ACIS 2023 Proceedings   Australasian (ACIS)

12-2-2023

Navigating the Terrain of Large Language Models in Higher Education- A systematic literature review

Shipra Chhina
Federation University Australia, Australia, s.chhina@federation.edu.au

Bhavna Antony
Federation University Australia, Australia, b.antony@federation.edu.au

Selena Firmin
Federation University Australia, Australia, s.firmin@federation.edu.au

Navigating the Terrain of Large Language Models in Higher Education- A systematic literature review

Full research paper

Shipra Chhina
Institute of Innovation, Science and Sustainability
Federation University
Mt Helen, Australia
Email: s.chhina@federation.edu.au

Bhavna Antony
Institute of Innovation, Science and Sustainability
Federation University
Mt Helen, Australia
Email: b.antony@federation.edu.au

Selena Firmin
Institute of Innovation, Science and Sustainability
Federation University
Mt Helen, Australia
Email: s.firmin@federation.edu.au

Abstract

As Large Language Models (LLM) emerge, opportunities for personalised learning are opening in education. LLMs are a valuable educational tool but raise ethical concerns regarding data privacy, consent, and potential bias reinforcement in higher education. Overusing AI-generated content can compromise critical thinking and problem-solving skills, resulting in less authentic learning. This creates a big question about using this disruptive model in education. To understand this, we attempt to review the use of LLMs in higher education through a systematic literature review utilising PRISMA approach. The findings reveal crucial insights into several challenges, but researchers are inclined to embrace LLMs for their benefits. The analysis reveals that higher education institutions must establish robust academic integrity policies and ensure AI-based assessments do not solely determine learning outcomes. A balanced approach that values social interaction, reflection, and collaboration is key to success in the AI world.

Keywords: Large Language Models, LLM, ChatGPT, Higher education

1   Introduction

"Technology is like a double-edged sword – it offers tremendous opportunities, yet we must tread cautiously as we navigate its incorporation into our lives." - Authors

Large language models (LLM) such as, GPT-3 and BERT, is an AI-based model that generates text like a human and answers questions accurately, among other tasks relating to language (Kasneci et al. 2023). Due to their ability to process vast amounts of text, generate human-like responses (Dergaa et al. 2023), LLMs are increasingly attractive for various applications in higher education (Rasul et al. 2023). Teachers and institutions must critically assess both the benefits and challenges associated with integrating LLMs into educational settings as they explore their potential for enhancing teaching and learning (Alqahtani et al. 2023; Rasul et al. 2023; Rudolph et al. 2023b).

LLMs like BERT and GPT-3 are acknowledged to have transformative potential in higher education (Alqahtani et al. 2023), but a comprehensive understanding of the benefits and challenges remains fragmented and scattered (Rasul et al. 2023). This paper presents a systematic literature review focused on the benefits and challenges of LLMs in higher education to fill this research gap. This analysis aims to identify and synthesise relevant peer-reviewed studies, reports, and academic articles that explore the multifaceted implications of adopting LLMs in higher education using a systematic approach. Therefore, we formulated the following research question for our study:

What are the benefits and challenges of using large language models in higher education?

There are several reasons why this type of literature review is essential. A review study can offer useful insights into using LLMs for researchers and practitioners (Liu et al. 2023a). The first advantage is that it allows us to explore the diverse range of applications of LLMs, from automated assessment and natural language processing to personalised learning and content generation (Bauer et al. 2023). A second benefit of LLMs is that they enable the examination of the diverse contexts and populations in which they have been implemented, providing insight into their effectiveness and adaptability across various educational settings (Liu et al. 2023a; Rudolph et al. 2023b). Thirdly, through the systematic literature review, we can uncover potential challenges and limitations of LLM usage, such as ethical considerations, biases, and privacy and security concerns (Rasul et al. 2023). As a result of consolidating and analysing existing research, we identify trends and gaps in the literature regarding LLMs in higher education. In addition to contributing to educational research, this paper will guide educators, administrators, and policymakers in integrating LLMs into their educational practices. This systematic literature review aims to give future research endeavours a solid foundation and foster evidence-based discussions about leveraging large language models in higher education.

The need for this study emerged from the literature itself. Liu et al. (2023a) suggested that future research needs summarising the limitations and embrace the potential opportunities associated with LLMs like ChatGPT. Exploring the challenges and advantages of innovative applications of AI in education can revolutionise the future of education (Alqahtani et al. 2023). Sullivan et al. (2023) argued that it is an important debate that we need further research to explore the ways ChatGPT has been discussed in higher education. Rudolph et al. (2023b) also suggested a need for research on the effects of AI-based tools in learning and teaching. It is important to understand both the affordances and the challenges of LLMs and conversation style generative AI in the higher education context (Eager and Brunton 2023). Moreover, not much has been done in this space to explore the benefits and challenges from a systematic perspective (Rasul et al. 2023). Therefore, the need of this literature review emerges even more currently where educators and researchers are exploring ways of ethical incorporation of LLMs in their learning and teaching model.

The remaining paper is organised as follows; the next section presents a comprehensive literature review on the related concepts to this study. The following section then explains the literature review method. The subsequent two sections present the analysis of literature findings on the challenges and benefits of LLM in higher education. Then the next section entails implications and future research directions. Finally, the last section sheds light on this study's contribution and concluding remarks.

2   Literature review

The following sections provide a comprehensive literature review to understand LLMs and how they relate to higher education.

2.1 What are LLMs?

In artificial intelligence and natural language processing, Large Language Models (LLMs) can ingest massive amounts of text data and generate human-like responses (De Angelis et al. 2023). Using deep learning techniques and neural networks, these models learn to mimic syntactic patterns, semantics and seemingly grasp language context (Limna et al. 2023). By providing new avenues for automation like versatility and ability to understand languages, personalisation, and efficient data processing, LLMs such as GPT-3 (Generative Pre-trained Transformer 3) and BERT (Bidirectional Encoder Representations from Transformers) have revolutionised several industries (Perkins 2023). Through natural language processing, LLMs assist in medical record analysis, drug discovery, and patient care in healthcare (Iftikhar 2023; Tang et al. 2023; Trajanov et al. 2023). Their applications in finance include sentiment analysis, financial forecasting, and fraud detection (Yang et al. 2023). In the e-commerce and media industries, LLMs play a key role in customer service, content creation, and language translation (Wu et al. 2023a). Since LLMs have the potential to transform learning experiences and aid researchers in a wide range of domains, they have attracted significant interest in higher education and research (Rasul et al. 2023; Rudolph et al. 2023b). For example, students are drawn towards ChatGPT for paraphrasing, quick idea generation, syntactical programming errors and a lot more writing aids that doesn't require in depth understanding of the concept. LLMs can enhance teaching by enhancing content generation, providing instant feedback, and supporting personalised learning pathways (Tapalova and Zhiyenbayeva 2022). In academic literature analysis, LLMs are used to analyse vast amounts of textual data, analyse sentiments, and uncover patterns (Liao et al. 2023).

In late 2022, ChatGPT attracted substantial attention, marking an important turning point in LLM research (Cotton et al. 2023). Although researchers began to recognise LLM's potential before this, their implementation in education and research had not yet been fully explored (Yan et al. 2023). In light of the fact that educators and researchers are just now tapping into this technology, there is a need for comprehensive studies that assess its impact on student learning, academic writing, and knowledge dissemination (Sallam 2023). It is also essential to further investigate potential biases, data privacy issues, and ethical considerations surrounding using LLMs in academic environments (Perkins 2023). In examining the potential for LLMs in higher education and research, we find that more research is needed to fill the knowledge gaps. In the next section, we conducted a literature review on using LLMs in higher education to understand the gaps better.

2.2 Use of LLMs in higher education

Recently, Large Language Models (LLMs) have gained momentum in higher education, offering a range of opportunities to educators, researchers, and students. The potential impact of LLMs on teaching, learning, and research has been explored by scholars in various academic contexts (Limna et al. 2023; Perkins 2023; Rasul et al. 2023; Rudolph et al. 2023b).

Several educational institutions are investigating LLMs to enhance content creation, automated grading, and personalised learning. With LLMs' natural language processing capabilities, real-time feedback and adaptive learning pathways can be provided, enabling a more individualised approach to learning (Tapalova and Zhiyenbayeva 2022). Students are eager to engage in interactive discussions with teachers when using LLMs to generate teaching materials, create assessments, and generate assessment materials (Eager and Brunton 2023). Furthermore, LLMs provide researchers with valuable tools for analysing large amounts of academic literature, identifying trends, and developing insights for their research (Lund et al. 2023). Through LLM-powered virtual tutors, students benefit from personalised learning and immediate support (Jeon and Lee 2023).

Implementing and maintaining LLM-based solutions is one of the primary concerns for educational institutions (Milano et al. 2023). It requires significant computational resources and expertise to develop and refine LLM models, making it financially challenging (Liu et al. 2023a). A reliance on external platforms or proprietary LLMs may also raise privacy concerns and concerns about vendor lock-in (Fraiwan and Khasawneh 2023). It may also be difficult for educational institutions to integrate LLMs into existing curricula (Rudolph et al. 2023a), as customising and aligning with specific learning outcomes can take considerable time. Several researchers have raised concerns about bias related to training data in LLMs used for academic literature analysis (Sallam 2023). Due to the vast amount of text available on the internet, LLMs may perpetuate existing biases, resulting in biased results (Hosseini and Horbach 2023). Researchers are unable to understand the reasoning behind LLM's outputs due to the opaque nature of LLM decision-making, often referred to as the "black box" problem (Hosseini and Horbach 2023; Sallam 2023).

Furthermore, researchers are concerned that plagiarism detection systems based on LLMs may flag legitimate academic work as plagiarised due to similar phrasing (Ajevski et al. 2023). There has been concern expressed by teachers who use LLMs in their classrooms about the loss of human interaction and personalised feedback (Jeon and Lee 2023). It is also important for teachers to avoid overreliance on LLM-generated content (Abd-Alrazaq et al. 2023), since it may not meet the needs of their diverse student populations. Moreover, as LLM technology advances, previously generated content may become obsolete, requiring constant adaptations and updates (Meskó and Topol 2023).

Although LLMs in higher education have demonstrated their transformative potential, significant challenges and benefits remain associated with their implementation. Many studies look at specific applications of LLMs or their technical aspects but do not examine their impact from an overall perspective (Fan et al. 2023). More comprehensive studies are needed to address these gaps to explore the challenges educators and institutions face when integrating LLMs into education (Rasul et al. 2023). Further scrutiny is needed concerning data privacy, algorithmic biases, and academic integrity. A study should also examine whether LLMs are useful as virtual tutors and how they affect student motivation, academic performance, and critical thinking (Eager and Brunton 2023). By studying the experiences of educators and students in the adoption of LLMs in higher education, we can gain valuable insights into their long-term sustainability and effectiveness.

3   Systematic Review Method

A systematic literature review tests hypotheses, assesses prior studies' consistency, and summarises results (Higgins et al. 2019). It is an effective way to uncover the yet unexplored issues in the literature through a time-framed, process-oriented and effective review of previous work. Additionally, it ensures that all relevant literature for a domain is considered. A systematic literature review is the most appropriate approach to understanding the current trends, challenges and benefits and future research (Chhina et al. 2019), on the use of LLMs in higher education.

Our approach was a combination of two approaches. The review uses a step-by-step methodology to produce reproducible and systematic results (Okoli and Schabram 2010) and a hermeneutic framework (Boell and Cecez-Kecmanovic 2014). With this iterative approach, we could expand pre-existing ideas and concepts within the literature and understand them better. To begin, you must clearly define the purpose of the review, which includes identifying its specific objectives and goals (Okoli and Schabram 2010). In this study, we identified the purpose to explore the recent trends in the challenges and benefits of using LLMs in higher education. By identifying themes in the literature, we aimed to establish a framework for classifying these challenges, benefits and research directions. The second step involves developing a research protocol and training participants. An outline of the literature review process is included in the protocol (Keele 2007). Researchers formulate research questions during this stage, the review protocol is developed, and reviewers are trained to ensure consistency and rigor (Okoli and Schabram 2010). Our plan examined prominent databases such as Scopus (Elsevier), ScienceDirect, Web of Science, and Springer and Google Scholar. Our next step was to create a roadmap for the review process. The third and fourth steps involved the framework, which included literature search and practical screening to select literature (Okoli and Schabram 2010). We explored a number of frameworks for conducting systematic literature reviews in this study and identified the hermeneutic framework, which is rooted in hermeneutic philosophy and serves as a powerful theoretical basis for explaining and understanding the review process (Boell and Cecez-Kecmanovic 2014). Even though a hermeneutic framework includes two circles—an inner circle for searching and acquiring and a wider one for analysing and interpreting—for this study, only the search and acquisition circle was used, as it enables the identification of relevant literature rigorously.

Literature acquisition and analysis become increasingly important as we gather more data (Boell and Cecez-Kecmanovic 2014). To capture the latest trends in the literature, we searched the aforementioned major databases from 2018 to 2023. However, as the topic gained attention in late 2022, most of the research included in this study is from 2023, as shown in Figure 1. Figure 2 below describes the complete selection process adopted from PRISMA. In this systematic review, we closely adhered to PRISMA guidelines, the most common peer-reviewed methodology (Page et al. 2021). Forty-one results from the selected databases and Google Scholar talk about using and implementing LLMs in the higher education industry. We refined our search strategy and continued the hermeneutic circle for identifying new literature sources of potential interest (Boell and Cecez-Kecmanovic 2014). However, as this area is quite recent, we could not locate more papers. A quality assessment of this type of literature is crucial, and the researchers should cite the source of the work (Rothstein and Hopewell 2009). For this study, the researchers ensured that all the source literature is appropriately cited with complete information. Using this approach, we were able to focus on selecting the best literature to use for our study. The next step

was to qualitatively extract the data to identify the challenges and benefits related to LLMs use and application in higher education. The following section of the paper presents a framework for classifying the challenges and benefits based on thematic patterns identified in the literature.
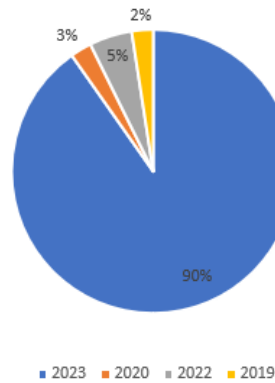

Figure 1: Distribution of the reviewed studies over the years from 2019- to date
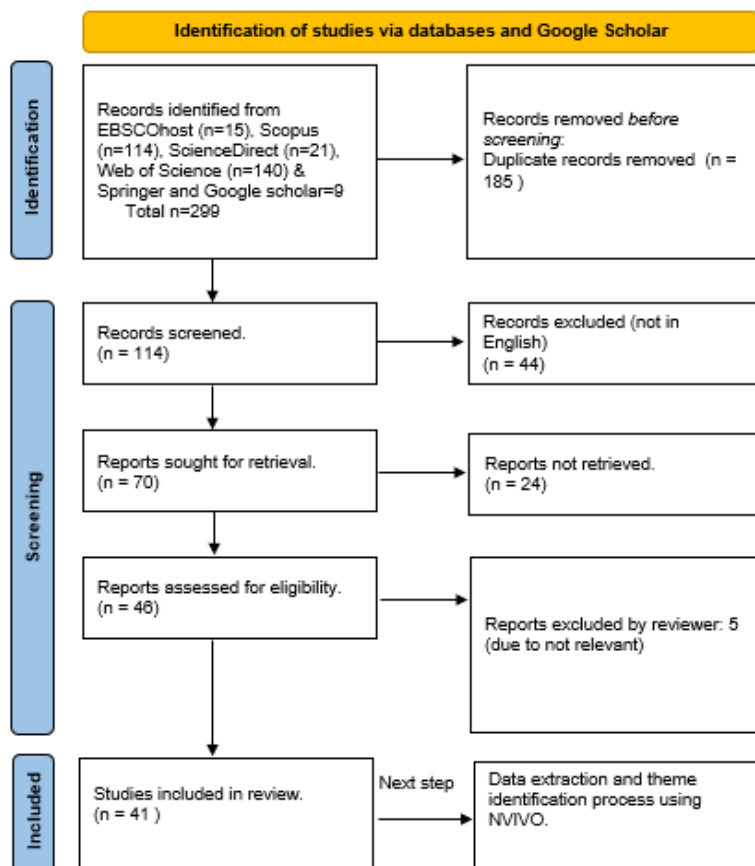

Figure 2: Paper selection process for the review. Framework adapted and modified from PRISMA (PRISMA 2020)

The final 41 studies were imported into NVivo for thematic analysis. NVivo is a renowned tool for qualitative data analysis. The idea was to assess these studies and identify the emerging themes for challenges, benefits and suggested opportunities for using LLM in higher education. The next section elaborates on the identified themes and the analysis of the findings.

4   Findings and Discussion

It has been both a source of enthusiasm and apprehension to introduce Large Language Models (LLMs) to higher education. Despite the potential for LLMs to revolutionise teaching and learning, educational institutions also face many challenges that must be considered carefully. Here, we summarise the key findings about the benefits and challenges of using LLMs in higher education, shedding light on critical aspects for educators, researchers, and administrators to consider.

4.1   Challenges of using LLM in higher education

Below is the list of challenges accumulated from the analysis of the challenges described by the gathered studies for this systematic literature review. Table I in Appendix A shows the distribution of these challenges in the studies.

1. **Danger to Authentic Learning:** An authentic learning experience is based on principles that can be used to guide institutions in designing curriculum that aligns student learning experiences with the world in which they will live and work (McKenzie et al. 2002). There is a risk of authentic learning experiences being undermined by the over-reliance on content generated by LLMs (Malik et al. 2023). As a result, students may become dependent on AI-generated answers, which hinders their ability to think critically, solve problems, and be creative. By relying heavily on AI-generated responses, we risk compromising authentic learning based on real-world problems and deeper understanding (Chan and Tsi 2023). Despite their efficiency, AI-generated responses may lack a deeper understanding of underlying principles (Dergaa et al. 2023). It can result in a lack of in-depth knowledge and critical insights, both critical for success in academic and real-world settings. A shift in the role of educators in undergraduate education may be brought about by the increased use of LLMs for content generation and automated feedback (Rudolph et al. 2023b). Despite AI's value, it should not replace human instructors but complement them (Rasul et al. 2023). Reducing instructor involvement may negatively impact students' ability to learn and interact meaningfully.

2. **Academic misconduct:** Misconduct can arise from a number of causes. As LLM-generated content is anonymous, students may mistake it for their work, resulting in the unintentional misappropriation of LLM-generated content as their own. (Liu et al. 2023c). The anonymity associated with the content also increases the likelihood of them submitting AI-generated responses without acknowledging their source (Susnjak 2022) and without attribution (Perkins 2023). Easy access to these systems and the increasing demands for high-quality work may also contribute to students setting aside ethical concerns and turning to AI-generated content to help with assessments (Cotton et al. 2023).

    Using LLMs extensively for content generation, assessment, and feedback undermines academic integrity (Perkins 2023), while also leaving educators unable to accurately gauge students' abilities, and eroding trust in assessment and evaluation processes (Rasul et al. 2023). Institutions of higher learning rely on honesty, fairness, and accountability and seek to fairly assess student's learning through assessments. Students that have misused generated content may be perceived as more proficient than they really are, thus resulting in inaccurate assessments of their abilities. This not degrades trust in academic institutions, but also negatively impacts the reputations of both the student and the institution.

    To address the ethical implications of AI integration, institutions must have robust policies and frameworks for academic integrity. Students should be educated about the importance of originality in academic work by institutions with clear policies against plagiarism. Students should also be challenged to demonstrate their understanding of subject matter through assessments and assignments that promote critical thinking.

3. **Unreliable and biased:** While LLMs are powerful, they can still produce inaccurate or entirely falsified results (Van Dis et al. 2023). Educators must verify AI-generated content before using it in their classrooms to prevent misinformation. AI contributions must be properly cited and acknowledged to maintain academic honesty and give credit where credit is due (Cotton et al. 2023). It is common for LLMs to be trained using large datasets that include biased and inaccurate information. For educators, monitoring AI-generated content is crucial for preventing false or biased information from reaching students. The large datasets used to train LLMs risk inadvertently producing references to non-existent sources or fabricated authors (Perkins et al. 2023). As a result, research papers and scholarly works seriously threaten their credibility and academic integrity (Malik et al. 2023). To ensure their accuracy and legitimacy,

all users – educators, researchers and students – should manually verify LLM-generated references. Preventing unintentional inclusion of fictitious information or references requires educators and researchers to adopt rigorous fact-checking procedures.

4. **Ethical and Equity Practices:** Introducing LLMs to higher education raises ethical concerns over data privacy, consent, and AI models potentially reinforcing biases (Wu et al. 2023b). Educators must be aware of equity concerns to ensure that all students have access to LLM resources. For training, LLMs utilise a vast amount of data, including text from various sources, including the Internet (Peng et al. 2023). The data must be handled with the utmost care. To ensure adequate protection of student data when integrating LLMs into education settings. Students need to be informed about how their data will be used by the service providers behind the LLMs with clear policies regarding data privacy. The technology and internet connections necessary to interact with LLMs may not be available to all students (Susnjak 2022). Educators must address these equity concerns to ensure that all students have equal access to LLM resources.

5. **Limited Evaluation of Graduate Skill Sets:** There is no comprehensive assessment or evaluation of graduate skills or requirements through an LLM program (Susnjak 2022). It is important for higher-level learning outcomes not to be solely dependent on AI-based assessments. LLMs can assist with some aspects of assessment, such as providing information and feedback, but they may not fully capture the complexity and depth of higher-level learning outcomes (Farrokhnia et al. 2023). A comprehensive evaluation of graduate skills requires a comprehensive assessment approach that incorporates a variety of assessment methods, including projects, presentations, research papers, and practical applications.

6. **Lack of Social Interaction:** AI-powered learning experiences may lack social interaction, which is crucial for students' emotional and social development (Rasul et al. 2023). Constructing meaning through reflection on experiences through collaboration and peer-to-peer interaction is vital. By collaborating and interacting with peers, students can discuss ideas, exchange perspectives, and construct meaning together. The social interactions enhance the learning process, providing multiple dimensions that AI alone may not be able to provide (Markauskaite et al. 2022).

LLMs in higher education present challenges that require thoughtful consideration and mitigation strategies. It is a concern that over-reliance on LLM-generated content may hinder students' ability to think critically and solve problems. A further threat to academic integrity arises when anonymous material generated by LLM is submitted as student work, resulting in academic misconduct. Undergraduate teaching can be adversely affected by the increased use of LLMs due to the reduced role of human instructors. Meaningful interactions between students and teachers may therefore be negatively affected. The results of LLMs may be unreliable or biased, which requires careful verification to avoid misinformation. Access to AI-generated answers may lead to academic dishonesty among students due to pressure to cheat. Data privacy, consent, and biases must be considered when integrating LLMs. A robust policy and framework are necessary to maintain academic integrity.

Further, AI-powered learning experiences lack social interaction and comprehensive evaluation of graduate skill sets, which poses challenges to students. Furthermore, AI-generated content may include fictitious references, threatening research papers' credibility and authenticity. To address these challenges, educators, institutions, and researchers must collaborate to maximize LLM benefits while mitigating their drawbacks.

4.2 Benefits of using LLM in higher education

Below is a list of benefits described by the studies gathered for this review. The distribution of these benefits is presented in the Table II in Appendix B.

1. **Easy to Use and Explain Confusing Ideas:** Students can better comprehend difficult topics by simplifying complex concepts through LLMs, such as ChatGPT (Chan and Tsi 2023). Educators also well, can use it to explain confusing concepts in a way that is more accessible. For example, ChatGPT, an Open AI LLM widely used by students, has proven effective in helping them grasp complex subjects. For example, on feeding the contents of this paper on ChatGPT, we asked it to suggest a title, and reading the whole text at once and analysing, the tool suggested the given title for this study. Students' understanding of complex ideas is enhanced, and their engagement with the material is heightened when LLMs like ChatGPT simplify complex ideas and provide clear explanations (Sallam et al. 2023). Further, LLMs allow students to interact seamlessly with technology, leading to an improved learning experience. A LLM such as

ChatGPT enhances students' creativity by helping them brainstorm ideas and organise their thoughts (Qadir 2023). For brainstorming with LLMs, students can present topics or ideas, and the model responds with related concepts, insights, and possible angles (Atlas 2023). By exploring multiple perspectives and possibilities, students can generate new ideas they might not have thought of otherwise. While LLMs can undoubtedly stimulate creativity, students' originality and individuality are essential to the creative process, and LLMs should not replace them.

2. **Smooth Out Awkward First Drafts:** LLMs assist students with writing by suggesting corrections, refinements, and suggestions (Sallam 2023). Learners can write more effectively for academic purposes by refining language and structure. By using LLMs for writing, students can improve the clarity and coherence of their academic writing by using the system's corrections and refinements (Perkins 2023). ChatGPT, for instance, can provide insight into grammar and syntax, suggest appropriate vocabulary choices, and suggest better ways to organise research papers (Huang and Tan 2023). In academic writing, LLMs can serve as the "basic building blocks" (Van Dis et al. 2023). As a result of providing accurate information and structuring content well, LLMs enable students to develop foundational writing skills (Dergaa et al. 2023). Using these models, students can strengthen their writing skills and proficiency in academic writing (Sullivan et al. 2023). Researchers can use LLMs as a starting point for further research by simply providing a prompt or question (Malik et al. 2023). It is common for language barriers to impede cross-cultural collaboration and knowledge dissemination. As a result of LLMs, students and researchers have been able to communicate and collaborate across linguistic barriers more easily (Latif et al. 2023). An LLM can also improve the clarity and coherence of written pieces by suggesting alternative sentence structures or synonyms (Atlas 2023). Thus, students have more time to conduct in-depth research, engage in critical thinking, and explore their academic subjects creatively when they spend less time on repetitive tasks like proofreading and grammar checks.

3. **Preparation for AI Integration:** Students exposed to LLMs in higher education are prepared for a future with widespread AI integration (Sallam et al. 2023). Digitally advanced workforces require students to understand AI technologies. Increasingly, technology-advanced workforces integrate AI into their processes (Rudolph et al. 2023b). LLMs prepare students for the technology-integrated workplaces, where took such as LLMs are in use (Krause 2023). Students studying LLMs, such as ChatGPT, have access to AI-driven language processing capabilities, and are familiar with the advantages and pitfalls of these systems. As these tools cannot read mind and lack the comprehensive understanding, using creative text prompts can produce the outcome that the user is looking for in a novel but in a less complex way (Haleem et al. 2022). These creative ways of text prompts should be incorporated in the curriculum. As a result of this exposure, students gain a better understanding of how AI-powered systems can understand, process, and generate human language and their limitations.

4. **Enhanced Learning through Personalisation:** The LLM can provide students with learning experiences tailored to their needs (Abd-Alrazaq et al. 2023). A more personalised and effective learning experience can be achieved by providing learners with customised content and feedback. It can be challenging for educators to cater to the individual needs of each student in traditional classroom settings due to the lack of time and resources. LLMs can create personalised learning materials, exercises, and quizzes based on the students' learning pace, prior knowledge, and interests (Kasneci et al. 2023). Even though these benefits are promising, it is important to balance personalisation and human interaction.

5. **Individualised Feedback:** In addition to providing instant feedback on assignments, LLMs can enhance students' learning outcomes by identifying areas for improvement (Alqahtani et al. 2023). One of the key advantages of using LLMs for this purpose is the ability to detect patterns and trends in students' writing and learning patterns (Liu et al. 2023a). Students can improve their reasoning and analytical abilities through LLMs' analysis of students' arguments and responses. It is extremely valuable to receive feedback on higher order thinking skills to foster a deeper understanding and growth in intellectual capacity.

6. **Automated Administrative Support:** By focusing on teaching and researching rather than administrative tasks, LLMs allow educators to make a significant contribution to society (Alqahtani et al. 2023). Teachers can also focus on creating innovative teaching methods and curricula by leveraging LLMs for administrative support (Kasneci et al. 2023). As a result, they can devote more time to stimulating critical thinking, fostering engagement, and enhancing

deeper understanding. Changing the focus can result in more meaningful and impactful student interactions, nurturing a passion for learning and enhancing subject comprehension.

7. **Innovative Assessment Activities:** Educators can use LLMs to design innovative assessment activities that promote students' critical thinking and problem-solving skills (Rasul et al. 2023). The LLM can provide immediate feedback to students as they respond to questions, guiding them towards the correct approach or encouraging them to explore alternative perspectives (Uchiyama et al. 2023). For example, educators can design virtual scenarios in which students must make decisions based on the information presented to them.

It has been shown that Large Language Models (LLMs) enhance student and teacher learning experiences. With LLMs, such as ChatGPT, students can better grasp complicated concepts, while educators can make confusing ideas more accessible. Additionally, LLMs provide students with suggestions and corrections, which helps them improve their writing skills. An LLM helps students develop essential digital literacy skills necessary in a future where AI is widely integrated. As foundational building blocks for academic writing, these models provide accurate information and improve content structuring. By brainstorming with ChatGPT, students are also able to enhance their creativity. In addition to producing higher-quality work in less time, LLMs allow students to focus on other aspects of their education. Individualising learning experiences and providing feedback enriches learning and improves learning outcomes. Through various research and writing tasks, LLMs assist students in developing critical thinking skills and problem-solving abilities. Higher education institutions can maximise the potential of AI-powered learning by utilising these benefits to create an adaptive and enriching learning environment for students.

5   Implications and future research directions

The implications and future research direction of integrating Language Model Models (LLMs) in higher education are multifaceted and call for a comprehensive approach to address the challenges and opportunities identified by this study. Following are the potential implications and suggested future research directions:

- Mitigation Strategies: To address LLMs' challenges in higher education, educational institutions should devise thoughtful mitigation strategies. As part of these strategies, guidelines and policies could be developed on the appropriate use of LLM-generated content, critical thinking exercises could be encouraged alongside LLM usage, and systems to detect academic misconduct could be implemented. We suggest future research on the framework development to address academic misconduct concerns raised by LLMs in education.

- Faculty Training and Development: Faculty members and educators must receive training and professional development on effectively integrating LLMs into their teaching practices. The goal of training should be to use LLMs to complement human instruction, not to replace it, and to maintain the integrity of the learning process. It is suggested to have some research attempts towards the training and development of educators for incorporating LLMs in their pedagogies.

- Ethical Considerations: Researchers need to conduct in-depth studies to determine whether LLM integration in higher education is ethical. Several issues should be explored, including data privacy concerns, the impact of AI on academic integrity, and potential biases in LLM-generated content.

- Social Interaction and Learning Outcomes: Further research should explore how AI-powered learning experiences affect social interaction and emotional development. It is crucial to understand how LLMs affect peer-to-peer interactions and collaboration to optimise the learning environment,

- Accuracy and Reliability: To minimise the risk of misinformation, researchers can enhance the verification process of AI-generated information. More research is needed to improve the accuracy and reliability of LLM-generated content.

- Evaluating LLM-Based Assessments: To understand the effectiveness and limitations of using LLMs to evaluate student performance, researchers should compare the results of AI-based assessments with those of traditional assessments.

- Cross-Disciplinary Impact: Examine the impact of LLMs across various academic disciplines. Understanding the nuances of LLM integration is crucial for successful implementation across different subjects.

Future studies should examine mitigation strategies, faculty training, ethical considerations to ensure successful integration, and the impact of LLMs on social interaction and skill development. It is important to seek ways to enhance the accuracy and reliability of LLM-generated content and assess student engagement, satisfaction, and long-term outcomes. As higher education institutions address these implications and focus on continuous improvement, they can create an adaptive, enriching learning environment that prepares students for an AI-enabled world.

6 Contribution and Conclusion

In this systematic literature review, we sought to shed light on critical aspects educators, researchers, and administrators must consider when implementing LLMs. The findings of this study provide valuable insights for optimising LLM use and contribute to growing knowledge on AI integration in education. It highlights several challenges associated with integrating LLMs into higher education, such as the risk of academic misconduct, the reduced role of instructors, unreliable and biased results, and the possibility of academic dishonesty. Privacy issues, consent issues, and possible biases in AI models also emerged as major challenges. Furthermore, LLM-generated content appears to contain fictitious references and lacks social interaction, which indicates deficiencies in evaluating graduate skill sets. However, transformative technology also holds several benefits for academia which cannot be ignored. Developing digital literacy skills is essential for a future where artificial intelligence is widely used. Individualising learning experiences and providing feedback enrich learning and enhance learning outcomes. LLMs also hold great potential for educators to incorporate innovative assessments using them.

The study proposes several research directions to address the challenges summarised in the literature so that the benefits cannot be overlooked. Universities should devise mitigation strategies, such as establishing guidelines and policies on LLM use, encouraging critical thinking alongside LLM-produced content, and creating systems to detect academic misconduct. The second factor is faculty development and training, which are essential to successfully integrating an LLM. LLM integration should be examined in depth in terms of ethical considerations. Research can focus on improving the verification process of AI-generated information to enhance the accuracy and reliability of LLM-generated content. The implications and suggestions presented in this study can help higher education institutions create engaging, adaptive, and enriching learning environments that prepare students for AI-enabled careers. For future generations to benefit from the potential of LLMs for revolutionising teaching and learning, educators, researchers, and administrators will have to collaborate to maximise AI's potential in education. Although every possible measure has been taken to consider every relevant study, the study does not ignore the possibility of human error. Further research is to overcome this limitation by including more studies in future.

7 References

Abd-Alrazaq, A., AlSaad, R., Alhuwail, D., Ahmed, A., Healy, P. M., Latifi, S., Aziz, S., Damseh, R., Alrazak, S. A., and Sheikh, J. 2023. "Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions," *JMIR Medical Education* (9:1), p. e48291. doi: https://doi.org/10.2196/48291

Ajevski, M., Barker, K., Gilbert, A., Hardie, L., and Ryan, F. 2023. "ChatGPT and the future of legal education and practice," *The Law Teacher*), pp. 1-13. doi: https://doi.org/10.1080/03069400.2023.2207426

Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., Alowais, S. A., Alshaya, O. A., Rahman, I., and Al Yami, M. S. 2023. "The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research," *Research in Social and Administrative Pharmacy*. doi: https://doi.org/10.1016/j.sapharm.2023.05.016

Álvarez-Álvarez, C., and Falcon, S. 2023. "Students' preferences with university teaching practices: analysis of testimonials with artificial intelligence," *Educational Technology Research and Development*, pp. 1-16. doi: https://doi.org/10.1007/s11423-023-10239-8

Atlas, S. 2023. "ChatGPT for higher education and professional development: A guide to conversational AI,". https://digitalcommons.uri.edu/cba_facpubs/548

Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., Fischer, M. R., and Fischer, F. 2023. "Using natural language processing to support peer-feedback in the age of artificial intelligence: A cross-disciplinary framework and a research agenda," *British Journal of Educational Technology*. doi: https://doi.org/10.1111/bjet.13336

Bhayana, R., Krishna, S., and Bleakney, R. R. 2023. "Performance of ChatGPT on a radiology board-style examination: Insights into current strengths and limitations," *Radiology*, p. 230582. doi: https://doi.org/10.1148/radiol.230582

Boell, S. K., and Cecez-Kecmanovic, D. 2014. "A hermeneutic approach for conducting literature reviews and literature searches," *Communications of the Association for Information Systems* (34:1), p. 12. doi: https://doi.org/10.17705/1CAIS.03412

Chan, C. K. Y., and Tsi, L. H. 2023. "The AI Revolution in Education: Will AI Replace or Assist Teachers in Higher Education?," *arXiv preprint arXiv:2305.01185.* doi: https://doi.org/10.48550/arXiv.2305.01185

Chhina, S., Chadhar, M., Vatanasakdakul, S., and Chetty, M. 2019. "Challenges and opportunities for Blockchain Technology adoption: A systematic review," *Proceedings of the Australasian Conference on Information Systems, Perth, Australia*, pp. 9-11. https://aisel.aisnet.org/acis2019/81

Cochran, K., Cohn, C., Rouet, J. F., and Hastings, P. 2023. "Improving Automated Evaluation of Student Text Responses Using GPT-3.5 for Text Data Augmentation," *International Conference on Artificial Intelligence in Education*: Springer, pp. 217-228. doi: https://doi.org/10.1007/978-3-031-36272-9_18

Cotton, D. R., Cotton, P. A., and Shipway, J. R. 2023. "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*), pp. 1-12. doi: https://doi.org/10.1080/14703297.2023.2190148

Crawford, J., Cowling, M., and Allen, K.-A. 2023a. "Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI)," *Journal of University Teaching & Learning Practice* (20:3), p. 02. doi: https://doi.org/10.53761/1.20.3.02

Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J.-A., Middleton, R., and Wilson, G. S. 2023b. "Artificial Intelligence and Authorship Editor Policy: ChatGPT, Bard Bing AI, and beyond," *Journal of University Teaching & Learning Practice* (20:5), p. 1. doi: https://doi.org/10.53761/1.20.5.01

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., and Rizzo, C. 2023. "ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health," *Frontiers in Public Health* (11), p. 1166120. doi: https://doi.org/10.3389/fpubh.2023.1166120

Dergaa, I., Chamari, K., Zmijewski, P., and Saad, H. B. 2023. "From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing," *Biology of Sport* (40:2), pp. 615-622. doi: https://doi.org/10.5114/biolsport.2023.125623

Dobslaw, F., and Bergh, P. 2023. "Experiences with Remote Examination Formats in Light of GPT-4," *arXiv preprint arXiv:2305.02198*). doi: https://doi.org/10.48550/arXiv.2305.02198

Eager, B., and Brunton, R. 2023. "Prompting higher education towards AI-augmented teaching and learning practice," *Journal of University Teaching & Learning Practice* (20:5), p. 02. doi: https://doi.org/10.53761/1.20.5.02

Fan, L., Li, L., Ma, Z., Lee, S., Yu, H., and Hemphill, L. 2023. "A bibliometric review of large language models research from 2017 to 2023," *arXiv preprint arXiv:2304.02020*. doi: https://doi.org/10.48550/arXiv.2304.02020

Farrokhnia, M., Banihashem, S. K., Noroozi, O., and Wals, A. 2023. "A SWOT analysis of ChatGPT: Implications for educational practice and research," *Innovations in Education and Teaching International*, pp. 1-15. doi: https://doi.org/10.1080/14703297.2023.2195846

Fraiwan, M., and Khasawneh, N. 2023. "A Review of ChatGPT Applications in Education, Marketing, Software Engineering, and Healthcare: Benefits, Drawbacks, and Research Directions," *arXiv preprint arXiv:2305.00237*. doi: https://doi.org/10.48550/arXiv.2305.00237

Giannos, P. 2023. "Evaluating the limits of AI in medical specialisation: ChatGPT's performance on the UK Neurology Specialty Certificate Examination," *BMJ Neurology Open* (5:1). doi: https://doi.org/10.1136%2Fbmjno-2023-000451

Haleem, A., Javaid, M., and Singh, R. P. 2022. "An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* (2:4), p. 100089. doi: https://doi.org/10.1016/j.tbench.2023.100089

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., and Welch, V. A. 2019. *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons. doi: https://doi.org/10.1002/9780470712184

Hosseini, M., and Horbach, S. P. 2023. "Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other Large Language Models in scholarly peer review," *Research Integrity and Peer Review* (8:1), p. 4. doi: https://doi.org/10.1186/s41073-023-00133-5

Huang, J., and Tan, M. 2023. "The role of ChatGPT in scientific communication: writing better scientific review articles," *American Journal of Cancer Research* (13:4), p. 1148. PMCID: PMC10164801

Iftikhar, L. 2023. "Docgpt: Impact of chatgpt-3 on health services as a virtual doctor," *EC Paediatrics* (12:1), pp. 45-55. Available here.

Jalil, S., Rafi, S., LaToza, T. D., Moran, K., and Lam, W. 2023. "Chatgpt and software testing education: Promises & perils," *2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*: IEEE, pp. 4130-4137. doi: https://doi.org/10.1109/ICSTW58534.2023.00078.

Jeon, J., and Lee, S. 2023. "Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT," *Education and Information Technologies*), pp. 1-20. doi: https://doi.org/10.1007/s10639-023-11834-1

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., and Hüllermeier, E. 2023. "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences* (103), p. 102274. doi: https://doi.org/10.1016/j.lindif.2023.102274

Keele, S. 2007. "Guidelines for performing systematic literature reviews in software engineering." Technical report, ver. 2.3 ebse technical report. ebse. Available here.

Koutcheme, C. 2023. "Training Language Models for Programming Feedback Using Automated Repair Tools," *International Conference on Artificial Intelligence in Education*: Springer, pp. 830-835. doi: https://doi.org/10.1007/978-3-031-36272-9_79

Krause, D. 2023. "ChatGPT and Generative AI: The New Barbarians at the Gate," *Available at SSRN 4447526*. doi: https://dx.doi.org/10.2139/ssrn.4447526

Latif, E., Mai, G., Nyaaba, M., Wu, X., Liu, N., Lu, G., Li, S., Liu, T., and Zhai, X. 2023. "Artificial general intelligence (AGI) for education," *arXiv preprint arXiv:2304.12479*. doi: https://doi.org/10.48550/arXiv.2304.12479

Leinonen, J., Denny, P., MacNeil, S., Sarsa, S., Bernstein, S., Kim, J., Tran, A., and Hellas, A. 2023. "Comparing code explanations created by students and large language models," *arXiv preprint arXiv:2304.03938*. doi: https://doi.org/10.48550/arXiv.2304.03938

Liao, W., Liu, Z., Dai, H., Xu, S., Wu, Z., Zhang, Y., Huang, X., Zhu, D., Cai, H., and Liu, T. 2023. "Differentiate chatgpt-generated and human-written medical texts," *arXiv preprint arXiv:2304.11567*. doi: https://doi.org/10.48550/arXiv.2304.11567

Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., and Chocksathaporn, P. 2023. "The use of ChatGPT in the digital era: Perspectives on chatbot implementation," *Journal of Applied Learning and Teaching* (6:1). doi: https://doi.org/10.37074/jalt.2023.6.1.32

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., and Liu, Z. 2023a. "Summary of chatgpt/gpt-4 research and perspective towards the future of large language models," *arXiv preprint arXiv:2304.01852*. doi: https://doi.org/10.1016/j.metrad.2023.100017

Liu, Z., He, X., Liu, L., Liu, T., and Zhai, X. 2023b. "Context matters: A strategy to pre-train language model for science education," *arXiv preprint arXiv:2301.12031*. doi: https://doi.org/10.48550/arXiv.2301.12031

Liu, Z., Yao, Z., Li, F., and Luo, B. 2023c. "Check Me If You Can: Detecting ChatGPT-Generated Academic Writing using CheckGPT," *arXiv preprint arXiv:2306.05524*. doi: https://doi.org/10.48550/arXiv.2306.05524

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., and Wang, Z. 2023. "ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing," *Journal of the Association for Information Science and Technology* (74:5), pp. 570-581. doi: https://doi.org/10.1002/asi.24750

Malik, A., Khan, M. L., and Hussain, K. 2023. "How is ChatGPT transforming academia? Examining its impact on teaching, research, assessment, and learning," *Examining its Impact on Teaching, Research, Assessment, and Learning (April 9, 2023)*. doi: https://dx.doi.org/10.2139/ssrn.4413516

Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Shum, S. B., and Gašević, D. 2022. "Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI?," *Computers and Education: Artificial Intelligence* (3), p. 100056. doi: https://doi.org/10.1016/j.caeai.2022.100056

Marron, L. 2023. "Exploring the Potential of ChatGPT 3.5 in Higher Education: Benefits, Limitations, and Academic Integrity," in *Handbook of Research on Redesigning Teaching, Learning, and Assessment in the Digital Era*. IGI Global, pp. 326-349.

McKenzie, A. D., Morgan, C. K., Cochrane, K. W., Watson, G. K., and Roberts, D. W. 2002. "Authentic learning: What is it, and what are the ideal curriculum conditions to cultivate it in," *Quality Conversations: Proceedings of the 25th HERDSA Annual Conference, Perth, Western Australia*: Citeseer, pp. 426-433. doi: https://doi.org/10.4018/978-1-6684-8292-6.ch017

Meskó, B., and Topol, E. J. 2023. "The imperative for regulatory oversight of large language models (or generative AI) in healthcare," *NPJ Digital Medicine* (6:1), p. 120. doi: https://doi.org/10.1038/s41746-023-00873-0

Milano, S., McGrane, J. A., and Leonelli, S. 2023. "Large language models challenge the future of higher education," *Nature Machine Intelligence* (5:4), pp. 333-334. doi: https://doi.org/10.1038/s42256-023-00644-2

Mills, A., Bali, M., and Eaton, L. 2023. "How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process," *Journal of Applied Learning and Teaching* (6:1). doi: https://doi.org/10.37074/jalt.2023.6.1.34

Ndukwe, I. G., Amadi, C. E., Nkomo, L. M., and Daniel, B. K. 2020. "Automatic grading system using sentence-BERT network," *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II 21*: Springer, pp. 224-227. doi: https://doi.org/10.1007/978-3-030-52240-7_41

Okoli, C., and Schabram, K. 2010. "A guide to conducting a systematic literature review of information systems research,". http://sprouts.aisnet.org/10-26

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., and Brennan, S. E. 2021. "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews," *bmj* (372). doi: https://doi.org/10.1136/bmj.n160

Peng, B., Galley, M., He, P., Cheng, H., Xie, Y., Hu, Y., Huang, Q., Liden, L., Yu, Z., and Chen, W. 2023. "Check your facts and try again: Improving large language models with external knowledge and automated feedback," *arXiv preprint arXiv:2302.12813*. doi: https://doi.org/10.48550/arXiv.2302.12813

Perkins, M. 2023. "Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond," *Journal of University Teaching & Learning Practice* (20:2), p. 07. doi: https://doi.org/10.53761/1.20.02.07

Perkins, M., Roe, J., Postma, D., McGaughran, J., and Hickerson, D. 2023. "Game of Tones: Faculty detection of GPT-4 generated content in university assessments," *arXiv preprint arXiv:2305.18081*. doi: https://doi.org/10.48550/arXiv.2305.18081

PRISMA. 2020. " PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only."

Qadir, J. 2023. "Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education," *2023 IEEE Global Engineering Education Conference (EDUCON)*: IEEE, pp. 1-9. doi: https://doi.org/10.1109/EDUCON54358.2023.10125121.

Rasheed, H. A., Zenkert, J., Weber, C., and Fathi, M. 2019. "Conversational chatbot system for student support in administrative exam information," *ICERI2019 Proceedings*: IATED, pp. 8294-8301. doi: https://doi.org/10.21125/iceri.2019.1974

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., and Heathcote, L. 2023. "The role of ChatGPT in higher education: Benefits, challenges, and future research directions," *Journal of Applied Learning and Teaching* (6:1). Available here.

Rothstein, H. R., and Hopewell, S. 2009. "Grey literature," *The handbook of research synthesis and meta-analysis* (2), pp. 103-125. Available here.

Rudolph, J., Tan, S., and Tan, S. 2023a. "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?," *Journal of Applied Learning and Teaching* (6:1). Available here.

Rudolph, J., Tan, S., and Tan, S. 2023b. "War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education," *Journal of Applied Learning and Teaching* (6:1). Available here.

Sallam, M. 2023. "The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations," *medRxiv*), p. 2023.2002. 2019.23286155. doi: https://doi.org/10.1101/2023.02.19.23286155

Sallam, M., Salim, N., Barakat, M., and Al-Tammemi, A. 2023. "ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations," *Narra J* (3:1), pp. e103-e103. doi: https://doi.org/10.52225/narra.v3i1.103

Sullivan, M., Kelly, A., and McLaughlan, P. 2023. "ChatGPT in higher education: Considerations for academic integrity and student learning,". doi: https://doi.org/10.37074/jalt.2023.6.1.17

Susnjak, T. 2022. "ChatGPT: The end of online exam integrity?," *arXiv preprint arXiv:2212.09292*. doi: https://doi.org/10.48550/arXiv.2212.09292

Tang, R., Han, X., Jiang, X., and Hu, X. 2023. "Does synthetic data generation of llms help clinical text mining?," *arXiv preprint arXiv:2303.04360*. doi: https://doi.org/10.48550/arXiv.2303.04360

Tapalova, O., and Zhiyenbayeva, N. 2022. "Artificial Intelligence in Education: AIEd for Personalised Learning Pathways," *Electronic Journal of e-Learning* (20:5), pp. 639-653. Available here.

Teel, Z. A., Wang, T., and Lund, B. 2023. "ChatGPT conundrums: Probing plagiarism and parroting problems in higher education practices," *College & Research Libraries News* (84:6), p. 205. doi: https://doi.org/10.5860/crln.84.6.205

Trajanov, D., Trajkovski, V., Dimitrieva, M., Dobreva, J., Jovanovik, M., Klemen, M., Žagar, A., and Robnik-Šikonja, M. 2023. "Review of Natural Language Processing in Pharmacology," *Pharmacological Reviews* (75:4), pp. 714-738. doi: https://doi.org/10.1124/pharmrev.122.000715

Uchiyama, S., Umemura, K., and Morita, Y. 2023. "Large Language Model-based System to Provide Immediate Feedback to Students in Flipped Classroom Preparation Learning," *arXiv preprint arXiv:2307.11388*. doi: https://doi.org/10.48550/arXiv.2307.11388

Van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., and Bockting, C. L. 2023. "ChatGPT: five priorities for research," *Nature* (614:7947), pp. 224-226. doi: *https://doi.org/10.1038/d41586-023-00288-7*

Wu, J., Gan, W., Chen, Z., Wan, S., and Lin, H. 2023a. "Ai-generated content (aigc): A survey," *arXiv preprint arXiv:2304.06632*. doi: https://doi.org/10.48550/arXiv.2304.06632

Wu, X., Duan, R., and Ni, J. 2023b. "Unveiling Security, Privacy, and Ethical Concerns of ChatGPT," *arXiv preprint arXiv:2307.14192*. doi: https://doi.org/10.48550/arXiv.2307.14192

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., and Gašević, D. 2023. "Practical and ethical challenges of large language models in education: A systematic literature review," *arXiv preprint arXiv:2303.13379*. doi: https://doi.org/10.48550/arXiv.2303.13379

Yang, H., Liu, X.-Y., and Wang, C. D. 2023. "FinGPT: Open-Source Financial Large Language Models," *arXiv preprint arXiv:2306.06031*. doi: https://doi.org/10.48550/arXiv.2306.06031

Appendix A

Table I: Distribution of challenges of using LLMs in higher education across the studies reviewed.

Challenges	Source
Danger to Authentic Learning	McMurtrie (2023), Rasul et al. (2023), Rudolph et al. (2023b), Jalil et al. (2023), Alqahtani et al. (2023), Crawford et al. (2023b), Eager and Brunton (2023), Crawford et al. (2023a), Perkins (2023), Milano et al. (2023), Susnjak (2022), Qadir (2023), Peng et al. (2023), Markauskaite et al. (2022), Malik et al. (2023), Leinonen et al. (2023), Dergaa et al. (2023),
Academic misconduct	Rasul et al. (2023), Mills et al. (2023), Dobslaw and Bergh (2023), Cochran et al. (2023), Marron (2023), Jalil et al. (2023), Alqahtani et al. (2023), Eager and Brunton (2023), Crawford et al. (2023a), Perkins (2023), Kasneci et al. (2023), Susnjak (2022), Qadir (2023), Malik et al. (2023), Huang and Tan (2023), Farrokhnia et al. (2023), Cotton et al. (2023),
Unreliable and biased	McMurtrie (2023), Teel et al. (2023), Bhayana et al. (2023), Jalil et al. (2023), Alqahtani et al. (2023), Eager and Brunton (2023), Álvarez-Álvarez and Falcon (2023), Bauer et al. (2023), Crawford et al. (2023a), Kasneci et al. (2023), Milano et al. (2023), Rasheed et al. (2019), Qadir (2023), Farrokhnia et al. (2023),
Ethical and Equity Practices	Rasul et al. (2023), Koutcheme (2023), Alqahtani et al. (2023), Crawford et al. (2023b), Crawford et al. (2023b), Eager and Brunton (2023), Bauer et al. (2023), Crawford et al. (2023a), Perkins (2023), Kasneci et al. (2023), Milano et al. (2023), Rasheed et al. (2019), Qadir (2023), Malik et al. (2023), Liu et al. (2023c), Cotton et al. (2023)
Limited Evaluation of Graduate Skill Sets	McMurtrie (2023), Cochran et al. (2023), Marron (2023), Jalil et al. (2023), Alqahtani et al. (2023), Eager and Brunton (2023), Crawford et al. (2023a), Perkins (2023),
Lack of Social Interaction	Rasul et al. (2023), Eager and Brunton (2023),

Appendix B

Table II: Distribution of benefits of using LLMs in higher education across the studies reviewed.

Benefits	Source
Easy to Use and Explain Confusing Ideas	McMurtrie (2023), Dobslaw and Bergh (2023), Rudolph et al. (2023b), Bhayana et al. (2023), Alqahtani et al. (2023), Crawford et al. (2023b), Eager and Brunton (2023), Álvarez-Álvarez and Falcon (2023), Crawford et al. (2023a), Perkins (2023),
Smooth Out Awkward First Drafts	Rasul et al. (2023), Giannos (2023), Cochran et al. (2023), Alqahtani et al. (2023), Crawford et al. (2023b), Dergaa et al. (2023), Álvarez-Álvarez and Falcon (2023), Kasneci et al. (2023), Liu et al. (2023c), Liu et al. (2023b), Latif et al. (2023), Huang and Tan (2023)
Preparation for AI Integration	McMurtrie (2023), Rasul et al. (2023), Dobslaw and Bergh (2023), Rudolph et al. (2023b), Giannos (2023), Bhayana et al. (2023), Marron (2023), Jalil et al. (2023), Alqahtani et al. (2023), Crawford et al. (2023b), Eager and Brunton (2023), Álvarez-Álvarez and Falcon (2023), Bauer et al. (2023), Crawford et al. (2023a), Perkins (2023), Kasneci et al. (2023), Malik et al. (2023), Krause (2023),
Enhanced Learning through Personalisation	Rasul et al. (2023), Rudolph et al. (2023b), Giannos (2023), Cochran et al. (2023), Crawford et al. (2023b), Eager and Brunton (2023), Crawford et al. (2023a), Kasneci et al. (2023), Milano et al. (2023), Qadir (2023), (Cotton et al. 2023)
Individualised Feedback	Rasul et al. (2023), Mills et al. (2023), Rudolph et al. (2023b), Alqahtani et al. (2023), Álvarez-Álvarez and Falcon (2023), Crawford et al. (2023a), Perkins (2023), Kasneci et al. (2023), Ndukwe et al. (2020), Rasheed et al. (2019), Uchiyama et al. (2023), Qadir (2023), Farrokhnia et al. (2023),
Automated Administrative Support	Rasul et al. (2023), Rudolph et al. (2023b), Alqahtani et al. (2023), Crawford et al. (2023b), Bauer et al. (2023), Ndukwe et al. (2020), Perkins et al. (2023), Farrokhnia et al. (2023),
Innovative Assessment Activities	Rasul et al. (2023), Dobslaw and Bergh (2023), Rudolph et al. (2023b), Cochran et al. (2023), Alqahtani et al. (2023), Crawford et al. (2023b), Eager and Brunton (2023), Álvarez-Álvarez and Falcon (2023), Bauer et al. (2023), Crawford et al. (2023a), Perkins (2023), Kasneci et al. (2023), Milano et al. (2023), Ndukwe et al. (2020), Rasheed et al. (2019), Liu et al. (2023b), Uchiyama et al. (2023), Perkins et al. (2023), Malik et al. (2023), Cotton et al. (2023),

Acknowledgements

Copyright