A Conceptual Framework for Identifying Risks of AI Integration in Higher Education

Asma Ali Mosa Eltharif
Computer Department
Education Faculty, Universality of Benghazi, Libya

Reem Ali Salem Abdalla
Computer Department
Education Faculty, Universality of Benghazi, Libya

ABSTRACT
The integration of Artificial Intelligence (AI) into Higher Education (HE) presents a dual-edged sword, offering transformative potential while introducing significant risks. This paper synthesizes existing literature to examine the multifaceted challenges associated with AI adoption in academic settings, including threats to academic integrity, algorithmic biases, data privacy concerns, and the erosion of critical human skills. This paper proposes the HE AI Risk Ecology (HE-AIRE) model, a conceptual framework that categorizes these risks into four interconnected layers: pedagogical, technological-ethical, institutional-operational, and socioeconomic-cultural. By highlighting systemic interdependencies, the model underscores the need for holistic strategies to mitigate risks while harnessing AI's benefits. The study calls for robust ethical frameworks, equitable implementation, and proactive policy interventions to ensure AI aligns with the core values of HE.

Keywords
Artificial Intelligence (AI), Higher Education (HE), Risks, Ethical Frameworks, Technology in Education.

1. INTRODUCTION

Artificial Intelligence (AI) technologies have become widely used in various aspects of life. With the development of computing and information processing techniques, artificial intelligence (AI) has been widely applied in multiple sectors. In healthcare, AI systems demonstrate improved diagnostic accuracy and enable personalized treatment plans [55]. But it's not just medicine; AI's playing a bigger role in environmental efforts too, whether it's modeling climate patterns or optimizing how we manage resources [84]. Over in finance and economics, the technology's proving equally transformative. Businesses are leaning on AI for everything from risk analysis to streamlining operations, and research like Cao [14] backs this up.

The education sector has similarly adopted these technologies, where AI applications are transforming teaching and learning processes. Known as Artificial Intelligence in Education (AIEd), a field encompassing intelligent tutoring systems, educational robotics, learning analytics, and adaptive learning platforms. These technologies are reshaping how learners interact with content and educators manage instruction [65, 118]. These AI tools process large datasets, identify complex patterns, and generate predictive insights - capabilities that are becoming fundamental to modern education systems.

The integration of artificial intelligence (AI) into HE has evolved from theoretical discussion to widespread implementation, with recent studies documenting its institutional adoption [118], pedagogical impacts [65,51], and operational transformations [83,109]. This technological shift is particularly evident in HE, where artificial intelligence is influencing teaching, administration, and student support. Personalized learning environments powered by AI are helping tailor instruction to individual student needs, enhancing engagement and performance [52]. Administrative processes are increasingly automated, streamlining tasks like grading, scheduling, and resource allocation [118]. Furthermore, predictive analytics are being used to identify students at risk and support enrollment and retention strategies [54]. Beyond these changes, AI introduces a broader pedagogical shift—transforming traditional educational experiences into algorithmically guided interactions, also it can give a new perspective to Teaching-Learning in HE [83].

However, as AI in education applications continues to develop and improve, discussing its potential benefits and perceived risks becomes crucial [118]. The increasing demand for professionals in the area of information systems and technologies from different economic and social sectors has generated new requirements for HE institutions. Artificial intelligence (AI) plays a crucial role in enhancing the pedagogical process, making learning more efficient and effective, improving the quality of teaching and learning, and expanding the range of available resources [65]. By adapting the teaching process to the personal characteristics and individual needs of students, AI can personalize the learning process to match individual student characteristics, such as learning pace, cognitive strategies, and knowledge gaps. This allows for dynamic adjustments in teaching, enabling timely responses to change in students' needs and interests [52]. This responsiveness enables educators to identify skills, factors, and behaviors associated with learning difficulties, allowing for timely interventions tailored to each student. Consequently, both educators and students can focus more on their primary roles: developing knowledge and fostering critical thinking skills, as knowledge alone does not inherently change behavior [83].

Thus, Artificial Intelligence (AI) is revolutionizing HE by enabling innovative teaching methods, improving learning outcomes, and streamlining administrative tasks [118]. In HE, institutions are increasingly incorporating AI-powered tools into their operations. Technologies such as adaptive learning platforms, predictive analytics for student retention, and AI-driven chatbots for academic advising are no longer experimental—they are becoming part of the digital infrastructure in many universities [118]. These applications are part of the broader field of Artificial Intelligence in Education (AIEd), which seeks to personalize and optimize the learning experience [65]. However, alongside these advancements, the integration of AI presents challenges that require critical examination [109].

This paper investigates the impact of artificial intelligence (AI) in HE by identifying and analyzing key risks and their implications for educators, students, and institutions. Through a review of existing literature, the study aims to develop a

systematic understanding of these challenges and proposes a conceptual framework to articulate the risks associated with AI integration in HE.

2. PROBLEM AND RESEARCH GAP

Artificial intelligence (AI) has become increasingly present in HE, bringing with it remarkable opportunities but also stirring deep concerns. On the positive side, AI has been praised for its role in creating personalized learning experiences and streamlining administrative tasks [118]. Yet, the rapid integration of AI has raised serious questions about academic integrity, fairness, and the preservation of education's fundamental values [108].

While a growing number of studies highlight the potential of AI, there is still a noticeable lack of comprehensive frameworks that explore its risks in a holistic manner [50]. Most research tends to zoom in on individual issues—like bias in algorithms or incidents of academic dishonesty—without looking at how these risks might overlap and even reinforce each other [66]. As a result, the broader, systemic nature of the challenges is often overlooked.

Although discussions about AI's ethical implications in education are common [108], practical guidance for universities remains limited. Institutions are often left without clear strategies to address these emerging risks in real-world settings [117].

A significant gap remains in the form of a unified approach that identifies various risks, illustrates their interconnections, ranks them by potential impact, and provides concrete, targeted actions for stakeholders. In the absence of such a model, efforts to manage AI risks risk remaining fragmented and ineffective.

In response, this study introduces the HE AI Risk Ecology (HE-AIRE) model, offering a structured, action-oriented framework designed to help universities better understand and navigate the complex landscape of AI-related risks.

3. UNDERSTANDING ARTIFICIAL INTELLIGENCE

3.1 Artificial Intelligence Definition

Artificial intelligence (AI) broadly refers to the ability of machines to simulate human intelligence processes, including learning, reasoning, and self-correction. According to Russell and Norvig [87], AI encompasses a wide range of subfields such as machine learning, natural language processing, computer vision, and robotics, all contributing to its vast and growing range of applications across different sectors, including education.

Historically, the concept of AI was first formally introduced by John McCarthy in 1955, who defined it as "the science and engineering of making intelligent machines" [70]. This foundational definition reflects the interdisciplinary nature of AI, blending principles from computer science, mathematics, cognitive science, and engineering to create systems capable of exhibiting intelligent behavior.

More recent perspectives have emphasized AI's interactive and functional abilities. For example, Luckin [65] describes AI not only as a system that simulates cognitive functions but also as a tool capable of collaborating, communicating, and adapting in dynamic environments. Similarly, European Commission [35] defines AI systems as "software (and possibly hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected data, and taking actions."

In machine learning, a major branch of AI, three primary forms of learning have emerged. The first, supervised learning, involves teaching machine patterns in a labeled dataset, enabling it to classify new, unseen data [3]. In contrast, unsupervised learning analyzes datasets without predefined labels, aiming to discover hidden patterns or groupings within the data [74]. Finally, reinforcement learning allows machines, or agents, to learn optimal behaviors through trial and error by interacting with their environment and receiving feedback in the form of rewards or penalties [99].

These forms of learning form the backbone of many AI systems deployed today, including those transforming the educational landscape.

3.2 Artificial Intelligence in HE

Artificial intelligence has steadily become a critical tool across multiple sectors, and education is no exception. Within HE, AI is increasingly used to develop what scholars call "intelligent educational systems" platforms that leverage AI to personalize learning experiences, automate administrative tasks, and enhance student support [118].

One of the most promising applications of AI in education is personalized learning. Adaptive platforms, such as Carnegie Learning's MATHia, tailor instruction to each student's pace, strengths, and weaknesses, providing a learning experience that traditional methods struggle to match [50]. Administrative automation is another significant benefit, with AI systems now capable of handling routine tasks like grading assignments, managing schedules, and answering frequently asked student queries through chatbots like Ivy.ai.

Additionally, AI is employed in predictive analytics, where systems such as Civitas Learning analyze student data to identify individuals at risk of dropping out or underperforming, allowing for timely and targeted interventions [93]. Beyond academics, AI tools help students manage their time, reduce administrative burdens, and even monitor mental health indicators [65].

Despite these advantages, scholars have cautioned that the deployment of AI in HE should be approached thoughtfully. Without careful design and implementation, AI could unintentionally reinforce biases or erode trust between students and institutions [108]. Therefore, integrating AI in universities requires not only technical sophistication but also ethical foresight and inclusive governance frameworks.

4. METHODS

The study employed a literature review to identify and explore the risks and drawbacks of integrating AI technologies into higher education institutions, categorizing them within an integrated framework encompassing relevant key concepts.

The framework presents the risks and drawbacks of using AI in HE institutions across different main themes. The study was based on two main questions:

1. What are the risks associated with the use of AI applications in HE institutions?
2. How do these risks interact with educational, technological, institutional, and social dimensions?

4.1 Data Sources

The sources and references of the study varied, as most of them were extracted from internationally accredited and well-known

databases such as Scopus, Web of Science, Google Scholar, SpringerLink, and Elsevier, in addition to reports issued by international organizations such as the World Economic Forum and UNESCO. Table 1 and Figure 1 show the data sources and the usage percentage corresponding to each source.

Table 1. The data sources

Source	Number of references	Percentage
Google Scholar	20	40%
Scopus	12	24%
Web of Science	8	16%
SpringerLink	5	5%
UNESCO	5	5%
Total	**50**	**100%**

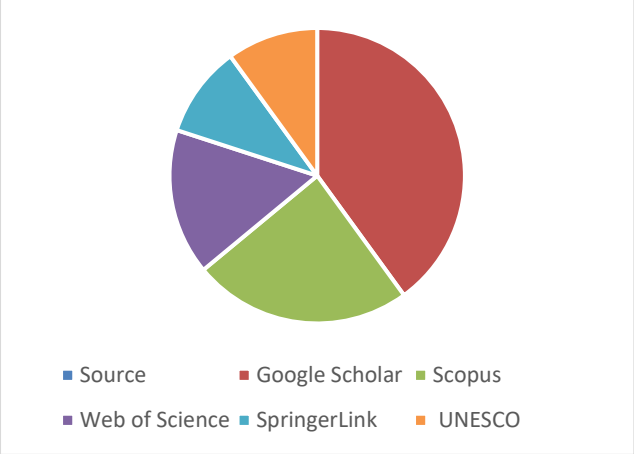

Fig 1: Data Source Chart

4.2 Inclusion and Exclusion Criteria

The research materials were chosen according to the following standards:

• **Included studies**: Publications from 2015 to 2024 that explicitly examine the risks and challenges linked to artificial intelligence (AI) in HE settings.

• **Excluded studies**: Works focusing on general or vocational education were not considered, nor were those that highlight only the advantages without addressing potential risks.

4.3 Analysis Method

This study used a thematic analysis approach to identify the challenges and risks arising from the integration of artificial intelligence in HE. Through careful review of previous scientific studies on the same topic, several key concepts were identified. These concepts were grouped according to the HE AI Risk Environment (HE-AIRE) model. This model divides risks into four areas:

- Educational risks, which address the impact of AI on teaching methods, the quality of learning, and student engagement.
- Technical ethical risks, including privacy issues, algorithmic bias, and ethical concerns regarding data use.
- Institutional risks, covering issues universities may face when adopting AI, such as policymaking and staff preparation.
- Economic risks, related to the financial costs of using AI and the risk of increasing inequality between rich and poor institutions.

This approach allowed for a deeper understanding of how these risks connect and overlap. It also showed the importance of handling them in a balanced and thoughtful way.

4.4 Develop Framework Concepts

This model provides a structured approach to analyzing and examining the risks associated with the integration of AI in HE, organizing them according to a framework organized into key levels. The framework primarily aims to guide the safe and ethical integration of AI in HE institutions. It focuses on four components of AI risks in HE: educational, technological, institutional, and economic. This model allows academics to anticipate challenges and formulate policies to improve the effectiveness of AI technologies.

5. REVIEW OF LITERATURE ON RISKS

Despite the increasing use of artificial intelligence, its application in HE institutions is not without risks. This study aims to explore the potential risks associated with the use of artificial intelligence in this field. Many studies have highlighted the challenges and risks that may arise from the integration of artificial intelligence in HE. To achieve this, a systematic search was conducted in previous studies that addressed the risks of artificial intelligence, with the aim of identifying the most prominent challenges that academic institutions may face when using this technology.

Ayala-Pazmiño [5] systematically evaluates the risks posed by AI integration in educational contexts, emphasizing ethical, pedagogical, and equity concerns. The study identifies the following key risks:

- **Threats to Academic Integrity:** AI tools facilitate cheating through automated essay generation and assignment completion, undermining authentic learning [58]. This necessitates redesigning assessments to prioritize creativity and critical thinking over rote outputs [59].
- **Widening Educational Inequalities:** Disparities in AI access exacerbate the "digital divide," leaving underserved students at a systemic disadvantage [62]. Socioeconomic gaps may deepen as privileged groups leverage AI tools more effectively [75].
- **Dehumanization of Learning Environments:** Over-automation reduces meaningful teacher-student interactions, which are crucial for socioemotional development [85]. AI-driven feedback systems cannot replicate the mentorship and adaptability of human educators [48,89].
- **Data Privacy and Algorithmic Bias:** Massive data collection by AI systems risks misuse and commercial exploitation [75]. Biased training data may perpetuate discriminatory outcomes in grading and resource allocation [64].
- **Erosion of Critical Human Skills:** Overdependence on AI diminishes creativity, handwriting proficiency, and collaborative problem-solving [98]. Machine-generated content may replace opportunities for original thought and expression [40,111].

Özer [80]conducted a systematic study evaluating the potential benefits and risks of AI integration across educational systems, examining perspectives from students, teachers, and administrators. Through a comprehensive review of existing literature, the study identified significant risks spanning ethical, pedagogical, and operational dimensions that necessitate urgent

mitigation. These include:

- **Privacy and Data Protection Concerns:** The use of AI systems in education raises critical issues regarding the security and misuse of student and teacher data. Unauthorized commercial exploitation of such data could restrict individuals' lifelong freedoms [56,102].
- **Exacerbation of Educational Inequalities:** Disparities in digital and AI literacy between and within countries may lead to unequal access to AI tools, reinforcing existing educational inequalities [21,42]. The "Matthew effect" in education suggests that advantaged groups may further benefit from AI, while disadvantaged groups fall further behind [79,29].
- **Erosion of Critical Thinking and Academic Integrity:** Overreliance on AI-generated content risks diminishing students' critical thinking, problem-solving, and independent research skills [57,97]. Additionally, AI tools like ChatGPT facilitate academic dishonesty, including plagiarism, undermining the value of original work [43,95].
- **Bias and Discrimination in AI Outputs:** AI systems trained on biased datasets may perpetuate and institutionalize societal biases related to gender, race, and culture, affecting fairness in educational assessments [63,69].
- **Inadequate Assessment and Evaluation Methods:** Traditional assessment frameworks struggle to accurately measure student contributions when AI tools are involved, potentially rewarding unethical use and masking learning deficiencies [17,37].
- **Depersonalization and Reduced Human Interaction:** While AI can personalize learning, excessive automation may reduce meaningful teacher-student interactions, negatively impacting mentorship and socioemotional development [30,86].
- **Job Displacement and Changing Teacher Roles:** The automation of teaching tasks risks misrepresenting teacher workloads, leading to potential job cuts despite the increasing complexity of their roles in managing AI-integrated classrooms [25,41].
- **Dependence on Unreliable AI-Generated Content:** AI systems like ChatGPT sometimes produce inaccurate or outdated information, requiring students to verify outputs—a skill many may lack [17,73].

The systematic literature review by Alpaydin [7] identifies several critical risks associated with the integration of generative artificial intelligence (GAI) in HE. These risks, drawn from empirical studies published between 2023 and 2024, span academic integrity, pedagogical challenges, ethical concerns, and institutional implications. Below is a synthesis of the key risks articulated by the authors:

- **Threats to Academic Integrity:** The widespread use of GAI tools, particularly ChatGPT, raises significant concerns about plagiarism and the authenticity of student work. Studies highlight that AI-generated content can produce passable responses to assessments, undermining traditional evaluation methods [76,7]. For instance, ChatGPT's ability to generate essays, code, and exam answers challenges the validity of assessments, necessitating revised academic integrity policies [2,94].
- **Erosion of Critical Thinking and Authentic Learning:** Overreliance on GAI tools may diminish students' critical thinking and problem-solving skills. Research indicates that students using ChatGPT for assignments risk bypassing deep engagement with course material, leading to superficial learning outcomes [92,8]. Faculty express concerns that AI dependence could reduce students' ability to "think independently" [91].
- **Bias and Ethical Dilemmas:** GAI systems can perpetuate biases present in their training data, disadvantaging marginalized student populations [8]. Ethical concerns also arise around transparency, as AI-generated content may lack accountability or reproduce harmful stereotypes [71]. Additionally, the use of AI for grading or feedback introduces questions about fairness and algorithmic bias [34].
- **Challenges for Assessment Design:** Educators face difficulties in distinguishing between student-authored and AI-generated work. Studies show that even experienced markers struggle to identify AI-produced text reliably [47]. This undermines assessment validity and necessitates redesigning assignments to emphasize creativity and higher-order thinking [76,7].
- **Depersonalization of Education:** While GAI can personalize learning through adaptive tools, excessive automation risks reducing meaningful student-instructor interactions. Faculty warn that AI-driven feedback may lack the nuance of human mentorship, potentially isolating learners [107,7].
- **Institutional and Policy Gaps:** Many institutions lack clear policies to govern GAI use, creating ambiguity for stakeholders. Without ethical guidelines, risks such as data privacy violations, intellectual property disputes, and labor displacement (e.g., AI replacing administrative roles) remain unaddressed [16,8].
- **Detection and Enforcement Limitations:** Current AI-detection tools (e.g., Turnitin) exhibit limited accuracy, especially against adversarial techniques that evade detection [82]. This arms race between AI-generated content and detection methods complicates academic oversight [23].

The study conducted by Akinwalere and Ivanov [1] concluded that the use of artificial intelligence (AI) in HE, despite its promising potential, involves a number of risks and challenges that must be addressed consciously and responsibly by educational institutions. Below are the key risks identified by the researchers:

- **Unintended Negative Outcomes:** One of the most prominent potential risks is that the use of AI systems may lead to unexpected adverse effects, despite the positive intentions of their developers. These outcomes often stem from reliance on outdated or unrepresentative data, which reduces the effectiveness of these systems when applied in different contexts [1].
- **Lack of Data Inclusivity:** The study highlighted that the prevalence of biased data, as seen in facial recognition technologies, results in lower accuracy when dealing with certain demographic groups. For example, research has shown that facial recognition accuracy is significantly higher for light-skinned men compared to dark-skinned women, which could negatively impact surveillance or assessment practices in classrooms [81].
- **Limited Accuracy and Reliability:** Most AI systems rely on statistical predictions based on correlation rather than causation, making it difficult for educators to distinguish between genuine and misleading patterns. This increases the risk of making educational decisions based on superficial or inaccurate conclusions [90].
- **Ineffectiveness of Predictive Outputs:** In some cases, the information provided by AI systems is not practically useful. For instance, the study mentioned a predictive system that informed professors that 80% of students would fail a particular course but offered no concrete suggestions to address the issue, leaving educators unsure of how to utilize this information [1].
- **Practical Implementation Challenges:** The study

demonstrated that the responsible use of AI requires comprehensive training for teachers and administrators, along with clear protocols defining when to rely on AI outputs and when to defer to human professional judgment. This is essential to avoid hasty or unfair decisions [1].
- **Profound Philosophical and Educational Implications:** The study warns that excessive reliance on AI could erode the human aspects of education and diminish the teacher's pedagogical and interactive role. The researchers emphasized that the teacher's role is irreplaceable and that AI should be viewed as a supportive tool that enhances—rather than replaces—the educational process [1].

The study by Dhawan and Batra [28] concludes that the use of artificial intelligence in HE is not without challenges and risks that require careful understanding and responsible management. Among the most prominent challenges are:

- **Data Privacy and Security Concerns:** AI systems rely on vast amounts of personal and sensitive data, making institutions vulnerable to breaches and misuse. This is particularly critical in the education sector, which has demonstrated relatively weak cybersecurity performance compared to other industries [33].
- **High Technological Costs:** Implementing AI requires substantial investments in infrastructure, data management, and algorithm development. These costs can strain the financial capacity of many HE institutions, especially in developing regions [45].
- **Digital Illiteracy Among Students and Educators:** A significant number of students and faculty members lack the necessary technical skills to effectively interact with AI systems, creating a digital divide and impeding equitable access to technological benefits [4].
- Shortage of Qualified Experts: Institutions often face difficulties in recruiting and retaining skilled professionals capable of developing and managing AI tools. This talent gap slows down effective implementation and affects the overall quality of AI-driven initiatives [67].
- **Bias in Algorithms:** AI systems may produce biased or discriminatory outcomes if trained on skewed or unrepresentative data. A well-known example is Amazon's AI recruiting tool, which was found to favor male applicants due to biased historical data [24].
- **Immature Nature of Technology and Lack of Human Skills:** Despite its speed and efficiency, AI still lacks essential human abilities such as empathy, creativity, critical thinking, and emotional intelligence. These limitations can be problematic in educational settings where human interaction is crucial.
- **Institutional Implementation Challenges:** Successful integration of AI requires robust digital infrastructure, well-defined strategic plans, and adequate training programs. Without these foundational elements, institutions risk poor outcomes and failed implementations [78].

The integration of artificial intelligence (AI) in HE presents significant transformative potential but is accompanied by several risks, as outlined by Zaman [113]. These risks span ethical, pedagogical, and operational domains, necessitating careful mitigation strategies to ensure responsible adoption.

- **Bias and Discrimination in AI Systems:** AI algorithms may perpetuate or amplify biases present in training data, leading to discriminatory outcomes in grading, admissions, or personalized learning recommendations [113]. For instance, historical biases in datasets could disadvantage marginalized student populations, reinforcing inequities in educational access and outcomes [112,13].
- **Data Privacy and Security Concerns:** The extensive data collection required for AI-driven tools raises critical privacy issues. Unauthorized access, misuse of sensitive student data, and insufficient transparency in data handling pose risks to student confidentiality [113]. Institutions must implement robust safeguards to ensure the protection of personal data and uphold trust in AI-driven educational environments. Without clear privacy frameworks, the risk of unauthorized data exploitation and breaches increases significantly, undermining both institutional credibility and student rights [112,96].
- **Erosion of Human Interaction:** Overreliance on AI risks diminishing the role of educators, reducing opportunities for mentorship, emotional support, and nuanced feedback essential for holistic student development [113]. AI cannot replicate the interpersonal dynamics vital for fostering critical thinking and creativity [112,49].
- **Academic Integrity Challenges:** AI-powered tools, such as automated essay generators or problem-solving applications, may facilitate plagiarism and undermine authentic learning [113]. This complicates the assessment of student competency and erodes academic standards [112,83].
- **Digital Divide and Accessibility Inequities:** While AI can enhance accessibility for some learners, disparities in technology access may exacerbate inequalities. Students from under-resourced backgrounds or institutions may lack the infrastructure to benefit from AI-driven education, widening the digital divide [113,106].
- **Implementation and Operational Barriers:** Deploying AI systems requires substantial investment in infrastructure, faculty training, and ongoing maintenance, which may be prohibitive for resource-constrained institutions [112]. Poorly designed implementations could lead to technological obsolescence or ineffective integration [113,119].
- **Ethical and Regulatory Gaps:** The lack of standardized ethical frameworks for AI in education raises concerns about accountability, transparency, and the potential for unintended consequences, such as algorithmic opacity or misuse of predictive analytics [112,31].

The integration of artificial intelligence (AI) in HE presents significant risks, as identified by Bennett and Abusalem [10]. These risks span academic integrity, pedagogical effectiveness, equity, and ethical concerns, necessitating proactive institutional responses.

- **Threats to Academic Integrity:** The ease of generating high-quality text, code, and multimedia via AI tools like ChatGPT raises concerns about fraudulent submissions. Studies indicate that AI-generated essays can achieve passable grades [100], undermining the credibility of assessments and devaluing qualifications if unchecked [11].
- **Bias and Inaccuracy in AI Outputs:** AI systems may propagate biases or inaccuracies due to reliance on pre-2021 data and unvetted sources [88]. Prompt phrasing can inadvertently introduce bias, as demonstrated by contrasting responses to politically charged questions [10].
- **Erosion of Critical Thinking:** Overreliance on AI for content generation risks diminishing students' higher-order cognitive skills, such as analysis and creativity, shifting learning outcomes toward superficial engagement [11].
- **Privacy and Intellectual Property Concerns:** The use of AI tools involves data collection, raising issues about student privacy, ownership of AI-generated content, and unauthorized use of academic work [10].
- **Depersonalization of Education:** AI-driven automation may reduce human interaction, potentially isolating students and eroding mentorship opportunities [11].

- **Exacerbation of Inequities:** Disparities in access to technology (e.g., reliable internet or AI tools) could widen educational gaps for marginalized students [10].
- **Ethical and Alignment Challenges:** Tensions between AI's efficiency and human values—such as transparency, empathy, and ethical judgment—pose risks if institutional policies fail to address these conflicts [19,11].

The paper by Tundrea [104] identifies five key ethical risks arising from the integration of artificial intelligence (AI) in HE, each with significant implications for academic integrity, equity, and human agency.

- **Perception of Equality:** The increasing sophistication of AI-driven educational tools, such as virtual teaching assistants and adaptive learning systems, risks blurring the distinction between human educators and machines. For instance, AI tutors like Georgia Tech's "Jill Watson" or China's Yixue system have demonstrated capabilities comparable to—or surpassing—human instructors in certain tasks [103]. This may lead to the erroneous perception that AI possesses moral equivalence to humans, undermining the relational and empathetic aspects of education central to student development (Bryson, 2010).
- **Data Privacy:** AI systems in HE rely on extensive data collection, including student behavior, engagement metrics, and even social media activity, raising concerns about surveillance and exploitation. The paper highlights that AI's ability to infer sensitive attributes (e.g., emotional states, political views) from seemingly neutral data exacerbates privacy risks [104,27]. Without robust safeguards, institutions risk violating student autonomy and enabling misuse by third parties [36].
- **Moral Agency:** Delegating decision-making to AI—such as admissions, grading, or dropout predictions—introduces ambiguity about accountability. The paper argues that AI lacks genuine moral reasoning, yet its recommendations may influence critical academic outcomes. For example, an AI system predicting student dropout risk (e.g., London South Bank University's model) could perpetuate biases if not transparently designed [103,101]. Ethical frameworks, like MIT's "Moral Machine," are proposed to crowdsource moral guidelines for AI systems [104,72].
- **Moral Deskilling:** Overreliance on AI for ethical or academic decisions may erode human judgment. The paper warns that outsourcing decisions to machines—such as scholarship allocations or disciplinary actions—could diminish educators' capacity to engage with complex moral dilemmas [103]. This deskilling effect threatens the foundational role of universities in fostering critical thinking and ethical reasoning.
- **Data Bias and Misclassification:** AI systems trained on non-representative or historically biased data may replicate or amplify inequalities. Examples include discriminatory admissions algorithms or grading systems that disadvantage marginalized groups [104,12,20]. The paper advocates for "intentional transparency" in algorithmic design and independent audits to mitigate bias [53].

The paper by Zanetti et al. [115] examines the potential risks of artificial intelligence (AI) in education, focusing on ethical, pedagogical, and operational challenges. The authors identify several key risks, supported by interdisciplinary literature, which are summarized below.

- **Bias and Algorithmic Discrimination:** AI systems in education are susceptible to biases embedded in their design, training data, or human interactions. The authors categorize these biases into five types:
  - Dataset bias: Limited or unrepresentative training data perpetuates stereotypes [39].
  - Associations bias: AI reinforces societal prejudices (e.g., gender or racial stereotypes) through flawed data patterns [120].
  - Automation bias: Overreliance on AI-driven decisions without human oversight [18].
  - Interaction bias: Human users may intentionally or unintentionally distort AI learning [116].
  - Confirmation bias: AI amplifies pre-existing beliefs, reducing diversity in educational outcomes [9].

Such biases can manifest in Intelligent Tutoring Systems (ITS), where culturally or linguistically insensitive algorithms disadvantage certain student groups [77].

- **Data Privacy and Security Concerns:** The collection and use of sensitive student data (e.g., emotional states, learning patterns) raise ethical issues. Despite frameworks like the Delicate checklist [38] and GDPR compliance, risks persist, including:
  - Unauthorized profiling of students via Learning Analytics [32].
  - Exploitation by third-party vendors or malicious actors [115].
- **Academic Integrity and Pedagogical Impact**
  - Plagiarism and critical thinking erosion: Generative AI tools may undermine originality [116].
  - Emotional and developmental risks: Prolonged interaction with AI, such as emotion-recognizing robots, could impair relational skills or impulse control [110].
  - Inequitable access: ITS and AI-driven tools may widen gaps due to socioeconomic disparities in technology adoption [105].

Long-Term Sociocultural Effects: The authors hypothesize that AI-mediated learning might alter cognitive development, particularly when virtual reality replaces experiential learning [6]. Additionally, biased AI outputs could normalize harmful stereotypes among students [60].

6. RESULTS AND DISCUSSION

A comprehensive review and thematic analysis of scholarly sources reveal that the risks associated with integrating Artificial Intelligence (AI) into Higher Education (HE) are multidimensional and interconnected. Four core thematic dimensions emerged from the literature as central to understanding these risks: pedagogical, technological-ethical, institutional-operational, and socioeconomic-cultural. Table 2 shows how previous studies have contributed to shaping each of the four dimensions.

Table 2. Mapping Reviewed Studies to the Four Thematic Dimensions

Dimension	Brief Description	Representative Studies
Pedagogical	Refers to the erosion of critical thinking, creativity, and authentic learning due to overreliance on AI tools.	[7], [10], [56] ,[92], [40]
Technological	Encompasses algorithmic bias, the inaccuracy of AI-	[64], [115], [1], [17], [73]

	generated outputs, and limitations of detection systems.	
Institutional	Covers the lack of AI governance policies, infrastructural challenges, and gaps in strategic implementation.	[15], [104], [113], [44], [78], [26].
Sociocultural	Involves issues such as the digital divide, inequitable access to AI tools, and the amplification of social bias.	[22], [42], [79], [62], [29], [11]

Pedagogical risks are prominent in literature, particularly concerning the erosion of critical thinking and authentic learning. Studies indicate that tools like ChatGPT enable students to produce seemingly original work without genuine intellectual engagement, raising serious questions about academic integrity and the effectiveness of traditional assessment methods [8,10]. Beyond plagiarism, there is apprehension that overreliance on AI may reduce students to passive consumers rather than active learners, threatening the core mission of higher education to foster independent and critical thought [57,91].

Technological-ethical risks encompass algorithmic bias, data privacy concerns, and the opacity of AI systems. Several studies highlight how biases in training data can produce discriminatory outcomes in grading, admissions, or personalized learning recommendations [64,116]. Moreover, the widespread collection of student data for AI analytics raises significant privacy concerns, with surveys revealing that over 70% of students are uneasy about how their data might be used [41]. The lack of transparency in AI decision-making processes further compounds these ethical challenges, leaving little room for human oversight or accountability [1].

Institutional-operational risks stem from gaps in governance, inadequate faculty training, and resource constraints. Many universities lack comprehensive policies to guide the ethical use of AI, resulting in fragmented and inconsistent practices [7,16]. Financial barriers, particularly in lower-resourced regions, hinder the deployment of AI infrastructure and limit access to training for faculty and staff [45,112]. This institutional unpreparedness heightens the risk of misusing AI tools or failing to harness their potential effectively.

Socioeconomic-cultural risks are closely tied to issues of equity and access. The digital divide continues to create disparities between students and institutions with ample resources and those with limited technological capacity [21,80]. Moreover, AI tools often lack cultural sensitivity, risking marginalization of diverse learner populations by imposing uniform, Western-centric pedagogical models [115,49]. Concerns also persist about AI's potential to erode the human dimension of education, reducing educators' roles to mere supervisors of automated systems rather than active facilitators of learning [25].

Across these four dimensions, the risks are not isolated but intricately linked. Technological flaws can directly influence pedagogical outcomes, while institutional gaps exacerbate socioeconomic inequalities. These interconnected findings underscore the need for a holistic understanding of AI integration in HE.

Therefore, to translate these insights into practical guidance, the following section introduces the HE AI Risk Ecology (HE-AIRE) model—a conceptual framework that organizes these dimensions into an interconnected system, providing a comprehensive lens for managing AI-related risks in higher education.

7. Proposed Conceptual Framework: The HE AI Risk Ecology (HE-AIRE) Model

This study systematically examines the multifaceted risks of AI in higher education through the development of the HE AI Risk Ecology (HE-AIRE) model as in Figure 2. Based on comprehensive literature analysis, this ecological framework organizes AI-related risks into four interconnected domains (as identified in Table 2), highlighting their systemic nature and mutual dependencies.

1. Core Pedagogical Layer

This innermost layer captures risks that directly affect teaching, learning, and assessment processes.

- Academic Integrity Threats: Tools like ChatGPT have enabled students to generate content with minimal effort, compromising the authenticity of assessments and encouraging plagiarism [7,11].
- Decline in Critical Thinking: Overreliance on AI-generated responses can reduce opportunities for students to engage in deep reasoning and problem-solving [56,92].
- Assessment Misalignment: Traditional grading systems struggle to differentiate between human- and AI-generated work, which challenges the validity of student evaluations [46,17].

2. Technological-Ethical Layer

This layer includes the risks embedded in the design, deployment, and operation of AI systems.

- Algorithmic Bias and Discrimination: AI systems often replicate social biases from training data, potentially disadvantaged marginalized groups [64,116, 68].
- Data Privacy Violations: The use of learning analytics and surveillance-based AI tools raises concerns about data misuse and student autonomy [103,75,113].
- Lack of Explainability and Accountability: Institutions often use AI systems as "black boxes," leading to opaque decision-making processes with limited opportunities for human oversight [1,15].

3. Institutional-Operational Layer

Here, the focus is on institutional preparedness and governance capabilities.

- Policy and Governance Gaps: Many institutions lack formal policies to regulate AI use in education, resulting in inconsistent practices and ethical blind spots [7,31].
- Faculty Training and Digital Literacy: Without adequate professional development, educators may misuse or underutilize AI tools, exacerbating risks [28,78].
- Implementation Challenges: Technological infrastructure, budget constraints, and staff shortages impede effective AI integration, particularly in low-resource contexts [44,119].

4. Socioeconomic-Cultural Layer

This outermost layer reflects the broader social, economic, and cultural environment in which AI in HE is embedded.

- Digital Divide and Inequity: Differential access to AI

technologies can widen existing gaps between well-resourced and marginalized institutions or learners [61,80,42].

- Cultural Insensitivity: AI tools may lack cultural adaptability, failing to recognize diverse learning styles or values [115,49].
- Labor Displacement and Role Erosion: Automation threatens to marginalize human educators by reducing their roles to facilitators of AI-generated instruction [26].

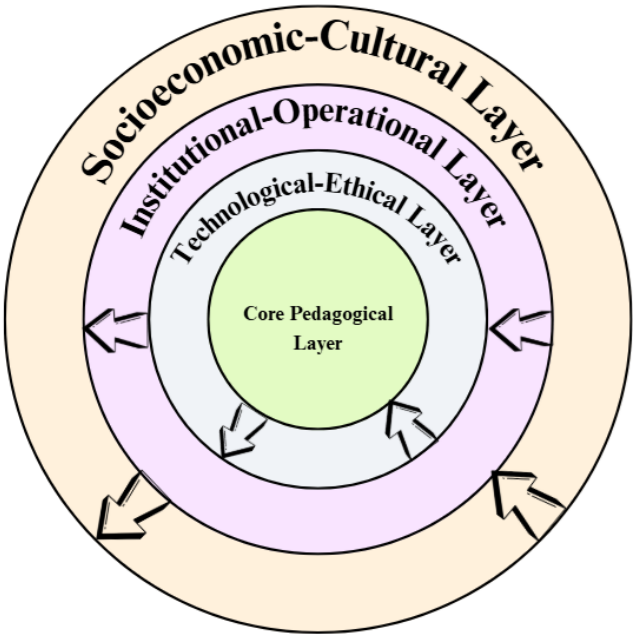

Fig 2: HE AI Risk Ecology (HE-AIRE) Model

The HE AI Risk Ecology (HE-AIRE) Model conceptualizes the interconnected nature of AI-related risks in HE by organizing them into four concentric layers. Each layer represents a distinct but interdependent domain—Core Pedagogical, Technological-Ethical, Institutional-Operational, and Socioeconomic-Cultural. The concentric design reflects the dynamic interplay among these domains, emphasizing that risks are systemic rather than isolated.

Bidirectional arrows between the layers represent feedback loops, illustrating how vulnerabilities in one area can intensify risks in others.

1. Core Pedagogical Layer ↔ Technological-Ethical Layer

The pedagogical and technological-ethical layers are closely linked. On one hand, educators shape how AI tools are implemented in the classroom, influencing technological norms and expectations. On the other hand, the ethical design and limitations of these tools affect pedagogical integrity.

For instance, if instructors incorporate AI-based tools for automated essay grading, the system's algorithmic limitations could introduce unintentional bias—such as favoring certain writing styles. This affects grading fairness and, by extension,

student outcomes. At the same time, pedagogical reliance on such tools may influence the future development of similar systems, pushing for more transparency and adaptability. As noted by Kasneci et al. [57] and [91], overreliance on AI in student work can erode critical thinking skills, while ethical risks related to algorithmic design [64,114] highlight the need for alignment between pedagogy and technology.

2. Technological-Ethical Layer ↔ Institutional-Operational Layer

AI tools used in higher education operate within institutional settings, and the governance of their deployment is influenced by both technical capabilities and ethical considerations. Institutions must establish guidelines to ensure responsible AI use, while technical developments often challenge the adequacy of those institutional responses.

Take, for example, facial recognition systems introduced to monitor attendance or prevent cheating. In the absence of robust data protection policies, such tools may lead to ethical breaches and institutional liability. Conversely, well-defined policies and adequate training can shape how ethically sound technologies are selected and deployed. The literature confirms these concerns: the absence of institutional policy [7,31] and gaps in faculty readiness [28,78] can worsen the ethical shortcomings of AI tools [1].

3. Institutional-Operational Layer ↔ Socioeconomic-Cultural Layer

The relationship between institutional structures and broader societal dynamics is mutual. Socioeconomic and cultural factors affect how institutions adopt and regulate AI tools, while institutional decisions can reinforce or mitigate existing inequities.

Consider a university located in a low-income region that lacks the infrastructure to support sophisticated AI platforms. This limitation reflects broader digital divides. However, if the institution actively seeks inclusive strategies—such as providing offline-compatible AI tools or investing in digital literacy, it can help reduce inequality and foster more equitable access to technology. This interplay is reinforced in the literature, where disparities in access [61,80], cultural insensitivity in AI design

[49], and implementation barriers [45,19] converge to shape institutional strategies and outcomes.

These bidirectional relationships highlight the ecological complexity of AI risks in higher education. By acknowledging the reciprocal influence between each layer, the HE-AIRE model moves beyond isolated interventions and toward holistic, systems-aware responses. This perspective is essential for policymakers, educators, and technologists aiming to ensure responsible and inclusive AI integration in educational environments.

8. CONCLUSION

The study demonstrates that artificial intelligence offers significant benefits to higher education—such as personalized learning, streamlined administrative processes, and enhanced decision-making simultaneously introduces complex challenges. To investigate the risks associated with AI in higher education, the research addressed the primary question concerning potential risks in universities and colleges through a comprehensive literature review. Findings indicate that these risks are systemic and multidimensional, encompassing educational, ethical, institutional, social, and cultural aspects.

For instance, tools like ChatGPT have complicated the ability to determine whether students produce their own work, adding pressure to assessment practices and learning evaluations [8,10]. Furthermore, excessive reliance on AI may encourage students to bypass critical thinking and accept machine-generated content without scrutiny [56].

Ethical concerns also emerge regarding the opaque functioning of AI systems. These technologies can reflect biases inherent in their training data and often operate as "black boxes," obscuring how decisions are reached [64,112]. Such risks are amplified in contexts lacking robust governance frameworks and adequate training [44]. Moreover, disparities in access to AI technologies persist, disadvantaging students in under-resourced regions [41].

Another significant issue involves AI's potential to alter the role of educators, with some instructors perceiving a shift from mentorship to technical oversight [25]. This development raises questions regarding whether AI serves as a tool to support education or risks replacing critical human elements, such as personal interaction, creativity, and equity.

To address how these risks intersect across educational, technological, policy, and societal spheres, the HE AI Risk Ecology (HE-AIRE) model has been proposed. This model illustrates the interconnected nature of these challenges and underscores the necessity for holistic consideration. Moving forward, higher education institutions must establish clear policies, enhance digital competencies, and ensure that AI adoption genuinely serves the core mission of education.

9. LIMITATIONS AND FUTURE WORK

This study is conceptual in nature and relies primarily on a comprehensive literature review to identify and synthesize the risks associated with AI integration in higher education. While the HE AI Risk Ecology (HE-AIRE) model provides a structured framework for understanding these risks, it has not yet been empirically tested across diverse institutional contexts or datasets.

A significant avenue for future research involves conducting empirical evaluations of the HE-AIRE model using various datasets and real-world scenarios. For instance, quantitative studies could collect data from different universities to examine how the identified risk dimensions manifest in practice. Qualitative investigations, such as case studies or interviews with stakeholders across multiple institutions, could further validate and refine the framework's applicability and comprehensiveness.

Such empirical assessments would not only strengthen the robustness of the HE-AIRE model but also provide valuable insights into contextual variations, enabling higher education institutions to develop tailored strategies for responsible and equitable AI integration.

Funding

This study was not supported by any funding

Compliance with Ethical Standards

Conflict of Interest

The authors declare that they have no conflict of interest.

10. REFERENCES

[1] Akinwalere, S. N., & Ivanov, V. (2022). Artificial intelligence in higher education: Challenges and opportunities. International Journal of Educational Technology in Higher Education, 19(1), 15.

[2] Alexander, K., Savvidou, C., & Alexander, C. (2023). Who wrote this essay? Detecting AI generated writing in second language education in higher education. *Teaching English with Technology, 23*(2), 25–43.

[3] Alpaydin, E. (2020). *Introduction to machine learning* (4th ed.). MIT Press.

[4] Annual Status of Education Report Rural Provisional, (2018*), ACER Centre New Delhi,* 1-344.

[5] Ayala-Pazmiño, M. (2023). Artificial intelligence in education: exploring the potential benefits and risks. *Digital Publisher CEIT*, *8*(3), 892-899.

[6] Barr, R. (2013). Memory constraints on infant learning from picture books, television on touchscreens. *Child Development Perspectives*, *7*(4), 105–110.

[7] Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, *15*(11), 676.

[8] Batista, P., Nikolic, M., & Alexander, B. (2024). *Critical Risks of Generative AI in Higher Education: A Systematic Review*.

[9] Bebbington, K., MacLeod, C., Ellison, M., & Fay, N. (2017). The sky is falling: Evidence of negativity bias in the social transmission of information. *Evolution and Human Behavior*, *38*(1), 92–101.

[10] Bennett, L., & Abusalem, A. (2024). Artificial Intelligence (AI) and Its Potential Impact on the Future of Higher Education. *Athens Journal of Education*, *11*(3), 195-212.

[11] Bennett, S., & Abusalem, A. (2024). *Artificial Intelligence and Academic Integrity Challenges in Higher Education*.

[12] Bhate, S. (1987). Prejudice against doctors and students from ethnic minorities. *British Medical Journal, 294*(6575), 838.

[13] Blodgett, S. L., & Madaio, M. (2021). Risks of AI foundation models in education. arXiv preprint arXiv:2110.10024.

[14] Cao, L. (2020). AI in finance: A review. *Available at SSRN*, *3647625*, 1.

[15] Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*(1), 38.

[16] Chan, E. (2023). *Policy Gaps and Ethical Challenges in AI Use for Higher Education Institutions*.

[17] Chiu, T. K. F. (2023). The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 1-17.

[18] Chou, J., Murillo, O., Ibars, R. (26 settembre 2017). How to Recognize Exclusion in AI. *Medium.*

[19] Christian, B. (2020). *The Alignment Problem: Machine Learning and Human Values.* W.W. Norton.

[20] Clarke, B., Fokoue, E., & Zhang, H. H. (2009). *Principles and theory for data mining and machine learning.* Springer.

[21] Cotton, D. R. E., Cotton, P. A., & Kneale, P. (2023). *Digital Inequality and Access to AI Technologies in Education*.

[22] Cotton, D. R., Cotton, P. A., & Shipway, J.R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12.

[23] Currie, G., & Barry, K. (2023). ChatGPT in nuclear medicine education. *Journal of Nuclear Medicine Technology, 51*(4), 247–254.

[24] Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women. Reuters Technology News, URL:*

[25] De Cremer, D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. *Harvard Business Review*, 18.

[26] De Cremer, D., de Sousa Jabbour, A. B. L., & Jabbour, C. J. C. (2021). *The Future of Work and AI in Education*.

[27] Deane, M. (2018). *AI and the future of privacy. Towards Data Science*.

[28] Dhawan, S., & Batra, A. (2021). *Emerging Challenges of AI Integration in Higher Education*.

[29] Dignum, V. (2021). *Ethics in Artificial Intelligence: Balancing Efficiency and Human Values*.

[30] Dignum, V. (2021). The role and challenges of education for responsible AI. *London Review of Education*, 19(1), 1-11.

[31] Doleck, T., Lemay, D. J., Basnet, R. B., & Bazelais, P. (2020). Predictive analytics in education: A comparison of deep learning frameworks. Education and Information Technologies, 25, 1951–1963.

[32] Drachsler, H., & Greller, W. (2016). Privacy and analytics – it's a DELICATE issue. A checklist to establish trusted learning analytics. *6th Learning Analytics and Knowledge Conference 2016*, Edinburgh, UK.

[33] Education Cybersecurity Report. (2018), *https://security scorecard.com/resources/2018 education-report.*

[34] Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20*(1), 57.

[35] European Commission. (2021). *Ethics guidelines for trustworthy AI.*

[36] European Parliament. (2019). *EU guidelines on ethics in artificial intelligence: Context and implementation* (Briefing No. PE 640.163).

[37] Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1–15.

[38] Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest editorial: Ethics and privacy in learning analytics. *Journal of Learning Analytics*, *3*(1), 5–15.

[39] Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., & Crawford, K. (2019). *Datasheets for Datasets.*

[40] Gee, J. P. (2018). What video games must teach us about learning and literacy. Palgrave Macmillan.

[41] Grassini, S. (2023). *Digital Divide and AI Literacy Inequalities in Global Education*.

[42] Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*, 692.

[43] Halaweh, M. (2023). *ChatGPT in education: Strategies for responsible implementation*. Bastas: Tokyo, Japan.

[44] Handa, C. (2019). *Technological Costs and Barriers to AI Adoption in Universities*.

[45] Handa, U. (2019). *Is Artificial Intelligence Development Expensive?* Cynoteck, URL:

[46] Hassoulas, A., Powell, N., Roberts, L., Umla-Runge, K., Gray, L., & Coffey, M. J. (2023). Investigating marker accuracy in differentiating between university scripts written by students and those produced using ChatGPT. *Journal of Applied Learning and Teaching, 6*(2), 71–77.

[47] Hassoulas, A., Watermeyer, R., & Burrows, R. (2023). *Detection Challenges of AI-Generated Academic Work*.

[48] Hattie, J. (2012). Visible learning for teachers: Maximizing impact on learning. Routledge.

[49] Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570.

[50] Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

[51] Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education, 32*(3), 504-526.

[52] Holstein, K., McLaren, B. M., & Aleven, V. (2020).

Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. Artificial Intelligence in Education, 12163, 339–351.

[53] IEEE Standards Association. (2019). *IEEE standards for artificial intelligence affecting human well-being.* IEEE Transmitter.

[54] Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. Educational Technology Research and Development, 68, 1961–1990.

[55] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H. & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, *2*(4).

[56] Kasneci, E., et al. (2023). *Impacts of AI on Academic Skills and Critical Thinking*.

[57] Kasneci, E., Sesler, K., Kuchemann, S., Bannert, M., Dementieva, D. et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.

[58] Kavale, K. A., & Forness, S. R. (2019). Cheating and plagiarism in schools and colleges. Psychology Press.

[59] Kulkarni, C., Cambre, J., Kotturi, Y., Bernstein, M. S., & Klemmer, S. R. (2015). Peer and self assessment in massive online classes. ACM Transactions on Computer-Human Interaction (TOCHI), 22(2), 1-30.

[60] Levendowski, A. (2018). How copyright law can fix artificial intelligence's implicit bias problem. *Washington Law Review*, *93*(2), 579–630.

[61] Liao, C., et al. (2021). *AI Access Disparities and the Reinforcement of Inequalities*.

[62] Liao, Y., Huang, R., Sun, C., & Li, X. (2021). Artificial intelligence in education: Opportunities and challenges from a learning science perspective. Frontiers in Education, 6, 1-7.

[63] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*, 410.

[64] López-Pérez, M.V., Pérez-Rodríguez, M.A., & Gutiérrez-Santiuste, E. (2020). Artificial intelligence in education: A review. Frontiers in Psychology, 11, 135.

[65] Luckin, R. (2018). *Machine learning and human intelligence: The future of education for the 21st century*. UCL Institute of Education Press.

[66] Luckin, R. (2021). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour, 5*(2), 146–152.

[67] Ma, Yizhi & Siau, Keng L. (2018). *Artificial Intelligence Impacts on Higher Education".* MWAIS 2018 Proceedings, 42.

[68] Mbakwe, A. B., Lourentzou, I., Celi, L. A., Mechanic, O. J., & Dagan, A. (2023). ChatGPT passing USMLE shines a spotlight on the flaws of medical education. *PLoS Digital Health*, 2, e0000205.

[69] Mbakwe, A., et al. (2023). *Bias and Fairness in AI-Driven Educational Tools*.

[70] McCarthy, J. (1955). *Proposal for the Dartmouth summer research project on artificial intelligence.*

[71] Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences, 13*(9), 856.

[72] MIT Media Lab. (2019). *Moral machine: Human perspectives on machine ethics* [Interactive platform].

[73] Mogali, S. R. (2023). Initial impressions of ChatGPT for anatomy education. *Anatomical Sciences Education*, 1–4.

[74] Murphy, K. P. (2012). Machine learning: A probabilistic perspective. MIT Press.

[75] National Education Association. (2020). Preparing educators for AI.

[76] Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education, 48*(4), 559 614.

[77] Nye, B. D. (2014). Intelligent Tutoring Systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, *25*(2), 177–203.

[78] Ocaña-Fernandez, Y., Valenzuela-Fernandez, L., & GarroAburto, L. (2019). Artificial Intelligence and its Implications in Higher Education. *Propósitosy Representaciones,* 7(2), 536-568.

[79] Özer, M. (2023). The Matthew effect in Turkish education system. *Bartın University Journal of Faculty of Education*, 12(4), 704-712.

[80] Özer, M. (2024). Potential benefits and risks of artificial intelligence in education. *Bartın University Journal of Faculty of Education*, *13*(2), 232-244.

[81] Perera, M., & Aboal, D. (2020). The Impact of a Mathematics Computer-Assisted Learning Platform on Students' Mathematics Test Scores.

[82] Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024). Detection of GPT 4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics, 22*(1), 89–113.

[83] Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. Research and Practice in Technology Enhanced Learning, 12(1), 22.

[84] Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Slavin Ross, A., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A. and Luccioni, A. (2019). Tackling Climate Change with Machine Learning. *arXiv e-prints*, arXiv-1906.

[85] Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta analytic approach. Review of

Educational Research, 81(4), 493-529.

[86] Rudolph, J., Tan, S., Tan, S. C. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 1-22.

[87] Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

[88] Scharth, M. (2023). *'How to perfect your prompt writing for ChatGPT, Midjourney and other AI generators'. The Conversation.*

[89] Sclater, N., & Peasgood, A. (2018). Artificial intelligence in education: Promises and implications for teaching and learning. British Journal of Educational Technology, 49(4), 745-760.

[90] Sharma, Y. (2011). *Boost to university-industry AI research collaboration*. University World News.

[91] Shimizu, H., et al. (2023). *AI Dependence and the Erosion of Critical Thinking Skills*.

[92] Shimizu, I., Kasai, H., Shikino, K., Araki, N., Takahashi, Z., Onodera, M., Kimura, Y., Tsukamoto, T., Yamauchi, K., Asahina, M., et al. (2023). Developing medical education curriculum reform strategies to address the impact of generative AI: Qualitative study. *JMIR Medical Education, 9*, e53466.

[93] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30–40.

[94] Singh, M. (2023). Maintaining the integrity of the South African university: The impact of ChatGPT on plagiarism and scholarly writing. *South African Journal of Higher Education, 37*(2), 203–220.

[95] Skavronskaya, L., Hadinejad, A., & Cotterell, D. (2023). Reversing the threat of artificial intelligence to opportunity: A discussion of ChatGPT in tourism education. *Journal of Teaching in Travel & Tourism*, 23, 253–258..

[96] Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529.

[97] Sok, S., & Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. *Cambodian Journal of Educational Research*, 3(1), 110-121.

[98] Sulik, M. J., Huerta, M., & Ziegler, J. C. (2017). The development of handwriting speed and legibility in grades 1-9. Journal of educational psychology, 109(6), 800-813.

[99] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

[100] Terwiesch, C. (2023*). Would Chat GPT Get a Wharton MBA? Mack Institute for Innovation Management.* Wharton University Pennsylvania.

[101] Tickle, L. (2015). How universities are using data to stop students dropping out. *The Guardian*.

[102] Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A. et al. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*, 15.

[103] Tundrea, A. (2020). *Ethical Risks of AI in Education: A Comprehensive Analysis*.

[104] Tundrea, E. (2020). *Artificial Intelligence in Higher Education: challenges and opportunities.* INTED2020 Proceedings, 2041-2049.

[105] VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221.

[106] Vasoya, N. H. (2023). The role of parents and educators in managing the risks of artificial intelligence. Asian Journal of Education and Social Studies, 41(4), 1–5.

[107] Watermeyer, R., Phipps, L., Lanclos, D., & Knight, C. (2023). Generative AI and the automating of academia. *Postdigital Science and Education, 6*(2), 446–466.

[108] Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235.

[109] Williamson, B., 2023. The social life of AI in education. *International Journal of Artificial Intelligence in Education*.

[110] World Economic Forum (2019). *The Global Risk Report 2019*.

[111] World Economic Forum. (2018). The Future of Jobs Report 2018. Retrieved from

[112] Zaman, B. U. (2023). Transforming education through ai, benefits, risks, and ethical considerations. *Authorea Preprints*, *1*, 1-26.

[113] Zaman, M. (2023). *The Risks of Bias, Privacy, and Inequality in AI Applications for Higher Education*.

[114] Zanetti, M. (2018). Pregiudizio ed etichettamento: il ruolo dell'insegnante nello sviluppo di comportamenti devianti. *Formazione & Insegnamento*, *16*(2), 193–204.

[115] Zanetti, M., et al. (2020). *Bias, Privacy, and Long-Term Implications of AI in Educational Environments*.

[116] Zanetti, M., Rendina, S., Piceci, L., & Cassese, F. P. (2020). Potential risks of artificial intelligence in education. *Form@ re-Open Journal per la formazione in rete*, *20*(1), 368-378.

[117] Zawacki-Richter, O., & Jung, I. (Eds.). (2023). *Handbook of open, distance and digital education*. Springer.

[118] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

[119] Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. Complexity, 2021, Article 8812542, 1–18.

[120] Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist - it's time to make it fair. Nature, 559, 324–326.