities, introduces acute privacy dilemmas. The potential for these technologies to be harnessed for surveillance, tracking, and profiling individuals without their consent presents a stark challenge to privacy norms. This concern is compounded by deploying “secret models” that can impose arbitrary punishments, affecting individuals at crucial junctures of their lives, from education and employment to sentencing and credit acquisition (O’Neil, 2016). The integration of AI in law enforcement and judicial processes heralds a future where algorithmic decision-making obscures accountability. Public failures with vast costs and irreparable losses, such as “robo-debt” in Australia, the Dutch “toeslagenaffaire” or the British Post Office scandal (Horizon), already serve as a warning over extraordinary risks of algorithmic decision-making (Baset, 2023; Alon-Barkat & Busuioc, 2023; McGuire & Renaud, 2023). The refrain ‘the algorithm told me to do it’ epitomises the abdication of human responsibility, ushering in an era of automated decision-making characterised by minimal oversight (Crawford, 2021), which comes with uncontrollable consequences.

The escalation of surveillance capabilities, alongside the burgeoning collaboration between private contractors, law enforcement, and the tech sector, signifies a profound expansion of surveillance infrastructure. This convergence, often fuelled by clandestine agreements and financial incentives, redraws the contours of civic life, fortifying the centres of power with technologies imbued with the logics of policing and militarisation (Crawford, 2021).

Bias

An intersectional analysis reveals that capital, race, and gender significantly contribute to algorithmic oppression, with these categories being politically, culturally, and socially constructed rather than biological (Lindgren, 2023a; Crawford, 2021). The act of classification itself centralises power by determining which differences are acknowledged, effectively perpetuating inequalities.

Investigations into high-tech systems highlight their role in automating inequality, particularly against poor and working-class individuals (Eubanks, 2017). Noble (2019) exposes that search engines’ AI algorithms, under the guise of neutrality, can draw on sexist and racist stereotypes, particularly affecting representations of African-American women. Instances of racial bias in AI include Google Photos misidentifying black individuals as ‘gorillas’, Facebook’s AI labelling them as ‘primates’, and Amazon’s facial recognition tool falsely associating members of the Congressional Black Caucus with criminal mugshots (Vincent, 2018; Mac, 2021; Singer, 2018). Furthermore, the standard female voices of digital assistants such as Siri, Alexa, and Cortana reinforce sexist norms by implying that women should be perpetually available and subservient (Lindgren, 2023a). The Stanford Institute for Human-
Centered AI’s 2021 AI Index Report reveals that across the world, only 16 per cent of tenure-track faculty positions with a focus on AI are held by women (Zhang et al., 2021).

Broussard (2018) critically reflects on the demographic homogeneity of those leading AI development, suggesting that the dominance of a small group of men with specific ideological leanings and a disregard for social norms poses significant risks to societal well-being and equity:

*We have a small, elite group of men who tend to overestimate their mathematical abilities, who have systematically excluded women and people of color in favor of machines for centuries, who tend to want to make science fiction real, who have little regard for social convention, who don’t believe that social norms or rules apply to them, … and who have adopted the ideological rhetoric of far-right libertarian anarcho-capitalists. What could possibly go wrong?* (p. 85)

**Sustainability**

In the Anthropocene, humanity stands at the precipice of ecological oblivion, a dire consequence of our collective inertia (Tan & Rudolph, 2023). The Anthropocene, catalysed by the Great Acceleration of the 1950s, marks an unprecedented onslaught on the biosphere, characterised by skyrocketing carbon dioxide levels, methane emissions, and a cascade of environmental calamities – from melting ice caps to decimated forests (Attenborough, 2020; Ripple et al., 2017; Wagler, 2011). This relentless march towards environmental degradation is not merely a symptom of industrialisation but a testament to a systemic disregard for the planet’s finite resources, fuelled by a voracious appetite for consumption and a myopic vision of progress (Tan & Rudolph, 2023). We are ensnared in the sixth mass extinction, with biodiversity loss occurring at rates 100 to 1,000 times the natural extinction rate (Pimm et al., 1995, 2014), signalling a catastrophic departure from the Holocene’s stability. Our quotidian existence, ensconced in unsustainable practices, propels us towards a future where the Earth’s capacity to sustain life, as we know it, dwindles (Klein, 2014; Schneider-Mayerson, 2020).

It is meaningful to adopt a comprehensive lifecycle approach to AI that unveils its multifaceted ecological costs (Brevini, 2023). The lifecycle begins with extractivism, where procuring rare metals and minerals essential for AI hardware unfolds (Smart, 2017). This phase is deeply intertwined with technocolonialism (Madianou, 2021), which reflects enduring colonial legacies and inequalities within the global supply chains of AI. The extraction processes can be exemplified by the surge in global demand for lithium (a crucial element in the production of rechargeable batteries) and its extensive environmental and social costs (Crawford, 2021; Lindgren, 2023a). There are “repressed stories of acid-bleached rivers and deracinated landscapes and the extinction of plant and animal species” (Crawford, 2021, p. 36).

The production and operational phases of AI introduce additional significant environmental burdens. The training of AI models incurs high environmental costs, a consequence of the substantial energy consumption and emissions generated by converged communication systems integral to AI functionalities (Brevini, 2023). These systems not only necessitate vast amounts of energy but also contribute to material toxicity and the proliferation of electronic waste (Brevini & Murdoch, 2017). Cloud computing, an integral part of machine learning and GenAI systems, had a greater carbon footprint than the airline industry in 2022; notably, a single data centre was spending as much electricity as 50,000 homes (Monserrate, 2022). Obviously, since 2022, we have registered an extraordinary expansion of cloud computing, AI, and machine learning systems, which require significantly higher energy consumption and have a massive environmental impact.
A critical aspect often overlooked is the water footprint of AI, particularly concerning data centres. These infrastructures require continuous, large-scale water supplies for cooling purposes, further exacerbating AI's environmental impact (Brevini, 2023). The lifecycle culminates in the disposal phase, where the issue of electronic waste emerges as a stark manifestation of global inequities. The burden of e-waste is disproportionately offshored to developing economies, transforming these nations into digital dumping grounds for more affluent countries (Brevini & Murdock, 2017). The briefly sketched-out lifecycle analysis of AI, from extractivism to e-waste, unveils the profound and pervasive environmental costs accompanying AI's technological advancements. It underscores the necessity for a paradigm shift towards more sustainable and equitable practices within the AI domain, challenging Big Tech and governments to reconcile the technological imperatives of AI with the pressing need to mitigate its environmental impact.

**Regulatory and Policy Challenges**

Regulatory and policy challenges pose threats at intergovernmental, country and university levels. The rapid development of AI technologies outpaces the formulation of regulatory frameworks. The lack of clear regulations poses a risk of unchecked AI development and deployment, potentially leading to harmful outcomes. Amidst efforts to regulate AI, including the EU’s ambitious AI act and anticipated executive orders in the US, there is a push for global cooperation, given AI’s borderless nature. However, the rapid evolution of AI, particularly large language models (LLMs), complicates regulatory goals and methods. While tech companies advocate for narrow regulation focusing on extreme risks, broader concerns remain about AI’s societal impacts. Proposals range from model-specific regulations to broader application-based approaches, with discussions on the appropriate regulatory body and the potential for an Intergovernmental Panel on Climate Change (IPCC)-like entity for AI (The Economist, 2023b, 2023c).

The European Union aims to regulate AI through a pioneering regulatory framework proposed in April 2021, focusing on ensuring safe, transparent, traceable, non-discriminatory, and environmentally friendly AI applications (European Parliament, 2023). This framework categorises AI systems by risk levels, introducing specific regulations accordingly, with a ban on AI posing unacceptable risks, such as cognitive behavioural manipulation and social scoring. High-risk AI systems affect safety or fundamental rights. They require rigorous assessment and registration. GenAI must meet transparency standards, while limited-risk AI systems should ensure user awareness. A provisional agreement on this AI Act was reached in December 2023 (European Parliament, 2023). While the EU’s framework is laudable, there is much reason to be sceptical. A brilliant cartoon in The Economist magazine by KAL says it perhaps best: During a recent AI summit, the cartoonist imagined that the UK, Europe, the US and China paid lip service by saying, “We declare that AI poses a potentially catastrophic risk to humankind”, while silently thinking “…and I cannot wait to develop it first” (The Economist, 2023d).

The public introduction of GenAI tools like ChatGPT has significantly impacted higher education institutions’ assessment practices, raising concerns about academic integrity. Moorhouse et al.'s (2023) findings indicate that nearly half of the world’s top 50 universities have published guidelines focusing on academic integrity, assessment design advice, and student communication. Moorhouse et al. (2023) advocate for the integration of GenAI into educational assessments, necessitating a new skill set for instructors: generative artificial intelligence assessment literacy.
Conclusions

GenAI, while opening new vistas in higher education, necessitates vigilant governance, ethical scrutiny, and comprehensive regulatory oversight to ensure societal welfare. Amid the unprecedented pace of AI development, typified by vast information consumption capabilities and potential limitlessness (Suleyman & Bhaskar, 2023), the imperative for cultivating critical AI literacy among educators and students has become paramount. This endeavour aims to enhance graduate employability and adapt to the evolving job market by integrating critical AI literacies within curricula (Rasul et al., 2023).

The discourse on GenAI’s integration into higher education heralds a transformative era, necessitating experimentation alongside a cautious approach to risk management (Tan et al., 2024). The recommendations for embedding AI in educational frameworks underscore the critical balance between leveraging technological advances and safeguarding academic integrity, creativity, and ethical standards (Crawford et al., 2023a; Gimpel et al., 2023; Mills, 2023). Amidst these significant challenges, it is at least a piece of good news that The World Economic Forum (2023) anticipates growth in education sector jobs, suggesting an increase in teaching roles facilitated by digital advancements. This projection underscores educators’ need to engage with AI tools, potentially transforming pedagogical strategies and student engagement.

In addressing the broader implications of AI in higher education, it is essential to engage in multi-stakeholder dialogues to forge guidelines that reflect a nuanced understanding of AI’s capabilities and limitations. Educators are encouraged to foster environments that promote critical thinking and digital literacy, preparing students to navigate and critique the AI-augmented landscape responsibly (Gimpel et al., 2023; Popenici, 2023a). We advocate for a tripartite focus on pedagogical innovation, regulatory frameworks, and cultivating an innovation-centric culture within higher education. At the same time, we underline the importance of securing an ongoing critical analysis of the ethical and instrumental aspects of AI use in education at all levels, provided by independent centres of expertise. This approach aims to harness AI’s potential to address accessibility, equity, sustainability and quality in education while ensuring ethical usage and mitigating biases (Tan, 2023).

In conclusion, incorporating AI within higher education is potentially revolutionary as it heralds a paradigm shift in pedagogical methodologies, teaching and learning and academic work. This seismic shift brings with it a double-edged sword: the promise of unprecedented learning enhancements and the peril of crucially important ethical dilemmas. It is imperative, therefore, that the academic community – educators and learners alike – cultivates a robust critical AI literacy. Such critical AI literacy is the cornerstone upon which we can commit to not just navigating but critically shaping the trajectory of AI in higher education, ensuring that it aligns with the overall aims of education and the bedrock principles of equity, integrity, and human-centric values.

Acknowledgements

Thanks to Shannon Tan for creating Figure 1 specifically for this commentary.

Conflict of Interest

We declare that we have no actual or perceived conflicts of interest.
References


Singha, S. & Singha, R. (2024). Revolutionizing content creation and curriculum development through Gen AI. In R. Sharma & A. Bozkurt (Eds.), *Transforming education with Gen AI: Prompt engineering and synthetic content creation* (pp. 261-280). IGI Global. https://doi.org/10.4018/979-8-3693-1351-0.ch013


