

Agentic AI Divide in Higher Education

Christina Micheal Samivaloo¹, Latha Haridass², Izawati Ngadni³, Fadzil Abdul Hanid⁴, Zainudin Hassan⁵

Corresponding Author: Izawati Ngadni

^{1,2,3}UNITAR International University

^{4,5}Universiti Teknologi Malaysia (UTM), Malaysia

HOW TO CITE:

Christina Micheal Samivaloo, Latha Haridass, Izawati Ngadni, Fadzil Abdul Hanid, Zainudin Hassan (2026). Agentic AI Divide in Higher Education. International Journal of Special Education, 41(24s), 639-650

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT:

Higher education institutions are expanding their use of agentic artificial intelligence (AI), and fundamental questions about equity, privacy, and fairness have emerged. Unlike earlier rule-based systems, agentic AI autonomously manages complex tasks including adaptive instruction, academic risk prediction, and institutional decision-making. This systematic literature review synthesises findings from 28 peer-reviewed studies (2021–2026) to examine how agentic AI creates or deepens inequalities among marginalised student populations. Three intersecting challenges are identified: (i) an AI literacy divide that prevents equitable engagement even when access exists; (ii) data governance deficits that erode student trust through opaque surveillance practices; and (iii) algorithmic bias that can silently reproduce structural inequities. Results confirm that access alone is insufficient; meaningful and fair use requires targeted AI literacy training, student-centred data governance frameworks, and bias-mitigation strategies embedded in institutional policy. This review calls for governance models that position students as rights-bearing participants and for longitudinal research extending beyond high-resource institutional contexts.

Keywords: Agentic Artificial Intelligence, Higher Education, AI Literacy, Algorithmic Fairness, Data Governance

1. Introduction

Artificial intelligence has evolved into agentic systems capable of autonomous, multi-step decision-making, including adapting instruction, predicting performance, managing placements, and triggering interventions in higher education (Karran et al., 2024). The adoption of AI in higher education has accelerated markedly since 2021, driven by the widespread availability of large language models and intelligent tutoring platforms. Universities now deploy AI across

learning management systems, automated assessment, student-support chatbots, and early-warning analytics dashboards (Nguyen, Gardner, & Sheridan, 2022). Recent research in open, distance and digital education also highlight AI's potential for personalised learning and administrative efficiency, while emphasising challenges related to infrastructure, training, ethical concerns, and data privacy (Karim et al., 2024). Yet literature consistently shows that these tools do not serve all students equally.

Students from marginalised backgrounds, including first-generation college students, racial and ethnic minorities, and those from low-income households, often face structural barriers that limit the benefits they derive from AI-enhanced environments (Shoval, 2025).. This review examines three interconnected concerns: the second-level AI literacy divide, where differences in AI skills affect equitable benefits (OECD, 2024); privacy and data governance, including surveillance and informed consent risks (Yan et al., 2023); and algorithmic bias, whereby historically biased data can reproduce existing inequalities (Baker & Hawn, 2022). By integrating these dimensions within the specific context of agentic AI, this review synthesises 28 studies to identify evidence, gaps, and governance priorities for responsible and equitable AI deployment in higher education.

1.1 Research Objectives

This review is guided by three research objectives:

- RO1: To investigate how differential agentic AI literacy shapes academic outcomes and belonging among marginalised students in higher education.
- RO2: To examine how student-centred data governance affects marginalised students' trust and their willingness to engage with agentic AI tools.
- RO3: To identify bias-mitigation strategies that ensure fair algorithmic decision-making for marginalised students in higher education.

1.2 Scope and Significance

The review covers peer-reviewed empirical and conceptual literature published between 2021 and March 2026 that addresses agentic AI in higher education and refers to at least one of three constructs: equity/digital divide, data privacy/governance, or algorithmic bias/fairness. Studies focused exclusively on K-12 settings, on general AI literacy without a higher-education context, or on non-agentic AI applications were excluded.

The significance of this integrated review is both scholarly and practical. Scholars lack a consolidated synthesis of how these three ethical dimensions interact in the deployment of agentic

AI. Practitioners, educators, institutional policymakers, and AI developers require integrated guidance to move beyond tool adoption towards responsible, equitable governance. This review addresses both needs.

2. 2. Methodology

A systematic literature review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to ensure transparency, comprehensiveness, and replicability (Sarkis-Onofre et al., 2021).

2.1 Search Strategy

A structured search was conducted across five academic databases: IEEE Xplore, Education Resources Information Centre (ERIC), PubMed, EBSCOhost, and Google Scholar. Twenty-one Boolean search queries were developed across three conceptual clusters:

- Higher education context: “higher education”, “university*”, “tertiary education”
- Student agency and agentic AI: “agentic AI”, “student agency”, “learner agency”, “autonomy”, “self-regulation”
- Equity, privacy, and fairness: “algorithmic fairness”, “algorithmic bias”, “data privacy”, “learning analytics”, “AI ethics”, “digital divide”, “ChatGPT”, “adaptive learning”, “online proctoring”

The search was limited to peer-reviewed, full-text articles in English published between January 2021 and March 2026. The date range was chosen to capture literature on contemporary agentic AI capabilities while ensuring a sufficient volume of peer-reviewed articles.

2.2 Eligibility Criteria

Inclusion criteria required that studies: (i) be published in peer-reviewed journals; (ii) be available in full text; (iii) be written in English; and (iv) directly address at least one of the three constructs of equity/digital divide, data privacy/governance, or algorithmic bias/fairness in the context of AI in higher education. Studies focused exclusively on K-12 settings, on general AI without reference to higher education, or on non-peer-reviewed sources were excluded.

2.3 Screening and Selection

A two-stage screening process was used. In stage

one, two independent reviewers screened titles and abstracts against the eligibility criteria. Conflicts were resolved by a third reviewer, and reasons for exclusion were recorded. In stage

two, full texts of potentially eligible studies were retrieved and assessed for inclusion. Table 1 summarises the selection process.

Table 1. PRISMA study selection summary

Phase	Starting (N)	Removed	Reason for Removal	Final (N)
Identification	—	—	Databases searched: IEEE Xplore (224), ERIC (127), PubMed (118), EBSCOhost (218), Google Scholar (326)	1,013
Screening	1,013	655	Duplicate papers; missing abstracts	358
Eligibility (relevance screen)	358	105	Low semantic relevance to search queries	253
Eligibility (full-text review)	253	203	No full script/findings; equity, privacy, or fairness not addressed	50
Inclusion	50	22	Excluded to retain highest-quality papers only	28

2.4 Data Extraction

Data were extracted using a standardised matrix table that captured: study title; author(s) and year; research design; primary objectives; and key findings on equity, privacy, or fairness. Two reviewers independently extracted the data; disagreements were resolved through discussion with a third reviewer.

2.5 Quality Appraisal

The 28 included studies were appraised using design-appropriate tools. Systematic and scoping reviews were appraised using JBI criteria (Aromataris et al., 2015), PRISMA-ScR (Tricco et al., 2018), AMSTAR 2 (Shea et al., 2017), and ROBIS (Whiting et al., 2016). Quantitative and mixed-methods studies were evaluated using CASP Cross-Sectional (Critical Appraisal Skills Programme [CASP], 2024) and the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018). Narrative reviews were assessed using SANRA (Baethge et al., 2019). Quality ratings ranged from moderate to very high across the corpus.

2.6 Data Synthesis

A thematic synthesis approach was applied to all 28 studies. Themes were identified inductively from the data extraction matrices and then

mapped deductively to the three research objectives. Three overarching themes emerged: (i) digital inequalities and the AI literacy divide; (ii) student privacy and data governance deficits; and (iii) algorithmic bias and the shift from unknown to known bias. These themes are reported in the Results and Discussion section.

3. Results

This section presents the outcomes of the systematic selection process, the quality appraisal of the included studies, and the three themes identified through thematic synthesis. All findings are grounded in the 28 articles retained after PRISMA-guided screening.

3.1 Study Selection

Twenty-one Boolean search queries, grouped into eight topic clusters, were executed across five academic databases. The PRISMA four-stage process governed selection. The identification phase returned 1,013 records. After screening, 655 duplicates and records lacking abstracts were removed, leaving 358 papers. The eligibility assessment excluded 105 records for low semantic relevance and a further 203 for insufficient coverage of equity, privacy, or fairness, producing a quality-ranked pool of 50. A final inclusion assessment removed 22

papers for unavailable full text or peripheral relevance, yielding a corpus of 28 articles. The distribution across databases was as follows: Google Scholar (n = 326), IEEE Xplore (n = 224), EBSCOhost (n = 218), ERIC (n = 127), and PubMed (n = 118).

3.2 Quality Assessment

Overall methodological quality across the 28 included studies ranged from moderate to very high, depending on research design. Each study was appraised using a tool appropriate to its design. Table 2 summarises the appraisal tool assignments, included studies, and quality ratings.

Table 2. Quality appraisal summary of 28 included studies

Research Design	Appraisal Tool	Studies Included	Rating
Systematic review	JBIR Systematic Review (Aromataris et al., 2015)	Baker & Hawn (2022) Al-kfairy et al. (2024) Kerman et al. (2024) Asri (2024) Khan et al. (2025) Chinta et al. (2024)	Moderate–High
Scoping review	PRISMA-ScR (Tricco et al., 2018)	Pierrès et al. (2024) du Plooy et al. (2024) Gordon et al. (2024) Lartey et al. (2025) Mangalindan & Wyatt (2025) Weidener & Fischer (2023)	Moderate–High
Narrative review	SANRA (Baethge et al., 2019)	Francis et al. (2025) Deckker & Sumanasekara (2025) Baker & Hawn (2022)	Moderate–High
Mixed methods	MMAT (Hong et al., 2018)	Ifenthaler et al. (2024) Laufer et al. (2021) Al-Zahrani (2024) Chan & Hu (2023) Shoval (2025)	Moderate–High
Qualitative	CASP Qualitative Checklist (2024)	Xu et al. (2025) Modén et al. (2024)	High
Quantitative	CASP Cross-Sectional & MMAT (2024)	Pellas (2023) Jiang & Pardos (2021)	High

		Mangal & Pardos (2024)	
Umbrella review	ROBIS + AMSTAR 2	El Arab et al. (2025)	High
Systematic mapping	JBİ Appraisal Checklist (2020)	Pham et al. (2025)	High
Bibliometric / science mapping	JBİ Appraisal Checklist (2020)	Mac Fadden et al. (2024)	High

Quality across design categories is discussed below.

3.2.1 *Systematic, Scoping, and Umbrella Reviews*

Most review studies were rated as high quality. They adhered to established reporting guidelines (PRISMA, PRISMA-ScR, or JBİ), employed comprehensive multi-database search strategies, and used independent dual-reviewer screening to minimise selection bias (Lartey et al., 2025; Mangalindan & Wyatt, 2025; Al-kfairy et al., 2024; Chinta et al., 2024; Pierrès et al., 2024; du Plooy et al., 2024; Gordon et al., 2024). A smaller subset received moderate ratings because they omitted formal quality checklists, reported only single-reviewer screening, or did not fully document their search protocol (Weidener & Fischer, 2023; Asri, 2024; Khan et al., 2025).

3.2.2 *Quantitative Studies*

Quantitative studies achieved very high-quality ratings (Pellas, 2023; Jiang & Pardos, 2021; Mangal & Pardos, 2024). The strongest work drew on datasets comprising nearly two million course enrolment records, applied rigorous missing-data procedures to minimise bias, and used advanced statistical methods, including adversarial learning algorithms and fairness metrics, to confirm the reliability of the findings (Jiang & Pardos, 2021; Mangal & Pardos, 2024). Survey-based studies employed well-validated instruments and established structural validity before reporting results (Pellas, 2023).

3.2.3 *Mixed Methods and Qualitative Studies*

Mixed-methods and qualitative studies were generally strong (Shoval, 2025; Modén et al.,

2024; Xu et al., 2025; Laufer et al., 2021). Methodological strengths included detailed audit trails, structured thematic coding procedures, and inter-rater reliability checks (Shoval, 2025; Modén et al., 2024; Xu et al., 2025). Studies that used multiple simultaneous data sources produced particularly robust findings through triangulation (Ifenthaler et al., 2024; Laufer et al., 2021). However, several mixed-methods studies received moderate ratings because convenience sampling limited the transferability of their findings (Al-Zahrani, 2024; Chan & Hu, 2023; Ifenthaler et al., 2024).

3.2.4 *Narrative and Bibliometric Reviews*

Quality in narrative and bibliometric reviews depended on the transparency of source-selection procedures. Deckker and Sumanasekara (2025) received high ratings for clearly documenting search terms, inclusion criteria, and the cultural representativeness of the sampled literature. Baker and Hawn (2022), Francis et al. (2025), and Mac Fadden et al. (2024) received moderate ratings because of limited documentation of search terms or reliance on a single database, which may have introduced coverage bias.

3.3 *Thematic Analysis*

Thematic synthesis across the 28 studies yielded three core themes, each aligned with one of the three research objectives. Table 3 provides a structured overview of the themes, core findings, evidential basis, and institutional implications, followed by the narrative elaboration below.

Table 3. Thematic synthesis: core themes, findings, and implications

Theme	Core Finding	Key Evidence	Implication
T1: AI Literacy Divide	Access alone does not ensure equitable benefit; AI literacy, the capacity to engage	Shoval (2025); Pellas (2023); Mac Fadden et	Mandatory AI literacy curricula; equity-centred

	critically with AI, is the decisive variable.	al. (2024); Modén et al. (2024)	rollout strategies targeting marginalised students.
T2: Data Governance Deficit	Institutional governance prioritises efficiency over student autonomy; opacity around data collection generates psychological distress and erodes trust.	Xu et al. (2025); Deckker & Sumanasekara (2025); Yan et al. (2023); UNESCO (2021)	Shift to student-centred, participatory consent models (e.g., OCAP framework); mandatory GDPR compliance.
T3: Algorithmic Bias	Accuracy does not imply fairness; systems may systematically disadvantage students based on protected characteristics even when performing well overall.	Baker & Hawn (2022); Jiang & Pardos (2021); Mangal & Pardos (2024); Chinta et al. (2024)	Adopt intersectionality-aware bias auditing (ABROCA+); apply pre-, in-, and post-processing debiasing strategies.

3.3.1 Theme 1: Digital Inequalities and the Second-Level AI Literacy Divide

Across 12 of the 28 studies, findings identified a second-level digital divide in which AI literacy, rather than access to technology, determines students' ability to benefit from AI. Mac Fadden et al. (2024) linked AI adoption inequalities to disparities in race, class, and geography, while Shoval (2025) found that minority and first-generation students were less likely to transfer AI skills beyond instructor-led activities due to limited pedagogical support. Pellas (2023) similarly identified structured AI training as a stronger predictor of effective and equitable AI use than demographic characteristics, while Modén et al. (2024) showed that insufficiently operationalised fairness can reproduce existing inequalities. Thus, equitable AI adoption requires more than universal access; embedding equity, diversity, and inclusion (EDI) into implementation and developing both technical and critical AI capabilities are essential (Lartey et al., 2025; Mangalindan & Wyatt, 2025).

3.3.2 Theme 2: Student Privacy and Data Governance Deficits

Seventeen of the 28 studies examined student privacy and data governance, showing that institutional frameworks often prioritise operational efficiency and risk management over student autonomy and rights (Deckker & Sumanasekara, 2025). Xu et al. (2025) found that inadequate transparency and perceived surveillance contributed to student anxiety,

discomfort, and reduced authentic engagement, while Yan et al. (2023) warned that continuous AI-enabled monitoring may normalise surveillance and erode privacy and autonomy. Although UNESCO (2021) and the EU AI Act (European Commission, 2024) emphasise transparency, accountability, and human oversight, implementation gaps remain. The evidence therefore supports student-centred governance through participatory and dynamic consent, cultural responsiveness, and the OCAP principles of Ownership, Control, Access, and Possession (Deckker & Sumanasekara, 2025), alongside stronger compliance with GDPR requirements for data minimisation, purpose limitation, and the right to explanation (Francis et al., 2025; Xu et al., 2025).

3.3.2 Theme 3: From Unknown Bias to Known Bias: Algorithmic Fairness

Fifteen of the 28 studies examined algorithmic bias and fairness, emphasising the need to make bias visible, measurable, and actively managed. Baker and Hawn (2022) identified bias across grading, dropout prediction, and academic placement, disproportionately affecting students by race, gender, and socioeconomic status, while Jiang and Pardos (2021) showed that excluding race from models does not prevent bias due to proxy variables. Mangal and Pardos (2024) highlighted intersectional disparities and demonstrated that ABROCA+ and adversarial learning can reduce predictive inequalities. Chinta et al. (2024) proposed a three-stage

debiasing approach involving pre-processing, in-processing, and post-processing, while Baker and Hawn (2022) and Pham et al. (2025) stressed the need for context-specific, transparent, and externally verifiable fairness measures. Collectively, the findings show that AI literacy influences who benefits, data governance shapes privacy risks, and algorithmic bias affects whose outcomes are advantaged or disadvantaged, making integrated approaches essential for equitable agentic AI implementation.

4. Discussion

This section interprets the thematic findings against the three pre-specified research objectives, examines convergences and divergences with prior literature, and derives evidence-based implications for educators, students, and institutional policymakers.

4.1 RO1: Unequal AI Literacy and Its Consequences for Marginalised Students

The findings confirm and extend the OECD's (2024) concept of a second-level digital divide by identifying an AI literacy divide involving students' evaluative, critical, and strategic capacities to engage with agentic AI. Shoval (2025) found that minority students recognised AI's value but were less able to apply it independently, while Pellas (2023) identified targeted training as a key predictor of equitable AI use. Thus, access alone is insufficient, and AI implementation must address the capabilities and needs of marginalised students (Modén et al., 2024). AI literacy inequity is therefore a pedagogical, structural, and governance challenge requiring coordinated policy, curriculum reform, and resource allocation. However, its relationship with academic belonging remains underexplored, with Shoval (2025) suggesting potential effects on students' engagement and persistence, warranting further longitudinal research.

4.2 RO2: Student-Centred Data Governance and the Architecture of Trust

The findings reveal a systemic gap between institutional data governance and student needs. Agentic AI relies on extensive behavioural, academic, and sometimes biometric data collection, while limited transparency can

generate anxiety, distrust, and reduced engagement (UNESCO, 2021; Yan et al., 2023; Xu et al., 2025). Current governance models often prioritise institutional risk management over student empowerment (Deckker & Sumanasekara, 2025; UNESCO, 2021; European Commission, 2024). Student-centred governance therefore requires transparent data practices, meaningful consent, and a right to explanation, supported by participatory and culturally responsive approaches such as the OCAP principles of Ownership, Control, Access, and Possession (Deckker & Sumanasekara, 2025). While the EU AI Act provides a regulatory foundation (European Commission, 2024), equivalent standards are needed more broadly to strengthen student trust, autonomy, and engagement with agentic AI.

4.3 RO3: Bias Mitigation and the Institutionalisation of Algorithmic Fairness

The review highlights a critical distinction between accuracy and fairness, as AI systems may achieve high predictive accuracy while systematically disadvantaging demographic groups (Baker & Hawn, 2022; Jiang & Pardos, 2021). The shift from "unknown bias" to "known bias" is therefore essential for equitable AI deployment, supported by tools such as ABROCA+ and the FBU framework for identifying and managing algorithmic disparities (Mangal & Pardos, 2024; Chinta et al., 2024). Intersectional fairness is also critical, as separate demographic analyses may overlook compounded disadvantages (Mangal & Pardos, 2024). Chinta et al.'s (2024) three-stage framework—pre-processing, in-processing, and post-processing—provides a practical roadmap for bias mitigation, while regular audits are necessary to ensure fairness as student populations, curricula, and educational contexts evolve.

4.4 Implications for Educators, Students, and Institutions

For educators, the findings call for pedagogical and attitudinal shifts by embedding AI literacy into teaching to develop students' critical evaluation, awareness of AI limitations, and understanding of ethical implications. The use of

AI-supported assessment and feedback tools also demonstrates the potential for AI to support personalised feedback and students' academic development when integrated appropriately into teaching practice (Chong & Dit, 2025). Educators should also recognise that equal access to AI tools does not necessarily ensure equitable outcomes. Educators should also recognise that equal access to AI tools does not necessarily ensure equitable outcomes. For students, responsible AI use requires not only technical skills but also awareness of data privacy, algorithmic bias, and governance processes influencing AI-supported decisions. At the institutional level, equitable agentic AI integration requires coordinated action through **(i)** curriculum policies that establish AI literacy as a cross-disciplinary competency, **(ii)** participatory data governance, **(iii)** procurement requirements for bias auditing and intersectional fairness, and **(iv)** regular independent audits of deployed AI systems. These measures should be implemented collectively to support sustainable and equitable AI adoption.

4.5 Limitations

Several limitations constrain the scope and generalisability of this review. First, restricting publications to English may have excluded relevant evidence from non-Anglophone contexts, while excluding grey literature, institutional reports, vendor documentation, and policy briefs may have limited coverage of practical implementation challenges (Nguyen, Zawacki-Richter, & Biggs, 2022; Zawacki-Richter et al., 2019; Deckker & Sumanasekara, 2025). Second, the literature is concentrated in high-resource settings, particularly North America, Western Europe, and Australasia, leaving lower-income contexts underrepresented (OECD, 2024; Modén et al., 2024). The rapid evolution of agentic AI and limited longitudinal research further restrict understanding of how equity, privacy, and fairness outcomes change over time (Jiang & Pardos, 2021; Al-kfairy et al., 2024). Finally, inconsistent definitions of AI literacy, fairness, equity, and student agency limit cross-study comparison; future reviews should therefore adopt pre-registered protocols and clearly defined constructs to strengthen synthesis

and comparability.

4.6 Future Research Directions

Five priorities emerge from the limitations and gaps identified in this review:

- Empirical studies with marginalised student populations as primary participants, using qualitative and participatory methods to capture lived experiences of AI-mediated inequity that survey and experimental methods cannot adequately represent.
- Research in low-resource and non-Western institutional contexts to test the generalisability of current governance and bias-mitigation models and to surface context-specific challenges that are obscured in high-resource settings.
- Longitudinal studies tracking equity, privacy, and fairness outcomes across multiple academic years to capture how AI literacy divides, governance practices, and algorithmic bias profiles evolve as institutional AI use matures.
- Co-design research with students, particularly those from marginalised backgrounds, in developing governance frameworks, AI literacy curricula, and fairness audit protocols, consistent with the participatory consent principles identified in Theme 2.
- Cross-national comparative research examining how different legal regimes (e.g., the GDPR in Europe, FERPA in the United States, and emerging AI regulation in the Global South) shape institutional governance practices and their implications for equity.

5. Conclusion

This systematic literature review examined how agentic, autonomous, and multi-step AI systems affect marginalised students in higher education across equity, privacy, and algorithmic fairness. The synthesis of 28 peer-reviewed studies shows that AI access alone does not ensure equitable outcomes; without targeted AI literacy support, agentic AI may widen achievement gaps by reinforcing differences in students' capacity to

use and critically evaluate AI. The findings also highlight risks from inadequate data governance and algorithmic bias, which can undermine privacy, trust, engagement, and fairness. Equitable AI integration therefore requires institutionally supported AI literacy, participatory data governance, continuous bias auditing, and transparent fairness measures, including ABROCA+ and the FBU framework. Overall, agentic AI adoption is not merely a technical issue but a governance and educational justice imperative requiring sustained institutional commitment to student rights, autonomy, and accountability.

6. Acknowledgements

The authors gratefully acknowledge UNITAR International University, particularly the Faculty of Education and Humanities, for providing the academic environment, library resources, and database access that supported the completion of this systematic review. The authors also extend their appreciation to Universiti Teknologi Malaysia (UTM) for the collegial support provided by the contributing authors. No external funding was received for this study.

7. Author Declaration

Author Contributions. Christina Micheal Samivaloo conceptualised the review, designed and executed the search strategy, led data extraction and quality appraisal, conducted the thematic synthesis, and wrote the original draft. Latha Haridass supervised the study, guided the methodological design and PRISMA protocol, and critically reviewed and edited successive drafts. Izawati Ngadni, as corresponding author, validated the data extraction and thematic synthesis, coordinated manuscript revisions, and

managed correspondence with reviewers. Fadzil Abdul Hanid and Zainudin Hassan contributed to the quality appraisal of included studies, provided critical feedback on the discussion and implications, and reviewed the manuscript for intellectual content. All authors read and approved the final manuscript for publication.

Originality and Competing Interests. The authors confirm that this manuscript is original work, has not been published previously, and is not under consideration elsewhere. The authors state there is no conflict of interest.

Data Availability Statement. Data availability is not applicable to this paper as no new data were created or analysed in this study. The literature corpus underlying the synthesis is fully cited in the References section.

Ethics Declaration. Not applicable. This systematic review did not involve human participants, human tissue, or animal subjects.

Use of Generative AI. In accordance with journal policy, the authors disclose the use of generative AI tools in the preparation of this manuscript. A generative AI tool was used in a limited, non-substantive capacity to support language refinement (grammar, clarity, and academic tone), APA 7th edition reference formatting, and formatting of the manuscript in line with the journal's template requirements. The tool was not used to generate, analyse, or interpret research data, to conduct the literature search or study screening, or to develop the conceptual arguments, findings, or conclusions reported in this review, all of which remain the original intellectual work of the authors. The authors reviewed, edited, and take full responsibility for the accuracy, integrity, and originality of the final manuscript.

References

- Al-kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., & Alfandi, O. (2024). Ethical challenges and solutions of generative AI: An interdisciplinary perspective. *Informatics*, 11, Article 58. <https://doi.org/10.3390/informatics11030058>
- Al-Zahrani, A. M. (2024). Unveiling the shadows: Beyond the hype of AI in education. *Heliyon*, 10, Article e30696. <https://doi.org/10.1016/j.heliyon.2024.e30696>
- Aromataris, E., Fernandez, R., Godfrey, C., Holly, C., Kahlil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132–140.

- Asri. (2024). Utilisation of artificial intelligence in improving student achievement. *PARADOKS: Jurnal Ilmu Ekonomi*, 7(4), 462–466.
- Baethge, C., Goldbeck-Wood, S., & Mertens, S. (2019). SANRA—A scale for the quality assessment of narrative review articles. *Research Integrity and Peer Review*, 4, Article 5. <https://doi.org/10.1186/s4639-650-639-650>
- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052–1092. <https://doi.org/10.1007/s4639-650-0639-650>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, Article 43. <https://doi.org/10.1186/s4639-650-0639-650>
- Chinta, S. V., Wang, Z., Yin, Z., Hoang, N., Gonzalez, M., Quy, T. L., & Zhang, W. (2024). *FairAIEd: Navigating fairness, bias, and ethics in educational AI applications* (arXiv:2407.18745). arXiv. <https://doi.org/10.48550/arXiv.2407.18745>
- Chong, Y., & Dit, T. J. (2025). Exploring CoGrader to enhance student feedback & academic growth. *Asian Journal of Social Science Research*, 7(2). <https://hdl.handle.net/20.500.11937/99303>
- Critical Appraisal Skills Programme. (2024). *CASP cross-sectional studies checklist*. CASP.
- Deckker, D., & Sumanasekara, S. (2025). Student-centric ethical frameworks for AI-driven education: Participatory consent and data ownership in a global perspective. *International Journal of Science and Research Archive*, 15(3), 1491–1503. <https://doi.org/10.30574/ijrsra.2025.15.3.1920>
- Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21), Article e39630. <https://doi.org/10.1016/j.heliyon.2024.e39630>
- El Arab, R. A., Al Moosa, O. A., Abuadas, F. H., & Somerville, J. (2025). The role of AI in nursing education and practice: Umbrella review. *Journal of Medical Internet Research*, 27, Article e69881. <https://doi.org/10.2196/69881>
- European Commission. (2024). *Artificial Intelligence Act: Risk-based regulatory framework for AI*. European Commission.
- Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, 81, Article 14048. <https://doi.org/10.3389/bjbs.2024.14048>
- Gordon, M., Daniel, M., Ajiboye, A., Uraiby, H., Xu, N. Y., Bartlett, R., Hanson, J., Haas, M., Spadafore, M., Grafton-Clarke, C., Gasiea, R. Y., Michie, C., Corral, J., Kwan, B., Dolmans, D., & Thammasitboon, S. (2024). A scoping review of artificial intelligence in medical education: BEME Guide No. 84. *Medical Teacher*, 46(4), 446–470. <https://doi.org/10.1080/0142159X.2024.2314198>
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018*. Canadian Intellectual Property Office.
- Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J., & Shimada, A. (2024). Artificial intelligence in education: Implications for policymakers, researchers, and practitioners. *Technology, Knowledge and Learning*, 29, 1693–1710. <https://doi.org/10.1007/s1639-650-0639-650>
- Jiang, W., & Pardos, Z. A. (2021). Towards equity and algorithmic fairness in student grade prediction. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 608–617). Association for Computing Machinery. <https://doi.org/10.1145/3461702.3462623>
- Karran, A. J., Charland, P., Martineau, J. T., Ortiz de Guinea, A., Lesage, A. M., Sénéchal, S., & Léger, P. M. (2024). Multi-stakeholder perspectives on responsible artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 6, Article 100169. <https://doi.org/10.1016/j.caeai.2024.100169>
- Karim, H. H., Abd Hamid, N. B., & Che Rus, M. F. (2024). Exploring the impact and challenges of using artificial intelligence in an open, distance and digital education university: A conceptual paper. *Asian Journal of Social Science Research*, 6(2), 40–57. <https://doi.org/10.5281/zenodo.14602832>
- Kerman, T. N., Banihashem, S. K., Karami, M., Er, E., van Ginkel, S., & Noroozi, O. (2024). A systematic review

-
- of peer feedback in online higher education. *Education and Information Technologies*, 29, 763–813.
- Khan, S., Mazhar, T., Shahzad, T., Khan, M. A., Ur Rehman, A., Saeed, M. M., & Hamam, H. (2025). Harnessing AI for sustainable higher education: Ethical considerations, operational efficiency, and future directions. *Discover Sustainability*, 6, Article 23. <https://doi.org/10.1007/s4639-650-0639-650>
- Lartey, S., Cummings, G., Richter, S., & Montgomery, C. (2025). Equity, diversity and inclusion in nursing educational institutions: A scoping review. *Journal of Advanced Nursing*. Advance online publication. <https://doi.org/10.1111/jan.17061>
- Laufer, M., Leiser, A., Deacon, B., Perrin de Brichambaut, P., Fecher, B., Kobsda, C., & Hesse, F. (2021). Digital higher education: A divider or bridge builder? Leadership perspectives on edtech in a COVID-19 reality. *International Journal of Educational Technology in Higher Education*, 18, Article 51. <https://doi.org/10.1186/s4639-650-0639-650>
- Mac Fadden, I., García-Alonso, E.-M., & López Meneses, E. (2024). Science mapping of AI as an educational tool exploring digital inequalities: A sociological perspective. *Multimodal Technologies and Interaction*, 8(12), Article 106. <https://doi.org/10.3390/mti8120106>
- Mangal, M., & Pardos, Z. A. (2024). Implementing equitable and intersectionality-aware ML in education: A practical guide. *British Journal of Educational Technology*, 55, 2003–2038. <https://doi.org/10.1111/bjet.13484>
- Mangalindan, K. E., & Wyatt, T. R. (2025). Mitigating implicit bias in medical education assessments. *Perspectives on Medical Education*. Advance online publication. <https://doi.org/10.5334/pme.1716>
- Modén, M. U., Ponti, M., Lundin, J., & Tallvid, M. (2024). When fairness is an abstraction: Equity and AI in Swedish compulsory education. *Scandinavian Journal of Educational Research*, 1–15. <https://doi.org/10.1080/00313831.2024.2349908>
- Nguyen, T., Gardner, L. A., & Sheridan, D. (2022). A framework for applying artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, Article 100065. <https://doi.org/10.1016/j.caeai.2022.100065>
- Nguyen, T., Zawacki-Richter, O., & Biggs, J. (2022). Personalized learning and agentic AI in higher education. *Educational Technology Research and Development*, 70(4), 1021–1045.
- Organisation for Economic Co-operation and Development. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. OECD Publishing. <https://doi.org/10.1787/15df715b-en>
- Pellas, N. (2023). Undergraduate students' attitudes toward AI-generated video content creation. *Smart Learning Environments*, 10, Article 57.
- Pham, N., Hung, P. N., & Nguyen-Duc, A. (2025). Systematic mapping of fairness for machine learning software in education. *The Journal of Systems & Software*, 219, Article 112244. <https://doi.org/10.1016/j.jss.2024.112244>
- Pierrès, O., Christen, M., Schmitt-Koopmann, F. M., & Darvishy, A. (2024). Could the use of AI in higher education hinder students with disabilities? A scoping review. *IEEE Access*, 12, 27810–27828. <https://doi.org/10.1109/ACCESS.2024.3365368>
- Sarkis-Onofre, R., Catalá-López, F., Aromataris, E., & Lockwood, C. (2021). How to properly use the PRISMA statement. *Systematic Reviews*, 10, Article 117. <https://doi.org/10.1186/s1639-650-01671-z>
- Selwyn, N., & Szili, A. (2025). The role of 'AI advocate' teachers in mediating technological change across Australian schools. *Teachers and Teaching*, 1–16. <https://doi.org/10.1080/13540602.2025.2596832>
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358, Article j4008. <https://doi.org/10.1136/bmj.j4008>
- Shoval, H. D. (2025). Artificial intelligence in higher education: Bridging or widening the gap for diverse student populations? *Education Sciences*, 15(5), Article 637. <https://doi.org/10.3390/educsci15050637>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson,
-

-
- M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M639-650>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO.
- Weidener, L., & Fischer, M. (2023). Teaching AI ethics in medical education. *Perspectives on Medical Education*. Advance online publication. <https://doi.org/10.5334/pme.954>
- Whiting, P., Savović, J., Higgins, J. P. T., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J., & Churchill, R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews. *Journal of Clinical Epidemiology*, 69(1), 225–234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>
- Xu, X., Liu, J., Zheng, R., Lei, V. N.-L., & An, Q. (2025). Learners' perception of data privacy when using AI language models: Reflective diary analysis of undergraduates in China. *Acta Psychologica*, 260, Article 105491. <https://doi.org/10.1016/j.actpsy.2025.105491>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*. Advance online publication. <https://doi.org/10.1111/bjet.13370>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), Article 39. <https://doi.org/10.1186/s4639-650-639-650>
-