International Journal of Basic and Applied Sciences, 14 (4) (2025) 159-167

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/3t4ghz09
Research paper

From Detection to Prevention: A Human-Centered Design (HCD) Approach to Mitigating AI Misuse in Learning

Ramlan Mustapha [1] *, Nurshahira Ibrahim [2], Mohd Nasir Ayub [1], Abu Bakar Jaafar [2], Mohd. Kamal Azman bin Jusoh [2], Maziah Mahmud [2]

[1] *Academy of Contemporary Islamic Studies, Universiti Teknologi MARA Pahang, Raub Campus, Malaysia*
[2] *Academy of Contemporary Islamic Studies, Universiti Teknologi MARA Pahang, Jengka Campus, Malaysia*
**Corresponding author E-mail: ramlan@uitm.edu.my*

Received: June 25, 2025, Accepted: July 29, 2025, Published: August 7, 2025

Abstract

The rapid integration of artificial intelligence (AI) tools in educational settings has created unprecedented challenges for academic integrity, with studies indicating that 89% of students are familiar with ChatGPT and 43% have used AI tools for academic work, while current detection-based approaches have proven inadequate, achieving only 26% accuracy in identifying AI-generated content. This study develops and validates a comprehensive Human-Centered Design (HCD) framework for mitigating AI misuse in educational contexts, shifting from reactive detection strategies to proactive prevention-oriented approaches that prioritize stakeholder engagement and ethical AI integration. Using a two-phase mixed-methods approach combining Nominal Group Technique (NGT) and Fuzzy Delphi Method (FDM) with seven experts, the research achieved exceptional expert consensus with an overall average agreement of 98% across all evaluated components. Eleven core elements were validated and organized into five interconnected dimensions: Core Design Principles (emphasizing learner agency and educator empowerment), Implementation Strategies (featuring progressive disclosure and reflection mechanisms), Assessment & Feedback Systems (incorporating authentic assessment and ethical analytics), Stakeholder Engagement (encompassing co-design processes), and Ethical Safeguards (focusing on bias detection and privacy protection). Four components achieved perfect 100% expert consensus, while the low-lowest-ranked element maintained an acceptable defuzzification value of 0.800, well above the 0.5 threshold required for validation. The validated HCD framework provides a robust, empirically supported approach for educational institutions to proactively address AI misuse while maintaining the benefits of AI integration, offering a paradigm shift from punitive detection methods to collaborative, prevention-focused strategies that enhance rather than restrict educational AI applications, thereby contributing to the emerging fields of Human-Centered Learning Analytics (HCLA) and Human-Centered Artificial Intelligence (HCAI) with practical guidance for educational stakeholders seeking responsible AI integration strategies.

Keywords: *Artificial Intelligence; Academic Integrity; Educational Technology; Fuzzy Delphi Method; Nominal Group Technique.*

1. Introduction

Artificial intelligence (AI) has quickly become a ubiquitous tool in classrooms, changing the way students learn and presenting educators with new possibilities and formidable obstacles. Shneiderman et al. (2021) and Lindenwood University (2024) argue that AI could improve educational productivity and create personalized learning experiences that meet the needs of individual students. However, they also warn that AI, if not properly governed and prone to algorithm bias, could limit human rights and lead to educational inequality. Concerns regarding data privacy and agency have been raised by the fast expansion of Learning Analytics (LA) and Artificial Intelligence in Education (AIED). One reason for this is that excluding stakeholders, such as students and teachers, from the design process can lead to mistrust and tools that aren't adequately aligned (Martinez-Maldonado et al., 2024). These days, schools face a double whammy: they need to leverage AI's revolutionary potential, but they also need to build robust systems to prevent it from being used unethically in classrooms.

The misuse of AI in education occurs in various forms, including blatant infractions like plagiarism and academic dishonesty, as well as more insidious impacts on skill development and creative thinking (Inspera, 2024; Wintmölle et al., 2024). Recent classifications of generative AI misuse indicate that malicious entities exploit AI systems to achieve specific and identifiable objectives, primarily aiming to manipulate public opinion via deepfakes, digital personas, and various types of fabricated media artefacts (Marchal et al., 2024). In educational contexts, these patterns of misuse transcend conventional academic integrity violations, presenting more intricate challenges such as the fabrication of fraudulent credentials, automated content generation that circumvents learning objectives, and the reinforcement of algorithmic biases that may intensify existing educational disparities. The rising instances of misuse underscore the pressing necessity for holistic strategies that tackle both technological weaknesses and the social environments in which educational AI systems function (Hinduja, 2024).

A methodological approach that places human wants, values, and capabilities at the center of technological design and implementation, human-centered design (HCD) emerges as a vital framework for addressing AI misuse in education. According to Giacomin (2014), human-centered design can be defined as "the process that places human needs and limitations in a higher priority than other targets during the design thinking and production differential stages. Systems that are strong, open, equitable, and in line with society's well-being are the result of AI development that takes a human-centered approach, which prioritizes the user's experience and ethical concerns (McKinsey, 2024). To overcome the problems caused by not considering stakeholder needs, not contextualizing pedagogically, and having low trust in new technology, the Human-Centered Learning Analytics (HCLA) and Human-Centered Artificial Intelligence (HCAI) research fields have recently emerged (Topali et al., 2024). This new way of thinking acknowledges that to have good AI governance in education, we need to put the requirements of students, teachers, and educational communities at the center of our design processes, rather than focusing solely on technology solutions.

The shift from detection-based to prevention-oriented strategies signifies a major transformation in the way educational institutions perceive and tackle AI misuse, evolving from reactive remediation to proactive design for ethical AI integration. A human-centered approach does not see AI as a cure-all for broken systems. Instead, it sees AI as a tool that needs to be used with care and thought about its context and possible effects. It also stresses how important humans are in finding patterns, giving meaning, and making important decisions about instruction (GSI Education, 2024). This approach requires a careful balance of the involvement of all stakeholders in the design and implementation of AI systems at every stage of the process. It is especially important to involve the target end-users, such as students, to find the right balance between human control and automation while looking at safety, reliability, and trustworthiness as key principles (Khosravi et al., 2024). By using HCD methods, schools can make full frameworks that not only find and fix problems with AI but also stop them from happening in the first place by carefully designing, involving stakeholders, and making learning spaces where AI is an empowering tool instead of a source of academic problems or moral concerns.

2. Literature review

2.1. Artificial intelligence and academic dishonesty

There has been a dramatic shift in the academic integrity environment due to the advent of advanced AI tools, especially generative AI platforms like ChatGPT, which have posed new and unprecedented challenges to conventional systems of educational evaluation. Artificial intelligence (AI) is changing the game in the realm of academic dishonesty, according to recent systematic literature reviews (Balalle & Pannilage, 2025; Ateeq et al., 2024). The field is also facing complex ethical and practical dilemmas due to technology's ability to both enable and detect academic misconduct. In today's fast-paced world of education, the presence of artificial intelligence (AI) presents major obstacles to the educational ecosystem's capacity to uphold academic integrity, which is essential for a high-quality education because it signifies honesty, trust, and ethical behaviour (Balalle & Pannilage, 2025). Studies examining the capabilities of different AI content detection tools in distinguishing between human and AI-authored content have shown substantial limitations in current detection methodologies (Elkhatat et al., 2023), which raises concerns about plagiarism and the potential challenges to academic integrity posed by AI-generated content, especially from models like ChatGPT. The first instance of academic dishonesty aided by ChatGPT was reported within two weeks after its launch (November 30, 2022; Cotton et al., 2023), prompting significant worries about student plagiarism due to the rapid public adoption of the tool.

Modern studies show that academic dishonesty with the help of AI takes many forms, from simple plagiarism to more complex forms of content creation that pass muster with current detection systems and institutional rules. According to surveys, almost 90% of students are familiar with ChatGPT, and 43% of college students have used it or a similar artificial intelligence tool for homework (89%), essays (53%), or at-home tests (48%). This shows that these tools are widely used in academic settings (Nerdy Nav, 2024; Forbes Analytics, 2024). Academic integrity concerns hurt usage, indicating that students are aware of the ethical implications but may be influenced by contextual pressures; on the other hand, time-saving features, academic stress, and peer influence are the motivations driving AI adoption in academic settings (Bin-Nashwan et al., 2023). Specifically, Dwivedi et al. (2023) found that AI tools are being classified as either assistive or substitutive. Assistive tools help students understand concepts better, while substitutive tools can take over the student's tasks, which could lead to integrity breaches. An especially worrying development that brings academic dishonesty into the digital era is the rise of AI-driven essay mills that provide custom essays, which are often unidentifiable by conventional plagiarism detectors (Khosravi et al., 2022).

Emerging as a crucial reaction to AI-facilitated academic misconduct, the development and implementation of AI detection technologies have been met with empirical evidence that consistently shows present detection systems have severe shortcomings in accuracy, dependability, and fairness. There is growing proof of the inaccuracy of AI detection technologies, but 68% of professors still use them to combat academic dishonesty—a 30 percentage point increase from the previous year (ArtSmart AI, 2024). Elkhatat et al. (2023) found that AI detection tools produce false positives and uncertain classifications when applied to human-written content, but they were more accurate when identifying content generated by GPT 3.5 than GPT 4. Other studies have shown that AI detection tools exhibit inconsistent performance. It is recommended that classifier results be regarded as supplementary information rather than definitive evidence because OpenAI's classifier correctly identifies 26% of AI-written text as "likely AI-generated" but incorrectly labels 9% of human-written text as AI-generated (Kirchner et al., 2023). Academic assessment equity and fairness are major concerns, as research has shown that AI detection tools are biased against non-native English writers and have a hard time dealing with AI content that has been edited or translated from another language (Liang et al., 2023; EdTech Digest, 2024). Research has shown that AI detection methods are completely ineffective when it comes to identifying AI-generated writing that has been edited by AI to make it seem more human (EdTech Digest, 2024), which is a major problem.

The literature reveals a growing consensus that addressing AI-facilitated academic dishonesty requires comprehensive institutional responses that move beyond technological detection toward proactive pedagogical and policy reforms centered on prevention, education, and authentic assessment design. Recent research emphasizes that while AI presents challenges to academic integrity, it should not be viewed solely as a threat but rather as a catalyst for reimagining educational practices and assessment methodologies (GSI Education, 2024; University of Pittsburgh Teaching Center, 2024). According to continuing studies by Stanford University on cheating among American high school students, which began before ChatGPT was released, 60 to 70% of pupils report having engaged in cheating behaviors at least once in the last month. A number that has remained stable or even declined marginally after the introduction of ChatGPT (Stanford GSE, 2024), indicating that AI could not be drastically changing cheating behaviors but rather giving new tools for already observed tendencies.

Educational institutions are increasingly adopting comprehensive approaches that include developing clear AI usage policies, implementing honor codes, providing faculty training on AI technologies, and designing assessments that emphasize critical thinking, creativity, and application rather than knowledge reproduction (Cornell CTI, 2024; Community College Daily, 2024). The literature strongly advocates for human-rights-by-design approaches to misconduct policies that centre human rights in policy development and implementation, ensuring procedural fairness, privacy protection, and recognition that not all use of AI tools automatically constitutes misconduct (Eaton, 2024). Future research directions identified across multiple systematic reviews include the need for longitudinal studies examining the long-term impact of AI on learning outcomes, development of more nuanced ethical frameworks for AI usage in education, and investigation of pedagogical approaches that can effectively integrate AI tools while maintaining academic integrity standards (Yusuf et al., 2024; MDPI Information, 2025).

Based on the above literature highlights, we can imagine that the issue of academic AI is not a matter of light. The generation of science and academic students needs to be more integrated into the knowledge of knowledge and not be dependent 100% on AI. Therefore, we think a human-centered design to resolve this issue needs to be built to guide in the future. Therefore, we will build a human-centered approach framework to solve the problem of abuse of AI in the academic and learning world. As a result, we formed research aims as follows:

2.2. Research aims

- Explore the constructs and elements of the Human-Centered Design Approach (HCD) based on experts' opinions.
- Verifying Human-Centered Design Approach (HCD content based on expert consensus)

3. Methodology

This study uses the Nominal Group Technique and Fuzzy Delphi Method approach to analyze the study data. This study is a systematic form of study that is based on expert income and is widely used in development studies (Mustapha et al., 2022).

3.1. Nominal group technique (NGT)

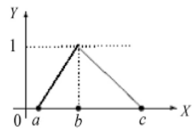
As a consensus methodology, the Nominal Group Technique was selected due to its methodical approach that combines qualitative and quantitative data collection in a group setting. This approach avoids the problems caused by group dynamics that are common in other methods like focus groups, brainstorming, and Delphi (Gallagher et al., 1993). Using the tried-and-true four-stage Nominal Group Technique (NGT) process, the methodology began with a silent generation phase. During this phase, participants were asked to reflect on and record their thoughts in response to the following structured question: "What are the most critical considerations for developing Islamic ethical frameworks for autonomous cybersecurity systems?" After that, there was a round of round-robin sharing, where everyone took turns presenting their ideas, but no one discussed them. After that, a round of clarification and debate ensued to ensure everyone was on the same page and eliminate any confusion. Afterwards, there was a vote and ranking phase where participants ranked their top five considerations anonymously, with the total scores representing the group's agreement. To examine data from the voting process, this study used NGT-Plus software. With this method, even the most reserved participants in a focus group may contribute ideas. Instead of focused topic probing, which might be required by a more organized survey approach at times, it is great at revealing themes of essential importance to respondents (Boddy, 2012).

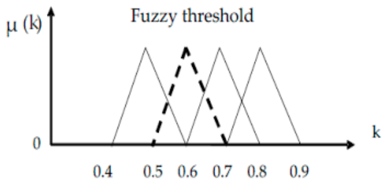
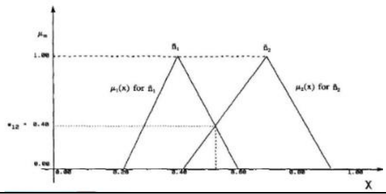
3.2. Fuzzy Delphi method

This research employs the Fuzzy Delphi Method (FDM). This study was selected because of its distinctive approach to acquiring expert consent for making a definitive judgment. This study employs two steps in the development of the study questionnaire, specifically through a literature review. This study comprises two phases in the development of the questionnaire items. The initial phase involves the researcher doing a brainstorming analysis to determine the necessary aspects for HCD in the NGT session, as referenced in Table 2. In the second phase, once all constructs and elements are gathered, the researcher constructs a 7-point expert questionnaire, which is then circulated to 7 experts with relevant experience and assessed using the Fuzzy Delphi Method (FDM).

The original Delphi Method, developed by Dalkey and Helmer in 1963, has undergone several revisions since then. Modernized and expanded upon from the original Delphi technique, the Fuzzy Delphi Method (FDM) comes highly recommended. The Delphi technique differs from FDM in that it handles decision-making uncertainty using probability theory rather than mathematical concepts. For this reason, Hsu, Lee, and Kreng (2010) proposed the Fuzzy Delphi Method (FDM), which combines fuzzy theory with conventional decision-making techniques to account for people's language preferences. Merging fuzzy theories with more conventional decision-making processes led to the development of the FDM. One of the first studies on using FDM for forecasting was done by Kaufmann and Gupta (1988), while Murray, Pipino, and Gigch (1985) and Mustapha and Darusalam (2022) were the first to suggest combining classical DM with fuzzy theory to reduce the latter's inherent imprecision. With the hope that computers would improve, the theory was expanded to incorporate FDM, the max-min algorithm, and fuzzy integration (Ishikawa, Amagasa, Shiga, Tomizawa, Tatsuta, & Mieno, 1993; Mustapha & Darusalam, 2022). According to Garai (2013), the inclusion of weights in the FDM version highlights the variety of expert knowledge and experience. Chang, Hsu, and Chang (2011) introduced a new variant of fuzzy decision-making (FDM) that uses fuzzy statistics and may stabilize iterative processes by accommodating continuous mathematically explicit membership functions. The steps for conducting FDM are as follows:

Table 1: Fuzzy Delphi Step

	Step	Formulation
1	Choosing a Suitable Expert: Seven Experts Participated in This Inquiry. An expert panel was convened to evaluate the significance of linguistic characteristics of the weight assigned to the assessment criteria by the relevant elements.	N/A
2	Triangle fuzzy numbers are used to capture all language aspects of decision-making. According to Hsieh, Lu, and Tzeng (2004), linguistic variables have been enhanced with fuzzy numbers. Triangular fuzzy numbers (m1, m2, m3) represent M1, M2, and M3, respectively. The lowest possible value is represented by minimum (m1), while the highest possible value is represented by maximum (m3). The original language variables are transformed into fuzzy numbers using the Fuzzy Scale, which is made from triangular fuzzy numbers.	$\mu_A(x)=\begin{cases}0, & x\le a\\ \frac{x-a}{b-a}, & a<x\le b\\ 1, & x=b\\ \frac{c-x}{c-b}, & b<x\le c\\ 0, & x\ge c.\end{cases}$ 

<table>
<tr><td>3</td><td>The Determination of Linguistic Variables and Average Responses. It is necessary to change all Likert scales to fuzzy scales once the selected expert has responded, before the researcher can calculate average replies and linguistic factors. Thus, we are averaging the opinions of all fuzzy numbers (Benitez, Martin, & Roman, 2007).</td><td><table>
<tr><th colspan="2">Linguistic variables</th><th>Fuzzy numbers</th></tr>
<tr><td>1</td><td>Very low (VL)</td><td>(0.0, 0.0, 0.1)</td></tr>
<tr><td>2</td><td>Low (L)</td><td>(0.0, 0.1, 0.25)</td></tr>
<tr><td>3</td><td>Medium low (ML)</td><td>(0.15, 0.3, 0.45)</td></tr>
<tr><td>4</td><td>Medium (M)</td><td>(0.35, 0.5, 0.65)</td></tr>
<tr><td>5</td><td>Medium high (MH)</td><td>(0.55, 0.7, 0.85)</td></tr>
<tr><td>6</td><td>High (H)</td><td>(0.8, 0.9, 1.0)</td></tr>
<tr><td>7</td><td>Very high (VH)</td><td>(0.9,1.0,1.0)</td></tr>
</table></td></tr>
<tr><td>4</td><td>The determination of the threshold value “d.” Methods for calculating the value of “d: to determine the level of agreement among specialists, a threshold value must be established (Thomaidis, Nikitakos, & Dounias, 2006). The distance between two fuzzy numbers can be determined using the formulas.</td><td>$d(\bar{m},\bar{n}) = \sqrt{\frac{1}{3}\left[(m1-n1)^2 + (m2-n2)^2 + (m3-n3)^2\right]}$</td></tr>
<tr><td>5</td><td>Identify the alpha cut aggregate level of fuzzy assessment. After a consensus amongst experts is reached, each item is given a fuzzy value (Mustapha et al., 2022). The following formula is used for the computation and determination of fuzzy values: This is the maximum value: Amax = (1)⁄4 (m1 + 2m2 + m3)</td><td>Amax = (1)⁄4 (m1 + 2m2 + m3)</td></tr>
<tr><td>6</td><td>Defuzzification system. The formula Amax = (1)⁄4 (a1 + 2 am + a3) is utilized in this procedure. According to Mustapha et al. (2022), when researchers employ average fuzzy numbers or average responses, the final score is a number between zero and one. Three formulas are involved in this process: i. A = 1/3 * (m1 + m2 + m3), ii. A = 1/4 * (m1 + 2m2 + m3), and iii. A = 1/6 * (m1 + 4m2 + m3). Where α-cut = (0 + 1) / 2 = 0.5, the A-cut value is the median value for '0' and '1'. The item will be rejected if the resultant A value is lower than the α-cut value = 0.5, as it does not represent an expert agreement. Bojdanova (2006) states that the alpha cut value needs to be more than 0.5. Tang & Wu (2010), who asserted that the α-cut value ought to exceed 0.5, lend credence to this idea.</td><td></td></tr>
<tr><td>7</td><td>Ranking process. Experts agree on defuzzification values, which are used to pick elements for placement; the most significant element is the one with the greatest value (Fortemps & Roubens, 1996).</td><td>$e_{ij} = \max_{x \geq y}\{\min(\mu_i(x), \mu_j(y))\}$ for all $i, j = 1, 2, 3, \cdots, m.$
</td></tr>
</table>

3.3. Experts' criteria

Booker and McNamara (2004) state that experts are those who have put in a lot of work to have the right combination of education, training, experience, accreditation, and acknowledgement from their peers (Nikolopoulos, 2004; Perera et al., 2012). A person is considered an "expert" when they have vast knowledge and years of experience in their chosen subject (Cantrill, Sibbald, & Buetow, 1996; Mullen, 2003). Fuzzy Delphi studies rely heavily on expert selection. If the study wants to be legitimate, trustworthy, and up to code, its experts need to be hand-picked using strict criteria (Mustapha & Darusalam, 2018). The researchers and specialists involved must own or have a thorough grasp of the field (Kaynak & Macauley, 1984). To ensure that the experts chosen for the study have the necessary knowledge and experience, the researcher sets stringent criteria, such as a minimum of seven years in the field. The first phase (NGT) and second phase (FDM) of this study were conducted with the participation of seven specialists.

3.4. Sampling

In this study, purposive sampling is used. To get a unanimous decision on a previously decided-upon topic, this tactic is ideal for the researcher. Purposive sampling is the most appropriate strategy for NGT and FDM, according to Hasson, Keeney, and McKenna (2000). This research benefited from the expertise of seven individuals. The qualifications and depth of knowledge of these experts led to their selection. The number of professionals needed for this study often falls between 5 and 10, assuming they are all uniform specialists. When homogeneity is present, a minimum of 10 to 15 experts can be used in a Delphi study (Adler & Ziglio, 1996).

3.5. Findings

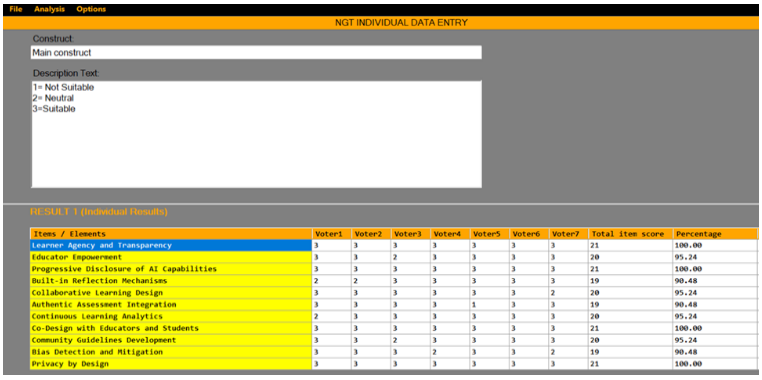
This study combines 2 methods: NGT and FDM. Phase 1, the NGT phase, is the brainstorming phase of the construct and element for mitigating AI misuse. The NGT phase (brainstorming) results are arranged in the following table:

Table 2: Construct and Elements for Mitigating AI misuse

Core component (Approach)	Elements
Learner Agency and Transparency	• Design AI tools that communicate their capabilities and limitations to students. • Provide visible indicators of AI assistance, allowing learners to make informed decisions about when and how to use these tools. • Build interfaces that encourage critical thinking rather than passive consumption.
Educator Empowerment	• Create AI systems that enhance rather than replace human instruction. • Provide educators with comprehensive dashboards showing how students interact with AI tools, enabling them to identify potential over-reliance or misuse patterns. • Include customizable controls that allow teachers to set appropriate boundaries for different learning contexts.
Progressive Disclosure of AI Capabilities	• Introduce AI features gradually, starting with basic assistance and expanding access as students demonstrate responsible use. This scaffolded approach helps build digital literacy alongside subject knowledge.
Built-in Reflection Mechanisms	• Integrate prompts that encourage students to explain their thinking process, justify their use of AI assistance, and reflect on their learning journey. This helps maintain metacognitive awareness and prevents mindless delegation to AI systems.

Collaborative Learning Design	• Structure AI interactions to promote peer discussion and collaborative problem-solving rather than isolated AI-human exchanges. • Design features that encourage students to share their AI-assisted work with classmates for feedback and discussion. • Assessment and Feedback System
Authentic Assessment Integration	• Develop assessment methods that work with AI tools rather than against them. • Focus on higher-order thinking skills, creative application, and real-world problem-solving that requires human judgment and contextual understanding.
Continuous Learning Analytics	• Implement ethical data collection that tracks learning progress without compromising privacy. • Use this information to identify students who may be struggling with appropriate AI use and provide targeted support. • Stakeholder Engagement
Co-Design with Educators and Students	• Involve teachers and learners directly in the design process through workshops, prototype testing, and iterative feedback cycles. This ensures the final system addresses real classroom needs and challenges.
Community Guidelines Development	• Facilitate collaborative creation of AI use policies that reflect shared values and educational goals. • Make these guidelines living documents that evolve with experience and changing technology. • Ethical Safeguards
Bias Detection and Mitigation	• Build in regular auditing processes to identify and address potential biases in AI recommendations or feedback. Ensure diverse representation in training data and testing scenarios.
Privacy by Design	• Implement strong data protection measures that give users control over their information while still enabling effective learning analytics and system improvement.

After the brainstorming process is complete, all core components and elements the voted process in the NGT process. The following are the results of the process voting that are analyzed using the NGT-PLUS software:



Items / Elements	Voter1	Voter2	Voter3	Voter4	Voter5	Voter6	Voter7	Total item score	Percentage
Learner Agency and Transparency	3	3	3	3	3	3	3	21	100.00
Educator Empowerment	3	3	2	3	3	3	3	20	95.24
Progressive Disclosure of AI Capabilities	3	3	3	3	3	3	3	21	100.00
Built-in Reflection Mechanisms	2	2	3	3	3	3	3	19	90.48
Collaborative Learning Design	3	3	3	3	3	3	2	20	95.24
Authentic Assessment Integration	3	3	3	3	1	3	3	19	90.48
Continuous Learning Analytics	2	3	3	3	3	3	3	20	95.24
Co-Design with Educators and Students	3	3	3	3	3	3	3	21	100.00
Community Guidelines Development	3	3	2	3	3	3	3	20	95.24
Bias Detection and Mitigation	3	3	3	2	3	3	2	19	90.48
Privacy by Design	3	3	3	3	3	3	3	21	100.00

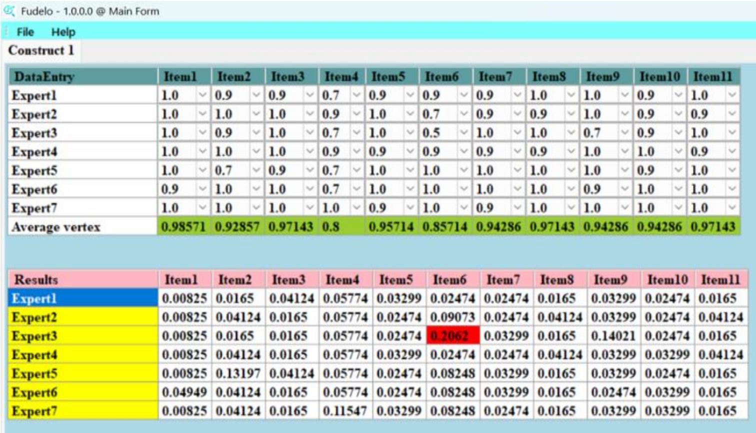
Fig. 1: NGT Plus Data Entry.

Table 3: NGT Plus Analysis Output (Voting Process)

Items/Elements	Voter1	Voter2	Voter3	Voter4	Voter5	Voter6	Voter7	Total item score	Percentage	Rank Priority	Voter Consensus
Learner Agency and Transparency	3	3	3	3	3	3	3	21	100	1	Suitable
Educator Empowerment	3	3	2	3	3	3	3	20	95.24	2	Suitable
Progressive Disclosure of AI Capabilities	3	3	3	3	3	3	3	21	100	1	Suitable
Built-in Reflection Mechanisms	2	2	3	3	3	3	3	19	90.48	3	Suitable
Collaborative Learning Design	3	3	3	3	3	3	2	20	95.24	2	Suitable
Authentic Assessment Integration	3	3	3	3	1	3	3	19	90.48	3	Suitable
Continuous Learning Analytics	2	3	3	3	3	3	3	20	95.24	2	Suitable
Co-Design with Educators and Students	3	3	3	3	3	3	3	21	100	1	Suitable
Community Guidelines Development	3	3	2	3	3	3	3	20	95.24	2	Suitable
Bias Detection and Mitigation	3	3	3	2	3	3	2	19	90.48	3	Suitable
Privacy by Design	3	3	3	3	3	3	3	21	100	1	Suitable

The comprehensive evaluations for the primary construct, as per the expert's assessment, are presented in Table 3. The study's findings indicate that all evaluated item percentages are appropriate for use. Dobbie et al. (2004), Mustapha et al. (2022), and Deslandes, Mendes, Pires, and Campos (2010) all determined that the proportion must exceed 70%. Analysis of the participants' responses indicates that the model's fundamental elements are appropriate for implementation.

After the NGT process was completed, we proceeded to the second phase of the model using the Fuzzy Delphi method. Data analysis was performed after expert questions were received from experts. All study data were analyzed using FUDELO software, specially created to analyze Fuzzy Delphi data. Data analysis is as follows:



DataEntry	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11
Expert1	1.0	0.9	0.9	0.7	0.9	0.9	0.9	1.0	1.0	0.9	1.0
Expert2	1.0	1.0	1.0	0.9	1.0	0.7	0.9	0.9	1.0	0.9	0.9
Expert3	1.0	0.9	1.0	0.7	1.0	0.5	1.0	1.0	0.7	0.9	1.0
Expert4	1.0	1.0	1.0	0.9	0.9	0.9	0.9	0.9	1.0	1.0	0.9
Expert5	1.0	0.7	0.9	0.7	1.0	1.0	1.0	1.0	1.0	0.9	1.0
Expert6	0.9	1.0	1.0	0.7	1.0	1.0	1.0	1.0	0.9	1.0	1.0
Expert7	1.0	1.0	1.0	1.0	0.9	1.0	0.9	1.0	1.0	1.0	1.0
Average vertex	0.98571	0.92857	0.97143	0.8	0.95714	0.85714	0.94286	0.97143	0.94286	0.94286	0.97143

Results	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11
Expert1	0.00825	0.0165	0.04124	0.05774	0.03299	0.02474	0.02474	0.0165	0.03299	0.02474	0.0165
Expert2	0.00825	0.04124	0.0165	0.05774	0.02474	0.09073	0.02474	0.04124	0.03299	0.02474	0.04124
Expert3	0.00825	0.0165	0.0165	0.05774	0.02474	0.2062	0.03299	0.0165	0.14021	0.02474	0.0165
Expert4	0.00825	0.04124	0.0165	0.05774	0.03299	0.02474	0.02474	0.04124	0.03299	0.03299	0.04124
Expert5	0.00825	0.13197	0.04124	0.05774	0.02474	0.08248	0.03299	0.0165	0.03299	0.02474	0.0165
Expert6	0.04949	0.04124	0.0165	0.05774	0.02474	0.08248	0.03299	0.0165	0.02474	0.03299	0.0165
Expert7	0.00825	0.04124	0.0165	0.11547	0.03299	0.08248	0.02474	0.0165	0.03299	0.03299	0.0165

Fig. 2: Data Entry in the FUDELO Software.

Table 4: Result

Results	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11
Expert1	0.008	0.017	0.041	0.058	0.033	0.025	0.025	0.017	0.033	0.025	0.017
Expert2	0.008	0.041	0.017	0.058	0.025	0.091	0.025	0.041	0.033	0.025	0.041
Expert3	0.008	0.017	0.017	0.058	0.025	0.206	0.033	0.017	0.140	0.025	0.017
Expert4	0.008	0.041	0.017	0.058	0.033	0.025	0.025	0.041	0.033	0.033	0.041
Expert5	0.008	0.132	0.041	0.058	0.025	0.082	0.033	0.017	0.033	0.025	0.017
Expert6	0.049	0.041	0.017	0.058	0.025	0.082	0.033	0.017	0.025	0.033	0.017
Expert7	0.008	0.041	0.017	0.115	0.033	0.082	0.025	0.017	0.033	0.033	0.017
Statistics	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11
Value of the item	0.014	0.047	0.024	0.066	0.028	0.085	0.028	0.024	0.047	0.028	0.024
Value of the "d" construct	0.03771										
Item < 0.2	7	7	7	7	7	6	7	7	7	7	7
% of item < 0.2	100%	100%	100%	100%	100%	85%	100%	100%	100%	100%	100%
Average of % consensus	98										
Defuzzification	0.986	0.929	0.971	0.800	0.957	0.857	0.943	0.971	0.943	0.943	0.971
Ranking	1	5	2	7	3	6	4	2	4	4	2
Status	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

This study reveals robust expert consensus on all 11 assessed items (see Table 4), as indicated by the fuzzy Delphi results table, with an overall average consensus of 98% and all items attaining acceptance status. The investigation engaged seven experts who assessed items utilizing fuzzy set theory, with values approaching 1.0 signifying greater agreement and values below 0.2 denoting robust consensus thresholds. All items, except Item 6, attained 100% unanimity (7 out of 7 experts rated below 0.2); however, Item 6 obtained 85% consensus (6 out of 7 experts). The defuzzification method, which translates fuzzy numbers into precise values for ranking, indicates that Item 1 attained the greatest ranking (defuzzification value: 0.986), followed by Item 3 (0.971) and Item 5 (0.957), while Item 4 ranked lowest (0.800), albeit remaining within acceptable limits. A threshold value of 0.2 and a construct validity measure of 0.03771 signify strong statistical reliability. This methodology conforms to recognized fuzzy Delphi techniques as described by Chang, Hsu, and Chang (2011) and Hasson, Keeney, and McKenna (2000), wherein expert consensus is attained through successive rounds until statistical convergence is achieved, generally necessitating consensus levels exceeding 75% for validation in social science research.

Based on this NGT and fuzzy result, the findings of this study have been certified by experts. We conclude that in the final output of the study, as in Figure 3:

Fig. 3: The Final HCD Approach.

This diagram depicts a thorough human-centered design framework aimed at preventing AI misuse in educational contexts, with a central blue circle symbolizing the primary emphasis on learners, educators, and educational results. Five interrelated components emanate from this center: Core Design Principles (red) highlight learner agency and transparency, alongside educator empowerment via explicit AI indicators and detailed dashboards; Implementation Strategies (green) include progressive disclosure of AI functionalities, integrated reflection mechanisms, and collaborative learning methodologies; Assessment & Feedback systems (orange) incorporate authentic assessment techniques and ethical learning analytics; Stakeholder Engagement (purple) encompasses co-design processes and the formulation of community guidelines; and Ethical Safeguards (dark blue) focus on bias detection and privacy protection. The dashed arrows linking all components signify ongoing feedback loops, illustrating that this framework functions as an integrated system in which each element informs and fortifies the others, ultimately ensuring that AI tools augment rather than supplant human learning while preserving educational integrity and ethical standards.

4. Discussion

The results of this study show that experts overwhelmingly agree on the proposed Human-Centered Design (HCD) framework for reducing AI misuse in schools. The fuzzy Delphi analysis shows that there is a remarkable 98% average agreement across all evaluated components. The two-phase methodology, which included the Nominal Group Technique (NGT) and the Fuzzy Delphi Method (FDM), successfully tested 11 core elements that were grouped into five related dimensions: Core Design Principles, Implementation Strategies, Assessment & Feedback Systems, Stakeholder Engagement, and Ethical Safeguards. It is especially impressive that elements like "Learner Agency and Transparency," "Progressive Disclosure of AI Capabilities," "Co-Design with Educators and Students," and "Privacy by Design" all got 100% expert agreement. Even the lowest-ranked element (Item 4) had a defuzzification value of 0.800, which is well above the 0.5 threshold for expert consensus. This strong validation fits well with the known fuzzy Delphi methods (Chang, Hsu, & Chang, 2011; Hasson, Keeney, & McKenna, 2000) and gives strong empirical proof that the framework can be used in real life to deal with the difficult problems that come up when AI is used in education. The importance of practical tactics for implementation, going beyond theoretical concepts, is shown by the near-unanimous acceptance of the Human-Centered Design framework among experts (98 percent agreement). This is especially true about "Privacy by Design" and other ethical protections. Data minimization practices that restrict student information collection to educationally necessary purposes (Cavoukian, 2009), algorithmic auditing protocols that routinely test for demographic bias in AI-powered grading systems (Barocas & Selbst, 2016), and transparent consent processes that clearly explain to students and parents how their data will be used and protected (Regan & Jesse, 2019) are specific mechanisms that are necessary for practical implementation of the framework's identified key ethical dimensions. As an example, frequent statistical monitoring of AI suggestion patterns across various student demographics might be used for bias detection. When differences reach specified levels, remedial steps could be necessary (Mehrabi et al., 2021). Dwork (2008) suggests using statistical noise to mask students' identities while keeping aggregate learning analytics intact; Pardo and Siemens (2014) suggest establishing student data review boards that include parents, educators, and privacy experts as a governance structure to further ensure privacy. According to Holstein et al. (2019), educational institutions should focus on creating standardized assessment tools, staff training programs, and incident response protocols to put these ethical principles into action. This will ensure that student data is adequately protected in real-world educational AI deployments, as there is unanimous agreement among experts on these safeguards.

The study shows that there is a lot of agreement among experts on the Human-Centered Design framework for avoiding AI misuse in schools (98% average agreement), but there will be a lot of obstacles to putting this framework into reality, so it's worth thinking about. One of the main challenges is the lack of funding for comprehensive AI literacy training, infrastructure upgrades, and ongoing professional development. Schools often have limited budgets and may not be able to afford these things, which are necessary for elements like "Co-Design with Educators and Students" and "Progressive Disclosure of AI Capabilities." Another potential roadblock to adoption could be pushback from teachers who aren't well-versed in AI tools. This is especially problematic because stakeholders' support is crucial for a smooth rollout, and some of these stakeholders may be wary of or overwhelmed by the rate of technological change. Although there is unanimous agreement among experts, schools without technical knowledge or with conflicting pedagogical agendas may find it difficult to implement the framework's "Learner Agency and Transparency" and "Privacy by Design" principles. Despite the strong validation of the theoretical framework, these implementation barriers highlight the need to carefully consider institutional capacity, change management strategies, and ongoing support systems to translate the eleven core elements across the five dimensions into sustainable practice. This will help bridge the gap between expert consensus and real-world educational contexts.

This study fills an important gap in the literature on educational technology by moving from reactive detection-based methods to proactive prevention-oriented strategies for reducing AI misuse. This work is important because it responds quickly to the urgent problems that schools are facing. Recent studies show that 89% of students know about ChatGPT and 43% have used AI tools for schoolwork (Nerdy Nav, 2024; Forbes Analytics, 2024). However, current AI detection systems are very limited, with accuracy rates as low as 26% for finding AI-generated content (Kirchner et al., 2023). The framework's focus on human-centered design principles directly addresses the growing concerns about stakeholder exclusion in AI system development, which Martinez-Maldonado et al. (2024) identify as a primary cause of mistrust and misalignment in educational AI tools. Moreover, the study advances the nascent domains of Human-Centered Learning Analytics (HCLA) and Human-Centered Artificial Intelligence (HCAI) by offering a systematic, empirically substantiated methodology that emphasizes learner autonomy, educator empowerment, and ethical principles over mere technological remedies, thus fostering the overarching initiative for responsible AI integration in education (Topali et al., 2024; Khosravi et al., 2024).

5. Future research directions and implementation implications

The validated HCD framework opens several promising avenues for future research and practical implementation. Primary research priorities should include longitudinal studies examining the framework's effectiveness in real educational settings, with particular attention to measuring long-term impacts on learning outcomes, academic integrity, and student digital literacy development, as suggested by recent systematic reviews (Yusuf et al., 2024; MDPI Information, 2025). Additionally, future investigations should explore the framework's adaptability across diverse educational contexts, including different cultural settings, age groups, and subject domains, building upon the cross-cultural validation approaches used in educational technology research (Ateeq et al., 2024). The development of practical implementation guidelines and training protocols for educators represents another critical research direction, particularly given the need for comprehensive faculty development programs identified in recent literature (Cornell CTI, 2024; University of Pittsburgh Teaching Center, 2024). Furthermore, researchers should investigate the integration of emerging technologies such as explainable AI and blockchain-based credentialing systems within the HCD framework while also examining the economic and policy implications of large-scale implementation. Finally, given the rapid evolution of AI capabilities, continuous validation and refinement of the framework through iterative expert panels and stakeholder feedback mechanisms will be essential to maintain its relevance and effectiveness in addressing evolving challenges in educational AI integration (Balalle & Pannilage, 2025; Eaton, 2024).

The results also show that the Human-Centered Design paradigm has far-reaching consequences for dealing with complex instances of AI abuse in classrooms, going well beyond the realm of typical academic dishonesty. Countering the threat of AI-driven essay mills, the framework's "Progressive Disclosure of AI Capabilities" and "Learner Agency and Transparency" components ensure students understand the limitations and appropriate usage boundaries of AI tools before they gain access to advanced features. These components achieved 100% expert consensus. To address deepfake concerns, the framework has implemented robust protection mechanisms such as algorithmic auditing protocols and "Privacy by Design" safeguards. Data minimization practices limit the collection of student biometric and personal information, which could be used to create deepfakes. Additionally, routine demographic bias testing protocols can be adjusted to detect patterns of synthetic content in student submissions (Mehrabi et al., 2021). To provide a human-centered early warning system that supplements the technological protections, the "Co-Design with Educators and Students" approach makes sure that both groups are actively involved in spotting emergent abuse trends. Instead of depending only on detection technologies—which, as the study points out, have accuracy rates as low as 26% for AI-generated content—the framework's stakeholder engagement dimension tackles the underlying cause of these complex misuse cases by encouraging digital literacy and an understanding of ethical AI. Academic integrity and the responsible use of AI's educational benefits can be achieved through a collaborative educational process that involves students, educators, and technology. This proactive and prevention-oriented approach turns the traditional game of cat-and-mouse between AI misuse and detection into something more constructive.

Acknowledgment

Here we would like to give UiTM Pahang a highlight for their support until the completion of this study. The same goes for all researchers who have made this study a success.

References

[1] Adler, M., & Ziglio, E. (Eds.). (1996). Gazing into the oracle. Bristol, PA: Kingsley
[2] ArtSmart AI. (2024, February 23). AI plagiarism statistics 2025: Transforming academic integrity. https://artsmart.ai/blog/ai-plagiarism-statistics/
[3] Ateeq, A., Alzoraiki, M., Milhem, M., & Ateeq, R. A. (2024). Artificial intelligence in education: Implications for academic integrity and the shift toward holistic assessment. *Frontiers in Education*, 9. https://doi.org/10.3389/feduc.2024.1470979.
[4] Balalle, H., & Pannilage, S. (2025). Reassessing academic integrity in the age of AI: A systematic literature review on AI and academic integrity. *Social Sciences & Humanities Open*, 11, 101299. https://doi.org/10.1016/j.ssaho.2025.101299.
[5] Benitez, J. M., Martín, J. C., & Román, C. (2007). Using Fuzzy Numbers for Measuring Quality of Service in the Hotel Industry. *Tourism Management*, 28, 544-555. https://doi.org/10.1016/j.tourman.2006.04.018.
[6] Bin-Nashwan, S. A., Al-Daihani, M., Al-Hajri, M., & Al-Harthy, A. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*, 75, 102370. https://doi.org/10.1016/j.techsoc.2023.102370.
[7] Boddy, C. (2012), "The Nominal Group Technique: an aid to Brainstorming ideas in research", *Qualitative Market Research*, Vol. 15 No. 1, pp. 6-18. https://doi.org/10.1108/13522751211191964.
[8] Bodjanova, S. (2006). Median Alpha-Levels of a Fuzzy Number. *Fuzzy Sets and Systems*, 157, 879-891. https://doi.org/10.1016/j.fss.2005.10.015.
[9] Booker, J. M., & McNamara, L. A. (2004). Solving Black Box Computation Problems Using Expert Knowledge Theory and Methods. *Reliability Engineering & System Safety*, 85, 331-340. https://doi.org/10.1016/j.ress.2004.03.021.
[10] Cantrill, J. A., Sibbald, B., & Buetow, S. (1996). The Delphi and Nominal Group Techniques in Health Services Research. International Journal of Pharmacy Practice, 4, 67-74. https://doi.org/10.1111/j.2042-7174.1996.tb00844.x.
[11] Chang, P. L., Hsu, C. W., & Chang, P. C. (2011). Fuzzy Delphi Method for Evaluating Hydrogen Production Technologies. *International Journal of Hydrogen Energy*, 36, 14172-14179. https://doi.org/10.1016/j.ijhydene.2011.05.045.
[12] Chen, M. Y., & Chen, B. T. (2014). Online Fuzzy Time Series Analysis Based on Entropy Discretization and a Fast Fourier Transform. *Applied Soft Computing*, 14, 156-166. https://doi.org/10.1016/j.asoc.2013.07.024.
[13] Chen, S. M. (2002). Forecasting Enrollments Based on High-Order Fuzzy Time Series. *Cybernetics and Systems*, 33, 1-16. https://doi.org/10.1080/019697202753306479.
[14] Chen, M. Y., & Chen, B. T. (2014). Online Fuzzy Time Series Analysis Based on Entropy Discretisation and a Fast Fourier Transform. *Applied Soft Computing*, 14, 156-166. https://doi.org/10.1016/j.asoc.2013.07.024.
[15] Clare, J., Walker, R., & Rundle, K. (2023). Linking artificial intelligence facilitated academic misconduct to existing prevention frameworks. *International Journal for Educational Integrity*, 19(1), 1-20. https://doi.org/10.1007/s40979-023-00142-3.
[16] Community College Daily. (2024, October 10). AI in education: Moving beyond academic dishonesty. https://www.ccdaily.com/2024/10/ai-in-education-moving-beyond-academic-dishonesty/.
[17] Cornell Center for Teaching Innovation. (2024). AI & academic integrity. https://teaching.cornell.edu/generative-artificial-intelligence/ai-academic-integrity.
[18] Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148.
[19] Dalkey, N., & Helmer, O. (1963). An Experimental Application of the Delphi Method to the Use of Experts. *Management Science*, 9, 458-467. https://doi.org/10.1287/mnsc.9.3.458.
[20] Deslandes, S. F., Mendes, C. H. F., Pires, T. D. O., & Campos, D. D. S. (2010). Use of the Nominal Group Technique and the Delphi Method to draw up evaluation indicators for strategies to deal with violence against children and adolescents in Brazil. *Revista Brasileira de Saúde Materno Infantil*,10, s29-s37. https://doi.org/10.1590/S1519-38292010000500003.
[21] Dobbie, A., Rhodes, M., Tysinger, J. W., & Freeman, J. (2004). Using a modified nominal group technique as a curriculum evaluation tool. *FAMILY MEDICINE*-KANSAS CITY-, 36, 402406.
[22] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management*, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642.
[23] Eaton, S. E. (2024). Future-proofing integrity in the age of artificial intelligence and neurotechnology: Prioritising human rights, dignity, and equity. *International Journal for Educational Integrity*, 20(1), 1-18. https://doi.org/10.1007/s40979-024-00175-2.
[24] EdTech Digest. (2024, November 5). Beyond AI detection: Rethinking our approach to preserving academic integrity. https://www.edtechdigest.com/2024/11/05/beyond-ai-detection-rethinking-our-approach-to-preserving-academic-integrity/.
[25] Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the authenticity of ChatGPT responses: A study on text-matching capabilities. *International Journal for Educational Integrity*, 19(1), 1-17. https://doi.org/10.1007/s40979-023-00137-0.
[26] Forbes Analytics. (2024). AI plagiarism statistics and trends in education. *Forbes Education*.
[27] Fortemps, P., & Roubens, M. (1996). Ranking and defuzzification methods based on area compensation. *Fuzzy Sets and Systems*. https://doi.org/10.1016/0165-0114(95)00273-1.

[28] Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*. https://doi.org/10.1046/j.1365-2648.2000.01567.x.

[29] GSI Education. (2024, May 2). AI in education: Embracing a human-centered approach to learning. https://www.gsineducation.com/blog/ai-in-education-embracing-a-human-centered-approach-to-learning

[30] Garai, A. (2013). Weighted Intuitionistic Fuzzy Delphi Method. *Journal of Global Research in Computer Science*, 4, 38-42.

[31] Kirchner, J. H., Ahmad, L., Aaronson, S., & Leike, J. (2023). *New AI classifier for indicating AI-written text*. OpenAI Blog. https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text

[32] Khosravi, H., Martinez-Maldonado, R., Shum, S. B., Gasevic, D., Siemens, G., Slade, S., ... & Hatala, M. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers & Education*, 191, 104-118.

[33] Hinduja, S. (2024). Lessons learned from ten generative AI misuse cases. *Cyberbullying Research Centre*.

[34] Hsu, Y. L., Lee, C. H., & Kreng, V. B. (2010). The Application of Fuzzy Delphi Method and Fuzzy AHP in Lubricant Regenerative Technology Selection. Expert Systems with Applications, 37, 419-425. https://doi.org/10.1016/j.eswa.2009.05.068.

[35] Gallagher, M., Hares, T., Spencer, J., Bradshaw, C., & Webb, I. (1993). The nominal group technique: A research tool for general practice? *Family Practice*, 10(1), 76-81. https://doi.org/10.1093/fampra/10.1.76.

[36] Giacomin, J. (2014). What Is Human Centred Design? *The Design Journal*, 17(4), 606-623. https://doi.org/10.2752/175630614X14056185480186.

[37] GSI Education. (2024, May 2). AI in education: Embracing a human-centred approach to learning. https://www.gsineducation.com/blog/ai-in-education-embracing-a-human-centered-approach-to-learning.

[38] Inspera. (2024, April 25). 7 examples of AI misuse in education. https://www.inspera.com/ai/examples-of-ai-misuse-in-education/.

[39] Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R., & Mieno, H. (1993). The Max-Min Delphi Method and Fuzzy Delphi Method via Fuzzy Integration. *Fuzzy Sets and Systems*, 55, 241-253. https://doi.org/10.1016/0165-0114(93)90251-C.

[40] Khosravi, H., Martinez-Maldonado, R., Shum, S. B., Gasevic, D., Siemens, G., Slade, S., ... & Hatala, M. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers & Education*, 191, 104-118.

[41] Lindenwood University. (2024, December 4). Human-centred AI: What is human-centric artificial intelligence? https://online.lindenwood.edu/blog/human-centered-ai-what-is-human-centric-artificial-intelligence/.

[42] Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 100779. https://doi.org/10.1016/j.patter.2023.100779.

[43] Marchal, N., Xu, R., Elasmar, R., Gabriel, I., Goldberg, B., & Isaac, W. (2024). Generative AI misuse: A taxonomy of tactics and insights from real-world data. *Tech Policy Press*. https://www.techpolicy.press/researchers-develop-taxonomy-of-generative-ai-misuse-from-real-world-data/.

[44] Martinez-Maldonado, R., Gasevic, D., Echeverria, V., Fernandez-Nieto, G., Swiecki, Z., & Buckingham Shum, S. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers & Education*, 204, 105-123.

[45] McKinsey. (2024, December 20). The case for human-centred AI. *McKinsey Digital*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-case-for-human-centered-ai.

[46] MDPI Information. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information*, 16(4), 296. https://doi.org/10.3390/info16040296.

[47] Mustapha, R., Ibrahim, N., Mahmud, M., Jaafar, A. B., Ahmad, W. A. W., & Mohamad, N. H. (2022). Brainstorming the Students' Mental Health after Covid-19 Outbreak and How to Curb from Islamic Perspectives: Nominal Group Technique Analysis Approach. *International Journal of Academic Research in Business and Social Sciences,* 12(2), 90–99 https://doi.org/10.6007/IJARBSS/v12-i2/12367.

[48] Mustapha, R., & Darusalam, G. (2022). *Pendekatan Kajian Rekabentuk dan Pembangunan dalam Kajian Kontemporari.* Universiti Malaya Press.

[49] Murray, T. J., Pipino, L. L., & Van Gigch, J. P. (1985). A Pilot Study of Fuzzy Set Modification of Delphi. *Human Systems Management*, 5, 76-80. https://doi.org/10.3233/HSM-1985-5111.

[50] Mullen, P. M. (2003). Delphi: Myths and Reality. *Journal of Health Organization and Management*, 17, 37-52. https://doi.org/10.1108/14777260310469319.

[51] Nerdy Nav. (2024, January 5). ChatGPT cheating statistics & impact on education (2024). https://nerdynav.com/chatgpt-cheating-statistics/.

[52] Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers & Education*, 168, 104-116.

[53] Stanford Graduate School of Education. (2024). What do AI chatbots mean for students and cheating? https://ed.stanford.edu/news/what-do-ai-chatbots-really-mean-students-and-cheating.

[54] Thomaidis, N. S., Nikitakos, N., & Dounias, G. D. (2006). The Evaluation of Information Technology Projects: A Fuzzy Multicriteria Decision-Making Approach. *International Journal of Information Technology & Decision Making*, 5, 89-122. https://doi.org/10.1142/S0219622006001897.

[55] Topali, P., Ortega-Arranz, A., Rodríguez-Triana, M. J., Er, E., Khalil, M., & Akçapınar, G. (2024). Designing human-centered learning analytics and artificial intelligence in education solutions: A systematic literature review. *Behaviour & Information Technology*, 43(6), 1071-1098. https://doi.org/10.1080/0144929X.2024.2345295.

[56] University of Pittsburgh Teaching Center. (2024). Generative AI: Encouraging academic integrity. https://teaching.pitt.edu/resources/encouraging-academic-integrity/.

[57] Wintmölle, F., Meier, M., & Maier, C. (2024). Artificial intelligence (AI) misuse: Perpetrators' moral justifications. *ICIS 2024 Proceedings*, 7.

[58] Yusuf, M., Quispe-Prieto, S., Kerestedjieva, S., Zhai, X., Rienties, B., & Whitelock, D. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, 12(2), e3489. https://doi.org/10.1002/rev3.3489.