5-2024

Ethical imperatives in AI-driven educational assessment: Framework and implications

Ming Soon Tristan LIM
Singapore Management University, tristanl.2021@engd.smu.edu.sg

Ethical Imperatives in AI-Driven Educational Assessment: Framework and Implications

Tristan LIM

Submitted to School of Computing and Information Systems in partial fulfillment of the requirements for the Degree of Doctor of Engineering

Dissertation Committee

Swapna Gottipati (Supervisor/ Chair)
Associate Professor of Information Systems (Education)
Singapore Management University

Michelle Cheong (Co-Supervisor)
Professor of Information Systems (Education)
Singapore Management University

Prof. David Lo (Committee Member)
OUB Chair Professor of Computer Science
Singapore Management University

Soujanya Poria (External Examiner)
Assistant Professor of Computer Science
Singapore University of Technology and Design

Singapore Management University
2024

I hereby declare that this Doctor of Engineering (EngD) dissertation is my original work, and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in this dissertation.

This EngD dissertation has also not been submitted for any degree in any university previously.

Tristan

Tristan LIM

4 May 2024

Abstract

This dissertation embarks on an extensive exploration of the ethical challenges emerging from the integration of AI in educational assessments. It uncovers the complex interplay between AI and the ethical imperatives these technologies pose within educational assessments.

Amidst the rapid development of AI-enabled educational technologies, such as Ubiquitous, Adaptive, and Immersive technologies, this research identifies a notable gap in literature specifically concerning the ethical imperatives and implications of AI in educational assessments. Addressing this gap, the dissertation has three primary objectives: to comprehend and analyze the underpinning educational technologies driving assessments, to elucidate the intricate relationship between AI, ethics, and educational assessments, and to develop a comprehensive theoretical framework addressing the ethical challenges inherent in AI implementations in assessments.

The dissertation contributes to the research field by offering a nuanced examination of AI's role in educational assessments and its ethical ramifications. It introduces a robust framework to guide educators, policymakers, and researchers through the ethical complexities of AI implementation. This study not only bridges the literature gap but also provides actionable insights for the practical application of AI in educational settings, emphasizing the need for ethical consideration at every stage of the assessment pipeline.

The dissertation highlights the dynamic trajectories of educational technologies, stressing the rising importance of adaptive technologies and the transformative role of immersive and ubiquitous technologies in assessments. It underscores the necessity of ethical vigilance in AI

applications and validates a generalizable framework for ethically grounded AI-enabled assessments.

The dissertation opens pathways for future exploration, suggesting the need for interdisciplinary methodologies, longitudinal studies, deeper analysis of learners' AI understanding, and practical applications of the study’s insights. It calls for a collaborative, informed approach among various stakeholders in education to responsibly harness AI's potential, ensuring its integration not only advances educational practices but does so with ethical integrity and pedagogical effectiveness.

Table of Contents

Acknowledgement

Foremost, I extend my deepest gratitude to my mother, whose enduring faith in education and the boundless opportunities it presents has been a guiding light in my life. Her encouragement and belief in the potential of a little boy have been instrumental in my perseverance and success.

I am immensely grateful to my wife, Christine, whose steadfast support has been my pillar of strength throughout this endeavor. Her selfless dedication to our family, especially in nurturing our two wonderful daughters Angel and Mirabelle with boundless love and energy, has allowed me to devote myself to this research.

My heartfelt thanks go to my esteemed supervisors, Prof. Swapna Gottipati, Prof. Michelle Cheong, and Prof. David Lo. Their invaluable support, encouragement, and belief in my work have been pivotal in this journey. Many thanks also go to the helpful pointers and support from Prof. Soujanya Poria. I am particularly appreciative of their guidance, which has been instrumental in enhancing my understanding of research quality and in shaping the development of my work to its current comprehensive and pertinent state.

As I stand at the beginning of a long research journey, I am excited about the prospect of continuing my collaborations with esteemed colleagues from SMU, SUTD and other universities. The connections and friendships I have forged along this path have enriched this experience, making it both fruitful and meaningful.

I recognize that this dissertation is not just a reflection of my efforts, but a combination of the support, guidance, and inspiration provided by all those mentioned above, and for that, I am eternally grateful.

Chapter 1. Introduction

In this dissertation, titled "*Ethical Imperatives in AI-Driven Educational Assessment: Framework and Implications*," we examine the transformative role of Artificial Intelligence (AI) in educational assessment practices. This work investigates both the opportunities and challenges presented by AI's integration into this domain, with a focus on the profound ethical dimensions that arise. As AI reshapes traditional assessment methods, it introduces a range of ethical complexities that demand thoughtful analysis and proactive measures.

Central to this dissertation is the interplay between cutting-edge AI technologies and the ethical imperatives they prompt in educational assessments. Notwithstanding the extensive research on the broader impact of technology in education (e.g., Bozkurt, 2020), specifically AI in education (AIED) (e.g., Chen et al., 2020), and ethical imperatives tied to AIED (e.g., Borenstein & Howard, 2021), there is a noticeable gap in literature specifically addressing the ethical imperatives of AI on educational assessment. These assessments, encompassing grading, feedback, and student performance evaluation (González-Calatayud, Prendes-Espinosa & Roig-Vila, 2021; Sánchez-Prieto et al., 2020), entail unique ethical challenges distinct from other AI applications in education (Bearman & Luckin, 2020). Issues such as fairness in grading algorithms and their impact on student self-esteem and motivation are critical to assessments, especially given their significant consequences on students' academic and career trajectories (e.g., Surahman & Wang, 2022). This gap highlights the need for comprehensive frameworks to navigate the ethical intricacies of AI in educational settings.

This dissertation aims to illuminate these critical ethical dilemmas, striving to balance technological advancements with ethical responsibility. It narrates the potential of AI in transforming educational assessments and critically examines the ethical boundaries shaping this change.

The dissertation has three primary objectives: first, to explore key educational technologies – Ubiquitous, Adaptive, and Immersive – and establish AI-driven adaptive technology as a growing domain, highlighting ethical imperatives as a central theme in technology-enabled educational assessments. Second, to expand the current understanding of the interactions between AI technology, ethics, and educational assessments, propose an ethics framework for AI-driven assessments, and discuss implications and measures that can be taken to address ethical imperatives in AI-driven assessments. Third, to validate and establish a robust theoretical framework that addresses the identified ethical challenges. This effort seeks to fill the existing literature gap and provide practical guidance for educators, policymakers, and researchers.

By offering an exploration of the interplay between AI technology and ethics in educational assessments, this dissertation makes a significant contribution to the field. It proposes a generalizable ethical framework, guiding stakeholders in addressing the complexities of AI in education. The dissertation provides a thorough analysis of technological impacts and a detailed examination of ethical considerations, charting a responsible and effective path for AI's integration in educational assessment practices.

1.1 Purpose of Study

The overarching purpose of this dissertation is to conduct an in-depth examination of the ethical imperatives arising from the integration of AI in educational assessments. This study is designed to:

1. Assess the extensive influence of various technologies, broadly classified as Ubiquitous, Adaptive, and Immersive technologies, on educational assessments. This involves a detailed analysis of how these technologies are reshaping assessment methodologies, with a focus on identifying pivotal technological (such as AI) and thematic research areas (such as ethical imperatives) within this domain.

2. To identify, explore, and critically analyze the ethical imperatives that emerge in tandem with the deployment of AI in educational assessments. This includes a thorough exploration of ethics and AI implementation in educational assessments, the implications brought forth, the proposal of an AI ethics framework for assessments, and the measures to manage ethical dilemmas and complexities introduced by AI.

3. Validate and establish a theoretically-grounded ethical framework for AI in educational assessments. This framework aims to be adaptable and relevant across various educational contexts, providing a guide for ethical decision-making and implementation of AI technologies in assessment practices.

1.2 Research Questions

These overarching research questions are formulated to align with and expand upon the study's purpose, guiding the investigation through its three pivotal chapters:

1. **Chapter 2 - Technological Evolution and its Influence on Educational Assessment Practices**

   RQ1: In what ways are Ubiquitous, Adaptive, and Immersive technologies transforming educational assessment practices?

   This question aims to dissect the transformative impact of educational technologies, classified broadly as Ubiquitous, Adaptive, and Immersive technologies, on educational assessments. By exploring how Ubiquitous, Adaptive, and Immersive technologies are being integrated and utilized, the research seeks to uncover key trends and themes in the evolving methodologies of educational assessments. The analysis will involve trend, bibliometric, and network analyses to systematically map out the advancements and patterns emerging in this area. This exploration is crucial for understanding the trajectory of technological evolution in educational settings and its implications on assessment practices, identifying AI as a key area of research focus.

   RQ2: What common themes underlie the current integration of these technologies in educational assessments?

   This question looks at understanding and inferring common themes that underlie the implementation of Ubiquitous, Adaptive, and Immersive technologies in assessments, through the outputs of network analyses. This question is pivotal in identifying how ethical consideration, among others, is an important theme in technology-enabled educational assessments. By examining these themes, the research contributes to an understanding of how ethical imperatives are to be aligned with advancements in educational technologies in assessment practices.

2. **Chapter 3 - Ethical Dimensions of AI in Educational Assessments**

RQ3: What are the Key AI Use Cases relating to Assessments?

This question explores the various applications of AI in educational assessments, utilizing network analysis and topic modeling to identify dominant trends and areas of focus in the literature. Understanding the range of AI applications and their prevalence in research is useful for comprehending the scope of AI applications that need to be addressed.

RQ4: What are the Key Ethical Principles Arising from the AI Implementations Relating to Assessments?

This question investigates the main ethical principles arising from AI use in assessments, using network analysis and topic modeling. The goal is to identify and categorize the ethical principles most commonly implicated in AI educational assessments, useful for comprehending the scope of ethical issues that need to be addressed.

RQ5: What are the Key Themes Inherent in the Consideration of Ethical Imperatives in Educational Assessments?

This question aims to identify and analyze the key themes between AI applications and ethical considerations in educational assessments, as found in the literature. By employing network analysis and topic modeling, the goal is to map out a generalizable framework, that will facilitate informed and ethical AI integration strategies in educational assessments.

RQ6: What are Solutions and Interventions that were Proposed to Address Key Ethical Imperatives, and their Associated Underpinning Theories?

This question looks to identify and recommend mitigating solutions and intervention measures that can be put in place to address ethical issues, by looking at proposed and/or implemented actions in existing literature. This inquiry will contribute to a better understanding of the current solutions landscape, offering insights into existing strategies and suggesting directions for future research and application.

3. **Chapter 4 - Formulating and Validating an Ethical Framework in AI-Enabled Educational Assessments**

RQ7: How do we Validate the Triadic Theoretical Framework using SEM Analysis?

This question aims to validate the triadic theoretical framework, which is pivotal in mapping the assessment pipeline and ethics elements in AI-enabled educational assessments. The validation process uses Structural Equation Modeling (SEM) analysis to scrutinize the framework's efficacy in encapsulating both the operational and ethical dimensions of AI in educational assessments. This includes an in-depth examination of how the framework integrates and reflects learners' perceptions, thereby ensuring that the framework is not only theoretically sound but also practically relevant and responsive to the needs and views of the learners.

RQ8: What are the Relationships that Emerge between the Stages of the Assessment Pipeline, Key Ethical Imperatives, Output Variables, and Learner Perceptions?

This inquiry is crucial to understanding the dynamic interplay between these components. It will provide insights into how each stage of the assessment pipeline interacts with ethical considerations, learner feedback, and various output variables. This exploration aims to uncover patterns and correlations that can inform the development of more effective, ethical, and learner-centric AI-enabled educational assessments.

1.3 Literature Review

1.3.1 Introducing Educational Technology in Educational Assessments

The evolution of educational technologies can be traced through the myriad tools, platforms, and systems developed over the decades to amplify and support the educational journey. From basic teaching aids such as chalkboards to the sophisticated, digitally-enabled platforms powered by AI, the core tenet of these technologies remains unchanged: enhancing the quality, efficiency, and personalization of the learning experience to cater to the diverse needs of learners (Collins & Halverson, 2018). Parallel to this technological advancement, assessment practices have also evolved. Assessment practices, which refer to the methods and tools educators employ to evaluate, document, and interpret students' academic proficiency and skills mastery, play a pivotal role in the academic journey. They inform critical decisions, be it in grading, student placement, progression, or curriculum adjustments, shaping the overall educational trajectory (Bearman et al., 2016; Brookhart, 2014; Taylor & Nolen, 2005).

In the 1960s and 1970s, the world of educational assessment was dominated by the use of mainframe computers, which were sizable machines that required extensive infrastructure. Schools and colleges would send batches of answer sheets to centers equipped with these computers, which would then process the results and dispatch them back. This was a significant leap from manual checking, offering a glimpse into the potential of computer-aided assessments, albeit in a rudimentary form (Suppes, 1966).

Throughout the 1980s, the educational assessment landscape underwent significant transformations. The introduction of computer-assisted testing, propelled by the widespread availability of personal computers, permitted more standardized testing environments and timely feedback. These tests predominantly feature basic multiple-choice or fill-in-the-blank questions (Bunderson, Inouye, & Olsen, 1988). Concurrently, the 1980s also saw the emergence of Integrated Learning Systems. These systems were among the early attempts at creating holistic computer-based environments tailored for both learning and assessment. Students engaged with these platforms by working on exercises, obtaining immediate feedback, and subsequently undergoing evaluations. A hallmark of these systems was their adaptability; they adjusted the difficulty level of content in real-time, based on individual student performance, setting the stage for more personalized learning experiences (Becker, 1991).

Transitioning into the late 1990s, the technological advancements of the digital age began to manifest more prominently within educational assessments. As computers transitioned from being exclusive, high-cost institutional tools to common household commodities, standardized testing also evolved. This ubiquity allowed standardized testing to make a pivotal shift from paper to digital. Not only did this facilitate quicker result processing, but it also paved the way for a richer, multimedia-based testing experience, incorporating elements like images, audio clips, and even immersive videos into assessments (Russell, 1999).

The dawn of the 21st century brought with it a proliferation of internet-connected platforms, reshaping various facets of educational assessments. One prominent innovation was the emergence of Learning Management Systems (LMS) like Moodle, Blackboard, and Canvas (Coates, James, & Baldwin, 2005). These platforms not only streamlined course content delivery but also integrated versatile assessment tools. Features such as online quizzes and automated feedback became standard, while the ability to conduct peer assessments harnessed the power of collaborative learning, digitally transforming a traditionally manual evaluation method (Topping, 2009). Alongside these advancements, e-portfolios started gaining traction. As digital repositories, they enabled students to continuously document and reflect upon their academic journey. This approach provided educators with a holistic perspective of student growth and learning processes over time (Barrett, 2007). Simultaneously, the late 2000s observed the intertwining of gaming principles with educational content, birthing the concept of game-based assessments. These interactive evaluations immersed students in simulated environments, challenging their problem-solving abilities, critical thinking, and collaborative skills, thus providing a nuanced understanding of their competencies in a dynamic setting (Shute et al., 2009).

In the backdrop of these technological shifts, the three core technology groups – ubiquitous, adaptive, and immersive technologies – have emerged as the most prominent and influential in modern education. Each has uniquely revolutionized assessment practices, offering a combination of continuous learning opportunities, personalized evaluation experiences, and realistic, hands-on simulations. Ubiquity technologies, marked by the rise of smartphones, tablets, and other portable devices, made learning and assessment a continuous, anywhere-anytime activity. Following closely was the rise of adaptive technologies, leveraging the power of algorithms and machine learning to craft personalized assessment experiences. Systems could now analyze a student's past performance, identify areas of strength and weakness, and adjust the assessment's difficulty accordingly. The immersive experience was further augmented by immersive technologies, with VR, AR and MR creating environments where students could be evaluated in highly realistic simulations, offering a practical and hands-on form of evaluation.

The journey of integrating technology into assessment practices is not just a matter of operational efficiency or advancing with the times. Deep-seated within this integration are educational theories that champion more personalized, immersive, and engaging learning experiences.

Central to many technology-integrated assessments is the Constructivist Learning Theory, proposed by Piaget (1954). This theory posits that learners build knowledge based on their experiences. Modern technologies, such as adaptive learning platforms and immersive simulations, offer environments where students can actively construct knowledge through exploration, reflection, and application. Within these technological environments, assessments become more than just gauges of recall; they delve into deeper cognitive processes, gauging understanding and practical application.

Drawing a parallel with adaptive technologies is Vygotsky's concept of the Zone of Proximal Development (1978). Zone of Proximal Development defines the gap between what a learner can do without help and what they can achieve with guidance. Adaptive assessment tools, with their ability to provide real-time feedback, often function within a student's Zone of Proximal Development, dynamically adjusting challenges to align with the learner's current capabilities.

However, the explosion of technology in learning also brings forth challenges. One significant concern is the potential of overwhelming students with excessive information. The Cognitive Load Theory, introduced by Sweller (1988), addresses this concern, suggesting that learners have a limited cognitive processing capacity. As such, assessment tools, especially those employing adaptive technologies, must be optimized to balance cognitive load, ensuring students are neither under-challenged nor stretched beyond their limits. Adaptive technologies reflect the principles of Differentiated Instruction. Tomlinson (2001) posited that educators should tailor instruction to meet individual needs. Modern adaptive assessment platforms epitomize this, adjusting in real-time to each student's performance and presenting challenges that are neither too easy nor too difficult.

Adding another layer of depth to the discussion is the Situated Learning Theory by Lave & Wenger (1991). This theory asserts that learning is most potent when embedded within students' activities, contexts, and cultures. Immersive and ubiquitous technologies craft environments where assessments are 'situated' in lifelike settings or remotely in contextual settings, making the evaluation process more contextually relevant and authentic.

The discussion on technology-enhanced assessments cannot be complete without addressing feedback. Grounded in the Feedback Theory is the belief that timely, specific, and actionable feedback can significantly improve learning and rectify misconceptions (Hattie & Timperley, 2007). Many modern technologies are designed to offer real-time, personalized feedback, reinforcing this pedagogical principle.

The push towards technology in assessments is not just a trend but is deeply rooted in established educational theories. These theories, including Constructivist Learning Theory, Situated Learning Theory and Zone of Proximal Development, not only guide the design and implementation of technology-enhanced assessments but also ensure their pedagogical soundness and effectiveness in contributing to the learning process.

1.3.2 Ethical Imperatives in AI-Driven Educational Assessment: Key Terms and Definitions

With the expanding influence of AI, building upon the foundation laid in the previous chapter on educational technology in assessment, this section transitions into the domain of AI-driven assessments. While technological advancements have streamlined and enhanced assessment methods, they have also introduced ethical imperatives that require thorough exploration (e.g., Borenstein & Howard, 2021).

AI-driven educational assessment refers to the application of AI technologies, on a formative or summative basis, in scaffolding and evaluating student learning, performance, and progress (Ouyang, Dinh & Xu, 2023). It encompasses the use of algorithms and machine learning techniques to analyze educational data, personalize learning experiences, and automate grading and feedback processes. AI-driven educational assessments, characterized by their use of big data and sophisticated AI algorithms, have revolutionized traditional assessment methods. Techniques such as natural language processing enable deeper analyses of student responses, offering personalized feedback in adaptive learning pathways (e.g., Rudolph, Tan & Tan, 2023). However, as these systems become more autonomous, the ethical implications surrounding their use become increasingly significant (Nguyen et al., 2023; Chaudhry and Kazim, 2022).

Ethical imperatives in the context of AI in education refer to the ethical obligations and considerations that must guide the design, development, and deployment of AI technologies (Gill, 2021). This entails ensuring that every aspect of AI-driven educational assessment – from data collection and algorithm design to decision-making and outcomes – is proactively grounded by foundational commitment to ethical considerations, beyond mere compliance or superficial ethical assurances. These imperatives encompass a broad spectrum of issues, including but not limited to, data privacy, security, bias and fairness, transparency, and accountability in AI decision-making processes (e.g., Hakami and Hernández-Leo, 2020; Nguyen et al., 2023; Memarian & Doleck, 2023). For instance, with AI systems processing vast amounts of personal data, the risk of data breaches and misuse is a significant concern. The literature highlights cases where inadequate data protection measures in educational technologies have led to privacy violations (Pontual Falcão et al., 2022). Further, AI algorithms are only as unbiased as the data they are trained on. Studies have shown instances where AI in assessments has inadvertently perpetuated biases, leading to unfair outcomes for certain student groups (Martín Núñez and Lantada, 2020; Latham and Goltz, 2019). In addition, the 'black box' nature of many AI systems poses challenges in understanding how decisions are made. This lack of transparency can hinder accountability, especially when AI-driven assessments influence critical educational decisions (Nazaretsky, Cukurova and Alexandron, 2022; Tlili et al., 2019).

These necessitate a rigorous ethical framework to guide the development and deployment of these technologies. A *framework*, in this context, refers to a structured, theoretical construct that provides guidelines or principles for the ethical development and implementation of AI in educational assessments (Hughes, Davis & Imenda, 2019). The development of a robust ethical framework is paramount to the implementation and governance of AI in educational assessments. This framework should be comprehensive, encompassing guidelines and principles (e.g., Li & Gu, 2023) that address the multifaceted ethical challenges posed by AI. Key components of an ethical framework include principles of fairness, accountability, transparency, and inclusivity, among others (Ashok et al., 2022). Incorporating user perspectives, such as learners or educators, is instrumental in formulating an effective framework (e.g., Holmes et al., 2021).

Practical applications of AI in educational assessment bring to light real-world ethical challenges. The literature discusses a wide range of implications, from the positive impacts on stewardship and resource allocation (Borenstein & Howard, 2021) to concerns about equity, accessibility, and the digital divide (Casas-Roma & Conesa, 2021). *Implications* here refer to the potential outcomes, effects, or impacts that AI-driven educational assessments have on various aspects of education, including pedagogy, policy, and student experience. Case studies and examples of AI implementation in diverse educational settings reveal a spectrum of mitigation and intervention programs and activities, from consent for data use (e.g., Costas-Jauregui et al., 2021) to the interpretation and application of AI-generated insights (e.g., Shabaninejad et al., 2022). Developing institutional policies and regulations that govern the ethical use of AI in education is important. This includes guidelines on data use, privacy protections, and standards for AI development (Pontual Falcão et al., 2022; Costas-Jauregui et al., 2021; Gedrimiene et al., 2020).

This literature review has provided an exploration of the ethical imperatives in AI-driven educational assessment. By critically examining the intersection of AI technology and ethics, it lays a foundation for the subsequent chapters, which will explore the development and application of an ethical framework in this context. The literature review underscores that ethical considerations in AI-driven educational assessment are not static but an ongoing journey. As AI technology evolves, so too must our understanding and frameworks for ethical use. This requires continuous research, stakeholder engagement, and an openness to adapt and revise ethical guidelines as new challenges arise.

1.4 Organization of Dissertation

The remaining chapters of the dissertation are structured as follows:

1. *Chapter 2:* "Technological Evolution and its Influence on Educational Assessment " - This chapter sets the stage by examining how emerging technologies such as Ubiquitous, Adaptive, and Immersive technologies are reshaping educational assessment practices. It identifies adaptive learning technologies, driven by extensive growth in AI, as a key area of research. It also introduces ethical responsibility as a common theme that arises alongside these technological changes, setting the tone for the in-depth discussions that follow.

2. *Chapter 3:* "Ethical Imperatives of AI in Educational Assessments" - Building on the technological context established in Chapter 2, this chapter dives into the ethical challenges and considerations engendered by AI in assessments. It offers a systematic literature mapping of ethical issues and AI implementations in assessments, identifies key themes and implications, and draws up a preliminary framework to generalize the key themes for researchers and practitioners.

3. *Chapter 4*: "Formulating and Validating an Ethical Framework in AI-Enabled Educational Assessments" - Synthesizing insights from the technological and ethical discussions in the preceding chapters, this chapter presents the development and validation of a comprehensive ethical framework. This framework is designed to guide ethical decision-making and implementation in AI-driven educational assessments.

4. *Chapter 5*: "Conclusion" – This chapter provides a synthesis of the dissertation's key findings and contributions. It revisits the initial research objectives, and highlights the pivotal insights gained from the exploration of technological evolution and ethical imperatives in AI-enabled educational assessments, underscoring the significance of the developed ethical framework. It also outlines the agenda for future research, suggesting potential avenues for further investigation and development in this evolving domain.

This dissertation adopts a methodical approach where each chapter builds upon the previous, ensuring a seamless narrative flow. Chapter 2 lays the foundational understanding of technological impact in educational assessments, leading to an exploration of emergent ethical imperatives. Chapter 3 focuses and delves deeper into these ethical imperatives of AI, dissecting the implications and measures to manage ethical challenges. Chapter 4 synthesizes the insights gained to construct and validate a comprehensive ethical framework, addressing the intricacies uncovered in earlier chapters. Finally, Chapter 5 closes the dissertation with a reiteration of research objectives, synthesis of overall findings, discussion of limitations, and setting a course for future research.

Through this structured approach, the dissertation ensures a robust and thorough coverage of the critical intersection between AI technology and ethical imperatives in educational assessments. The research questions, closely tied to the study's overarching purpose, enable a focused, in-depth examination of each facet of this intersection, contributing significantly to both the academic and practical realms in the field of education technology.

Chapter 2. Technological Evolution and its Influence on Educational Assessment

This research examines the integration and trajectory of ubiquitous, adaptive, and immersive technologies in educational assessments, based upon qualitative predictions from Horizon Report (HR) (from 2011 to 2023 by Johnson et al., 2011, 2012, 2013, 2014, 2015, 2016, Becker et al., 2017, 2018, Alexander et al., 2019, Brown et al. 2020 and Pelletier et al. 2021, 2022, 2023) and bibliometric statistics. Through bibliometric and network analyses, we identify key educational technological trends and their interconnectedness within the academic domain.

The results underscored the ascendance of adaptive technologies for real-time feedback, the evolving role of immersive technologies in reshaping assessment paradigms, and the pervasive reach of ubiquitous technologies in crafting contextually anchored evaluations. The extensive growth of adaptive technologies, driven by research innovations in AI, looks to play a crucial role in future technology-enabled educational assessments.

Research also identified common themes across use of educational technologies. The ethical imperatives of educational technology deserve particular emphasis. As we integrate more sophisticated technologies into educational assessments, the responsibility to address and manage ethical challenges becomes more critical.

Grounded in pedagogical underpinnings, the study presents pressing research gaps, theoretical and practical insights, positioning itself as a reference for researchers and practitioners in enhancing educational technology-infused assessment strategies.

2.1 Introduction

The rapid adoption of technology in education has brought with it a transformative effect on assessment practices. As technology-enabled learning continues to evolve, it becomes crucial to understand how these developments are influencing assessment redesigns and implementation (Ertmer and Ottenbreit-Leftwich, 2013; Mishra, 2020; Robertson and Barber, 2017). However, while there exists a plethora of studies on technology's broader impact on education, a distinct gap is observed in the literature when it comes to the nuanced influence of specific technological trends on assessment practices.

The intricate dynamics between technological advancements and their ramifications on the design, delivery, and evaluation of assessments presents a pressing research problem: *As educational technology influences shift, how are assessments adapting to these changes, and what does the future hold?* In this research, we embark on a journey to bridge this gap and address the posited problem. We explore the evolution and impact of critical educational technology groups on assessment practices, leveraging the predictions made in HR, coupling them with bibliometric and network analyses. We delve into the predictions of HR, analyzing their intersections with assessment practices. We also discuss common themes that underlie these emerging technology groups, and in particular, ethical imperatives that may intertwine with educational technology adoptions in assessments.

Numerous bibliographic and reference sources offer insights into the potential trajectories of technology in education. Among these is HR. Published by Educause since 2004, HR stands out for its consistent yearly predictions on emerging global educational technologies, covering diverse areas such as teaching, learning, and assessments. Given its long-standing reputation and comprehensive coverage, HR serves as an invaluable resource for this study. Building upon the foundational works of Martin et al. (2011) and Dubé and Wen (2022), we delved into recent literature to explore the nexus between technological trends and their implications on assessment practices. Unlike previous works that predominantly focused on the broad impact of technological trends on education, our study hones in on the specific domain of assessment, aiming to shed light on its evolving landscape.

Through analyses of HR's predictions, we have distilled three dominant educational technology groups that stand at the forefront of reshaping assessment practices, forming the core of our study:

1. *Ubiquitous Technology:* Encompassing technologies such as mobile devices, Internet of Things (IoT) and robotics that facilitate pervasive learning experiences, transcending temporal and spatial boundaries.

2. *Adaptive Technology*: Centered around the paradigm of personalizing learning experiences, leveraging the power of data-driven insights and AI.

3. *Immersive Technology*: Encompassing the realm of games and gamification, and mixed reality (MR) (alongside its derivative technologies virtual reality (VR) and augmented reality (AR)) that craft rich, immersive, and interactive learning environments.

These groups have been identified as the torchbearers of the next wave of assessment innovations, offering glimpses into the future trajectories of assessments in educational settings.

This study investigates the following research questions:

1. *RQ1: In what ways are Ubiquitous, Adaptive, and Immersive technologies transforming educational assessment?*

   This question aims to dissect the transformative impact of educational technologies, classified broadly as Ubiquitous, Adaptive, and Immersive technologies, on educational assessments. By exploring how Ubiquitous, Adaptive, and Immersive technologies are being integrated and utilized, the research seeks to uncover key trends and themes in the evolving methodologies of educational assessments. The analysis will involve trend, bibliometric, and network analyses to systematically map out the advancements and patterns emerging in this area. This exploration is crucial for understanding the trajectory of technological evolution in educational settings and its implications on assessment practices, identifying AI as a key area of research focus.

2. *RQ2: What common themes underlie the current integration of these technologies in educational assessments?*

   This question looks at understanding and inferring common themes that underlie the implementation of Ubiquitous, Adaptive, and Immersive technologies in assessments, through the outputs of network analyses. This question is pivotal in identifying how ethical imperatives, among others, is an important theme in technology-enabled educational assessments. By examining these themes, the research contributes to an understanding of how ethical imperatives are to be aligned with advancements in educational technologies in assessment practices.,

Through network analyses, we have unearthed key research themes inherent to these technology groups, providing insights on existing research areas and emerging research gaps for researchers and educators alike. From these key research themes, common themes are further distilled; these include ethical complexities that underscore the deployment of education technologies in real-world assessment settings. This study, in essence, aims to chart the course for the future, highlighting the intersections of educational technology and assessment, and the myriad possibilities they herald.

The remainder of the study is structured in five sections: (i) section 2.2 delves into the literature review and conceptual framework of the study; (iii) section 2.3 details the methodology; (iv) section 2.4 reviews the results of the study, by providing a visual representation of the evolution of the main predicted technology groups, performing an in-depth network analysis of each technology group, detailing the breakdown of the trends of each technology group and discussing the common themes arising from the technology groups; (v) section 2.5 discusses implications and limitations of the study; and (vi) section 2.6 presents the conclusion, which provides direction for future research.

2.2 Methodological Approach and Conceptual Framework

2.2.1 Methodological Approach

Bibliometric analysis is a well-established quantitative method used to study the patterns of publication within a given field, topic, or set of topics over time (Hood & Wilson, 2001). It offers a robust mechanism to analyze the magnitude of research activity, the influential works and authors, and the overall trends in a specific domain (Mingers & Leydesdorff, 2015).

In this study, bibliometric analysis was particularly applied to the HR's predictions to gauge the volume and trajectory of research surrounding its identified technologies. This approach was grounded on the principle that the frequency and pattern of publications on a specific technology could offer insights into its acceptance, adaptation, and significance in the academic community (Zhang, Thijs, & Glänzel, 2016; Daim, 2006). Using Google Scholar, a recognized tool for comprehensive academic search, the study analyzed publications, citations, and keyword occurrences related to the identified technologies from 2011 to 2023. This timeframe aligned with the HR's period of interest, ensuring contemporaneity and relevance.

Network analysis, beyond its traditional applications in social science, has been increasingly recognized for its utility in mapping and understanding the relational patterns within large datasets (Borgatti & Halgin, 2011). In the context of educational technologies, it can unveil the relationships between various technologies, their applications, and their impact on educational outcomes (Brandes & Erlebach, 2005).

For this study, once the technologies were identified and coded from the HR, a network analysis was conducted to discern patterns, clusters, and relationships among them. Using the *CorTexT* platform, both Author Keywords and Indexed Keywords sourced from Scopus were processed, ensuring a rich and diverse dataset for analysis. The Louvain hierarchical community detection algorithm was employed to optimize modularity, a choice backed by its efficiency and accuracy in handling large-scale networks (Blondel et al., 2008).

Combining both bibliometric and network analyses presents a holistic methodological approach. While bibliometric analysis identifies and quantifies the trends, network analysis maps the intricate web of relationships, providing depth to the understanding of the evolution of educational technologies. This dual approach, grounded in the foundational principles of bibliometric and network analyses, offers a comprehensive and theoretically sound blueprint that other researchers can replicate. The coupling of quantitative trend analysis with relational mapping has been recognized for its potential in providing both breadth and depth in understanding academic landscapes (Börner, Chen, & Boyack, 2003; Leydesdorff & Rafols, 2012). This ensures the rigor and relevance of the study's findings.

2.2.2 Research Gap and Conceptual Framework

Over the years, the rapid development of educational technologies has garnered significant attention from researchers, educators, and policymakers. As technology becomes an integral part of the learning process, understanding its impact on assessment practices is essential.

While there is a wealth of research on the implementation and efficacy of various educational technologies, there remains a conspicuous gap in understanding how predictions from established reports, like HR, align with the evolution of assessment practices. Most studies either focus on the application of isolated technologies or delve into educational assessments without leveraging predictive reports as a foundational lens (Means et al., 2010; Selwyn, 2021). The absence of a systematic exploration of HR's predictions in the context of current and emerging assessment practices underscores the need for this research.

For the study's conceptual framework, two pillars are identified:

- *Historical Evolution of Technology in Education:* Drawing on the works of Collins and Halverson (2018) and Selwyn (2021), this pillar traces the trajectory of technological interventions in education, offering a backdrop to evaluate the HR's predictions.

- *The Interplay of Technology and Assessment Practices:* Grounded in insights from Black and Wiliam (1998) and Bennett (2002), this pillar delves into the nexus between technology and assessment, exploring how emerging technologies might reshape assessment methodologies to be more adaptive, relevant, and efficient.

By positioning this study at the confluence of these pillars, the objective is to provide an understanding of the prospective trajectory of educational assessments influenced by HR's predictions. This conceptual framework, while anchored in recognized literature, brings a novel perspective by intertwining HR's predictive essence with the ever-evolving realm of educational assessments.

2.3 Methodology

To obtain an understanding of the trajectory of educational technologies as presented in the HR from 2011 to 2023, this study delves into the technological metatrends. The primary aim is to trace the educational technologies that have impacted assessment practices over the stipulated period. The study focused on HR reports from 2011 to 2023. Each HR report employs a methodology that surveys experts to discern trends in the short term (year of the report), mid-term (2-3 years post-report), and long term (4-5 years post-report). However, from 2020 onwards, Educause altered this methodology, shifting from specific time horizons to broader discussions on emergent educational technologies and practices, as articulated by expert panelists.

Given the expansive nature of the research field, this study employed coding for each technology prediction and subsequently visualized these codes within technology groupings. Network analysis was subsequently executed to discern key research themes within each cluster, distilling principal research areas. A literature search, facilitated by Scopus, was conducted due to its informative content tagging by professional indexers, which was invaluable for network analysis (Campedelli, 2021).

This study utilized a blend of the methodologies applied by Martin et al. (2011), Dubé and Wen (2022), and Lim, Gottipati and Cheong (2022). The steps are as follows:

1. *Trend Analysis from HR:*
    - Capture educational technology and practice trends from HR within the stipulated period 2011-2023.

- Visually represent these trends, clustering technologies based on their inherent similarities, and provide a detailed view for each cluster.
- Exclude trends unrelated to assessments.
- Analyze the dominant metatrends and juxtapose them with assessment practices evident in the existing literature.

2. *Bibliometric Analysis from Google Scholar:*

- Measure research impact by determining the publication count for each pertinent technology group from 2011 to 2023 via Google Scholar.
- To enable year-to-year comparison, a weighting factor (WF_i) was introduced, defined by Eq. 1:

$$WF_i = \frac{\bar{p}}{p_i} = \frac{\frac{1}{N}\sum_{i=2011}^{2023} p_i}{p_i} \tag{1}$$

Where:

$\bar{p}$ is the mean publication count from 2011 to 2023.

p_i is the publication count for year *i*.

i = {2011, 2012… 2023}.

N is the total number of years.

- By employing this weighting factor, we ensure that years with fewer publications on assessment-related topics are not disproportionately disadvantaged.

3. *Network Analysis via CorTexT:*

- For each identified technology group, use the *CorTexT* platform (Breucker et al., 2016) to conduct text parsing and network mapping.
- Extract major thematic representations from both Author Keywords and Indexed Keywords sourced from Scopus.

The essence of network analysis is to visualize keyword themes in a clustered layout. Each keyword associates with specific counterparts based on proximity metrics. We employed the Louvain hierarchical community detection algorithm (Aynaud, 2020), which optimizes modularity. It assesses optimal linkage densities, considering intra-cluster and inter-cluster connections, and is efficient for large-scale networks.

2.4 Results and Discussion

2.4.1 Trend Analysis from HR

Using HR data, we identified the primary technology groups shaping assessment practices. Figure 1 depicts the trend analysis: the vertical axis corresponds to the HR report year, while the horizontal axis indicates the prediction period for technologies. It is noteworthy that before 2020, HR's methodology forecasted technology trends for the reporting year, as well as mid-term (two-to-three years) and long-term (four-to-five years) outlooks. However, starting in 2020, HR shifted to discussing only the current year's emerging technology trends, eliminating extended future predictions. The technology clusters are color-coded as follows: (i) ubiquitous technologies (in orange), (ii) adaptive technologies (in green), and (iii) immersive technologies (in blue).

The HR analysis reveals a consistent upward trajectory for adaptive technologies, with their mentions increasing over time. Recently, the emphasis on adaptive technologies has surpassed other technology groups. Immersive technologies, on the other hand, have had a steady, albeit sporadic, presence between 2012 and 2022. Ubiquitous technologies experienced peaks in HR mentions during two distinct periods – 2011 to 2012 and 2015 to 2019 – but remained relatively muted during other times. Detailed trend breakdowns are discussed in sections 2.4.3 to 2.4.5.

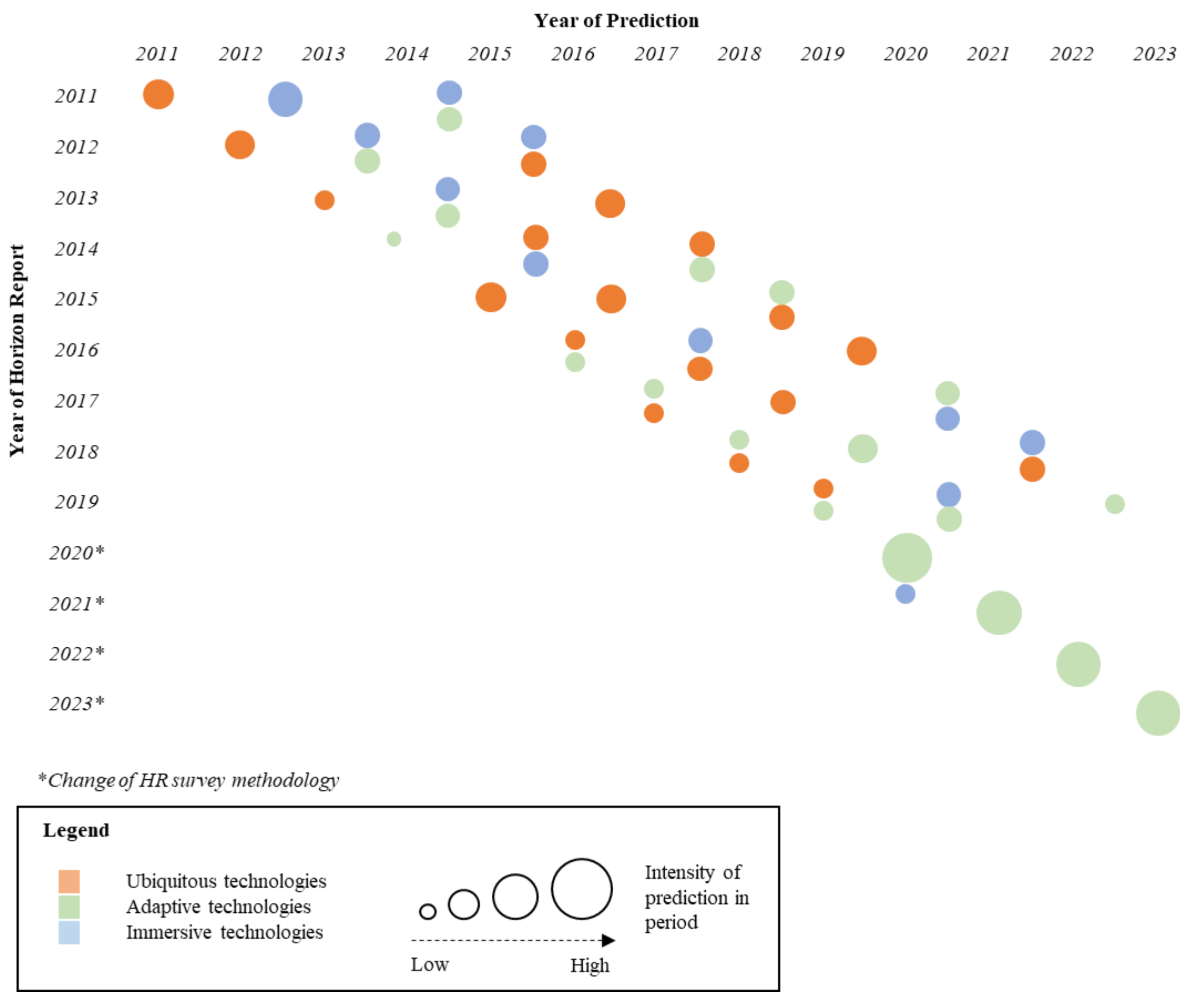


Figure 1: Education technologies impacting assessment practices based on HR from 2011 to 2023

2.4.2 Bibliometric Analysis from Google Scholar

A bibliometric analysis was conducted to determine the influence of the specified technology group on assessment practices.

Table 1 presents the annual count of assessment-related papers sourced from Google Scholar for each year under review. The corresponding weighting factors were calculated using Eq. (1).

Year, *i*	**Number of papers available, p_i**	**Weighting factor, WF_i**
2011	2650	1.61509434
2012	3030	1.412541254
2013	3270	1.308868502
2014	3570	1.198879552
2015	3650	1.17260274
2016	3710	1.153638814
2017	3950	1.083544304
2018	4220	1.014218009
2019	4710	0.908704883
2020	4890	0.875255624
2021	5730	0.746945899
2022	5560	0.769784173
2023	6700	0.63880597

Table 1: Number of published papers between 2011 to 2023, along with their weighting factor.

Table 2 presents the annual publication counts for each technology group, sourced from Google Scholar over each year under review. The data is adjusted using the weighting factor from Eq. (1). Figure 2 provides a visual depiction of the data presented in Table 2.

Year, *i*	**Ubiquitous**	**Adaptive**	**Immersive**
2011	291	556	1060
2012	358	570	958
2013	351	630	978
2014	334	644	908
2015	416	678	1012
2016	445	727	942
2017	425	759	902
2018	411	752	991
2019	419	783	926
2020	479	842	1015
2021	495	986	1003
2022	528	1063	1195
2023	466	1330	1101

Table 2: Number of published papers in each technology group between 2011 to 2023, adjusted by weighting factor in Eq. (1).

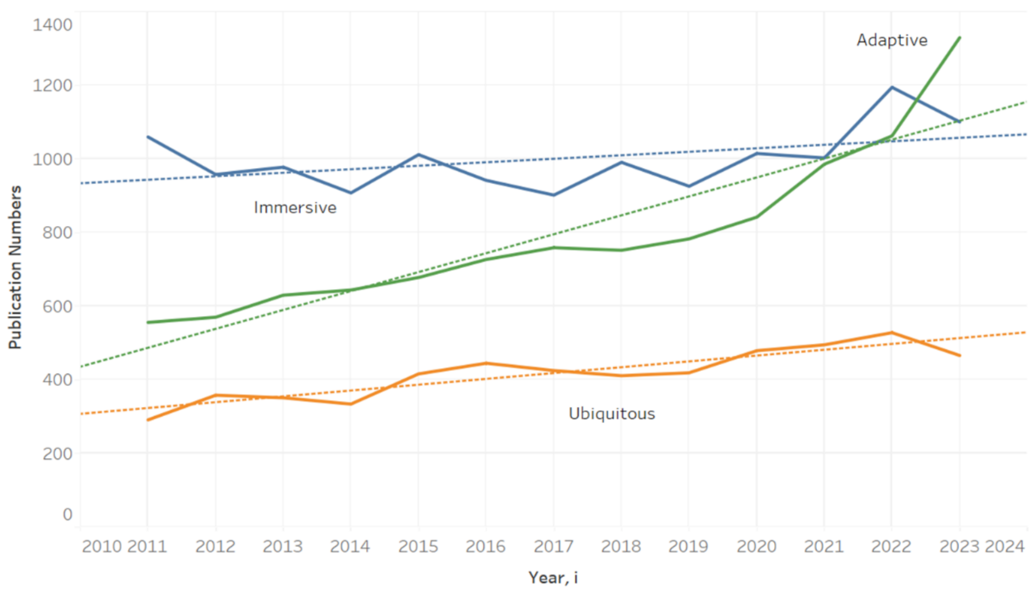


Figure 2: Number of published papers in each technology group between 2011 to 2023, adjusted by weighting factor in Eq. (1).

2.4.3 Addressing RQ1: Adaptive Technologies: Trend Evolution and Network Analysis

Adaptive instruction dynamically adjusts the curriculum, learning environments, or assessments to cater to individual student needs (Shute, Lajoie & Gluck, 2000). Central to this is the principle of AI, which seeks to create machines that can emulate human-like attributes such as cognition, learning, decision-making, and adaptation to environments (Chen, Chen, & Lin, 2020). A notable subset of AI, particularly relevant to adaptive instruction, is generative AI, which can create entirely new data that is mostly coherent and contextually relevant. In the context of adaptive instruction, generative AI could, for instance, craft personalized learning materials or questions based on a student's previous interactions, ensuring each learner receives assessment or feedback content that is tailored not just to their level, but also to their learning style and preferences.

Figure 3 details the evolutionary journey of adaptive assessments in both HR and published literature. Early in the prediction period, around 2014, there was a discernible emphasis on data-driven assessments. These encompassed predictive AI applications designed to glean teaching insights, monitor student progression, and refine learning methodologies. More recently, advancements in generative AI and adaptive virtual assistants have catalyzed an uptick in personalized and authentic assessment interactions, thereby enriching the student learning experience and outcomes. By 2023, adaptive technologies have become a focal point in HR and academic publications. Notably, while these technologies constituted roughly half of the research compared to the immersive technologies domain in 2011, their growth trajectory since has been nothing short of remarkable, as visualized in Figure 2. The growing emphasis of adaptive technologies in HR is matched by the steeply growing publication numbers in the scholarly community (Figure 3).

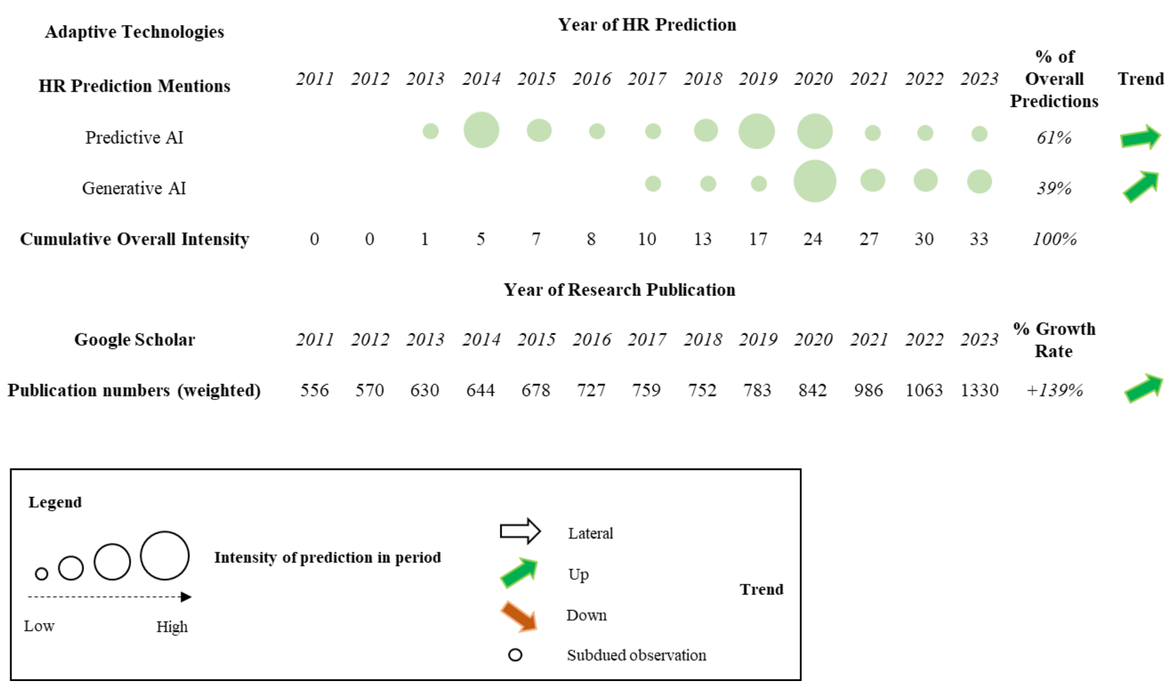


Figure 3: Breakdown of trends in adaptive technologies

The versatility of adaptive technologies spans a broad spectrum of assessments, evaluating a diverse set of multimodal learning evidence – from digital reports and code submissions to video presentations. Network analysis, depicted in Figure 4, segments the existing literature on adaptive technology-backed assessment practices into eleven distinct clusters, namely: (i) intervention and feedback, (ii) data, (iii) prediction, (iv) personalization, (v) behavior, (vi) theory, ethics and law, (vii) design, (viii) affect, (ix) administration, (x) grading, and (xi) robotics. A detailed exploration of these research clusters and examples of their associated literature are presented in Table 3.

Cluster Name	Cluster Description	Example(s) of Related Literature
Intervention and Feedback	*Bridging Analytics and Pedagogy* –Where predictive analytics meet actionable feedback, offering a roadmap for timely interventions and personalized student guidance.	*Cognii*, an AI-based Educational Technologies company, and Florida International University, partnered for the roll out of *Cognii VLA* to Information Systems Management students (King, 2019). *Cognii VLA* is an intelligent AI tutoring system that provides subject matter-based assessments and instantaneous chatbot-style feedback using natural language conversations to learners, while providing pedagogical insights and analytics to faculty members. Lim et al. (2022, 2023) addressed the issues of shortage of effective feedback, lack of freedom of choice, and lack of consideration of multimodality of digital education, using an analytics-enabled *GHMA* assessment format.
Data	*The Backbone of AI in Education* –Dive deep into the raw materials of AI-driven decisions,	Martinez-Maldonado et al. (2019) discussed challenges with collocated interaction data that can be used for predicting collaborative assessment activities.

Cluster Name	Cluster Description	Example(s) of Related Literature
	unraveling the intricacies of data processing and utilization.	
Prediction	*Anticipating the Learner's Journey* –Harnessing machine learning to foresee student pathways, opening avenues for pre-emptive actions in the learning process.	Ferreira-Mello et al. (2019) reviewed predictive text mining approaches that were applied in education, including assessments.
Personalization	*Crafting Unique Learning Experiences* – Tailoring assessment interactions for individual learners, marrying data insights with human-centric design.	Khosravi et al. (2022) utilized a few case studies relating to adaptive educational systems, to illustrate how learner feedbacks from assessments can be personalized and explainable, which can bolster feedback effectiveness.
Behavior	*Decoding Learner Psyche* – Going beyond scores to understand student behaviors, driving strategies to enhance engagement and learning outcomes.	Shabaninejad et al. (2022) developed an analytics platform known as *Student Inspection Facilitator*, to interpret learning behavior and identify at-risk learners.
Theory, Ethics and Law	*Navigating the Ethical, Philosophical and Legal Maze* – Charting the unexplored waters of AI's ethical, philosophical, and legal implications in education. A	Bearman and Luckin (2020) proposed evaluative judgement (or judgements on quality of work) and personal epistemology (or "*meta-knowing*"), as new models of assessment in an AI-enabled world.

Cluster Name	Cluster Description	Example(s) of Related Literature
	must for forward-thinking educators.	
Design	*Sculpting the Future of Assessments* – Reimagining assessment designs for the digital age, blending efficacy with modern media and tools.	Balderas et al. (2018) proposed a scalable assessment that applied learning analytics on Wiki-based assessment tasks. Mott et al. (2016) proposed a modular learning infrastructure, that included a digital learning environment, recommendation engines and competency-based assessments. The infrastructure incorporated data standards for interoperability, to support data exchanges in the infrastructure workflow. In an undergraduate sports technology-related course, Liu & Zhu (2020) designed an online analytics-based personalized adaptive learning evaluation tool, integrated with an intelligent teaching assistant.
Affect	*Emotions at the Forefront of AI* – A deep dive into the emotional dimensions of learning, offering a new lens to view affective AI's potential in social-emotional assessments.	Nazari, Shabbir and Setiawan (2021) used an AI-based digital writing assistant for formative assessment and feedback, to engage students with a variety of affective characteristics. The authors found a promotion of behavioral engagement and attitudinal technology acceptance of the writing tool.
Administration	*Guarding the Fortress of Learning* – Ensuring the seamless and secure administration of	Williams (2019) discussed the convergence of AI and blockchain technology, for instance, in secure automated

Cluster Name	Cluster Description	Example(s) of Related Literature
	adaptive assessments, from authentication to academic integrity.	credentialling of learners' assessments in authentic learning environment. Amigud et al. (2018) developed a cloud-based application known as *OpenProctor* which utilizes machine learning techniques to prove authorship assurance and identity of learners, using their textual contents. This can be used to manage academic integrity in remote authentic assessments.
Grading	*Redefining Evaluation in the AI Era* – Moving beyond traditional grading paradigms, leveraging AI's precision and scalability for unbiased evaluation.	Richardson and Clesham (2021) discussed the efficacy of automated grading systems to facilitate assessment, and shared a case study of automated grading using the *Intelligent Essay Assessor* scoring engine.
Robotics	*The Human-AI Interface in Learning* – Where tangible robotics meets abstract AI, creating immersive and responsive assessment environments.	Khairy et al. (2022) shared the value and the algorithmic characteristics of the use of assistive humanoid robots in the delivery of oral assessments. Lee and Yeo (2022) developed an AI-based chatbot that simulates open-ended, authentic interactions, for responsive teaching to promote arithmetic skills.

Table 3: Research cluster breakdown and related literature for adaptive technology-supported assessment

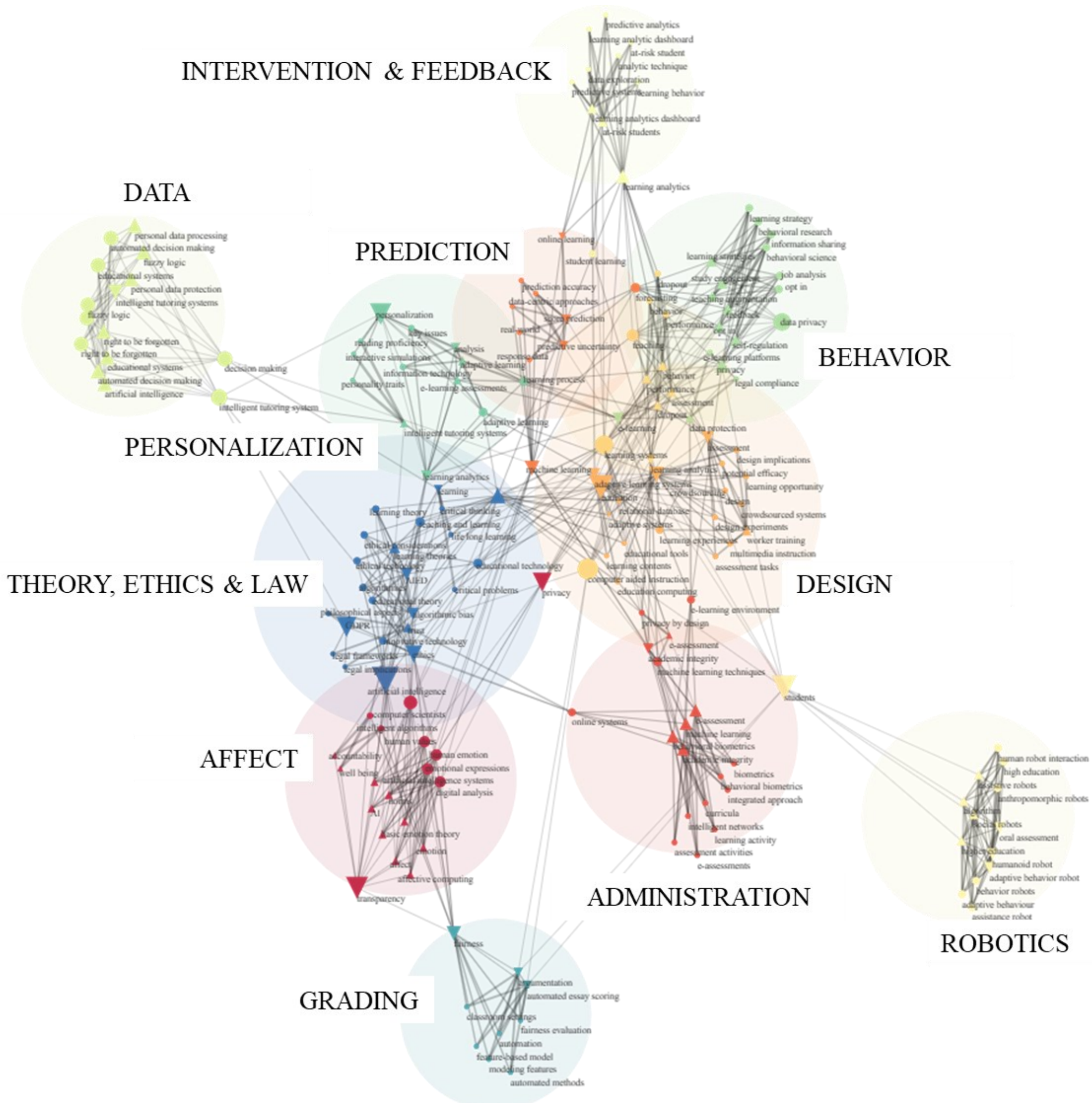


Figure 4: Network analysis on adaptive technology-supported assessment practices

Adaptive assessment brings about assessment environments that are not static but malleable, molding itself in real-time to the learner's needs and progress. Such dynamism ensures that learning is not just personalized in content but also in pacing. Drawing from the examples in the literature, we see instances where AI-driven systems, like Cognii VLA, are not just assessing but also offering real-time feedback, embodying principles reminiscent of Vygotsky's Zone of Proximal

Development (Vygotsky & Cole, 1978). Here, the learner is continuously nudged, challenged, and supported, ensuring that their learning trajectory moves forward.

Adaptive technologies have also shown promise in their capability to navigate complex human emotions and behaviors (e.g., Nazari, Shabbir, and Setiawan's (2021) exploration of engaging students with diverse affective characteristics). The interplay between AI and affective dimensions of learning underscores a crucial pivot - the shift from assessing mere cognitive outcomes to understanding, and catering to, the emotional journeys of learners.

The underpinnings of these systems lie in the presence of significant amounts of educational data, detailing not just learners' academic progress but also their behaviors, hesitations, and preferences. It is imperative to approach its collection and utilization with an ethical compass, ensuring the sanctity of individual privacy. The transformative potential of adaptive technologies also extends to the administrative aspects of assessments. Ensuring academic integrity in an increasingly digital environment is a challenge, but AI, as suggested by the works of Williams (2019) and Amigud et al. (2018), offers solutions that blend security with scalability. Such applications highlight the breadth of AI's impact, ranging from pedagogical interventions to administrative robustness.

The ethical dimensions of AI's integration into assessment practices, particularly in grading, present an ongoing debate. However, the advantages presented by adaptive technologies, especially in their precision in measuring tailored cognitive abilities and providing deep insights into learner progression, are compelling (Chassignol et al., 2018). Given the trajectory of current developments, it is reasonable to foresee an increasingly integral role for AI and data-driven evaluations in upcoming assessment paradigms.

2.4.4 Addressing RQ1: Immersive Technologies: Trend Evolution and Network Analysis

Immersive technologies allow the crafting of digital environments where learners not only interact but actively engage with digital entities, be it through avatars or embedded objects (e.g., Gaspar et al., 2020 and Dawley & Dede, 2014). Manifestations include immersive simulators (e.g., flight or medical simulators) and virtual worlds blending augmented and mixed realities, like the Lord of the Ring Online. The intrinsic value of these technologies is their ability to provide an unobtrusive yet potent medium for assessments. In particular, the emergence of stealth assessments, highlighted by Shute (2011), encapsulates the essence of this evolution. These assessments are founded on the evidence-centered design framework (Mislevy, Steinberg & Almond, 2003), ensuring that evaluations are both evidence-based and ongoing, seamlessly integrated into the immersive environments. For instance, "*Use Your Brainz*" (Shute, Rahimi & Emihovich, 2017) evaluates problem-solving competencies, while "Train B&P" (Li, Cheng & Liu, 2013) and "Triage Trainer" (Knight et al., 2010) assess computational problem solving and triage training, respectively.

Immersive technologies emerged as the second most prominent area of research interest at 2023, as visualized in Figure 5. While there is unmistakable interest surrounding the capabilities of immersive technologies in assessments, as evidenced by the consistent publication numbers (ranging around 900 to 1100 annually), there is also an emerging plateau. HR's consistent but relatively moderate emphasis on immersive environment-related education and assessments reinforces the flatlining bibliometric statistics. This saturation, while not entirely unexpected, might stem from the challenges the domain faces, such as technology (im)maturity, nuances of system testing specialization, and the complexities of skill transfer.

Delving deeper into the evolution of immersive technology trends, we note varied sub-trends over the years. The earlier period, particularly 2013 to 2015, saw gamification at the forefront, spotlighting the integration of game elements (e.g., badges, leaderboards) in formative assessments to invigorate student learning. This trend was gradually overshadowed by the advent of extended reality and human-computer interaction, especially pronounced in 2020 to 2021. As extended reality technologies, combined with intuitive user interfaces, gain traction, we foresee richer, high-touch assessment designs that cater to competency and skills-based pedagogies. The affordability of equipment, advancements in extended reality, and breakthroughs in network capabilities, such as 6G and WiFi 7, promise potent immersive experiences, even for distant learners.

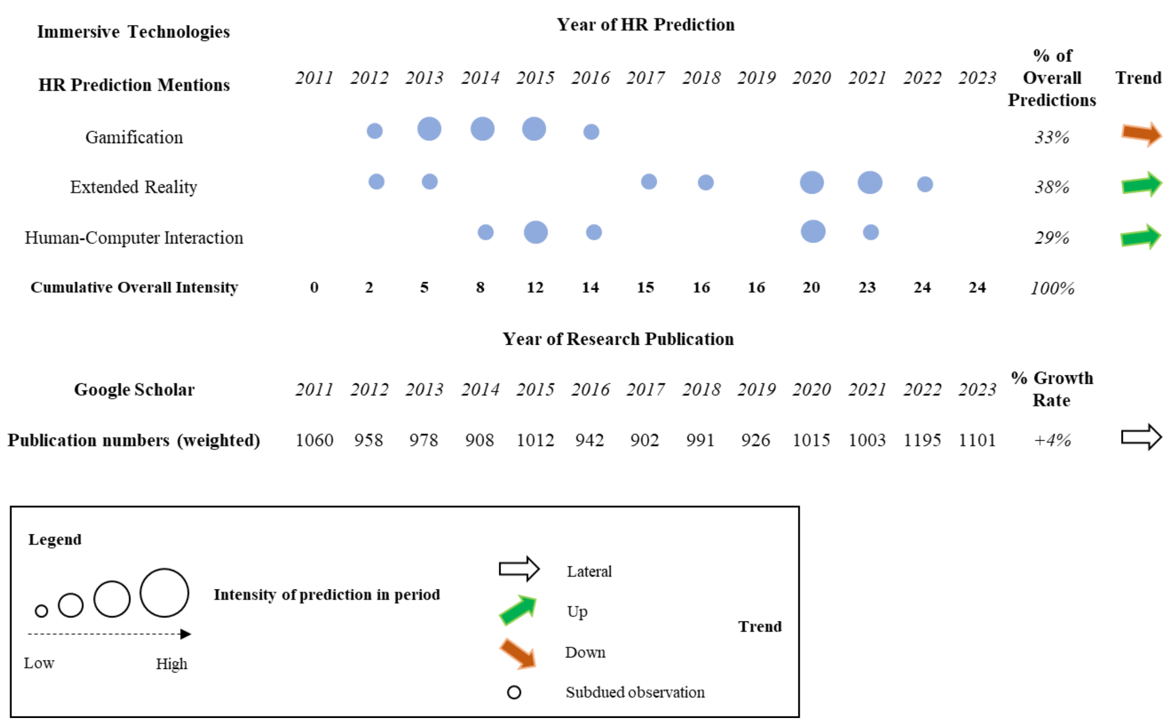


Figure 5: Breakdown of trends in immersive technologies

Network analysis reveals eight distinct clusters for assessment practices backed by immersive technology, detailed in Figure 6, namely: (i) vocation, (ii) interactivity, (iii) scenario, (iv) development, (v) program design, (vi) assessment design, (vii) engagement, and (viii) grading. Table 4 offers a breakdown of these research clusters and examples of their corresponding literature.

Cluster Name	**Cluster Description**	**Example(s) of Related Literature**
Vocation	*Real-world Simulations for Professional Training* – Dive into immersive assessments tailored for vocational training, blending realism with digital innovations.	To assess vocational cross-cultural competence of Special Operations Forces, Mateo et al. (2017) developed an authentic, mission-centric scenario-based training, in a simulated village with mock foreign players. To assess vocational competence for managerial training in a Brazilian bank, Cechella, Abbad and Wagner (2021) applied situational gamified assessments, incorporating a digital feedback system.
Interactivity	*Engaging Students at Every Turn* – Unlock the potential of interactivity in immersive environments, fostering active learning and self-assessment.	Robinson et al. (2021) utilized authentic and interactive immersive vignettes as a tool for formative assessments for a bioengineering ethics class. Philippe et al. (2020) shared that immersive and interactive virtual reality learning environments can help promote active student-centred learning and self-assessment.

Cluster Name	Cluster Description	Example(s) of Related Literature
Scenario	*Storytelling Meets Immersive Learning* – Harness the power of scenarios in immersive assessments, guiding learners through narrative-driven experiences.	Keast (2018) incorporated the virtual reality world of *Second Life*, as a formative assessment environment to promote scenario-based learning interactivity for online music courses.
Development	*Building Tomorrow's Immersive Learning Environments* – From concept to reality, explore the intricacies of developing immersive systems for authentic assessment.	Gerard et al. (2022) developed a prototype for a gamified virtual reality game to promote authentic assessing of crime scene evaluation and evidence-derived inferences in criminology courses.
Program Design	*Blueprints for Immersive Learning* – Delve into the architecture of immersive assessments, from data handling to algorithmic nuances.	Steynberg, van Biljon and Pilkington (2020) investigated the design aspects of virtual reality learning environment as a situated learning assessment space.
Engagement	*Magnetizing Learner Attention* – Unearth strategies to captivate learners in immersive settings, enhancing motivation, cognition, and creativity.	In safety management education for the construction industry, Fang and Goh (2022) demonstrated that virtual reality or mixed reality simulation-based activities help improve learner motivation.

Cluster Name	Cluster Description	Example(s) of Related Literature
Assessment Design	*Crafting Immersive Evaluation Landscapes* – A deep dive into pedagogical frameworks and design elements, paving the way for next-generation immersive assessments.	Ersozlu, Ledger and Hobbs (2021) studied pedagogies of practices as a framework for analyzing virtual simulation as a technological solution for assessments. Braunstein et al. (2022) developed a five-level taxonomy of social embedding, drawing on assessment research tied to the fostering of social dimensions in work-based learning.
Grading	*Measuring Success in the Immersive Realm* – Navigate the nuances of grading within immersive ecosystems, balancing traditional metrics with innovative methods.	Garvey (2022) reviewed grading and ungrading practices in gamified assessments (i.e., assessments infused with game elements) and game-based learning assessments (i.e., assessments using games as an environment for learning).

Table 4: Research cluster breakdown and related literature for immersive technology-supported assessment

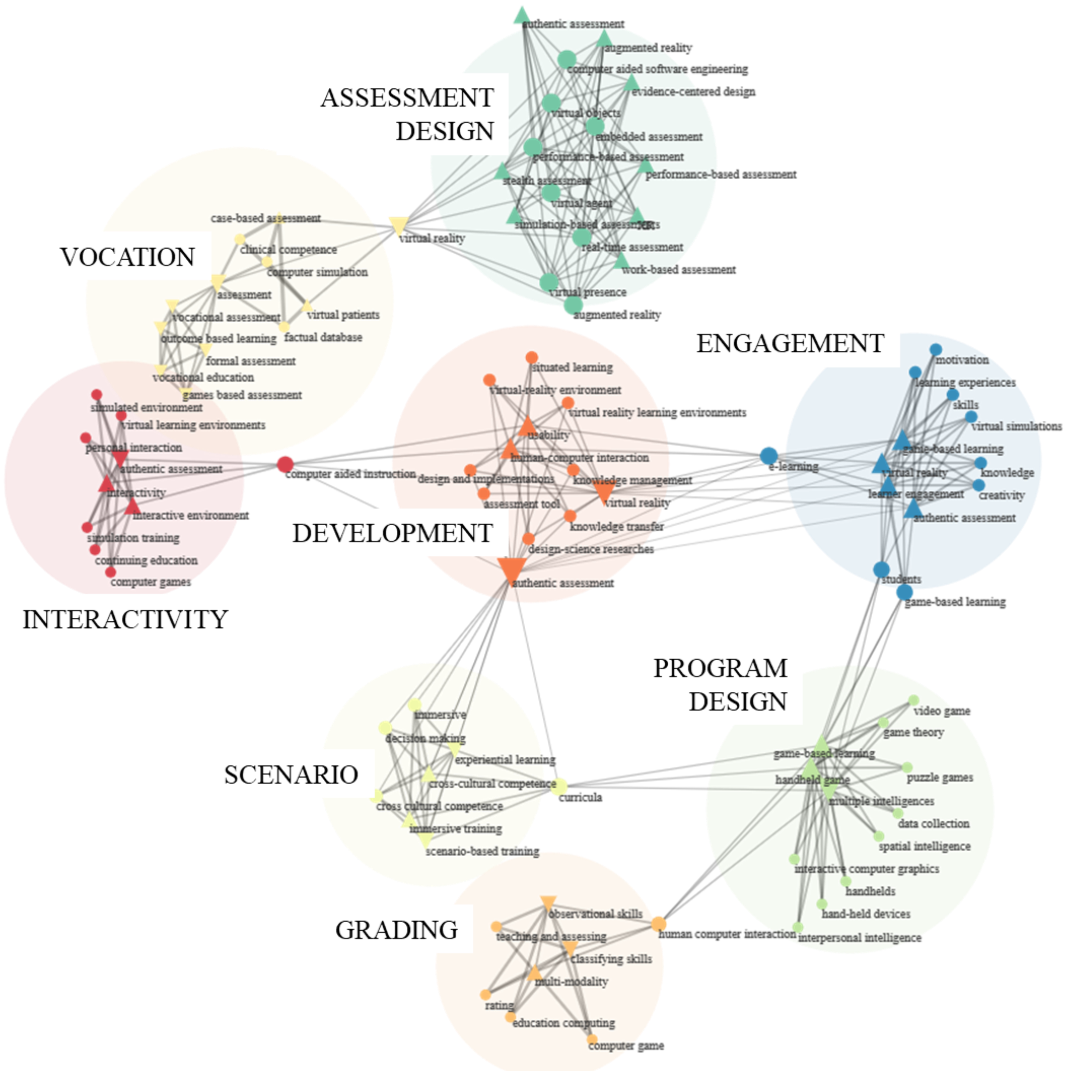


Figure 6: Network analysis on immersive technology-supported assessment practices

At its heart, immersive assessments break the barriers of a two-dimensional learning plane. They transport learners into intricately designed virtual realms, where they do not just respond but interact, engage, and influence. This is not about replicating real-world scenarios but amplifying them, adding layers of complexity, variability, and adaptability that a physical environment might not offer. The richness of these assessment environments – be it through the authenticity of a virtual vocational training scenario or the depth of an interactive bioengineering ethics class, as referenced in Table 4 – underscores the multiplicity of assessment avenues they open up.

What truly marks the distinction of immersive assessments is the shift from passive to active learning. Learners are no longer mere recipients of information or passive participants in a predetermined assessment task. They are active agents, navigating challenges, making decisions, and, importantly, learning from the consequences of those decisions. In these contexts, failure is not a dead-end but a learning opportunity, a moment to reflect, adapt, and iterate. This is underpinned by the tenets of Experiential Learning Theory, emphasizing the importance of experience in the learning process (Kolb, 2014). In immersive assessments, every decision, interaction, and outcome become a part of the learner's experience, adding layers to their understanding and skills. These are not mere testing grounds; they are rich, dynamic ecosystems where learning is continuous, iterative, and deeply contextual. Such a perspective aligns with the Constructivist theory, suggesting that learners actively construct knowledge from experiences, an idea that immersive environments appear tailor-made to support (Piaget, 1954).

The continuous stream of data such assessments generate – ranging from the overt actions learners take to the subtle hesitations they exhibit – offers a window into cognitive processes, decision-making patterns, and problem-solving strategies. This granularity has the potential to significantly augment learning evaluation, transforming it from a field that often dwells on outcome metrics to one that deeply understands the learning journey. The power of this data lies not just in its volume but in its potential to be leveraged for real-time interventions. Imagine a scenario where a learner's repeated hesitations at a particular juncture in a simulation triggers an immediate, context-specific support mechanism – be it in the form of a hint, a resource, or a scaffolded challenge. This continuous, real-time interventions align well with the principles of formative assessment in a micro or nano-second scale (and perhaps in a sub-conscious level), where feedback is immediate and iterative, enabling learners to recalibrate their strategies and understanding (Black & Wiliam, 1998). This dynamism, rooted in data, also embodies the principles of Vygotsky's Zone of Proximal Development, where feedback from adaptive learning can consistently challenge learners formatively just beyond their current level of competence, ensuring optimal growth (Vygotsky & Cole, 1978).

Harnessing the full potential of this data-centric approach requires robust technological infrastructure. Current LMS and educational environments may need significant overhauls to accommodate the high granularity and velocity of data from immersive environments. The sheer volume and diversity of data demands sophisticated algorithms and models to discern meaningful patterns, which may be alleviated to some extent by generative AI. Specifically, generative AI excels in creating complex, realistic simulations and scenarios that evolve in real-time based on the learner's interactions and decisions. This unique capability means it can generate entirely new content or questions that are tailored to the learner's current understanding, skills, and even misconceptions. Such a task is beyond the reach of traditional algorithms, as it requires the AI to not only analyze the data but to creatively construct educational experiences that are both relevant and challenging for each student. This adaptability ensures that learning is deeply personalized, engaging, and effective, bridging gaps in knowledge and skill in ways previously unattainable.

However, the deployment of such AI tools raises significant ethical considerations. The ethics of data – its collection, storage, and usage – emerges as a significant concern, especially in educational settings. There is a pressing need for transparency in how these algorithms operate, ensuring that educators and learners alike understand the logic behind automated interventions (Selwyn, 2019).

It is evident that the initial outlays in immersive technologies can be balanced by long-term benefits (Engelbrecht, Lindeman & Hoermann, 2019). The cost-effective adaptability of content, combined with reduced maintenance costs, portability, and minimal spatial demands, makes a compelling case for the incorporation of immersive technologies in educational assessments.

2.4.5 Addressing RQ1: Ubiquitous Technologies: Trend Evolution and Network Analysis

Ubiquitous computing, coined by Weiser (1991), envisions a world where computational resources enhance both human and environmental capabilities, offering authentic, situated digital assessment experiences.

Ubiquitous technologies ranked third in publication interest as of 2023. Publications numbers were relatively weaker throughout the period, matched by HR's subdued mentions of ubiquitous applications in assessments. This said, publications grew 53% over the study period, rising from 291 in 2011 to 446 in 2023, as depicted in Figure 7.

Early in the 2010s, mobile devices, exemplifying the melding of ubiquitous applications, sensors, and network infrastructure, were heralded as game-changers for delivering authentic assessments. By the mid-2010s, excitement peaked around the makerspace movement, 3D printing assessments, and the leveraging of IoT and wearable technologies for hypersituated authentic assessments. More recently, advancements in IoT, paired with pervasive robotics and AI, have enabled groundbreaking authentic assessments in sectors like ecology, transport, and medicine.

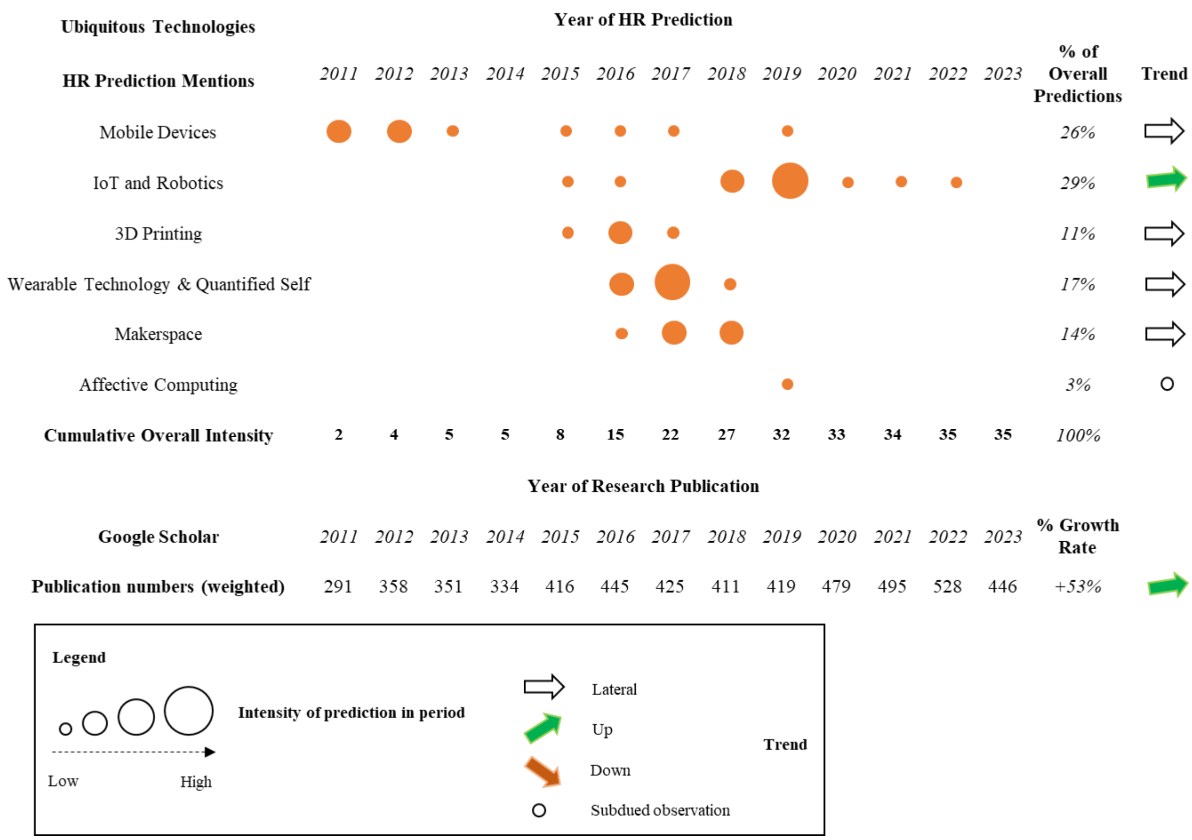


Figure 7: Breakdown of trends in ubiquitous technologies

Two prominent applications of ubiquitous learning technology in assessment are hypersituated assessments and makerspace-based assessments, detailed below:

Hypersituation: Hypersituated assessments are enabled by IoT, robotics, and wearable technologies. Johnson et al. (2015) describe hypersituation as leveraging ubiquitous technologies to amplify learning through continuous interactions with smart objects in learners' environments. These devices, part of blended learning, integrate real-world smart objects to provide valuable assessment feedback. Examples include the use of humanoid robotic patients for clinical evaluations or IoT systems for surgical technical skill assessments at the National Autonomous University of Mexico (Johnson et al., 2015).

Makerspace: This movement has rejuvenated interest in higher education through community-oriented spaces powered by ubiquitous technologies, for instance, ubiquitous fabrication tools (e.g., 3D printers), microprocessor-based mini-computers (e.g., Raspberry Pi), microcontroller boards (e.g., Arduino), and other electronic hardware, circuitry gadgets, manufacturing tools and software applications. Alternative terms with similar community-oriented physical tinkering space concepts are Hackerspace and Fab Lab (Hira & Hynes, 2018). Makerspaces emphasize hands-on, experiential learning, fostering creativity and innovation. Assessment methods in makerspaces range from practical evaluations using interviews and artifacts to knowledge tests (Lin et al, 2020).

Network analysis reveals eight distinct clusters for assessment practices backed by ubiquitous technology, detailed in Figure 8, namely: (i) software and platform, (ii) modality, (iii) development, (iv) design, (v) pedagogy, (vi) assessment strategies, (vii) administration, and (viii) geospatial. Table 5 offers an in-depth breakdown of these clusters and associated literature.

Cluster	Cluster Description	Example(s) of Related Literature
Software and Platform	*Beyond the Device: Learning Everywhere* – Unravel the power of mobile platforms in delivering authentic assessments, merging daily technologies with advanced methodologies.	Tong, An and Zhou (2020) used a messaging super-app known as WeChat to develop and implement an authentic technology-mediated assessment task-based language teaching. Tepper, Bishop and Forrest (2020) developed an online Student Clinical ePortfolio for the Bond University Medical Program, utilising a mobile-enabled, secure, digital platform available on multiple devices from any location allowing a range of clinically relevant assessments “at the patient’s bedside”. This can help provide evidence of multiple student-patient interactions and procedural skill competency. Prendes-Espinosa, Gutiérrez-Porlán and García-Tudela (2021) shared digital collaborative platforms that has been developed for group ubiquitous e-assessment implementation.
Modality	*Multifaceted Learning Interactions* – Delve into the diverse modalities of ubiquitous tools, capturing rich, multimodal assessment responses from learners.	Soto and Ambrose (2016) shared the use of screencasts in mobile devices to conduct multimodal formative assessments, that allowed educators to gain greater insights for targeted learning interventions. Richards (2012) explored a mobile device-based formative assessment to evaluate understanding of algebraic inequalities in daily tasks, through learners’ multimodal responses.
Development	*Crafting Ubiquitous Learning Ecosystems* – Chart the course of developing omnipresent	Candra et al. (2019) developed ubiquitous assessments using the five Borg and Gall stages, namely: (i) evaluation of product requirements, (ii) development of early-stage

Cluster	Cluster Description	Example(s) of Related Literature
	assessment systems, with a glimpse into enhancing them using AI.	products, (iii) validation of product from domain experts; (iv) test on small-scale field environment, and (v) test on large-scale field environment.
Design	*Building Inclusive Learning Spaces* – Focus on the design and accessibility elements that make ubiquitous assessments truly universal.	Macy, Macy and Shaw (2018) studied universal design elements of ubiquitous learning environment for assessments, that extended to learners with learning disability. Using the m-AssIST model, Santos, Cook and Hernández-Leo (2015) outlined the critical emerging properties that were useful to analyze and design mobile assessment activities.
Pedagogy	*Teaching in the Age of Ubiquity* – Unpack pedagogical methods tailor-made for the world of ubiquitous assessments, fostering genuine, hands-on learning.	dos Santos et al. (2016) used problem-based learning in a software engineering course to engineer mobile application solutions in an authentic assessment format. To improve community college library instruction, Viars, Cullen and Stalker (2017) applied an experiential learning technique on a collaborative rubric-driven assessment, based on an iPad, a topic and a database. The task completion accuracy and experience gained were comparatively better than previous non-ubiquitous assessment techniques.
Assessment Strategies	*Blueprints for Next-Gen Evaluation* – Explore a range of design approaches for assessment in a ubiquitous world, from	Martin et al. (2019) shared ubiquitous online assessment strategies of award-winning faculties, including the design of rubrics and timely feedback.

Cluster	Cluster Description	Example(s) of Related Literature
	format choices to feedback mechanisms.	Keast (2018) shared grading rubrics pertinent to assessments of online music courses. Formative assessments designed using a mobile application has helped students learn to compose and perform songs.
Administration	*Guarding Remote Digital Assessments* – Navigate the administrative challenges in ubiquitous assessments, ensuring security, integrity, and support.	Killam et al. (2022) explored how academic integrity can be maintained through open-web take-home exams.
Geospatial	*The World as a Classroom* – Immerse in the geospatial applications of ubiquitous assessments, transforming physical spaces into rich learning landscapes.	Fletcher, Kickbusch and Huijser (2022) developed a field-based project assessment using mobile devices for an introductory geospatial information science module, allowing authentic features, study areas and aerial photos that form the basis of the assessments.

Table 5: Research cluster breakdown and related literature for ubiquitous technology-supported assessment

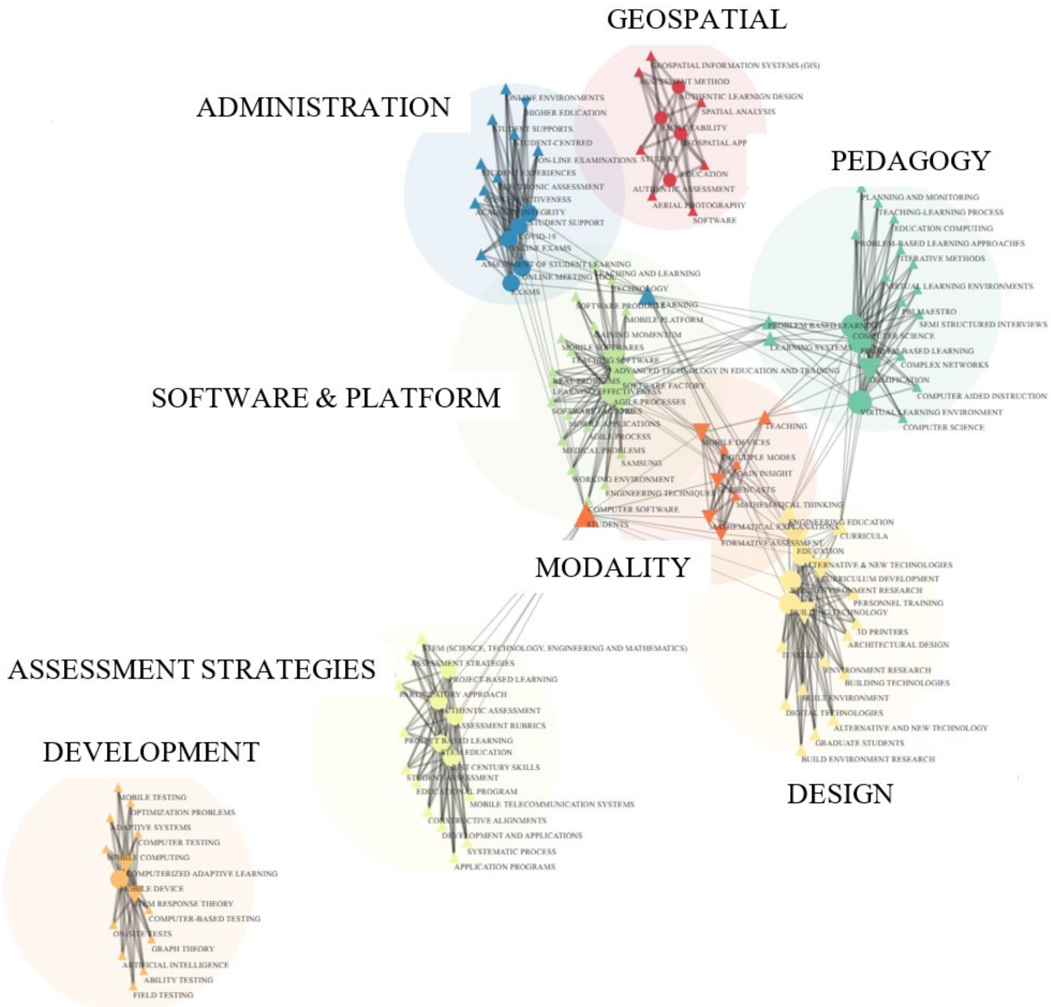


Figure 8: Network analysis on ubiquitous technology-supported assessment practices

Ubiquitous technologies have brought about a rethinking of educational assessments, pushing the boundaries beyond traditional classroom-bound assessments towards more dynamic, integrated, and contextually-rich forms. HR reports illustrate a growing interest in the fusion of mobile learning, IoT, and other ubiquitous technologies in education. Such integrations mark a departure from conventional assessment methodologies, opening avenues for remote, hands-on, experiential, and contextual forms of learner evaluation.

Grounded in the theoretical constructs of situated learning, this evolution emphasizes the importance of authentic, real-world contexts in learning experiences. Knowledge is not just acquired but lived, with assessments serving as integrated, contextual experiences rather than detached evaluative instances.

While platforms and software have been harnessed for their potential to dissolve the barriers between daily technologies and advanced pedagogical methodologies, the true essence of this transformation lies in the modalities they enable. The capacity to engage learners through diverse channels remotely beyond classroom settings – be it visual, auditory, or tactile – expands the horizon for capturing rich, multifaceted learner responses, providing a more holistic view of their understanding and capabilities. Moreover, the advent of geospatial applications in assessments exemplifies the transformative power of ubiquitous technologies. The physical world, augmented by digital layers, becomes a vast learning landscape, providing opportunities for authentic, context-rich assessments.

Growing research emphasis on inclusive design, universal access, and the logistics of assessment administration accentuates the complexities involved when designing ubiquitous assessments. The concept of inclusive design advocates for creating learning environments and resources that cater to the diverse needs of all learners, including those with disabilities (Burgstahler & Cory, 2010). In the context of ubiquitous assessments, this means ensuring that the technology utilized is accessible to everyone, regardless of their physical, cognitive, or socio-economic constraints. Universal access goes a step further, emphasizing not just accessibility but also usability, ensuring that all learners can effectively engage with the assessment tools and platforms (Seale, 2013). The logistics of administering ubiquitous assessments add another layer of complexity. With assessments potentially taking place anytime, anywhere, and on a range of devices, issues such as connectivity, compatibility, data security, and even the digital literacy of learners come to the fore (Crompton, 2013). This logistical challenge is further compounded when considering the diverse learning contexts, from formal institutional settings to informal, real-world scenarios.

While challenges to adopting ubiquitous learning technology assessments exist – potential cheating in remote settings, shifting social dynamics, and implementation costs (Sophonhiranrak, 2021; Asiimwe & Khan, 2013) – the benefits far outweigh the drawbacks. The upward trajectory in research interest underscores the growing enthusiasm around assessments linked to ubiquitous technologies. The journey ahead, while promising, demands a thoughtful, nuanced approach, ensuring that the essence of learning and assessment remains rooted in authenticity, context, and sound pedagogical principles. This domain is poised for greater exploration in the coming years.

2.4.6 Addressing RQ2: Common Themes Underlying the Use of Educational Technologies

These technologies, though differing in their specific applications and designs, converge on several key themes that are crucial in shaping the future of education. We explore these themes in Table 6 to understand their interconnectedness and implications.

Common Themes	Description	Implications
Ethical, Legal, and Accessibility Considerations	As these technologies develop, they bring forth important ethical imperatives and legal considerations, particularly around explainability, trust, data privacy and equitable access. Ensuring that these technologies are accessible to diverse learners, including those with disabilities, is also a key theme, promoting inclusivity in education and bridging educational gaps.	For *Researchers:* This theme necessitates research into ethical AI design, legal frameworks for data use in education, and development of accessible technology solutions. It also calls for studies on the impact of these technologies on diverse and marginalized groups to ensure equitable educational opportunities. For *Practitioners:* Educators and administrators need to be aware of and comply with ethical guidelines and legal requirements related to the use of these technologies. They also must advocate for and implement accessible educational technologies in their classrooms. This includes encouraging the provision of appropriate resources and support for students with disabilities, ensuring that no student is disadvantaged by the use of technology. For *Policy Makers:* The theme underscores the need for developing and enforcing policies that address the ethical use of technology in education, data privacy, and accessibility standards. Policymakers play a crucial role in creating frameworks that guide the responsible and equitable use of educational technologies. For *Technology Developers:* Developers of educational technologies must prioritize ethical considerations, data

Common Themes	Description	Implications
		privacy, and accessibility from the design phase. This involves creating user-centered designs that cater to a diverse range of needs and abilities, ensuring that technologies are not only innovative but also inclusive and fair.
Enhanced Personalization and Shift Toward Student-Centered Learning	At the heart of these technologies is the ability to personalize educational experiences. Adaptive technology achieves this through AI and analytics that tailor learning paths according to individual learner profiles. Immersive technology, through VR and AR, offers personalized, realistic scenarios for learners. Ubiquitous technology ensures personalization by making learning accessible across various devices and contexts, suiting the learner's lifestyle and preferences.	For *Researchers:* This theme calls for research into effective personalization algorithms, understanding how personalized learning impacts student engagement, retention, and outcomes. It also invites studies on the long-term effects of student-centered learning environments on various aspects of educational development. For *Practitioners:* Educators need to adapt their teaching methods to leverage these technologies effectively. This involves understanding each student's learning journey and using technology-driven insights to guide personalized learning experiences. Educators also need to foster environments where students are encouraged to take charge of their own learning, reflecting the shift towards student-centered approaches. For *Policy Makers:* Policymakers need to recognize and support the shift towards personalized, student-centered learning in educational policy and funding. This might involve creating frameworks that encourage the adoption of technologies in schools and ensuring that policies are in place

Common Themes	**Description**	**Implications**
		to support equitable access to these technologies. Administrators should advocate for and support the integration of these technologies into the curriculum. This includes encouraging the provision of resources and infrastructure necessary for personalized, student-centered learning and ensuring that teachers are adequately supported in this transition. For *Technology Developers:* Developers should focus on creating adaptive, immersive, and ubiquitous learning tools that are flexible and customizable to a wide range of learning needs and preferences. This includes incorporating user feedback into the design process to ensure that the technologies meet the diverse requirements of learners.
Increased Engagement and Enhancement of Active and Experiential Learning	These technologies are revolutionizing engagement in educational settings. Adaptive technology keeps learners engaged with interactive, AI-driven content that responds to their inputs. Immersive technology captivates learners through realistic simulations and interactive environments. Ubiquitous technology maintains	For *Researchers:* This theme opens avenues for research into how different types of engagement and experiential learning activities influence learning outcomes. Studies can explore the effectiveness of various interactive and immersive techniques in improving critical thinking, problem-solving skills, and overall academic performance. For *Practitioners:* Educators need to integrate these technologies into their teaching strategies to maximize student engagement and promote active learning. This might involve designing interactive lessons, utilizing VR/AR for simulations,

Common Themes	Description	Implications
	engagement by allowing learners to interact with educational content anytime, anywhere, thereby embedding learning into their daily lives. By incorporating real-world scenarios, problem-solving tasks, and interactive content, these technologies promote the development of higher-order thinking skills such as critical thinking, analysis, synthesis, and creativity.	and creating opportunities for learners to engage with content outside the traditional classroom setting. For *Policy Makers:* Policy makers should advocate for and support educational initiatives that prioritize engagement and experiential learning. This might include funding for technology integration in schools and development of curriculum guidelines that emphasize active learning. They should encourage the provision of necessary infrastructure and resources and create policies that encourage active and experiential learning approaches. For *Technology Developers:* Developers should focus on creating user-friendly, engaging, and interactive educational technologies. This includes designing immersive and interactive content that is not only educationally valuable but also appealing and stimulating to learners.
Data-Driven Insights and Analytics	The ability to collect and analyze data is a common thread. This data is used in adaptive technology to refine learning paths and feedback, in immersive technology to track engagement and learning outcomes in simulations, and in ubiquitous	For *Researchers:* This theme invites exploration into the most effective ways to collect, analyze, and interpret educational data. Research can focus on the development of sophisticated algorithms for personalized learning, studies on the impact of data-driven decisions on educational outcomes, and the ethical implications of data use in education.

Common Themes	Description	Implications
	technology to understand learning patterns across various contexts and platforms.	For *Practitioners:* Educators need to be skilled in interpreting data provided by these technologies to inform their teaching strategies. This requires a fundamental understanding of data analytics and its application in educational settings. Educators can use these insights to identify areas where students struggle and adapt their teaching methods accordingly. For *Policy Makers:* Policymakers play a critical role in creating guidelines and regulations around the ethical use of data in education. They should establish policies that protect student privacy while also promoting the effective use of data to enhance educational outcomes. Administrators should facilitate the integration of data-driven technologies in educational institutions. This includes investing in the necessary infrastructure and training for staff to effectively utilize these insights. Administrators also need to ensure that data collection and use adhere to ethical and privacy standards. For *Technology Developers:* Developers should focus on creating technologies that not only collect and analyze data effectively but also present it in a user-friendly manner to educators and learners. This involves ensuring data accuracy, reliability, and relevance to educational objectives.

Common Themes	Description	Implications
Professional Development and Teacher Support	As educational technology evolves, so too must the professional development of educators. Understanding and effectively implementing these technologies requires ongoing teacher training and support, ensuring that educators are equipped to leverage the latest tools, trends, and pedagogical strategies to enhance learning.	For *Researchers:* This theme calls for research into effective professional development models for educators in the context of emerging educational technologies. Studies can explore the most effective methods for training teachers, the impact of professional development on teaching practices, and the long-term effects on student learning outcomes. For *Practitioners:* Educators must actively engage in professional development opportunities to stay abreast of the latest technological advancements and pedagogical strategies. They need to be open to exploring new teaching methodologies and adapting their instructional approaches to leverage the potential of these technologies fully. For *Policy Makers:* Policymakers should advocate for and allocate funding towards professional development programs in educational technology. They should also develop policies that recognize and support the evolving role of educators in a technology-enhanced learning environment. Administrators need to prioritize and facilitate ongoing professional development and support for teachers. This includes allocating resources for training programs, providing time for teacher learning and collaboration, and fostering a culture of continuous improvement and innovation in teaching practices.

Common Themes	Description	Implications
		For *Technology Developers:* Developers should consider the needs and capabilities of educators when designing educational technologies. This involves creating intuitive, user-friendly technologies and providing comprehensive training materials and resources to support educators in using these tools.

Table 6: Common themes across educational technologies

The common themes identified – ethical, legal, and accessibility considerations; enhanced personalization and a shift toward student-centered learning; increased engagement and enhancement of active and experiential learning; data-driven insights and analytics; and professional development and teacher support – represent the pillars upon which the future of education will be built.

Each theme is interdependent, suggesting that advancements in one area will likely influence and drive progress in others. For instance, the ethical and legal considerations around technology use in education impact how personalized and student-centered learning technologies are developed and implemented without infringing the rights of humans. Similarly, the push towards increased engagement and active learning is intrinsically linked to the capabilities of data-driven technologies to provide insights that inform and shape these experiences.

The implications of these themes span across multiple stakeholders – including researchers, practitioners, policymakers, and technology developers – each playing a vital role in shaping the future of education. For researchers, there is a wealth of areas to explore that will provide deeper insights into the efficacy and impact of these technologies. Practitioners must embrace continuous professional development to effectively integrate these tools into their teaching. Policymakers and administrators are tasked with creating supportive frameworks and policies that facilitate ethical, equitable, and effective use of educational technologies. And finally, technology developers must continue to innovate while remaining steadfastly committed to creating ethically responsible and pedagogically sound tools.

2.4.7 Ethical Imperatives as a Key Theme Underlying the Use of Educational Technologies

Ethical imperatives form a key pillar in this discussion, underscoring the importance of responsible and equitable technology use in educational settings. The rapid advancement in adaptive, immersive, and ubiquitous technologies brings forth complex ethical dilemmas. Issues of data privacy, the explainability of AI decisions, trust in technology, and equitable access are paramount. These concerns are not just theoretical; they have practical implications that touch every aspect of educational technology deployment, from design to implementation.

Ethical imperatives extend into enhanced personalization and student-centered learning. As educational experiences become more tailored to individual needs through AI and analytics, the question of how much data is ethically permissible to collect and analyze becomes increasingly significant. The balance between personalization and privacy is delicate and requires careful navigation. Similarly, the drive towards increased engagement through interactive and immersive technologies raises questions about the ethical implications of such deep involvement in digital environments. The potential impact on cognitive and social development, especially in younger learners, calls for a cautious and thoughtful approach.

Educators and administrators must be vigilant in ensuring that the use of these technologies aligns with ethical teaching practices and fosters a safe, inclusive, and equitable learning environment. Policymakers and educational leaders play a crucial role in shaping and establishing clear guidelines and policies that govern the use of data, protect student privacy, and ensure that technology use is equitable and accessible to all students. Additionally, ongoing dialogue and collaboration between educators, technologists, policymakers, and the wider community are essential to navigate the ethical complexities presented by these emerging technologies.

As we venture further into integrating educational technologies in education, our compass must be firmly set on ethical considerations. It is through this ethical lens that we can harness the full potential of educational technologies to create learning experiences that are not only innovative and effective but also respectful, inclusive, and equitable for all learners.

2.5 Implications and Limitations

2.5.1 Theoretical Implications

The theoretical implications of a study reflect the broader contributions it makes to existing scholarly literature, frameworks, and conceptual understanding of a subject. Here are the theoretical implications that can be gleaned from this study:

- *Implications for Methodological Approach*

  Predictive reports, such as HR, offer prospective roadmaps for the trajectory of educational technologies. However, the real-world evolution of these technologies can diverge from predictions. By juxtaposing HR's predictions against actual technological advancements, this research underscores the theoretical significance of continuously revisiting and recalibrating predictive models. It is not merely about the accuracy of forecasts but about understanding the dynamics that influence the divergence. This offers a dual theoretical insight: first, the inherent fluidity and unpredictability of technological evolution in educational contexts, and second, the imperative for predictive models to be flexible, adaptive, and open to iterative refinement.

  In addition, the adoption of bibliometric and network analyses in this study is not just a methodological choice; it reflects an epistemological stance. By quantifying trends through bibliometric analysis and mapping relational patterns through network analysis, this research underscores the interconnectedness of the academic landscape. It posits that understanding educational technologies is not just about identifying isolated trends but about discerning the intricate web of relationships that bind these trends together. Such an approach echoes systems thinking, suggesting that the field of educational technology is a complex, interrelated system where changes in one node (or technology) can ripple across the entire network. The implication here is twofold: methodologically, it highlights the need for holistic, integrative research approaches; epistemologically, it calls for a systems-oriented understanding of educational technologies, recognizing their interdependent nature.

- *Implications for Educational Technology and Assessment*

  The integration of emergent technologies into assessment practices has necessitated a rethinking of traditional assessment paradigms. Historically, assessments were static, episodic events designed to measure knowledge at specific intervals. With the integration of immersive and ubiquitous technologies, however, this study suggests that assessments are evolving into dynamic, continuous processes. The transition from binary modes of "real vs. virtual" to a spectrum of realities challenges conventional notions of authenticity in assessments. This shift aligns with the principles of situated learning, wherein assessments become more contextually bound, capturing learning as it happens within both real and virtual environments (Lave & Wenger, 1991). The implication is profound: it calls for educators, policymakers, and curriculum designers to reconceptualize assessment strategies, moving beyond traditional methodologies to embrace more fluid, integrative, and contextually relevant approaches.

  Drawing from critical pedagogy (Freire, 1970) and actor-network theory (Latour, 2005), the integration of AI in assessments raises profound ethical and philosophical questions.

Critical pedagogy underscores the power dynamics in educational processes, and the introduction of AI amplifies ethical concerns, especially in terms of surveillance, data privacy, and learner autonomy. On the other hand, with AI systems gaining prominence in dynamically adapting and personalizing assessments, the traditional role of educators comes into question. This not only represents a methodological shift but also a philosophical one. The interplay between AI-driven assessments and human pedagogical interventions echoes the principles of actor-network theory, suggesting a reconfiguration of 'agency' in educational settings. Both educators and AI systems emerge as co-constructors in the assessment landscape, collaboratively influencing and shaping the learning outcomes. This co-construction necessitates a renewed focus on ethical pedagogy, ensuring that while leveraging AI's capabilities, the humanistic, ethical, and philosophical foundations of education remain at the core of educational assessments (Latour, 2005).

2.5.2 Practical Implications

Academic Implications and Research Gaps

Our findings present the evolving interplay between educational technology and educational assessments, emphasizing areas such as trend evolution and cluster analysis that may have been underrepresented or overlooked in prior academic literature. This study can serve as a foundational reference for future academic pursuits, guiding researchers towards areas ripe for in-depth exploration, whether it be in the realms of immersive, adaptive, or ubiquitous learning technologies.

As these technologies continue to shape educational assessments, it is imperative to distill critical research gaps from the existing literature to inform future research and practice (Table 7).

Education Technology	Emerging Research Areas	Key Research Gaps
Adaptive technologies	Ethical imperatives and implications	As assessments draw from behavioral analytics, biometrics, and AI-driven algorithms, it becomes essential to rigorously address issues related to data privacy, the integrity of consent, and broader ethical considerations, especially when focus is drawn towards deepening personalization.
	Advancing toward hyper-personalized evaluations	Current adaptive technologies are progressively moving towards individualized educational experiences. The confluence of biometrics with AI-driven methodologies paves the way for assessments aligned to both academic and physiological nuances. This evolution raises crucial inquiries regarding the pedagogical consequences of assessments responsive to real-time emotional and physiological parameters.
	Immediate feedback and cognitive implications	The integration of advanced systems, exemplified by platforms like Cognii VLA, accentuates the shift towards real-time feedback in assessments. The academic discourse needs to focus on the ramifications of such instantaneous feedback on learners' cognitive processes and overall psychological frameworks.
	Metacognition in AI-mediated learning environments	The rise of AI-mediated feedback mechanisms introduces a layer of metacognitive reflection for learners. Researchers should delve into the implications of continuous AI interactions on learners' self-awareness and introspection, with a particular emphasis on cognitive science principles and instructional methodologies.
	Decentralization of assessments through blockchain	The intersection of AI and blockchain technology posits the potential for a decentralized assessment framework. Such a landscape, characterized by

Education Technology	Emerging Research Areas	Key Research Gaps
		transparent and verified networks, prompts scholars to consider the implications for institutional roles and the integrity of educational outcomes.
Immersive technologies	Dimensions of immersive learning	Immersive technologies serve beyond creating mere simulated environments; they establish alternative dimensions for learning and evaluation. As the distinction between tangible and virtual diminishes, there arises a need to study the potential alterations in learners' assessment perceptions, particularly in terms of stress and performance in virtual settings.
	Time manipulation in assessments	The incorporation of extended reality within immersive technologies presents distinct time dynamics. Traditional assessments, confined by time limitations, contrast with immersive evaluations that can modify temporal experiences. The impact of such temporal variations on educational outcomes necessitates further research attention.
	Biotechnological convergence in assessments	Merging biotechnologies with immersive platforms introduces an innovative assessment paradigm. Environments responsive to physiological indicators such as heart rate or ocular movements suggest the emergence of emotionally-responsive evaluations. This intersection warrants detailed exploration, especially concerning ethical considerations and individualized assessment potential.
	Sociocultural considerations in virtual settings	Virtual settings, whether intentionally or inadvertently, integrate certain cultural and societal elements. The influence of these sociocultural elements on diverse learners within immersive assessments requires rigorous investigation, especially concerning the cultural implications of virtual evaluations.

Education Technology	Emerging Research Areas	Key Research Gaps
	Identity in immersive assessments	The use of avatars and self-representations in immersive technologies raises questions concerning identity and performance during evaluations. The interplay between a learner's tangible self and their virtual representation, and its subsequent effect on assessment perception and feedback, is an area ripe for academic scrutiny.
Ubiquitous technologies	Cognitive load analysis	Integrating sensors within ubiquitous devices to assess a student's cognitive load can facilitate adaptive assessment modulation. Physiological markers, such as pupil dilation or skin conductance, may serve as indicators of cognitive strain, allowing for real-time adjustment of assessment challenges.
	Multimodal sensory engagement	Broaden assessment modalities beyond the auditory and visual domains. Employing tactile or olfactory devices could offer differentiated assessment experiences, particularly in specialized areas of study like botany or chemistry.
	Physiological feedback for metacognition	Continuous physiological monitoring during learning activities can offer students insights into their emotional and physiological responses, fostering a deeper understanding of their learning process and promoting self-regulated learning techniques.
	Assessment of collaborative capacities	Utilizing ubiquitous technology to document and evaluate a student's collaborative endeavors, both in digital and physical domains, can offer a comprehensive understanding of their teamwork and communication competencies.
	Time-dependent performance analysis	Employ machine learning techniques to evaluate students' interaction patterns, determining optimal times for assessment based on their historical performance metrics at different periods.

Education Technology	Emerging Research Areas	Key Research Gaps
	Lifestyle-centric evaluation	Integrate assessment parameters within daily routines, evaluating competencies such as financial literacy from shopping patterns or geographical knowledge from travel behaviors.
	Geo-contextual assessments	Harness geospatial data to offer location-specific assessments, providing questions or challenges relevant to a student's immediate environment, especially during field trips or experiential learning activities.
	Emotion-based assessment modulation	Utilize sensors to detect students' emotional states, and modify assessment parameters accordingly, ensuring that the emotional conditions are conducive for optimal performance.

Table 7: Emerging research areas and research takeaways

Adaptive technologies bring forth considerations about immediate feedback and its cognitive ramifications, with a growing emphasis on hyper-personalized evaluations informed by both academic and physiological parameters. Immersive technologies challenge traditional academic notions, proposing a shift from mere simulated environments to alternate dimensions of learning and assessment. The temporal dynamics and sociocultural nuances introduced by these technologies call for rigorous research exploration. Ubiquitous technologies underscore the importance of cognitive load analysis, offering an opportunity to leverage sensors for real-time assessment adjustments based on physiological markers.

Practitioner Implications

Our research underscores pivotal implications for practitioners involved in educational assessments. Educational institutions, from primary to tertiary levels, can leverage the technological insights derived from this study to optimize their assessment techniques. By strategically integrating emerging technologies into their assessment frameworks, these institutions can enhance the efficacy and precision of their evaluation methods. Additionally, in light of the rapid technological advancements, there is an evident need for continuous professional development for educators. This will ensure that they remain abreast of current technological advancements and are proficient in incorporating them into assessment methodologies.

Societal Implications

From a societal perspective, the implications of our research extend to a diverse audience, including learners and educators. Technologies, notably those with adaptive functionalities, offer the potential to devise more tailored and inclusive assessment environments. These refined environments are capable of addressing varied learning profiles and assessment needs. Moreover, with the proliferation of immersive and ubiquitous technologies, the scope of educational assessments has expanded beyond traditional settings. This shift underlines the importance of accessibility and flexibility, allowing individuals to engage with assessment processes irrespective of temporal or spatial constraints.

Policy Implications

At a broader policy level, our research findings provide valuable insights that can inform and influence educational policy design and implementation. Equipped with the detailed insights from this research, policymakers can formulate comprehensive and adaptable assessment policies that effectively integrate technological advancements. As the significance and complexity of educational technology grow, it is crucial for governmental agencies and policy-making bodies to make judicious decisions regarding resource allocation to foster innovation in the realm of educational assessments. Furthermore, as technology becomes an integral component of the assessment experience, the development of stringent policies addressing ethical considerations, especially pertaining to artificial intelligence, data privacy, and potential biases in technological tools, is imperative.

2.5.3 Limitations

Firstly, while Scopus and Google Scholar serve as robust repositories for peer-reviewed literature (Fahimnia, Sarkis, and Davarzani, 2015; Rodrigues et al., 2014), they are not exhaustive. Benefits can arise from the use of other notable databases, such as Web of Science, ACM, IEEE Xplore, EBSCO Host, Wiley, SAGE Journals, and Taylor and Francis. Campedelli (2021) suggests that the overlap in publication titles between Scopus and Web of Science might be approximately 50% to 60%. Therefore, considering both databases could enhance the breadth of literature review. While this study offers a first assessment on this subject matter, subsequent research may consider the inclusion of Web of Science and other pertinent databases to ensure a thorough examination of the subject.

Secondly, while ubiquitous, adaptive, and immersive technologies have the potential to reimagine assessment practices, it is essential to understand that these technologies are mere instruments. Successful utilization of these technologies in assessment design, from both instrumentalist and relational viewpoints, should not be misconstrued as techno-positivism. It is paramount that the adoption of such technologies in assessment practices is approached with situated and disciplined circumspection for applicative usefulness, as emphasized by An and Oliver (2021). Grounded in the theoretical constructs of Technological Pedagogical Content Knowledge (or TPACK), any paradigm shift is less about the technology itself and more about the interplay of technological affordances, pedagogical strategies, and content delivery.

Crafting assessments integrating such technologies require a delicate balance. On one hand, the technology needs to be robust, reliable, and versatile, and on the other, the pedagogical underpinnings of the assessment need to be sound, ensuring that the technology serves to enhance, rather than overshadow, the learning experience. This intertwining of technology and pedagogy is at the heart of the TPACK framework. The TPACK framework, as proposed by Mishra and Koehler (2006), emphasizes the interplay between three core components: technological knowledge, pedagogical knowledge, and content knowledge. In the context of assessments, the TPACK framework suggests that effective assessments should not just leverage the best of available technology but should also be rooted in effective pedagogical strategies, all while ensuring that the content being assessed is relevant and meaningful. The framework serves as a reminder that while technology can greatly augment the assessment experience, it should always be in service of pedagogical aims and content goals (Koehler, Mishra & Cain, 2013).

2.6 Chapter Conclusion

This study embarked on a journey to understand the evolving landscape of educational technologies, specifically focusing on ubiquitous, adaptive, and immersive technologies, and their implications for educational assessments. By juxtaposing HR's predictions against actual technological advancements, our findings elucidated the dynamic and often unpredictable trajectories of these technologies in educational contexts.

Our research emphasized the prominence of adaptive technologies heralding the future of real-time, personalized feedback. Immersive technologies, on the other hand, are blurring the boundaries between tangible and virtual, redefining the dimensions of learning and assessment. Ubiquitous technologies, as underscored by our study, are pushing the envelope by integrating sensors and multiple modalities, offering assessments that are deeply rooted in learners' daily lives and environments.

The exploration of common themes – encompassing ethical, legal, and accessibility considerations; a shift towards enhanced personalization and student-centered learning; the promotion of increased engagement through active and experiential learning; the leveraging of data-driven insights and analytics; and the emphasis on professional development and teacher support – collectively form the core pillars of modern educational practices. They reflect the dynamic interplay between technology and pedagogy, underpinning the transformation and future direction of educational environments.

Central to this discussion is the paramount importance of ethical imperatives in the deployment and use of educational technologies. As we progress into an era where technology deeply integrates into education, the ethical implications encompassing data privacy, the explainability of AI, and equitable access, among others, cannot be overemphasized. These considerations are critical not only in shaping the development and application of technologies but also in ensuring that their integration into educational settings is conducted responsibly and justly.

The theoretical implications of our work highlighted the need for continuous recalibration of predictive models and advocated for a systems-oriented understanding of educational technologies. The study also resonated with key pedagogical paradigms, emphasizing the role of situated learning, critical pedagogy, and actor-network theory in the context of technology-integrated assessments. From a practical standpoint, our findings illuminate potential avenues for academic exploration, emphasizing research gaps and emerging areas in the domains of adaptive, immersive, and ubiquitous technologies.

This study's unique approach of contrasting predictive HR reports with actual technological evolutions offers fresh insights into the dynamics of technological adoption in educational settings. The synthesis of bibliometric and network analyses underscored the interconnected nature of academic trends, emphasizing the need for holistic research approaches. Moreover, the detailed cluster analyses presented in our study provide a guide for educators, practitioners, and policymakers, towards innovative, ethical, and effective assessment strategies. Our discussions on the theoretical constructs of TPACK further reinforced the intertwined nature of technology, pedagogy, and content, emphasizing that technology, though pivotal, should augment and not eclipse pedagogical aims.

While our study offers a snapshot of the current educational technology landscape, it is essential to recognize the continuously evolving nature of this field. Future research should consider the inclusion of other databases, such as Web of Science, ACM, and IEEE Xplore, to ensure a broader and more encompassing literature review. There is also a burgeoning need to delve deeper into the ethical, societal, and philosophical ramifications of these technologies, especially as AI and biometrics become more entrenched in educational settings. As technology continues its rapid advancement, subsequent studies could also focus on the real-world implementation and efficacy of the emerging assessment paradigms highlighted in our research. The horizon is replete with technological opportunities, and with careful navigation, the essence of learning and assessment can remain rooted in authenticity, context, and sound pedagogical principles.

Building on the insights from Chapter 2's exploration of the technological evolution within education, Chapter 3 delves into the critical domain of ethical imperatives in AI-driven assessments. It extends the dialogue to a crucial examination of how ethical considerations must underpin the integration of educational technologies such as AI, ensuring that these advancements foster equity and integrity in educational assessments.

Chapter 3. Ethical Imperatives of AI in Educational Assessments

The previous chapter highlighted the extensive growth of adaptive educational technology, and in particular, AI's role in augmenting educational assessments. Managing of ethical imperatives, a core theme that transcended all technological implementations in educational settings, also emerged. By maintaining a steadfast focus on these ethical principles, we can ensure that the integration of technology in educational assessments serves to enhance, rather than detract from, the overall quality and equity of learning experiences.

In a societal institution as fundamental as education, where teaching practitioners or researchers apply AI in academic processes such as assessments, it is important to study the divide between what may be ethically permissible and not permissible.

This study applied a systematic literature mapping methodology to scour extant research, so as to holistically structure the landscape into explicit topical research clusters. Through topic modelling and network analyses, research mapped ten key ethical principles to five research archetypical domains, and reviewed the contribution and intensity of these ethical principles in each thematic domain. The study extended this review, by mapping out ethics programs and activities that can be applied in practice, alongside their relevant underpinning theories.

The findings of this research look to provide researchers and practitioners the insights into the application methods of AI in assessments, and in particular, in terms of their intertwined ethical challenges and how these challenges may be addressed, for follow up studies.

3.1 Introduction

AIED is the machine mimicry of human-like consciousness and behavior to achieve educational goals, through the use of technology that allows digital systems to perform tasks commonly associated with intelligent beings.

Of the three pillars of education, assessment exists as an important component, alongside pedagogy and curriculum (Hill and Barber, 2014). Within the AIED domain, Chaudhry and Kazim (2022) scoured the landscape and concluded that assessment is one of the four key sub-domains in AIED, alongside learning personalization, automated learning systems, and intelligent learning environments. In an educational context, assessment refers to '*any appraisal (or judgment or evaluation)… of work or performance*' (Sadler, 1989). The infusion of AI in assessments has grown significantly in recent years. Research on assessments related to digital education in the higher education landscape showed that AI and adaptive learning technologies have tripled between 2011 to 2021 and is likely to surpass immersive learning technologies as a prime research area in the near future (Lim, Gottipati and Cheong, 2022, p. 5). Among stakeholders, there is a consensus positive view that "*AI would provide a fairer, richer assessment system that would evaluate students across a longer period of time and from an evidence-based, value-added perspective*" (Luckin, 2017).

Infusion of AI in assessments also brings along its own set of concerns. AI implementation comes with technical and operational issues relating to system implementation. Arguably, these challenges have relatively lesser grey areas to contend with, than the complication of navigating the parameters and boundaries of ethics. Evaluators, as practitioners of assessments, will need to acknowledge, respect, and uphold ethical principles that may plague the implementation of an AI-based assessment.

The research objective of this study is to examine the landscape of AI-related ethical imperatives for educational assessments, through the lens of a systematic literature mapping approach. A systematic literature mapping study is a study concerned with the mapping and structuring of a topical research area, the identification of gaps in knowledge, and the examination of possible research topics (Petersen, Vakkalanka and Kuzniarz, 2015). The research novelty and value of this work lies in the notable lack of research providing a holistic inspection and review of the aforementioned landscape.

This study investigates the following research questions:

1. *RQ3: What are the Key AI Use Cases relating to Assessments?*

   This question explores the various applications of AI in educational assessments, utilizing network analysis and topic modeling to identify dominant trends and areas of focus in the literature. Understanding the range of AI applications and their prevalence in research is useful for comprehending the scope of AI applications that need to be addressed.

2. *RQ4: What are the Key Ethical Principles Arising from the AI Implementations Relating to Assessments?*

   This question investigates the main ethical principles arising from AI use in assessments, using network analysis and topic modeling. The goal is to identify and categorize the ethical principles most commonly implicated in AI educational assessments, useful for comprehending the scope of ethical issues that need to be addressed.

RQ5: What are the Key Themes Inherent in the Consideration of Ethical Imperatives in Educational Assessments?

This question aims to identify and analyze the key themes between AI applications and ethical considerations in educational assessments, as found in the literature. By employing network analysis and topic modeling, the goal is to map out a generalizable framework, that will facilitate informed and ethical AI integration strategies in educational assessments.

3. *RQ6: What are Solutions and Interventions that were Proposed to Address Key Ethical Imperatives, and their Associated Underpinning Theories?*

   This question looks to identify and recommend mitigating solutions and intervention measures that can be put in place to address ethical issues, by looking at proposed and/or implemented actions in existing literature. This inquiry will contribute to a better understanding of the current solutions landscape, offering insights into existing strategies and suggesting directions for future research and application.

The significance of this research is, through a systematic meta-analysis of existing literature in the field, (i) understand and consolidate knowledge regarding what was previously explored relating to AI-based assessment methods and their interconnected ethical issues, (ii) provide a comprehensive and integrated inquiry into the association of the ethical problems faced, and the mitigation and intervention techniques applied to solve these problems, and (iii) identify potential future research topics in the field.

Results of this study identified five key research archetypical themes, with presence across the system layers of cognitive, information and physical domains of an AI-based assessment pipeline, namely: (i) AI system design and check for assessment purposes; (ii) AI-based assessment construction and rollout; (iii) data stewardship and surveillance; (iv) administration of assessments using AI systems; and (v) AI-facilitated assessment grading and evaluation. Ten AI ethics principles epitomize the key ethics considerations across each of the five research themes, each manifesting varying levels of importance.

This study provides a comprehensive treatment of this subject matter to date. We hope the findings of this research can provide researchers and practitioners the insights into the application methods of AI in assessments, especially in terms of their intertwined ethical imperatives and how these challenges may be addressed, for follow up studies.

The remainder of the study is organized as follows: (i) section 3.2 introduces the background of AIED, in relation to ethics and assessments, supported by a survey of the state-of-the-art; (ii) section 3.3 discusses the systematic literature mapping approaches undertaken, explains the machine learning methods utilized, discusses research validity and repeatability issues, and highlights limitations to the research; (iii) section 3.4 presents the tables and graphic visualizations from charting, coding, topic modelling, and network analyses, and provides in-depth analyses of the data.; (iv) section 3.5 aims to formalize the results into actionable formats that can be used by practitioners and researchers, and discuss practical and theoretical implications of the findings; and last but not least, (v) section 3.6 summarizes the key findings, impact of study, and closes with proposed future work that can be studied by practitioners and researchers.

3.2 Literature Review

3.2.1 AIED

Depending on the context in educational technology-related papers, the term 'AI' in education is commonly used broadly and interchangeably with the term 'adaptive learning technology'. It also serves as an umbrella term that includes 'learning analytics', 'educational data mining', 'educational data science', 'teaching analytics', 'data-driven decision-making in education', and 'big data in education' (Romero and Ventura, 2020).

It is useful to note that 'intelligence' in AI exists as a continuum. Chassang et al. (2021) describes how AI can be mapped into Bloom's taxonomy of learning, with 'Crystallized Intelligence' describing lower order thinking skills and 'Fluid Intelligence' describing higher order thinking skills. The former includes mainly supervised learning (or target-based prediction), with "*encoding capacity, middle-long term memory and ability to access memorized data in a logical way*", while the latter includes a higher level of intelligence abstraction, with "*the ability to solve new problems, use logics in new situations and identify patterns without necessarily having the prior experience of similar information or problems*".

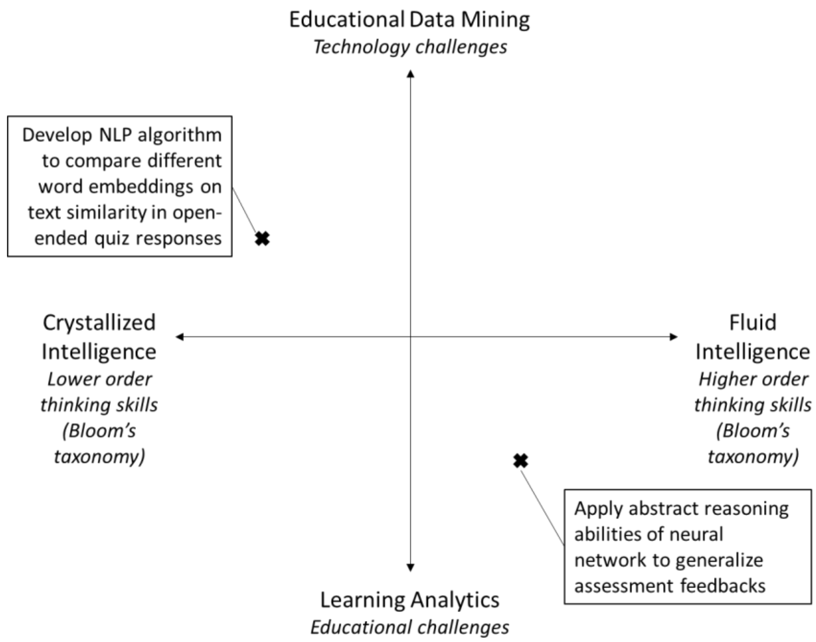


Figure 9: Relationship between educational data mining and learning analytics within AIED

AIED is a superset that encompasses terms such as educational data mining or learning analytics, depending on its specific use cases (Figure 9). Educational data mining is relatively more focused on the technological challenges of developing and applying data mining techniques in education. For instance, the use of educational data mining that involves developing a natural language processing algorithm to compare different word embeddings on text similarity in open-ended quiz responses is a use case of AI described as Crystallized Intelligence. On the other hand, learning analytics is more focused on the educational challenges of data-driven decision-making, through the use of predictive models. For instance, the use of learning analytics applying the abstract reasoning abilities of neural network to generalize assessment feedback will be a use case of AI described as Fluid Intelligence. There are many other use cases of AI, including visual data analytics for AI-integrated communication dashboards and recommender systems for formative assessment questions etc.

From a technology point of view, under the sub-fields of AI are machine learning, and in turn, deep learning and reinforcement learning. Machine learning is a field of inquiry that seeks to understand and create methods to leverage data and learn to make decisions. Deep learning leverages on a class of machine learning methods (specifically neural networks) to identify representation elements of a dataset, so as to learn features and perform tasks. Reinforcement learning seeks to utilize machine learning methods for intelligent agents to take actions in an environment to optimize some notion of reward.

A multitude of learning techniques and algorithms exist under each of these sub-fields. For instance, collaborative analytics refers to machine learning procedures performed to measure metrics tied to interdependent student relationships to predict collaboration dynamics in group assessments. A use case in Martinez-Maldonado et al. (2021) used natural language processing techniques, including a latent semantic analysis algorithm, to study the progression of collaborative critical thinking skills using online forum data. This computer-supported collaborative learning research, was underpinned by theoretical model of communities of enquiry, positing that meaningful learning on online forums occurs when there exist high levels of social, cognitive and teaching presence.

Generative AI models, which are designed to create new content or simulations, are increasingly being explored in educational settings. These models can generate personalized learning materials, simulate complex scenarios for assessment, and provide adaptive feedback, bridging the gap between traditional assessments and dynamic, individualized learning experiences.

AIED has grown tremendously in research intensity in recent years. It is observed that there is a significant increase in the number of research papers cited in Google Scholar; research publications increased from 1,739 papers in the decade of 1990 to 1999, to 22,060 papers in the most recent decade ending this year, representing more than twelve folds of increase.

3.2.2 AIED and Ethics

At present, although AI has yet to achieve comparability mimicking human levels of consciousness, it is still of urgent and paramount importance to consider ethical issues in AI applications, including AIED.

Aside from concerns regarding infringement of relevant laws and AI crimes (Sibai, 2020), ethical threats may exist, for instance, in the forms of systemic inequality and discrimination against marginalized learner groups in AI-driven assessments. Chaudhry and Kazim (2022) emphasizes that "*risks of AI going wrong have increased significantly for all stakeholders including, ed-tech companies, schools, teachers and learners…. a lot more work needs to be done on ethical AI in learning contexts to mitigate these risks.*" Prioritizing ethics is crucial to ensure the wellbeing of students, educators and other stakeholders involved in AIED.

Ethics is a branch of philosophy. At its core are the concepts of "good" or "bad", and "right" or "wrong". Ethics is closely tied to the study of values (i.e., axiology) and the study of taste and beauty (i.e., aesthetics). Ethics can be subdivided into three core research areas, which can be further divided into a multitude of sub-branches. These three core research areas are, namely:

1. *Metaethics:* Study of nature (i.e., moral ontology), meaning (i.e., moral semantics), and the scope and knowledge to defend or support (i.e., moral epistemology) moral judgments.

- *Moral ontology*: This investigates the nature of moral judgments. For instance, applying moral relativism in moral ontology, we may ask the question: "*Is the right or wrong of plagiarism necessarily contextualizable to societal conventions?*"
- *Moral semantics*: This evaluates the meaning and implications to the meaning of moral judgment. For instance, applying ethical naturalism in moral semantics, it may hold that the cognitivist ethical proposition of "*ensuring fairness in assessments is an ethically good act*" may be reducible and supervene into the natural property of "*maximizing happiness*", as embraced by utilitarianism.
- *Moral epistemology*: This studies and justifies moral knowledge. For instance, applying ethical rationalism in moral epistemology, the ethics of academic integrity is tied to the moral truths of justice and fairness, which are known by reasoning alone *a priori*.

2. *Normative ethics:* Study of the moral rules and standards that guide how individuals, institutions and societies should behave in a moral sense.

- *Virtue ethics*: This emphasizes the inherent disposition of an individual, and not specific actions. 'Good' in this context is the development of practical wisdom, and the flourishing of individual character and wellbeing. As such, morality becomes a holistic personal development process. Virtual ethics argues that virtuous individuals through good motivations can make good moral choices. However, the downsides are that individuals may not agree on what is good, and the existence of value plurality is reduced.

Here, it can be argued that ethics training for stakeholders implementing assessments in AIED achieves *phronesis* (i.e., acquiring practical wisdom to make 'good' decisions), and is sufficient in itself as an ethics stewardship measure.

- *Deontological ethics*: This emphasizes on an individual's rights and duties, including the presence of natural, absolute rights (i.e., natural rights theory), the presence of human rationality and inviolable moral laws (i.e., Kantian categorical imperative), and the morality of good actors arising from unbiasedness behind a veil of ignorance (i.e., contractualism). 'Good' in this context is the fulfilment and discharge of moral duties. As such, morality is focused on intention and obligation. Deontology argues that clear moral intuition and boundaries exist, even in cross cultural settings, although downsides are the lack of flexibility and possibility of conflict between human rights and moral duties.

  Here, ethics should be viewed from the lens of human rights, rationality and unbiasedness. For instance, the right to privacy may be viewed as an inviolable moral standard.

- *Consequentialism*: This emphasizes that the outcome of an action defines the morality of an action. Utilitarianism promotes actions that maximize happiness for the greatest number of people. Intellectualism promotes actions that encourage and cultivate knowledge. Situational ethics promotes *moral particularism,* which focuses on contextualizing actions that seek to engender love. 'Good' in this context is the actions that promote 'ideal' outcomes. As such, morality becomes results focused. Consequentialism argues a practical approach that is multi-perspective and objective, although the downsides may be an over-endorsement of value pluralism and the presence of adverse intended motivation behind an action that may use consequentialism to inappropriately justify their course of action.

Here, in the case of AIED, it can be argued that the impact to the relevant stakeholders should be measured, using appropriate metrics, as an ethics stewardship measure.

3. *Applied ethics:* Study of the practical application of philosophical tools to examine and provide solutions to real-world morality issues.

   This can be applied on scopes of digital ethics, defined as the "*attempt to guide human conduct in the design and use of digital technology*", and in narrower terms, AI ethics, defined as the "*attempt to guide human conduct in the design and use of artificial automata or artificial machines, or computers in particular, by rationally formulating and following principles or rules that reflect basic individual and social commitments and our leading ideals and values*" (Hanna and Kazim, 2021). On the subject of definition, it is useful to highlight the difference between AI ethics and ethical AI. Siau and Wang (2020) clarifies the former as "*principles, rules, guidelines, policies and regulations related to AI*", and the latter as "*AI that performs and behaves ethically*". The former relates to the behavior of humans, whereas the latter relates to the behavior of AI systems.

This study does not seek to argue the meta-ethics and normative ethics tied to assessments applying AIED. In this study, we take a more practical approach by considering fundamental ethical principles (e.g., fairness and trust) that inform the design, regulation and the use of AIED in assessments. These are ethics principles which provide concrete property instantiations of applied ethics, as opposed to abstract moral universals (Stringer, 2018). The study makes explicit these principles by describing applied instances of these principles found in existing peer-reviewed literature. In addition, the study cites practical solutions and mitigation measures that can be used to uphold these principles. It should be highlighted that while the study seeks to provide a generalizable approach to the consideration of ethical imperatives of AI in assessments, the application in specific domains (e.g., medicine and healthcare) may vary in breadth (e.g., safeguard human safety during assessments in medical field training) and are outside the scope of this study.

From the real-world practical application standpoint, we note that there is a tradeoff between the agenda of advancing AI technology, and the governance and stewardship of the use of AI in an ethical manner. Many professional and governmental bodies have pushed for responsible AI governance and stewardship. Siau and Wang (2020) identified eight institutions that have drawn up such ethics guidelines to facilitate the adoption, development and embracing of AI, including professional bodies such as the Institute of Electrical and Electronics Engineers (IEEE, 2019), and government bodies, such as the Australian government's Department of Industry, Science and Resources (Australian Government, 2019). Specifically, on the subject of AIED, Nguyen et al.

(2022) studied five such relevant guidelines, including UNESCO (2021)'s adoption on AI ethics guidelines with applications on AIED, and European Parliament (2021)'s report on AIED. These said, Siau and Wang (2020) notes that companies and institutions are presently more heavily weighted towards growing AI capabilities, with lower focus on ethical considerations.

Hinderance to the lack of focus on AI ethics in practice could stem from several reasons. Firstly, among nascent studies that investigate this tradeoff, Bessen, Impink, and Seamans (2022) studied the cost of integrating ethics in AI development, from a data management perspective. We note that further AI ethics research on cost-benefit analyses can be useful to help balance this tradeoff, and advance AI ethics governance and stewardship. Secondly, while the present AI ethics guidelines, especially the ones related to AIED studied by Nguyen et al. (2022), seek to address AI ethics issues, we note the lack of specificity on AIED applications, for instance, in assessments. Stahl, Timmermans, and Mittelstadt (2016) shares how ethics discourses should be "*focused on particular technologies to have practical importance*." In turn, Whittlestone et al. (2019) argues that these guidelines are "*not specific enough to be action guiding.*" This lack of idiosyncrasy and relevance can deter actionable applications.

To our best knowledge, a systematic literature mapping on ethical dimensions of the application of AI in assessments is lacking. This study aims to address the latter to enhance real-world adoption in the sub-domain of assessments within AIED.

3.2.3 AIED and Educational Assessment

There are different applications of AI in assessment practices. Sánchez-Prieto et al. (2020) presents a systematic literature review on AI-driven assessments, and subdivided assessments into three themes, namely:

1. *Assessment of student behavior:* This includes the contextualization of assessment delivery, and the prediction of assessment outcomes.
2. *Assessment of student sentiment:* This includes the personalization of feedback, and the analysis of socio-emotional elements.
3. *Assessment of student achievement:* This includes the automation of grading, and the categorization or profiling of students using data from assessment performance.

In a similar study, González-Calatayud, Prendes-Espinosa and Roig-Vila (2021) identified thematic uses of AI in assessments, mainly in individual or group adaptive formative assessments, automated grading and personalized feedback.

There exist different AI-driven assessment types, such as individual or group cognitive assessments and socio-emotional assessments. In addition, there exist different AI use cases across the assessment development and delivery pipeline, including but not limited to, assessment construction, curation and delivery, proctoring, grading, learning intervention and assistance, and feedback (assessment pipeline). From a technology system perspective, the AI system development pipeline includes "*the decision to start collecting data till the point when the machine learning model is deployed in production*" (Chaudhry and Kazim, 2022).

Different assessment types and areas of the assessment pipeline can be associated with different AI ethical concerns. For instance, in relation to assessment type, a socio-emotional assessment may be constructed in a manner that performs a semestral long behavioral tracking surveillance on learners, resulting in privacy infringement and anxiety disorders among students. In another example, in relation to the assessment pipeline, specifically regarding the design of an AI assessment system, data and/or modelling deficiencies may perpetuate stigmatization of minority group students, resulting in negative learning and psychological impact.

In this study, we seek to discuss the ethical dimensions of the application of AI in assessments across the assessment development and delivery pipeline, taking into account different assessment types.

In summary, in this literature review, the discourse on *AIED* provides an overview of the field of inquiry, the discourse on *AIED and ethics* provides the context and scope of ethics considerations within AIED, and the discourse on AIED and assessment looks at the applications of AI in assessment practices. The next section discusses the details of the methodology applied in this systematic literature mapping approach.

3.3 Methodology

In this study, we apply the systematic literature mapping approach. The study was conducted using the research methodology in Kabudi, Pappas and Olsen (2021), building upon the guidelines as proposed by Petersen, Vakkalanka and Kuzniarz (2015). We apply the methodology undertaken by both studies as follows, namely: (i) search and selection, (ii) data extraction, (iii) classification and analysis, and (iv) evaluation of validity.

PRISMA approach, or the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses approach*, was employed as a guideline to conduct the search and selection phase (Moher et al., 2009). In accordance with the recommended methodology as part of the PRISMA-P checklist, details including the eligibility criteria, sources of information, search protocol, research records, data items and synthesis of data are described in the following sub-sections.

NVivo11, EndNote X9 and Excel spreadsheets were used for information organization. Further information extraction, data visualization, and machine learning tools and techniques are described in the following sub-sections.

3.3.1 Search and Selection

As AIED researchers stem from a variety of fields publishing across a wide range of publications, literature search was conducted using Scopus, an interdisciplinary rigorously curated database covering the widest range of disciplines (240 disciplines) relative to similar citation databases, with contents including over 87 million publication items, 1.8 billion cited references, 17 million author profiles, 94,000 affiliation sources and 7,000 publishers. On average, each paper indexed on Scopus has 10% to 15% more citations than similar databases (Elsevier, 2022), which implies a more extensive systematic literature mapping analysis.

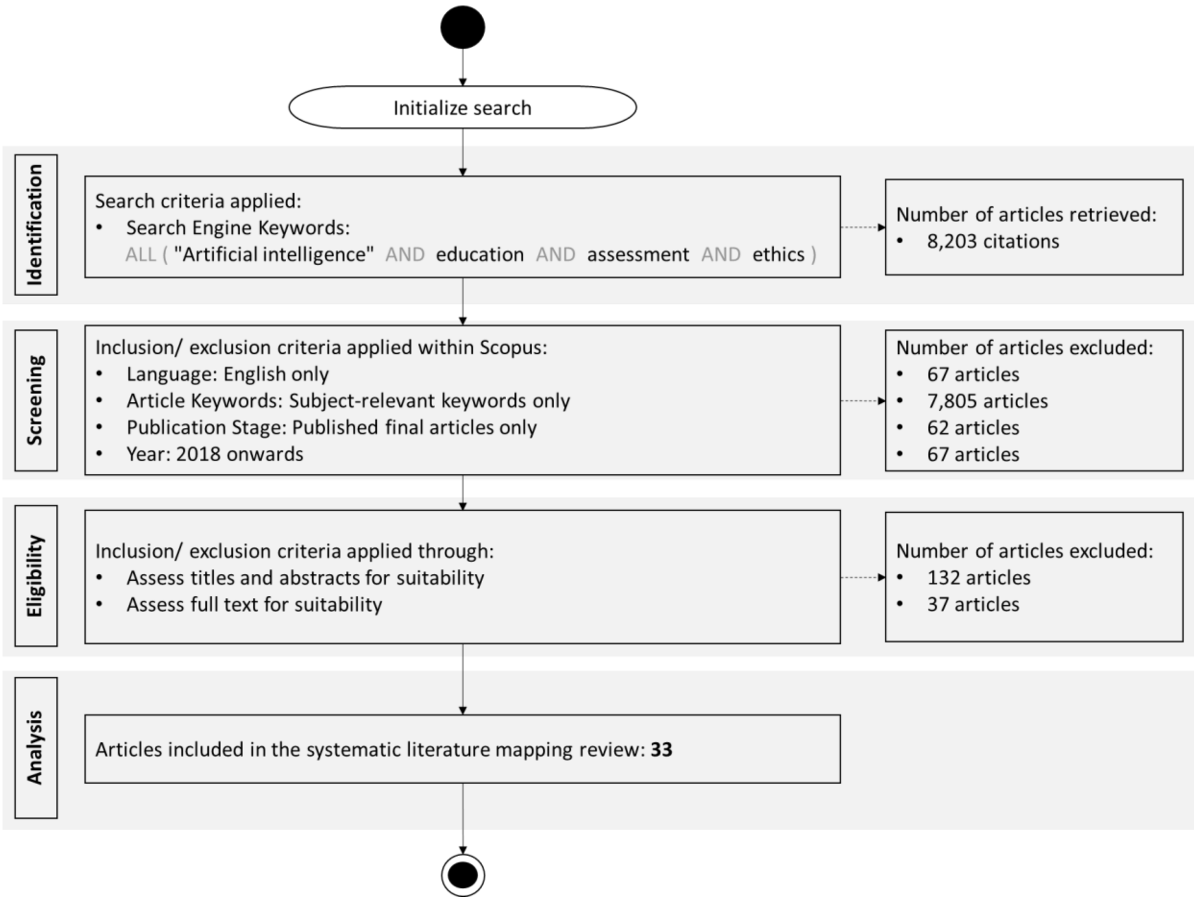


Figure 10: PRISMA - The systematic mapping process

The first stage of PRISMA, or the *identification* stage, identifies the possible papers to be considered using the Scopus search engine. The search entry was as follows: *ALL ("Artificial intelligence" AND education AND assessment AND ethics)*. This stage identified a corpus of 8,203 papers.

The second stage of PRISMA, or the *screening* stage, looks at excluding inappropriate and unrelated papers. This stage reduced the corpus count to 202. Search applied the following inclusion criteria:

1. *Language:* Only articles written in English language were included. This step omitted 67 articles.
2. *Keywording:* Only articles with subject-relevant keywords coded by Scopus for indexing purposes (also known as *Indexed Keywords* by Scopus) were included. Subject-relevant keywords included, and are not limited to: *Education, University, Higher Education, Learning Environment(s), Learning System(s), E-learning, Online Learning, Education Computing, Intelligent Tutoring System(s), Computer Aided Instruction, Learning Analytic(s), Curricul(a/um), Teaching, Learning, Learning Process(es), Collaborative Learning, Student(s), Academic Performance(s), Ethic(s), Ethical Consideration(s), Ethical Issue(s), Ethical Technolog(y/ies), Fairness, Data Privacy, Trust, Moral(s), Moralit(y/ies), Perception(s)*. Excluded keywords included, and are not limited to: *Medical Ethic(s), Bioethic(s), Clinical Stud(y/ies), Patient Simulation(s), Doctor-Patient Relationship(s), Risk Assessment(s), Human Resource Management, Software Engineering, Electronic Assessment(s)*. This step omitted 7,805 articles.
3. *Publication Stage:* Only peer-reviewed final articles published in scientific venues (e.g., books, journals and conferences) were included, for rigority of selection. This step omitted 62 articles.

4. *Year of Publication:* Only articles published in 2018 and beyond were included, to ensure recency of literature. Rigorous peer-reviewed articles would have reviewed key prior related literature within their respective papers. This step omitted 67 articles.

The third stage of PRISMA, or the *eligibility* stage, requires scanning title and abstracts, and full papers to identify relevant eligible articles. This stage yielded a final corpus count of 33 articles. Search applied the following inclusion criteria:

1. *Assess Titles and Abstracts for Suitability:* Only relevant titles and abstracts were included. There should be explicit and direct references to the subject matter. This step omitted 132 articles.

2. *Assess Full Papers for Suitability:* Only relevant full papers were included. An additional inclusion criterion here was that all articles should have their full text accessible for analysis. This step omitted 37 articles.

A summary of the PRISMA approach is shown in Figure 10.

3.3.2 Data Extraction

As a citation engine, data in Scopus is highly structured and robustly tagged, delivering metadata for analytical purposes, including (i) author(s), (ii) document title, (iii) affiliation(s), (iv) year, (v) publication, (vi) volume, issue and page source, (vii) citation, (viii) document type, (ix) keywords, and (x) digital object identifier (DOI), among others.

The final pool of 33 primary studies were thoroughly analyzed to answer the research questions of this study (refer to **Appendix I**). Information that was extracted from Scopus included: (i) citation information, such as author(s), title, year, publication, and citation count etc., (ii) bibliographical information, such as affiliation(s), and publisher etc., (iii) abstract, (iv) keywords, and (v) references.

3.3.3 Classification and Analysis

Using the data extracted from Scopus, the study utilized *Tableau Desktop Professional* version 2021.1.20 to perform exploratory data analyses to have a better understanding of the research scape. Tableau platform allows powerful conversion of complex computations into appealing data visualizations.

With the Scopus extracted data, research utilized a corpus analysis platform *CorTexT* (Breucker et al., 2016) to perform text parsing, and a first pass of topic modelling and network mapping, so as to identify major thematic representations of corpuses comprising of Author Keywords and Indexed Keywords. This allowed us to perform machine learning for pattern recognition, utilizing unsupervised text mining techniques on these keywords to identify useful patterns.

Using the Python Library *pyLDAvis* (Sievert and Shirley, 2014), topic modelling generated a topic representation of the keyword corpus' textual fields using the Latent Dirichlet Allocation method, which allowed a visualization of the most relevant words fitting to the topic. Here, each topic was defined as a keyword probability distribution, and each document was defined as a topic probability distribution. Given the total number of topics defined, the topic model was inferred by probabilistically assigning topics to documents, and positioned in 2D according to a multi-dimensional scaling algorithm for visualization purposes.

While topic modelling provided a sense of the latent themes from the underlying keywords, research further performed network analyses to visualize thematic keyword representations in a clustering format, where each keyword was grouped with distinct members, and linked via proximity measures. The Louvain hierarchical community detection algorithm was used (Aynaud, 2020). This algorithm is based on modularity optimization, where the optimal linkage densities are measured, taking into account within-cluster and between-cluster linkages. Louvain algorithm is efficient on large networks.

The first pass of topic modelling and network analyses above allowed the identification of distinct sub-themes of AI application areas and ethical issues. With the key sub-themes of AI application areas and ethical issues identified as a priori, each article was thoroughly evaluated and coded to classify the following: (i) application areas where AI is used in assessments (e.g., assessment curation and personalized feedback etc.), and the (ii) type of ethical issues relevant to AI-based assessments as cited in paper (e.g., fairness and explainability etc.). This would allow us to address RQ3 and RQ4.

The study then undertook a thorough review of the full papers, and provided further analyses to tabulate the following: (i) breakdown of each type of ethical issues identified in each paper (e.g., how explainability of AI systems is an important ethical consideration in assessments etc.), and (ii) breakdown of mitigation and intervention methods for each ethical issue as highlighted in each paper (e.g., applying data sanitization to reduce risk of discriminatory decision making from AI systems etc.). The paper-by-paper breakdown, as shown a tabular format in **Appendix II**, can provide useful research value-add to practitioners and researchers.

Using the coded sub-themes of AI application areas and ethical issues, research undertook the second pass of topic modelling and network analyses. The topic modelling and network analyses outputs would be used to guide the identification of the major research themes to address RQ5. From here, we performed further analyses to address the mitigation techniques and underpinning theories in RQ6.

3.3.4 Evaluation of Validity

In the application of systematic literature mapping, it was useful to consider the following types of validity to ensure that the methodology was robustly constructed. These included (i) descriptive validity, (ii) interpretive validity, (iii) theoretical validity, and (iv) generalizability (Petersen and Gencel, 2013). Detailed reporting of the systematic mapping methodology process, including the evaluation of validity, helps improve repeatability of the study.

1. *Descriptive validity*

   This describes the extent to which there existed objective and accurate observations. To lower the risk of this threat, a data extraction and coding spreadsheet was designed to support data recording. This provided objectification of the data extraction process, and allowed interventive correction to ensure accuracy, if required. As such, this risk was considered under control.

2. *Interpretive validity*

   This describes, given the data extracted and coded, the validity of the conclusions drawn. A key threat might be researcher biasness. This was alleviated by ensuring that no primary papers authored by the authors were included in the primary papers extracted, which reduced threats in interpretation.

3. *Theoretical validity*

   This describes the prospect of being able to capture what was purported to be captured. Research looked to ensure that the thematic phenomena identified in the study represented the patterns of the real world. Scopus provided a strong integration with major publishers, and its wide interdisciplinary focus ensured the lowering of probability of missing key research information. In terms of paper screening, careful curation of keywords, selection of final peer-reviewed papers published in scientific venues, and the recency of literature ensured that the literature reviewed was accurate, peer-reviewed and timely. Extensive in-depth reviews were also made in the full texts to ensure that each paper included was suitable. For quality assessment and to reduce potential biases, the methodology and data extraction process were checked by an independent external reviewer, with subject matter-relevant background.

4. *Generalizability*

   This describes the external validity (i.e., generalizability on the basis of repeatability and extendibility of results from this study to other research), and internal validity (i.e., causal effect between application of AI in assessments and their related ethical issues). In the presence of a wide range of similar ethical discourse on different AI applications, the classification and analysis methodology should not result in major threats for both internal and external validity. However, it is acknowledged that external validity may be influenced by factors such as domain-specificity (Leslie, 2019), cultural differences (Awad et al, 2018), and sample size (Khan et al., 2022), and future studies can help alleviate external validity concerns.

3.3.5 Limitations

Although Scopus is a robust database of digital records for peer-reviewed literature to map and survey specialized scientific areas (e.g., Fahimnia, Sarkis and Davarzani (2015), Rodrigues et al., (2014)), we recognize that Scopus is not the only one available. There are other valid alternatives, including Web of Science, ACM, IEEE Xplore, EBSCO Host, Wiley, SAGE Journals, and Taylor and Francis, among others. Among these databases, there are arguments brought forth by Campedelli (2021) that the degree of overlap between the publication titles in both databases may be closer to 50% to 60%, hence including both databases may have value. However, the discriminant feature that supported the selection of Scopus is that its informative tagging of all papers by professional indexers using Indexed Keywords existed at a higher frequency and provided a richer pool of content for each item, in particular, for textual mining purposes, as compared to Keyword Plus from Web of Science. As a first assessment on this subject matter, this should suffice. For follow-up works, it will be useful to consider integrating Web of Science and/or other relevant databases to provide a comprehensive scan of this landscape.

Secondly, throughout this study, there were no assumptions made on the intrinsic value of thematic diversity. This work utilizes unsupervised machine learning techniques to search for latent topics embedded in existing literature, and through which seeks to act as a key building block upon which future research can be applied. The study of possible inhibited or dysfunctional states within this thematic diversity, plausibly due to scholarly or technology inertia, or the lack of infrastructure or skills resulting in resistance to state-of-the-art adoption, are outside of the scope of this study. It may be useful for future research to quantify the value of this thematic diversity, in terms of (i) its operationalization impact as a segregable assessment pipeline component, and (ii) the extent to which ethics (or the lack thereof), either in isolation or in combination, impact upon the assessment pipeline component. Drawing from Klinger, Mateos-Garcia and Stathoulopoulos (2020), it may be useful to apply Weitzman (1993)'s economic valuation of ecological diversity, taking into account the cost-benefit analysis of preserving diversity and the threshold below which the archetypical research theme becomes unsustainable.

3.4 Findings

This section presents the findings based on an analytical investigation of selected published primary papers identified as relevant to the study.

3.4.1 Exploratory Data Analysis

Research undertook exploratory data analyses to explore where the studies discussing ethical issues relating to AI-based assessments arise from.

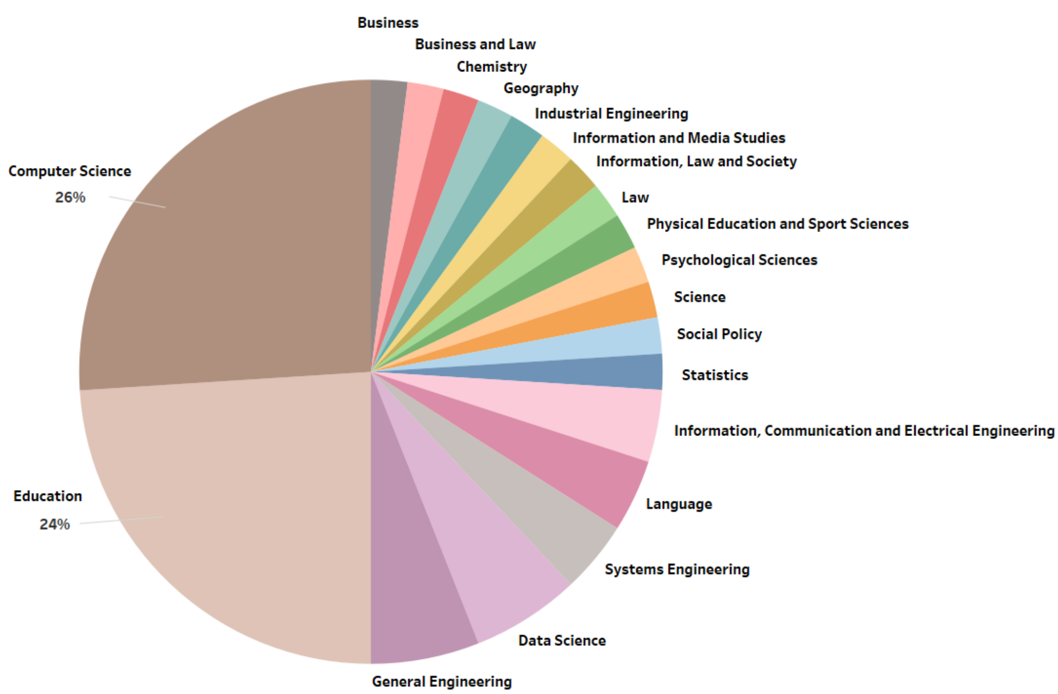


Figure 11: Breakdown of disciplines tied to authors' affiliated department

In Figure 11, it was noted that, through observing the authors' affiliated departments, the papers emerged from researchers across a wide range of disciplines. About half of the researchers originate from Computer Science and Education departments, with Computer Science leading marginally. The remainder stemmed from eighteen disciplines, ranging from soft pure disciplines such as Language, to hard applied disciplines like Engineering. This suggests that research pertaining to AI-integrated assessments and ethics do not necessarily emanate from the Education departments, but from a broad spectrum of disciplines.

In Figure 12, it was noted that most papers originate from The University of Melbourne, Australia (4); Carnegie Mellon, United States (3); University of Oulu, Finland (3); and University of Eastern Finland, Finland (3). The remainder are institutions that published two papers; the yellow circles denote institutions that published one paper. Papers from The University of Melbourne originated from Law, Psychological Sciences, and Engineering schools. Papers from Carnegie Mellon University originated from the Language Technologies Institute, Eberly Center for Teaching Excellence, and Human-Computer Interaction Institute. Papers from Oulu University originated from the Geography, Chemistry and Education departments. Papers from University of Eastern Finland originated from the Computing and Education Sciences departments. These provided an overview of the disparateness of affiliation and departments.

Subject matter is a complex multi-faceted issue that spans, among others, pedagogy, technology, and psychology domains. Taking cues from present research work, it may be useful for the nature of such research to have more pluralistic cross-disciplinary collaborations (e.g., computer science, education, social science etc.) in research and development work, to achieve sounder theoretically underpinned methodology approaches and more stakeholder inclusivity (Raji, Scheuerman and Amironesei, 2021).

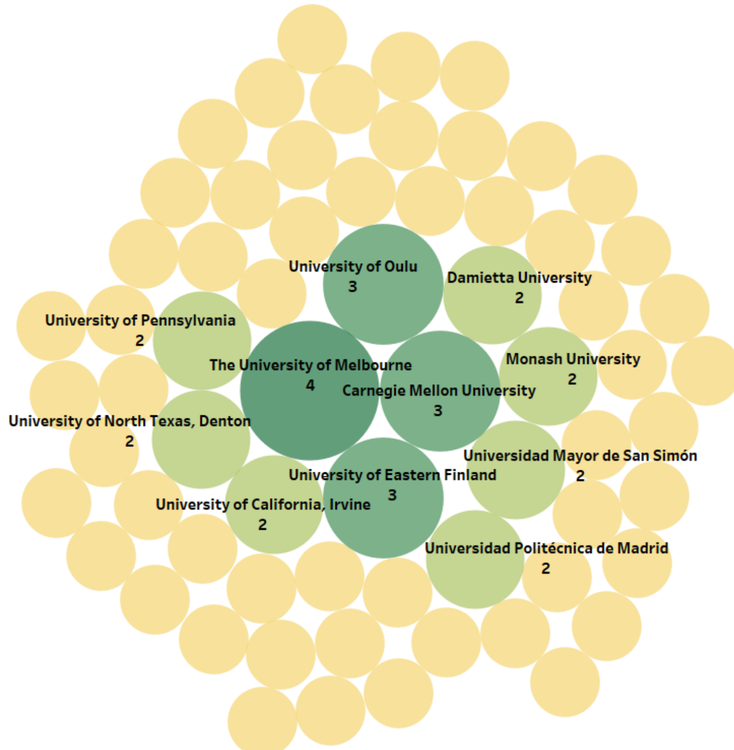


Figure 12: Breakdown of affiliated institutions

In Figure 13, it was noted that the top five countries where the papers originated were the United States, Australia, Brazil, Finland, and Spain (tied at fourth). Bozkurt (2020) found that countries that have historically dominated in educational technology research stemmed from the United States, United Kingdom, and Taiwan. In a different sub-domain in education, ethics-related education, for instance, in the field of healthcare, is also dominated by United States and Taiwan (Andersson et al., 2022). From these lists, aside from the United States which continued to lead in publication numbers, United Kingdom and Taiwan appear to be laggards. There are still room for different countries to play leading roles in this research aspect, as innovators, early adopters and early majority.

In Figure 14, it was noted that the top publication venues were Computer Science venues, namely (i) 4 conference papers published in Springer's *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, (ii) 3 conference papers published in Association for Computing Machinery's (ACM) *ACM International Conference Proceeding Series*, and (iii) 3 journal papers published in Institute of Electrical and Electronics Engineers' (IEEE) *IEEE Access*.

As a proportion, 55% of the papers were published in Computer Science venues such as those listed above, 24% were published in Educational Technology venues, such as the *British Journal of Educational Technology, International Journal of Artificial Intelligence in Education*, and *Journal of Learning Analytics*; and 15% were published in Education journals (mainly domain-specific education) such as the *International Journal of Information and Communication Technology Education, Journal of Information Systems Education*, and *International Journal of Engineering Education*. This suggests that Computer Science venues play a leading role (and may hold suitable target audience who can and are interested to participate) in the discourse of ethics in AI-based assessment practices.

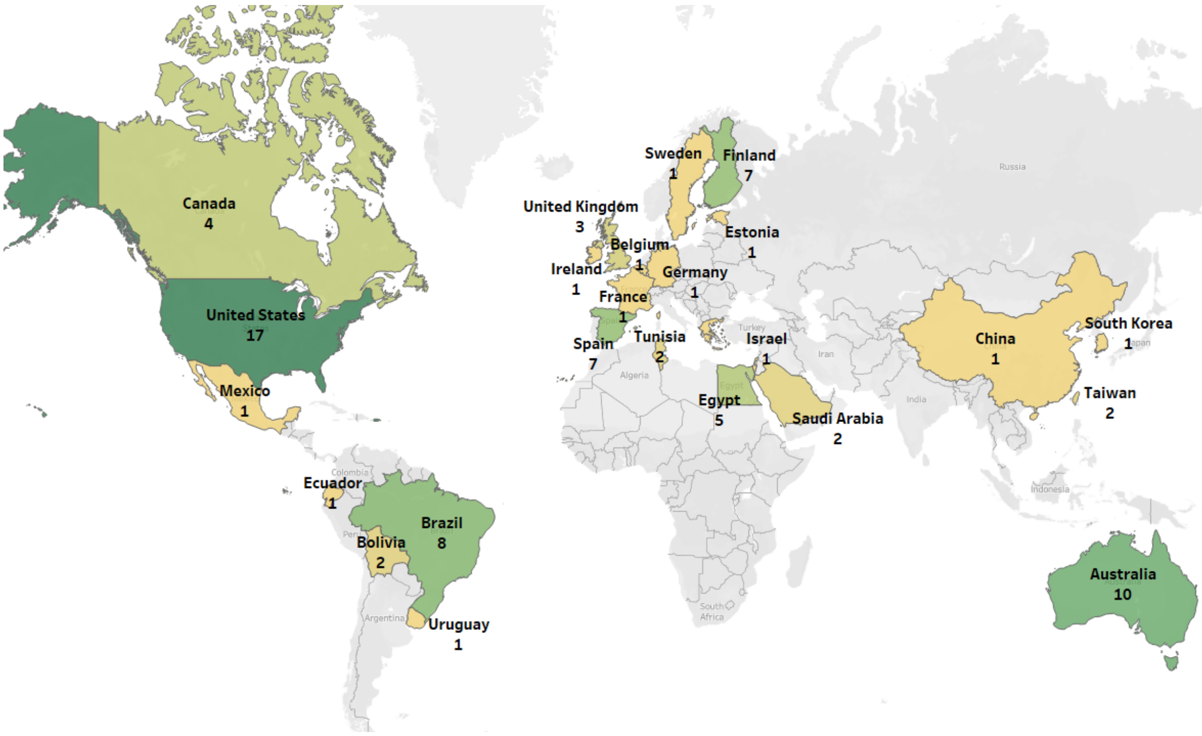


Figure 13: Breakdown of author locations

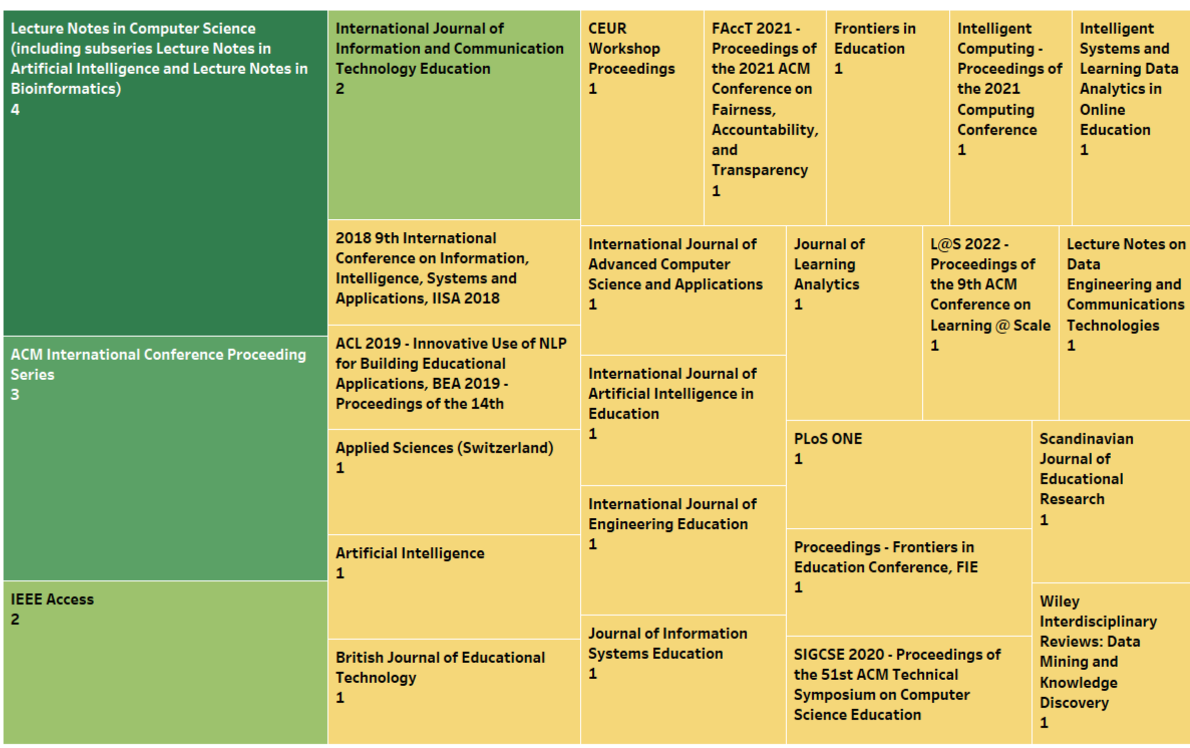


Figure 14: Breakdown of publication source

On an overall basis, the landscape is still at the early phases of the technology adoption curve, with enough room for researchers beyond the education domain. Present research is marginally led by the computer science discipline. However, research in this area is highly interdisciplinary, spanning from soft pure to hard applied disciplines, including but not limited to, education, computer science, educational and technology philosophy and psychology, social policy, and law. When applied on specific domains, it may also involve domain specific knowledge, such as sport science. Hence, a cross-pollination of ideas through inter-departmental collaborations can be highly beneficial to advance in this field.

3.4.2 First Pass of Topic Modelling and Network Analyses

Topic modelling was performed, where the optimal number of topics were generated using a model with the highest topic coherence. Further, we performed network analyses to identify topic clusters. These allowed us to recognize patterns in an unsupervised machine learning approach.

From this first pass of topic modelling, ten latent topics were identified. For instance, in Figure 15, we observed the latent topic of AI modelling and predictive analytics that might be relatively more closely linked to ethical issues of explainability and fairness. This mirrored well with the network analyses visualization in Figure 16. Here, we observed a more granular fourteen latent topic clusters, with the clusters of *Modelling*, and *Predictive Analytics and At-Risk Student*s, situated in close proximity in the cluster diagram. The higher granularity of the outputs allowed us to identify distinct sub-themes of AI application areas and ethical issues. The top keywords and latent topics of topic modeling are shown in Table 8.

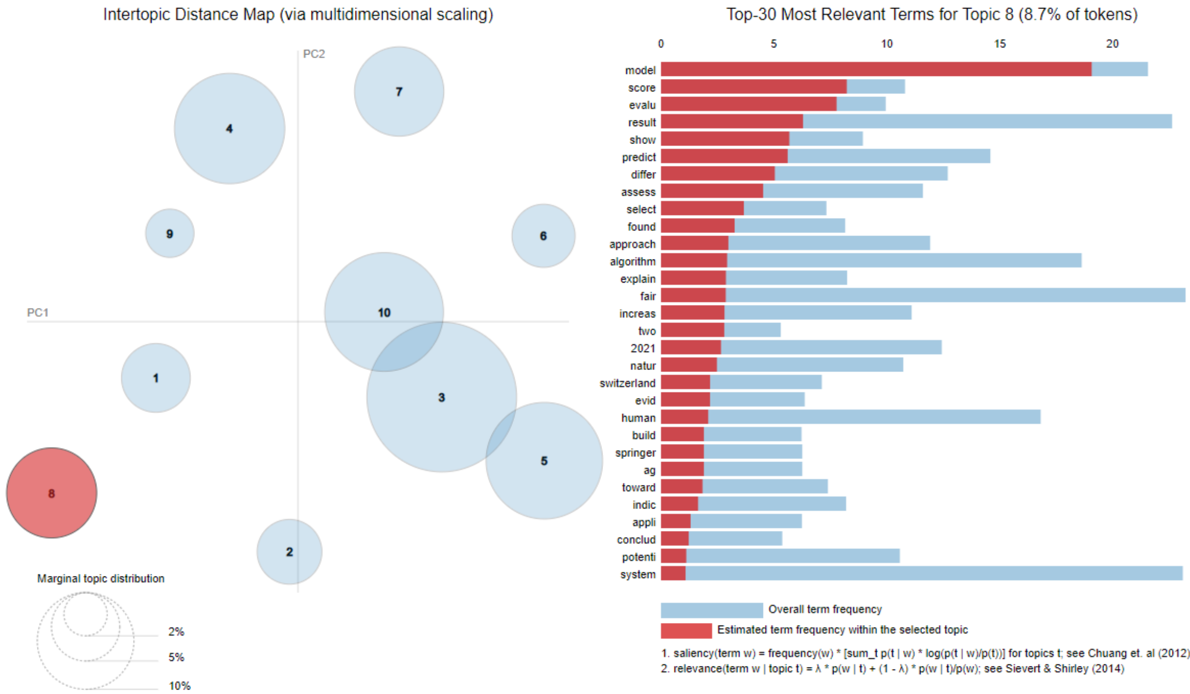


Figure 15: Topic modelling of keyword corpuses

Latent Topic	**Percentage of Tokens**	**Top Keywords**
Engineering systems	5.0%	Build; Design; Model; Technology; System
Automated grading	14.5%	Feedback; Response; Evaluation; Grading; Automation
Intelligent tutoring and feedback	23.9%	AI; Teacher; Intelligent; Support; Feedback
Predictive analytics and at-risk students	13.0%	Predict; Improve; Support; Measure; Environment
Explainable AI	4.5%	Explainable; Fair; Ethic; Algorithm; Review
Forecasting	4.3%	Forecast; Model; Automation; Learning; Predict

Adaptive robots	4.3%	Adapt; Environment; Tool; Robot; System
Modeling	8.7%	Model; Score; Evaluation; Result; Predict
Security and cheating	2.50%	Privacy; Data; Institution; Risk; Ethics
Assessment tasks	15.10%	Curation; Delivery; Task; Construct; Generate

Table 8: First pass of topic modelling – Latent topic and top keywords

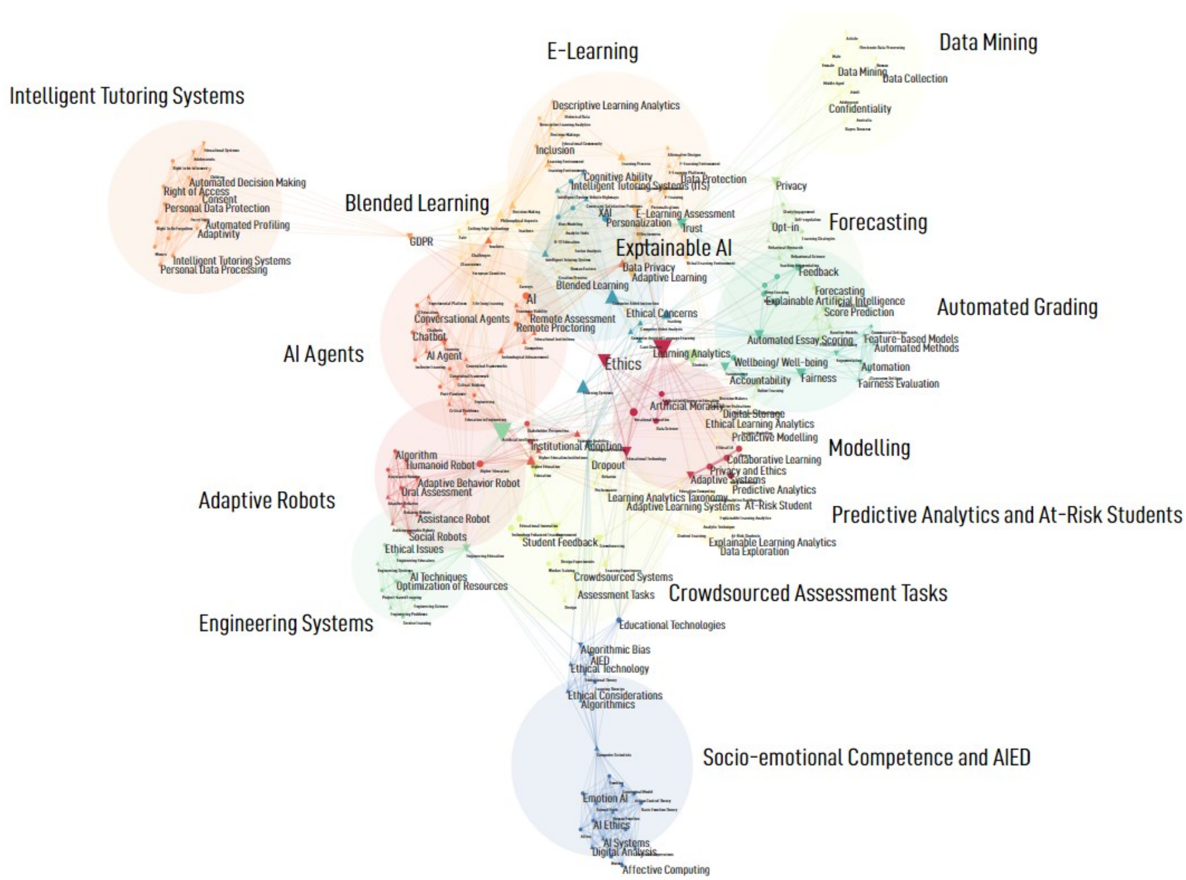


Figure 16: Network analysis of keyword corpuses

Through the review of the first pass of topic modelling, network analyses outputs, and full paper reviews, the study extensively identified fourteen sub-themes of AI application areas and ten sub-themes of ethical issues. We populate them in Table 9 and Table 10, respectively.

3.4.3 RQ3: What are the Key AI Use Cases Relating to Assessments?

As shown in Table 9, the fourteen sub-themes of AI application areas were, namely:

1. *AI-based assessment construction, curation, or delivery*

   AI can be a useful tool to: (i) construct assessments organically through, e.g., the use of generative AI to generate draft question samples for an assessment practitioner's review (Last and Danon, 2020); (ii) construct assessments collaboratively through e.g., crowdsourcing of assessment tasks (Ahn et al., 2021); and (iii) curate and deliver personalized assessment (Gupta and Chen, 2022; Heo & Lee, 2019) through e.g., curating multiple choice format formative assessments generated via AI conversational agents (Pereira, 2016).

   This form of AI utilization is cited in 45% of the primary studies. This is the third highest cited AI utilization form.

2. *AI-based socio-emotional assessment*

   AI can be used to: (i) assess non-cognitive psycho-emotional behavior qualities, such as persistence and grit, initiative and adaptability etc., through ambient intelligence (Stark and Hoey, 2021; Williamson, 2021); and (ii) discover socio-emotional patterns that may be predictive of assessment performance (Peña-Ayala, 2018).

   This form of AI utilization is cited in 6% of the primary studies.

3. *AI-based group assessment using collaborative analytics*

AI can be used to: (i) assess collaborative dynamics in group projects through the application of analytics on learners' rich multimodal interaction data; and (ii) model interdependent relationships to support students' collaboration dynamics through adaptive methods (Schneider, Dowell and Thompson, 2021).

This form of AI utilization is cited in 3% of the primary studies.

4. *AI-derived opportunities for learning intervention or assistance*

AI can be useful for: (i) identifying and supporting learning intervention opportunities in formative assessments (Shabaninejad et al., 2021; White et al., 2021); and (ii) providing scaffolding assistance in assessments, e.g., with the help of AI-driven hints through Intelligent Tutoring Systems (Conati et al., 2021; Latham and Goltz, 2019).

This form of AI utilization is cited in 52% of the primary studies. This is the second highest cited AI utilization form.

5. *AI-generated personalized feedback*

AI can be a useful tool to personalize feedback in terms of: (i) clarifying approaches to attempt a new assessment; (ii) reviewing assessment performance; and/or (iii) recommending takeaways from assessments to improve learners' demonstration of competence at their current or future workplace (Gupta and Chen, 2022; Merikko et al., 2022).

This form of AI utilization is cited in 55% of the primary studies. This is the highest cited AI utilization form.

6. *AI-based predictive analytics*

AI can be a useful tool to predict assessment outcomes, to support educators' focus on learning scaffolding (Chounta et al., 2022; Kim et al., 2021).

This form of AI utilization is cited in 15% of the primary studies. This is the fifth highest cited AI utilization form.

7. *AI-based teaching evaluation*

The term '*assessment*' in an educational context is extended to the assessing of educators as a formal evaluation process to review teaching effectiveness and perform in classrooms (Tlili et al., 2018).

This form of AI utilization is cited in 9% of the primary studies.

8. *AI-facilitated response and grading*

AI can be a useful tool for: (i) adaptive understanding and responding to learners in formative assessments (Khairy et al., 2022); and (ii) automated grading of structured (e.g., multiple choice questions) and non-structured (e.g., open-ended or essay questions) assessment responses (Litman et al., 2021; Kumar and Boulanger, 2020).

This form of AI utilization is cited in 30% of the primary studies. This is the fourth highest cited AI utilization form.

9. *AI-based proctoring*

AI can be a useful tool to proctor on-site (e.g., in classrooms or exam halls) or remotely (e.g., online assessment on Massive Open Online Courses, or MOOCs) to lower the risk of cheating in quizzes and exams (Elshafey et al., 2021).

This form of AI utilization is cited in 6% of the primary studies.

10. *AI-based authentication and security gateways for the conducting of assessments*

AI can be used to build authentication and security measures in assessment systems, so that the risk of personal and educational data leakage and exploitation can be lowered (Kiennert et al, 2019).

This form of AI utilization is cited in 3% of the primary studies.

11. *AI-backed plagiarism detection in assignment submissions*

AI can be a useful tool to perform plagiarism detections, so that assignment submissions that present someone else's works or ideas without acknowledgement or consent can be detected (Kiennert et al, 2019).

This form of AI utilization is cited in 3% of the primary studies.

12. AI-integrated communication dashboard for assessment outputs

An AI-integrated communication of assessment outputs can be designed in an inclusive format that promotes: (i) clear automated communication and feedback of assessments, accounting for the nuanced sensitivity of communications that may be absent in poorly designed AI-based text generators or audio-visual responses; and (ii) automation of inclusiveness actions. For instance, for learners with disabilities (such as motor disabilities), there may be an automation of increased time limit catered to complete an assessment (Costas-Jauregui et al., 2021).

This form of AI utilization is cited in 3% of the primary studies.

13. Autonomous intelligent agents for assessment purposes

This relates to the care given to the design of AI algorithms to create intelligent agents that make autonomous decisions for assessment purposes (Casas-Roma and Conesa, 2021; Hakami and Hernández-Leo, 2020).

This form of AI utilization is cited in 9% of the primary studies.

14. AI-based behavioral tracking for assessment purposes

AI can be used to track spatio-temporal data of assessment behavior: (i) to predict learning intervention opportunities; and/or (ii) for assessment evaluation purposes (White et al., 2021).

This form of AI utilization is cited in 3% of the primary studies.

Paper	Applications of AI Related to Assessments														Note
	Construction, curation, or	**Socio-emotional assessment**	**Group assessment**	**Learning intervention/**	**Behavioral tracking**	**Personalized feedback**	**Predictive analytics**	**Teaching evaluation**	**Grading and response**	**Proctoring**	**Authentication and security**	**Plagiarism detection**	**Communication dashboard**	**Autonomous intelligent agent**	
Gupta and Chen (2022)	•														- Telegram chatbot @dawebot, an AI-enabled conversational agent, provides formative assessments via multiple choice questions in any subject (Pereira, 2016)
Chounta et al. (2022)	•					•	•								-
Deho et al. (2022)				•			•								-

Shabaninejad et al. (2022)				•		•									-
Nazaretsky, Cukurova and Alexandron (2022)	•			•		•									-
Pontual Falcão et al. (2022)				•		•									-
Merikko et al. (2022)				•		•									-
Khairy et al. (2022)	•														-
Megahed, Abdel-Kader and Soliman (2022)									•						-
Conati et al. (2021)	•														- AI-driven hints on assessments delivered through Intelligent

															Tutoring System
González-Calatayud, Prendes-Espinosa and Roig-Vila (2021)	•			•		•			•						-
White et al. (2021)					•										- Tracking data of assessment behaviour applied for learning intervention
Ahn et al. (2021)									•						-
Stark and Hoey (2021)		•													-
Papa and Jackson (2021)														•	-
Kim et al. (2021)						•									-
Litman et al. (2021)									•						-

Casas-Roma and Conesa (2021)				•		•		•	•					•	-
Costas-Jauregui et al. (2021)				•		•							•		-
Elshafey et al. (2021)										•					-
Schneider, Dowell and Thompson (2021)			•												- Application of collaborative analytics in group assessments
Gedrimiene et al. (2020)	•			•		•									-
Kumar and Boulanger (2020)						•	•		•						-
Khosravi, Sadiq and Gasevic (2020)	•														-

Martín Núñez and Lantada (2020)	•			•		•			•						-
Hakami and Hernández-Leo (2020)				•		•		•	•					•	-
Mougiakou, Papadimitriou and Virvou (2019)	•			•		•			•						-
Mayfield et al. (2019)	•			•		•			•						-
Latham and Goltz (2019)	•			•		•									- AI conversational agent - AI-driven hints on assessments delivered through Intelligent

															Tutoring System
Tlili et al. (2019)	•			•		•									-
Kiennert et al. (2019)										•	•	•			-
Peña-Ayala (2018)	•	•		•		•	•								-
Tlili et al. (2018)	•			•		•		•							-
Total Count	**15**	**2**	**1**	**17**	**1**	**18**	**5**	**3**	**10**	**2**	**1**	**1**	**1**	**3**	
Percentage	*45%*	*6%*	*3%*	*52%*	*3%*	*55%*	*15%*	*9%*	*30%*	*6%*	*3%*	*3%*	*3%*	*9%*	

Table 9: Breakdown of sub-themes of AI application areas by paper

3.4.4 RQ4: What are the Key Ethical Principles Arising from the AI Implementations Relating to Assessments?

As shown in Table 10, the ten sub-themes of ethical principles were, namely:

1. *Inclusivity*

   This ethics principle relates to inclusive and accessibility considerations applied to AI systems to meet different student needs in a personalized environment at scale.

Inclusiveness is concerned about exhibiting empathy towards sensitive learner conditions, such as health, disabilities and learning disorders (e.g., pregnancy, visual handicap, or dyslexia), gender, race, prior education backgrounds (e.g., non-native speakers), and socio-economic backgrounds. In the design of the AI system, actions are taken to respect the diversity of learners and ensure that prejudices, stereotypes, discrimination and biasness do not creep into assessments (Gupta and Chen, 2022; Martín Núñez and Lantada, 2020; Tlili et al., 2019).

Inclusiveness is concerned about considering the sensitivity of communication and feedback generated by AI systems, so that learners are not negatively impacted by AI-generated textual comments or audio-visual responses when they clarify approaches to a new assessment or receive takeaways from completed assessments (Costas-Jauregui et al., 2021).

Inclusiveness is concerned about lowering the chances of conformity, peer pressure and segregation that may be reinforced because of AI generated decisions, which can negatively impact both educators and learners (Gedrimiene et al., 2022).

This ethics principle is cited in 18% of the primary studies.

2. *Fairness*

This ethics principle relates to fair, equitable and appropriate assessment practices that should be perpetuated by AI systems. It is noted that the definition of fairness is plagued with problems of subjectivity, contextualization and cultural-specificity (Hakami and Hernández-Leo, 2020).

Fairness is concerned about the treatment of data and algorithmic bias to ensure diversity, equity, non-prejudice and non-favoritism towards learners' sensitive attributes, so that needs of minority groups are not disadvantaged or underrepresented. This overlaps with the concept of inclusivity. (Gupta and Chen, 2022; Deho et al., 2022; Megahed, Abdel-Kader and Soliman, 2022; Casas-Roma and Conesa, 2021; Latham and Goltz, 2019; Tlili et al., 2019).

Tied to this discourse are the ethical concepts of (Mayfield et al., 2019): (i) *Allocation harm:* This relates to the equitable distribution of resources of learning, such that the possibility of differential outcome distributions generated by AI systems are minimized; (ii) *Representational harm:* This relates to the stereotyping bias perpetuated by data and/or algorithm, resulting in the marginalizing of groups of learners.

Fairness is concerned about the unintended labelling or profiling of learners, which can affect their learning journey and well-being (Peña-Ayala, 2018).

Fairness is concerned about, in the context of socio-emotional assessments, how applying universal assumptions on emotional states is harmful, due to different cultural context of emotional interactions and norms (Stark and Hoey, 2021).

Fairness is concerned about the ad hoc implementation of AI systems, in the absence of: (i) standard code of practices and ethics; and (ii) befitting monitoring and accounting mechanisms, which may impact the implementation of fair, equitable and appropriate assessment practices (Tlili et al., 2018).

This ethics principle is cited in 42% of the primary studies. This is the joint-second highest cited ethical issue, alongside explainability.

3. *Accountability*

This ethics principle relates to the responsible discharge of AI ethics when designing and delivering AI systems, depending on the roles and contexts, in a consistent manner. Stakeholders who construct, operate and use AI systems should be accountable for AI systems and decisions.

Accountability is concerned about the moral obligation for institutions to reflect and act, given that it has access to data that may know and understand how students learn (Costas-Jauregui et al., 2021). Students are data subjects who are not generally able to influence the handling of data in an ethical manner (Gedrimiene et al., 2020). Care should be applied when overseeing sensitive data. It was noted that between 2007 and 2011, there were 133 incidents linked to educational institutions unintendedly disclosing sensitive learner information (Stiles, 2012). Such incidents can lead to reputational, legal and/or financial liabilities (Tlili et al., 2018).

Accountability is concerned about the processes where relevant stakeholders provide reasons and take responsibilities for the actions of decisions influenced by AI algorithms (Hakami and Hernández-Leo, 2020). There should be proper consent, and non-maleficence academic interventions. For instance, in the development of AI systems for socio-emotional assessment, designers of AI systems should recognize and be accountable to the fact that there may exist diverse human attitudes to emotions to ensure fairness. Further, AI developers should also be cognizant that care should be applied to the underlying data, as data can be misused and emotions can be harnessed as a social phenomenon (Stark and Hoey, 2021).

Accountability is concerned about demonstrating compliance with relevant regulations and guidelines (Latham and Goltz, 2019). However, it is noted that there may exist challenges relating to such compliance. For instance, there may exist a lack of interoperability of regulatory guidelines on misuse of private information (e.g., European Union and Latin America), and a lack of clarity regarding whether if the institution or the students own the data that are shared by the students (Costas-Jauregui et al., 2021). Furthermore, for online courses offered worldwide, informed consent for use of data can be affected by the data protection regulations where the learner is domiciled. This greatly increases the difficulty of compliance efforts (Tlili et al., 2018).

Accountability is concerned about the availability of avenues for redress that are fair and unprejudiced due to the adverse use of AI systems to account for detrimental individual or societal effects.

This ethics principle is cited in 30% of the primary studies. This is the fourth highest cited AI utilization form.

4. *Accuracy*

This ethics principle relates to the reliability and validity of assessments when an AI system is applied. In the presence of possible biasness or errors introduced by data and AI algorithms, which may compromise the reliability and validity of assessments, there should exist measures to establish, log, communicate, diagnose and mitigate the biasness or errors.

Accuracy is concerned about poor data quality. For instance, due to the distributed nature of online learning, data collected may be incomplete or erroneous. Linked to this latter notion is the right for data subjects to check and rectify data collected, so that data inputs are accurate. Poor data quality can negatively impact AI-driven decisions (Tlili et al., 2018).

Accuracy is concerned about inappropriate data inputs. For instance, predictive models with imbalanced dataset (e.g., gender) may generate less effective predictions (e.g., for minority gender). The possibility of discriminatory and unfair practice extends to other socio-economic demographic information, such as ethnicity, underrepresented groups etc. (Chounta et al., 2022). In another example, in the implementation of collaborative analytics, it is vital to capture multimodal data that best represents and assesses the collaborative interactions of students, otherwise the AI-generated outputs cannot be relied upon (Schneider, Dowell and Thompson, 2021).

Accuracy is concerned about inaccurate understanding and interpretation of learner responses in assessments. For instance, adaptable humanoid robots may not understand or interpret the responses of learners in an oral assessment well. Correct answer rate, if affected by learner's pronunciation or robot's lack of contextual understanding, can affect confidence in the AI system (Khairy et al., 2022). Research by Ahn et al. (2021) also showed that automated grading of learners' work, which contains complex data, rich semantic meaning and idiosyncratic and local nuances, may not be well graded by present computational approaches that utilize metrics such as counts of parts of speech and essay length as proxies for writing complexity and quality. Further, present systems are generally rigid in formulation of tasks and grading. For instance, the rejection of lexicon and grammar of minority dialects. This limits the choice of tasks, types of acceptable answers and styles of writing (Mayfield et al., 2019).

Accuracy is concerned about the validity and reliability of an assessment instrument. Kim et al. (2021) shared that, to allow for shorter assessments in online learning, it is imperative that the question set reduction guided by AI algorithms is done in such a way where the reduced set is able to approximate the original assessment's evaluation of learning. This can help ensure reliability and trust in assessment instruments.

Accuracy is concerned with the treatment of prediction errors and biases. For instance, in the case of cheating detection, the treatment of non-cheating cases that were falsely detected as cheating (or false positives), and cheating cases that were not detected (or false negatives). The false positives and negatives will have to be reduced to improve accuracy rates (Kiennert et al., 2019) .

Accuracy is concerned about the possibility of inaccurate predictions, due to the "gaming" of AI systems. For instance, there exist a possibility for students to modify their behavior and "game" the AI system, when they have the knowledge that they are assessed by the AI system, and the knowledge of the parameters of the AI model. This can create inaccurate AI decisions when assessing students (Tlili et al., 2018).

This ethics principle is cited in 27% of the primary studies. This is the fifth highest cited AI utilization form.

5. *Auditability*

This ethics principle relates to permitting independent third-party reviewers to audit, analyze and report findings relating to the usage and design of data and AI algorithms in assessments.

Auditability is concerned about the understanding, validating, reviewing and improving of the AI system applied, so that there are appropriate transparency, traceability and utilization of data and AI algorithms, and appropriate validity and reliability of assessment instruments. However, it is noted that challenges may arise if algorithms are proprietary (Tlili et al., 2019; Casas-Roma and Conesa, 2021).

This ethics principle is cited in 9% of the primary studies.

6. *Explainability*

This ethics principle relates to the lowering of opacity relating to data, AI algorithms and AI-driven decisions, the justification of its use, and the communication of details in a non-technical easy-to-understand manner to relevant stakeholders (Kumar and Boulanger, 2020; Casas-Roma and Conesa, 2021).

Explainability is concerned with transparency of the design of AI systems. Transparency is tied to information availability, accessibility conditions, possibility of pragmatic decision-making assistance, and user knowledge (Nazaretsky, Cukurova and Alexandron, 2022). It is important for AI-based assessments to be developed in an explainable and transparent manner to safeguard trust and fairness with human stakeholders. For instance, for AI recommender systems, why are some assessment questions recommended over others (Chounta et al., 2022). "Black-box" AI recommendations, which provide low or no insights into recommendation rationales (Abdi, 2020), may be plagued with biases and confounding problems (Bastani, Bastani and Kim, 2018; Khosravi et al., 2021), resulting in unjustified actions and discrimination (Papa and Jackson, 2021). However, it is noted that full transparency may be harmful, as users can "game" the system to their benefit and the detriment of others (Hakami and Hernández-Leo, 2020). Furthermore, challenges may arise from disclosure of proprietary algorithms or trade secrets (Latham and Goltz, 2019).

Explainability is concerned about the use of explanations to gain insights into the behavior of AI systems. Present explainability approaches include global and local approaches. The former synthesizes and uncovers qualities of inputs that affect model behaviors on a global basis, whereas the latter looks to explain the model's behavior to a specific input. Another more recent approach is to leverage on generative capabilities of models to self-explain a human-understandable explanation for input-output responses (Bommasani et al., 2021).

Explainability is concerned about tradeoffs relating interpretability and complexity of AI systems. The design of the AI algorithms may be potentially complex. For instance, for collaborative analytics, the use of interdependent modelling to assess group dynamics and outcomes may raise challenges on explainability of AI systems. Interdependent models (which look at students' influences on one another over time) may be significantly more complex than independent models (which look at students as isolated events). There is a need to assess the trade-off between model complexity and explainability, and ascertain if simpler models are sufficient to model dynamic interdependence (Schneider, Dowell and Thompson, 2021). Deho et al. (2022) suggests that interpretable models (e.g., logistic regression) may provide less unfairness as compared to complex fairness-aware models, with robust accuracy results (Kung and Yu, 2020). This supports the notion of using interpretable and explainable models for AI in assessments.

Explainability is concerned about the absence of theoretical basis to justify the development and use of AI system for assessment purposes. González-Calatayud, Prendes-Espinosa and Roig-Vila (2021) cites a lack of pedagogical underpinning and AI training, which affects the meaningful development of assessments with pedagogical reference models when AI is applied.

This ethics principle is cited in 42% of the primary studies. This is the joint-second highest cited ethical issue, alongside fairness.

7. *Privacy*

This ethics principle relates to the protection of data subjects against injurious effects from the use of personal information applied in AI systems, without unduly affecting regulatory compliance tied to privacy and restricting AI development.

Privacy is concerned about the governance of end-to-end data stewardship, including data collection, storage, disclosure, sharing, security and disposal when applied to AI assessments (Chounta et al., 2022). Kiennert et al. (2019) highlights the importance on the management of sensitive data, such as authentication and biometric samples (e.g., data collected for password, voice recognition, facial recognition and/or keystroke detection). Leakage of these data can cause risk of harm. AI systems should be secure and not vulnerable to tampering. Stark and Hoey (2021) highlight how individuals are sensitive about data sharing and utilization pertaining to their emotions and emotional expressions.

For example, the Facebook emotional contagion study (Kramer, Guillory and Hancock, 2014) was criticized for manipulating emotive content of users. Safeguarding of trust and confidence in the governance of data stewardship are important for stakeholder security, privacy and risk of harm (White et al., 2021).

Privacy is concerned about the implementation of fair data stewardship practices such as notice, access, and choice (Mougiakou, Papadimitriou and Virvou, 2019). Explicit consent should be obtained from data subjects, such that users should be given the right to maintain control over data usage, control the purpose and extent of usage, be granted the option to modify the usage and context, and be given the right to opt in and out of participation. It is noted that consent involving minors can be challenging, as this may require both the students' and their parents' consents (Latham and Goltz, 2019). Students who wish to join or withdraw from certain AI-influenced activities may be allowed to do so, especially vulnerable groups such as students with learning disorders, language barriers, or students who come from lower socio-economic backgrounds (Gedrimiene et al., 2020). Educator may also exercise the right to opt in or out of participation in AI-driven teaching evaluation (Tlili et al., 2018). The possibility to opt in or out may result in data gaps that can affect accuracy of results and research outcomes, indirectly isolate and reveal outcomes of those who opt in or out, and affect discharge of institutional duty to enhance learning experience for students (Tlili et al., 2019). Merikko et al. (2022) finds that learners are open to sharing data related to demographics and learning performance, but are apprehensive about sharing when it comes to their online behavior, sensitive or process data. Further, the more personal and granular the data are, the less likely the learners will share them. Furthermore, learners who were not performing well are less likely to share their performance data. This may be tied to help-seeking avoidance, as seeking help may be a sign of weakness and a threat to self-esteem (White and Bembenutty, 2013).

Privacy is concerned about constant surveillance arising from AI use (Megahed, Abdel-Kader and Soliman, 2022). There may exist a possibility of violation to individuals' rights to privacy when too much data surveillance exists, especially when data is used beyond academic purposes, for control and surveillance to modify human behavior (Pontual Falcão et al., 2022). Mayfield et al. (2019) discusses the undesirable anxieties and behavioral change related to constant surveillance.

This ethics principle is cited in 55% of the primary studies. This is the highest cited ethical issue.

8. *Trust*

This ethics principle relates to the placing of confidence on (i) AI systems and the (ii) provision of data to achieve assessment objectives. The former is a characteristic of the human-machine relationship formed with an AI system. Low trust is largely linked to the lack of human properties (e.g., lack of affect, emotions, pedagogical intuition) in AI systems (Nazaretsky, Cukurova and Alexandron, 2022). The latter is related to the preservation of privacy.

Trust is concerned about the ability to rely on AI systems to make decisions and provide feedback. Pontual Falcão et al. (2022) suggests discomfort among students and educators of AI-driven decision making that involves ranking, sorting and classifying individuals, that may reflect political interests, social values, and risks of omissions or biases. In addition, the same paper shares that learners and educators were not confident that incorporating AI in assessments can result in improvements of feedback quality.

Trust is concerned about the presence of a clear and global consensus of purposes and specifications of AI systems. For instance, with regards to socio-emotional assessments, Stark and Hoey (2021) highlights a lack of consensus objective agreement on emotion at a global level as an issue. This is because, in the absence of a consensus, the large variation in the implied social and ethical responsibilities have normative implications for AI systems e.g., while considering ethical values such as accountability and fairness when assessing socio-emotional qualities in learners. This affects trust on AI systems.

Trust is concerned about the autonomy and control that stakeholders have on AI systems. For instance, Pontual Falcão et al. (2022) cites educators' discomfort at lack of autonomy and control due to its use in the appraisal of teaching performance and excessive intrusion in learners' learning routine. Learners may also be worried that their autonomy and independent decision making may be deprived.

This ethics principle is cited in 12% of the primary studies.

9. *Human Centricity*

This ethics principle relates to the aim towards upholding human agency, dignity and autonomy, minimization of harm (and when necessary, weighed against a greater good), and equitable distribution of benefits.

Human centricity is concerned about agency and autonomy of users. This overlaps with the concept of trust. Users should not be impacted by profiling, ranking and personalizing derived from AI algorithms. Learning should not be viewed as "*product-oriented learning experiences*" (Duignan, 2020). There should be care applied, when it comes to AI algorithm manipulating learner behaviors and emotions (Papa and Jackson, 2021). AI systems should not negatively impact a learner's capacity to learn and his/ her level of autonomy to make learning decisions. There should exist a presence of reversible and clear processes, and the possibility to intervene for blocking, termination, correction and erasure (Mougiakou, Papadimitriou and Virvou, 2019).

Human centricity is concerned about the states of human wellbeing (e.g., psychological wellbeing and satisfaction). A well-designed AI system should seek to achieve positive states of human wellbeing.

This ethics principle is cited in 12% of the primary studies.

10. Academic Integrity

This ethics principle relates to dishonest and deceptive learner behavior to violate assessment rules and regulations.

Academic integrity is concerned about the identification of dishonest and deceptive assessment behaviors through the use of AI-based proctoring and plagiarism detection, both in physical venues and remote assessment platforms (Elshafey et al., 2021; Kiennert et al., 2019).

This ethics principle is cited in 6% of the primary studies.

Paper	AI Ethical Issue(s) as Cited in Paper									
	Inclusivity	**Fairness**	**Accountability**	**Accuracy**	**Auditability**	**Explainability**	**Privacy**	**Trust**	**Human Centricity**	**Academic Integrity**
Gupta and Chen (2022)	•	•								
Chounta et al. (2022)		•	•	•	•	•	•			
Deho et al. (2022)		•				•				
Shabaninejad et al. (2022)						•				
Nazaretsky, Cukurova and Alexandron (2022)			•			•		•		
Pontual Falcão et al. (2022)							•	•		
Merikko et al. (2022)							•			
Khairy et al. (2022)				•						
Megahed, Abdel-Kader and Soliman (2022)		•					•			
Conati et al. (2021)						•				
González-Calatayud, Prendes-Espinosa and Roig-Vila (2021)						•				
White et al. (2021)							•			
Ahn et al. (2021)				•						
Stark and Hoey (2021)		•	•				•	•		
Papa and Jackson (2021)		•				•			•	

Kim et al. (2021)				•						
Litman et al. (2021)		•		•		•				
Casas-Roma and Conesa (2021)		•			•	•	•		•	
Costas-Jauregui et al. (2021)	•		•				•	•		
Elshafey et al. (2021)										•
Schneider, Dowell and Thompson (2021)	•			•		•				
Gedrimiene et al. (2020)	•		•				•			
Kumar and Boulanger (2020)						•				
Khosravi, Sadiq and Gasevic (2020)			•				•			
Martín Núñez and Lantada (2020)	•						•			
Hakami and Hernández-Leo (2020)		•	•			•			•	
Mougiakou, Papadimitriou and Virvou (2019)							•		•	
Mayfield et al. (2019)		•		•			•			
Latham and Goltz (2019)		•	•			•	•			
Tlili et al. (2019)	•	•	•		•	•	•			
Kiennert et al. (2019)				•			•			•
Peña-Ayala (2018)		•					•			
Tlili et al. (2018)		•	•	•			•			
Total Count	**6**	**14**	**10**	**9**	**3**	**14**	**18**	**4**	**4**	**2**
Percentage	*18%*	*42%*	*30%*	*27%*	*9%*	*42%*	*55%*	*12%*	*12%*	*6%*

Table 10: Breakdown of sub-themes of ethical issues by paper

3.4.5 Second Pass of Topic Modelling and Network Analyses

Next, we utilize the keyword corpuses of fourteen sub-themes of AI application areas and ten sub-themes of ethical issues as an input, to perform the second pass of topic modelling, and network analyses.

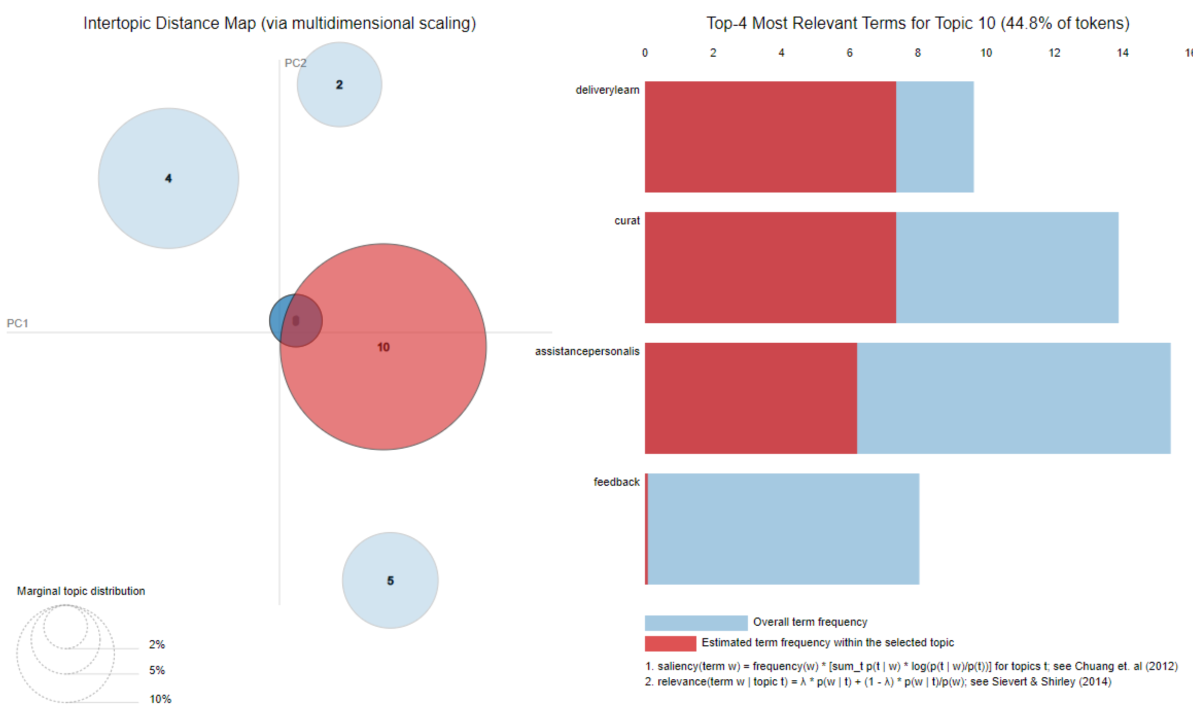


Figure 17: Topic modelling of corpuses involving AI application areas and related ethical principles

Latent Topic	**Percentage of Tokens**	**Top Keywords**
System design and check	*7.5%*	System; Design; Review
Data stewardship and surveillance	*2.9%*	Privacy; Sensitive; Data

Assessment construction and rollout	*44.8%*	Deliver; Curate; Personalize
Assessment administration	*9.6%*	Proctor; Plagiarism; Cheat
Grading and evaluation	*20.7%*	Evaluation; Feedback; Response

Table 11: Second pass of topic modelling – Latent topic and top keywords

Figure 18: Network analyses of corpuses involving AI application areas and related ethical principles

Research identified five topical archetypes via topic modelling. For instance, in Figure 17, we observed the dominant latent topic linked to AI-based assessment construction and rollout aspects. This mirrored well with the network analyses visualization in Figure 18. In the network analysis diagram, we observed a clear clustering of five topics, with *Assessment Construction and Rollout* similarly dominant in the cluster diagram. The top keywords and latent topics of topic modeling are shown in Table 11.

Understanding of these five key archetypical themes allows researchers and practitioners to breakdown the landscape into clear segments, to decide which area they would like to develop further insights and applications.

3.5 Discussion

Morley et al. (2020) emphasizes the importance of the translation of AI principles into the 'what' and 'how' of implementation. Building on the work of preceding sections, this section discusses actionable insights to make the addressing of AI ethics operable in the real world.

3.5.1 RQ5: What are the Key Themes Inherent in the Consideration of Ethical Imperatives in Educational Assessments?

Ontology can be defined as "*an explicit specification of a conceptualization*" (Gruber, 1993), geared towards a "*shared taxonomy of entities*" (Smith and Welty, 2001), as applied in information systems. This is opposed to the philosophical use of the concept of ontology as a nature of reality (Ashok et al., 2022). The investigation of AI systems can be considered a sub-field of information systems.

Ashok et al., (2022) describes three fundamental domains to conceptually represent the interweaving ethical elements and interrelationships inherent in the design and application of AI in digital technologies. This is theoretically underpinned by ontological frameworks of Ogden and Richards (1923), Popper (1979) and Project and Peirce (1998).

This triadic framework is a modular architecture of an assemblage of technological components that consist of the:

1. *Physical domain* (or the referent or object in semiotics): This includes the: (i) *device layer* which comprises a logical capability operating system layer, and the physical machinery

hardware layer; and the (ii) *network layer* which comprises the logical transmission network protocol layer and physical network transport layer.

Some relevant applications are author systems, intelligent tutoring shells, AI-integrated learning environments, educational robotics, and AI collaborative tools.

2. *Cognitive domain* (or the symbol or science in semiotics): This comprises the *content layer* where data is stored, created, mapped, manipulated, utilized, and shared.

   Some relevant examples are multimodal structured contents of text, and unstructured contents of images, sounds and videos of assessment submissions. This layer also provides the metadata and directory information of users, content tags, location stamps, time stamps, encoding and copyright etc.

3. *Information domain* (or the reference or interpretant in semiotics): This comprises the *service layer* which encompasses the functionality of the application and its interaction with users, underpinned by AI algorithms.

   Some relevant examples are use of knowledge representation for instructions, human factor and interface design, and AI-integrated visualization and graphics for feedback.

We extend the triadic ontological framework as described by Ashok et al., (2022) to model and visualize the systematic literature map of this study (Figure 19). We note that, over and above the triadic domains, Ashok et al., (2022) further describes a governance domain, defined by Floridi,

(2018) as *"the practice of establishing and implementing policies, procedures, and standards for the proper development, use, and management of the infosphere*." In our opinion, governance is a key consideration across all triadic domains, and hence, the governance domain is not explicitly illustrated in our framework.

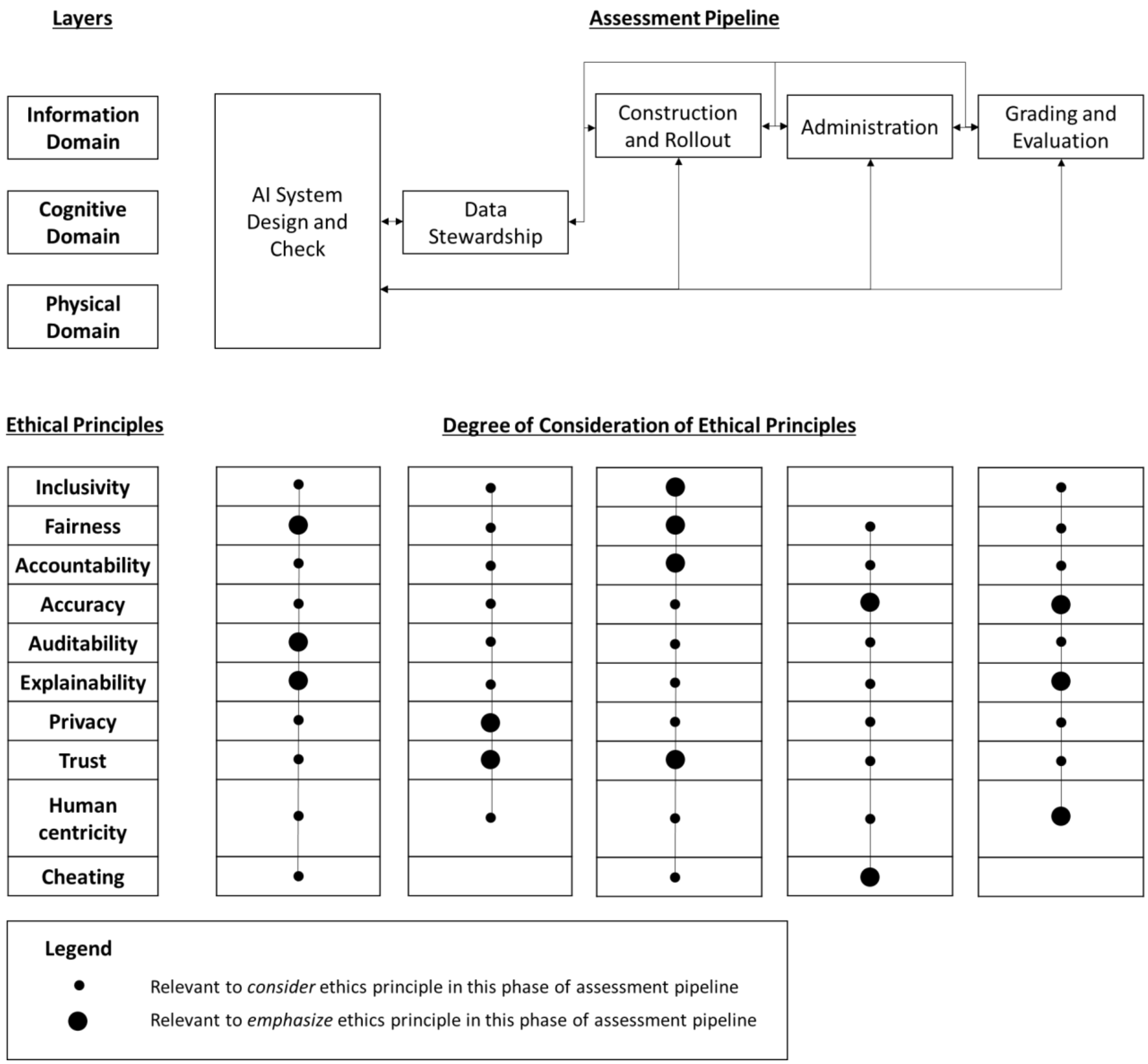


Figure 19: Visualization of the systematic literature map of key research themes

The five distinct clusters identified by topic modelling and network analyses in Figure 17 and Figure 18 are mapped to the triadic ontological framework in Figure 19, as follows:

1. *AI system design and check for assessment purposes*

   This phase extends across the physical, cognitive and information domains, and is involved with the design, implementation and maintenance of the AI system for system interactivity, robustness and security. From a predictive analytics point of view, the model constructed should be appropriate – upholding accuracy, inclusivity, accountability, privacy, trust and human centricity.

   At this phase, the overriding ethics imperatives are explainability and auditability. The AI system should be created with clear, easy-to-understand and transparent protocols, so that relevant stakeholders and independent third-party auditors can review the processes, perform interventions, mitigate issues, and enable redress in an event of negative outcomes that may arise. In addition, fairness is concerned about the treatment of algorithmic bias to ensure diversity, equity, non-prejudice and non-favoritism towards learners' sensitive attributes, so that needs of minority groups are not disadvantaged or underrepresented.

2. *Data stewardship and surveillance*

   This phase extends across the cognitive and information domains, and is involved with the governance and implementation of good data stewardship, and appropriate surveillance practices (if any).

At this phase, the overriding ethics imperative is privacy. One instance is behavioral surveillance, which may be a violation to human rights to privacy especially when data is used beyond academic purposes, for control and surveillance to modify human behavior. In addition, trust is also an important facet concerned about the preservation of privacy when sensitive data are disclosed.

3. *AI-based assessment construction and rollout*

This phase is predominantly situated in the information domain, and is involved with the construction, curation or delivery of assessment, the communication of evaluation and feedback with stakeholders via AI-integrated communication dashboards, and the carrying out of interventions and assistances to improve assessment and evaluation performance. Assessment and evaluation can be in the form of formative (or summative) individual (or group) cognitive (or socio-emotional) assessment. It can also be a form of teaching evaluation.

At this phase, the overriding ethics imperatives are inclusivity and fairness, so that appropriate and equitable assessments and evaluations are rolled out, embracing diversity, empathy and sensitivity towards the evaluated stakeholders. Furthermore, accountability is an important ethics consideration, as there should exist a responsible discharge of AI ethical principles and compliance with relevant rules and guidelines, when designing and delivering AI-driven assessments. In addition, there should exist trust and confidence on AI systems to achieve assessment and evaluation objectives.

4. *Administration of assessments using AI systems*

This phase is predominantly situated in the information domain, and is involved with the administration of assessment and evaluation, which may comprise authentication and security measures, proctoring and/or plagiarism detection.

At this phase, the overriding ethics imperatives are the overcoming of cheating violations, and the application of accuracy to correctly identify assessment candidates and cheating cases.

5. *AI-facilitated assessment grading and evaluation*

This phase is predominantly in the information domain, and is involved primarily with the interpretation of textual and/or audio-visual responses collected by AI systems, the evaluation of performance, and the provision of feedback. These may be performed by autonomous intelligent agents. From an educator's point of view, this phase may involve the evaluation of teaching effectiveness.

At this phase, the overriding ethics imperative is explainability, so evaluators can understand and adjudge if the grading and/or ranking is accurate and reliable. In addition, there is an element of human centricity. This largely relates to the agency and autonomy of human users, in the presence of AI-generated decisions, and the capacity to intervene for correction and redress.

There is an emphasis that the framework does not draw clear delineations when categorizing AI assessment use cases across triadic domains. For instance, the *Grading and Evaluation* research theme is predominantly arising from the cognitive domain. However, coding and rolling out a moral reasoning AI system for AI-generated decisions, evaluations, responses and feedback, a sub-item of this research theme, may straddle across all cognitive, information and physical domains. This said, the framework provides a guide to generalize observed phenomena.

3.5.2 RQ6: What are Solutions and Interventions that were Proposed to Address Key Ethical Imperatives, and their Associated Underpinning Theories?

This study provides a breakdown of mitigation and intervention programs and activities for ethical imperatives, and a non-exhaustive list of key theoretical concepts that can help underpin research in the selected area (Table 12 to Table 16). This endeavor can help provide actionable insights to address ethical issues impacting assessments in AIED, and the takeaways can be purposed as a thematic guide to future applied research.

1. AI system design and check for assessment purposes

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Physical, Information, Cognitive	**Examples of Tools and Techniques (Inclusivity, Fairness)** - Chatbot *Sammy* helps promote judgment-free inclusiveness to learners. Its inclusivity also extends to ubiquitous access, and is helpful to learners with disabilities (e.g., visual or hearing impaired) or learning disorders (Gupta and Chen, 2022). - Chatbot *CiSA* is designed to promote equity and social inclusivity for international students (Heo & Lee, 2019).	Deontological (Normative ethics)	- Compliance to behavioral rules
Physical, Information, Cognitive	**Governance (Fairness, Accountability, Trust, Explainability, Human centricity, Privacy, Inclusivity, Accuracy, Academic integrity, Auditability)** - Build an AI ethics framework, that will make explicit ethic issues (such as privacy, explainability, fairness etc.), so that users can anticipate and avoid ethical issues on AI systems (Shapiro and Blackman, 2020; AI HLEG, 2019). - Ethical AI practices should be defined and regulated at institutional and national levels, so that educators and students are protected by policies, as AI ethics are complex	Deontological (Normative ethics) ***Other relevant non-philosophical theories:*** - Psychology theories, e.g.: Cognitive Dissonance Theory	- Compliance to rules and standards

Domain	**Ethics Mitigation/ Intervention Programs and Activities**	**Theoretical Underpinning**	**Outcome**
	issues that may not be entirely understood or may be understood differently by stakeholders (Gedrimiene et al., 2020). Establish a detailed and clear constitution for AI ethics, highlighting ethical guidelines in the AI process, and the duties and rights of all stakeholders (Tlili et al., 2018). - Consensus around clear AI ethical principles are mixed and varied, as they stem from regulations (e.g., GDPR), laws (e.g., FERPA), standards (e.g., IEEE) and codes (e.g., Asilomar AI principles). It will be useful to establish a clear consensus set of principles (Latham and Goltz, 2019). - Trade-offs may exist in the application of principles when considered from the point of views of individuals, stakeholders or the society. As such, tensions may arise and the trade-offs will need to be managed (Latham and Goltz, 2019). For instance, definitions of fairness may be clarified, to reduce subjectivity and improve contextual awareness (Hakami and Hernández-Leo, 2020). - Centre AI ethics around underpinning theories, such as learning theories, to provide humanistic and social dimensions to the AI-mediated process (Papa and Jackson, 2021).	- Techno-Sociology theories, e.g.: Theory of Network Society	

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	- Useful to consider appointing Chief (Digital) Ethics Officer, and an institutional team dedicating time and effort to ethics governance, stewardship, and education (Borenstein and Howard, 2021). Andrews et al. (2022) details the profile and requirements of this office bearer. - Prospective AI technology may suffer from interpretive flexibility. From a technological standpoint, identify and chart evolution, interrelationship and non-linearity of technological development. Understand technical capacities, socio-interactional mechanisms and contextualized cultural relevance to institution (Schiff, 2021). For instance, when applying AIED in a metaverse setting, it is necessary to consider the possibilities of cyber-bullying? (Hwang and Chien, 2022) - Consider different cultural context of emotional interactions and norms, individual and collective subjective assessments, and changing global paradigms of emotional contexts and subjectivity. Due to ethical valences and social effects of the diversity of ethical opinions, it is useful to establish a global objective agreement and consider how e.g., cultural context differences can impact the design and deployment of AI systems (Stark and Hoey, 2021).		

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Physical, Information, Cognitive	**System Supervision, Audit and Intervention (Accountability, Explainability, Auditability, Fairness)** - Accountability begins from designers and developers of the AI system. Guiding questions to consider for good accountability practices (Hakami and Hernández-Leo, 2020): (i) Consequences of algorithmic decisions on societies and individuals. (ii) Influence of consequences and the number of people affected by the consequences. (iii) Degree of awareness on how AI algorithms drive decisions. (iv) Possibilities of occurrence of discrimination and bias, and how this can impact public perception. (v) Preventive strategies and techniques that can be put in place at the onset of system design. (vi) Maintenance strategies and techniques that can intervene AI system during deployment. (vii) Optimization strategies and techniques that can improve AI system post-deployment. - AI system to allow the possibility of intervention and corrective actions to enhance the automated process (Tlili et al., 2019). One way to improve "intervenability" is the clear	Deontological (Normative ethics) ***Other relevant non-philosophical theories:*** - Information theories, e.g.: Signaling Theory, Theory of Information Asymmetry, Theory of Voluntary Disclosure - Techno-Sociology theories, e.g.: Theory of Network Society	- Compliance to rules and standards

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	segmentation of AI system into key components. This includes, e.g., separation of registration, consent, interaction, assessment, grading, profiling and automated decision-making (Mougiakou, Papadimitriou and Virvou, 2019). - Ensure AI process is logged, tracked, interpreted, and checked by independent auditors (Casas-Roma and Conesa, 2021). AI system transparency is paramount. - Enable whistleblowing, and early warning systems, with systematic investigations, and measurement of ethics violations and organizational climate. Availability of avenues for redress should be fair and unprejudiced to account for detrimental individual or societal effects (Hoekstra and Kaptein, 2021). - No clear guidelines exist on who should be the stakeholder responsible for understanding, validating, reviewing, and improving the AI system. It will be useful to clarify if this stakeholder should be the educators, system designers, or administrators, or if learners should also be involved in this process (Tlili et al., 2019).		

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Physical, Information, Cognitive	**System Design and Construction (Fairness, Privacy, Trust, Explainability, Human centricity, Inclusivity, Accuracy)** - Integrate AI ethics in entire AI development pipeline (Megahed, Abdel-Kader and Soliman, 2022). - Clear taxonomies of ethics will be useful to guide AI system design choices and related research, so that attempts to address ethical issues can be holistic, rather than ad hoc (Mayfield et al., 2019). - Participatory or co-design process in the creation of AI systems, to reduce barriers of trust (Nazaretsky, Cukurova and Alexandron, 2022; Costas-Jauregui et al., 2021; Schneider, Dowell and Thompson, 2021). - Development of AI system would be best served as an inter-disciplinary approach, integrating disciplines such as Anthropology and Sociology (Costas-Jauregui et al., 2021). - Create transparent AI systems and avoid "black-box" AI solutions to improve trust (Peña-Ayala, 2018). However, too much transparency may also lead to information overload - a paradox known as "transparency paradox".	Deontological (Normative ethics) ***Other relevant non-philosophical theories:*** - Management Information Systems theories, e.g.: Socio-technical Theory, Task-technology Fit Theory, Cognitive Fit Theory - Techno-Sociology theories, e.g.: Technological Determinism and Social Constructivism Theories	- Compliance to behavioral rules

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	Nazaretsky, Cukurova and Alexandron (2022) suggests that it may be useful to study the extent to which transparency can benefit use of AI in assessments. - Ensure equitable access to fair opportunities and quality level in AI assessment environment (e.g., requirement of certain internet connection speed or hardware) (Casas-Roma and Conesa, 2021). - Beyond assessment performance metrics, it is useful to consider relevant AI system design factors such as socio-emotional aspects, self-regulation, cognitive load and inclusivity considerations (Hakami and Hernández-Leo, 2020).	- Technology theories, e.g.: Instrumental-ization Theory, Theory of Reasoned Action, Theory of Planned Behavior	
Physical, Information	**Development Team (Inclusivity, Accountability)** - Subject matter is a complex multi-faceted issue that spans, among others, pedagogy, technology, and psychology domains. It will be useful to include broad disciplinary expertise (e.g., computer science, education, social science) in development work, to achieve sounder theoretically underpinned methodology approaches and more stakeholder inclusivity (Raji, Scheuerman and Amironesei, 2021).	Deontological (Normative ethics)	- Compliance to behavioral rules

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	- Diverse AI development teams to identify and lower biases resulting from use of biased dataset or wrongly trained models (Martín Núñez and Lantada, 2020). - Promote proactive (not reactive) behavior when designing and managing AI systems, so as to ensure the safety of all stakeholders are accounted for (Tlili et al., 2018).		
Physical, Information	**Algorithm and Coding (Explainability, Fairness, Human centricity)** - Improve resilience to adversaries and black swan events. Monitor prediction, detect unexpected model functionality and malicious use. Ensure systemic safety from e.g., cyberattacks (Hendrycks et al., 2021). - Formalize existing human norms and values into expressive and flexible responsible coding, by using decision-theoretic logic programming to achieve value alignment (van Otterlo, 2017). - Use of ethics mitigation techniques and algorithms in the entire AI pipeline, including pre-processing for data, in-processing within the model, and post-processing of model results. For instance, for unfairness mitigation, Deho et al. (2022) suggests that goal of fairness may not be equal	Deontological (Normative ethics)	- Compliance to behavioral rules

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	treatment, but the "needed" treatment to booster learning success. Useful to study the factors and the extent to which the factors apply in terms of what is "needed".		
Physical, Information	**AI Modelling and Parameter Tuning (Explainability, Accuracy, Fairness)** - Use of ethics mitigation techniques and algorithms in modelling. For instance, Litman et al. (2021) shares unfairness mitigation techniques using fairer feature selection strategies, which may work, over and above mitigating imbalanced dataset and/or pre-training bias-free models. Further, the authors cite hybrid feature-based and neural network models, that combined accuracy with explainability. - Assess the trade-off between model complexity and explainability, and ascertain if simpler models are sufficient. For example, for collaborative analytics, it may be useful to apply simpler models to understand dynamical interdependence, rather than more complex models (Schneider, Dowell and Thompson, 2021). - Consider the use of local, global, and generative approaches to provide a human-understandable explanation	Deontological (Normative ethics)	- Compliance to behavioral rules

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	for a model's input-output responses (Bommasani et al., 2021). - Explainable AI, e.g., use of metrics such as **<u>SH</u>**apley **<u>A</u>**dditive ex**<u>P</u>**lanations (SHAP), can help assess trustworthiness of complex algorithms (including ensembles and deep learning models), such as feature-based multi-layer perceptron deep neural network. Explainable AI can play roles in parameter tuning to improve interpretability and generalizability (e.g., tuning of hidden layer depth), discover decision making process (e.g., simpler or more complex feature selection to make up the explanation), and provide granular personalized feedback to learners (customizable and trustworthy explanations to learners) (Kumar and Boulanger, 2020). - Assessment of affective states using affect-aware systems and multimodal data can also be tainted by biasness. Care should be placed on affect-detection systems (Mayfield et al., 2019).		
Physical, Information, Cognitive	**Human Oversight (Fairness, Trust, Explainability, Human centricity, Auditability, Privacy)** - Human-in-the-loop (i.e., active human oversight) (Deho et	Virtue theory (Normative ethics)	- Human's ethics consciousness, sensibility and analytical skills

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	al., 2022; Casas-Roma and Conesa, 2021; Kiennert et al., 2019). - Nazaretsky, Cukurova and Alexandron (2022) endorses active human oversight of AI decisions, as greater autonomy and control can (i) improve trust between educators and AI systems, (ii) reduce anxiety of educators' replacement by AI systems, and (iii) reduce errors where the educators are accountable.		applied to error and biasness detection, monitoring, evaluation, remedy, and prevention
Physical, Information, Cognitive	**Human Reliance on AI System (Trust)** - Experience relating to the use of AI may influence educators' and learners' opinions. To bridge expectations of AI use for educators and learners, it is important to improve data literacy levels and knowledge of tools available (Pontual Falcão et al., 2022). - Trust can be shaped by institutional commitment and context. Institutional expectations and support may be more explicitly shared to improve trust (Pontual Falcão et al., 2022).	Virtue theory and consequential-ism theory (Normative ethics) ***Other relevant non-philosophical theories:*** - Information theories, e.g.: Signaling Theory, Theory of Information Asymmetry, Theory of	- Character and moral uprightness - Maximize ethics consciousness and analytical skills

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
		Voluntary Disclosure	

Table 12: AI system design and check for assessment purposes: Breakdown of ethics mitigation and intervention programs and activities, and their key theoretical underpinning

2. Data stewardship and surveillance

Domain	**Ethics Mitigation/ Intervention Programs and Activities**	**Theoretical Underpinning**	**Outcome**
Information	**Examples of Tools and Techniques (Accuracy)** - Schneider, Dowell and Thompson (2021) proposes data collection tools that allow visualization of social interactions in real time, e.g., integrating computer vision and ubiquitous sensors.	Consequential-ism theory (Normative ethics)	- Maximize useful outcomes of accurate data
Information	**Governance (Fairness, Accountability, Trust, Explainability, Human centricity, Privacy, Inclusivity, Accuracy, Academic integrity, Auditability)** - Compliance with regulations and guidelines (e.g., Children's Online Privacy Protection Rule (COPPA), Family Educational Rights and Privacy Act (FERPA), EU General Data Protection Regulation (GDPR) (Mougiakou, Papadimitriou and Virvou, 2019; Mayfield et al., 2019). - Clear institutional regulation of data protection and access (Pontual Falcão et al., 2022; Costas-Jauregui et al., 2021), for instance, principles for ethical use of learner data published by the Open University (The Open University, 2014).	Deontological (Normative ethics) Philosophical theories of privacy, e.g.: Control Theory of Privacy ***Other relevant non-philosophical theories:*** - Psychology theories, e.g.: Cognitive Dissonance Theory	- Compliance to rules and standards

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	- Establish framework defining the following elements (Peña-Ayala, 2018): (i) *Stakeholders*: Data subjects, data recipients and data curators (ii) *Type of information:* Sensitive attributes, quasi-identifiers, explicit identifiers and auxiliary information (iii) *Data:* Student demographics, educators, courses, assessments, course evaluations and disciplinary actions (iv) *System architecture:* Data access layer, data publishing, statistical disclosure control, differential privacy mechanism, and anonymizer mechanism.	- Techno-Sociology theories, e.g.: Theory of Network Society	
Information	**Stewardship (Fairness, Accountability, Trust, Explainability, Human centricity, Privacy, Inclusivity, Accuracy, Academic integrity, Auditability)** - Quality and inclusive data management practices that promotes transparency in data collection, use and dissemination. Such practices should also adapt to diversity, and technological, social, and educational trends (Martín Núñez and Lantada, 2020). - Useful to include all privacy ethical and legal issues in the inception of system design. Authors proposed to express explicit consent request at the onset of data	Deontological (Normative ethics) Philosophical theories of privacy, e.g.: Control Theory of Privacy ***Other relevant non-philosophical theories:***	- Compliance to rules and standards

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	collection, detailing the purpose, granularity, modality, security policy and the persons that may be involved in the data processing process. Establish clear window periods for the collection and use of data. Record data access, learning intervention and impact of intervention in the design of AI systems (Costas-Jauregui et al., 2021; Tlili et al., 2019). - Ensure that all data subjects are protected from unfair data use (Casas-Roma and Conesa, 2021). Important to consider concerns on use of such data for profiling, tracking and behavioral shaping (Stark and Hoey, 2021). - All stakeholders are to understand the purpose, access, utilization boundaries and the interpretation possibilities of data (Gedrimiene et al., 2020).	- Behavioral Science theories, e.g.: Nudge Theory, Moral Paternalism Theory	
Information, Cognitive	**Availability of Choice (Privacy)** - Data subjects should have the right not to be subjected to AI based decision processing and profiling. Information regarding the logic of, for instance, profiling, should be shared with data subjects (Mougiakou, Papadimitriou and Virvou, 2019).	Virtue theory and deontological theory (Normative ethics)	- Character and moral uprightness - Individual rights to privacy

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	- Understand potential problems and solutions relating to learner opting in or out (Tlili et al., 2019). - Clearly explain all consequences of joining and withdrawal to students. When students are provided the ability to opt in or out, unfavorable consequences should not be placed on students (Gedrimiene et al., 2020). - Merikko et al. (2022) recommends that: (i) Request for data opt-in is made for a specific intervention, rather than requesting for general data consent. For instance, learners can select between personalized AI-driven feedback, general feedback, or no feedback, explicitly stating that only the former requires data disclosure. (ii) Decreasing or low opt-in rate be investigated for better learning intervention. For instance, if there are issues that raised suspicion, if there are reasons why learners opt-out, or if the learners understood the reasons for learning intervention.		
Information, Cognitive	**Collecting Data (Accuracy, Accountability, Privacy, Fairness)** - Use of plain, concise, and easy to understand language,	Deontological theory (Normative ethics)	- Individual rights to privacy

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	when requesting for data consent from data subjects (Mougiakou, Papadimitriou and Virvou, 2019). - Express explicit consent request at the onset of data collection by the AI system, at multiple levels and various interaction touchpoints with the AI system, seeking permission to improve academic developers' understanding of the process of learning on the platform (Costas-Jauregui et al., 2021; Peña-Ayala, 2018).	***Other relevant non-philosophical theories:*** - Psychology theories, e.g.: Self-Determination Theory	
Information	**Input Data (Accuracy)** - Ensure the data collected and used are accurate and clean, allowing stakeholders' access and rectification (Tlili et al., 2018). For instance, to enhance data accuracy, Khairy et al. (2022) proposed to train robots with wide pronunciation data samples, including native and foreign speakers. - Important to consider data context. For instance, for collaborative analytics, it is appropriate to understand and capture multimodal data that best measure collaborative interactions at various levels (e.g., group or individual), contexts (e.g., cultural), and time or phase of collaboration (Schneider, Dowell and Thompson, 2021).	Consequential-ism theory (Normative ethics)	- Maximize useful outcomes of data use

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Information	**Processing Data (Fairness)** - Casas-Roma and Conesa (2021) proposes to ensure data sanitization that sanitizes data of potential discriminatory decisions, such that neither data, model nor predictions affect vulnerable learners.	Consequential-ism theory (Normative ethics)	- Maximize useful outcomes of data use
Information	**Using Data (Accountability, Privacy)** - Mitigate adverse data use. For instance, countermeasures should be in place to prevent the use of data for learner profiling which may impact learner well-being, e.g., discriminatively to infer likelihood of cheating (Kiennert et al., 2019). - Users of data should endorse a consent form, which is updated periodically to include changes in purpose, scope and details of data usage. Use of data are to be done in the spirit of non-maleficence, such that learners' learning experiences and academic performances are not harmed (Khosravi, Sadiq and Gasevic, 2020).	Consequential-ism theory (Normative ethics)	- Maximize non-maleficence outcomes of data use
Information	**Output Data (Accountability, Fairness, Explainability, Auditability, Privacy)** - AI model decisions should be explainable, based on local, global, and generative approaches to provide a	Deontological (Normative ethics)	- Moral reasoning for AI decisions, evaluations,

Domain	**Ethics Mitigation/ Intervention Programs and Activities**	**Theoretical Underpinning**	**Outcome**
	human-understandable explanation for a model's input-output responses (Bommasani et al., 2021). For the latter generative approach, consider designing AI systems that have the ability to provide rationale for autonomous AI decisions (e.g., reasons behind AI recommender decisions) (Hakami and Hernández-Leo, 2020). - Consider utilizing moral reasoning AI systems, that accounts for potential ethical outcomes in AI generated decisions (Casas-Roma and Conesa, 2021). However, Bigman and Gray (2018) notes that humans have poor perceptions of machines making moral decisions. - Use counterfactual explanations to evaluate decisions. Counterfactual explanations allow users to "identify what would have needed to be different in order for the AI to have decided otherwise". This helps to rationalize the AI generated decision (Casas-Roma and Conesa, 2021). **-** In terms of practical examples of explainable AI, Khosravi et al. (2022) shares some useful tools, including the provision of: (i) A modeling workflow schema, detailing the model inputs, processes and outputs.		responses and feedbacks

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	(ii) A navigation explanation flow chart, detailing step-by-step explanations in the user interaction process. (iii) User feedbacks and recommendations, detailing clear rationale on AI-based decisions.		
Information	**Sharing Data (Privacy)** - Useful to clarify and explicitly share third party data sharing scenarios, e.g., institutions sharing with each other, external agencies, or companies to improve AI systems (Tlili et al., 2019). - Key factors that may lead to acceptance of data sharing includes (White et al., 2021): (i) Perception of respect of data collector to data privacy (ii) Trust in data collector (iii) Extent of which data collector benefits from data (iv) Sensitivity of data collection and use (v) Inherent risk of harm Authors share that there are no differences in judgment of acceptance across age, gender of educational levels, and there are no marked differences between educators and learners. - Explore use of emerging technologies including	Consequential-ism theory (Normative ethics)	- Maximize beneficial judgment of data sharing

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	blockchain and smart contracts to manage data securely (Costas-Jauregui et al., 2021).		
Information, Cognitive	**Surveillance (Explainability, Accountability, Privacy, Fairness)** - Pontual Falcão et al. (2022) states that data should not be used beyond academic purposes, for control and surveillance to modify human behavior. Authors share that learners were generally confident about data privacy safeguards that exist in institutions. However, there may exist gender-based differences in safety concerns of AI surveillance (Latham and Goltz, 2019).	Deontological theory (Normative ethics) ***Other relevant non-philosophical theories:*** - Behavioral Science theories, e.g.: Nudge Theory, Moral Paternalism Theory	- Individual rights to privacy

Table 13: Data stewardship and surveillance: Breakdown of ethics mitigation and intervention programs and activities, and their key theoretical underpinning

3. AI-based assessment construction and rollout

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Cognitive	**Examples of Tools and Techniques (Explainability, Inclusivity)** - Implementation of *Student Inspection Facilitator*, a context-independent learning analytics dashboard at the University of Queensland, which provides explainable recommendations to guide learning intervention for educators (Shabaninejad et al., 2022). - *Smart Ecosystem for Learning and Inclusion (SELI)* platform allows students to warn the AI system about their own disabilities. In the case of, for instance motor disability, the platform can grant more or unlimited time to complete a formative assessment (Costas-Jauregui et al., 2021).	Deontological (Normative ethics)	- Compliance to behavioral rules
Cognitive	**Education and Training (Explainability, Accuracy, Accountability, Privacy, Fairness)** - Training to stakeholders, in terms of both AI technology and relevant pedagogical reference models, to understand limitations, possibilities and characteristics of AI-driven assessments (González-Calatayud, Prendes-Espinosa and Roig-Vila, 2021).	Virtue theory and consequential-ism theory (Normative ethics) ***Other relevant non-philosophical theories:***	- Character and moral uprightness - Maximize ethics consciousness and analytical skills

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	- Training of all stakeholders (e.g., learners, educators and administrative staff) can improve trust, and competence to design ethical considerations into the AI systems, and/ or perform validation checks on ethical practices (Hakami and Hernández-Leo, 2020; Tlili et al., 2018).	- Learning theories, e.g.: Cognitive Social Learning Theory - Psychology theories, e.g.: Cognitive Load Theory, Anthro-pomorphism Theory - Techno-Sociology theories, e.g.: Activity Theory, Actor Network Theory - Technology theories, e.g.: Computer Supported Cooperative Work (CSCW) Theory	

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Cognitive	**AI-integrated communication dashboard (Inclusivity, Accountability, Privacy, Trust)** - Create intentionally designed agent representation of identities (e.g., race, appearance, voice, language, gender), rather than relying on data-driven agent representation of identities, which may lack the nuanced understanding of identities (e.g., intersection of marginalized groups). These pedagogical agents can influence students' perception of their own identity and belongingness (Mayfield et al., 2019). - Regarding the sensitivity of communication and feedbacks generated by the AI system, authors proposed to study such communications, with due inputs from educators, psychologists and communication experts (Costas-Jauregui et al., 2021).	Consequential-ism theory (Normative ethics) ***Other relevant non-philosophical theories:*** - Learning theories, e.g.: Kolb's Experiential Learning - Psychology theories, e.g.: Dual Coding Theory, Cognitive Theory of Multimedia Learning - Technology theories, e.g.:	- Maximize beneficial outcomes of communica-tion

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
		Theory of Human-Computer Interaction (HCI)	
Cognitive	**Learning Intervention and Assistance (Explainability, Inclusivity)** - Useful to consider an intervention strategy of education triage, which balances the impact of intervention, with the scope of care needed, the resources available and the number of learners requiring care (Tlili et al., 2019). - Incorporate explainability in intervention and assistance. Personalizing explainable AI-driven hints incorporated in an intelligent tutoring system, may improve learning when undertaking an assessment. This is due to modulation effects of user characteristics on perception and explanation of hints. Explainability of AI-driven hints, incorporating "why" and "how" explanations on how the hints are derived, are useful to effect positive learner perceptions (Conati et al., 2021).	Deontological (Normative ethics)	- Compliance to behavioral rules

Table 14: AI-based assessment construction and rollout: Breakdown of ethics mitigation and intervention programs and activities, and their key theoretical underpinning

4. Administration of assessments using AI systems

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Physical, Information, Cognitive	**Authentication and Security (Academic integrity)** - Privacy preservation protocols ensuring (Peña-Ayala, 2018): (i) Data publishing and third party sharing of sensitive data, that do not leak sensitive data. (ii) Disclosure control and data mining, that do not undermine the identification of individuals tied to sensitive data. - For authentication and security measures, it is useful to consider the following: (i) Implement AI authentication tools such as facial recognition and/or voice recognition (Elshafey et al., 2021). (ii) Pseudonymity to prevent linkability of sensitive data in an event of data leakage and exploitation (Kiennert et al., 2019). (iii) Use of malleable signatures that allows a counterparty to modify the signed information, so that it becomes unfeasible to distinguish between the original signature and the sanitized signature. This retains the validity of the signature but ensures unlinkability of the sensitive data (Kiennert et al., 2019).	Deontological (Normative ethics)	- Cheating detection, prevention, monitoring and evaluation

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
	- Allow sensitive data to be stored at decentralized dedicated entities, or Trusted Third Parties (TPP), for access and retrieval purposes. The focus of the design of AI systems for educational purposes are not aimed at guaranteeing treatment of sensitive data, but meant for e.g., proctoring an assessment. Establishing dedicated TPPs, alongside Public Key Infrastructure (PKI) and Certification Authorities (CA), may be useful to decentralize and improve privacy for data subjects (Kiennert et al., 2019).		
Physical, Information, Cognitive	**Proctoring (Academic integrity)** - For anti-cheating AI proctoring techniques (Elshafey et al., 2021), it is useful to consider: (i) Head pose estimation can be used to track assessment takers' attention. (ii) Gaze estimation to determine angle of students' gaze. (iii) Scene change detection to look for changes in background environment. (iv) Object detection to detect unauthorized objects in environment within camera view.	Deontological (Normative ethics)	- Cheating detection, prevention, monitoring and evaluation

Table 15: Administration of assessments using AI systems: Breakdown of ethics mitigation and intervention programs and activities, and their key theoretical underpinning

5. *AI-facilitated assessment grading and evaluation*

Domain	**Ethics Mitigation/ Intervention Programs and Activities**	**Theoretical Underpinning**	**Outcome**
Information, Cognitive	**Grading and evaluation (Accuracy, Explainability)** - Integrate explainable AI. For instance, it can be useful to apply SHAP at a rubric grading level, which provides robust explanations to individual predictions, while accounting for global factors affecting the performance of the AI model (Kumar and Boulanger, 2020). - Promote accuracy in grading. For instance, for assessment responses in minority dialects, allow flexibility in system design of assessment tasks and languages, provision of topic selection and choice that reflects culturally aligned opportunities, and collaborative sharing of work to receive feedback beyond AI generated responses (Mayfield et al., 2019). In another example, Kim et al. (2021) proposes an AI approach to reliably identify reduced assessment size, and approximate test scores, improving accuracy. - Consider alternative uses of AI that can help mitigate ethics issues in assessments. For instance, Ahn et al. (2021) shares that crowdsourcing of assessment grading can be accurate, in agreement with experts. Further, it provides learning value to the crowdsourced graders.	Consequential-ism theory (Normative ethics)	- Maximize beneficial outcomes of grading

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
Cognitive	**AI-generated responses and feedbacks (Inclusivity, Accountability, Privacy, Trust)** - Create intentionally designed agent representation of identities (e.g., race, appearance, voice, language, gender), rather than relying on data-driven agent representation of identities, which may lack the nuanced understanding of identities (e.g., intersection of marginalized groups). These pedagogical agents can influence students' perception of their own identity and belongingness (Mayfield et al., 2019). - Regarding the sensitivity of communication and feedbacks generated by the AI system, authors proposed to study such communications, with due inputs from educators, psychologists and communication experts (Costas-Jauregui et al., 2021).	Consequential-ism theory (Normative ethics) ***Other relevant non-philosophical theories:*** - Learning theories, e.g.: Kolb's Experiential Learning - Psychology theories, e.g.: Dual Coding Theory, Cognitive Theory of Multimedia Learning - Technology theories, e.g.:	- Maximize beneficial outcomes of communica-tion

Domain	Ethics Mitigation/ Intervention Programs and Activities	Theoretical Underpinning	Outcome
		Theory of Human-Computer Interaction (HCI)	

Table 16: AI-facilitated assessment grading and evaluation: Breakdown of ethics mitigation and intervention programs and activities, and their key theoretical underpinning

3.5.3 Theoretical Implications

Understanding the ethical imperatives of the application of AI in assessments across the assessment creation pipeline, in a conceptual framework as shown in Figure 19, are but the first steps toward informing a critical awareness and a more holistic approach in the governance, stewardship and regulation practices of ethics in this subject matter.

To steer this ethical discourse, building upon the framework proposed by Floridi (2018), it is useful to consider what are (i) politically feasible, (ii) culturally sustainable, (iii) socially acceptable, (iv) institutionally preferred, and (v) legally or rule enforceable, before elaborating and enforcing a system of rules to regulate agents' ethical behaviors. This requires a proactive and constructive normative cascade underpinned by theoretical and empirical foundations, rather than a reactive add-on afterthought that arise from societal backlashes, e.g., when stakeholders are negatively affected or when the public rejects legal recommendations.

A significant aspect of the theoretical discourse in this study discusses hard ethics (i.e., the morally good or bad, or right or wrong duties, behavior, responsibilities, values and rights). This is a fundamentally important discourse as hard ethics helps make and shape regulations. However, there also exists the plane of soft ethics, that embodies normative ground over and above hard ethics, as a "*post-compliance ethics*" (Floridi, 2018). Even if regulations may already exist in the right side of the moral against the immoral divide, regulations do not cover everything, and human agents will need to leverage on ethics evaluation to guide and perform self-regulation of morality, especially if competing interests or values arise that need to be assessed and adjudged. Only then can we achieve good moral citizenry towards a mature infosphere.

3.5.4 Practical Implications

The broad spectrum of mitigation/ intervention programs and activities for ethical imperatives discussed in this study, alongside the underpinning theories, form a practical base to inform the development and implementation of governance and stewardship solutions. The takeaways can be purposed as a thematic guide to future applied research.

From a stewardship perspective, in the design of such programs and activities, it is beneficial to consider resource commitments and impact assessments, starting from smaller scale ethics stewardship exercises. It is useful for each exercise to be approached from these perspectives: (i) build awareness, (ii) signal importance, (iii) engage stakeholders, (iv) provide (transferable) solutions, (v) establish custodians, (vi) develop, use, manage and coordinate solutions, (vii) establish conventions, (viii) institutionalize practices. This can help determine and control data and algorithmic processes for accuracy, privacy and human centricity, devise effective procedures for fair, inclusive, trustworthy and explainable decision making, lower cases of academic cheating, and identify accountabilities and audit procedures.

From a governance perspective, at the institutional level, there can be a Chief (Digital) Ethics Officer and an operational team focusing its time and efforts to oversee related governance, stewardship and education matters. Institutional decision making can be segregated into three top-down levels, namely strategic, tactical and operational. Strategic decisions steer the pivotal long-term vision and culture of ethics integration in the institution. Tactical decisions, informed by strategic guidance, are mid-term focused, and translate strategy into action plans, with the possibility of short-term tactical deviations to address new AI ethics challenges or digital trends that present ethical dilemmas. Operational decisions are the on-the-ground decisions that develop and apply tactical decisions. An institutional feedback culture should be encouraged, to empower stakeholders to raise concerns, especially when the subject matter lies in a sensitive cross section between education, technology, and ethical boundaries.

In the translation of principles to practice, there may also exist ethical risks that should be avoided. Floridi (2021) highlights the following "*ethics regunds*", namely: (i) *ethics shopping*, or the picking and choosing of ethics principles that are justified as a posteriori and retrofitted to pre-existing behaviors, (ii) *ethics bluewashing*, or the implementation of superficial or misleading measures to appear ethical, (iii) *ethics lobbying*, or the use of ethics to avoid or delay good and necessary regulation and enforcement, (iv) *ethics dumping*, or the export or import of unethical activities to a place with less strict regulations, and (v) *ethics shirking*, or the engaging of less ethical works over a period of time to lower the perceived resistance against such works. A misjudgment or misunderstanding can over time, in the lens of Socrates, lead to ethical malpractices.

3.6 Chapter Conclusion

Ethical check and balance should be put in place with the increasingly pervasive use of AI, especially when its growth trajectory appears seemingly aimed towards anthropomorphism, a reality further taking shape by recent advances in generative AI. Possible violations of fundamental human ethics in a societal institution as important as education should be looked upon with scrutiny.

In this study, we looked at how the design and use of AI in education, and in particular, assessments, can conform as closely as possible to basic ethical principles. We systematically investigated the key assessment components and ethical principles highlighted in existing literature, mapped them across the end-to-end assessment pipeline while accounting for different assessment types, and constructed a systematic literature mapping framework highlighting key archetypical research themes. The study took an additional step to raise potential mitigation or intervention programs and activities that can be applied in practice. The proposed systematic literature mapping framework allows researchers and practitioners to deep dive into key thematic research areas, while the latter step facilitates a practical implementation of ethics programs and activities in educational institutions.

Research identified five key archetypical research themes, namely (i) AI system design and check for assessment purposes, (ii) data stewardship and surveillance, (iii) AI-based assessment construction and rollout, (iv) administration of assessments using AI systems, and (v) AI-facilitated assessment grading and evaluation. Ten literature-derived ethical principles, namely, accuracy, privacy, human centricity, fairness, inclusivity, trust, explainability, academic integrity, accountability and auditability, were mapped to these research themes. The study summarizes and rationalizes the impact of each ethical element in each research area, and discusses theoretical and practical implications of the findings.

As highlighted in the limitations section, future work can extend the use of literature databank beyond Scopus, to include e.g., Web of Science, IEEE Xplore or EBSCO Host, in the systematic literature mapping exercise. Furthermore, to account for thematic diversity, the intrinsic value of the diversification of archetypical research themes can be studied.

While this study is based upon the subject of assessments, the ethical imperatives of the discourse has relevance beyond assessments, and can be applied to other areas of AIED. Other future works can contribute to the examination on the underpinning theories relating the ontological, semantics, and the epistemological deliberations and practical applications of ethics in this subject matter, across the spheres of philosophy, learning, psychology, sociology and technology. In addition, practical applications of the actionable insights in this study, in the form of strategic and operational frameworks or case studies, can be another pragmatic endeavor by practitioners and researchers.

Herwix et al. (2022) highlighted the importance of more serious and systematic engagement with the selection, framing and prioritization of ethical issues. There is an emphasis among the state-of-the-art for the need to be more aware, anticipatory, reflecting and informed about the variety of perspectives and contemporary debates concerning AIED ethics. In particular, the relevancy and idiosyncrasy to assessments in our study can help bring forward distinctive actionable applications in this realm.

Chapter 4. Formulating and Validating an Ethical Framework in AI-Enabled Educational Assessments

The rapid integration of AI in educational assessments has ushered in a new era of efficiency and accuracy, yet it has concurrently raised significant ethical concerns. This chapter addresses the critical need for a robust ethical framework in AI-assisted educational assessments.

It presents an in-depth analysis of the triadic ontological framework as discussed in the previous chapter (Lim, Gottipati, and Cheong, 2023), utilizing SEM to validate its applicability and effectiveness. The framework comprises three primary domains—physical, cognitive, and information—and incorporates five stages of the assessment pipeline: system design and check, data stewardship and surveillance, assessment construction and rollout, assessment administration, and grading and evaluation. Key ethical elements such as inclusivity, fairness, accountability, accuracy, auditability, explainability, privacy, trust, human centricity, and cheating mitigation strategies are integrated within this framework.

The research objectives focus on two main areas: (1) validating the triadic theoretical framework through SEM analysis, including an examination of how it reflects learners' perceptions, and (2) investigating the relationships between the stages of the assessment pipeline, ethical imperatives, output variables, and learner perceptions.

Findings from this study reveal significant insights into the interplay between AI-assisted educational assessment stages, ethical imperatives, and learner perspectives. These insights underscore the necessity for frameworks that are not only theoretically sound but also practically relevant and responsive to learner needs. The findings underscore the need for a holistic approach in embedding AI into educational assessments, balancing technological advancement with ethical responsibility and pedagogical effectiveness.

4.1 Introduction

The integration of AI in educational assessments has gained substantial momentum, heralding a new era of enhanced accuracy and efficiency in assessment processes. However, the proliferation of AIED has not been without its ethical challenges, raising critical concerns such as fairness, accountability, privacy, and trust (Nguyen et al., 2023; Memarian & Doleck, 2023). A notable gap in the consensus on ethical principles governing AI in assessments accentuates the necessity for a robust theoretical framework (e.g., Li & Gu, 2023). Such a framework should aim to steer the development and validation of ethical constructs in AI-enabled educational assessments.

Understanding the ethical imperatives of AI application in educational assessments is pivotal for effective governance and stewardship. To this end, as shared in the previous chapter, Lim, Gottipati, and Cheong (2023) proposed a triadic ontological framework that encapsulates the comprehensive architectural assemblage of AI-enabled educational assessment components. This framework draws theoretical support from seminal works by Ashok et al. (2022), Project and Peirce (1998), Popper (1979), and Ogden and Richards (1923). The current study extends this discourse by endeavouring to validate the triadic framework through SEM analysis, focusing on mapping the assessment pipeline and integrating key ethical elements in AI-enabled educational assessments.

The framework elaborates on three primary domains – physical, cognitive, and information – and incorporates five pivotal stages of the assessment pipeline: system design and check, data stewardship and surveillance, assessment construction and rollout, assessment administration, and grading and evaluation. Embedded within this framework are essential ethics elements such as inclusivity, fairness, accountability, accuracy, auditability, explainability, privacy, trust, human centricity, and strategies to mitigate cheating.

A further addition to this research is a segment dedicated to soliciting and evaluating learners' perceptions of AI ethics across these stages of educational assessments. This segment aims to capture the nuanced experiences and attitudes of learners, providing a learner-centric perspective that is often underrepresented in AI ethics discourse (e.g., Jang, Choi & Kim, 2022). By integrating survey and qualitative feedback mechanisms, this study seeks to distil actionable insights from learners' perspectives, enriching the ethical framework with ground-level data on user experience and expectations.

The technical contribution of this study lies in the rigorous validation of the triadic theoretical framework using SEM analysis (e.g., Wang, Sun & Chen, 2023). This methodology offers a comprehensive approach to understand the interplay between different stages of the assessment pipeline, key ethical imperatives, and resultant variables such as learner satisfaction, perceived learning efficacy, sense of academic support, and perceived instructor presence.

This study investigates the following research questions:

1. *RQ7: How do we validate the Triadic Theoretical Framework using SEM analysis?*

   This question aims to validate the triadic theoretical framework, which is pivotal in mapping the assessment pipeline and ethics principles in AI-enabled educational assessments. The validation process uses SEM analysis to scrutinize the framework's efficacy in encapsulating both the operational and ethical dimensions of AI in educational assessments. This includes an in-depth examination of how the framework integrates and reflects learners' perceptions, thereby ensuring that the framework is not only theoretically sound but also practically relevant and responsive to the needs and views of the learners.

2. *RQ8: What are the Relationships that Emerge between the Stages of the Assessment Pipeline, Key Ethical Imperatives, Output Variables, and Learner Perceptions?*
   This inquiry is crucial to understanding the dynamic interplay between these components. It will provide insights into how each stage of the assessment pipeline interacts with ethical considerations, learner feedback, and various output variables. This exploration aims to uncover patterns and correlations that can inform the development of more effective, ethical, and learner-centric AI-enabled educational assessments.

The remainder of this study is organized as follows: Section 4.2 presents a literature review on theory-underpinned key AI applications and ethics elements in AI-enabled educational assessments, focusing on the triadic ontological framework by Lim, Gottipati, and Cheong (2023). Section 4.3 describes the expanded methodology for validating the triadic theoretical framework using SEM analysis. Section 4 presents the results of the SEM analysis, including insights from learners' perspectives. Finally, Section 4.5 concludes the study with implications for the design and implementation of AI-enabled educational assessments, limitations, and directions for future research.

4.2 Literature Review

The integration of AI into educational assessments marks a "bar-raising" shift in learning and evaluation (Dede, Etemadi & Forshaw, 2021). Educational institutions would have to move beyond attempts to resist the use of AI tools (Hargreaves, 2023), as AI applications in educational assessments span a diverse range of beneficial functionalities, from personalized learning environments to sophisticated data analytics for performance evaluation (Rudolph, Tan & Tan, 2023; Lim, Gottipati & Cheong, 2023). These applications are grounded in a confluence of educational, learning, and technology theories, reflecting an evolving paradigm in educational methodologies.

Learner-facing AI applications encompass intelligent tutoring systems (ITS) and personalized learning systems (PLS), among others. These ITS and PLS leverage AI algorithms to dynamically adjust educational content and assessments in accordance with the specific needs and individual learning paces of students. As argued by Fariani, Junus & Santoso (2023) and Xie et al. (2019), these systems align with the principles of constructivist learning theory. Within this framework, learning is characterized as an active, contextualized process in which individuals construct knowledge rather than passively acquire it. The cognitive model of AI-powered personalized systems employs a 'theory of mind' approach, recognizing and accommodating the unique learning trajectories of each student. It also incorporates adaptive testing methods, including cognitive diagnostic assessment. Sun, Wu & Xu (2023) have demonstrated how adaptive testing can align with Bloom's Taxonomy, providing a more nuanced assessment of learners' knowledge and cognitive skills. This personalized approach has the potential to significantly enhance learning effectiveness, with research indicating improvements in learning outcomes of up to two standard deviations (Lee & Soylu, 2023).

Educator-facing AI applications encompass automated writing evaluation (AWE) systems (often referred to as automated essay scoring systems), and predictive AI for learning outcomes, among others. AWE has gained substantial traction, utilizing natural language processing to automate the assessment of written student responses and offer constructive feedback (Huang et al., 2023). This aligns with Mead's social interaction theory and Vygotsky's sociocultural theory, both emphasizing the pivotal role of scaffolded learning and feedback in skill assessment (Ding & Zou, 2024). AI applications extend further into predictive AI, where machine learning models are employed to forecast student performance and learning outcomes. As highlighted by Sghir, Adadi & Lahmer (2023), these predictive models analyze historical data to identify students at risk, enabling timely interventions. Dropout rates in e-learning environments can rise to as high as 80% (Anagnostopoulos et al., 2020). This approach is deeply rooted in the behaviorist evaluation of learning, underscoring the significance of early identification and intervention in formative assessments to reshape learning trajectories and optimize overall outcomes (Duin & Tham, 2020). AI applications can also be approached through the lens of the assessment development and delivery pipeline. In this context, underpinned by Ashok et al. (2022), Project and Peirce (1998), Popper (1979), and Ogden and Richards (1923), Lim, Gottipati & Cheong (2023) conducted a study utilizing network analysis and topic modeling to identify AI application areas throughout five stages of an assessment pipeline. This approach allows researchers and educators to critically examine the implications of AI at each assessment stage, including AI ethical issues. For instance, at the assessment construction stage, AI might raise concerns about algorithmic bias and fairness, which can be evaluated in light of ethical theories such as consequentialism and deontology. The surveillance aspect, when analyzed from an ethical standpoint, may invoke discussions on privacy and surveillance theories, such as the panopticon concept. This perspective facilitates a more holistic evaluation of AI's impact on educational assessments, and in particular, from an AI ethics perspective.

The application of AIED has precipitated a range of ethical imperatives that are critical to address for the responsible use of this technology (Nguyen et al., 2023). Ethics principles and frameworks do not emerge in isolation; rather, they exert a significant influence on research and development. The viewpoints articulated within them do not merely coexist with technological progress but actively shape research projects and methodological advancements, thus playing a pivotal role in defining expectations, values, and objectives within the sphere of technological development. This is evident in how issues like algorithmic bias and algorithmic fairness have now become routine topics of discussion for companies and institutions when implementing AI systems. The number of such guidelines have grown to as much as 200 in 2023 (Corrêa et al., 2023).

The exploration of ethical principles and frameworks in the context of AIED has been an area of growing academic interest. For instance, Holmes et al. (2021) explored ethical principles by surveying 60 leading AIED researchers and developed a ‘strawman’ draft ethics framework for AIED that mapped the ethics of algorithms in education, ethics of data used in AI and ethics of learning analytics. On a more granular level, Hong et al. (2022) further proposed an AIED data ethics framework that considered data processes from collection to disposal. It is beneficial for AI ethics discourse to be “*specific enough to be action guiding*” (Whittlestone et al., 2019). In the context of educational assessments, Lim, Gottipati & Cheong (2023) recognized the idiosyncratic relevance of considering fundamental ethical principles in assessments that would provide "*concrete property instantiations of applied ethics*".

The *Triadic AIED Assessment Framework* proposed by Lim, Gottipati, and Cheong (2023) (Figure 20) forms the cornerstone of this study. This framework, built upon a systematic literature survey, is an innovative approach to understanding and integrating AI in educational assessments, emphasizing the application across five key stages of the educational assessment pipeline across the physical, cognitive, and information domains.

The *physical domain* pertains to the tangible aspects of AI systems, including hardware and infrastructure essential for AI deployment in assessment settings. It encompasses the device and network physical components necessary for robust and secure systems capable of handling the demands of educational data processing, as well as the physical interface between the AI system and users. *Cognitive domain* focuses on the AI algorithms and data that drive the assessment processes. This content layer involves the development and application of intelligent algorithms capable of adapting to diverse learning styles and needs. This domain also encompasses the AI's ability to analyze and interpret data, providing insights into student learning and performance. *Information domain* is the service layer includes AI-underpinned user interface and interaction aspects of educational assessments.

In the triadic framework, the five key stages of assessment pipeline are, namely: (i) *AI system design and checks for assessment purposes,* (ii) *data stewardship and surveillance,* (iii) *AI-based* *assessment construction and rollout,* (iv) *administration of assessments using AI systems, and* (v) *AI-facilitated assessment grading and evaluation*. *AI system design and checks for assessment purposes* is directed towards the development of a robust and secure AI system, spanning across physical infrastructure, cognitive functionalities, and data management protocols (e.g., Lin, Huang & Lu, 2023). *Data stewardship and surveillance* focuses on data governance aspects and, if necessary, the implementation of surveillance measures (e.g., Williamson, Bayne & Shay, 2020). *AI-based assessment construction and rollout* revolves around leveraging AI capabilities to construct, deliver, and optimize assessments, fostering streamlined communication and formative feedback mechanisms (e.g., Dai & Ke, 2022). During the *administration of assessments using AI systems* stage, AI is instrumental in upholding the integrity of assessments through rigorous authentication and security measures, including proctoring and plagiarism detection (e.g., Surahman & Wang, 2022; Nigam et al., 2021). Finally, *AI-facilitated assessment grading and evaluation* play a central role in the interpretation of assessment responses, performance measurement, and the provision of insightful feedback (e.g., Ramesh & Sanampudi, 2022).

There exist different emphases of key ethical elements in each stage of the assessment pipeline. Each presenting unique challenges, considerations and emphases in the context of AI-enabled educational assessments, these ethical issues include:

Fairness: This ethics principle highlights the imperative of upholding fair, equitable, and appropriate assessment practices within AI systems, acknowledging the complexity in defining fairness due to subjectivity, context, and cultural nuances. Fairness, in this context, encompasses the elimination of data and algorithmic bias to ensure diversity, equity, and non-prejudice, preventing disadvantages for minority groups and aligning with inclusivity. Ethical considerations encompass "allocation harm" for equitable resource distribution, "representational harm" to combat bias marginalizing learner groups, and concern about unintended learner profiling (Mayfield et al., 2019). Additionally, it emphasizes the importance of avoiding universal emotional assumptions in socio-emotional assessments due to cultural differences (Stark & Hoey, 2021) and underscores the need for standard ethical codes and robust monitoring mechanisms for the effective implementation of fair assessment practices in AI systems (Tlili et al., 2018).

Inclusivity: This ethics principle underscores the significance of inclusivity and accessibility within AI systems for education, particularly in personalized, large-scale settings. It highlights key insights, such as the need to exhibit empathy towards learners' diverse conditions, including health, disabilities, gender, race, educational backgrounds, and socio-economic status. Additionally, it emphasizes the importance of sensitivity and supportiveness of AI-generated communication and feedback (Costas-Jauregui et al., 2021), and addresses the potential for AI decisions to perpetuate conformity, peer pressure, or segregation (Gedrimiene et al., 2020). In essence, the principle seeks to foster empathetic and inclusive AI systems that promote a diverse and equitable learning environment.

Accountability: This ethics principle emphasizes responsible AI system design and operation across various contexts and roles. It highlights the need for those involved in AI systems, particularly in education, to be responsible stewards of data, as students often lack influence in data handling (Gedrimiene et al., 2020). Decision-makers must provide clear reasons and take responsibility for AI-driven outcomes (Hakami and Hernández-Leo, 2020). Compliance with regulations can be challenging due to inconsistencies, and the ownership of shared data remains uncertain, posing obstacles to accountability efforts (Costas-Jauregui et al., 2021). Additionally, the principle underscores the importance of avenues for addressing the adverse consequences of AI system use, both at individual and societal levels.

Accuracy: This ethics principle underscores the importance of accuracy in AI assessments to maintain their reliability and validity. Key drivers of accuracy include the necessity of ensuring high-quality data inputs to prevent negative impacts on AI-driven decisions (Tlili et al., 2018), addressing imbalanced datasets to avoid discriminatory outcomes (Chounta et al., 2022), accurately interpreting learner responses, ability to handle prediction errors (Khairy et al., 2022), and possibly guarding against students' gaming of AI systems to their academic advantage (Tlili et al., 2018).

Auditability: This ethical principle highlights the necessity of allowing independent third-party assessors the authority to examine and report on the utilization and configuration of data and AI algorithms in assessment processes. Auditability pertains to comprehending, validating, and reviewing AI systems to ensure adequate traceability, transparency, and utilization of data and AI algorithms, as well as ensuring the credibility and dependability of assessment tools. It should be acknowledged, however, that difficulties may arise when dealing with proprietary algorithms, as indicated in the studies by Tlili et al. (2019) and Casas-Roma and Conesa (2021).

Explainability: This ethics principle underscores the vital need for transparency and comprehensibility in AI systems. It emphasizes making data, AI algorithms, and AI-driven decisions easily understandable for relevant stakeholders and justifying their use in a non-technical manner (Casas-Roma & Conesa, 2021). Achieving transparency in AI system design, information accessibility, and user comprehension is crucial for fostering trust and fairness among human stakeholders. It also advocates for clear rationales in AI recommendations while acknowledging the balance needed to protect proprietary algorithms (Latham and Goltz, 2019). Furthermore, it highlights the tradeoff between interpretability and complexity in AI systems, emphasizing the importance of simpler models when feasible, and the necessity of a sound theoretical basis for applying AI in assessments, incorporating pedagogical principles and AI training for meaningful implementation (González-Calatayud, Prendes-Espinosa & Roig-Vila, 2021).

Privacy: This ethics principle emphasizes the importance of safeguarding individuals' privacy and data protection in AI systems, from data collection to disposal (Chounta et al., 2022). It highlights the need to manage sensitive data securely and respect individuals' emotions and expressions in data usage to maintain trust. Obtaining explicit consent, especially from minors, and allowing flexibility for opting in or out of AI-related activities are crucial for fair data practices, despite potential challenges and data gaps. Constant surveillance resulting from AI use raises concerns about privacy infringements, and potential anxiety and behavioral changes resulting from such surveillance (Megahed, Abdel-Kader and Soliman, 2022).

Trust: This ethics principle emphasizes the significance of trust in AI systems and data utilization for assessments. Trust involves confidence in AI systems' decision-making abilities and feedback quality, with concerns arising from the absence of human-like attributes, potential biases, and doubts about improvement (Pontual Falcão et al., 2022). Furthermore, trust is also tied to the level of autonomy and control granted to educators and learners; excessive intrusion and reduced autonomy can erode trust in AI systems. Additionally, trust relies on a consensus regarding AI system purposes, particularly evident in socio-emotional assessments, where a lack of agreement can lead to ethical discrepancies affecting trust (Stark & Hoey, 2021).

Human Centricity: This ethics principle underscores the need to prioritize human agency and dignity. It emphasizes care on positive states of human wellbeing, user protection from AI manipulation of learner behaviors and emotions, and intervenability and reversibility of AI processes for correction, termination, erasure and blocking when learners' level of autonomy and/or capacity to learn are/is diminished (Mougiakou, Papadimitriou & Virvou, 2019).

Academic Integrity: This ethical principle focuses on uncovering and discouraging deceitful conduct by learners, especially when it comes to assessments. It entails the utilization of AI-powered monitoring and plagiarism detection techniques to spot instances of cheating, whether they occur in physical exam settings or in online assessment platforms (Elshafey et al., 2021; Kiennert et al., 2019).

The triadic theoretical framework provides a comprehensive and structured approach to understanding and evaluating AI applications in educational assessments. By encompassing the physical, cognitive, and information domains, and applying these across the various stages of the assessment pipeline, the framework ensures a holistic evaluation of AI tools, grounded in ethical considerations and practical effectiveness. Addressing these ethical imperatives is not just about mitigating assessment risks but also about harnessing the potential of AI in assessments in a manner that is equitable, responsible, and beneficial for all stakeholders. Beyond Lim, Gottipati & Cheong (2023), no prior study has studied AI ethics in the contextual level of educational assessments. This study will build upon the work of Lim, Gottipati & Cheong (2023) for framework validation, providing a structured approach to evaluating AI tools in educational assessment settings.

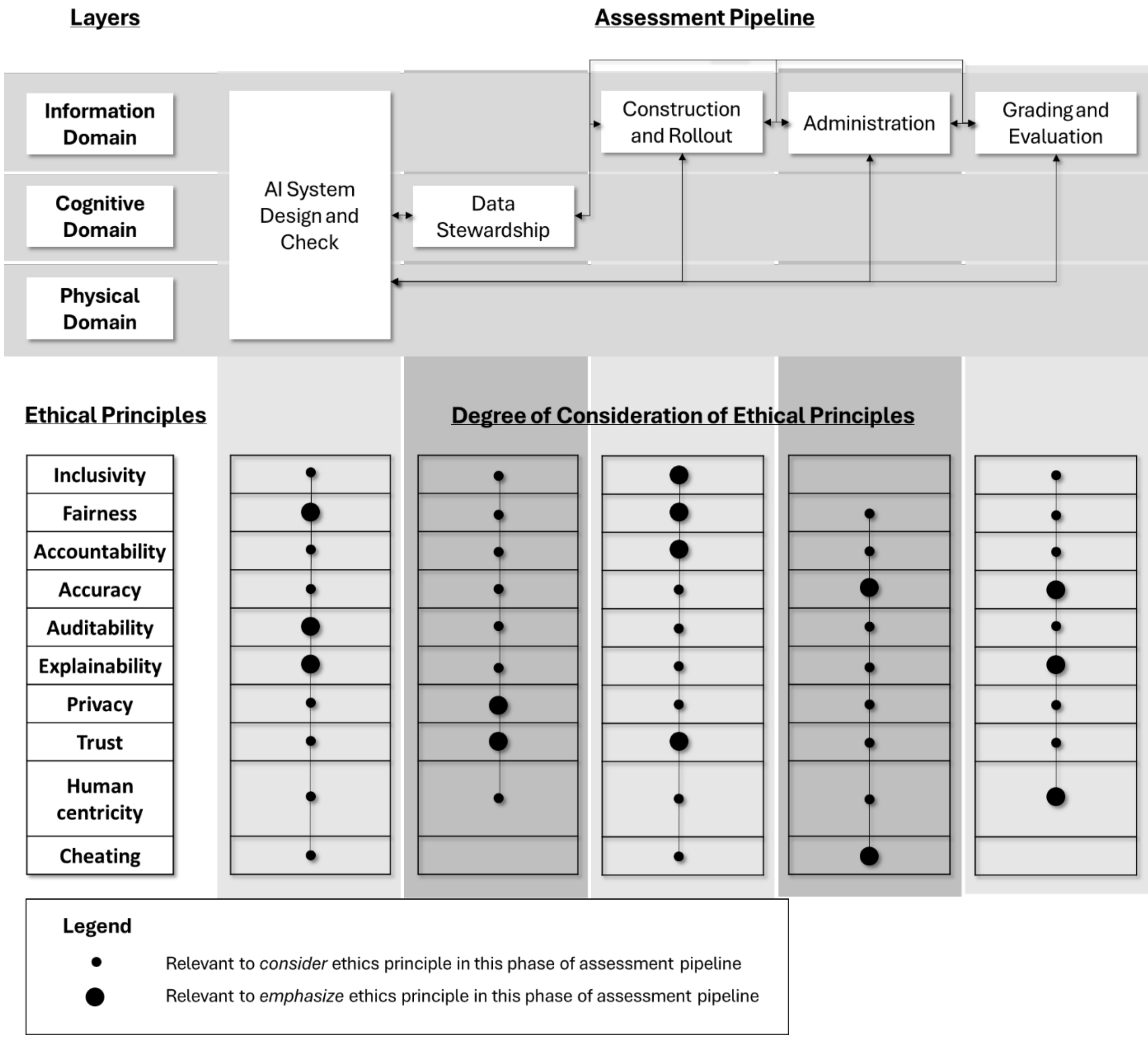


Figure 20: Triadic AIED Assessment Framework proposed by Lim, Gottipati, and Cheong (2023)

4.3 Methodology

4.3.1 Conceptual Framework and SEM

The previous chapter explained the Triadic AIED Assessment Framework as proposed by Lim, Gottipati & Cheong (2023) through a systematic literature mapping approach. In this section, we present the conceptual framework of our study (Figure 21), which aimed to validate the framework through SEM and empirical data examination of the relationships of endogenous ethic input variables, exogenous pedagogical outcome output variables, and the latent constructs of the stages of the assessment pipeline as conceptually introduced in Lim, Gottipati & Cheong (2023) (Lim et al., 2023). The investigation looked to understand the ethical principles that exert significant influence on each of these assessment pipeline stages. Furthermore, the study postulated that the consideration of ethical principles throughout the assessment pipeline is conducive to enhancing pedagogical outcomes. In considering pedagogical outcome output variables for validating the framework with SEM, the literature suggests focusing on aspects like communication, support, and presence in the learner-instructor interaction, as these factors are significantly influenced by AI systems in educational settings and can directly impact student outcomes like motivation, satisfaction, and achievement (Seo et al., 2021). In this study, theoretically grounded output variables include: (i) *Learner satisfaction*: Overall satisfaction with AI-assisted education (e.g., Kashive, Powale & Kashive, 2020); (ii) *Perceived learning efficacy*: Perceived effectiveness of AI in enhancing learning outcomes (e.g., Shi, Li & Zhang, 2024); (iii) *Sense of academic support*: Degree to which students feel supported academically by AI tools (e.g., Wu & Yang, 2022); and (iv) *Perceived instructor presence*: How the use of AI impacts the perceived presence and engagement of instructors (e.g., Bolick & da Silva, 2024).

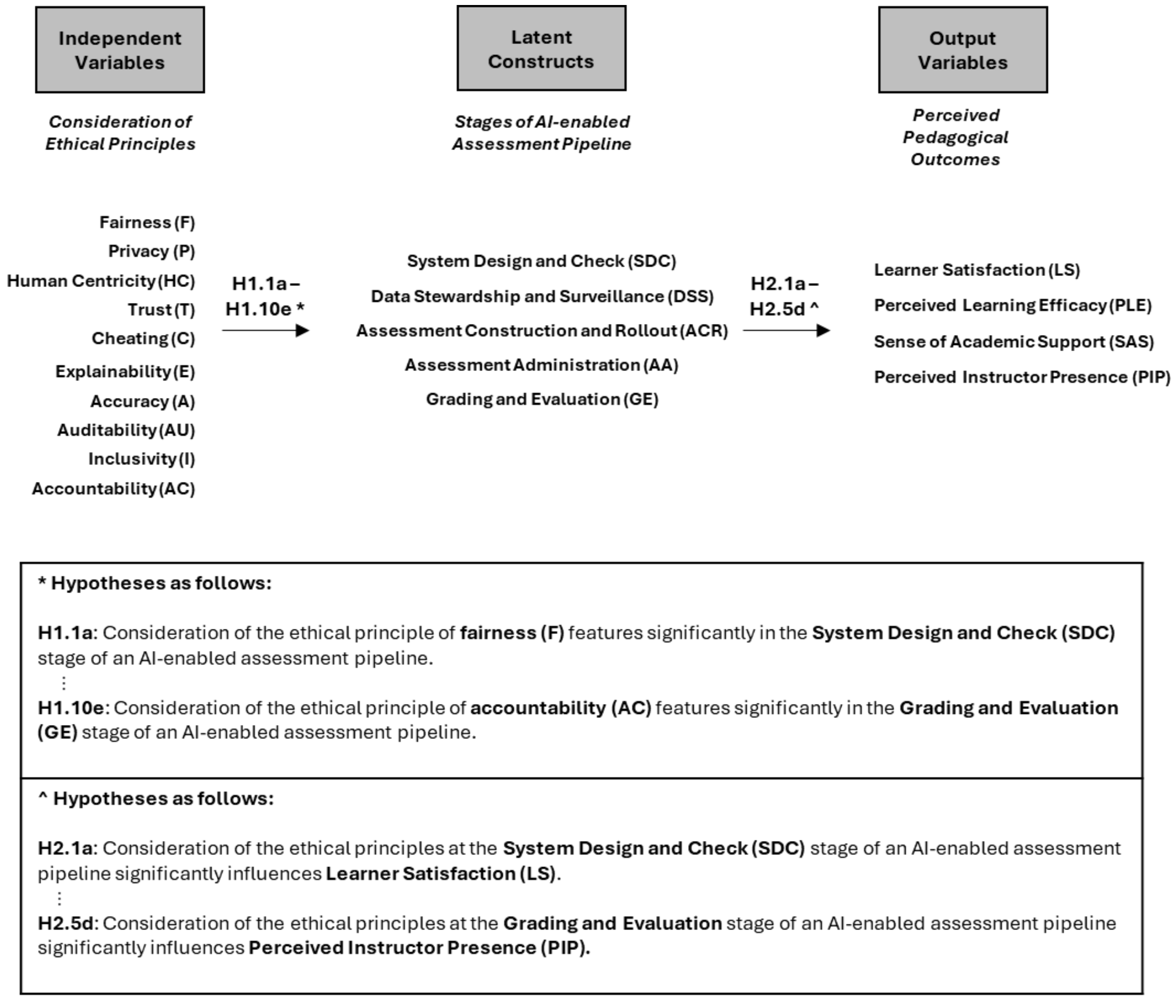


Figure 21: Conceptual framework to validate Triadic AIED Assessment Framework

4.3.2 Data Collection and Survey Instrument

In this pilot study, a sample size of 95 anonymized and voluntary undergraduate participants were included in the study. The age range of the participants was between 18 to 22; 47 males (49%) and 48 females (51%). All participants have varying levels of prior knowledge and exposure to AI assessment systems. To ensure survey participants understood the survey questions accurately, at the start of each survey conducted, survey participants were provided with definitions of the ethical principles as outlined by Lim, T., Gottipati, S., & Cheong, M. (2023), along with explanations of the SEM output variables related to the pedagogical outcomes.

The mixed-method research instrument devised in this study aimed to encapsulate user perceptions from a learner-centric perspective. The quantitative survey (refer to **Appendix III**) aspect comprised five items for each of the ten ethical dimensions for input variables, and two items for each of the four pedagogical outcomes for output variables, resulting in a total of 58 items. Responses were solicited using a 10-point Likert scale, ranging from (1) strongly disagree to (10) strongly agree. Qualitative open-ended responses were also added to capture user insights, context and nuances. Responses are coded with ("RXX"), where XX (e.g. 11) represents the ordering of the student respondent (i.e., 11th respondent of sample pool). Students participate in the anonymized survey on a voluntary basis. Students' grading would not be impacted in any way through participation (or non-participation) of the survey.

To ensure the content validity of the 58-item survey, a panel of six experts was enlisted for evaluation. Among these experts, three possessed specialized knowledge in the domain of AI, while the remaining experts were in the field of AI education. Each expert was requested to assess every item on a scale of (1) not relevant, (2) somewhat relevant, (3) quite relevant, and (4) relevant. Items garnering ratings of 3 or 4 were considered relevant, whereas those with ratings of 1 or 2 were deemed not relevant. The widely accepted Content Validity Index (CVI) was employed as the metric for content validity. Two distinct types of CVIs were computed: item-level CVIs (I-CVIs) and scale-level CVIs (S-CVI). The I-CVI was determined as the proportion of panel experts who assigned a rating of 3 or 4 to an item, effectively categorizing them as relevant. The S-CVI, as employed in this study, denotes the mean proportion of items rated 3 or 4 by the panel of experts. In addition, an assessment of the survey instrument's reliability was conducted through the computation of the Cronbach Alpha coefficient. To ascertain the suitability of employing factor analyses, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was examined.

4.4 Findings and Discussion

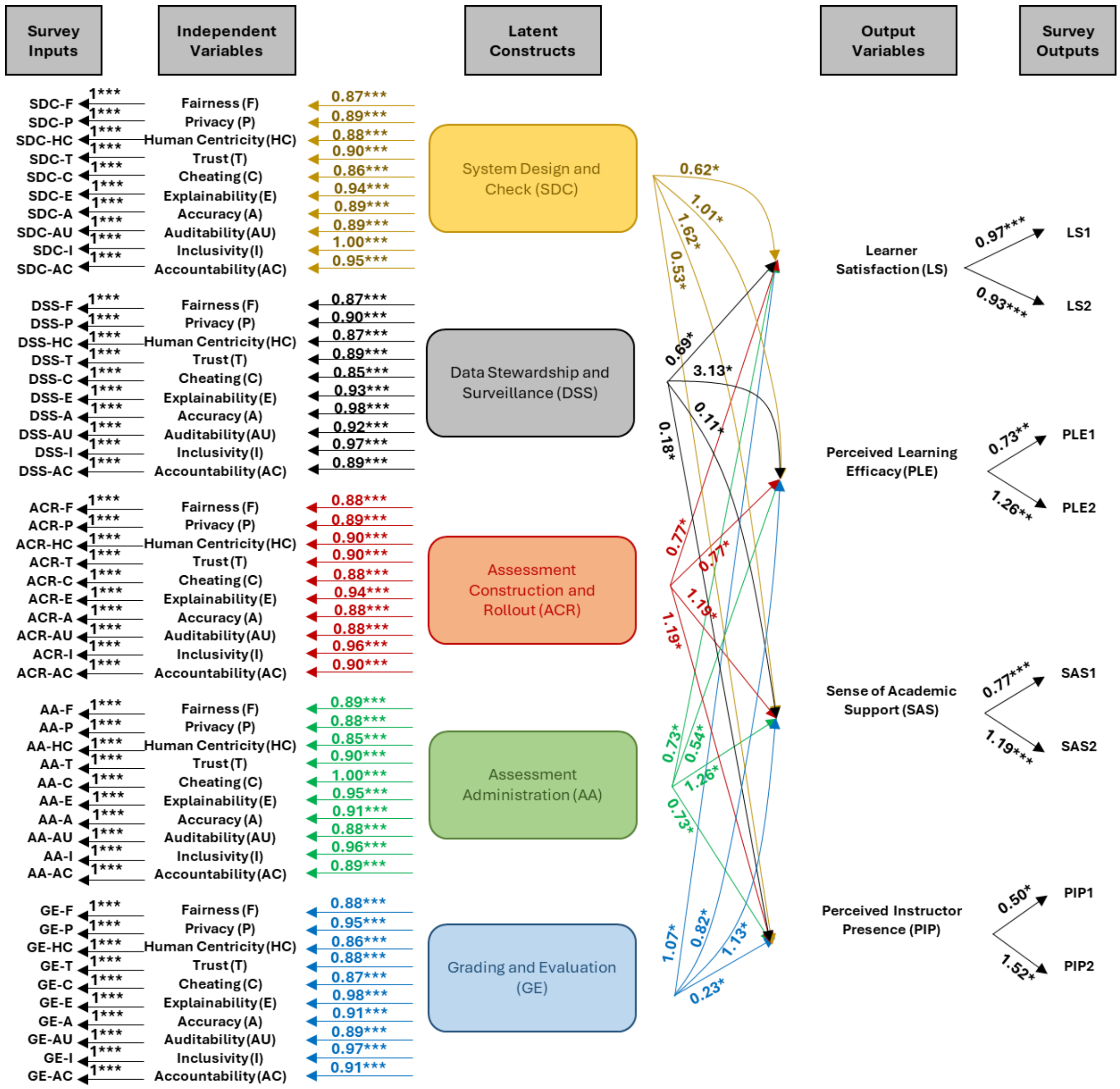
4.4.1 RQ7: Validation of Triadic Theoretical Framework through SEM Analysis

To ensure the robustness of our survey instrument, we followed established guidelines (Jang, Choi & Kim, 2022), which recommend an I-CVI value of at least 0.83 and an S-CVI value of at least 0.90, as indicators of content validity when assessed by a panel of six experts. Our findings, with S-CVI values ranging from 0.92 to 0.97, surpassed these thresholds, confirming the content validity of the survey instrument. Additionally, the survey instrument exhibited good internal consistency, with a Cronbach's Alpha coefficient of 0.95, indicating a high degree of reliability. Furthermore, the dataset's KMO measure of 0.92 suggested a well-structured dataset, supporting the suitability of conducting factor analyses.

Structural equation modeling, conducted using the R package *lavaan*, resulted in a model that demonstrated a good fit to the data. The statistical indices were as follows: χ^2(182) = 20.333; Comparative Fit Index (CFI) = 0.909; Tucker-Lewis Index (TLI) = 0.895; Standardized Root Mean Square Residual (SRMR) = 0.059; and Root Mean Square Error of Approximation (RMSEA) = 0.035. A model is considered to fit well when incremental fit measures (CFI and TLI), typically above 0.9 with 1 being optimal. The SRMR focuses on the residuals of the model, offering a direct measure of fit based on the observed and predicted data. The RMSEA, in contrast, provides an estimate of how well the model might fit the population's covariance matrix, adjusting for model complexity. Both of these measures indicating good fits (below 0.08) support the conclusion that the model is appropriately specified and represents the data well (Hu & Bentler, 1999).

Figure 22 provides a simplified depiction of the model.

Ethical principles exert significant influence on each of these assessment pipeline stages, at statistical significance of less than 0.001 level; in turn, the consideration of ethical principles throughout the assessment pipeline significantly enhances pedagogical outcomes, at statistical significance of less than 0.05 level. This clear association aligns with hypotheses H1.1a to H2.5d, validating the framework and demonstrating significant relationships between the endogenous ethic input variables, exogenous pedagogical outcome output variables, and the latent constructs of the stages of the assessment pipeline.



*Significance at the 0.05 level; **Significance at the 0.01 level; ***Significance at the 0.001 level.

Figure 22: Structural Equation Model

4.4.2 RQ8: Examination of Relationships between Assessment Pipeline Stages, Ethical Imperatives, Output Variables, and Learner Perceptions

The examination of relationships between assessment pipeline stages, ethics imperatives, pedagogical outcomes, and learner perceptions illuminates the complex interplay between technology, ethics, and human experience in assessment settings. By dissecting the hypotheses guided by SEM analysis and survey results, we aim to provide fundamental understanding of how AI in assessments influences ethical considerations and shapes learners' perceptions.

Fairness: This ethical principle features moderately across all assessment stages. Most students believed that the notion of fairness is important. Assessment systems should not favor specific students or prejudice marginalized groups, through intentional or unintended profiling or labeling of students. Students were not sure if AI systems would assess students from different backgrounds in a fair manner. Students cited concerns where *"if AI systems are built based on certain [learner profiles], would they discriminate against others?"* (R22). There should exist contextualization, objectivity and cultural specificity. Some students shared that "*explaining clearly how AI systems work would help*" (R46). Some students tied fairness to trust, sharing that fairness would lead to greater trust and "*acceptance of AI systems*" (R25). Fair AI systems would provide *learner satisfaction* and a perceived *sense of academic support*.

Inclusivity: This ethics principle is featured highly across all assessment stages. Students overwhelmingly endorse assessments and associated resources that foster diversity and equity. AI systems should be supportive and have unqualified "*accessibility to all*" (R33) (similar to results in Tlili et al., 2019). AI feedback should exhibit empathy and be "*judgment-free and welcoming*" (R56) (similar to results in Costas-Jauregui et al., 2021). While many students are uncertain about the possibility of AI systems achieving complete inclusivity, they place their trust in educational institutions to strive for it. Inclusive AI systems would provide *learner satisfaction* and a perceived *sense of academic support*.

Accountability: This ethics principle features highly in *system design and check* stage, and to lesser extents, *assessment construction and roll out*, *grading and evaluation,* and *data stewardship and surveillance* stages. Many students believed that educational institutions, and mainly "*the people in charge*" (R62) should take responsibility for AI issues (similar to results in Kong, Cheung & Zhang, 2023 and Gedrimiene et al., 2020). "*The creator of AI [assessment] systems are humans, not AI itself*" (R29). Most students trust that compliance with relevant educational guidelines and regulations should suffice, and educational institutions would demonstrate accountability. The presence of accountability in AI systems would provide a perceived *sense of academic support*.

Accuracy: This ethics principle is featured highly at the *data stewardship and surveillance* stage. Data quality is of overwhelming importance amongst the students (similar to results in Tlili et al., 2018), as it supports all other stages of the assessment pipeline. While students would prefer having the option to opt in or out of data collection, they understood that "*opting out may affect how representative the data is*" (R56), and impact accuracy in AI-based assessments (similar to results in Berendt, Littlejohn & Blakemore, 2020). Despite this, most students expressed confidence in AI's potential to enhance the validity and reliability of assessment instruments, mitigate prediction biases, and detect violations of assessment integrity. Leveraging appropriate data sources, data-driven insights can enhance objectivity and perceived credibility of assessments. Students note that AI systems maintain consistent accuracy levels and do not succumb to human fatigue or emotional states. AI systems *"don't get tired when grading our assessments*." (R47). "*Sometimes we can sense impatience or irritation in the voice of lecturers when we receive [assessment] feedback, so we stop asking"* (R10). Provided students cannot exploit AI systems to their advantage, they generally associate AI systems with accuracy, anticipating improved *learning efficacy* with AI systems.

Auditability: This ethical principle features moderately across all assessment stages. Students generally believed that AI system's processes and decisions should be auditable by independent third parties. As "*assessment outcomes can affect [a student's] learning, motivation and achievements*" , hence there should be "*some kind of regular audit trails and logging*" (R02) to assess the actions and decisions of AI systems. Most students trust that educational institutions "*will take care of such check and balances*" (R44). Auditable AI systems would provide a perceived *sense of academic support*.

Explainability: This ethics principle features highly across all assessment stages. Students believed that actions and decisions of AI systems should be clear and explainable *"in layman terms*" (R13). They stressed the importance of reliability and validity in assessment tools, as discrepancies or ambiguities stemming from unexplainable or unreliable AI actions can "*affect trust on AI systems*" (R48). This should include understanding the "*input, model and output of AI systems*" and "*how they change over time*" (R03). A lack of explainability may prompt students to "*challenge grading*" (R20) and increase instances of reevaluation (similar to results in Seo et al., 2021). Such AI systems might be more open to misunderstandings. Conversely, AI systems that prioritize user understanding and accessibility of information, grounded in pedagogical principles, foster a positive sense of *perceived learning efficacy*.

Privacy: This ethics principle is featured highly at the *grading and evaluation* stage, and moderately at all other stages. Students prioritized their autonomy in controlling and divulging their personal data (similar to results in Latham and Goltz, 2019), especially when they are tied to grading, evaluation and academic achievement (similar to results in Berendt, Littlejohn & Blakemore, 2020). Students preferred the flexibility to opt in or out of data collection processes. Nevertheless, they acknowledged the potential trade-off between privacy and the optimal delivery of educational experiences (similar to results in Tlili et al., 2019), showing a pragmatic rather than absolute stance on this issue. While students generally endorsed research-related behavioral surveillance "*provided anonymity is preserved*", they expressed discomfort with "*continuous educational monitoring*" (R60), including intrusive methods like biometric data collection such as eye-tracking and facial expression analytics, echoing surveillance concerns raised in Seo et al. (2021). A prevailing trust among learners was that educational institutions serve as "*protected space*" (R34) safeguarding privacy. An academic environment that supports privacy can lead to *learner satisfaction* and provide a perceived *sense of academic support*.

Trust: This ethics principle is featured moderately in the *assessment construction and roll out*, *system design and check*, *data stewardship and surveillance* stages, and to lesser extent, *assessment administration* stage. Students were comfortable interfacing with AI systems and believed that students' learning autonomy could be respected (in contrast to findings by Henne & Gstrein, 2023). They emphasized the importance of social interactions and human input, aligning with findings from Pontual Falcão et al. (2022), and Park and Kim (2020). While uncertain about AI's potential impact on reducing instructors' involvement in assessments, students expressed optimism that instructors can "*focus on more meaningful assessable content*" (R55). Students were neutral on whether AI auto-grading systems can fairly grade assessments with diverse learning backgrounds. There were concerns about AI's ability to appropriately interpret assessment responses, personalize and respect nuanced learning decisions of users. Students perceived that AI systems might lack contextual information and innately *"human social interaction skills and emotions*" (R39), and mistakenly adjudge an assessment response to be incorrect or learning decision to be inappropriate (similar to results in Khairy et al., 2022). Students also expressed apprehension about a possible lack of recourse to "*clarify misunderstandings when AI misinterprets [their] responses"* (R44). Most students in this study displayed moderate trust to AI systems. A trustworthy human-in-the-loop AI system can lead to *learner satisfaction* and high *perceived instructor presence*.

Human Centricity: This ethics principle is featured moderately in the *assessment construction and roll out* stage, and to lesser extents, in *system design and check* and *data stewardship and surveillance* stages. Learners generally believe that AI systems can facilitate academic help-seeking behavior while preserving learning autonomy. The anonymity provided by AI may reduce self-consciousness, thereby "*encouraging more questions*" (R23) in self-regulated learning (similar to results in Adams et al., 2023). There was a slight concern about the potential diminishment of learning agency and ownership due to overreliance on AI systems. This might foster a "*false sense of security*" (R42) on learning efficacy. This said, these concerns were offset by the benefits of enhancing efficiency (similar to results in Adams et al., 2023), promoting creativity (contrast with 'stifling of creativity' in Adams et al., 2023) and provision of just-in-time support (similar to results in Seo et al., 2021). Most learners perceived that human agency and dignity would be respected (contrast with reducing humans to 'objects' in Berendt, Littlejohn & Blakemore, 2020), and AI system could deliver a positive *sense of academic support*.

Academic Integrity: This ethical principle features highly in *assessment administration* stage, and to lesser extents *assessment construction and roll out*, and *system design and check* stages. Most students supported the notion of academic integrity for fair assessments, and believed that it is crucial to incorporate features to prevent students from unfairly cheating in AI-based assessments. Cheating is considered serious academic breaches of integrity and most students would be "*scared*" (R41) if they are penalized for plagiarism. Students shared that they "*try to submit early to check against Turnitin similarity detection software*" (R61), to avoid unintended plagiarism and revise problematic segments (similar to results in Stone, 2023). However, most students expressed confusion and a lack of understanding about AI plagiarism detection. Students expressed that it is "*vague how AI plagiarism works*" (R33). "*Unlike standard similarity detection where we get to see the sources of plagiarism, AI plagiarism does not do that*" (R31). In general, students were unsure if AI systems can fairly and confidently identify cheating without falsely accusing honest students, or being misled by cheating students. AI systems that can fairly and confidently identify cheating cases could deliver a positive *sense of academic support*.

4.4.3 Consideration of Learner Perspective in AI Ethics and its Impact on the Framework

The concept of human-centered AI design emerges as a crucial approach to addressing ethical imperatives within education, and in particular, in educational assessments. By prioritizing human values, needs, and experiences, educational institutions can cultivate ethical reasoning, empathy, and social responsibility to harness AI's potential in assessment creation. This human-centered ethos reflects a broader trend towards humanizing technology within educational contexts.

Here are some key insights: Firstly, the discourse surrounding trust in AI systems reveals a complex interplay between confidence and skepticism. Despite displaying a level of moderate trust, students concurrently harbor reservations regarding AI's inherent limitations, biases, and potential errors. This duality underscores the nuanced nature of trust within technological frameworks, where assurance coexists with doubt. It underscores the imperative for transparency, accountability, and ongoing scrutiny to foster trust. Secondly, AI explainability plays a critical role of transparency in fostering pedagogical trust and understanding. By clarifying AI processes and decisions, educational institutions seek to empower learners with insights into assessment mechanisms, promote pedagogical transparency and help learners understand their learning trajectories. Students expressed concerns about potential biases in AI systems and emphasize the importance of fairness and inclusivity. This highlights a critical challenge in AI-based assessments: ensuring that algorithms are unbiased and do not perpetuate discrimination or inequity. Addressing these concerns requires careful attention to algorithmic design, data quality, and transparency. Thirdly, the discussion around privacy underscores the complex trade-offs between data privacy and the delivery of optimal educational experiences. While students value autonomy and control over their personal data, they also recognize the potential benefits of data-driven insights and personalized learning experiences. This trade-off between privacy and personalization raises questions about how educational institutions should consider the ethical and practical dimensions of data usage in learning environments. Fourthly, the discussion on cheating detection illuminates the dynamics between maintaining academic integrity and building trust in AI-based assessment systems. While students support measures to prevent cheating, there are doubts about the fairness and accuracy of AI plagiarism detection. This underscores the importance of developing robust and transparent cheating detection mechanisms that inspire confidence among users.

In this study, we validate how ethics may serve as foundational pillars throughout the assessment pipeline. Each stage, from design to implementation, is imbued with ethical considerations, emphasizing the inseparable link between ethical principles and assessment practices. This association highlights the need for a holistic ethical framework that permeates every aspect of assessment development and administration.

This said, SEM analysis in this study identified areas of improvement to the framework. In particular, we revise the degree of consideration of ethical principles in each stage of the assessment pipeline, basing on actual user perceptions rather than systematic literature survey. Here, we identify ethical principles with factor loading of > 0.95 to represent primary ethical principles to consider at each assessment stage; $0.90 \leq 0.95$ to represent secondary considerations;

< 0.90 to represent ethical principles that are useful to consider, but not of key considerations. Figure 23 shows the revised framework.

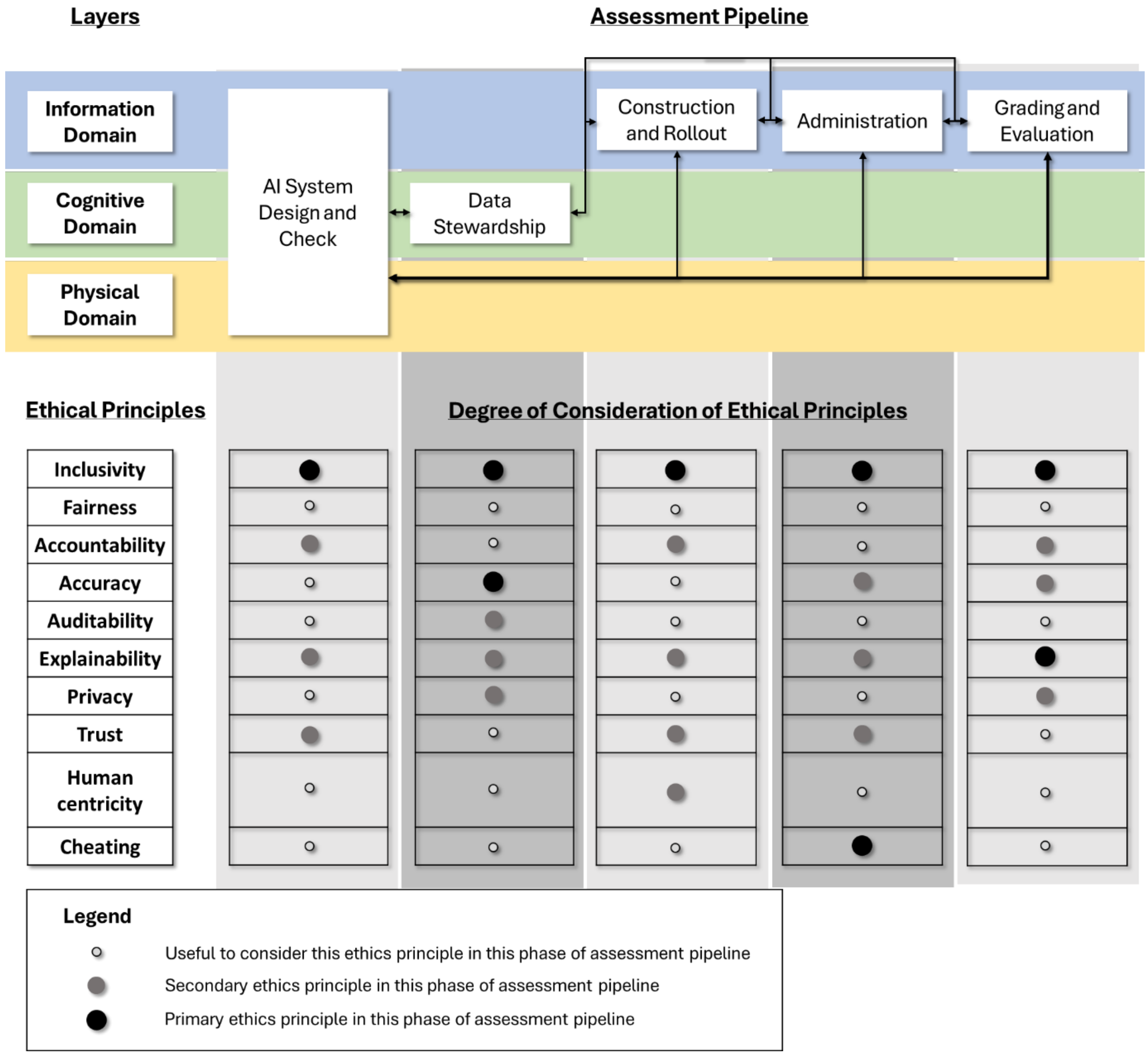


Figure 23: Revised Triadic AIED Assessment Framework based on learner perceptions

4.5 Implications, Limitations and Conclusion

4.5.1 Implications

By confronting ethical imperatives inherent in AI systems, such as fairness, inclusivity, and privacy, educational communities are prompted to engage in critical discussions about values, biases, and social implications. This association underscores AI's role not only as a technological tool but also as a catalyst for ethical consciousness and dialogue. The implications of this study are three-pronged, spanning pedagogical practices, technological development, and educational policy.

From a pedagogical standpoint, the research underscores the importance for educators and AI developers to deeply embed ethical considerations into AI design and application. This includes a heightened focus on inclusivity and explainability (ethical issues highlighted across *all* assessment stages), ensuring that AI-driven assessments are accessible and understandable to a diverse range of learners. The necessity to incorporate diverse perspectives and learning styles becomes paramount, aiming to mitigate biases and create equitable educational experiences. Furthermore, the emphasis on explainability extends to making AI processes transparent, enabling students to comprehend how their assessments are generated and evaluated. This clarity fosters a deeper understanding and confidence in educational assessments.

Technologically, the study highlights the need for AI systems in education to be developed with a core focus on ethical considerations. This entails the creation of algorithms that are not only inclusive but also transparent in their functioning. Generative AI systems enhanced by human feedback have demonstrated the capacity for inclusive and empathetic interactions (e.g., Yu et al., 2023). Furthermore, the dynamic nature of both AI technology and educational needs mandates continuous monitoring and refinement of these systems. A significant technological implication is the enhanced focus on AI explainability. This involves crafting user interfaces and feedback mechanisms that are comprehensible and accessible, catering to a wide range of users.

On the policy front, the findings of the study advocate for the development of comprehensive and robust frameworks governing the ethical application of AI in educational settings. There is also a pronounced need for professional development programs aimed at educators and administrators, focusing on the ethical dimensions of AI in education. These programs should equip them with the knowledge and skills to effectively navigate the potentials and limitations of AI in assessments. Lastly, the study points towards the necessity for regulatory oversight in the educational use of AI, ensuring adherence to ethical standards and addressing any legal and ethical concerns that may arise.

4.5.2 Limitations

This study, while providing significant insights into the ethical imperatives of AI in educational assessments, is not without its limitations. Firstly, the reliance on learner perceptions as the primary data source may introduce a degree of subjectivity. While these perceptions are invaluable for understanding user experience, they may not fully capture the entire spectrum of ethical complexities in AI-driven assessment systems, for instance, through educators, educational administrators, AI system designers and developers, and the public among others. Furthermore, learners' understanding of AI technologies and their underlying principles may vary, potentially influencing their responses and the subsequent analysis. There is also room to expand the study beyond the pilot sample size of 95 participants.

Another limitation lies in the scope of the study. The research focused predominantly on higher education environments, limiting the generalizability of the findings to other educational contexts, such as K-12. Additionally, the cultural and geographical diversity of the study's participants was restricted, which may impact the universality of the results. The ethical considerations in AI educational assessments can vary significantly across different cultures and educational systems, and this study may not fully account for these variations.

4.5.3 Chapter Conclusion

The exploration of ethical imperatives in AI-driven educational assessments reveals a multifaceted research area where technology, pedagogy, and policy intersect. AI, in this context, emerges not just as a technological tool but as a catalyst for human-centered ethical discourse and consciousness in educational assessments.

From a pedagogical standpoint, this research emphasizes the need for educators and AI developers to prioritize ethical considerations, advocating for, among others, inclusivity and explainability in AI-driven assessments to cater to diverse learner needs. Technologically, the study sheds light on the necessity of developing AI systems with an inherent focus on ethical principles, ensuring continuous adaptation to the evolving educational landscape. On the policy front, it calls for comprehensive frameworks and professional development programs to guide and inform the ethical use of AI in education.

However, the study acknowledges its limitations, including potential subjectivity in learner perceptions and the limited scope in terms of participant diversity and educational contexts. These limitations highlight the need for a broader, more inclusive approach in future research.

Aside from studies that can help overcome the inherent limitations of this study, suggestions for future research include the following:

- Conducting longitudinal studies could offer deeper insights into how perceptions and impacts of AI in educational assessments evolve over time, particularly as technology and educational practices continue to advance.

- Investigate the depth of learners' understanding of AI technologies and principles would provide valuable context to their perceptions and experiences. This could involve assessing their knowledge base and how it influences their views on ethical issues.

- Adopting interdisciplinary research methodologies that bring together expertise from education, technology, ethics, and policy studies could yield more holistic and nuanced insights into the ethical dimensions of AI in education.

While this study offers significant insights into the ethical imperatives of AI in educational assessments, it also opens the door to a plethora of research opportunities. These future directions not only promise to deepen our understanding but also guide the responsible and effective integration of AI in educational settings. The implications of this study call for a holistic and concerted approach in embedding AI into educational assessments. Addressing these implications is crucial for stakeholders in the education sector to fully harness AI's potential in a manner that is ethically responsible, pedagogically enriching, and technologically progressive.

Chapter 5. Conclusion and Future Works

5.1 Conclusion

This dissertation, "*Ethical Imperatives in AI-Driven Educational Assessment: Framework and Implications*," represents a comprehensive inquiry into the ethical dimensions at the intersection of AI and educational assessment.

This dissertation has three primary objectives: Firstly, it aims to analyse foundational educational technologies – Ubiquitous, Adaptive, and Immersive technologies – that are critical in integrating technology into educational assessments. It focuses on highlighting the rapid growth of AI-driven adaptive technology and identifying the crucial role of ethical imperatives in this technological integration. Secondly, the study endeavours to enrich and expand the current understanding of the complex interaction between AI technology, ethical imperatives, and educational assessments. Thirdly, it strives to validate and establish a comprehensive, theory-based framework to effectively tackle the identified ethical imperatives. This research not only addresses a significant gap in the current academic discourse but also provides practical, actionable guidance for educators, policymakers, and academic researchers.

The significance of this research lies in its contribution to filling a notable gap in existing literature, particularly concerning the ethical imperatives brought forth by AI in educational assessments. By proposing a robust ethical framework, this dissertation has provided practical guidance for educators, policymakers, and researchers, thereby facilitating a responsible and effective integration of AI in educational assessment practices.

5.2 Limitations

The overall limitations of the research presented across the earlier chapters can be distilled into several key themes, reflecting the constraints within the selection of databases, the consideration of wider cultural, educational and stakeholder contexts, and the methodological approach. The limitations are broadly synthesized as below:

- Future studies can aim for a more inclusive literature review by integrating a broader range of databases beyond Scopus and Google Scholar. This expanded approach provides a more exhaustive exploration of existing research, potentially uncovering additional insights and perspectives that were not captured in the initial study. Diversifying literature sources could mitigate the risk of overlooking critical studies and enhance the study's foundational knowledge base.
- Expanding the research to include a broader array of educational environments and a more diverse participant pool can be useful. This would involve conducting studies across different educational levels, such as K-12, and incorporating participants from varied cultural and geographical backgrounds. Such diversification would enhance the study's relevance and applicability to a wider range of educational contexts, offering insights that are more generalizable.
- Future research could incorporate views from a wider range of stakeholders, including educators, educational administrators, AI developers, and policymakers. This approach can provide a more holistic understanding of the ethical implications of AI in education, capturing a broader spectrum of concerns and expectations regarding AI-driven assessments.

- While the initial research employed unsupervised machine learning and text mining techniques for thematic analysis, there is room for methodological diversification. Future studies could explore additional qualitative and quantitative methods to assess the impact of thematic diversity and the ethical considerations within educational assessments. This might include case studies, comparative analyses, and participatory research methods to offer more nuanced insights into the integration of AI and technology in educational contexts.

By addressing these limitations, future research can build upon the initial study's foundations, offering richer, more comprehensive insights into the integration of AI in educational assessments. This approach not only broadens the scope of the research but also deepens the understanding of how AI can be leveraged to enhance educational outcomes ethically and effectively.

5.3 Future Works

This section builds upon the foundational insights from earlier chapters and proposes future research directions. These proposed studies aim to bridge the gap between theoretical frameworks and practical applications in the realm of AI-driven educational assessments, with an emphasis on ethical imperatives. The future works will converge interdisciplinary research methodologies, conduct longitudinal studies, assess learners' understanding of AI technologies, and implement actionable insights in practical settings.

5.3.1 Interdisciplinary Research Methodologies

Interdisciplinary research methodologies, which converge expertise from education, technology, ethics, and policy studies, are pivotal for comprehensively understanding the ethical imperatives of AI in education. This approach recognizes that the challenges and opportunities presented by AI in educational assessments are not just technological issues, but are deeply intertwined with pedagogical theories, ethical principles, and policy frameworks. By drawing on diverse disciplines, we can develop more nuanced and effective ethical frameworks that cater to the multifaceted nature of AI-driven educational assessments. This holistic understanding is crucial for ensuring that AI is used in a way that is beneficial, fair, and respectful of students' rights and diverse needs.

The proposed research involves the establishment of interdisciplinary research teams dedicated to exploring the ethical implications of AI in education. These teams would consist of experts from various fields, including: (i) Education Specialists: To provide insights into pedagogical approaches, learning theories, and the practical realities of classroom and assessment environments; (ii) Technologists and AI Researchers: To offer expertise in the latest AI developments, data science methodologies, and technical aspects of AI implementation in educational settings; (iii) Ethicists: To guide the discussion on moral and ethical considerations, ensuring that AI applications align with ethical standards and societal values; (iv) Policy Experts: To contribute insights on educational policies, regulatory frameworks, and compliance with national and international standards.

These interdisciplinary teams would work on two main areas. These include: (i) developing dynamic ethical governance models, and (ii) creating global ethical standards for diverse educational contexts. For the former, these models would be designed to evolve with the rapidly changing landscape of AI technology. They would include mechanisms for ongoing monitoring, evaluation, and adaptation of ethical guidelines as new AI technologies and applications emerge. The governance models would also incorporate feedback loops from educators, students, and other stakeholders, ensuring that they remain relevant and responsive to users' needs and concerns. Pertaining to the latter, recognizing that AI in education is a global phenomenon with local contexts, the research would aim to develop ethical standards that are universally applicable yet adaptable to various cultural and educational settings. This would involve comparative studies of different educational systems, collaboration with international educational bodies, and a deep understanding of cross-cultural ethical considerations. The goal is to establish a set of standards that promote the responsible and equitable use of AI in education worldwide.

By integrating these diverse perspectives and expertise, the research aims to provide comprehensive, actionable solutions that address the ethical imperatives of AI in education. This interdisciplinary approach is not only essential for advancing our understanding of AI's ethical dimensions but also crucial for the practical implementation of AI technologies in a way that enhances educational outcomes and upholds ethical principles.

5.3.2 Longitudinal Impact Studies

Understanding the long-term effects of ethical imperatives in AI-driven assessments is pivotal for comprehensively evaluating the impact of these technologies on learners. The complexities inherent in the ethical use of AI in educational contexts necessitate in-depth, longitudinal studies. These studies are essential for tracking and understanding how students' perceptions, attitudes, and responses to the ethics of AI in educational assessments evolve over time. The primary goal is to observe the enduring effects of ethical AI practices on the overall learning experience, encompassing psychological, social, and academic dimensions.

The proposed research involves conducting extensive, multi-year studies that systematically track and analyze the psychological, social, and academic impacts of the ethical use of AI in educational assessments. The key components of this research include: (i) Study design and population: Implementing longitudinal study designs with diverse student populations across various educational levels and disciplines. This would involve repeated observations of the same variables over extended periods, potentially spanning several years. (ii) Psychological impact assessment: Investigating how ethical AI-driven assessments affect students' mental health, stress levels, motivation, and attitudes towards learning. This aspect would consider variables such as students' confidence in the fairness and accuracy of AI assessments, anxiety levels related to technology use, and perceptions of privacy and data security. (iii) Social impact analysis: Examining the social implications of ethical AI in education, including student interactions, collaborative learning environments, and the broader educational community's response. This component would look at changes in student-teacher dynamics, peer interactions, and the overall classroom culture in the presence of AI-driven tools. (iv) Academic performance tracking: Assessing the impact of ethical AI-driven assessments on academic outcomes. This would involve analyzing changes in student performance, engagement levels, and learning outcomes over time, providing insights into the effectiveness and implications of AI-driven educational practices. (v) Development of ethical AI tools: Part of the research would focus on creating and implementing AI tools that emphasize ethical principles such as inclusivity and accessibility. These tools would be designed to cater to diverse learning needs, including those of students with disabilities, and would be integrated into the longitudinal studies to assess their impact. (vi) Data collection and analysis methods: Utilizing a variety of data collection methods, including surveys, interviews, academic records, and observational studies. Advanced analytical techniques, such as longitudinal data analysis, mixed methods approaches, and machine learning algorithms, would be employed to interpret the collected data. (vii) Ethical and regulatory compliance: Ensuring all research activities comply with ethical standards and regulatory requirements, particularly concerning data privacy, consent, and the welfare of participants.

This research is expected to contribute significantly to the field of educational technology by providing in-depth insights into the long-term effects of ethical AI in educational assessments. The findings could inform the development of more effective, ethical, and student-centered AI-driven assessment tools. Additionally, this research could guide policymakers and educational institutions in implementing AI technologies responsibly and effectively. Moreover, the longitudinal nature of the study would provide a comprehensive view of the evolving landscape of AI in education, capturing changes and trends that shorter-term studies might miss. This would be instrumental in shaping future educational policies and practices, ensuring they are aligned with the best interests of students and the broader educational community.

5.3.3 Assessing Learners' AI Literacy

In the context of establishing a framework for ethical imperatives in AI-driven educational assessments, with theoretical constructs built upon learner perceptions, learners' understanding and knowledge of AI ethics and AI technologies (such as generative AI) play a pivotal role in shaping their perceptions and experiences. The concept of AI literacy extends beyond mere awareness of AI's functionality; it encompasses an understanding of its underlying principles, ethical implications, and its broader impact on learning and assessment processes. This literacy is crucial, as it influences how they interact with, respond to, and critically evaluate AI-driven tools and methodologies in educational settings. Beyond learners, for the purposes of framework construction, the understanding of this literacy can also extend to educators, system designers, education administrators, among others.

The proposed research involves a multi-faceted approach to assess and enhance AI literacy. This research will be divided into several key phases: (i) Initial assessment and baseline study: The first phase involves conducting comprehensive surveys and interviews to gauge the current level of AI literacy among the involved stakeholders, such as students and educators. This will include assessments of their understanding of AI concepts, their awareness of ethical issues related to AI, and their ability to critically engage with AI-driven technologies. (ii) Development of educational interventions: Based on the findings from the initial assessment, the second phase focuses on designing targeted educational interventions. These interventions could include specialized curricula, workshops, and interactive learning modules aimed at enhancing AI literacy. The content will cover key areas such as AI ethics, data privacy, algorithmic bias, and the role of AI in personalized learning. (iii) Pilot implementation and feedback loop: The educational interventions will be piloted in select educational institutions. Feedback from participants will be collected systematically to refine and adapt the interventions. This iterative process ensures that the educational content remains relevant, engaging, and effective in enhancing AI literacy. (v) Longitudinal study and impact analysis: Following the implementation of these interventions, a longitudinal study will be conducted to evaluate their long-term impact. This study will assess changes in AI literacy levels over time and examine how increased literacy affects perceptions and interactions with AI-driven educational assessments. The study will also explore the broader impacts on educational outcomes and ethical considerations. (vi) Dissemination and scaling: The final phase involves disseminating the successful interventions on a larger scale. Collaborations with educational bodies, policy makers, and AI technology providers will be sought to integrate these interventions into broader educational frameworks and policies. The goal is to ensure that AI literacy becomes an integral component of education systems, preparing learners and educators to navigate the evolving landscape of AI-driven education effectively and ethically.

Constructing a framework for ethics in AI-driven educational assessments, with theoretical constructs built upon stakeholders' perceptions, requires the assessing of AI literacy over time. Assessing and enhancing AI literacy helps foster a critical understanding and ethical consciousness among educational stakeholders, further ensuring that such frameworks are robustly constructed.

5.3.4 Dynamic Network Model

The current research on AI ethics primarily outlines the importance of individual ethical principles statically across different stages of the AI-driven assessment pipeline. Future work can focus on developing a dynamic model that visualizes and analyzes the interrelationships among ethical principles, such as inclusivity, fairness, accountability, and others. This model aims to provide a more holistic and operational framework that can guide the development, deployment, and continuous evaluation of AI systems in a way that reflects the interconnected nature of these principles.

To achieve this, future research may utilize a network graph where nodes represent ethical principles and directed edges symbolize influential relationships and dependencies. This graph would be developed based on comprehensive literature reviews, survey instruments, expert interviews, and case studies. Empirical data would be collected to understand: (i) how nodes that represent each principle are mapped; (ii) how edges that denote significant interactions (e.g., fairness impacts trust) are drawn; and (iii) how edges specify if certain interactions are stronger or more critical in specific contexts. It would be useful to identify which principles are most central or have the most connections in different stages of the assessment pipeline, using tools like centrality measures in network analysis (e.g., degree, betweenness, closeness). Qualitative data from stakeholder interviews, literature reviews, or case studies would be helpful to provide context to the connections and assess the strength of each relationship. We can also apply simulations to observe how manipulation of various nodes (principles) within the network causes resultant changes in the system's ethical behavior. This approach will help in understanding the elasticity and resilience of the ethical framework under different operational stresses. To allow for model changes across time, algorithms can be created that can dynamically adjust the weight of certain principles based on real-time data and feedback, ensuring that the AI system remains aligned with current ethical standards throughout its lifecycle.

To validate this framework, future research may consider: (i) Testing the model against multiple real-world scenarios to ensure it reliably reflects the real-world ethical considerations in the contextual environment; (ii) Continuously refining the model through feedback from a panel of interdisciplinary experts, including ethicists, technologists, and legal advisors; and (iii) Engaging with the broader AI and ethics community through workshops to gather a wide range of feedback and foster discussions on improving the model.

This proposed future work looks to map out the interconnected relationships between ethical principles, and understand how these relationships evolve over time and under different conditions. The development of a dynamic network model provides a more realistic and flexible approach to ensuring that AI systems considers the real-world interconnectedness of ethical principles, while adhering to ethical standards throughout their operational lifecycle.

5.3.5 Implementing Actionable Insights

The transition from theoretical ethics frameworks to tangible, real-world applications is crucial for the meaningful impact of AI-driven educational assessments. This transition involves converting theoretical insights, drawn from research on ethical imperatives in AI education, into practical strategies and tools. Such a translation is necessary to bridge the gap between conceptual understanding and actionable practice, ensuring that ethical principles are not just theoretical ideals but active components in the design, implementation, and evaluation of AI technologies in educational settings.

Firstly, this can involve the development of practice-oriented strategic and operational frameworks. The primary goal of this research avenue is to create practice-oriented frameworks grounded in the ethical principles outlined in previous chapters. These frameworks would serve as blueprints for educators and policymakers, guiding the ethical integration and practical application of AI in educational settings. The key objectives would be to: (i) To distill ethical principles into actionable strategies and protocols; (ii) To align these strategies with existing educational standards and technological capabilities; (iii) To ensure these frameworks are adaptable to different educational contexts and learner needs; and (vi) To develop case studies that exemplify the practical application of the theoretical frameworks. These case studies would be drawn from diverse educational settings, showcasing the implementation of ethical AI practices in real-world scenarios.

Secondly, to facilitate the adoption of these frameworks and guidelines, this research also proposes the organization of workshops and training programs for educators, administrators, and policymakers. These programs would focus on enhancing understanding of AI technologies, ethical considerations, and the practical application of the developed frameworks. The key objectives would be: (i) To provide hands-on training and resources for effective implementation of ethical AI practices; (ii) To foster a community of practice for continuous learning and adaptation of AI in education; (iii) To encourage feedback and collaborative refinement of the frameworks and guidelines.

Thirdly, an essential aspect of this research is to inform and influence educational policy at various levels, advocating for the integration of ethical considerations in AI-driven educational assessments. The key objectives would be: (i) To engage with policymakers and educational leaders to integrate ethical AI principles into educational policies; (ii) To provide policy briefs and recommendations based on research findings; and (iii) To advocate for the establishment of standards and regulations governing the use of AI in educational assessments.

The implementation of actionable insights from theoretical research into AI-driven educational assessments aims to transform ethical frameworks from academic discourse into practical tools and guidelines, ensuring that AI in education advances in a manner that is ethically sound, educationally beneficial, and policy-compliant. The successful execution of these strategies would mark a significant step forward in aligning technological innovation with ethical imperatives, ultimately enhancing the quality and integrity of education in the AI era.

5.4 Closing Remarks

The dissertation unveiled the dynamic and often unpredictable trajectories of educational technologies, emphasizing the growing prominence of adaptive technologies and the revolutionary role of immersive and ubiquitous technologies in educational assessments. The research underscored the ethical imperatives posed by AI and the necessity of addressing them in all stages of the assessment pipeline. The study also proposed and validated a generalizable framework that can be utilized in when developing and rolling out ethically underpinned AI-enabled assessments.

As we venture into an era marked by rapid technological advancements in education, this dissertation underscores the criticality of navigating the ethical imperatives with diligence and foresight. The integration of AI into educational assessments offers tremendous potential but also poses significant ethical challenges that must be addressed holistically. This research serves as a foundational reference point for stakeholders in education to engage with AI responsibly, ensuring that its integration enhances the educational experience in a manner that is also pedagogically effective and ethically sound.

References

Abdi, S., Khosravi, H., Sadiq, S., & Gasevic, D. (2020, March). Complementing educational recommender systems with open learner models. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 360-365). ACM.

Abu-Rasheed, H., Weber, C., & Fathi, M. (2023). Context Based Learning: A Survey of Contextual Indicators for Personalized and Adaptive Learning Recommendations – A Pedagogical and Technical Perspective. *Frontiers in Education*, *8*, p. 1210968. Frontiers. DOI: https://doi.org/10.3389/feduc.2023.1210968.

Adams, D., Chuah, K. M., Devadason, E., & Azzis, M. S. A. (2023). From novice to navigator: Students' academic help-seeking behaviour, readiness, and perceived usefulness of ChatGPT in learning. *Education and Information Technologies*, 1-18.

Ahn, J., Nguyen, H., Campos, F., & Young, W. (2021, April). Transforming everyday information into practical analytics with crowdsourced assessment tasks. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 66-76). ACM.

Alexander, B., Ashford-Rowe, K., Barajas-Murph, N., Dobbin, G., Knott, J., McCormack, M., Pomerantz, J., Seilhamer, R. & Weber, N. (2019). *Educause Horizon Report: 2019 Higher Education Edition.* Louiseville, CO: Educause.

Amigud, A., Arnedo-Moreno, J., Daradoumis, T., & Guerrero-Roldan, A. E. (2018). Open proctor: An academic integrity tool for the open learning environment. In *Advances in Intelligent Networking and Collaborative Systems: The 9th International Conference on*

Intelligent Networking and Collaborative Systems (INCoS-2017) (pp. 262-273). Springer International Publishing.

An, T., & Oliver, M. (2021). What in the world is educational technology? Rethinking the field from the perspective of the philosophy of technology. *Learning, Media and Technology*, *46*(1), 6-19.

Anagnostopoulos, T., Kytagias, C., Xanthopoulos, T., Georgakopoulos, I., Salmon, I., & Psaromiligkos, Y. (2020). Intelligent predictive analytics for identifying students at risk of failure in Moodle courses. In *Intelligent Tutoring Systems: 16th International Conference, ITS 2020, Athens, Greece, June 8–12, 2020, Proceedings 16* (pp. 152-162). Springer International Publishing.

Andersson, H., Svensson, A., Frank, C., Rantala, A., Holmberg, M., & Bremer, A. (2022). Ethics education to support ethical competence learning in healthcare: An integrative systematic review. *BMC Medical Ethics, 23,* 29. DOI: https://doi.org/10.1186/s12910-022-00766-z.

Andrews, D., Leitner, P., Schön, S., & Ebner, M. (2022). Developing a Professional Profile of a Digital Ethics Officer in an Educational Technology Unit in Higher Education. In *International Conference on Human-Computer Interaction* (pp. 157-175). Springer, Cham.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, *62*, 102433.

Asiimwe, E. N., & Khan, S. Z. (2013). Ubiquitous computing in education: A SWOT analysis by students and teachers. In *12th World Conference on Mobile and Contextual Learning (mLearn 2013),* 3, 18. Hamad bin Khalifa University Press (HBKU Press).

Australian Government. (2019). Australia's Artificial Intelligence Ethics Framework. Department of Industry, Science and Resources. [Online]. Canberra, Australia. Retrieved: https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework [Accessed 20 Nov 2022].

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., ... & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59-64.

Aynaud, T. (2020). Python-louvain. Louvain algorithm for community detection. [Online]. Retrieved: https://github.com/taynaud/python-louvain. [Assessed 15 Nov 2022].

Balderas, A., Palomo-Duarte, M., Dodero, J. M., Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2018). Scalable authentic assessment of collaborative work assignments in wikis. *International Journal of Educational Technology in Higher Education*, *15*(1), 1-21.

Barra, E., López-Pernas, S., Alonso, Á., Sánchez-Rada, J. F., Gordillo, A., & Quemada, J. (2020). Automated assessment in programming courses: A case study during the COVID-19 era. *Sustainability*, *12*(18), 7451.

Barrett, H. C. (2007). Researching electronic portfolios and learner engagement: The REFLECT initiative. *Journal of Adolescent & Adult Literacy*, *50*(6), 436-449.

Bastani, H., Bastani, O., & Kim, C. (2018). Interpreting predictive models for human-in-the-loop analytics. *arXiv preprint arXiv:1705.08504*.

Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: developing the assessment design decisions framework. *Teaching in Higher Education, 21*(5), 545-556. DOI: https://doi.org/10.1080/13562517.2016.1160217.

Bearman, M., Luckin, R. (2020). Preparing University Assessment for a World with AI: Tasks for Human Intelligence. In: Bearman, M., Dawson, P., Ajjawi, R., Tai, J., Boud, D. (eds) *Re-imagining University Assessment in a Digital World. The Enabling Power of Assessment,* vol 7. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-41956-1_5.

Becker, H. J. (1991). How computers are used in United States schools: Basic data from the 1989 IEA computers in education survey. *Journal of Educational Computing Research*, *7*(4), 385-406.

Becker, S. A., Brown, M., Dahlstrom, E., Davis, A., DePaul, K., Diaz, V., & Pomerantz, J. (2018). *NMC Horizon Report: 2018 Higher Education Edition.* Louisville, CO: Educause.

Becker, S. A., Cummins, M., Davis, A., Freeman, A., Hall, C. G., & Ananthanarayanan, V. (2017). *NMC Horizon Report: 2017 Higher Education Edition.* Louisville, CO: Educause.

Bennett, R. E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *Journal of Technology, Learning, and Assessment, 1*(1).

Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: Learner choice and fundamental rights. *Learning, Media and Technology*, *45*(3), 312-324.

Bessen, J., Impink, S. M., & Seamans, R. (2022, July). The cost of ethical AI development for AI startups. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 92-106). ACM.

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21-34.

Bingley. DOI: https://doi.org/10.1108/978-1-83909-850-520201012/full/html.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, *5*(1), 7-74.

Blanchard, E. G. (2012). On the weird nature of ITS/AIED conferences. In *International Conference on Intelligent Tutoring Systems* (Vol. *7315*, pp. 280–285). Springer.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008. DOI: 10.1088/1742-5468/2008/10/P10008.

Bohemia, E., & Davison, G. (2012). Authentic learning: the gift project. *Design and Technology Education: An International Journal, 17*(2), 49-61. DOI: https://ojs.lboro.ac.uk/DATE/article/view/1731.

Bolick, A. D., & da Silva, R. L. (2024). Exploring Artificial Intelligence Tools and Their Potential Impact to Instructional Design Workflows and Organizational Systems. *TechTrends*, *68*(1), 91-100.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Borenstein, J., & Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. *AI and Ethics*, *1*(1), 61-65.

Borgatti, S. P., & Halgin, D. S. (2011). On network theory. *Organization Science, 22*(5), 1168–1181.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. Annual Review of Information Science and Technology, 37(1), 179-255.Brandes, U., & Erlebach, T. (Eds.). (2005). *Network analysis: Methodological foundations*. Springer-Verlag.

Bozkurt, A. (2020). Educational technology research patterns in the realm of the digital knowledge age. *Journal of Interactive Media in Education*, *2020*(1).

Braunstein, A., Deutscher, V., Seifried, J., Winther, E., & Rausch, A. (2022). A taxonomy of social embedding-A systematic review of virtual learning simulations in vocational and professional learning. *Studies in Educational Evaluation*, *72*, 101098. DOI: https://doi.org/10.1016/j.stueduc.2021.101098.

Breucker, P., Cointet, J., Hannud Abdo, A., Orsal, G., de Quatrebarbes, C., Duong, T., Martinez, C., Ospina Delgado, J. P., Medina Zuluaga, L. D., Gómez Peña, D. F., Sánchez Castaño, T. A., Marques da Costa, J., Laglil, H., Villard, L., & Barbier, M. (2016). CorTexT Manager (version v2). [Online]. Retrieved: https://docs.cortext.net [Accessed 24 Jan 2023].

Brookhart, S. M. (2014). *How to design questions and tasks to assess student thinking*. ASCD.

Brown, M., McCormack, M., Reeves, J., Brooks, D. C., & Grajek, S. (2020). *2020 Educause Horizon Report: Teaching and Learning Edition.* Louiseville, CO: Educause.

Bunderson, C. V., Inouye, D. K., & Olsen, J. B. (1988). The four generations of computerized educational measurement. *ETS Research Report Series*, *1988*(1), i-148.

Burgstahler, S. E., & Cory, R. C. (Eds.). (2010). *Universal design in higher education: From principles to practice*. Harvard Education Press.

Campedelli, G. M. (2021) Where are we? Using Scopus to map the literature at the intersection between artificial intelligence and research on crime. *Journal of Computational Social Science, 4,* 503–530.

Candra, O., Islami, S., Syamsuarnis, S., Elfizon, E., Hastuti, H., Habibullah, H., & Eliza, F. (2019). Validity of development on authentic assessment tool of curriculum 2013 based in information technology. *International Journal of Scientific and Technology Research*, *8*(12), 265-267.

Casas-Roma, J., & Conesa, J. (2021). A literature review on artificial intelligence and ethics in online learning. *Intelligent Systems and Learning Data Analytics in Online Education*, 111-131. DOI: https://doi.org/10.1016/B978-0-12-823410-5.00006-1.

Cechella, F., Abbad, G., & Wagner, R. (2021). Leveraging learning with gamification: An experimental case study with bank managers. *Computers in Human Behavior Reports*, *3*, 100044. DOI: https://doi.org/10.1016/j.chbr.2020.100044.

Charyton, C. (2013). *Creative engineering design assessment: background, directions, manual, scoring guide and uses*. Springer Science & Business Media.

Chassang, G., Thomsen, M., Rumeau, P., Sèdes, F., & Delfin, A. (2021). An interdisciplinary conceptual study of Artificial Intelligence (AI) for helping benefit-risk assessment practices. *AI Communications, 3*(4), 1-26.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science, 136*, 16-24.

Chaudhry, M. A., & Kazim, E. (2022). Artificial Intelligence in Education (AIEd): a high-level academic and industry note 2021. *AI and Ethics*, *2*(1), 157-165.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8,* 75264-75278.

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, 100002.

Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, *32*(3), 725-755.

Coates, H., James, R., & Baldwin, G. (2005). A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary education and management*, *11*(1), 19-36.

Collins, A., & Halverson, R. (2018). *Rethinking education in the age of technology: The digital revolution and schooling in America*. Teachers College Press.

Conati, C., Barral, O., Putnam, V., & Rieger, L. (2021). Toward personalized XAI: A case study in intelligent tutoring systems. *Artificial Intelligence*, *298*, 103503.

Corrêa, N. K., Galvão, C., Santos, J. W., Del Pino, C., Pinto, E. P., Barbosa, C., ... & de Oliveira, N. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, *4*(10).

Costas-Jauregui, V., Oyelere, S. S., Caussin-Torrez, B., Barros-Gavilanes, G., Agbo, F. J., Toivonen, T., Motz, R., & Tenesaca, J. B. (2021, October). Descriptive analytics dashboard for an inclusive learning environment. In *2021 IEEE Frontiers in Education Conference (FIE)* (pp. 1-9). IEEE.

Crompton, H. (2013). A historical overview of mobile learning: Toward learner-centered education. In Z. L. Berge & L. Y. Muilenburg (Eds.), *Handbook of mobile learning* (pp. 3-14). Routledge.

Dai, C. P., & Ke, F. (2022). Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review. *Computers and Education: Artificial Intelligence*, 100087.

Daim, T. U., Rueda, G., Martin, H., & Gerdsri, P. (2006). Forecasting emerging technologies: Use of bibliometrics and patent analysis. Technological forecasting and social change, 73(8), 981-1012.

Dawley, L. & Dede, C. (2014). Situated learning in virtual worlds and immersive simulations. In *Handbook of research on educational communications and technology,* 723-734. Springer, New York.

Dede, C. Etemadi, A., & Forshaw, T. (2021). *Intelligence augmentation: Upskilling humans to complement AI* [White Paper]. The Next Level Lab, Harvard Graduate School of Education.
https://pz.harvard.edu/sites/default/files/Intelligence%20Augmentation-%20Upskilling%20Humans%20to%20Complement%20AI.pdf.

Deho, O. B., Zhan, C., Li, J., Liu, J., Liu, L., & Duy Le, T. (2022). How do the existing fairness metrics and unfairness mitigation algorithms contribute to ethical learning analytics? *British Journal of Educational Technology, 53,* 822–843.

Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies*, 1-53.

dos Santos, S. C. (2016). PBL-SEE: An authentic assessment model for PBL-based software engineering education. *IEEE Transactions on Education*, *60*(2), 120-126.

Dubé, A. K., & Wen, R. (2022). Identification and evaluation of technology trends in K-12 education from 2011 to 2021. *Education and information technologies*, *27*(2), 1929-1958.

Duignan, P. A. (2020). Navigating the future of learning: the role of smart technologies. *Leading Educational Systems and Schools in Times of Disruption and Exponential Change: A Call for Courage, Commitment and Collaboration,* pp. 125–137. Emerald Publishing Limited,

Duin, A. H., & Tham, J. (2020). The current state of analytics: Implications for learning management system (LMS) use in writing pedagogy. *Computers and Composition*, *55*, 102544.

Dzaldov, B. S. (2022). Designing for Meaningful Synchronous and Asynchronous Discussion in Online Courses. In: MacKinnon, K., Wilton, L., Murphy, S., Dzaldov, B. S., Wattar, D., DesRochers, J., & Mann, A. (eds). *Authentic Assessment in Online Learning,* chapter 4. Pressbooks. Retrieved: https://ecampusontario.pressbooks.pub/designingforonlinediscussion/chapter/chapter-4-integrating-alternative-forms-of-communication-in-online-discussion/ [Accessed 15 Jan 2023].

Eicher, B., Polepeddi, L., & Goel, A. (2018). Jill Watson doesn't care if you're pregnant: Grounding AI ethics in empirical studies. *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 88-94). ACM.

Elsevier (2022). How Scopus works, Scopus contents. [Online]. Retrieved: https://www.elsevier.com/solutions/scopus/how-scopus-works/content [Assessed 15 November 2022].

Elshafey, A. E., Anany, M. R., Mohamed, A. S., Sakr, N., & Aly, S. G. (2021). Dr. Proctor: A multi-modal AI-based platform for remote proctoring in education. In *International Conference on Artificial Intelligence in Education* (pp. 145-150). Springer, Cham.

Engelbrecht, H., Lindeman, R. W., & Hoermann, S. (2019). A SWOT analysis of the field of virtual reality for firefighter training. *Frontiers in Robotics and AI, 6*(101). DOI: https://doi.org/10.3389/frobt.2019.00101.

Ersozlu, Z., Ledger, S., & Hobbs, L. (2021). Virtual simulation in ITE: technology driven authentic assessment and moderation of practice. In *Authentic Assessment and Evaluation Approaches and Practices in a Digital Era* (pp. 53-68). Brill.

Ertmer, P. A., and Ottenbreit-Leftwich, A. (2013). Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning. *Computers and Education, 64*, 175-182.

European Parliament (2021). Report on artificial intelligence in education, culture and the audiovisual sector (2020/2017(INI)). Committee on Culture and Education. [Online].

Retrieved: https://www.europarl.europa.eu/doceo/document/A-9-2021-0127_EN.html [Accessed 20 Nov 2022].

Fahimnia, B., Sarkis, J., & Davarzani, H. (2015). Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics*, *162*, 101–114. DOI: https://doi.org/10.1016/j.ijpe.2015.01.003.

Fahimnia, B., Sarkis, J., & Davarzani, H. (2015). Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics*, *162*, 101–114. DOI: https://doi.org/10.1016/j.ijpe.2015.01.003.

Fariani, R. I., Junus, K., & Santoso, H. B. (2023). A Systematic Literature Review on Personalised Learning in the Higher Education Context. *Technology, Knowledge and Learning*, *28*(2), 449-476.

Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(6), e1332. https://doi.org/10.1002/widm.1332.

Fletcher, A., Kickbusch, S., & Huijser, H. (2022). Authentic learning using mobile applications and contemporary geospatial information requirements related to Environmental Science. *Journal of Geography in Higher Education*, *46*(2), 185-203.

Floridi, L. (2018). Soft ethics and the governance of the digital. *Philosophy & Technology*, *31*(1), 1-8.

Floridi, L. (2021). Translating principles into practices of digital ethics: Five risks of being unethical. In *Ethics, Governance, and Policies in Artificial Intelligence* (pp. 81-90). Springer, Cham.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1).

Fox-Turnbull, W. (2015). *Conversations to support learning in technology education.* In *The future of technology education* (pp. 99-120). Springer, Singapore.

Freire, P. (1970). *Pedagogy of the oppressed*. New York: Continuum.

Garvey, G. (2022). Perspective Chapter: Ungrading, Grading Contracts, Gamification and Game-Based Learning. In *Active Learning-Research and Practice for STEAM and Social Sciences Education* (pp. 167-197). IntechOpen, London. DOI: http://dx.doi.org/10.5772/intechopen.105967.

Gaspar, H., Morgado, L., Mamede, H., Oliveira, T., Manjón, B., & Gütl, C. (2020). Research priorities in immersive learning technology: the perspectives of the iLRN community. *Virtual Reality, 24*(2), 319-341.

Gedrimiene, E., Silvola, A., Pursiainen, J., Rusanen, J., & Muukkonen, H. (2020). Learning analytics in education: Literature review and case examples from vocational education. *Scandinavian Journal of Educational Research*, *64*(7), 1105-1119.

Gerard, F. J., Mentzelopoulos, M., Economou, D., Khalish, Z., Ingram, J., & Ashley, E. (2022). Work-In-Progress – CrimOPS – Gamified Virtual Simulations for Authentic Assessment in Criminology. In *2022 8th International Conference of the Immersive Learning Research Network (iLRN)* (pp. 1-3). IEEE. DOI: https://doi.org/10.23919/iLRN55037.2022.9815942.

Gill, K. S. (2021). Ethical encounters. *AI & Society*, *36*, 1-7.

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied Sciences*, *11*(12), 5467.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199–220.

Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational technology research and development, 52*(3), 67-86. DOI: https://doi.org/10.1007/BF02504676.

Gupta, S., & Chen, Y. (2022). Supporting inclusive learning using chatbots? A chatbot-led interview study. *Journal of Information Systems Education*, *33*(1), 98-108.

Hakami, E., & Hernández Leo, D. (2020). How are learning analytics considering the societal values of fairness, accountability, transparency and human well-being: A literature review. In Martínez-Monés A, Álvarez A, Caeiro-Rodríguez M, Dimitriadis Y (Eds.)*, Learning Analytics Summer Institute Spain 2020* (pp. 121-41). Aachen: CEUR. Valladolid, Spain.

Halabi, A. K. (2021). Pivoting authentic assessment to an accounting podcast during COVID-19. *Accounting Research Journal*, *34*(2), 156-168.

Hanna, R., & Kazim, E. (2021). Philosophical foundations for digital ethics and AI ethics: a dignitarian approach. *AI and Ethics*, *1*(4), 405-423.

Hargreaves, S. (2023). 'Words Are Flowing Out Like Endless Rain Into a Paper Cup': ChatGPT & Law School Assessments. *The Chinese University of Hong Kong Faculty of Law Research Paper*, (2023-03).

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112.

Hendrycks, D., Carlini, N., Schulman, J., & Steinhardt, J. (2021). Unsolved problems in ML safety. *arXiv preprint arXiv:2109.13916*.

Henne, T., & Gstrein, O. J. (2023). Governing the 'datafied' school: Bridging the divergence between universal education and student autonomy. In Zwitter, A., Gstrein, O.J. (Eds.), *Handbook on the Politics and Governance of Big Data and Artificial Intelligence* (pp. 395 – 427). Elgar.

Heo, J., & Lee, J. (2019, July). CiSA: An inclusive chatbot service for international students and academics. In *International Conference on Human-Computer Interaction* (pp. 153-167). Springer, Cham.

Herrington, J., Reeves, T. C., & Oliver, R. (2007). Immersive learning technologies: Realism and online authentic learning. *Journal of Computing in Higher Education, 19*(1), 80-99.

Herwix, A., Haj-Bolouri, A., Rossi, M., Tremblay, M. C., Purao, S., & Gregor, S. (2022). Ethics in information systems and design science research: Five perspectives. *Communications of the Association for Information Systems*, *50*(1), 589-616.

High-Level Expert Group on Artificial Intelligence (AI HLEG). (2019). *Ethics guidelines for Trustworthy AI.* European Commission. [Online]. Retrieved https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html [Accessed 22 Nov 2022].

Hill, P., and Barber, M. (2014). *Preparing for a renaissance in assessment.* Pearson, London.

Hira, A., and Hynes, M. M. (2018). People, means, and activities: a conceptual framework for realizing the educational potential of makerspaces. *Educational Research International, 2018*(6923617). DOI: https://doi.org/10.1155/2018/6923617.

Hoekstra, A., & Kaptein, M. (2021). The integrity of integrity programs: Toward a normative framework. *Public Integrity*, *23*(2), 129-141.

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*, 1-23.

Hong, Y., Nguyen, A., Dang, B., & Nguyen, B. P. T. (2022, July). Data Ethics Framework for Artificial Intelligence in Education (AIED). In *2022 International Conference on Advanced Learning Technologies (ICALT)* (pp. 297-301). IEEE.

Hood, W. W., & Wilson, C. S. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, *52*, 291-314.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal 6*(1):1-55.

Hu, Q., & Rangwala, H. (2020). Towards fair educational data mining: A case study on detecting at-risk students. In *Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)* (pp. 431–437). ERIC.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), 112-131.

Hughes, S., Davis, T. E., & Imenda, S. N. (2019). Demystifying theoretical and conceptual frameworks: A guide for students and advisors of educational research. *J Soc Sci*, *58*(1-3), 24-35.

Hussar, W.J., and Bailey, T.M. (2020). *Projections of Education Statistics to 2028* (NCES 2020-024). U.S. Department of Education, Washington, DC: National Center for Education Statistics.

Hwang, G. J., & Chien, S. Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. *Computers and Education: Artificial Intelligence*, 100082.

Ibrahim, M.M., Malik, M., & Avianti, R.A. (2022). Lecturers' perceptions of authentic assessment in times of COVID-19 pandemic: a case of Indonesian universities. *Journal of Applied Research in Higher Education*. DOI: https://doi.org/10.1108/JARHE-02-2022-0041.

IEEE. (2019). Ethically aligned design. [Online]. Retrieved: https://ethicsinaction.ieee.org/ [Accessed 20 Nov 2022].

Jang, Y., Choi, S., & Kim, H. (2022). Development and validation of an instrument to measure undergraduate students' attitudes toward the ethics of artificial intelligence (AT-EAI) and analysis of its difference by gender and experience of AI education. *Education and Information Technologies*, *27*(8), 11635-11667.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., and Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition.* Austin, Texas: The New Media.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., and Ludgate, H. (2013). *NMC Horizon Report: 2013 Higher Education Edition.* Austin, Texas: The New Media Consortium.

Johnson, L., Adams Becker, S., Estrada, V., and Freeman, A. (2015). *NMC Horizon Report: 2015 Higher Education Edition*. Austin, Texas: The New Media Consortium.

Johnson, L., Adams Becker, S., Estrada, V., Freeman, A. (2014). *NMC Horizon Report: 2014 Higher Education Edition.* Austin, Texas: The New Media Consortium.

Johnson, L., Adams, S., and Cummins, M. (2012). *The NMC Horizon Report: 2012 Higher Education Edition.* Austin, Texas: The New Media Consortium.

Johnson, L., Smith, R., Willis, H., Levine, A., and Haywood, K., (2011). *The 2011 Horizon Report.* Austin, Texas: The New Media Consortium.

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, *2*, 100017.

Kashive, N., Powale, L., & Kashive, K. (2020). Understanding user perception toward artificial intelligence (AI) enabled e-learning. *The International Journal of Information and Learning Technology*, *38*(1), 1-19.

Keast, D. A. (2018). Teaching Reflections on Two Decades of Online Music Courses. In *Pedagogy Development for Teaching Online Music* (pp. 227-243). IGI Global. DOI: https://doi.org/10.4018/978-1-5225-5109-6.ch011.

Khairy, D., Alkhalaf, S., Areed, M. F., Amasha, M. A., & Abougalala, R.A. (2022). An algorithm for providing adaptive behavior to humanoid robot in oral assessment. *International Journal of Advanced Computer Science and Applications, 13*(9). DOI: http://dx.doi.org/10.14569/IJACSA.2022.01309119.

Khan, A. A., Badshah, S., Liang, P., Waseem, M., Khan, B., Ahmad, A., ... & Akbar, M. A. (2022, June). Ethics of AI: A systematic literature review of principles and challenges. In *Proceedings of the International Conference on Evaluation and Assessment in Software Engineering 2022* (pp. 383-392).

Khosravi, H., Demartini, G., Sadiq, S., & Gasevic, D. (2021, April). Charting the design and analytics agenda of learnersourcing systems. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 32-42). ACM.

Khosravi, H., Kitto, K., & Williams, J. J. (2019). RiPPLE: A Crowdsourced Adaptive Platform for Recommendation of Learning Activities. *Journal of Learning Analytics, 6*(3), 91-105.

Khosravi, H., Sadiq, S., & Gasevic, D. (2020, February). Development and adoption of an adaptive learning system: Reflections and lessons learned. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 58-64). ACM.

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., ... & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *3*, 100074.

Kiennert, C., De Vos, N., Knockaert, M., & Garcia-Alfaro, J. (2019). The influence of conception paradigms on data protection in e-learning platforms: A case study. *IEEE Access*, *7*, 64110-64119.

Killam, L. A., Luctkar-Flude, M., Brune, S., & Camargo-Plazas, P. (2022). Redefining Cheating on Written Exams: A Shift Toward Authentic Assessment to Promote Universal Design for Learning in the Context of Critical Caring Pedagogy. *Advances in Nursing Science, 45*(3), E127-E143.

Kim, J. H., Baek, J., Hwang, C., Bae, C., & Park, J. (2021, June). Condensed discriminative question set for reliable exam score prediction. In *International Conference on Artificial Intelligence in Education* (pp. 446-450). Springer, Cham.

Kimbell, R. (2012). Evolving project e-scape for national assessment. International *Journal of Technology and Design Education, 22*(2), 135–155.

King, S. V. (2019). Artificial Intelligence Tutor Provides Personalized Learning. *Florida International University: Insider News.* [Online]. Retrieved: https://insider.fiu.edu/artificial-intelligence-cognii/. [Accessed 24 Oct 2022].

Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2020). A narrowing of AI research? *arXiv preprint arXiv:2009.10385*.

Knight, J. F., Carley, S., Tregunna, B., Jarvis, S., Smithies, R., de Freitas, S., Dunwell, I., & Mackway-Jones, K. (2010). Serious gaming technology in major incident triage training: a pragmatic controlled trial. *Resuscitation, 81*(9), 1175-1179.

Koehler, M. J., Mishra, P., & Cain, W. (2013). What is technological pedagogical content knowledge (TPACK)? *Journal of Education*, *193*(3), 13-19.

Koh, K. (2017). Authentic assessment. *Oxford Research Encyclopedia of Education*. DOI: https://doi.org/10.1093/acrefore/9780190264093.013.22.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT Press.

Kong, S. C., Cheung, W. M. Y., & Zhang, G. (2023). Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. *Educational Technology & Society*, *26*(1), 16-30.

Koretsky, M. D., McColley, C. J., Gugel, J. L., & Ekstedt, T. W. (2022). Aligning classroom assessment with engineering practice: A design-based research study of a two-stage exam with authentic assessment. *Journal of Engineering Education*, *111*(1), 185-213.

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788-8790.

Kumar, V., & Boulanger, D. (2020, October). Explainable automated essay scoring: Deep learning really has pedagogical value. In *Frontiers in Education* (Vol. 5, p. 572367). Frontiers Media SA.

Kung, C., & Yu, R. (2020). Interpretable models do not compromise accuracy or fairness in predicting college success. In *Proceedings of the Seventh ACM Conference on Learning@ Scale* (pp. 413–416). ACM.

Lam, T. Y., & Dongol, B. (2020). A blockchain-enabled e-learning platform. *Interactive Learning Environments, 30* (7), 1229-1251. DOI: https://doi.org/10.1080/10494820.2020.1716022.

Last, M., & Danon, G. (2020). Automatic question generation. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(6), e1382.

Latham, A., & Goltz, S. (2019, June). A survey of the general public's views on the ethics of using AI in education. In *International Conference on Artificial Intelligence in Education* (pp. 194-206). Springer, Cham.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory.* Oxford University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Learning Technology Standards Committee (LTSC). (2007). The Learning Object Metadata standard. *IEEE.* [Online]. Retrieved: https://www.ieeeltsc.org/working-groups/wg12LOM/lom

Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education*, *191*, 104646. DOI: https://doi.org/10.1016/j.compedu.2022.104646.

Lee, J., & Soylu, M. Y. (2023). *ChatGPT and Assessment in Higher Education* [White Paper]. Center for 21st Century Universities, Division of Lifetime Learning, Georgia Institute of Technology. https://c21u.gatech.edu/papers/chatgpt-and-assessment-higher-education.

Leslie, D. (2019). Understanding artificial intelligence ethics and safety. *arXiv preprint arXiv:1906.05684*.

Leydesdorff, L., & Rafols, I. (2012). Interactive overlays: A new method for generating global journal maps from Web-of-Science data. *Journal of Informetrics, 6*(2), 318-332.

Li, S., & Gu, X. (2023). A Risk Framework for Human-centered Artificial Intelligence in Education. *Educational Technology & Society*, *26*(1), 187-202.

Li, Z. Z., Cheng, Y. B., & Liu, C. C. (2013). A constructionism framework for designing game-like learning systems: Its effect on different learners. *British Journal of Educational Technology, 44*(2), 208-224.

Lim[1], T., Gottipati, S. & Cheong, M. (2022). Authentic Assessments for Digital Education: Learning Technologies Shaping Assessment Practices. In *Proceedings of the 30th International Conference on Computers in Education (ICCE 2022)*. 1, p. 587-592. Kuala Lumpur, Malaysia. ISBN: 978-986-972-149-3.

Lim, T., Gottipati, S., & Cheong, M. (2023). Ethical Considerations for Artificial Intelligence in Educational Assessments. In *Creative AI Tools and Ethical Implications in Teaching and Learning* (pp. 32-79). IGI Global.

1

Lim, T., Gottipati, S., Cheong, M., Ng, J. W. & Pang, C. (2022). Assessment Design for Digital Education: An Analytics-based Authentic Assessment Approach. In *2022 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*, Hong Kong.

Lim, T., Gottipati, S., Cheong, M., Ng, J. W., & Pang, C. (2023). Analytics-enabled authentic assessment design approach for digital education. *Education and Information Technology, 28*, 9025-9048. DOI: https://doi.org/10.1007/s10639-022-11525-3.

Lin, C. C., Huang, A. Y., & Lu, O. H. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. *Smart Learning Environments*, *10*(1), 41.

Lin, Q., Yin, Y., Tang, X., Hadad, R., & Zhai, X. (2020). Assessing learning in technology-rich maker activities: A systematic review of empirical research. *Computers & Education, 157*(103944). DOI: https://doi.org/10.1016/j.compedu.2020.103944.

Litman, D., Zhang, H., Correnti, R., Matsumura, L. C., & Wang, E. (2021, June). A fairness evaluation of automated methods for scoring text evidence usage in writing. In *International Conference on Artificial Intelligence in Education* (pp. 255-267). Springer, Cham.

Liu, Y., & Zhu, T. (2020). Individualized New Teaching Mode for Sports Biomechanics Based on Big Data. *International Journal of Emerging Technologies in Learning (iJET), 15*(20), 130-144.

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour, 1*, 0028. DOI: https://doi.org/10.1038/s41562-016-0028.

Macy, M., Macy, R., & Shaw, M. (2018). Bringing the ivory tower into students' homes: Promoting accessibility in online courses. *Ubiquitous Learning*, *11*(1), 13-21.

Martin Nunez, J. L., & Diaz Lantada, A. (2020). Artificial intelligence aided engineering education: State of the art, potentials and challenges. *International Journal of Engineering Education*, *36*(6), 1740-1751.

Martin, F., Ritzhaupt, A., Kumar, S., & Budhrani, K. (2019). Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation. *The Internet and Higher Education*, *42*, 34-43.

Martin, S., Diaz, G., Sancristobal, E., Gil, R., Castro, M., & Peire, J. (2011). New technology trends in education: Seven years of forecasts and convergence. *Computers & Education, 57*(3), 1893-1906.

Martinez-Maldonado, R., Gaševic, D., Echeverria, V., Fernandez Nieto, G., Swiecki, Z., & Buckingham Shum, S. (2021). What Do You Mean by Collaboration Analytics? A Conceptual Model. *Journal of Learning Analytics*, *8*(1), 126-153.

Martinez-Maldonado, R., Kay, J., Buckingham Shum, S., & Yacef, K. (2019). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. *Human–Computer Interaction*, *34*(1), 1-50.

Mateo, J. C., McCloskey, M. J., Leishman, C. D., & Federe, M. (2016). Scenario-Based Practical Exercises to Train and Assess General Cross-Cultural Competence for Special Operations Forces. In *Advances in Cross-Cultural Decision Making. Proceedings of the AHFE 2016 International Conference on Cross-Cultural Decision Making (CCDM)* (pp. 159-170). Springer. DOI: https://doi.org/10.1007/978-3-319-41636-6_13.

Mayfield, E., Madaio, M., Prabhumoye, S., Gerritsen, D., McLaughlin, B., Dixon-Román, E., & Black, A. W. (2019, August). Equity beyond bias in language technologies for education. In *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 444-460). Association for Computational Linguistics.

Mayfield, E., Madaio, M., Prabhumoye, S., Gerritsen, D., McLaughlin, B., Dixon-Román, E., & Black, A. W. (2019, August). Equity beyond bias in language technologies for education. In *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 444-460). Association for Computational Linguistics.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-analysis and Review of Online Learning Studies*. US Department of Education [Online]. Retrieved from https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf. [Accessed 14 Oct 2023].

Megahed, N. A., Abdel-Kader, R. F., & Soliman, H. Y. (2022). Post-pandemic education strategy: Framework for artificial intelligence-empowered education in engineering (AIEd-

Eng) for lifelong learning. In *International Conference on Advanced Machine Learning Technologies and Applications* (pp. 544-556). Springer, Cham.

Memarian, B., & Doleck, T. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI), and higher education: A systematic review. *Computers and Education: Artificial Intelligence*, 100152.

Merikko, J., Ng, K., Saqr, M., & Ihantola, P. (2022). To opt in or to opt out? Predicting student preference for learning analytics-based formative feedback. *IEEE Access*, *10*, 99195-99204.

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European journal of operational research*, *246*(1), 1-19.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record*, *108*(6), 1017-1054.

Mishra, S. (2020). Technology Applications in Education: Policy and Prospects. In: Mishra, S. & Panda, S. (eds). *Technology-enabled Learning: Policy, Pedagogy and Practice*, Chapter 2, 19-32. Commonwealth of Learning.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary research and perspectives, 1*(1), 3-62.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. DOI: https://doi.org/10.1371/journal.pmed.1000097.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics, 26*(4), 2141–2168.

Mott, J., Nyland, R., Williams, G., Atkinson, M., & Ceglia, A. (2017). The next-generation CBE architecture: A learning-centric standards-based approach. In *Handbook of research on competency-based education in university settings* (pp. 134-156). IGI Global.

Mougiakou, E., Papadimitriou, S., & Virvou, M. (2018, July). Intelligent tutoring systems and transparency: The case of children and adolescents. In *2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA)* (pp. 1-8). IEEE.

National Research Council. (1996). *National science education standards.* National Academies Press.

Nazaretsky, T., Cukurova, M., & Alexandron, G. (2022). An instrument for measuring teachers' trust in AI-based educational technology. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (pp. 56-66). ACM.

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*, *7*(5), e07014. DOI: https://doi.org/10.1016/j.heliyon.2021.e07014.

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, *28*(4), 4221-4241.

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021). A systematic review on AI-based proctoring systems: Past, present and future. *Education and Information Technologies*, *26*(5), 6421-6445.

Ogden, C. R., & Richards, I. A. (1923). *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism.* London Routledge & Kegan Paul.

Ouyang, F., Dinh, T. A., & Xu, W. (2023). A systematic review of AI-driven educational assessment in STEM education. *Journal for STEM Education Research*, *6*(3), 408-426.
Papa, R., & Jackson, K. M. (2021). Enduring questions, innovative technologies: Educational theories interface with AI. In *Intelligent Computing* (pp. 725-742). Springer, Cham.

Park, C., & Kim, D. G. (2020). Perception of instructor presence and its effects on learning experience in online classes. *Journal of Information Technology Education: Research*, *19*, 475-488. Popper, K. (1979). *Three Worlds*. University of Michigan.

Pelletier, K., Brown, M., Brooks, D. C., McCormack, M., Reeves, J., & Arbino, N. (2021). *2021 Educause Horizon Report: Teaching and Learning Edition.* Boulder, CO: Educause.

Pelletier, K., McCormack, M., Reeves, J., Robert, J., and Arbino, N. (2022). *2022 Educause Horizon Report: Teaching and Learning Edition*. Boulder, CO: Educause.

Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., and Grajek, S. (2023). *2023 Educause Horizon Report: Teaching and Learning Edition*. Boulder, CO: Educause.

Peña-Ayala, A. (2018). Learning analytics: A glance of evolution, status, and trends according to a proposed taxonomy. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(3), e1243.

Pereira, J. (2016). Leveraging chatbots to improve self-guided learning through conversational quizzes. *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 911-918). ACM.

Petersen, K., & Gencel, C. (2013). Worldviews, research methods, and their relationship to validity in empirical software engineering research. In *2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement* (pp. 81-89). IEEE.

Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, *64*, 1-18.

Philippe, S., Souchet, A. D., Lameras, P., Petridis, P., Caporal, J., Coldeboeuf, G., & Duzan, H. (2020). Multimodal teaching, learning and training in virtual reality: a review and case study. *Virtual Reality & Intelligent Hardware*, *2*(5), 421-442.

Piaget, J. (1954). The construction of reality in the child. *New York: Basic Books.* DOI: http://dx.doi.org/10.1037/11168-000.

Pontual Falcão, T., Lins Rodrigues, R., Cechinel, C., Dermeval, D., Harada Teixeira de Oliveira, E., Gasparini, I., ... & Ferreira Mello, R. (2022, March). A Penny for your Thoughts: Students and Instructors' Expectations about Learning Analytics in Brazil. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (pp. 186-196).

Popper, K. (1979). *Three worlds*. University of Michigan.

Prendes-Espinosa, M. P., Gutiérrez-Porlán, I., & García-Tudela, P. A. (2021). Collaborative work in higher education: tools and strategies to implement the e-assessment. *Workgroups eAssessment: Planning, Implementing and Analysing Frameworks*, 55-84.

Project, P. E., & Peirce, C. S. (1998). *The essential peirce* (Volume 2). Indiana University Press.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.

Raji, I. D., Scheuerman, M. K., & Amironesei, R. (2021, March). You can't sit with us: Exclusionary pedagogy in ai ethics education. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 515-525). ACM.

Ramesh, D., & Sanampudi, S. K. (2022). An automated essay scoring systems: a systematic literature review. *Artificial Intelligence Review*, *55*(3), 2495-2527.

Richards, R. (2012). Exploring Formative Assessment with a Project Using Mobile Phones. *Ubiquitous Learning: An International Journal*, *4*(2), 91-102.

Richardson, M., & Clesham, R. (2021). Rise of the machines? The evolving role of Artificial Intelligence (AI) technologies in high stakes assessment. *London Review of Education, 19*(1), 1-13.

Roberston, L., & Barber, W. (2017). New directions in assessment and evaluation: Authentic assessment in fully online learning communities. *Education Research*, *11*(3), 249-262.

Robinson, C. J., Driskel, L., Blauvelt, E., & Perry, L. (2021, July). Class Exercises Involving Ethical Issues Reinforce the Importance and Reach of Biomedical Engineering (and the Impact of the Coronavirus on Teaching Strategy and Measures of Assessment). In *2021 ASEE Virtual Annual Conference Content Access*. DOI: https://doi.org/10.18260/1-2—36796.

Rodrigues, S. P., van Eck, N. J., Waltman, L., & Jansen, F. W. (2014). Mapping patient safety: A large-scale literature review using bibliometric visualisation techniques. *BMJ Open*, *4*(3), e004468. DOI: https://doi.org/10.1136/bmjopen-2013-004468.

Rodrigues, S. P., van Eck, N. J., Waltman, L., & Jansen, F. W. (2014). Mapping patient safety: A large-scale literature review using bibliometric visualisation techniques. *BMJ Open*, *4*(3), e004468. DOI: https://doi.org/10.1136/bmjopen-2013-004468.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), e1355.

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

Russell, M. K. (1999). *Testing on computers: A follow-up study comparing performance on computer and on paper*. Boston College.

Sadler, D. R. (1989). Formative assessment in the design of instructional systems. *Instructional Science 18*, 119–144.

Sánchez-Prieto, J. C., Gamazo, A., Cruz-Benito, J., Therón, R., & García-Peñalvo, F. J. (2020). AI-driven assessment of students: Current uses and research trends. In *P. Zaphiris & A. Ioannou* (Eds.), Learning and Collaboration Technologies. Design, Experiences. 7th International Conference, LCT 2020, Held as part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I (pp. 292-302). Springer Nature. DOI: https://doi.org/10.1007/978-3-030-50513-4_22.

Santos, P., Cook, J., & Hernández-Leo, D. (2015). M-AssIST: Interaction and scaffolding matters in authentic assessment. *Journal of Educational Technology & Society*, *18*(2), 33-45.

Satu, M. S., Roy, S., Akhter, F., & Whaiduzzaman, M. (2018). IoLT: an IOT based collaborative blended learning platform in higher education. In *2018 International Conference on Innovation in Engineering and Technology (ICIET)* (pp. 1-6). IEEE.

Schiff, D. (2021). Out of the laboratory and into the classroom: the future of artificial intelligence in education. *AI & society*, *36*(1), 331-348.

Schneider, B., Dowell, N., & Thompson, K. (2021). Collaboration analytics—current state and potential futures. *Journal of Learning Analytics*, *8*(1), 1-12.

Schultz, M., Young, K., K. Gunning, T., & Harvey, M. L. (2022). Defining and measuring authentic assessment: a case study in the context of tertiary science. *Assessment & Evaluation in Higher Education*, *47*(1), 77-94.

Seale, J. K. (2013). *E-learning and disability in higher education: accessibility research and practice*. Routledge.

Seilhamer, R., Chen, B., Bauer, S., Salter, A. & Bennett, L. (2018). Changing Mobile Learning Practices: A Multiyear Study 2012–2016. *Educause Review.* [Online]. Retrieved from https://er.educause.edu/articles/2018/4/changing-mobile-learning-practices-a-multiyear-study-2012-2016. [Accessed 24 Oct 2022].

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. John Wiley & Sons.

Selwyn, N. (2021). *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury Publishing.

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. *International Journal of Educational Technology in Higher Education*, *18*(1), 1-23.

Sghir, N., Adadi, A., & Lahmer, M. (2023). Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022). *Education and Information Technologies*, *28*(7), 8299-8333.

Shabaninejad, S., Khosravi, H., Abdi, S., Indulska, M., & Sadiq, S. (2022, June). Incorporating Explainable Learning Analytics to Assist Educators with Identifying Students in Need of Attention. In *Proceedings of the Ninth ACM Conference on Learning@ Scale* (pp. 384-388). DOI: https://doi.org/10.1145/3491140.3528292.

Shapiro, J., & Blackman, R. (2020). Four steps for drafting an ethical data practices blueprint. *TechCrunch.* [Online]. Retrieved*: https://techcrunch.com/2020/07/24/four-steps-for-an-ethical-data-practices-blueprint/* [Assessed 15 November 2022].

Shi, S. J., Li, J. W., & Zhang, R. (2024). A study on the impact of Generative Artificial Intelligence supported Situational Interactive Teaching on students' 'flow' experience and learning effectiveness – a case study of legal education in China. *Asia Pacific Journal of Education*, 1-27.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer games and instruction, 55*(2), 503-524.

Shute, V. J., Lajoie, S. P., & Gluck, K. A. (2000). Individualized and group approaches to training. *Training and retraining: A handbook for business, industry, government, and the military*, 171-207.

Shute, V. J., Ventura, M., Bauer, M., & Zapata-Rivera, D. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning. *Serious games: Mechanisms and effects*, *2*, 295-321.

Shute, V., Rahimi, S., & Emihovich, B. (2017). Assessment for learning in immersive environments. In *Virtual, augmented, and mixed realities in education* (pp. 71-87). Springer, Singapore.

Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: ethics of AI and ethical AI. *Journal of Database Management*, *31*(2), 74-87.

Sibai, F. N. (2020, June). AI crimes: a classification. In *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)* (pp. 1-8). IEEE.

Sievert, C., & Shirley, KE. (2014). LDAvis: A method for visualizing and interpreting topics, *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces* (pp. 63–70). Baltimore, Maryland, USA.

Smith, B., & Welty, C. (2001). Ontology: Towards a new synthesis. *Formal Ontology in Information Systems*. ACM Press.

Sophonhiranrak, S. (2021). Features, barriers, and influencing factors of mobile learning in higher education: A systematic review. *Heliyon, 7*(4), e06696. DOI: https://doi.org/10.1016/j.heliyon.2021.e06696.

Soto, M., & Ambrose, R. (2016). Screencasts: Formative assessment for mathematical thinking. *Technology, Knowledge and Learning*, *21*(2), 277-283.

Stahl, B. C., Timmermans, J. O. B., & Mittelstadt, B. D. (2016). The ethics of computing: A survey of the computing-oriented literature. *ACM Computing Surveys, 48*(4), 1–38.

Stark, L., & Hoey, J. (2021). The ethics of emotion in artificial intelligence systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 782-793). ACM.

Steynberg, J., Van Biljon, J., & Pilkington, C. (2020). Design Aspects of a Virtual Reality Learning Environment to Assess Knowledge Transfer in Science. In *Innovative Technologies and Learning: Third International Conference, ICITL 2020, Porto, Portugal, November 23–25, 2020, Proceedings 3* (pp. 306-316). Springer International Publishing. DOI: https://doi.org/10.1007/978-3-030-63885-6_35.

Stone, A. (2023). Student perceptions of academic integrity: a qualitative study of understanding, consequences, and impact. *Journal of Academic Ethics*, *21*(3), 357-375.

Stringer, R. (2018). Realist ethical naturalism for ethical non-naturalists. *Philosophical Studies*, *175*(2), 339-362.

Sun, S., Wu, X., & Xu, T. (2023). A Theoretical Framework for a Mathematical Cognitive Model for Adaptive Learning Systems. *Behavioral Sciences*, *13*(5), 406.

Suppes, P. (1966). The uses of computers in education. *Scientific American*, *215*(3), 206-223.

Surahman, E., & Wang, T. H. (2022). Academic dishonesty and trustworthy assessment in online learning: a systematic literature review. *Journal of Computer Assisted Learning*, *38*(6), 1535-1553.

Švábenský, V., Vykopal, J., Čeleda, P., Tkáčik, K., & Popovič, D. (2022). Student assessment in cybersecurity training automated by pattern mining and clustering. *Education and Information Technologies*, *27*(7), 9231-9262.

Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice, 18*(4), 433-449. DOI: https://doi.org/10.1080/0969594X.2011.582838

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285.

Taylor, C. S., & Nolen, S. B. (2005). *Classroom assessment: Supporting teaching and learning in real classrooms*. Pearson.

Taylor, D. L., Yeung, M., Bashet, A. Z. (2021). Personalized and Adaptive Learning. In: Ryoo, J., Winkelmann, K. (eds) *Innovative Learning Environments in STEM Higher*

Education. SpringerBriefs in Statistics. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-58948-6_2.

Tepper, C., Bishop, J., & Forrest, K. (2020). Authentic assessment utilising innovative technology enhanced learning. *The Asia Pacific Scholar*, *5*(1), 70-75.

The Open University. (2014). *Policy on Ethical use of Student Data for Learning Analytics*. [Online]. Retrieved: https://help.open.ac.uk/documents/policies/ethical-use-of-student-data/files/22/ethical-use-of-student-data-policy.pdf [Accessed: 15 Nov 2022].

Tlili, A., Essalmi, F., Jemni, M., & Chen, N. S. (2018). A complete validated learning analytics framework: Designing issues from data preparation perspective. *International Journal of Information and Communication Technology Education (IJICTE)*, *14*(2), 1-16.

Tlili, A., Essalmi, F., Jemni, M., & Chen, N. S. (2019). A complete validated learning analytics framework: designing issues from data use perspective. *International Journal of Information and Communication Technology Education (IJICTE)*, *15*(3), 42-59.

Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms*. ASCD.

Tong, P., An, I. S., & Zhou, Y. (2020). Holistic and dynamic: teacher-researcher reflections on operating mobile-assisted learning tasks supported by WeChat for Chinese as a foreign language. *Instructional Science*, *48*, 729-763.

Topping, K. J. (2009). Peer assessment. *Theory into practice*, *48*(1), 20-27.

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2021). Recommendation on the ethics of artificial intelligence. UNESDOC Digital Library. [Online]. Retrieved: https://unesdoc.unesco.org/ark:/48223/pf0000380455 [Accessed 20 Nov 2022].

van Otterlo, M. (2017). From algorithmic black boxes to adaptive white boxes: Declarative decision-theoretic ethical programs as codes of ethics. *arXiv preprint arXiv:1711.06035*.

Viars, K., Cullen, M. A., & Stalker, A. R. (2017). Handheld Learning: Authentic Assessment Using iPads. In *The Experiential Library* (pp. 73-85). Chandos Publishing.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.

Wang, S., Sun, Z., & Chen, Y. (2023). Effects of higher education institutes' artificial intelligence capability on students' self-efficacy, creativity and learning performance. *Education and Information Technologies*, *28*(5), 4919-4939.

Weiser, M. (1991). The computer for the 21st century. *Scientific American, 265*(3), 94–104.

Weitzman, M. L. (1993). What to preserve? An application of diversity theory to crane conservation. *The Quarterly Journal of Economics*, *108*(1), 157-183.

White, J. P., Dennis, S., Tomko, M., Bell, J., & Winter, S. (2021). Paths to social licence for tracking-data analytics in university research and services. *PloS one*, *16*(5), e0251964.

White, M. C., & Bembenutty, H. (2013). Not all avoidance help seekers are created equal: Individual differences in adaptive and executive help seeking. *Sage Open*, *3*(2), 2158244013484916.

Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). *Ethical and societal implications of algorithms, data, and artificial intelligence: A roadmap for research*. London: Nuffield Foundation.

Wiley, D. A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. *The instructional use of learning objects, 2830*(435), 1-35.

Williams, P. (2019). Does competency-based education with blockchain signal a new mission for universities? *Journal of higher education policy and management*, *41*(1), 104-117.

Williamson, B. (2021). Psychodata: disassembling the psychological, economic, and statistical infrastructure of 'social-emotional learning.' *Journal of Education Policy, 36*(1), 129-154.

Williamson, B., Bayne, S., & Shay, S. (2020). The datafication of teaching in Higher Education: critical issues and perspectives. *Teaching in Higher Education*, *25*(4), 351-365.

Wu, S. Y., & Yang, K. K. (2022). The Effectiveness of Teacher Support for Students' Learning of Artificial Intelligence Popular Science Activities. *Frontiers in Psychology*, *13*, 868623.

Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, *140*, 103599.

Yang, F., & Goh, Y. M. (2022). VR and MR technology for safety management education: An authentic learning approach. *Safety science*, *148*, 105645. DOI: https://doi.org/10.1016/j.ssci.2021.105645.

Yu, P., Xu, H., Hu, X., & Deng, C. (2023, October). Leveraging generative AI and large Language models: a Comprehensive Roadmap for Healthcare Integration. In *Healthcare* (Vol. 11, No. 20, p. 2776). MDPI.

Zhang, L., Thijs, B., & Glänzel, W. (2011). The diffusion of H-related literature. *Journal of Informetrics, 5*(4), 583-593.

Zhao, X., Li, X., Wang, J., & Shi, C. (2020). Augmented reality (AR) learning application based on the perspective of situational learning: high efficiency study of combination of virtual and real. *Psychology*, *11*(9), 1340-1348.

Ziker, C., Ydo, E., Zapata-Rivera, D., Hillier, M., & Casale, M. (2020, June). Special Session—Challenges and Opportunities for Assessment in XR. In *2020 6th International Conference of the Immersive Learning Research Network (iLRN)* (pp. 421-423). IEEE.

Appendix I

Paper	Document Type	Year	Citation Count	Publication Source	Publisher
Gupta and Chen (2022)	Original Article	2022	3	Journal of Information Systems Education	ISCAP - Information Systems and Computing Academic Professionals
Chounta et al. (2022)	Original Article	2022	4	International Journal of Artificial Intelligence in Education	Springer
Deho et al. (2022)	Review Article	2022	3	British Journal of Educational Technology	John Wiley and Sons Inc
Shabaninejad et al. (2022)	Conference Paper	2022	-	L@S 2022 - Proceedings of the 9th ACM Conference on Learning @ Scale	Association for Computing Machinery, Inc
Nazaretsky, Cukurova and Alexandron (2022)	Conference Paper	2022	2	ACM International Conference Proceeding Series	Association for Computing Machinery
Pontual Falcão et al. (2022)	Conference Paper	2022	1	ACM International Conference Proceeding Series	Association for Computing Machinery
Merikko et al. (2022)	Original Article	2022		IEEE Access	Institute of Electrical and

					Electronics Engineers Inc.
Khairy et al. (2022)	Original Article	2022	-	International Journal of Advanced Computer Science and Applications	Science and Information Organization
Megahed, Abdel-Kader and Soliman (2022)	Book Chapter	2022	1	Lecture Notes on Data Engineering and Communications Technologies	Springer Science and Business Media Deutschland GmbH
Conati et al. (2021)	Original Article	2021	12	Artificial Intelligence	Elsevier B.V.
González-Calatayud, Prendes-Espinosa and Roig-Vila (2021)	Review Article	2021	9	Applied Sciences (Switzerland)	MDPI AG
White et al. (2021)	Original Article	2021	2	PLoS ONE	Public Library of Science
Ahn et al. (2021)	Conference Paper	2021	-	ACM International Conference Proceeding Series	Association for Computing Machinery
Stark and Hoey (2021)	Conference Paper	2021	11	FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency	Association for Computing Machinery, Inc

Papa and Jackson (2021)	Conference Paper	2021	-	Intelligent Computing - Proceedings of the 2021 Computing Conference	Springer Nature
Kim et al. (2021)	Conference Paper	2021	-	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	Springer Science and Business Media Deutschland GmbH
Litman et al. (2021)	Conference Paper	2021	-	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	Springer Science and Business Media Deutschland GmbH
Casas-Roma and Conesa (2021)	Book Chapter	2021	1	Intelligent Systems and Learning Data Analytics in Online Education	Elsevier
Costas-Jauregui et al. (2021)	Conference Paper	2021	2	Proceedings - Frontiers in Education Conference, FIE	Institute of Electrical and Electronics Engineers Inc.
Elshafey et al. (2021)	Conference Paper	2021	4	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	Springer Science and Business Media Deutschland GmbH
Schneider, Dowell and	Original Article	2021	13	Journal of Learning Analytics	UTS ePRESS

Thompson (2021)					
Gedrimiene et al. (2020)	Original Article	2020	16	Scandinavian Journal of Educational Research	Routledge
Kumar and Boulanger (2020)	Original Article	2020	8	Frontiers in Education	Frontiers Media S.A.
Khosravi, Sadiq and Gasevic (2020)	Conference Paper	2020	29	SIGCSE 2020 - Proceedings of the 51st ACM Technical Symposium on Computer Science Education	-
Martín Núñez and Lantada (2020)	Original Article	2020	4	International Journal of Engineering Education	Tempus Publications
Hakami and Hernández-Leo (2020)	Conference Paper	2020	2	CEUR Workshop Proceedings	CEUR-WS
Mougiakou, Papadimitriou and Virvou (2019)	Conference Paper	2019	3	2018 9th International Conference on Information, Intelligence, Systems and Applications, IISA 2018	Institute of Electrical and Electronics Engineers Inc.
Mayfield et al. (2019)	Conference Paper	2019	17	ACL 2019 - Innovative Use of NLP for Building Educational Applications, BEA 2019 - Proceedings of the 14th Workshop	-

Latham and Goltz (2019)	Conference Paper	2019	5	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	Springer Verlag
Tlili et al. (2019)	Original Article	2019	3	International Journal of Information and Communication Technology Education	IGI Global
Kiennert et al. (2019)	Original Article	2019	7	IEEE Access	Institute of Electrical and Electronics Engineers Inc.
Peña-Ayala (2018)	Original Article	2018	42	Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery	-
Tlili et al. (2018)	Original Article	2018	8	International Journal of Information and Communication Technology Education	IGI Global

Appendix II

Paper	AI Ethical Issue(s) as Cited in Paper	Breakdown of AI Ethical Issue(s) as Cited in Paper	Breakdown of Mitigation/ Intervention Methods for AI Ethical Issue(s) Cited in Paper
Gupta and Chen (2022)	a. Inclusivity b. Fairness	**a. Inclusivity, b. Fairness** - Inclusivity in an AI-enabled conversational agent represents the meeting of different student needs (e.g., disadvantaged students) in a personalized learning environment at scale. For example, virtual teaching assistant *Jill Watson* (today known as *Sylla*) at Georgia Tech was cited to lack capability to understand healthcare or racial sensitive conditions such as learner's pregnancy or minority group backgrounds (Eicher, Polepeddi and Goel, 2018). The aspects of ensuring non-prejudice and non-favoritism towards a learner's sensitive attributes fit well with the concept of fairness.	**a. Inclusivity, b. Fairness** - Chatbot *Sammy* helps promote judgment-free inclusiveness to learners. Its inclusivity also extends to ubiquitous access, and is helpful to learners with disabilities (e.g., visual or hearing impaired) or learning disorders. - Chatbot *CiSA* is designed to promote equity and social inclusivity for international students (Heo & Lee, 2019).
Chounta et al. (2022)	a. Accountability b. Accuracy c. Auditability	**a. Accountability** - Accountability risks exist if dependency on AI undermines	-

d. Explainability e. Fairness f. Privacy	educators' role in assessment (e.g., AI-automated formative assessments that diminishes role of educators in learning scaffolding), especially if the outcomes of AI are harmful, erroneous or inappropriate. **b. Accuracy** - Predictive model with imbalanced dataset (e.g., gender) is less effective (e.g., for minority gender). The possibility of discriminatory and unfair practice extends to other socio-economic demographical information, such as ethnicity, underrepresented groups etc. **c. Auditability, d. Explainability** - These relate to the importance of AI-automated assessments to be developed in an explainable and transparent manner to safeguard trust and fairness with human stakeholders. E.g., for AI recommender systems, why are some assessment questions recommended over others.

		e. Fairness - This relates to the importance of appropriateness and fairness in assessment. **f. Privacy** - This relates to ethical risks on learner data collection, storage, purpose of use and extent of use when applied to AI assessments.	
Deho et al. (2022)	a. Fairness b. Explainability	**a. Fairness** - Ensuring non-prejudice and favoritism towards a learner's sensitive attributes. For example, an intelligent tutoring system oversampled learners from WEIRD (white, industrialized, rich, and democratic) countries relative to non-WEIRD countries (Blanchard, 2012). In another example, models that predicted poor performance discriminated against African Americans (Hu and Rangwala, 2020). **b. Explainability**	**a. Fairness** - Fairness-minded human-in-the-loop (i.e., active human oversight). - Use of unfairness mitigation techniques and algorithms in the entire AI pipeline, including pre-processing for data, in-processing within the model, and post-processing of model results. - Authors suggested that goal of fairness may not be equal treatment, but the "needed" treatment to booster learning success. Useful to study the factors and the extent to which the factors apply in terms of what is "needed".

		- Research suggests that interpretable models (e.g., logistic regression) may provide less unfairness as compared to complex fairness-aware models, with robust accuracy results (Kung and Yu, 2020). This supports the notion of using interpretable and explainable models for AI in assessments.	
Shabaninejad et al. (2022)	a. Explainability	**a. Explainability** - This relates to the overcoming of "black-box" AI recommendations, which provide no insights into recommendation rationales (Abdi, 2020), and may be plagued with biases and confounding problems (Bastani, Bastani and Kim, 2018; Khosravi et al., 2021).	**a. Explainability** - Implementation of *Student Inspection Facilitator*, a context-independent learning analytics dashboard at the University of Queensland, which provides explainable recommendations to guide learning intervention for educators.
Nazaretsky, Cukurova and Alexandron (2022)	a. Trust b. Explainability c. Accountability	**a. Trust** - Research cites lack of human properties (e.g., lack of affect, emotions, pedagogical intuition) as main reason for low AI trust among educators. **b. Explainability**	**a. Trust** - Participatory or co-design process in the creation of AI systems, to reduce barriers of trust. **b. Explainability** - Create transparent AI systems and avoid "black-box" AI solutions.

		- Research cites lack of transparency on how AI makes decisions contributing to low trust among educators. Transparency is tied to information availability, accessibility conditions, possibility of pragmatic decision-making assistance, and user knowledge. **c. Accountability** - This relates to the possibility for educators who construct, operate and use AI to be accountable and responsible for AI systems and decisions.	However, too much transparency may also lead to information overload - a paradox known as "transparency paradox". Authors suggest that it may be useful to study the extent to which transparency can benefit use of AI in assessments. **c. Accountability** - Active human oversight of AI decisions, as greater autonomy and control can (i) improve trust between educators and AI systems, (ii) reduce anxiety of educators' replacement by AI systems, and (iii) reduce errors where the educators are accountable.
Pontual Falcão et al. (2022)	a. Privacy b. Trust	**a. Privacy** - This relates to the possibility of violation to individuals' rights to privacy when too much data surveillance exist, especially when data is used beyond academic purposes, for control and surveillance to modify human behavior.	**a. Privacy** - Clear institutional and national regulation of data protection and access. **b. Trust** - Experience relating to the use of AI may influence educators' and learners' opinions. To bridge

	- Results shared that learners were confident of data privacy safeguards that exist in institutions.	expectations of AI use for educators and learners, it is important to improve data literacy levels and knowledge of tools available.
	b. Trust - Research suggests discomfort of AI-driven decision making that involves ranking, sorting and classifying individuals, that may reflect political interests, social values, and risks of omissions or biases. - Research cites educators' discomfort at lack of autonomy and control due to its use in the appraisal of teaching performance and excessive intrusion in learners' learning routine. Learners may also be worried that their autonomy and independent decision making may be deprived. - Research cites that learners voiced that it may not be fair to impose the obligation of learning support and intervention to educators, and that learners were not confident that incorporating AI in assessments can result in improvements of feedback	- Trust can be shaped by institutional commitment and context. Institutional expectations and support may be more explicitly shared to improve trust.

		quality. Educators may have a lack of belief that feedback quality may improve as timely and quality feedback is time consuming, due to the need to balance their workload and institutional expectations.	
Merikko et al. (2022)	a. Privacy	**a. Privacy** - Research shares that learners are open to sharing data related to demographics and learning performance but are apprehensive about sharing when it comes to their online behavior, sensitive or process data. Further, the more personal and granular the data are, the less likely the learners will share them. - Research shares that learners who were not performing well are less likely to share their performance data. This may be tied to help-seeking avoidance, as seeking help may be a sign of weakness and a threat to self-esteem (White and Bembenutty, 2013). - Research was not able to predict data opt-in behavior, based on	**a. Privacy** - Clear guidelines to data sharing, including the type of data shared and the purpose for which it is used. It is recommended that: (i) Request for data opt-in is made for a specific intervention, rather than requesting for general data consent. For instance, learners can select between personalized AI-driven feedback, general feedback, or no feedback, explicitly stating that only the former requires data disclosure. (ii) Decreasing or low opt-in rate be investigated for better learning intervention. For instance, if there are issues that raised suspicion, if there are reasons why learners opt-out, or if the learners understood the reasons for learning intervention.

		learning engagement, learning strategies and/or willingness for data sharing. However, research was limited by limited sample size.	
Khairy et al. (2022)	a. Accuracy	**a. Accuracy** - This relates to the accuracy of the adaptable humanoid robot in understanding or interpreting the responses of learners in an oral assessment. Correct answer rate, if affected by learner's pronunciation or robot's lack of contextual understanding, can affect confidence in the AI system.	**a. Accuracy** - This is a technology issue. Authors proposed to train robots with wide pronunciation data samples, including native and foreign speakers.
Megahed, Abdel-Kader and Soliman (2022)	a. Privacy b. Fairness	**a. Privacy** - This relates to data privacy, surveillance and security issues arising from AI use. **b. Fairness** - This relates to prejudice and discriminatory issues arising from AI use.	**a. Privacy, b. Fairness** - Integrate AI ethics in entire AI development pipeline.
Conati et al. (2021)	a. Explainability	**a. Explainability** **-** Research shares that lack of	**a. Explainability** - Explainability of AI-driven hints,

		explainability reduces trust, perceived usefulness, and intention to use AI system again.	incorporating "why" and "how" explanations on how the hints are derived, are useful to effect positive learner perceptions. - Personalizing explainable AI-driven hints incorporated in an intelligent tutoring system, may improve learning when undertaking an assessment. This is due to modulation effects of user characteristics on perception and explanation of hints.
González-Calatayud, Prendes-Espinosa and Roig-Vila (2021)	a. Explainability	**a. Explainability** - Research cites lack of transparency of AI decision-making algorithms. - Research also cites lack of pedagogical underpinning and AI training, which affects the meaningful development of assessments with pedagogical reference models when AI is applied.	**a. Explainability** - Training to stakeholders, in terms of both AI technology and relevant pedagogical reference models, to understand limitations, possibilities and characteristics of AI-driven assessments.
White et al. (2021)	a. Privacy	**a. Privacy** - Research shares that risks of collecting tracking data for AI use include data breaches, loss of	**a. Privacy** - Key factors that may lead to acceptance of data sharing includes: (i) Perception of respect of data

		privacy, and community backlash. Safeguarding of trust and confidence are important for stakeholder security, privacy and risk of harm.	collector to data privacy (ii) Trust in data collector (iii) Extent of which data collector benefits from data (iv) Sensitivity of data collection and use (v) Inherent risk of harm Research shares that there are no differences in judgment of acceptance across age, gender of educational levels, and there are no marked differences between educators and learners.
Ahn et al. (2021)	a. Accuracy	**a. Accuracy** - Research shares that automated grading of learners' work, which contains complex data, rich semantic meaning, and idiosyncratic and local nuances, may not be well graded by present computational approaches, that utilize metrics such as counts of parts of speech and essay length as proxies for writing complexity and quality. - Further uncertainties and	**a. Accuracy** - Research shares that crowdsourcing of assessment grading can be accurate, in agreement with experts. Further, it provides learning value to the crowdsourced graders.

		challenges exist, as it requires significant corpus of training data that may not be useful to create assessments tailored to local contexts. - While pretrained models of natural language processing showed promise, utilizing and finetuning comes at high computational costs.	
Stark and Hoey (2021)	a. Privacy b. Fairness c. Accountability d. Trust	**a. Privacy** - Research highlights how individuals are sensitive about data sharing and utilization pertaining to their emotions and emotional expressions. For example, the Facebook emotional contagion study (Kramer, Guillory and Hancock, 2014) were criticized for manipulating emotive content of users. **b. Fairness** - Research highlights how universal assumptions on emotional states are harmful, due to different cultural context of emotional interactions and norms.	**a. Privacy** - Important to consider concerns on use of such data for profiling, tracking and behavioral shaping. **b. Fairness, c. Accountability** - Consider different cultural context of emotional interactions and norms, individual and collective subjective assessments, and changing global paradigms of emotional contexts and subjectivity. **d. Trust** - Due to ethical valences and social effects of the diversity of ethical opinions, it is useful to establish a global objective agreement and

		c. Accountability - Research highlights lack of accountability of AI developers, relating development of AI systems with recognition of diverse human attitudes to emotions, and the possibility that emotions can be harnessed as a social phenomenon. **d. Trust** - Research highlights lack of consensus objective agreement on emotion at a global level as an issue as the large variation in the implied social and ethical responsibilities have normative implications for AI systems, e.g., when considering ethical values such as accountability and fairness.	consider how e.g. cultural context differences can impact the design and deployment of AI systems.
Papa and Jackson (2021)	a. Fairness b. Explainability c. Human centricity	**a. Fairness** - Research cites biasness introduced to AI algorithm, as humans can have implicit bias. **b. Explainability** - Research cites that AI algorithms	**a. Fairness, b. Explainability, c. Human centricity** - Ensure human-in-the-loop in AI systems. - Formalize existing human norms and values into expressive and flexible responsible coding, by

		may be opaque, creating unjustified actions and discrimination.	using decision-theoretic logic programming to achieve value alignment (van Otterlo, 2017).
		c. Human centricity - Research cites that agency and autonomy of users should not be impacted by profiling, ranking, and personalizing derived from AI algorithms. Learning should not be viewed as "product-oriented learning experiences" (Duignan, 2020). Further, there should be care applied, when it comes to AI algorithm manipulating learner behaviors and emotions.	- Build an AI ethics framework, that will make explicit ethic issues (such as privacy, explainability, fairness etc.), so that users can anticipate and avoid ethical issues on AI systems (Shapiro and Blackman, 2020). - Centre AI ethics around learning theories, to provide humanistic and social dimensions to the AI-mediated process.
Kim et al. (2021)	a. Accuracy	**a. Accuracy** - Research shares that, to allow for shorter assessments in online learning, it is imperative that the question set reduction is done in such a way where the reduced set can approximate the original assessment's evaluation of learning. This can help ensure reliability and trust on assessment instrument.	**a. Accuracy** - Research proposes an AI approach to reliably identify reduced assessment size, and approximate test scores.

Litman et al. (2021)	a. Fairness b. Explainability c. Accuracy	**a. Fairness** - Research found small but significant algorithmic bias with respect to demographics, such as socio-economic status, race and gender. - In AI-driven systems, tradeoff may exist between e.g.: (i) Fairness and reliability (increasing fairness reduced reliability) (ii) Fairness and explainability (increasing accuracy may reduce explainability) **b. Explainability, c. Accuracy** - Research cited how feature-based models are more explainable, while neural network models are more accurate.	**a. Fairness** - Research shares fairness mitigation techniques using fairer feature selection strategies, which may work, over and above mitigating imbalanced dataset and/or pre-training bias-free models. **b. Explainability, c. Accuracy** - Research shares hybrid feature-based and neural network models, that combined accuracy with explainability.
Casas-Roma and Conesa (2021)	a. Fairness b. Explainability c. Auditability d. Privacy e. Human centricity	**a. Fairness** - This relates to compromising diversity and equity of learners, such that minority groups are disadvantaged, and their needs underrepresented.	**a. Fairness, b. Explainability, c. Auditability, d. Privacy, e. Human centricity** - Ensure human-in-the-loop in AI systems. - Utilize moral reasoning AI systems, that account for potential

	b. Explainability - This relates to opacity of AI generated decisions. **c. Auditability** - This relates to lack of traceability of AI algorithm process. **d. Privacy** - This relates to consent and confidentiality, and monitoring of learner behaviors and habits. **e. Human centricity** - This relates to differential access and availability of technology and resources, such that risk of digital divide and technology exclusion arises. Further, there may exist a risk of forming unfair groups due to classification and profiling.	ethical outcomes in AI generated decisions. - Utilize a set of AI ethics guidelines, that makes explicit ethic issues (such as privacy, explainability, fairness etc.), so that users can anticipate and avoid ethical issues on AI systems (AI HLEG, 2019). **a. Fairness** - Ensure data sanitization that sanitizes data of potential discriminatory decisions, such that neither data, model nor predictions affect vulnerable learners. **b. Explainability** - Use counterfactual explanations to evaluate decisions. Counterfactual explanations allow users to "*identify what would have needed to be different in order for the AI to have decided otherwise*". This helps to rationalize the AI generated decision. **c. Auditability**

			- Ensure AI process is logged, tracked, interpreted, and checked by independent auditors. AI system transparency is paramount. **d. Privacy** - Ensure that data subjects are protected from unfair data use. **e. Human centricity** - Ensure equitable access to fair opportunities and quality level in AI assessment environment (e.g., requirement of certain internet connection speed or hardware).
Costas-Jauregui et al. (2021)	a. Inclusivity b. Privacy c. Trust d. Accountability	**a. Inclusivity** - This relates to inclusivity in an AI-integrated visualization dashboard for learners with learning disabilities. It is estimated that 34% of students in the US between 3 and 21 have some form of learning disabilities (U.S. Department of Education, 2020). - This also relates to the risk of excluding students due to AI categorization. Further, system-	**a. Inclusivity** - *Smart Ecosystem for Learning and Inclusion (SELI)* platform allows students to warn the AI system about their own disabilities. In the case of, for instance motor disability, the platform can grant more or unlimited time to complete an assessment. - Authors proposed to make AI system available for institutional and/or regulatory evaluations.

	induced exclusion, and ignorance of cultural diversity of students' learning in platform construction, have ramifications to the validity and reliability of data. - It is useful to consider the sensitivity of communication and feedbacks generated by the AI system, so that learning can be enhanced. **b. Privacy** - This relates to the collection and utilization of private data. - Authors suggest that there may exist a lack of interoperability of regulatory guidelines (e.g. EU and Latin America), and clarity regarding whether if the institution or the students own the data that are shared by the students. **c. Trust** - This relates to clarity of purpose and specifications of the AI system. **d. Accountability** - This relates to the moral	- Development of AI system would be best served as an inter-disciplinary approach, integrating disciplines such as Anthropology and Sociology. - Regarding the sensitivity of communication and feedbacks generated by the AI system, authors proposed to study such communications, with due inputs from educators, psychologists and communication experts. **b. Privacy** - Authors proposed to study the eight essential principles for ethical use of learner data published by the Open University (The Open University, 2014). - Establish clear window periods for the collection and use of data. SELI platform is also exploring use of emerging technologies including blockchain and smart contracts to manage data securely. - Useful to include all privacy ethical and legal issues in the inception of system design. Authors

		obligation for institutions to reflect and act, given that it has access to data that may know and understand how students learn.	proposed to express explicit consent request at the onset of data collection, detailing the purpose, duration, security policy and the persons that may be involved in the data processing process. **c. Trust** - Include educators and students in the co-creation process of the AI system. **d. Accountability** - Clear institutional policy to identify how institutions are accessing data that can impact students learning. - Record data access, learning intervention and impact of intervention in the design of AI systems.
Elshafey et al. (2021)	a. Academic integrity	**a. Academic integrity** - This relates to the concern of cheating using remote assessment platforms.	**a. Academic integrity** - Implement AI authentication tools such as facial recognition and/or voice recognition. - For anti-cheating AI techniques: (i) Head pose estimation can be

			used to track assessment takers' attention. (ii) Gaze estimation to determine angle of students' gaze. (iii) Scene change detection to look for changes in background environment. (iv) Object detection to detect unauthorized objects in environment within camera view.
Schneider, Dowell and Thompson (2021)	a. Explainability b. Accuracy c. Inclusiveness	**a. Explainability** - This relates to the potential complexity of interdependent modelling and assessing group dynamics and group outcomes, which may raise challenges on explainability of AI systems. Interdependent models (which looks at students' influences on one another over time) may be significantly more complex than independent models (which looks at students as isolated events). **b. Accuracy** - This relates to the capturing of multimodal data that best assess the collaborative interactions of	**a. Explainability** - Assess the trade-off between model complexity and explainability, and ascertain if simpler models are sufficient. - It may be useful to apply simpler models to understand dynamical interdependence, rather than more complex models. **b. Accuracy** - Understand and capture multimodal data that best measure collaborative interactions at various levels (e.g., group or individual), contexts (e.g., cultural), and time or phase of collaboration. - Authors proposes new data

		students. **c. Inclusiveness** - This relates to the development of tools that can respond to unique needs of users.	collection tools that allow visualization of social interactions in real time, e.g., integrating computer vision and ubiquitous sensors. **c. Inclusiveness** - Include educators and students in the co-creation process of the AI system.
Gedrimiene et al. (2020)	a. Inclusivity b. Privacy c. Accountability	**a. Inclusivity** - This relates to conformity, peer pressure and segregation that may be reinforced, as a result of AI generated decisions, which can impact both educators and learners. **b. Privacy** - This relates to the right of students to join or withdraw from certain AI-influenced activities, especially vulnerable groups such as students with learning disabilities, language barriers, or students who come from lower socio-economic backgrounds. **c. Accountability**	**a. Inclusivity** - Ethical AI practices should be defined and regulated at institutional and national levels, so that educators and students are protected by policies, as AI ethics are complex issues that may not be entirely understood or may be understood differently by stakeholders. **b. Privacy** - Clearly explain all consequences of joining and withdrawal to students. When students are provided the ability to opt in or out, unfavorable consequences should

		- This relates to how students are data subjects who are not able to influence the handling of data in an ethical manner.	not be placed on students. **c. Accountability** - All stakeholders are to understand the purpose, access, utilization boundaries and the interpretation possibilities of data.
Kumar and Boulanger (2020)	a. Explainability	**a. Explainability** - This relates to the opacity of AI generated decisions.	**a. Explainability** - Use **SH**apley **A**dditive ex**P**lanations (SHAP) at a rubric grading level, which provides robust explanations to individual predictions, while accounting for global factors affecting the performance of the AI model. - Explainable AI can help assess trustworthiness of complex algorithms (including ensembles and deep learning models), such as feature-based multi-layer perceptron deep neural network. - Explainable AI can play roles in parameter tuning to improve interpretability and generalizability (e.g., tuning of hidden layer depth), discover decision making process (e.g., simpler or more complex

			feature selection to make up the explanation), and provide granular personalized feedback to learners (customizable and trustworthy explanations to learners).
Khosravi, Sadiq and Gasevic (2020)	a. Privacy b. Accountability	**a. Privacy** - This relates to the collection and utilization of private data. **b. Accountability** - This relates to the consent, and non-maleficence academic interventions in the AI process.	**a. Privacy** - Express explicit consent request at the onset of data collection by the AI system, seeking permission to improve academic developers' understanding of the process of learning on the platform. **b. Accountability** - Endorse consent form, which is updated to include changes in purpose, scope and details of research. - Researchers are to conduct research in the spirit of non-maleficence, such that learners' learning experiences and academic performances are not harmed.
Martín Núñez and Lantada (2020)	a. Inclusivity b. Privacy	**a. Inclusivity** - This relates to social, race or gender prejudices and stereotypes	**a. Inclusivity** - Diverse AI development teams to identify and lower biases resulting

		that may be perpetuated by AI systems. **b. Privacy** - This relates to poor data collection and sharing practices.	from use of biased dataset or wrongly trained models. **b. Privacy** - Quality and inclusive data management practices that promotes transparency in data collection, use and dissemination. Such practices should also adapt to diversity, and technological, social, and educational trends.
Hakami and Hernández-Leo (2020)	a. Fairness b. Explainability c. Accountability d. Human centricity	**a. Fairness** - This relates to conscious or unconscious treatment of data and algorithm, resulting in biased algorithm outcomes. - Definition of fairness is plagued with subjective and contextual issues. **b. Explainability** - This relates to presence of transparency to observe the motives and logic underlying autonomous algorithmic decisions and actions. However, full transparency may be harmful, as users can "game" the	**a. Fairness, b. Explainability, c. Accountability, d. Human centricity** **-** Creation and use of code of ethics and guidelines for individuals and institutions. **a. Fairness** - Importance of clarifying definitions of fairness, to reduce subjectivity and improve contextual awareness. **b. Explainability** - Greater stakeholder knowledge and understanding of AI systems

	system to their benefit and the detriment of others. **c. Accountability** - This relates to the processes where relevant stakeholders provide reasons and take responsibilities for the actions of decisions influenced by AI algorithms. **d. Human centricity** - This relates to states of human wellbeing (e.g. psychological wellbeing, satisfaction). - This also relates to capacity to learn and level of autonomy to make learning decisions.	and their underlying mechanisms. - Design AI systems that have the ability to provide rationale for autonomous AI decisions (e.g. reasons behind AI recommendations). **c. Accountability** - Accountability begins from designers and developers of the AI system. - Guiding questions to consider for good accountability practices: (i) Consequences of algorithmic decisions on societies and individuals. (ii) Influence of consequences and the number of people affected by the consequences. (iii) Degree of awareness on how AI algorithms drive decisions. (iv) Possibilities of occurrence of discrimination and bias, and how this can impact public perception. (v) Preventive strategies and techniques that can be put in place at the onset of system design. (vi) Maintenance strategies and

			techniques that can intervene AI system during deployment. (vii) Optimization strategies and techniques that can improve AI system post-deployment. **d. Human centricity** - Beyond assessment performance metrics, it is useful to consider relevant AI system design factors such as socio-emotional aspects, self-regulation, cognitive load and inclusivity considerations.
Mougiakou, Papadimitriou and Virvou (2019)	a. Privacy b. Human centricity	**a. Privacy** - This relates to fair data security practices such as choice, notice, security, and access. **b. Human centricity** - This relates to the presence of reversible and clear processes, and the possibility to intervene for blocking, termination, correction and erasure.	**a. Privacy, b. Human centricity** - Paper has emphasis on ethics compliance with guidelines and regulations (e.g., EU General Data Protection Regulation (GDPR), United Nations Convention on the Rights of the Child etc.). **a. Privacy** - Use of plain, concise, and easy to understand language, when requesting for data consent from data subjects. This includes communicating significance,

			consequences, and period of data collection. - Data subjects should have the right not to be subjected to AI based decision processing and profiling. Information regarding the logic of, for instance, profiling, should be shared with data subjects. **b. Human centricity** - Clear segmentation of AI system into key components to improve "intervenability" of system. This includes, e.g., separation of registration, consent, interaction, assessment, grading, profiling and automated decision making.
Mayfield et al. (2019)	a. Fairness b. Accuracy c. Privacy	**a. Fairness** - This relates to biased modeling along demographic classifications like age, race and gender. - Biased modeling is exacerbated by amplification effect of machine learning on real-world outcomes (e.g. recidivism prediction with racial biasness in judicial hearings), or disproportionate prediction error	**a. Fairness, b. Accuracy, c. Privacy** - Clear taxonomies of ethics will be useful to guide research and AI system design choices, so that attempts to address ethical issues can be holistic, rather than ad hoc. **a. Fairness** - Intentionally designed agent

	(e.g. errors in facial recognition, in particular for female with darker skin tones). - Topics of fairness include: (i) Allocation harm: This relates to the equitable distribution of resources of learning, such that the possibility of differential outcome distributions generated by AI systems are minimized. (ii) Representational harm: This relates to the stereotyping bias perpetuated by data and/or algorithm, resulting in the marginalizing of groups of learners. **b. Accuracy** - This relates to the lack of reliability using AI systems, as present systems are generally rigid in formulation of tasks and grading. For instance, the rejection of lexicon and grammar of minority dialects. This limits the choice of tasks, types of acceptable answers, and tolerable styles of writing. **c. Privacy**	representation of identities (e.g. race, appearance, voice, language, gender), rather than relying on data-driven agent representation of identities, which may lack the nuanced understanding of identities (e.g. intersection of marginalized groups). These pedagogical agents can influence students' perception of their own identity and belongingness. - Beyond grading of students' text responses, assessment of affective states using affect-aware systems and multimodal data can also be tainted by biasness. Care should be placed on affect-detection systems. **b. Accuracy** - Flexibility in system design of assessment tasks and languages, provision of topic selection and choice that reflects culturally aligned opportunities, and collaborative sharing of work to receive feedback beyond AI generated responses. This can help to promote accuracy in grading,

		- This relates to undesirable anxieties and behavioral change related to constant surveillance.	e.g., responses in minority dialects. **c. Privacy** - Compliance with regulations and guidelines (e.g., Children's Online Privacy Protection Rule (COPPA), Family Educational Rights and Privacy Act (FERPA), EU General Data Protection Regulation (GDPR)).
Latham and Goltz (2019)	a. Fairness b. Accountability c. Privacy d. Explainability	**a. Fairness** - This relates to source data bias, and lack of objectivity of AI algorithms exhibiting inherent biasness of racism, sexism, and other discriminations. **b. Accountability** - This relates to the requirement to demonstrate compliance with relevant regulations and guidelines. **c. Privacy** - This relates to the explicit consent to be obtained from data providers. - Consent involving minors can be challenging, as this may require	**a. Fairness, b. Accountability, c. Privacy, d. Explainability** - Consensus around clear AI ethical principles are mixed and varied, as they stem from regulations (e.g., GDPR), laws (e.g., FERPA), standards (e.g., IEEE) and codes (e.g., Asilomar AI principles). It will be useful to establish a clear consensus set of principles. - There may exist gender-based differences in safety concerns of AI surveillance. - Trade-offs may exist in the application of principles when considered from the points of views of individuals, stakeholders or the

		both the students' and their parents' consents. **d. Explainability** - This relates to the right to receive explanations for automated decisions. - Challenges may arise from disclosure of proprietary algorithms or trade secrets.	society. As such, tensions may arise and there may be no set ways to manage these trade-offs.
Tlili et al. (2019)	a. Inclusivity b. Accountability c. Explainability d. Auditability e. Privacy f. Fairness	**a. Inclusivity** - This relates to inclusivity pertaining to disability. For example, an AI-based assessment system should account for longer time for reading for visually handicapped learners. **b. Accountability, c. Explainability, d. Auditability** - This relates to the understanding, validating, reviewing, and improving of AI algorithm applied, so that there are appropriate validity and transparency of algorithms. Challenges may arise if algorithms are proprietary.	**a. Inclusivity** - Intervention strategy of education triage, which balances the impact of intervention, with the scope of care needed, the resources available and the number of learners requiring care. **b. Accountability, c. Explainability , d. Auditability** - No clear guidelines exist on who should be the stakeholder responsible for understanding, validating, reviewing, and improving the AI system. It will be useful to clarify if this stakeholder should be the educators, system

		e. Privacy - This relates to the right to opt-in or out. The possibility to opt in or out may result in data gaps that can affect accuracy of results and research outcomes, indirectly isolate and reveal outcomes of those who opt in and affect discharge of institutional duty to enhance learning experience for students. **f. Fairness** - This relates to profiling-based discrimination arising from AI systems.	designers, or administrators, or if learners should also be involved in this process. **e. Privacy** - Specify details of the data sharing, data granularity, data retention period and multimodality of data records collected. - Understand potential problems and solutions relating to learner opting in or out. - Useful to clarify third party data sharing scenarios, e.g., institutions sharing with each other, external agencies, or companies to improve AI systems. **f. Fairness** - AI system to allow the possibility of intervention and corrective actions to enhance the automated process.
Kiennert et al. (2019)	a. Accuracy b. Privacy c. Academic integrity	**a. Accuracy** - This relates to the treatment of false positives and false negatives in cheating detection.	**a. Accuracy** - Allow human checks to override system outputs.

	b. Privacy - This relates to the management of sensitive data, such as authentication and biometric samples (e.g., data collected for password, voice recognition, facial recognition and/or keystroke detection).	**b. Privacy** - Pseudonymity to prevent linkability of sensitive data in an event of data leakage and exploitation. - Countermeasures in place to prevent the use of data for learner profiling which may impact learner well-being, e.g., discriminatively to infer likelihood of cheating. - Use of malleable signatures that allows a counterparty to modify the signed information, so that it becomes unfeasible to distinguish between the original signature and the sanitized signature. This retains the validity of the signature but ensures unlinkability of the sensitive data. - Allow sensitive data to be stored at decentralized dedicated entities, or Trusted Third Parties (TPP), for access and retrieval purposes. The focus of the design of AI systems for educational purposes are not aimed at guaranteeing treatment of sensitive data, but meant for e.g., proctoring an assessment.

			Establishing dedicated TPPs, alongside Public Key Infrastructure (PKI) and Certification Authorities (CA), may be useful to decentralize and improve privacy for data subjects.
Peña-Ayala (2018)	a. Privacy b. Fairness	**a. Privacy** - This relates to freedom from intrusion, biasness, or interference, relating to the collection, usage, and analysis of confidential data. **b. Fairness** - This relates to unintended labeling of students that can affect their learning journey and well-being.	**a. Privacy** - Ensure privacy preservation, including: (i) Data publishing and third party sharing of sensitive data, that do not leak sensitive data. (ii) Disclosure control and data mining, that do not undermine the identification of individuals tied to sensitive data. - Establish framework defining the following elements: (i) Stakeholders: Data subjects, data recipients and data curators (ii) Type of information: Sensitive attributes, quasi-identifiers, explicit identifiers and auxiliary information (iii) Data: Student demographics, educators, courses, assessments, course evaluations and disciplinary

			actions (iv) System architecture: Data access layer, data publishing, statistical disclosure control, differential privacy mechanism, and anonymizer mechanism. **b. Fairness** - Increase transparency of AI process to learners to improve trust. - Ensure consent is received for data collection at multiple levels and various interaction touchpoints with the AI system.
Tlili et al. (2018)	a. Accuracy b. Accountability c. Privacy d. Fairness	**a. Accuracy** - This relates to the possibility for students to modify their behavior and game the AI system, with the knowledge of being assessed. - This also relates to the right for data subjects to check and rectify data collected, so that data inputs are accurate. - Further, due to the distributed nature of online learning, data collected may be incomplete or erroneous. Poor data quality can	**a. Accuracy, b. Accountability, c. Privacy, d. Fairness** - Training of all stakeholders (e.g. learners, educators and administrative staff) can improve trust, and competence to design ethical considerations into the AI systems, or perform validation checks on ethical practices. - Establish a detailed and clear constitution for AI ethics, highlighting ethical guidelines in the AI process, and the duties and

	negatively impact AI-driven decisions. **b. Accountability** - This relates to the care when managing sensitive data. It was noted that between 2007 and 2011, there were 133 incidents linked to educational institutions unintendedly disclosing sensitive learner information (Stiles, 2012). Such incidents can lead to reputational, legal and/or financial liabilities. **c. Privacy** - This relates to educators' right to opt in or out of participation in AI-driven teaching evaluation. - This is also relating to issues pertaining to informed consent. For instance, for online courses offered worldwide, informed consent for use of data can be affected by the data protection regulations where the learner is domiciled. **d. Fairness**	rights of all stakeholders. - Ensure the data collected and used are accurate and clean, allowing stakeholders' access and rectification. - Promote proactive (not reactive) behavior when designing and managing AI systems, so as to ensure the safety of all stakeholders are accounted for.

	- This relates to profiling-based discrimination arising from AI systems. - This also relates to ad hoc implementation of AI systems by educators, without a standard code of practices and ethics, which may affect fairness.

Appendix III

Code	Likert Survey Questions
SDC-F	How important do you think it is for these AI assessment systems to avoid favoring or disadvantaging any group of students?
DSS-F	How important is it for all students' data to be treated fairly in these systems, such that there are no unintended labelling or profiling of learners?
ACR-F	How crucial do you think it is for AI to ensure fair treatment for every student when communicating assessment feedback with non-prejudice and non-favoritism?
AA-F	How critical is it to have an unbiased process when AI is used to administer assessments for cheating violations?
GE-F	How vital do you believe fairness is when AI evaluates different types of student responses without problems of subjectivity, contextualization and cultural-specificity ?
SDC-P	How crucial do you think protecting students' personal information is in the design of these AI systems?
DSS-P	How important is the responsibility of handling of personal data during AI-assisted assessments?
ACR-P	How essential do you find privacy measures in the construction and personalization of AI-assisted assessments?
AA-P	How significant do you find safeguarding privacy when AI systems are used to monitor assessments?
GE-P	How critical do you find maintaining data security and privacy during the AI grading process?
SDC-HC	How important is it that AI in assessments should be designed to avoid manipulating learner behaviors and emotions?
DSS-HC	How important do you think it is for AI assessments to maintain and enhance students' psychological wellbeing?

ACR-HC	In AI-assisted assessment construction, how crucial is it to avoid profiling and classifying students unfairly?
AA-HC	How significant is it for AI systems to ensure your wellbeing during the assessment process?
GE-HC	How vital do you find transparent and reversible processes in AI systems for grading and evaluation?
SDC-T	How important are clarity and consensus of purposes of designing AI systems in educational assessments?
DSS-T	How significant is your trust in AI systems to preserve your autonomy and control over your data?
ACR-T	When an AI-assessment is developed, how important is it to trust the assessment to respect your autonomy in learning?
AA-T	How critical do you find trust in AI systems for maintaining your autonomy and control during assessment?
GE-T	In the grading process, how crucial is it for you to trust AI systems for accurate and unbiased evaluation?
SDC-C	How important do you think it is for AI assessments to be designed to have robust mechanisms to detect dishonest behaviors?
DSS-C	How significant do you find the role of AI in maintaining integrity in assessment environments?
ACR-C	How important is it for AI tools to ensure the fairness and integrity of the assessment process during assessment construction and rollout?
AA-C	How critical do you believe is the role of AI in ensuring a cheating-free environment during exams?
GE-C	How vital is it for AI grading systems to have safeguards against the influences of cheating?
SDC-E	How crucial is transparency about how AI systems make decisions in the design phase?
DSS-E	How significant do you find the need for transparency in AI data management and surveillance processes?

ACR-E	How crucial do you think it is for AI tools to be transparent and understandable in their construction and rollout of assessments?
AA-E	How critical do you believe is the need for AI systems to clearly explain their actions and decisions during the administration, invigilation and monitoring of assessments?
GE-E	How vital do you find the aspect of explainability in ensuring fairness and transparency in AI-based grading?

SDC-A	How important do you think it is for AI assessments to appropriately handle prediction errors and biases, such that predicted AI decisions are accurate and cannot be gamed.
DSS-A	How crucial is the accurate understanding and interpretation of data collected, like student responses, in AI-driven assessments?
ACR-A	When AI tools are used to create and roll out assessments, how vital do you believe is their ability to make accurate and unbiased predictions about student performance?
AA-A	How crucial do you find the role of AI in ensuring the accuracy of administration, invigilation and monitoring of assessments, particularly in preventing false positives and negatives?
GE-A	How significant do you believe is the need for accuracy in AI systems to ensure fair and reliable grading?

SDC-AU	How crucial do you find the transparency and traceability of AI systems in the design phase for assessments?
DSS-AU	How significant do you believe is the role of auditability in ensuring responsible data stewardship in AI systems?
ACR-AU	How important do you think is the role of independent auditing in maintaining the integrity of AI-assisted assessment construction?
AA-AU	How critical do you find the need for transparency and the possibility of auditing in AI systems when AI detects security breaches and assessment violations?
GE-AU	How vital is it for AI-enabled grading to be transparent and subject to independent review for validity and reliability?

SDC-I	How crucial is it for AI in assessments to consider and respect cultural, gender, and socioeconomic diversity?
DSS-I	When AI systems monitor assessments, how important is it that they are sensitive to the diverse backgrounds and needs of students and do not reinforce stereotypes or discrimination?
ACR-I	In developing AI-based assessments, how important is it that AI generated decision should not induce and reinforce conformity, peer pressure and segregation that may negatively impact learners?"
AA-I	How critical is it for AI systems to avoid reinforcing any form of peer pressure or segregation during the administration, invigilation and monitoring of assessments?
GE-I	How vital is it that AI grading systems should not be impacted by prejudices, stereotypes, discrimination and biasness ?

SDC-AC	How crucial do you find the role of insitutions in ensuring responsible compliance with ethical guidelines and ethical handling of student data during the AI--enabled assessment design phase?
DSS-AC	How important do you think it is for AI systems to provide avenues for redress and correction in case of data misuse?
ACR-AC	How crucial do you think it is for institutions to be accountable for fair and non-discriminatory assessment practices of AI systems?
AA-AC	How critical do you find the need for availability of avenues for redress due to adverse decisions and actions from AI systems, in the process of administration, invigilation and monitoring of AI-enabled assessments?
GE-AC	How important is it for institutions utilizing AI grading systems to demonstrate responsibilities for the actions of decisions influenced by AI algorithmss and provide fair redress mechanisms?

LS1	Are you (or would you be) satisfied with the use of AI tools in your assessment experiences?
LS2	To what extent do AI tools in assessments met (or might meet) your expectations for fair and effective evaluation?

PLE1	Do you believe that AI-assisted assessments have contributed (or might contribute) positively to your understanding of the subjects and academic growth?
PLE2	How effective do you find (or would you) find AI tools in providing meaningful feedback on your assessments?
SAS1	How well do (or might) AI assessment tools provide you with academic support and guidance?
SAS2	Do you feel that AI tools in assessments adequately assist (or might adequately assist) you in identifying and overcoming learning challenges, enhancing your study and learning strategies?
PIP1	Do you feel that the use of AI tools in assessments negatively affects (or might negatively affect) your instructors' involvement in your learning process?
PIP2	To what extent do AI tools in assessments diminish the role of instructors in guiding and evaluating your performance?

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
AIED	Artificial Intelligence in Education
AR	Augmented Reality
AWE	Automated Writing Evaluation
CFI	Comparative Fit Index
CVI	Content Validity Index
HR	Horizon Report
I-CVI	Item-level Content Validity Index
ITS	Intelligent Tutoring System
IoT	Internet of Things
KMO	Kaiser-Meyer-Olkin
LMS	Learning Management System
MR	Mixed Reality
PLS	Personalized Learning System
RMSEA	Root Mean Square Error of Approximation
SRMR	Root Mean Square Residual
S-CVI	Scale-level Content Validity Index
SEM	Structural Equation Modeling
TPACK	Technological Pedagogical Content Knowledge
TLI	Tucker-Lewis Index
VR	Virtual Reality

List of Publications

Below is a record of papers presented or published by the author during the course of Doctor of Engineering that are related to the dissertation:

No.	Publication Title	Publication Venue	Publication Type	Publication Status	Citation Count (Google Scholar)	Reference
1	Authentic Assessments for Digital Education: Learning Technologies Shaping Assessment Practices	30th International Conference on Computers in Education (ICCE 2022) (Tier 2; SMU Whitelist)	Conference	Published	5	Lim, Gottipati, Cheong (2022)
2	Educational Technologies and Assessment Practices: Evolution and Emerging Research Gaps	Reshaping Learning with Next Generation Educational Technologies[1] (Book)	Book Chapter	Accepted; In Press.	-	-
3	Assessment Design for Digital Education: An Analytics-based Authentic Assessment Approach	2022 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE) (Tier 2; SMU Whitelist)	Conference	Published	1	Lim et al. (2022)
4	Analytics-enabled authentic assessment design approach for digital education	Education & Information Technologies (Tier 1; h5-index 91; Ranked #2 in Google Scholar Metrics in Education	Journal	Published	2	Lim et al. (2023)

		Technology; Ranked #1 in Google Scholar Metrics in Education)				
5	Artificial Intelligence in Today's Education Landscape: Understanding and Managing Ethical Issues for Educational Assessment	Singapore Rising Scholars Conference 2023	Conference	Presentation	N/A	-
6	Ethical Considerations for Artificial Intelligence in Educational Assessments	Creative AI Tools and Ethical Implications in Teaching and Learning (Book)	Book Chapter	Published	7	Lim, Gottipati, Cheong (2023)
				Total Citations	**15**	

[1] Book project link: *https://www.igi-global.com/book/reshaping-learning-next-generation-educational/327368.*