ASM Sc. J., 20(1), 2025

https://doi.org/10.32802/asmscj.2025.1917

Ethical Design of AI for Education and Learning Systems

W.Y. Leong[1*] and J.B. Zhang[2]

[1]*INTI International University, Malaysia*

[2]*National Technological University, Singapore*

The increasing integration of artificial intelligence (AI) in education presents both significant opportunities and critical ethical challenges. This paper explores the ethical design of AI for education and learning systems, focusing on key principles such as transparency, privacy, fairness, and accountability. AI technologies hold the potential to revolutionise personalised learning, assistive technologies, and administrative efficiency. However, issues such as bias, data privacy, and the potential reduction in human interaction require careful attention. Ethical AI systems in education should be designed to mitigate bias by using diverse and representative datasets, protect user privacy by securing sensitive student data, and ensure inclusivity by accommodating diverse learning needs, including those of students with disabilities. Additionally, transparency in AI processes is critical to fostering trust among students, educators, and parents. Continuous feedback loops, collaboration with stakeholders, and clear policies on the use of AI are also necessary to align AI tools with educational values and goals. The paper concludes by recommending best practices for ethically implementing AI in educational settings, emphasising the need for cross-disciplinary collaboration and ongoing evaluation to enhance the fairness, accountability, and inclusivity of AI-driven educational systems.

Keywords: AI in education; ethical AI; transparency; inclusivity; privacy; bias mitigation; personalised learning

I. INTRODUCTION

The integration of Artificial Intelligence (AI) in education offers transformative potential, from personalised learning experiences to administrative automation and adaptive assessments. However, alongside these benefits come significant ethical concerns. As AI systems are increasingly deployed, their design must adhere to ethical principles that prioritise transparency, fairness, privacy, and inclusivity to ensure positive educational outcomes and prevent harm (Table 1). This study explores the ethical dimensions of AI tools in education, providing an overview of the core challenges, risks, and mitigation strategies. AI-based educational technologies often analyse vast datasets to personalise learning paths and automate feedback, yet they may inadvertently introduce bias or reduce human interaction (Russell, 2020). With education being a key factor in societal development, the ethical design of AI is essential to ensure equity and trust in digital learning environments (European Commission, 2022; IEEE, 2024).

The research objective of Ethical AI Design in Education is transparency and accountability. The systems should clearly explain how AI decisions are made, allowing stakeholders to understand and trust outcomes (DataCamp, 2024). Based on privacy and data security, the sensitive student data must be protected through strong governance frameworks and informed consent (European Commission, 2022). The AI must be designed to support all learners, including those with disabilities or from underrepresented backgrounds, by using diverse training datasets (Litslink, 2023). The role of teachers should be complemented, not replaced, ensuring

*Corresponding email: waiyie@gmail.com

that technology enhances interpersonal connections in education (IEEE, 2024).

Table 1. AI Applications and Their Ethical Considerations

AI Application in Education	Example	Potential Ethical Issues
Personalised Learning Systems	Adaptive learning platforms	Bias in algorithms, lack of transparency
Automated Grading	AI grading essays and assessments	Loss of fairness, risks to academic integrity
AI-Powered Chatbots and Virtual Tutors	Chat-based learning assistants	Decreased human interaction, privacy concerns
Assistive Technologies	Speech-to-text, text-to-speech tools	Potential bias, accessibility challenges

Several challenges arise when designing and implementing ethical AI in education is on bias and discrimination. Algorithms may perpetuate biases, leading to discriminatory learning outcomes. AI systems may collect personal data without sufficient safeguards, compromising student privacy (Kay, 2018). In terms of over-reliance on Technology, students may become overly dependent on AI systems, weakening critical thinking and problem-solving skills. The digital divide will lead to unequal access to AI-based tools may widen the gap between privileged and disadvantaged learners (DataCamp, 2024).

II. LITERATURE REVIEW

The ethical design of AI for education has evolved alongside the development of artificial intelligence and its application across different sectors. This section outlines key milestones and developments that have shaped the conversation around ethics and AI in education, from the early conceptualisation of intelligent systems to the establishment of ethical frameworks in the 21st century.

AI research began in the mid-20th century, focusing on machine learning and problem-solving algorithms. In the 1960s, rule-based systems were introduced to offer simple tutoring solutions in education, like the PLATO system used for computer-based learning. However, these systems were limited in scope, and the notion of ethical considerations—such as bias or privacy—was not yet part of the discussion.

In the 1980s, Intelligent Tutoring Systems (ITS) emerged, which aimed to provide personalised learning experiences by adapting content to the learner's progress. Although these systems showed promise, concerns about equity and fairness were raised, as only a limited number of schools had access to this technology. Researchers began recognising the need for inclusive design to avoid widening the educational divide. This period also introduced discussions about accountability for system errors in student assessments.

The widespread availability of the internet and digital tools in the early 2000s led to a boom in e-learning systems and adaptive learning technologies. As AI became more embedded in educational platforms, researchers and policymakers began focusing on the ethical implications of these tools. In 2004, the IEEE launched working groups to develop guidelines for AI applications, including those in education.

Around the same time, universities and research institutions began publishing guidelines on privacy and fairness to protect student data and ensure that AI tools did not perpetuate bias. These guidelines emphasised transparency, calling for clear documentation of how AI models operate and influence learning outcomes.

In the 2010s, advances in machine learning and big data analytics allowed AI tools to become more sophisticated, supporting personalised learning paths and automated assessments. Technologies like chatbots, virtual assistants, and automated grading systems became popular in educational settings. However, these advances sparked debates about student privacy, academic integrity, and the potential over-reliance on AI tools, which could diminish critical thinking skills (DataCamp, 2024).

The European Union's General Data Protection Regulation (GDPR), enacted in 2016, was a landmark event in the discussion on AI ethics, requiring organisations to adopt stricter data governance policies (UNESCO, 2021). Educational institutions were prompted to reassess their use of AI tools, focusing on consent, transparency, and accountability.

By the 2020s, the adoption of AI in education had grown significantly, driven by advancements in natural language processing (NLP) and the demand for online education during the COVID-19 pandemic. The ethical use of AI became a major focus, with institutions publishing guidelines and policies to ensure responsible use of these technologies.

In 2022, the European Commission published its guidelines on ethical AI use in education, highlighting the need for inclusivity, privacy protection, and fairness (European Commission, 2022).

The IEEE released reports advocating for transparency and accountability in AI tools used in higher education, ensuring that students and educators could trust these technologies (IEEE, 2024).

Today, the conversation on ethical AI in education is more relevant than ever. With the rapid rise of Generative AI models—such as ChatGPT—educators are exploring new ways to incorporate AI ethically into the classroom (Gulson, 2020). However, concerns about bias, loss of human interaction, and the digital divide continue to dominate the discussion.

Efforts are now being made to develop global standards for ethical AI, focusing on collaborative design involving educators, developers, and policymakers. The future of AI in education depends on balancing technological innovation with ethical responsibility, ensuring that AI systems support equitable, inclusive, and privacy-compliant learning environments.

The history of ethical AI design in education reflects a gradual shift from early experimentation with intelligent systems to the development of comprehensive frameworks for responsible AI use. Each era has introduced new technologies and challenges, prompting the creation of guidelines that prioritise privacy, fairness, transparency, and inclusivity (Leong, 2024a). As AI continues to evolve, the ethical principles that guide its design will play a critical role in ensuring sustainable, student-centred learning systems.

The increasing adoption of artificial intelligence (AI) in education offers numerous benefits, including personalised learning experiences, enhanced administrative efficiency, and assistive technologies for students with disabilities (Leong, 2023). However, these advancements bring challenges that require a deliberate focus on ethical design. This literature review explores various studies and frameworks addressing the ethical dimensions of AI in education.

Key Themes in Ethical AI Design for Education is transparency and accountability (Table 2). Transparency is vital to maintaining trust among stakeholders. AI tools should openly disclose their data sources, algorithms, and decision-making processes (Holmes, 2016; IEEE, 2024). Schools and educational institutions must also establish accountability frameworks to address potential AI failures.

In terms of bias mitigation and inclusivity, studies highlight the risks of biased algorithms perpetuating stereotypes, negatively impacting marginalised communities (DataCamp, 2024). Ethical AI must use diverse datasets and actively address potential biases to ensure equitable learning outcomes (European Commission, 2022).

Privacy concerns are paramount when dealing with sensitive student data. Research emphasises the need for robust data governance frameworks to protect against misuse (IEEE, 2024).

While AI can streamline educational processes, over-reliance may weaken interpersonal relationships, impacting social learning and trust between students and teachers (Selwyn, 2019).

Table 2. Key Studies on Ethical AI in Education

Study	Focus	Key Findings
European Commission (2022)	Ethical AI Guidelines for Educators	Emphasises inclusivity and privacy as core principles.
DataCamp (2024)	Benefits and Risks of AI in Education	Highlights potential biases and privacy risks.
IEEE (2024)	AI in Higher Education	Advocates for transparency and accountability in AI use.

Figure 1 summarises the most common ethical challenges in AI for education and highlights recommended solutions from the literature. This horizontal bar graph highlights the severity of key ethical challenges—such as bias, privacy issues, loss of human interaction, and the digital divide—on a scale from 1 to 10. It also presents corresponding

solutions, such as bias mitigation and data governance, which address these challenges effectively. This visualisation emphasises the importance of balancing technological benefits with ethical considerations in education.

AI tools such as speech-to-text and text-to-speech systems provide significant benefits for students with disabilities (Holmes, 2019). These tools ensure accessibility and foster equitable participation. Research demonstrates that adaptive learning platforms enhance student engagement by tailoring content to individual learning styles, though they also raise concerns about privacy (European Commission, 2022). Organisations and institutions are beginning to publish ethical guidelines for AI use in education to address challenges and promote responsible adoption (IEEE, 2024).

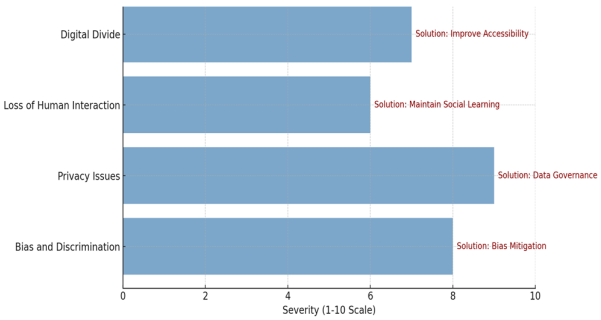

Figure 1. Common Ethical Challenges in AI for Education and Corresponding Solutions

III. METHODOLOGY

Figure 2 illustrates the conceptual framework for ethical AI design in education. It highlights the interaction between data input, algorithmic processing, and ethical oversight mechanisms. This framework outlines the flow of data from input to AI processing, ensuring transparency, fairness, and privacy through ethical oversight mechanisms. The system integrates continuous feedback from students to enhance personalised learning, maintain accountability, and support inclusive education.

Case studies from higher education institutions deploying AI-based grading systems report mixed outcomes (IEEE, 2024). While administrative efficiency improved, there were concerns about algorithmic fairness and academic integrity. Feedback from both students and teachers, emphasised the need for greater transparency and inclusivity in AI operations (European Commission, 2022).

Figure 2. Conceptual Framework for Ethical AI Design in Education

The introduction of AI in education is reshaping how learning is delivered and assessed. However, without careful attention to ethics, these technologies could exacerbate existing inequalities and introduce new challenges. Ethical design ensures AI not only enhances learning outcomes but also aligns with educational values such as equity, privacy, and fairness (Leong, 2024b). This paper advocates for a collaborative approach, involving developers, educators, and policymakers, to design AI systems that respect ethical norms and maximise educational benefits.

A. Case Study: Implementation of Ethical AI in a University

This section presents the methodology for implementing ethical AI design in educational settings, followed by a case study demonstrating the application of these principles. The methodology includes data collection strategies, system design frameworks, and ethical evaluation techniques, while the case study focuses on real-world AI implementation, challenges, and outcomes.

The proposed methodology (Table 3) follows a systematic approach to ensure that AI technologies in education are designed, implemented, and evaluated with ethical considerations in mind.

Table 3. Ethical AI Design Framework Steps

Step	Description
Requirement Analysis	Identify stakeholders and ethical requirements.
Data Collection and Governance	Ensure datasets are diverse, unbiased, and privacy-compliant.
Algorithm Development	Implement bias mitigation techniques and transparency measures.
Ethical Oversight Mechanism	Integrate continuous monitoring for ethical compliance.
Feedback Loop and Evaluation	Collect feedback from students, educators, and administrators.

Data Governance Policies: Secure management of student data through informed consent.

Diverse Datasets: Include data from varied socio-economic, demographic, and cultural backgrounds to mitigate bias.

Algorithmic Auditing: Regular checks to identify and eliminate bias in predictive models.

Inclusion Metrics: Monitor how well the system supports diverse learning styles and needs.

Dashboard for Educators: Provide transparency through a dashboard that tracks AI decisions and outcomes in real-time.

Student Surveys: Gather feedback to understand the perceived fairness and effectiveness of the AI system.

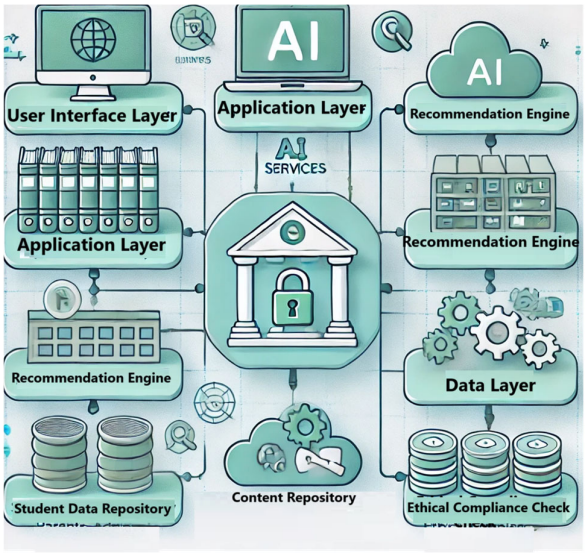

Figure 3. System Architecture for Ethical AI-Powered Learning Platform

This case study presents the deployment of an AI-based personalised learning platform in a large public university (Leong, 2024c). The platform aimed to enhance student engagement by adapting course content to individual learning preferences while also providing automated grading and feedback. The objective is to improve academic outcomes through personalised learning, reduce the workload on instructors through automated grading and ensure ethical compliance through transparency, inclusivity, and privacy measures.

Figure 3 is the system architecture diagram showcasing how the platform integrates data input, personalised learning algorithms, and ethical oversight mechanisms. This diagram outlines how data flows through the system—from input, processing by AI algorithms, and personalised learning interfaces, to ethical oversight mechanisms ensuring transparency, privacy, and fairness. Continuous feedback loops ensure ongoing improvements and alignment with ethical principles.

Table 4. Comparison of Metrics Before and After AI Implementation

Metric	Pre-AI Implementation	Post-AI Implementation
Student Engagement Rate	60%	85%
Instructor Grading Time	10 hours per week	4 hours per week
Perceived Fairness (Survey)	Not Measured	90% Positive Response
Bias Detection Instances	Not Applicable	3 Instances Detected and Corrected

Table 4 demonstrates that the AI system improved student engagement and reduced instructors' grading workload. However, the bias detection audit identified three instances of bias, which were addressed by updating the algorithm (Table 5).

Student were asked to use AI tool to assist in their writing and learning, Figure 4 shows the increase in student engagement rate before and after the AI platform was

implemented. The graph illustrates the engagement rate before and after the implementation of the AI-powered learning platform. It shows a steady increase in engagement after the AI deployment, highlighting the platform's positive impact on student involvement and learning outcomes over time.

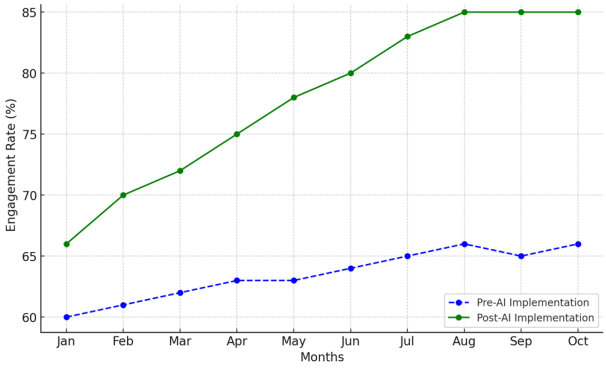

Figure 4. Engagement Rate Improvement Over Time

Table 5. Challenges Encountered and Corresponding Solutions

Challenge	Solution
Bias in Early Algorithm Models	Introduced diverse datasets and auditing tools.
Privacy Concerns	Implemented strict data governance policies.
Resistance from Educators	Provided training on AI tools and ethical use.

The feedback from students and instructors emphasised the importance of transparency and real-time monitoring. Students appreciated the personalised content but expressed concerns about relying too much on AI for grading. Instructors highlighted the need for clear policies on AI use to maintain academic integrity.

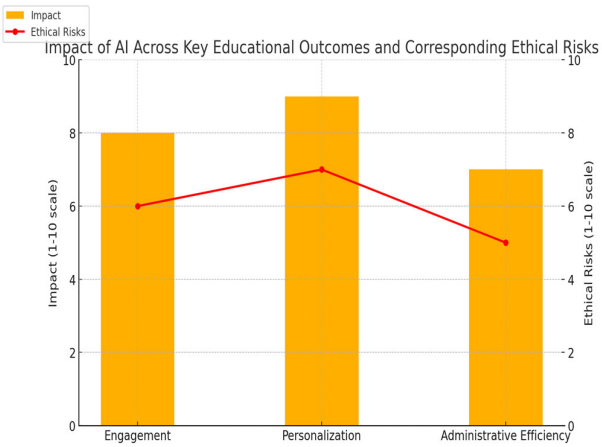

Figure 5. Impact of AI across key educational outcomes and corresponding ethical risks

Figure 5 shows AI's impact on several educational outcomes (engagement, personalisation, and administrative efficiency), along with identified ethical risks for each area. The bar chart shows the positive impact of AI on engagement, personalisation, and administrative efficiency, while the line graph overlaid highlights the corresponding ethical risks in each area. This demonstrates that while AI can significantly enhance educational outcomes, it also introduces risks—such as reduced privacy, algorithmic bias, and decreased human interaction—that need to be carefully managed for ethical AI implementation.

The implementation of an ethical AI-powered learning platform shows promising results in terms of student engagement and efficiency gains. However, the challenges encountered—such as bias and privacy issues—underscore the need for continuous monitoring, training, and ethical oversight. The platform's success demonstrates the value of combining technological innovation with ethical responsibility. This methodology and case study provide real-world insights into the design and deployment of ethical AI systems in education. The findings highlight the importance of feedback, transparency, and continuous improvement, offering a roadmap for institutions seeking to leverage AI responsibly in learning environments.

IV. CHALLENGES AND LIMITATIONS

While AI offers tremendous potential to transform education, implementing AI systems in an ethical and responsible manner presents various challenges. In Table 6,

based on the comments from students, the challenges of AI in education were recorded.

Table 6. Summary of Key Challenges in Ethical AI Design for Education

Challenge	Description
Bias and Discrimination	AI models may perpetuate stereotypes and biases from training data.
Privacy and Data Governance	AI systems often collect personal data, raising concerns about data security.
Loss of Human Interaction	Over-reliance on AI may reduce meaningful interactions between students and educators.
Resistance from Educators	Educators may resist adopting AI due to lack of trust or technical knowledge.
Digital Divide	Unequal access to AI technologies could worsen educational inequalities.

Figure 6 compares the severity of these challenges based on feedback from educational institutions using AI. The horizontal bar chart highlights the severity of key challenges, including bias and discrimination, privacy issues, loss of human interaction, resistance from educators, and the digital divide. Privacy issues are shown as the most severe challenge, underscoring the need for robust data governance frameworks, while bias and resistance also rank high, requiring focused interventions.

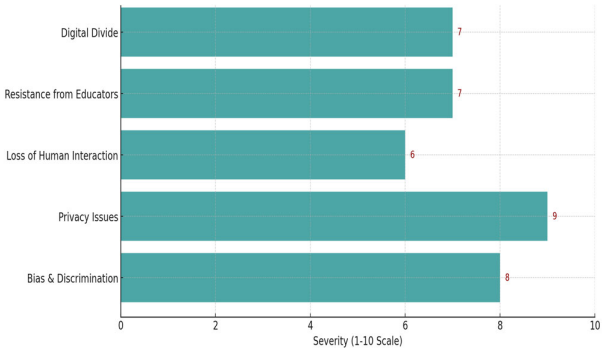

Figure 6. Severity of Challenges in AI Adoption in Education (Scale of 1-10)

Table 7. Challenges and Mitigations

Bias and Discrimination	
Problem	Algorithms trained on biased data can perpetuate inequality.
Solution	Use diverse datasets and perform regular audits to identify bias (Floridi, 2019).
Privacy and Data Governance	
Problem	AI systems collect large amounts of student data, which can be misused.
Solution	Implement privacy-by-design principles and enforce robust data governance policies (European Commission, 2022)
Loss of Human Interaction	
Problem	Students may engage more with AI than with teachers or peers, weakening interpersonal skills (DataCamp, 2024).
Solution	Balance AI tools with human-led activities to ensure social learning continues.

Figure 7 is a conceptual diagram illustrating the relationship between key challenges and possible solutions. This diagram shows how key challenges—such as bias, privacy issues, loss of human interaction, resistance from educators, and the digital divide—can be addressed through specific mitigation strategies, including bias audits, data governance policies, and hybrid learning models. This visualisation highlights the interconnectedness of challenges and solutions, emphasising a structured approach to ethical AI implementation.

Figure 7. Challenges and Mitigations in Ethical AI Design for Education

In Table 8, the proposed solutions for institutions to tackle AI challenges in education are highlighted. Bias in AI algorithms could be minimised by implementing Diverse Data Collection Protocols to collect data from various demographic, cultural, and regional backgrounds to ensure representativeness. Regular bias audits and conduct periodic reviews of AI algorithms for bias and discriminatory patterns. A more inclusive AI teams could encourage diversity in the teams developing and deploying AI systems. Adopt strong data governance policies, establish clear policies on data usage, access, and sharing, aligned with legal frameworks like GDPR. Use privacy-enhancing technologies, employ methods like anonymisation, differential privacy, and secure multi-party computation.

Ensure students and parents are informed about data collection and have the option to opt-out.

Table 8. Proposed solutions for Institutions to AI Challenges

Key Challenge	Solution Implemented	Result
Bias in Algorithm	Diverse datasets and bias audits	Reduction in biased outcomes
Privacy Concerns	Data governance frameworks	Improved data protection policies
Loss of Human Interaction	Hybrid learning model	Enhanced student-teacher interaction

A survey conducted among 10 universities implementing AI tools revealed the following insights. Figure 8 depicts feedback from institutions regarding their challenges and solutions after AI adoption. This bar chart summarises the feedback from educational institutions that adopted AI, highlighting key insights:

60% reported a positive impact on student engagement.

30% identified privacy concerns as a major challenge.

10% resolved bias issues after implementation.

25% overcame initial resistance from educators through training and support.

The chart underscores both the benefits and challenges of adopting AI in education, emphasising the need for continuous monitoring and support.

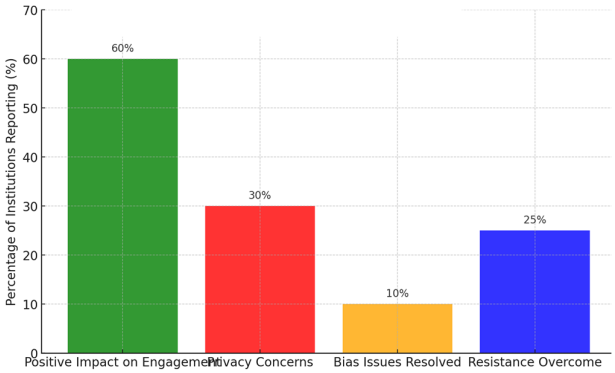

Figure 8. Feedback on AI Adoption Challenges and Solutions

The challenges of implementing AI ethically in education require comprehensive solutions that address bias, privacy, and accessibility (Zhang, 2024). Balancing AI tools with human interaction is essential to maintain a healthy learning environment. Overcoming these challenges demands collaboration among developers, educators, and policymakers to design and deploy systems responsibly (World Economic Forum, 2022).

V. CONCLUSIONS

The ethical design of AI for education and learning systems is essential for ensuring that technological advancements align with the core values of fairness, transparency, privacy, and inclusivity. As AI continues to transform educational practices through personalised learning, automated assessments, and adaptive systems, institutions must adopt a balanced approach that integrates ethical oversight mechanisms at every stage.

AI systems must utilise diverse datasets and conduct regular algorithm audits to avoid perpetuating bias. Inclusive design ensures that educational tools cater to the needs of all students, including those from underrepresented or marginalised groups.

Transparency in AI processes is essential to build trust among educators, students, and parents. Clear documentation of how AI algorithms function and influence decision-making is crucial for maintaining accountability. Implementing real-time dashboards provides users with insights into AI operations.

Robust data governance frameworks must be implemented to secure student information. Informed consent policies are

essential to ensure students and parents understand how data is used and stored. Adhering to data protection laws, such as GDPR, further reinforces privacy protections.

AI tools should complement, not replace, human interactions. While AI can improve efficiency, educators play a vital role in fostering critical thinking, social skills, and emotional intelligence. Hybrid learning models, which combine AI tools with traditional teaching methods, offer a promising way to maintain this balance.

The ethical design of AI is not a one-time effort but requires continuous monitoring, feedback, and improvements. Stakeholders, including educators, students, and parents, must be involved in evaluating the effectiveness and fairness of AI systems over time.

In conclusion, the responsible deployment of AI in education depends on a holistic framework that integrates technological innovation with ethical principles. Collaboration between developers, educators, and policymakers is crucial to creating systems that enhance learning outcomes without compromising ethical standards. By focusing on transparency, accountability, privacy, and inclusivity, educational institutions can leverage AI effectively to support student-centred, equitable, and sustainable learning environments. This paper underscores the importance of continuous adaptation to emerging challenges, such as bias, privacy concerns, and digital inequality. Future research should explore new strategies and frameworks to address these issues while advancing AI capabilities to support the evolving needs of learners.

VI. REFERENCES

DataCamp 2024, AI in Education: Benefits, Challenges, and Ethical Considerations.

European Commission 2022, Ethical Guidelines on AI and Data in Education.

Floridi, L 2019, Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Springer.

Gulson, KN & Witzenberger, K 2020, 'The Ethics of AI in Education: Who Is Accountable?', Critical Studies in Education, vol. 61, no. 4, pp. 405–421.

Holmes, W, Bialik, M & Fadel, C 2019, Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. OECD Education Working Papers.

IEEE 2024, Transformative Potentials and Ethical Considerations of AI Tools in Higher Education.

Kay, J, & Luckin, R 2018, 'Rethinking Learning Systems with Artificial Intelligence: Ethical Challenges and Opportunities', Proceedings of the ACM Conference on Learning Analytics and Knowledge (LAK).

Leong, WY, Leong, YZ & Leong, WS 2023, 'Virtual reality in education: case studies and applications', IET Conference Proceedings, pp. 186-187. DOI: 10.1049/icp.2023.3332.

Leong, WY 2024a, 'Enhancing Practical Skills Training Through Virtual Reality in TVET Education', 2024 International Conference on TVET Excellence & Development (ICTeD), Malaysia, 16-17 Dec 2024.

Leong, WY 2024b, 'Fostering Creative Thinking Through Immersive Virtual Reality Environments in Education, Educational Innovations and Emerging Technologies', vol. 4, no. 3, pp.8-25.

Leong, WY & Zhang, JB 2024, 'Revolution in University Ranking System', ASM Science Journal, vol. 19.

Luckin, R, Holmes, W, Griffiths, M & Forcier, LB 2016, Intelligence Unleashed: An Argument for AI in Education. Pearson.

Russell, S & Norvig, P 2020, Artificial Intelligence: A Modern Approach (4th Edition). Pearson.

Selwyn, N 2019, 'Should Robots Replace Teachers? AI and the Future of Education', Polity Press.

UNESCO 2021, AI in Education: A Guide for Policymakers.

World Economic Forum 2022, Ethics by Design in Artificial Intelligence: Recommendations for Policymakers in Education.

Zhang, HL & Leong, WY 2024, 'Transforming Rural and Underserved Schools with AI-Powered Education Solutions', ASM Science Journal, vol. 19.