

Beyond Code: Elucidating A Human-Centric Approach to Ai Ethics

Zaheril Zainudin¹, Syaza Hazirah Mohd Hashim¹, Nor Asiah Ismail¹, Haslizah Jibin¹, Nurkhidayah Baharuddin^{1*}

¹Faculty of Education and Liberal Studies, City University Malaysia, Petaling Jaya, 46100, MALAYSIA

*Corresponding author email: dr.zaheril@city.edu.my

Received 29 Jul 2025; Accepted 18 Aug 2025; Available online 20 Sep 2025

Abstract: This study examines the necessity of adopting a human-centric approach to artificial intelligence (AI) ethics, emphasizing the integration of empathy, dignity, and moral accountability within AI-driven educational ecosystems. Using a mixed-methods design, the research collected quantitative survey data from 80 Malaysian primary school teachers and qualitative interview insights from seven participants. Findings reveal that teachers generally perceive AI as a beneficial tool that enhances efficiency and learning experiences, yet they express concern about ethical issues such as privacy, transparency, and algorithmic bias. The participants strongly advocate for AI policies and systems that prioritize human values and civilizational continuity. Thematic analysis highlights six core themes: AI as a practical assistant, human–AI collaboration, safeguarding human intellect, societal responsibility, ethical safeguards, and fostering creativity. The study concludes that ethical AI integration must move beyond technical compliance toward relational accountability, inclusive governance, and pedagogical empowerment. It recommends developing national ethical frameworks, professional AI literacy programs, and participatory design models that ensure technology remains aligned with human welfare and educational integrity. Finally, this paper contributes to global discourse by proposing a holistic framework for embedding human values in AI ethics

Keywords: Artificial intelligence ethics, human-centric design, educational technology, teacher perspectives, ethical governance

1. Introduction

1.1 Background of Study

In the modern era, technology has become an important part of our daily lives, shaping not only our communication (Rogers, 1986) but also our understanding of the ethical values (Namukasa, 2025) that form the basis of civilization (Misra, 2022). In fact, experts and thinkers agree that humans are the most important element for the survival of any civilization (Alegre, 2024). The formation of a civilization is closely tied to human development (Matcas, 2016), and the efforts of a society to build its civilization contribute to its progress (Mitra, 2025). Many scholars view artificial intelligence (AI)’s potential as a major asset for societal development (UNDP, 2025) and for achieving the goals of a developed nation (Brandao, 2025). Indeed, the high level of public trust in AI’s ability to enhance life is largely driven by its capacity to simplify daily tasks and provide greater convenience (McIntyre et al., 2025). Thus, the rapid progress of modern technology, particularly in AI systems, has brought both significant advantages (Tai, 2020) and challenges (Sakubu, 2025) to humanity.

*Corresponding author: dr.zaheril@city.edu.my

<https://jthkss.com/> All right reserved.

The advancement of modern technology, especially in AI, has a major impact on how society defines and determines its culture (UNESCO, 2025). The rapid advancement of AI has made it a central component of modern life, influencing everything from communication and education to healthcare and administration (Sari and Purwanta, 2021). Nevertheless, this study believes that the connection between AI technology and human-centric approach has also changed our understanding of the balance between the utilization of technology and human civilization. With the current progress of AI, it is possible to create a balance between using AI and preserving core human ethical values and principles (Fan, 2024). This balance allows us to leverage AI technology's capabilities while promoting concepts such as compassion, empathy, and respect for human dignity (Kumar and Rangoli, 2025). The link between technology and human values is very important in the age of AI (Alalaq, 2025). Indeed, this technological advancement is seen as a significant force for improving the quality of human life (Alsaleh, 2024), offering the promise of more precise and rapid outcomes compared to traditional methods (Vernyuy, 2024).

Despite several promising benefits, however, the unguided implementation of AI presents significant ethical challenges that can negatively impact human civilization (Mehra, 2025). The core problem lies in the potential for AI to undermine fundamental human values if not developed with a human-centric approach in mind (Chong et al., 2025). Without careful consideration, AI systems could exacerbate inequalities or infringe upon personal rights (Sabah et al., 2025). Therefore, the rise of AI has sparked a crucial conversation about the need for a delicate balance between leveraging this technology and preserving core human principles, such as empathy, compassion, and personal dignity (Nikitenko et al., 2025). In the modern era, eventually, the link between technology particularly AI and civilizational values is more important than ever.

Thus, the purpose of this paper is to elucidate a human-centered approach to AI ethics. The primary goal is to investigate whether human values might be efficiently incorporated into AI systems to build an equitable and ethical society. The research problem revolves around the urgent demand for an equilibrium strategy: one that properly utilizes AI's enormous capability while also upholding and enriching human well-being and cultural norms. The research paper addresses the issues of individuals who are skeptical of AI's ability to incorporate values, suggesting that prioritizing ethics in the design stage is critical for establishing trust in society as well as guaranteeing that technology benefits mankind. Hence, to achieve this aim, the paper addresses the following research questions:

- a. What are the teachers' perspectives on the potential benefits of AI?
- b. What are the teachers' perspectives on ethical and societal implications of AI?
- c. What are the teachers' perspectives on AI and human values?
- d. What recommendations can be proposed to support the human-centric approach to AI ethics?

1.2 Literature Review

Fundamentally, the ethical issues surrounding AI have developed across advancements in technology, progressing from theoretical discourse to practical, real-life dilemmas (Annet, 2025). The philosophical origins of AI ethics could be linked back to theorists which debated the origins of intellect and awareness, as well as the moral consequences of developing sentient beings (Akova, 2023). However, as AI becomes more integrated into everyday usage, attention has turned from hypothetical concerns to concrete ethical frameworks (Lee et al., 2025). Scholars such as Creswell (2014) in quantitative research design, as well as others who have documented the rise of AI's influence on various sectors, have provided the methodological and contextual foundation for such debates (Sari and Purwanta, 2021).

Current AI ethics frameworks often center on a set of core principles designed to guide the development and deployment of AI systems (Papagiannidis et al., 2025). These principles typically include fairness (Radanliev, 2025), ensuring that AI decisions do not perpetuate or amplify societal biases (Ferrara, 2023); accountability, establishing who is responsible for an AI system's actions; and transparency (Cheong, 2024), making AI's decision-making processes understandable to humans (Jowader, 2025). Organizations like IBM and UNESCO have contributed significantly to these frameworks, emphasizing the importance of ethical guidelines in AI development (Gade, 2025). IBM's approach to AI ethics focuses on three pillars: the purpose of AI, the data used to train it, and the process of its development (Kirova et al., 2025), while UNESCO (2022) has provided a comprehensive recommendation on the ethics of AI, highlighting the need for global cooperation.

Despite these efforts, a critical gap exists in many current approaches. A significant portion of the literature and current frameworks tend to be technology-focused, prioritizing the technical aspects of AI rather than its impact on human well-being (Yuxuan and Wan Hussain, 2025). They often focus on how to make the AI system itself more "ethical" (reducing algorithmic bias) without deeply exploring how technology interacts with and affects human values on a broader societal level (Hanna et al., 2025). This narrow perspective often fails to fully address public concerns about the moral and safety implications of AI (Nawi, 2020). Research has shown that a lack of public understanding about AI can lead to uncertainty and apprehension regarding its potential risks (Brauner et al., 2023).

This gap highlights the need for a human-centric approach to AI ethics. While current frameworks are necessary, they are not sufficient. A human-centric model goes beyond technical principles to consider the preservation and enhancement of human dignity, compassion, and civilizational values (Gilbert and Gilbert, 2024). It recognizes that humans, as the primary actors, must be the ones to influence and shape the future of AI (Kirk et al., 2025). This approach acknowledges that while AI can make future predictions based on existing data, the human intellect remains crucial for

evaluating and refining these predictions (Akinagbe, 2024). The literature suggests that for AI to truly be a tool for societal enrichment, it must be developed with a strong foundation in moral and ethical values determined by society itself (Wiese et al., 2025). This is because the capability to perfect creations is more confidently entrusted to humans.

In the end, an analysis of the existing research indicates a shift from conceptual debates to principle-based frameworks. Yet most contemporary frameworks are largely technology-focused, leaving a significant gap in the subject matter. This study attempts to address the gap by proposing explaining a human-centered approach to AI ethics, one which emphasizes the incorporation of essential human values into all aspects of AI development and application. This strategy is vital for establishing public trust, assuring responsible AI utilization, and developing an environment wherein AI really benefits mankind.

2. Methodology

This study employs mixed methods to investigate the primary school teachers' perspectives between AI and human-centric values. This approach, as defined by Creswell (2014), prioritizes the collection and analysis of numerical data to understand an occurrence in real setting. The primary objective is to determine and explain patterns and connections within the data, specifically focusing on the extent of AI's influence on human-centric principles. The study's design is explanatory sequential design by collecting and analyzing quantitative data first and then using qualitative data to help explain the quantitative results (Creswell and Clark, 2017). At the same time, this study gathers data at a single point in time to capture a snapshot of public opinion and perceptions (Cohen et al., 2007).

The data collection process was conducted using a Google Form survey, a method chosen for its efficiency in gathering data from a dispersed population and its suitability for systematic analysis (Lochmiller and Lester, 2015). The survey consisted of six specific questions designed to assess teachers' perspectives and potential benefits of AI, its ethical and societal implications of AI, besides its integration with human-centric values. The questions were carefully put on aligning with the research objectives, ensuring that the collected data directly addressed the key questions of the study.

The quantitative sample for survey consisted of 80 respondents, selected using simple random sampling (probability) to provide a balanced and insightful perspective. While the qualitative participant for interview comprised 4 men and 3 women using purposive sampling (non-probability), a ratio intended to ensure balance and a comprehensive understanding of feedback from participants (Creswell, 2014). The target population was the public-school teacher, with a specific focus on exposure to AI technology in their daily teaching task. The majority of respondents were aged 30 to 40 (45 individuals), with smaller numbers from 40-50 (30 individuals) and 50-and-above (5 individuals) age groups. This age distribution was designed to capture a range of perspectives, from those who grew up with AI to those who have witnessed its more recent integration into society. The purposeful selection of this sample aligns with similar studies that examine teacher perception of AI and its societal impact (Razak and Norman, 2025; Sivanganam et al., 2025).

The collected data was processed using descriptive statistical analysis and thematic analysis. This involved calculating percentages, frequencies, mean and standard deviation to summarize the responses to each survey question beside creating themes based on the research questions. This form of analysis allowed for a clear and direct interpretation of the data, providing a solid foundation for the subsequent discussion and conclusions.

2.1 What are the teachers' perspectives on the potential benefits of AI?

Table 1. Teachers' Perspectives on the Potential Benefits of AI

Statement	Frequency	Percentage	Mean	SD
AI can significantly improve the quality of human life.	61	76.3	3.94	0.89
AI helps me complete tasks more efficiently.	65	81.3	4.02	0.81
AI simplifies my daily tasks and routines.	57	71.3	3.79	0.93
AI's solutions are more accurate than traditional methods.	47	58.8	3.54	1.01
Benefits of AI outweigh its potential risks.	42	52.5	3.51	1.06
AI provides convenient access to information and services.	64	80.0	4.01	0.83
AI can positively impact education.	68	85.0	4.07	0.84

The descriptive analysis of teachers' responses reveals a generally positive orientation toward the potential benefits of AI in educational and societal contexts. Table 1 presents the distribution of responses across seven key benefit statements, with mean scores ranging from 3.51 to 4.07, indicating moderate to high levels of agreement. The highest-rated item is "AI can positively impact education" ($M = 4.07$, $SD = 0.84$) that underscores teachers' recognition of AI's strategic relevance in domains closely aligned with their professional values (Muhammed and Kamsin, 2025). Similarly, the strong endorsement of AI's efficiency ($M = 4.02$) and informational accessibility ($M = 4.01$) suggests that teachers perceive AI as a facilitator of both pedagogical and administrative tasks (Xue et al., 2025).

Conversely, the relatively lower mean scores for statements concerning AI's accuracy ($M = 3.54$) and risk-benefit balance ($M = 3.51$) indicate a nuanced stance. While a majority still agreed with these statements, the elevated standard deviations ($SD = 1.01$ and 1.06 , respectively) point to greater variability in perceptions that possibly reflect ethical concerns, trust issues, or limited exposure to AI-integrated systems (Govindarajoo et al., 2025). Largely, the data suggest that Malaysian secondary school teachers exhibit a cautiously optimistic view of AI (Qil et al., 2025), valuing its practical benefits while remaining mindful of its limitations (Zulkarnain and Yunus, 2023). These findings align with global trends in educational technology adoption, where enthusiasm is tempered by calls for critical engagement and professional development (Panjani, 2024).

2.2 What are the teachers' perspectives on ethical and societal implications of AI?

Table 2. Teachers' Perspectives on Ethical and Societal Implications of AI

Statement	Frequency	Percentage	Mean	SD
Concerned about AI violating personal privacy	56	70.0	3.80	0.96
AI may deepen societal inequality	51	63.8	3.65	1.00
AI systems should be transparent and understandable	68	85.1	4.21	0.81
Possible to integrate human civilization values into AI	59	73.8	3.89	0.89
Unsure about AI's future due to ethical concerns	54	67.6	3.73	0.94
Developers/authorities have provided sufficient risk information	28	35.0	3.02	1.17
Concerned AI lacks moral/ethical considerations for responsible use	59	73.8	3.86	0.91

The data in Table 2 reveal a nuanced and ethically engaged stance among Malaysian secondary school teachers regarding the societal ramifications of AI. While respondents generally acknowledge the transformative potential of AI, their views are tempered by concerns over transparency, equity, and moral accountability (Zhang et al., 2025). The highest-rated item is "AI systems should be transparent and understandable" ($M = 4.21$, $SD = 0.81$) that demonstrates a strong normative expectation among teachers for ethical accountability and system explainability (Govindarajoo et al., 2025). This aligns with global calls for interpretable AI, particularly in high-stakes domains such education (Davidson, 2024). Concerns about privacy ($M = 3.80$), moral accountability ($M = 3.86$), and ethical uncertainty ($M = 3.73$) further suggest that teachers are not merely passive adopters of technology but active evaluators of its societal implications (Dilek and Baran, 2025). The relatively high standard deviations across these areas indicate diverse perspectives, possibly shaped by varying levels of exposure, digital literacy, and institutional support (Todino, 2025). Notably, the lowest-rated item is "Developers/authorities have provided sufficient risk information" ($M = 3.02$, $SD = 1.17$) that points to a perceived gap in communication and governance. This finding highlights the need for more robust policy frameworks, stakeholder engagement, and professional development initiatives to bridge the trust deficit between teachers and AI developers (Hohma and Lutge, 2023). In sum, the data reflects a critical yet constructive stance: teachers recognize the ethical complexities of AI and advocate for systems that are transparent, inclusive, and aligned with human values (Fu and Weng, 2024). These insights are vital for informing AI integration strategies in education are not only technically sound but socially responsible (Vias et al., 2025).

2.3 What are the teachers' perspectives on AI and human values?

Table 3. Teachers' Perspectives on AI and Human Values

Statement	Frequency	Percentage	Mean	SD
Balance between AI and human values can be achieved	66	82.6	4.09	0.81
AI can serve humanity without compromising empathy and compassion	60	75.0	3.91	0.89
Highlighting human nature ensures harmonious coexistence with AI	63	78.8	4.00	0.85
AI and human values are mutually exclusive and cannot coexist	21	26.3	2.65	1.21
Future AI policies should prioritize human values	70	87.5	4.28	0.74
Trust that developers will prioritize human well-being and civilizational values	47	58.8	3.55	1.05
Optimistic about AI's future to serve humanity	65	81.3	4.09	0.84

The data presented in Table 3 reflect a predominantly optimistic and value-conscious orientation among Malaysian primary school teachers regarding the integration of AI with humanistic principles. The responses suggest that teachers

perceive AI not as a threat to human values, but as a tool whose ethical alignment depends on intentional design and policy stewardship (Kucukuncular and Ertugan, 2025). The highest-rated item is “Future AI policies should prioritize human values” ($M = 4.28$, $SD = 0.74$) that signals a clear normative expectation among teachers for governance frameworks that safeguard empathy, dignity, and civilizational continuity (Georgiadis and Arvanitidou, 2024). This finding aligns with global discourses on responsible AI, where human-centricity is increasingly viewed as a prerequisite for sustainable innovation (Sigfrids et al., 2023). Teachers also express confidence in the possibility of harmonizing AI with compassion and empathy ($M = 3.91$), and in the role of human nature as a guiding principle for coexistence ($M = 4.00$). These perspectives suggest that teachers envision AI not merely as a technological advancement, but as a sociotechnical system that must be anchored in ethical reflection and cultural sensitivity (Promsiri, 2025). Interestingly, the statement “AI and human values are mutually exclusive and cannot coexist” received the lowest mean score ($M = 2.65$, $SD = 1.21$), indicating broad disagreement with the notion of inherent incompatibility. This rejection of binary thinking reflects a more integrative and hopeful stance, where AI is seen as capable of complementing rather than displacing human value (George, 2024). However, the moderate score for trust in developers ($M = 3.55$, $SD = 1.05$) reveals a degree of skepticism regarding the intentions and accountability of AI stakeholders. This ambivalence underscores the need for transparent, participatory, and ethically grounded development processes that actively involve teachers and other end-users (Harjatanaya et al., 2025).

2.4 What recommendations can be proposed to support the human-centric approach to AI ethics?

Table 4: Thematic Analysis of Human-Centric Approach to AI Ethics

Theme	Subtheme	Coded Excerpts
1. AI as a Practical Assistant in Daily Tasks	Efficiency in Teaching Tasks	“I used ChatGPT to draft lesson plans it saved me hours.” “AI helped me summarize student feedback quickly.” “It made me feel more in control of my workload.” “I felt less stressed during exam season.”
	Emotional Impact	“AI gave me a rough draft, but I had to refine it to suit my students.”
2. Human-AI Collaboration in Idea Generation	AI as a Spark, not a Solution	“It’s like having a brainstorming partner.” “AI lacks context only I know what works in my classroom.”
	Human Judgment as Essential	“It’s useful, but not final.” “AI should support not replace teachers.”
3. Safeguarding Human Intellect	Augmentation over Replacement	“We need to teach students how to think, not just use tools.”
	Curriculum Integration	“Critical thinking must be embedded in AI literacy.” “We should design tasks that require human insight.”
4. Societal Responsibility in AI Ethics	Cultural Embedding	“AI must reflect our values, not just global norms.” “Local educators should be part of AI design.”
	Policy and Oversight	“We need clear guidelines on ethical use.” “Ministries should regulate how AI is used in schools.”
5. Ethical Concerns and Safeguards	Privacy and Surveillance	“I worry about student data being misused.” “Who controls the information AI collects?”
	Bias and Fairness	“AI might reinforce stereotypes.” “We need transparency in how decisions are made.”
6. Fostering Creativity and Critical Thinking	Pedagogical Strategies	“Let students critique AI outputs.” “Use AI to provoke deeper questions, not just answers.”
	Empowerment through Inquiry	“Creativity is our edge; AI can’t replicate that.” “We must teach students to challenge AI.”

Thematic analysis of interview data yielded six interrelated themes that illuminate Malaysian primary school teachers’ nuanced perspectives on AI in education. These themes reflect a balance between pragmatic engagement, ethical reflection, and pedagogical intentionality as follows:

Theme 1: AI as a Practical Assistant in Daily Tasks - Teachers consistently described AI as a time-saving tool that enhances task efficiency, particularly in lesson planning and feedback synthesis. The emotional relief associated with AI utilization such as reduced stress and increased control. This suggests that AI is perceived not merely as a technical aid but as a contributor to professional well-being (Kucukuncular and Ertugan, 2025).

Theme 2: Human-AI Collaboration in Idea Generation - Teachers framed AI as a generative partner rather than a final authority. While AI was valued for its capacity to spark ideas, teachers emphasized the indispensability of human judgment in contextualizing and refining outputs. This reflects a collaborative epistemology where AI augments but does not replace teacher tasks (Flavins et al., 2025).

Theme 3: Safeguarding Human Intellect - Participants expressed strong convictions about preserving human cognitive primacy. AI was seen as a tool for augmentation, not substitution, with calls to embed critical thinking into AI literacy curriculum. This theme underscores the pedagogical imperative to cultivate discernment and intellectual autonomy in students (Ng, 2025).

Theme 4: Societal Responsibility in AI Ethics - Teachers advocated culturally grounded AI design and policy oversight. The emphasis on local educator involvement and ethical regulation reflects a desire for participatory governance and context-sensitive implementation. This theme aligns with global discourses on inclusive and equitable AI development (Roche et al., 2021).

Theme 5: Ethical Concerns and Safeguards - Concerns about privacy, surveillance, and algorithmic bias were prevalent. Teachers questioned data governance and emphasized the need for transparency in AI decision-making. These reflections highlight the ethical tensions inherent in educational AI adoption and the need for robust safeguards (Dave, 2025).

Theme 6: Fostering Creativity and Critical Thinking - AI was seen as a pedagogical provocation rather than a solution. Teachers proposed strategies that use AI to stimulate inquiry, critique, and creative expression. This theme positions AI as a catalyst for deeper learning, reinforcing the irreplaceable role of human imagination and critical engagement (Diaz Noguera, 2024).

Eventually, these underlying perspectives represent teachers as contemplative practitioners who approach AI by means of the lenses of ethical pragmatist as well as pedagogical motivation. Their personal experiences argue for an equilibrium implementation of AI as one that improves teaching while not jeopardizing human values, ethical standards, or autonomy for students.

In summary, the findings show teachers as ethically involved agents which demand AI systems that are not merely practical yet ethically compatible. Their perspectives provide important insights into developing AI policy that respect humanity, create trust, along with foster inclusive innovations in education as well. The results of this study indicate the vital necessity for a human-centered approach to AI ethics. The facts, specifically the significant degree of general acceptance of an equitable human-AI future, elevates this notion from a theoretical aspiration to a practical necessity for fostering public trust as well as preserving social well-being. The discussion explains such results and implications for diverse stakeholders and subsequently proposes a framework for real-world implementation.

2.4 Data Analysis

The reviewed literature presents a multifaceted analysis of the transformation of traditional newspaper groups in the context of media convergence. A thematic approach was used to identify recurring patterns and insights, which were categorized into four main dimensions: content innovation, technological integration, organizational restructuring, and policy support.

First, content innovation has emerged as a key factor in ensuring audience engagement and relevance in the digital age. Second, the integration of advanced technologies such as artificial intelligence, big data, and cloud computing is a recurring theme. These technologies are used to optimize content production, enable precise audience targeting, and improve content distribution efficiency. Third, organizational restructuring is highlighted as a major challenge and a critical component of successful transformation. Resistance to change, lack of technical expertise, and rigid hierarchical structures were identified as significant barriers to innovation. Finally, policy support has played a pivotal role in facilitating the transformation of traditional newspaper groups. Government subsidies, regulatory frameworks, and infrastructure support have been instrumental in promoting the adoption of new technologies and the establishment of multi-platform ecosystems within media groups.

3. Results

Table 3 presents the findings from 21 studies, categorized into key areas: content innovation, technological integration, organizational restructuring, and policy support. Each study emphasizes the profound impact of media convergence on traditional newspaper groups, capturing both the potential opportunities and limitations.

In terms of content innovation (n=10), the studies highlight how traditional newspaper groups strengthen content innovation, enrich content formats, and enhance their communication and influence, while optimizing content production and user experience. Regarding technological integration (n=9), the studies explore how technological convergence is a key strategy for the transformation of traditional newspaper groups, using advanced technologies to achieve cross-media

content integration and efficient distribution, thereby improving communication effectiveness. For organizational restructuring (n=5), the studies demonstrate how traditional newspaper groups optimize their communication capabilities by integrating resources and establishing a cross-media matrix. However, all development strategies depend on policy support, which is closely tied to government regulation of public media in China. Policy support (n=6) is crucial, with these studies revealing that the establishment of new media matrices by traditional newspaper groups, leveraging the advantages of new media, relies on strong national policies and government funding, providing critical support for the transformation of newspaper groups.

3. Implications

The integration of quantitative survey data and qualitative thematic narratives offers a multidimensional understanding of Malaysian primary school teachers' perspectives on AI in education specifically human-centric approach to ethical concerns. The findings reveal significant implications for educational policy, professional teaching development, curriculum design, and ethical governance.

Policy and Governance Implications: Teachers' strong endorsement of AI's potential benefits particularly in enhancing efficiency, access to information, and societal well-being signals readiness for AI adoption in schools. However, their concerns regarding privacy, bias, and institutional transparency underscore the urgent need for robust regulatory frameworks (Cober et al., 2015; Govindarajoo et al., 2025). Policymakers must prioritize the development of national AI guidelines that address ethical utilization, data protection, and culturally responsive design (Floridi et al., 2018; Jobin et al., 2019). The low trust in developers and authorities to communicate risks ($M = 3.02$) suggests that top-down implementation strategies must be complemented by participatory governance involving educators as co-designers and evaluators (Dilek and Baran, 2025).

Pedagogical and Curriculum Implications: The qualitative themes reveal that teachers perceive AI as a pedagogical assistant rather than a replacement. Their emphasis on human judgment, creativity, and critical thinking suggests that AI integration should be framed as augmentation, not automation (Vias et al., 2025). Curriculum designers should embed AI literacy within broader frameworks of digital citizenship, emphasizing ethical reasoning, contextual interpretation, and inquiry-based learning (Luckin, 2018; Georgiadis and Arvanitidou, 2024). Strategies such as student-led critiques of AI outputs and value-based discussions can foster deeper engagement and intellectual autonomy (Ng, 2025).

Professional Development Implications: Teachers' cautious optimism and ethical awareness point to the need for targeted professional development programs. These should move beyond technical training to include interpretive, ethical, and pedagogical dimensions of AI utilization (Panjani, 2024). Workshops and collaborative design labs can empower teachers to experiment with AI tools while critically reflecting on their implications (UNDP, 2025). Given the variability in perceptions as evidenced by elevated standard deviations in risk-related items, differentiated support is essential to accommodate diverse levels of digital fluency and ethical literacy (Kucukuncular and Ertugan, 2025).

Socio-Cultural Implications: The call for AI systems to reflect local values and civilizational principles highlights the importance of cultural embedding in AI design. Teachers' narratives advocate for systems that are not only functional (Dave, 2025) but also aligned with national educational philosophies and community norms (UNDP, 2025). This suggests that AI deployment in Malaysian schools must be context-sensitive, inclusive, and dialogic (Vias et al., 2025) by engaging educators, students, and local stakeholders in shaping ethical and culturally resonant AI ecosystems (Seldon and Abidoye, 2018; Zawacki-Richter et al., 2019).

Collectively, the findings underscore a dual imperative: to embrace technological innovation while ensuring human-centric governance. This study's conceptual framework, therefore, provides stakeholders with the necessary guidance to align AI innovation with the transformative goals of Malaysian education, fostering an approach that is both technologically potent and ethically sound.

4. Recommendations

As AI become increasingly embedded in educational contexts, ethical considerations must extend beyond technical safeguards and algorithmic design. A human-centric approach to AI ethics foregrounds the lived experiences, values, and agency of educators and learners, emphasizing relational accountability over computational precision. This perspective challenges the dominant techno-solutionist paradigm by situating AI within broader socio-cultural and pedagogical ecosystems. Policy frameworks must evolve to reflect this shift. Ethical guidelines should not only regulate data privacy and algorithmic transparency but also ensure contextual equity and participatory governance. Mandating teacher involvement in AI procurement and implementation processes affirms their role as co-designers rather than passive adopters. Longitudinal studies are essential to assess AI's impact on professional identity, instructional autonomy, and systemic inclusion. Professional development must prioritize critical AI literacy, enabling educators to interrogate design

assumptions and advocate for culturally responsive tools. Reflective communities of practice can foster ethical dialogue and collaborative sense-making, reinforcing human judgment in AI-mediated environments. In classrooms, AI should support and not supplant the teacher's expertise. Human-AI collaboration models must preserve pedagogical intentionality and relational depth. Tools must be designed to recognize diverse learner identities and promote student agency, curiosity, and autonomy. Ultimately, ethical AI integration demands a paradigm shift: from code-centric compliance to human-centric care. By aligning policy, pedagogy, and professional empowerment, stakeholders can ensure that AI serves as a catalyst for equity, dignity, and transformative learning.

Table 5: Framework of Policy Matrix: Human-Centric AI Ethics in Education

Strategic Domain	Actionable Recommendation	Responsible Stakeholders	Intended Outcome
Policy & Governance	Establish ethical AI guidelines (privacy, transparency, equity)	Ministries of Education, Regulatory Bodies	Safeguarded rights and equitable AI deployment
	Mandate teacher involvement in AI procurement	School Leaders, Procurement Committees	Contextual relevance and pedagogical alignment
	Fund longitudinal studies on AI's impact	Research Councils, Government Agencies	Evidence-informed policy and sustained evaluation
Professional Development	Develop tiered AI literacy programs for educators	Teacher Training Institutes, EdTech Providers	Scalable capacity building across roles and contexts
	Integrate AI scenarios into pre-service teacher education	Universities, Accreditation Bodies	Critical engagement and future-ready educators
	Support reflective communities of practice	School Networks, Professional Associations	Collaborative innovation and ethical dialogue
Pedagogical Practice	Promote human-AI collaboration models in instruction	Curriculum Designers, Teachers	Augmented teacher judgment and relational pedagogy
	Ensure cultural and linguistic responsiveness in AI tools	EdTech Developers, QA Panels	Inclusive learning environments and identity affirmation
	Monitor AI's influence on student motivation and autonomy	Teachers, School Psychologists	Enhanced learner agency and engagement
Research & Evaluation	Adopt mixed methods approaches to evaluate AI in education	Academic Researchers, Think Tanks	Nuanced understanding of AI's educational impact
	Develop conceptual models linking AI use to teacher empowerment	Policy Analysts, Educational Researchers	Frameworks for sustainable and empowering integration
	Conduct cross-national comparative studies	International Consortia, UNESCO, OECD	Globally relevant insights and culturally responsive policy

5. Conclusion

The ethical integration of artificial intelligence in education must move beyond technical compliance and embrace a human-centric paradigm one that prioritizes relational accountability, contextual equity, and professional agency. Teachers are not merely end-users of AI systems; they are critical actors whose insights, values, and pedagogical expertise must inform every stage of AI design, deployment, and evaluation. By embedding ethical considerations into policy, professional development, and classroom practice, stakeholders can ensure that AI supports and not supplants the human judgment and instructional intentionality. This approach demands systemic commitment: national guidelines must reflect ethical pluralism and local realities; professional development must cultivate critical AI literacy; and classroom tools must be responsive to diverse learner identities. Research must illuminate the lived experiences of educators and learners, using mixed-methods and conceptual modeling to guide sustainable innovation. In the end, a human-centric ethic reframes AI not as a neutral tool, but as a socio-technical actor embedded in complex educational ecologies. By centering human dignity, cultural responsiveness, and collaborative design, we can shape AI systems that advance equity, foster empowerment, and enrich the moral fabric of education. This is not merely a technical challenge it is a collective ethical responsibility.

Acknowledgement

The authors would like to express his gratitude to the City University Malaysia for the support provided.

Conflict of Interest

The authors declare there is no conflict of interest

References

- Akinagbe, O. (2024). The future of artificial intelligence: Trends and predictions. *Mikailalsys Journal of Advanced Engineering International*, 1(3), 249-261. DOI: 10.58578/mjaei.v1i3.4125
- Akova, F. (2023). Artificially sentient beings: Moral, political, and legal issues. *New Techno Humanities*, 3(1), 41-48. <https://doi.org/10.1016/j.techum.2023.04.001>
- Alalag, A. S. (2025). Artificial Intelligence technologies and the humanities: Importance and challenges. *Artificial Intelligence and Electrical & Electronics Engineering*, 1(2), 1-8. https://www.researchgate.net/publication/396375915_Artificial_Intelligence_Technologies_and_the_Humanities_Importance_and_Challenges
- Alegre, S. (2024). *Human rights, robot wrongs: Being human in the age of AI*. London: Atlantic Books.
- Alsaleh, A. (2024). The impact of technological advancement on culture and society. *Scientific Report*, 14, 1-8. <https://doi.org/10.1038/s41598-024-83995-z>
- Annet, K. A. (2025). The ethics of AI: Philosophical perspectives. *Research Invention Journal of Research in Education*, 5(2), 65-71. DOI: 10.59298/RIJRE/2025/526571.
- Brandao, P. (2025). The impact of artificial intelligence on modern society. *AI*, 6(8), 1-29. 190. <https://doi.org/10.3390/ai6080190>
- Brauner, P., Glawe, F., Luca Liehner, G., Vervier, L. & Ziefle, M. (2025). Mapping public perception of artificial intelligence: Expectations, risk–benefit tradeoffs, and value as determinants for societal acceptance. *Technological Forecasting and Social Change*, 220, 1-17. <https://doi.org/10.1016/j.techfore.2025.124304>
- Cheong, B. C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamic*, 6, 1-11. DOI: 10.3389/fhumd.2024.1421273
- Chong, K., Tan, Y. & Zhou, X. (2025). Impacts of artificial intelligence development on humanity and social values. *Information*, 1-14, 16(9), <https://doi.org/10.3390/info16090810>
- Coher, R., Tan, E., Slotta, J., So, H. J., & Könings, K. D. (2015). Teachers as participatory designers: Two case studies with technology-enhanced learning environments. *Instructional Science*, 43(2), 203-228. <https://doi.org/10.1007/s11251-014-9339-0>
- Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education (6th ed.)*. London: Routledge.
- Creswell, J. W. (2014). *Research Design: Qualitative, quantitative, and mixed methods approaches (4th ed.)*. California: Sage Publications.
- Creswell, J. W., Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. California: Sage Publications.
- Dave, A. (2025). The ethical implications of AI in education. *Research Review Journal of Social Science*, 5(1), 1-8. DOI: 10.31305/rjss.2025.v05.n01.001
- Davidson, D. (2024). *AI in education: Revolutionizing learning and administration*. (n.p.): Pure Water Books.
- Diaz Noguera, M. D. Hervas-Gomez, C. & Vera, F. (2024). *Artificial intelligence and education enhancing human capabilities, protecting rights, and fostering effective collaboration between humans and machines in life, learning, and work*. Barcelona: Ediciones OCTAEDRO.
- Dilek, M. & Baran, E. (2025). AI literacy in teacher education: Empowering educators through critical co-discovery. *Journal of Teacher Education*, 76(3), 1-43. DOI: 10.1177/00224871251325083
- Fan, L. (2024). Artificial intelligence ethics: A dialogue between technological advances and human values. *International Journal of Education and Humanities*, 14(2), 260-265. DOI: 10.54097/tvqkkf40
- Ferrara, E. (2024). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci*, 6(1), 3. <https://doi.org/10.3390/sci6010003>
- Flavin, E., Hwang, S. & Morales, M. (2025). “Let’s ask the robot!”: Epistemic stance between teacher candidates toward AI in mathematics lesson planning. *Journal of Teacher Education*, 76(3), 262-279. <https://doi.org/10.1177/00224871251325079>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P. & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Fu, Y. & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, 7, 1-20. <https://doi.org/10.1016/j.caeai.2024.100306>
- Gade, K. (2025). The ethics of intelligence: Navigating AI’s impact on society. *ESP International Journal of Advancements in Science & Technology*, 3(1), 1-9. DOI: 10.56472/25839233/IJAST-V3I1P104

- George, A. S. (2024). Artificial intelligence and the future of work: Job shifting not job loss. *Partners Universal Innovative Research Publication*, 2(2), 17-37. DOI: 10.5281/zenodo.10936490
- Georgiadis, M. & Arvanitidou, P. (2024). The impact of principals' empathy on teachers' performance. *Proceedings*, 111(1), 1-10. <https://doi.org/10.3390/proceedings2024111021>
- Gilbert, C. & Gilbert, M. (2024). Strategic framework for human-centric ai governance: Navigating ethical, educational, and societal challenges. *International Journal of Latest Technology in Engineering Management & Applied Science*. 13(8), 132-14. DOI: 10.51583/IJLTEMAS.2024.130816
- Govindarajoo, M V. V., Nair, S. M., Sekhon, R. S., Wai, C. S., Hoong, L. C., Huat, T. B. & Okawa, T. (2025). Exploring teachers' views on benefits, ethical issues, and challenges in integrating AI tools in Malaysian schools. *Edelweiss Applied Science and Technology*, 9(9), 699-710. DOI: 10.55214/2576-8484.v9i9.9935
- Hanna, m. g., Pantanowitz, L., Jackson, B., Palmer, O., Visweswaran, S., Pantanowitz, J., Deebajah, M. & Rashidi, H. H. (2025). Ethical and bias considerations in artificial intelligence/machine learning. *Modern Pathology*, 38(3), 1-13. <https://doi.org/10.1016/j.modpat.2024.100686>
- Harjatanaya, T. Y., Ferary, D. & Sarhindi, I. (2025). Education for sustainable development using participatory methodologies in Indonesian higher education: The case of Satya Terra Bhinneka University. *International Journal of Sustainability in Higher Education*, 26(6), 1348-1368. <https://doi.org/10.1108/IJSHE-09-2024-0657>
- Hohma, E. & Lutge, C. (2023). From trustworthy principles to a trustworthy development process: The need and elements of trusted development of AI systems. *AI*, 4(4), 904-925. DOI: 10.3390/ai4040046.
- Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jowarder, R. A. (2025). The ethics of AI decision-making: Balancing automation, explainable AI, and human oversight. *International Journal of Science and Research Archive*, 14(03), 435-443. <https://doi.org/10.30574/ijrsra.2025.14.3.0677>
- Kirk, H.R., Gabriel, I., Summerfield, C., Vidgen, B. & Hale, S. S. (2025). Why human–AI relationships need socio affective alignment. *Humanities & Social Science Communications*, 12, 1-9. <https://doi.org/10.1057/s41599-025-04532-5>
- Kirova, V., Ku, C., Laracy, J. & Marlowe, T. (2023). The ethics of artificial intelligence in the era of generative AI. *Journal of Systemics, Cybernetics and Informatics*, 21(4), 42-50. DOI: 10.54808/JSCI.21.04.42
- Koliada, N., Abramova, O. & Drachuk, M. (2025). Developing the concept of digital humanism as human interaction with artificial intelligence. *Pakistan Journal of Life and Social Sciences*, 23(1), 238-248. <https://doi.org/10.57239/PJLSS-2025-23.1.0021>
- Kucukuncular, A. & Ertugan, A. (2025). Teaching in the AI era: Sustainable digital education through ethical integration and teacher empowerment. *Sustainability*, 17(16), 1-20. DOI: 10.3390/su17167405
- Kumar, M. R. & Rongali, S. K. (2025). Balancing AI and human collaboration. *World Journal of Advanced Research and Reviews*, 25(03), 562-567. <https://doi.org/10.30574/wjarr.2025.25.3.0556>
- Lee, L. A., Durakovic, B. & Yeralan, S. (2025). AI ethics: Contextual frameworks and domain-specific concerns. *Heritage and Sustainable Development Editorial*, 7(1), 1-14. <https://doi.org/10.37868/hsd.v7i1.1224>
- Lochmiller, C., Lester, J. N. (2015). *An introduction to educational research: Connecting methods to practice*. California: Sage Publications.
- Luckin, R. (2018). *Machine learning and human intelligence: The future of education for the 21st century*. London: UCL IOE Press.
- Matcas, V. L. (2016). *The human civilization*. Washington: Amazon Digital Services LLC.
- McIntyre, A., Conover, L. & Russo, F. (2025). A network approach to public trust in generative AI. *Philosophy & Technology*, 38(4), 1-31. <https://doi.org/10.1007/s13347-025-00974-6>
- Mehra, C. S. (2025). Ethical challenges and solutions in artificial intelligence development. *International Journal for Multidisciplinary Research*, 7(2), 1-10. <https://www.ijfmr.com/papers/2025/2/42527.pdf>
- Misra, S. (2022). Revitalizing human values in an age of technology. *Philosophies*, 7(6), 136, 1-11. <https://doi.org/10.3390/philosophies7060136>
- Mitra, S. (2025). *Another big bang: Evolution of civilisation*. Freedom all the way. New Delhi: BlueRose Publishers Pvt Ltd.
- Muhammed, S. & Kamsin, I. F. (2025). Teacher's acceptance and intention to use artificial intelligence technology in teaching and learning based on the UTAUT model. *International Journal of Information and Education Technology*, 15(7), 1428-1435. <https://www.ijiet.org/vol15/IJiet-V15N7-2344.pdf>
- Namukasa, S. (2025). The role of technology in shaping future generations' values. *Eurasian Experiment Journal of Humanities and Social Sciences*, 6(3), 110-116. https://www.researchgate.net/publication/390398719_The_Role_of_Technology_in_Shaping_Future_Generations'_Values

- Nawi, Aliff. (2020). Early exploration towards issues and impact the use of artificial intelligence technology towards human beings. *Asian Journal of Civilizational Studies*, 1(4), 24-33. <https://www.ajocs.com/index.php/ajocs/article/view/25>
- Ng, T. S. (2025). Cultivation of students' autonomous learning ability in the context of artificial intelligence. *Journal of Artificial Intelligence Practice*, 8(2), 39-44. DOI: 10.23977/jaip.2025.080205
- Nikitenko, V., Voronkova, V., Oleksenko, R., Kyvliuk, O., Klochek, L., Koliada, N., Abramova, O., & Drachuk, M. (2025). Developing the concept of digital humanism as human interaction with artificial intelligence. *Pakistan Journal of Life and Social Sciences*, 23(1), 238–248. <https://doi.org/10.55309/pjlss.2025.23.1.53>
- Panjani, A. (2024). AI (artificial intelligence) integration in education: Teachers' perspectives, professional development and policy recommendations. *Journal of Information Systems Engineering and Management*, 9(4), 138-145. DOI: 10.52783/jisem.v9i4s.10602
- Papagiannidis, E., Mikalef, P. & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, 34(2), 1-18. <https://doi.org/10.1016/j.jsis.2024.101885>
- Promsiri, T. (2025). AI and the psychology of educational disruption: Historical patterns and cognitive implications. *Acta Psychologica*, 260, 1-13. <https://doi.org/10.1016/j.actpsy.2025.105637>
- Qi, G. A., Hoon, E. C. L., Scott, N., Zaini, N. H., Lin, W. F. & Na, K. S. (2025). Perceptions of the value of artificial intelligence (AI) in education. *Innovative Teaching and Learning Journal*, 9 (1), 274-283. <https://doi.org/10.11113/itlj.v9.177>
- Radanliev, P. (2025). AI ethics: Integrating transparency, fairness, and privacy in AI development. *Applied Artificial Intelligence*, 39(1). 1-41. <https://doi.org/10.1080/08839514.2025.2463722>
- Razak, N. & Norman, H. (2025). The use of artificial intelligence (AI) among teachers in implementing ICT-based education for the RBT subject. *International Journal of Academic Research in Progressive Education and Development*, 14(2), 1085-1094. DOI: 10.6007/IJARPED/v14-i2/25384
- Roche, C., Lewis, D., & Wall, P. J. (2021). Artificial intelligence ethics: An inclusive global discourse? *arXiv*, 8, 643-658. DOI: 10.48550/arXiv.2108.09959
- Rogers, E. M. (1986). *Communication technology: The new media in society*. New York: Free Press.
- Sabah, A. S., Hamouda, A. A., Abu-Nasser, B. S. & Abu-Naser, S. S. (2025). The intersection of AI and human rights: Challenges and opportunities. *International Journal of Academic Information Systems Research*, 9(4), 15-21. <http://ijeais.org/wp-content/uploads/2025/4/IJAISR250403.pdf>
- Sakubu, D. (2025). Challenges of artificial intelligence today and future implications for society and the world. *World Journal of Advanced Research and Reviews*, 26(01), 3045-3054. <https://doi.org/10.30574/wjarr.2025.26.1.1380>
- Sari, J. & Purwanta, E. (2021). The implementation of artificial intelligence in STEM-based creative learning in the society 5.0 era. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah*, 6, 433-440. DOI: 10.24042/tadris.v6i2.10135
- Seldon, A. & Abidoye, T. (2018). *The fourth education revolution: Will artificial intelligence liberate or infantilise humanity?* London: University of Buckingham Press.
- Sigfrids, A., Leikas J, Salo-Pontinen, H. & Koskimies, E. (2023) Human-centricity in AI governance: A systemic approach. *Frontiers in Artificial Intelligence*, 6, 1-9. DOI: 10.3389/frai.2023.976887
- Sivanganam, J., Md Yunus, M. & Mohd Said, N. E. (2025). Teachers' perceptions in using artificial intelligence (AI) in ESL classrooms. *International Journal of Academic Research in Progressive Education and Development*, 14(1), 2409-2427. <https://doi.org/10.6007/IJARPED/v14-i1/25034>
- Tai M. C. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal*, 32(4), 339–343. https://doi.org/10.4103/tcmj.tcmj_71_20
- Todino, M. D. (2025). Educational technologies. *Encyclopedia*, 5(1), 1-23. <https://doi.org/10.3390/encyclopedia5010023>
- United Nations Development Programme (UNDP). (2025). *A matter of choice: People and possibilities in the age of AI*. New York: United Nations Development Programme (UNDP).
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2022). *Recommendation on the ethics of artificial intelligence*. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO).
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2025). *Artificial intelligence and culture*. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO).
- Vernyuy, A. (2024). Impact of technological advancements on human existence. *International Journal of Philosophy*, 3(2), 54-66. 10.47941/ijp.1874.
- Vias, S., Derioh, M. M. G., Chandran, S., Ragunathan, S. & Ridwan S. R. (2025). An exploratory study of teachers' perceptions of social-emotional learning (SEL) strategies in reducing teacher burnout among selected government and international schools in Semenanjung Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 15(9), 1408-1418. <http://dx.doi.org/10.6007/IJARBS/v15-i9/26280>

- Wiese, L. J., Patil, I., Schiff, D. S. & Magana, A. J. (2025). AI ethics education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 8, 1-22. <https://doi.org/10.1016/j.caeai.2025.100405>
- Xue, L., Ghazali, N. & Mahat, J. (2025). Artificial intelligence (AI) adoption among teachers: A systematic review and agenda for future research. *International Journal of Technology in Education*, 8(3), 802-824. DOI: 10.46328/ijte.1191
- Yuxuan, Z. & Wan Hussain, W. M. H. (2025). Artificial intelligence (AI) awareness (2019–2025): A systematic literature review using the SPAR-4-SLR protocol. *Social Sciences & Humanities Open*, 12, 1-16. <https://doi.org/10.1016/j.ssaho.2025.101870>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 1-27. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhang, N., Lun, W. & Barrot, J. (2025). Primary teachers' acceptance and sustained adoption of AI powered learner corpora for writing instruction through TAM and ECM perspectives. *Scientific Reports*, 15(1), 1-25. DOI: 10.1038/s41598-025-21205-0
- Zulkarnain, N, S. & Yunus, M. (2023). Primary teachers' perspectives on using artificial intelligence technology I English as a second language teaching and learning: A systematic review. *International Journal of Academic Research in Progressive Education and Development*, 12(2), 861-875. DOI: 10.6007/IJARPED/v12-i2/17119