



**Journal of Educational
Psychology and Pedagogical
Sciences (JEPPS)**

**ISSN:2791-0393 (Print) eISSN:
2791-0407**

**Vol.6, No. 1, (Jan-Jun, 2026): 1-
14**

Submitted 04 Mar 2026

Accepted 20 Mar 2026

Published 30 Mar 2026

DOI: [https://doi.org/
10.52587/jepps.v6i1.137](https://doi.org/10.52587/jepps.v6i1.137)

<https://jepps.su.edu.pk/article/55>

OPEN ACCESS

Safeguarding Academic Integrity Through AI Literacy Program

Mehwish Raza ^{1*}

^{1*} Department of Education, Forman Christian College
University, Lahore. E-mail: mehwishraza@fccollege.edu.pk

ABSTRACT

The frequent reliance of students on generative AI tools in Pakistan's Higher Education Institutions (HEIs) has intensified issues concerning academic integrity due to unacknowledged AI authorship, cognitive offloading on AI tools, incorrect factual information and citations generated by AI. These challenges are amplified by limited digital literacy, insufficient institutional regulations directing AI use as a learning partner, and a fragmented understanding of ethical issues arising from excessive AI use among university students. This study addresses this gap by systematically developing an AI Literacy (AILP) for a public-sector university in Pakistan. The study reports development of Students' Learning Outcomes (SLOs) for AILP by using a sequential mixed-methods design across three research phases. Two evidence bases were generated by following a Systematic Literature Review (SLR) of selected literature (n=56). These included the key literacies essential for responsible interaction with AI and a lexicon of ethical issues often resulting in academic integrity challenges. Emerging themes were identified through subjecting the lexicon to bibliometric co-occurrence analysis using VOSViewer. The SLOs were drafted in line with backward design principles and validated by a team of experts in four iterative review cycles using Delphi review process. The final version of SLOs presented 86% consensus among reviewers and provided a curricular foundation for a hands-on, activity-based curriculum divided into four modules. The AILP will be delivered as a mandatory 0-credit course starting Fall26 through the Moodle learning management system, making this program Pakistan's first mandatory AI literacy course for university students, developed with an aim of creating more aware and proficient users of AI. This study provides a blueprint for contextualizing AI literacy into an actionable, scalable pedagogical intervention for students across age groups. This study is particularly generalizable for resource-constrained and policy-sparse educational environments.

Keywords: AI Literacy, Academic Integrity, bibliometric co-occurrence analysis, Delphi review

1. Introduction

Academic integrity has been a longstanding concept broadly attributed to a set of values

characterized by truthfulness, honesty, trust, fairness, respect, and responsibility in an academic setting, applicable to students, teachers, and researchers alike (Gunton, 2022). Academic integrity is the cornerstone of ethical practices and essential for maintaining the credibility of educational institutions. Well-articulated academic integrity policies help limit incidents of academic misconduct, such as plagiarism, collusion, and cheating. Plagiarism is considered the fundamental breach of academic integrity; its long history can be traced back to 80AD when Roman poet Martial reported the theft of his intellectual work. He used the Latin word *plagiarius*, meaning kidnapper, for the first time, which later became 'plagiarism' in English, as used by Ben Jonson in 1601. The unprecedented benefits of Artificial Intelligence (AI) have led to a rapid adoption by the Higher Education Institutions (HEIs) globally across diverse administrative and academic operations. HEIs in the Global South have also adopted a spectrum of AI tools for different operations, depending on the institution's individual digital maturity, which is determined by institutions' digital infrastructure, user competence, and well-articulated AI guidelines for active stakeholders (Maimela & Mbonde, 2025). In Pakistan, only a few private universities have smart institutional administrative systems, digital governance and pedagogy, and immersive assessment culture, while a large majority of universities are not prepared to systematically adopt, regulate, and foster a responsible use of AI (Azeem et al., 2025). Public sector universities, in particular, cater to large student populations while struggling with campus-wide reliable broadband, sufficient computing hardware, data-processing capacity, and essential digital competence (Farheen et al. 2025).

Despite limited access to AI tools on campuses, a large majority of students at public universities have personal subscriptions to Large Language Models (LLMs) such as ChatGPT, resulting in uneven access compared to faculty, who typically rely on the institutional resources. Faculty are generally unwilling to purchase independent subscriptions due to the high costs and the lack of institutional support (Farheen *et al.* 2025). Research asserts that students demonstrate high exposure and functional use of AI by actively using it as an academic assistant in various forms (Azeem et al., 2025). This situation is replicated in countries mirroring a similar digital context, such as Namibia, Brazil, and Thailand (Seo et al., 2025)

The uneven access to and familiarity with AI tools between faculty and students creates the Matthew effect, resulting in power imbalances and undermining academic integrity. Students increasingly use LLMs for assignment help, brainstorming, and even exam preparation, while many faculty do not have access to technologies that precisely track AI content in students' assignments (Cotton, Cotton, & Shipway, 2024; Inuwa-Dutse, 2023). This disparity increases unauthorised use of AI for academic tasks, which falls into the category of plagiarism. The accessibility limitations also hinder discussions between students and faculty about the functionalities and ethical ambiguities related to AI use (Chan, 2023). This epistemic and procedural asymmetry between faculty and students presents AI as a shortcut to learning, leading to over-reliance and cognitive offloading on AI tools. Studies indicate that over-reliance on AI can erode writing and critical thinking skills, the originality of ideas, individual creativity, and overall human intellectual development, which are the core purposes of education (Luckin et al., 2024). Research also suggests that AI-generated content, data, and citations may be inaccurate; therefore, users should verify their validity, reliability, and accuracy.

The limited accessibility of AI is further compounded by the lack of professional development opportunities for faculty to adapt pedagogy and assessment design to the rapid developments in AI. Studies indicate that process-oriented and reflective assessment promote higher-order skills and authentic responses. Additionally, faculty should also have specific tracking software to precisely trace AI-generated output in students' assignments (Suchithra & Arya, 2025).

Despite all contextual challenges, academicians cannot deny the transformative affordances of AI for both the students and faculty (Dignum, 2017). In addition to personalising learning, LLMs facilitate students' metacognition and self-regulation. In addition to several other benefits, AI-driven language models particularly offer valuable learning opportunities for learners of English as a second language by providing them with a range of writing styles, new vocabulary, correct grammatical expressions, and instant feedback (Ramdhani, Husnawadi, & Waliyadin, 2025). LLMs can assist students through various stages of their academic tasks, from idea generation to the final write-up (Cotton et al., 2024). LLMs also offer unique advantages to faculty to innovate curriculum, resources, assessment, and teaching strategies; thus, significantly reducing their workload (Sharma, 2025). AI tools provide insights into intricate aspects of student performance through data analytics, significantly accelerating the identification of unique learning needs and the provision of personalized support (Cotton et al., 2024). Both students and faculty benefit from improved efficiency and access to innovative resources, such as image and text generators, video-making tools, data analysis tools, literature review management systems, and bots for interacting with PDFs (Azeem et al., 2025). Higher education faculty around the world utilise freemium packages of various AI tools within existing academic operations to significantly enhance the efficiency of administrative and pedagogical processes (Chan, 2023).

Despite the significant academic advantages of AI, the development and training of state-of-the-art algorithms are predominantly navigated by the Global North, leaving a large majority in the Global South as mere consumers of this technology, fostering a dependency on Western infrastructure while lacking the necessary local regulatory frameworks to manage and adapt these innovations according to their needs and context (Maimela & Mbonde, 2025). AI has introduced a new form of digital colonialism, exerting influence over societies, cultures, and identities in the Global South, further exacerbating the existing power imbalances.

To avoid plagiarism, institutions globally have well-defined academic integrity principles to foster original thinking, reflection, and productivity in academic work (Gonsalves, 2025). It is of paramount importance for educational institutions to uphold high standards of academic integrity to maintain public trust in them as key drivers of social development. AI has heightened the threats to human agency and created more complexity in academic integrity issues by blurring the line between student effort and machine output, raising new forms of misconduct referred to as “AI-giarism,” which differs from conventional plagiarism but is equally harmful to learning and the development of human potential (Mijwil, et al., 2023).

Research asserts that users' digital literacy is the gateway to operationalizing the benefits of AI while adhering to institutional academic integrity policies. Digital literacy is demonstrated through users' ability to access, manage, understand, evaluate, integrate, communicate, and create information using digital technologies in a secure, ethical, and effective manner, with awareness of ethical issues and possible solutions (Figaredo & Stoyanovich, 2023). Unlike technologically advanced countries, nations struggling with digital inequalities and low literacy rates often lack high digital competence.

Technology misuses increase when educational institutions lack clear and well-defined policies for the ethical use of technology, including AI (Chan, 2023). Pakistan's public sector universities lack a federal policy on the ethical use of AI, leaving each institution to establish its own guidelines for regulating AI applications in academic and administrative operations. Faculty members have the autonomy to create course-specific AI policies to maintain high standards of academic integrity.

Given the complex landscape of digital inequalities between faculty and students, the absence of institutional policies, and the lack of digital infrastructure, raising awareness among students about the ethical use of AI becomes of paramount importance (Grassini, 2023).

A responsible use of AI necessitates users' critical literacy of prompt engineering, validation of information, understanding, and reacting to AI-driven ethical dilemmas, acknowledging AI use through citations, using AI to promote self-regulation, critical thinking, and creativity (Suchithra & Arya, 2025). All universities in Pakistan must strictly adhere to the degree roadmaps approved by the federal regulatory authority, Higher Education Commission (HEC); therefore, it is not a common practice to offer courses outside the scope of the approved degree roadmap. Like several other countries, in Pakistan, the courses related to digital competence are often restricted to Computer Science disciplines and remain largely inaccessible to students from other fields (Luckin et al., 2024).

There is evidence from universities in Kenya and El Salvador that have benefited from AI ethics modules, leading to improved quality research and academic performance (Ortiz et al., 2025). General Education courses on the ethical use of AI reflect institutional accountability by preparing aware users with the essential skills to proactively mitigate AI-induced academic integrity challenges. Such courses are beneficial for students, as they foster AI literacy and empower them to build a shared understanding of their university's academic integrity policies (Dignum, 2017). This study provides insights into the development and validation of Students' Learning Outcomes (SLOs) for a non-credit AI literacy course at a large public sector university in Pakistan. The "AI Literacy Program" (hereafter AILP) has been designed to enhance students' AI literacy and address the rising incidents of academic integrity violations caused by unethical AI use.

As a mandatory requirement for all new courses, AILP will be subjected to approval from the university's Academic Council. After these internal approvals, AILP will be offered as a compulsory 0-credit course through Moodle. AILP will be Pakistan's first official public university course aimed at promoting students' AI literacy, and it will be a mandatory requirement for students across all degree programmes.

The research is directed by the following two key questions:

1. Which essential user competencies are reported in the literature to direct a responsible and ethical use of AI that harnesses human potential in the context of developing economies?
2. What are the probable ethical challenges associated with AI-induced academic integrity issues?

2. Conceptual Framework

The AI Literacy Program (AILP) is a practical framework that integrates Foundational, Technical, Humanistic, and Ethical Literacies into a structured curriculum to enhance higher standards of Academic Integrity as a characteristic of an informed AI user. This study conceptualizes AI literacy as a multifaceted phenomenon of cognitive, digital, ethical, and psychological constructs. Users are expected to demonstrate high-level cognitive functions while interacting with AI tools, such as selecting the right tool, posing a clear command, evaluating AI-generated output, and contextualizing it for a specific academic discourse (Ng, Leung, Chu, & Qiao, 2021). They are expected to reflect on AI output to deepen their learning experience and regulate its use to avoid over-reliance on AI tools. The ethical understanding of AI positions users to demonstrate the epistemic value of a unique scholarly identity, which each individual is expected to consciously develop in an AI era (Bearman, et al., 2025).

The study builds on Long and Magerko's (2020) AI Literacy Framework and Backward Design Framework (Wiggins & McTighe, 2005). Both theoretical frameworks extend the competency structure for designing SLOs, keeping critical evaluation, metacognitive awareness, and responsible use of AI central to the spirit. Backward design ensures a cohesive alignment .

across SLOs, instructional activities, and assessment strategies.

3. Methodology

The research methods aiming at conceptualizing the SLOs for AI Literacy Program comprised of three phases. In the first phase, a Systematic Literature Review (SLR) was conducted in compliance with the PRISMA 2020 guidelines, ensuring transparency and academic rigor. The purpose of this review was to identify key user competencies and ethical challenges reported in the extant literature. In the next phase, the results of the SLR reporting key use competencies were operationalized in four themes: Technical Literacy of AI; AI Ethics 1: Bias and Transparency; AI Ethics 2: Accountability and Responsibility; and AI Ethics 3: Strategies to comply with Academic Integrity Standards. In the last phase, the SLR findings of ethical issues were subjected to a co-occurrence analysis using VOSViewer to map the range of user understanding for an ethical use of AI.

This study developed foundational SLOs for a 0-credit course suitable for multidisciplinary undergraduate students. The study was conducted in accordance with the university's ethical guidelines, following IRB approval.

For SLR, data was shortlisted from Scopus, Web of Science, Google Scholar, Research Rabbit, JStor, SpringerLink, ScienceDirect, and Elicit against the keywords: *AI ethics, academic integrity issues in AI use in higher education, AI misuse in developing countries, and AI competence and skills*.

The retrieved records were scanned against the inclusion and exclusion criteria presented in **Table 1**. Inclusion and Exclusion Criteria for SLR

Attribute	Inclusion Criteria	Exclusion Criteria
Timeline	2022 – 2025	Publications before 2022 or after September 2025
Geographic Origin	Developing economies (as defined by the World Bank)	developed economies without comparative insights from developing regions
Document Type	Peer-reviewed research article, book chapter, conference papers	Editorials, opinion pieces, blogs, newsletters, dissertations, government/NGO reports, or non-peer-reviewed materials
Focus	AI skills and competencies, and ethical challenges posed by AI, AI-driven academic integrity issues	General ICT or EdTech studies without a specific reference to AI
Language	English	languages other than English
Context	Higher Education	Non-academic sectors
Accessibility	The full version of the publication is available through Google Scholar, JStor, SpringerLink, ScienceDirect, and Elicit	The full version of the publication is not available through Google Scholar, JStor, SpringerLink, ScienceDirect, and Elicit

A total of 56 research articles were reviewed to establish the two significant evidence bases for this study. The review findings highlighted essential skills users need to effectively utilize AI while also improving their personal potential, as well as the ethical challenges that require formal education. The key SLOs for AILP were drafted based on these findings.

To validate the SLOs, a team of experts from the university was appointed using purposive sampling. The validation team included representatives of the Department of Education (n=4), Center for Teaching Excellence (n=2), and the Dean's Office of Academic and Student Affairs (n=1). The strategic selection of diverse participation on the validation team was intended to provide a broad spectrum of insights, ensuring the ecological validity and applicability of AILP.

The validation process for SLOs was planned into four Delphi review cycles, each projecting a distinct focus for the validation team. The review cycles were facilitated by Google Forms. A consolidation report was shared with all reviewers before holding an online Zoom meeting to further discuss points of contradiction and consensus. Each Zoom meeting, lasting 60 minutes, was recorded with the reviewers' consent and saved to the Cloud. The first two rounds of review validated the scope, clarity, and relevance of ethical challenges addressed in AILP. The second and third review cycles focused on the rephrasing of SLOs. Table 2 presents the review criteria and 5-point Likert scale used for validation by Delphi reviewers.

Table 1 Review Criteria for Delphi Review Cycles

Review Criteria	Review Scale
<i>Concise</i> – The SLO is expressed in a logical manner, free from redundancy and repetition	1 = Very Poor – Does not meet the requirement at all
<i>Accuracy</i> – The SLO is quantifiable. It precisely reflects the intended knowledge, skills, or attitudes.	2 = Poor – Meets the requirement to a limited extent; major improvement needed
<i>Clarity</i> – The SLO is understandable and unambiguous	3 = Fair – Somewhat acceptable but requires noticeable refinement
<i>Evidence of a Problem-Based Learning (PBL) Approach</i> – The SLO reflects an inquiry-based, real-world problem orientation, the application of knowledge, and/or student-centered problem-solving	4 = Good – Adequately meets the requirement with only minor issues
<i>Relevance to the Module Theme</i> – The SLO directly aligns with and contributes to the stated user competency	5 = Excellent – Fully meets the requirement with no issues
<i>Context</i> – The SLO is well-aligned to learners' local academic context	

The iteration and statistical feedback from each round were shared with experts through email. Anonymity in responses was safeguarded throughout the Delphi review cycles. AILP teaching content, micro-lectures, reading exercises, case studies, strategy development exercises, AI use simulations, and assessment objectives have been informed by the final iteration of SLOs. The estimated completion time for the entire course is approximately 20 hours.

4. Results

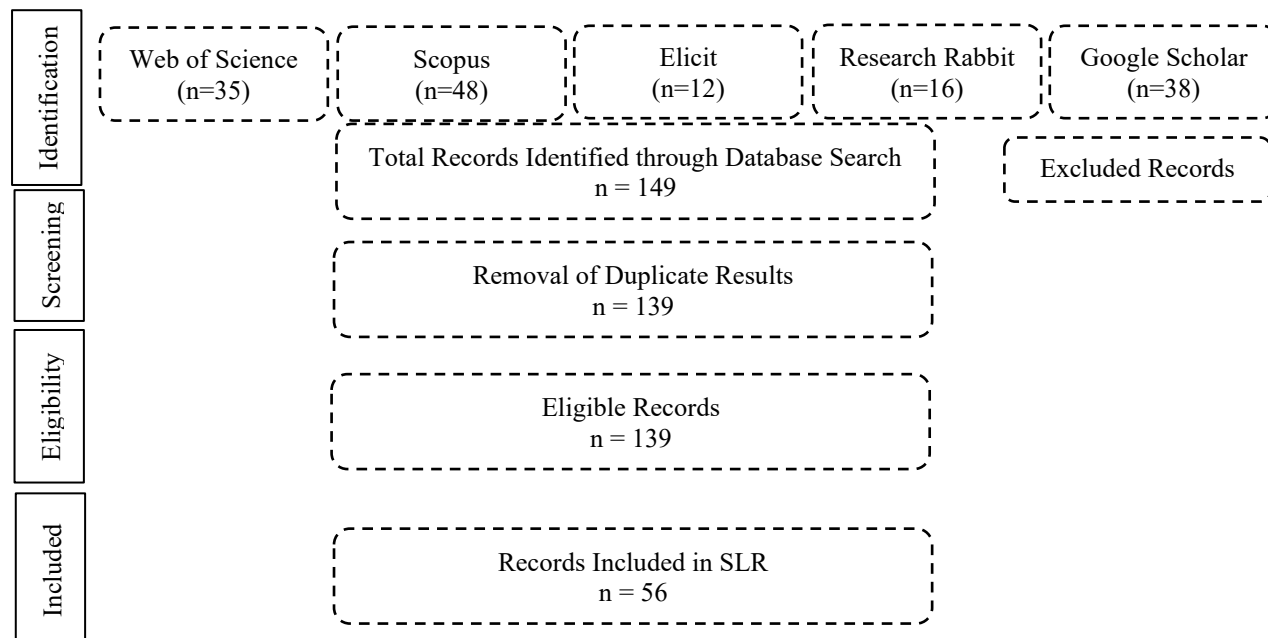
Phase 1: Systematic Literature Review

This review focuses on time period after the first AI-driven LLM, ChatGPT became accessible to users in the developing countries after November 2022. The initial search of the selected databases retrieved a total of 1200 records presenting findings predominantly from technologically advanced countries located in North America, Europe, Scandinavia, and Oceania. There has been a surge in publications since 2018, with a focus on AI-related competencies and ethical considerations, reflecting the growing need to address the ethical dilemmas that are a byproduct of AI's educational implications.

After removing duplicate records, the records were refined through abstract screening based on the inclusion and exclusion criteria. A full-text review of records led to further exclusion of

studies. Ultimately, 56 studies were included in the qualitative synthesis to explore nuanced perspectives on the necessary human skills for ethical collaboration with AI in higher education. Figure 1 illustrates the selection process for the records included in this study.

Figure 1 PRISMA flowchart - Literature Search Process



The demographic information of selected records included in this review is listed in Table 3.

Table 2 Demographic Information of Selected Records

Variable	Category	N	%
Year of Publication	2022	2	3 %
	2023	13	23 %
	2024	17	30 %
	2025	24	42 %
Origin	South Asia	32	57 %
	Africa	17	30 %
	Latin America	5	8 %
	Small Island States	2	3 %
Publication Type	Research paper	43	76 %
	Conceptual paper	5	8 %
	Review paper	8	14 %
Research Methodology	Qualitative methods	6	10 %
	Systematic reviews	24	42 %
	Quantitative methods	16	28 %
	Mixed-methods	7	12 %
	Policy reviews	3	5 %

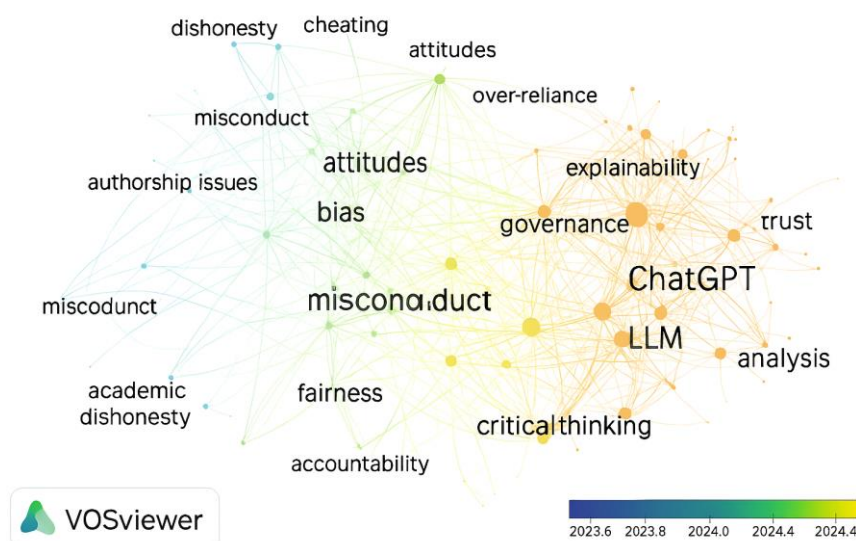
The selected records discussed the concept of AI Literacy from the perspective of developing economies, with a clear characteristic of a limited role in creating AI agents, yet a significant reliance on AI in academic discussions. The developing economies also exhibit deep-seated inequalities in access to stable internet, high-quality devices, and licensed tools. To effectively and responsibly engage with AI, this review highlights three key types of literacies, classified in Table 4, that empower users to leverage human potential when interacting with AI.

Table 3 Frequency of Sub-Literacies

Domain	Key Competencies	N	%
Foundational Literacy	Basic Digital Literacy	32	57
	Understanding of AI use	23	41
	Language proficiency	15	26
	Awareness of fairness, transparency, and accountability in academic discourse	18	32
	Contextual understanding	9	16
	Understanding of digital citizenship	10	17
Technical Literacy	AI Tool Proficiency	22	39
	Basic ICT Skills	35	62
	Understanding of AI capabilities, limitations, and ethical risks	24	42
	Prompt engineering	24	42
	Data literacy	22	39
	Ability to reshape AI for personal context	10	17
Humanistic Literacy	Critical thinking	26	46
	Ethical reasoning & Decision making	21	37
	Awareness of AI's societal implications	14	25
	AI's influence on identity, interpersonal, and emotional intelligence skills	11	19
	Information Evaluation	28	50
	Intercultural communication	16	28
	Life-long learning skills	17	30

The second investigative question concerned the ethical challenges of using AI in the academic context. A lexicon of AI-related academic integrity challenges was developed through co-occurrence analysis of abstracts and keywords of the included records by using VOSviewer for bibliometric mapping and text mining. To ensure relevance to the research questions, the emergent vocabulary surrounding AI and academic integrity was screened manually to remove irrelevant or generic terms such as 'Higher Education', AI, etc. The remaining terms were retained as entries in the lexicon. Through this iterative process, a total of 34 unique items were identified and included in the visualization of a co-occurrence network map presented in Figure 2. The proximity of nodes represents relatedness based on co-appearance of this lexicon in the literature. Clusters of terms were automatically detected by VOSviewer's clustering algorithm, providing a thematic grouping of threats including plagiarism-related risks, authorship integrity, misuse of generative AI, and fairness and accountability concerns.

Figure 2 Co-occurrence Network of AI Challenge Domains



The lexicon network of four emerging themes illustrated in Figure 2 identifies AI-driven ethical challenges to academic integrity. These are also presented in Table 5.

Table 4. AI Challenges Lexicon Themes

Theme	Lexicon
1	Bias, Fairness, Accountability
2	Privacy, Data protection
3	Equity, Limited access, Over-reliance
4	Authorship issues, Transparency

The findings listed in Tables 4 and 5 established the foundation for developing the SLOs for AILP. The SLOs were reviewed by experts in four Delphi rounds until 86% consensus was achieved. Table 6 outlines the first and final drafts of SLOs.

Table 5 Student Learning Outcomes - 1st & Final Draft Version

First Draft of SLOs	Final Draft of SLOs
Demonstrate foundational digital literacy to contextualize ethical use of AI in academic contexts	Demonstrate awareness of AI tools, their capabilities and limitations.
Demonstrate digital literacy to select, contextualise, and communicate with the most suitable AI tools for carrying out a diverse range of academic tasks while safeguarding academic integrity	Apply prompt engineering principles to create strategic prompts and critically evaluate AI-generated outputs for their alignment with a specific academic context by ensuring transparency, accuracy, and ownership.
Practice strategies to enhance fairness and accountability in AI-driven assessments and evaluations	Demonstrate critical thinking, ethical reasoning, and intercultural communication to foster lifelong learning, responsible decision-making, and the preservation of integrity in academic practices

Critically examine how reliance on AI affects equity in academic access and opportunity	Practice responsible authorship for transparent integration of AI tools to maintain accountability and uphold academic integrity standards
Design inclusive solutions that address digital divides in AI-supported learning	Critically assess the personal and societal impacts of AI, including the dangers of over-dependence and digital inequalities
Evaluate AI's societal implications for academic participation and lifelong learning	Demonstrate continuous skill development to understand developments in AI technologies, reshaping AI for personal context, and assessment of AI-generated content for accuracy, originality, and transparency

5. Discussion

The current study reports the development of a mandatory AI literacy module for a public-sector university in Pakistan to address rising violations of academic integrity principles resulting from students' unregulated use of Gen-AI. A rigorous Systematic Literature Review (SLR) was conducted to identify essential skills and user competencies for the responsible use of AI that enhances human potential and prepares users to address prevalent ethical challenges in academic AI applications. The majority of records included in the SLR were journal articles (42%), with a significant number ($n = 18$) authored by South Asian researchers. Other records presented findings from Uganda, Kenya, South Africa, Congo, Ethiopia, Namibia, and Nigeria. Findings from the Latin American context were derived from research carried out in Mexico, Colombia, and Brazil, while records from Thailand, Indonesia, and the Fiji Islands were also included. This uneven geographic distribution highlights the underrepresentation of many countries in the Global South and reflects significant digital divides across this region. Although the included records report ethical challenges arising from unregulated AI use, there is a notable scarcity of institutional governance practices and ethical frameworks developed within the unique context of the Global South as institutional responses to AI-induced threats to academic excellence. This imbalance highlights the unpreparedness of institutions and the missed opportunities to regulate AI to achieve the desired societal impact in developing economies.

Out of the 56 shortlisted records, the majority of 39% ($n = 22$) positioned *digital literacy* as the gateway to responsibly interact with AI tools. Digital literacy is conceptualized as the understanding of key computing operations and the ability to create, store, manage, retrieve, share, evaluate and integrate a range of digital information using a variety of digital platforms (Mijwil, et al., 2023). Digital literacy also conceptualizes the roles and responsibilities of users while navigating in a digital world (Long & Magerko, 2020). Building on the Digital Literacy, AI literacy is a combination of basic digital literacy and cognitive skills of critical thinking, reasoning, information evaluation, intercultural communication, and behavioural dispositions of lifelong learning (Long & Magerko, 2020). AI literacy extends beyond technical proficiency to a holistic engagement with AI tools to foster individual, societal, ethical, and practical values (Dignum, 2017). It is a comprehensive perspective that significantly shifts from pure technical application of AI to a more integrated deployment that prioritises human augmentation through collaborating with AI as a development partner. Research argues that AI literacy converges prompt engineering, awareness of domain-specific AI tools, device proficiency, and understanding of AI ethics with advanced competencies in algorithmic structuring (Suchithra & Arya, 2025). Another key user competence highlighted in the literature is the ability to understand, utilise, and critically evaluate

AI-generated output with the goal of personalizing and training machine learning algorithms, while also safeguarding individual privacy (Bearman, et al., 2025). The concept of AI literacy is evolving due to the rapid shifts in the affordances of AI and fragmentations across disciplines; however, users' proficiency in critical evaluation of AI-generated output is standardized across contexts (Ortiz, et al., 2025). Across records, it is emphasized that AI-generated content should be verified against reliable sources to develop quality control measures to identify false information, thereby maintaining academic integrity and research credibility. Literature also emphasizes the interdisciplinary appraisal of AI's societal implications, requiring users to be mindful of applying their computational knowledge to promote responsible social and human augmentation (Farheen, Cheema, Samee Ullah, & Bandeali, 2025). Mindfully using AI tools allows individuals to engage ethically with AI systems while also critically evaluating the broader impact on learning in a knowledge-based society (Inuwa-Dutse, 2023). Another significant user competence is to improve personal human potential through the active use of AI tools. The multimodal ability to use AI tools to promote individual creativity, critical reasoning, logic, and communication skills is likely to enhance overall human potential (Long & Magerko, 2020). Users' ability to strategically allocate administrative tasks to AI is crucial for optimizing human-AI collaboration to optimize individual performance and fostering creativity (Maimela & Mbonde, 2025).

In addition to ethical understanding and regulated use of AI tools, another important skill that greatly impacts AI performance is the efficiency of prompts. Prompt engineering is a specialized skill that sits at the cusp of linguistics, cognitive science, and algorithmic systems. An efficient prompt decomposes complex human intents into unambiguous, structured instructions, clearly sequenced to delineate human expectations for AI systems (Bearman, et al., 2025).

AI output is only as reliable as the datasets that support it. Literature emphasizes users' responsibility to report incorrect information to help improve dataset accuracy and dependability. Providing feedback to AI systems reflects attentive human oversight, a key disposition for critically engaging with AI systems (Raza, 2023). Another key responsibility of users is to maintain the provenance of intellectual property when utilizing AI-generated output. AI output is not original; it builds on thousands of datasets fed into the backend algorithms, created by humans over time. Proper attribution of intellectual credit is a legal requirement when using another scholar's work and also helps minimize the risk of false information (Chan, 2023).

Findings from the SLR also highlight the importance of intercultural communication, users' language skills, and questioning abilities in obtaining constructive feedback from AI tools, which likely encourages lifelong learning among users from developing countries. The findings of this SLR directed the development of three modules for the AI Literacy Course. These included: Foundational Literacy, Technical Literacy, and Humanistic Literacy.

The second objective, to identify AI-induced academic integrity challenges, uncovered four emerging themes presenting a broad range of ethical issues and dilemmas discussed in the selected literature. These include short-term concerns such as algorithmic bias, incorrect content, data privacy violations, fairness issues, accountability concerns, authorship disputes, and long-term risks of declines in creative thinking, critical reasoning, logical thinking, and deep learning.

AI algorithms reinforce social, cultural, and learning biases, raising concerns about the trustworthiness of AI-generated outputs, which often lack reliable sources. AI models are trained on datasets that underrepresent certain demographic or linguistic groups, such as those in Pakistan. Users in Pakistan encounter a language barrier to benefit from the known affordances of AI tools. The findings of this study show that AI output depends heavily on the quality of users' prompts and the backend algorithm, which do not always produce outputs directly relevant to the needs of users in Pakistan, thus challenging the fundamental principles of inclusion and fairness (Farheen

et al., 2025).

In the context of AI, fairness does not imply *a one-size-fits-all approach*; rather, it refers to factual accuracy within a specific context and the inclusion of diverse sociocultural perspectives, ensuring applicability and authenticity for users (Dignum, 2017). Privacy and data protection are other significant ethical concerns reported in the literature. AI models may ask users for personal information, and users in the Global South are more vulnerable to systemic data colonialism due to their limited ability to exercise rights, such as informed consent, access, or redress (Maimela & Mbonde, 2025).

The findings of SLR conceptualized key competencies for AI literacy and the fundamentals for ethical awareness. These evidence sets were conceptualized into six Students' Learning Outcomes (SLOs) setting the foundation of AILP. The SLOs were validated through four iterative Delphi reviews. The final draft of SLOs are reflexive and evaluative in nature. They project the notion of academic integrity as a virtue of users' AI literacy, aligning with Bearman et al. (2023), indicating that sustainable practices of using AI in academic contexts can't be achieved through rigorous rules of right and wrong, but rather a responsible engagement with AI is a consequence of users' critical engagement with AI-generated output and their individual decision on the accuracy and reliability of the output. It also aligns with the findings of Cotton et al. (2024) that detection approaches and limiting the use of AI put students at a disadvantage because they miss out on the benefits of the revolutionary technology and do not promise high standards of academic integrity. Research shows that users who understand transparency and authorship when using AI are better equipped to make responsible decisions than those who are simply told not to use AI.

Conclusion

Globally, the academic community is challenged by the potential implications of unregulated AI use, which demands an urgent need to establish guidelines and whistleblowing systems to uphold rigorous standards of academic integrity and ensure that new technologies align with institutional values and principles. The existing literature highlights several ethical and academic challenges posed by the rapid proliferation of AI; it largely overlooks the complex vulnerabilities faced by users in the Global South. This gap necessitates that researchers in this region explore the pedagogical, philosophical, psychological, and social implications of AI more deeply and develop systems contextualized to the region's unique digital and institutional infrastructure. Developing a contextualized AI Literacy Program (AILP) addresses this critical gap in the literature and contextualizes AI literacy for the unique needs of users in Pakistan. As AI continues to transform academic practices, instead of restricting its use and risking students in this region falling behind global trends, it is crucial to develop and implement empirically validated literacy programs to prepare students for future demands.

The implementation of AILP as a 0-credit course on Moodle across non-computing disciplines, without altering existing degree programs, is a scalable solution for Pakistan's HEIs to empower students to use AI responsibly. It is mandatory to evaluate the long-term effects of AILP on students' academic behaviour and examine the correlation between AI literacy and the decrease in AI-induced plagiarism cases at the university. The next frontier for higher educational institutional policies in Pakistan will be to create AI-resilient assessments and to continuously improve faculty digital pedagogy to ensure students stay aligned with global trends and market needs.

DECLARATION STATEMENTS

This paper is extracted from the first author's PhD thesis

Conflict of Interest

None to declare

Ethical Permissions

Ethical approval was obtained from relevant review board at University of Peshawar. Informed consent was secured from all participants prior to data collection.

Funding

None to declare

Author Contribution

TA conceptualized the study and collected the data. JA supervised the project. Both authors conducted the data analysis and wrote and approved the final version of this manuscript.

6. References

- Arias Ortiz, E., Castro Vergara, N., Forero, T., Nina, G., Giambruno, C., Pérez, M., & Segura, D. (2025). AI and Education: Building the Future Through Digital Transformation. Inter-American Development Bank. doi:<http://dx.doi.org/10.18235/0013500>
- Azeem, A., Siddiqui, G., & Ali, S. Z. (2025). Impact of Artificial Intelligence on Academic Integrity: Exploring Teachers' and Students' Perspectives. *Journal of Development and Social Sciences*, 6(1), 499-507. Retrieved 9 6, 2025, from <https://ojs.jdss.org.pk/journal/article/view/1339/1256>
- Bearman, M., Fawns, T., Corbin, T., Henderson, M., Liang, Y., Oberg, G., . . . Matthews, K. (2025). Time, emotions and moral judgements: how university students position GenAI within their study. *Higher Education Research & Development*, 1–15. doi:<https://doi.org/10.1080/07294360.2025.2580616>
- Chan, C. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*. doi:DOI:10.1186/s41239-023-00408-3
- Cotton, D., Cotton, P., & Shipway, R. (2024). Chatting and Cheating: Ensuring academic integrity in teh era of ChatGPT. *Innovations in education and Teaching International*, 61(2), 228 - 239.
- Dignum, V. (2017). Responsible Artificial Intelligence: Designing Ai for Human Values. *ITU Journal: ICT Discoveries*, 1-8. Retrieved 06 10, 2025, from https://www.itu.int/dms_pub/itu-s/opb/journal/S-JOURNAL-ICTF.VOL1-2018-1-P01-PDF-E.pdf
- Farheen, S., Cheema, A., Samee Ullah, R., & Bandiali, M. (2025). Equity and Bias in AI Educational Tools: A Critical Examination of Algorithmic Decision-Making in Classrooms. *The Critical Review of Social Sciences Studies*, 3(3). Retrieved 09 16, 2025, from <https://thecrsss.com/index.php/Journal/article/view/665/691>
- Figaredo, D. D., & Stoyanovich, J. (2023). Responsible AI literacy: A stakeholder-first approach. *Big Data & Society*, 10(2). doi:<https://doi.org/10.1177/20539517231219958>
- Gonsalves, C. (2025). Addressing student non-compliance in AI use declarations: implications for academic integrity and assessment in higher education. *Assessment & Evaluation in Higher Education*, 50(4), 592–606. doi:<https://doi.org/10.1080/02602938.2024.2415654>
- Grassini, S. (2023). Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences*, 13(7), 692. doi: <https://doi.org/10.3390/educsci13070692>.

- Gunton, L. (2022). Librarians as Drivers of Academic Integrity for Student Success at University. *Journal of the Australian Library and Information Association*, 71(2), 156-170. Retrieved 09 10, 2025, from <https://www.tandfonline.com/doi/full/10.1080/24750158.2022.2063101>
- Inuwa-Dutse, I. (2023). FATE in AI: Towards Algorithmic Inclusivity and Accessibility. 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, (pp. 1-14). doi:10.1145/3617694.3623233
- Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. doi:<https://doi.org/10.1145/3313831.3376727>
- Luckin, R., Rudolph, J., Grünert, M., & Tan, S. (2024). Exploring the future of learning and the relationship between human intelligence and AI. *Journal of Applied Learning & Teaching*, 7(1). Retrieved 09 10, 2025, from <https://jalt.journals.publicknowledgeproject.org/index.php/jalt/article/view/1659/767>
- Maimela, C., & Mbonde, P. (2025). Artificial intelligence in South African universities: curriculum transformation and decolonisation—aid or obstacle? *Frontiers in Sociology*. doi:<https://doi.org/10.3389/fsoc.2025.1543471>
- Mijwil, M., Hiran, K., Doshi, R., Dadhich, M., Al-Mistarehi, A., & Bala, I. (2023). ChatGPT and the Future of Academic Integrity in the Artificial Intelligence Era: A New Frontier. *Al-Salam Journal for Engineering and Technology*, 68(2). doi:<https://doi.org/10.55145/ajest.2023.02.02.015>
- Ng, D. T., Leung, J. K., Chu, S. k., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. doi:<https://doi.org/10.1016/j.caeai.2021.100041>
- Ramdhani, D., Husnawadi, I., & Waliyadin, A. (2025). Plagiarism and Academic Integrity in Indonesian Islamic Higher Education: Factors and Policy Implications. *EDUKASI: Jurnal Penelitian Pendidikan Agama Dan Keagamaan*, 23(2), 322-339. Retrieved 9 5, 2025, from <https://jurnaledukasi.kemenag.go.id/edukasi/en/article/view/2154/703>
- Raza, M. (2023). Implications of ChatGPT: Insights from the Higher Education of Pakistan. *Journal of Education & Humanities Research (JEHR)*, 16(2), 38-46.
- Seo, H., Paulo da Silva, M., Jona, B., Iqbal, A., Burkett, K., Khan, H., & Carreno, A. (2025). Comparative analysis of AI attitudes among JMC students in Brazil and the US: a mixed-methods approach. *Online Media and Global Communication*, 4, 187 - 215. Retrieved 09 10, 2025, from <https://api.semanticscholar.org/CorpusID:276042918>
- Sharma, D. (2025). Balancing Innovation and Integrity in the Pacific: Exploring the Dual Impact of AI on Tertiary Education in Fiji. *International Journal of Social Sciences, Language and Linguistics*, 20(3), 14-21. doi:10.55640/ijssll-05-09-03
- Suchithra, V. G., & Arya, C. (2025). The Study on Ethics and Biases in AI-Powered Education. *European Journal of Contemporary Education and E-Learning*, 3(2), 37-43. Retrieved 09 15, 2025, from <https://ejceel.com/index.php/journal/article/view/178/140>
- Wiggins, G., & McTighe, J. (2005). *Understanding by Design*. Ascd.