Article

Volume 11 (2025), pp. 27-50
American Journal of STEM Education: Issues and Perspectives

https://doi.org/10.32674/0x43s107

Artificial Intelligence and Higher Education: A Systematic Review

Arakhita Behera
Regional Institute of Education Ajmer, India

Pranzal Trivedi
Central University of Rajasthan, India

Suresh Kumar Patra
Central University of Rajasthan, India

Chandayya Makeni
Nagaland University, India

ABSTRACT

This systematic review explores the definitions and research surrounding Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) within higher education literature. The study aims to provide a comprehensive synthesis of FATE's evolution and the challenges and opportunities identified across the analyzed papers. The review encompasses 25 SCOPUS articles published between 2019 and 2023. FATE definitions were classified as either technical or descriptive, with some studies offering multiple definitions. Findings indicate a predominance of descriptive definitions, particularly for fairness, which are primarily quantitative. Ethics definitions are mostly qualitative, while fairness research tends to be quantitative. Future research could bridge the divide between experts and the public by integrating technical and descriptive definitions and combining qualitative and quantitative approaches.

Keywords: Accountability, Artificial Intelligence, Ethics, Fairness, Higher Education, Transparency

INTRODUCTION

The swift evolution of generative AI technologies, such as ChatGPT, has unlocked significant potential for educational innovation (Abbasi et al., 2025; Shahzad et al, 2025). These tools demonstrate proficiency in areas like coding, creative writing, and producing text and visuals (Denny et al., 2024). Since 2022, ChatGPT and similar AI systems have shown substantial promise for diverse applications, prompting interest in incorporating advancements from fields like computer vision and natural language processing into educational settings (Bengesi et al., 2024; Mai et al., 2024). However, the growing integration of AI in education has raised critical ethical concerns (Holmes et al., 2021).

Historically, human professionals managed quality assurance in higher education, but the rise of AI algorithms and data analytics has driven organizations to adopt AI for decision-making processes (Adams Becker et al., 2017; McDonald et al., 2025; Overono and Ditta, 2025). While general meanings of FATE exist, they often lack specificity for AI applications in higher education. As AI-driven decision-making becomes more common, examining FATE in this context is increasingly important. Analyzing prior studies provides valuable insights into explored themes, challenges, and lessons in artificial morality.

AI in education has been studied for nearly 30 years (Woolf, 2015), yet challenges and unrealized opportunities persist, partly due to the static nature of educational systems (Woolf et al., 2013). Although research output in this area has increased (Tang et al., 2021), studies often focus on qualitative rather than practical impacts. As AI becomes ubiquitous in education, ethical issues, such as flawed reasoning, biased data, and non-transparent decision-making, have emerged (Holmes et al., 2021). Extensive reviews highlight AI's potential for tailored teaching and learning (Woolf, 2015), including applications like pedagogical support, e-learning assessment profiling and institutional decision-making (Chen et al., 2020). Nevertheless, there is a clear gap in systematic reviews targeting FATE specifically within AI and university studies.

Bearman et al. (2022) caution that a poor understanding of AI's ethical implications could harm specific demographic groups or society broadly. Many AI systems function as opaque "black boxes," underscoring the need for robust ethical frameworks. This study addresses the scarcity of research on FATE in AI and education by synthesizing existing literature and offering a thorough analysis of FATE in higher education contexts. The research questions guiding this study are:

- Q1: Which FATE definitions—individually or collectively—have been discussed in research on artificial intelligence and higher education?

- Q2: How is FATE examined in research on higher education and artificial intelligence?

This work advances and applies FATE In the fields of AI and higher education. The methodologies portion describes our search strategy, analytical methodology, and theme classification. The results section discusses the findings of our two study objectives and gives a demographic summary of the examined works. The limitation and potential applications of FATE in AI-driven education are outlined after the discussion and conclusion section respectively.

ChatGPT and Generative AI in Education

The adoption of generative AI, notably ChatGPT, has ushered in a significant transformation in education. Adiguzel et al. (2023) underscore ChatGPT's capacity to reshape education by enabling personalized instruction, automating assessments, and delivering real-time feedback. Intelligent tutoring systems and adaptive learning platforms can customize content to meet individual student needs, boosting engagement and academic outcomes. Su and Yang (2023) introduce the IDEE framework—Identify objectives, determine automation level, ensure ethical use, and Evaluate effectiveness—as a guide for responsibly incorporating ChatGPT into educational settings.

ChatGPT supports AI literacy, aids language acquisition, and facilitates self-paced learning. It serves as a virtual tutor, assists with essay composition, and provides formative feedback (Su & Yang, 2023). Its conversational capabilities foster exploratory learning and critical discussions. Michel-Villarreal et al. (2023) suggest that ChatGPT promotes inclusivity by supporting non-native speakers, accommodating neurodiverse learners, and ensuring equitable educational opportunities. Educators also gain from its applications in lesson planning, automated feedback, and content creation.

Nevertheless, challenges remain. Adiguzel et al. (2023) highlight risks such as algorithmic bias, data privacy issues, and the potential to exacerbate educational disparities, stressing the need for teacher training and institutional backing. Academic integrity is a concern, as ChatGPT can bypass plagiarism detection tools and perform well on graduate-level exams, raising issues of overreliance and misuse (Michel-Villarreal et al., 2023). Guo et al. (2023) note that while ChatGPT is fluent, it sometimes lacks contextual understanding and critical depth compared to human experts. Ethical frameworks are essential, with Su and Yang (2023) advocating for bias monitoring, transparency, and accountability via the IDEE framework. Michel-Villarreal et al. (2023) propose a "thing ethnography" approach, treating AI as a collaborative participant in educational contexts to better understand its role.

The Use and Acceptance of AI in Educational Contexts

The adoption of AI in education is an expanding research field. Lavidas et al. (2024) investigated factors influencing students' intentions to use AI in humanities and social sciences, applying the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2). Their findings indicate that performance expectancy, habit, and hedonic motivation drive behavioural intention, while facilitating conditions and intention predict actual usage, highlighting both practical and emotional factors in AI acceptance (Lavidas et al., 2024). Liu et al. (2023) explored ChatGPT's impact on cognitive processes, noting its ability to boost creativity and reduce cognitive load, but cautioning that uncritical use may hinder deep learning and critical thinking.

Luger and Sellen (2016) examined user perceptions of intelligent personal assistants like Siri, revealing that while natural language interaction is valued, concerns about trust, transparency, and control remain—issues pertinent to AI adoption in education. The European Parliament (2024) emphasized AI's implications for democratic values, data protection, and bias mitigation, stressing the importance of regulatory and ethical frameworks in educational settings. Krouska et al. (2023) introduced an AI-driven recommendation system for e-learning, enhancing engagement and performance by delivering personalized learning paths.

Algorithmic Bias, Explainable AI, and Ethical Frameworks

Algorithmic bias in educational AI often stems from biased datasets and non-transparent models. Western scholars have critically addressed these challenges. For example, Raji et al. (2020) advocate for independent audits of black-box AI systems in decision-making to ensure accountability. Floridi and Cowls (2019) propose a five-principle ethical framework—beneficence, non-maleficence, autonomy, justice, and explicability—as a global foundation for AI ethics. In educational data mining, Slade and Prinsloo (2013) highlight the conflict between predictive analytics and student privacy, urging stronger ethical governance. These seminal works provide a robust conceptual basis for understanding FATE in educational AI systems.

Together, these studies indicate that while AI holds significant potential for education, its success depends on user trust, ethical design, and alignment with pedagogical goals. Tackling algorithmic bias, improving model transparency, and developing ethical frameworks are essential for unlocking AI's transformative capabilities in education.

RESEARCH METHOD

We systematically extracted FATE-related data and analyzed articles using a structured search process and thematic classification, as outlined below.

Search Process

The search was conducted exclusively on the Scopus database, selected for its extensive coverage, surpassing Web of Science by 20% in journal volume (De La Cruz-Lovera et al., 2019; Su et al., 2020; Giwa et al., 2022; Soni, Yadav and Patra; 2025). We focused on English-language articles published between 2019 and 2023, using the search query: (“Fairness” OR “Accountability” OR “Transparency” OR “Ethics” OR “FATE”) AND (“AI” OR “Artificial Intelligence”) AND (“Higher Education” OR “University”). The OR operator ensured inclusion of studies addressing one or more FATE components, while the AND operator narrowed the scope to AI in higher education.

Figure 1 presents the PRISMA flowchart, illustrating the search process. We filtered titles and abstracts from the 2,355 publications that were initially found in Scopus, including review articles about artificial intelligence and data science in higher education. The inclusion criteria included review articles that discussed the application of data science and artificial intelligence in higher education institutions. The criterion for exclusion 149 publications for retrieval were obtained after excluding research that were centered on the healthcare or secondary education sectors. After further screening, 186 papers were evaluated, with 39 unavailable. Of the remaining 51 papers, 35 addressed K–12 populations, 17 focused on academic integrity, and 9 had unsuitable research designs. During data extraction, 24 papers lacked relevant content for our research questions, leaving a final set of 25 papers.

Coding Procedure

Beyond generative AI, broader AI tools in education have prompted investigation into adoption, perception, and trust. To address our first research question, we have analyzed the meanings of FATE (Fairness, Accountability, Transparency, Ethics) terminology as presented or adopted in the reviewed studies. We coded and summarized the definitions or discussions of each FATE term found in the evaluated manuscripts. We extracted pertinent material on FATE concepts by searching for the phrases "Fair," "Accountability," "Transparency," and "Ethic" in the PDF versions of these articles. The collected material often referenced existing literature, while some studies introduced new or refined definitions of FATE terms, building on prior work. The papers included both proposed and cited definitions.

Figure 1

PRISMA Chart of Study Selection Process

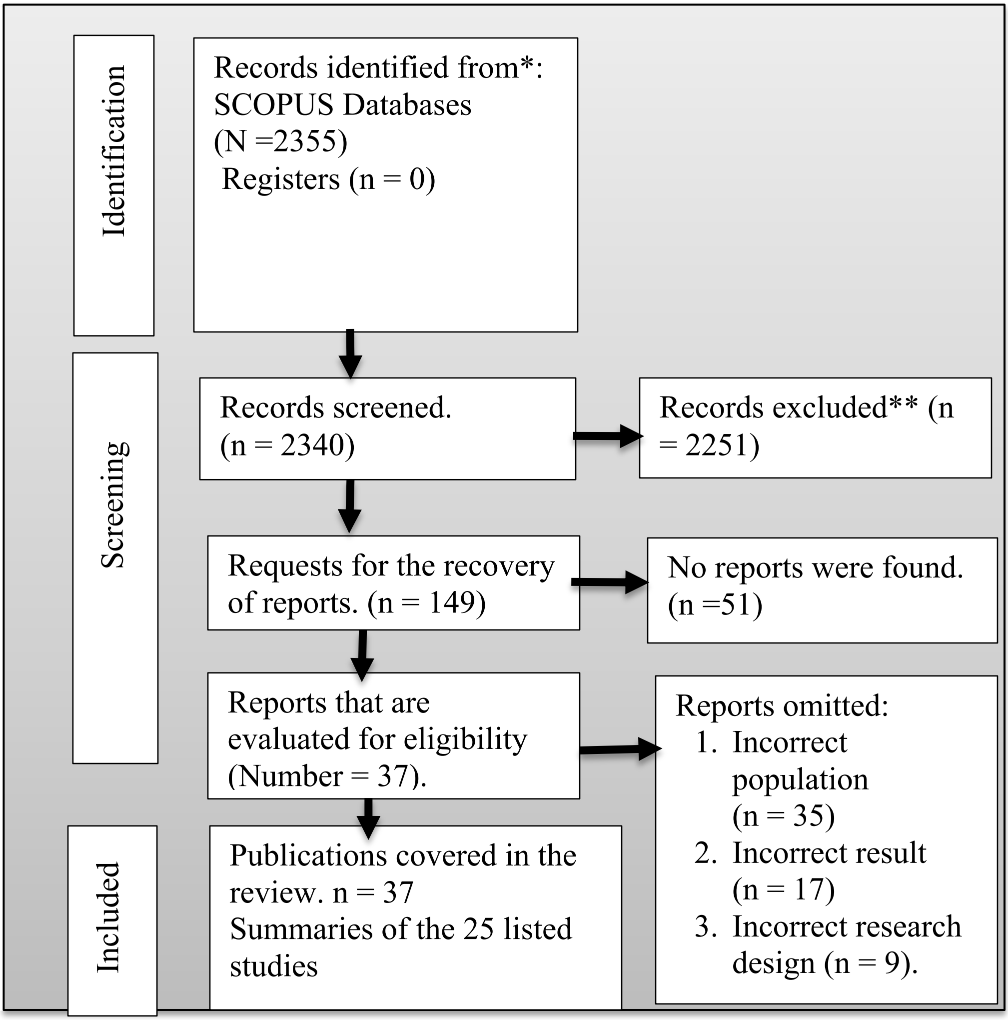


Figure 2

Publication year of reviewed literature

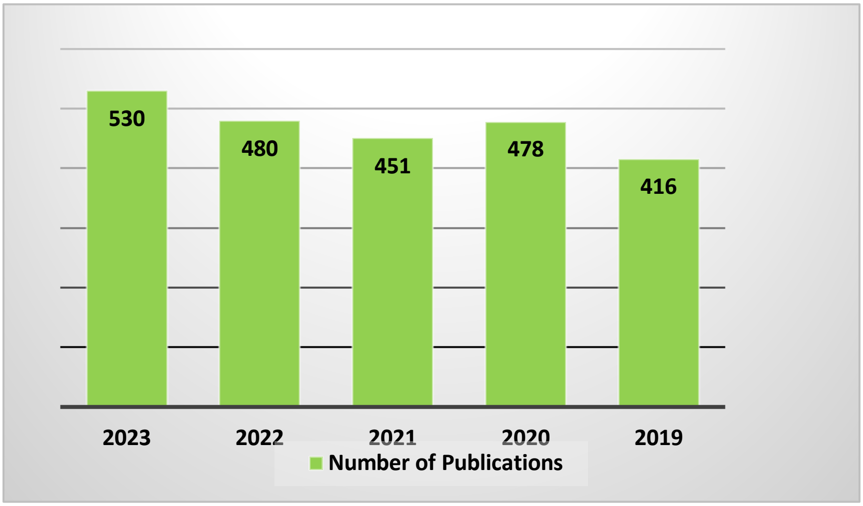
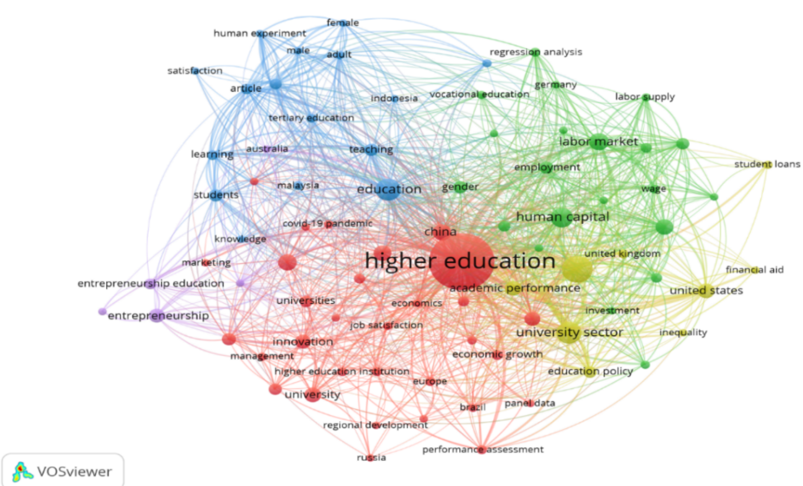


Figure 3

Predominant article title keywords



As expected, the phrases that appear most often include education, learning, AI, fairness, and generic terms pertaining to AI goals or features of higher education (Figure 3).

In order to answer the second enquiry, we classified and examined the different kinds of FATE-related studies in the evaluated articles, paying particular attention to the main FATE word that was covered in each. This allowed us to identify trends, challenges, and opportunities in FATE research, as well as distinguish between studies emphasizing qualitative versus quantitative approaches and those with more in-depth analyses. We primarily extracted this information from the abstract, methods, results, and conclusion sections, aiming to highlight the key concept or feature of each study rather than providing a detailed methodological overview.

The stages in our procedure were as follows:

1. Search SCOPUS using the designated input string.
2. Download complete records of all relevant papers.
3. Assess the relevance of each article by reviewing its title and abstract.
4. Thoroughly examine each article to verify relevance, using the secondary inclusion and exclusion criteria outlined in section 2.1.
5. Collect demographic data from the selected studies.
6. Perform the FATE term search as outlined in section 2.2.
7. Extract definitions of FATE concepts from each study and record them in an Excel spreadsheet.
8. Classify definitions as technical (containing advanced mathematical terms or jargon) or descriptive.
9. Provide a summary of the qualitative and technical definitions of every FATE word as presented in the examined research.
10. Identify the primary FATE term in each paper by noting the most frequently mentioned term and determining whether it was analyzed qualitatively or quantitatively.
11. Categorize studies as qualitative (e.g., literature reviews, reflective analyses, focus groups, open-ended interviews) or quantitative (e.g., Likert-scale surveys, empirical analyses).
12. Provide a summary of the qualitative and quantitative study types for each FATE term across the reviewed papers.
13. Outline the main FATE study methodologies and findings from each paper.
14. In the discussion section, consolidate suggestions and future research trajectories derived from the comprehensive assessment of the studies, including limits and suggested future endeavors for each major FATE term addressed.

To strengthen reliability, a second coder independently classified a subset of definitions. Discrepancies were discussed and resolved through consensus, ensuring consistency in interpretation.

RESULTS

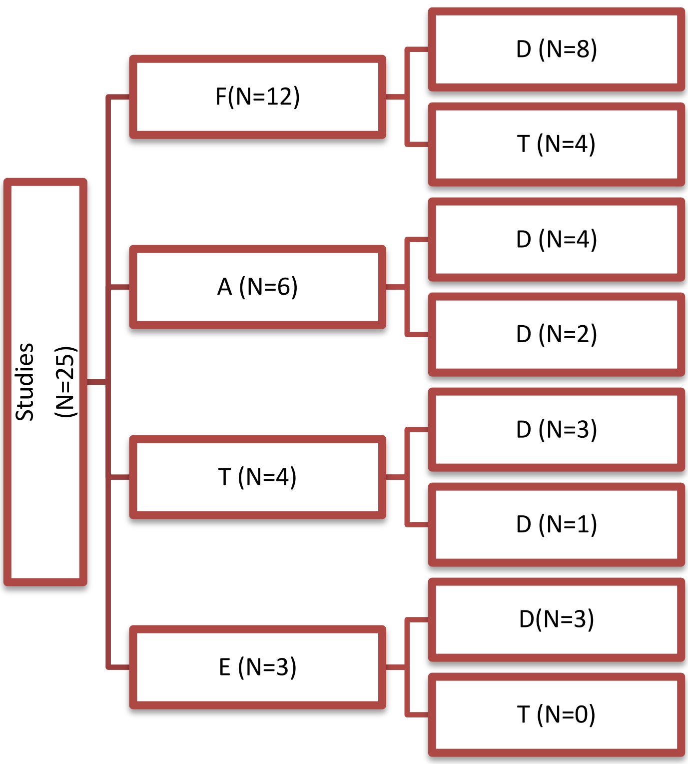
Our review of 25 peer-reviewed articles synthesized definitions and findings related to FATE in the context of artificial intelligence (AI) within higher education. Section 3.1 defines FATE terms, while Section 3.2 explores key themes, methodologies, findings, challenges, and future directions.

FATE definitions mentioned in the publications under review

Figure 4 illustrates the frequency of Fairness (F), Accountability (A), Transparency (T), and Ethics (E) in the reviewed studies, categorized as Technical (T) or Descriptive (D) definitions. Articles often focused on a primary FATE theme (e.g., Fairness) but included definitions of other FATE terms (e.g., Accountability). Below, we summarize each term's meaning, with detailed descriptive and technical definitions.

Figure 4

Frequency of each word of FATE in reviewed studies, distinguished by descriptive (D) and technical (T) definitions



Fairness

Fairness, the primary focus of 10 studies, emphasizes reducing bias and ensuring equitable AI outcomes in higher education, approached through descriptive and technical perspectives.

Fairness in AI for higher education aims to eliminate bias and promote equity. Shin et al. (2022) define fair algorithms as those avoiding discriminatory impacts, a view echoed by Kobis and Mehner (2021) and Kong et al. (2023), who prioritize protecting marginalized groups. Tenorio et al. (2023) stress avoiding biased AI outputs. Prinsloo and Slade (2017) propose five justice frameworks: utilitarian, rights-based, fairness/justice, common-good, and virtue. Casacuberta and Guersenzvaig (2019) apply Dreyfus' ethical expertise model, outlining a five-stage progression from novice to expert. Gorur et al. (2020) advocate for fairness literacy to foster understanding of social justice, inclusion, and privacy. Adiguzel et al. (2023) highlight risks of algorithmic bias amplifying educational inequities.

Fairness is quantified using statistical methods across the machine learning pipeline. Jiang and Pardos (2021) assess group fairness through metrics in dataset construction, model training, and inference. Mashhadi et al. (2022) apply batch classification to identify injustice hierarchies. Fang et al. (2020) develop mathematical models for individual and group fairness. Segal-Halevi et al. (2020) use ordinal algorithms for user-driven fairness evaluations. Pereira et al. (2021) employ SHAP to ensure equitable feature contributions based on additivity, consistency, and missingness criteria.

Accountability

Accountability, addressed in six studies, focuses on mechanisms to assign responsibility for AI outcomes in higher education.

Accountability involves holding AI creators, owners, and users responsible. Pagallo (2017) debates whether responsibility lies with data providers or developers. Ungerer and Slade (2022) identify accountability gaps in algorithmic systems. Bearman et al. (2022) explore AI's impact on authority dynamics. Kong et al. (2023) emphasize human accountability for AI decisions. Su and Yang (2023) integrate accountability into their IDEE framework for ethical AI use.

Transparency

Transparency, examined in four studies, emphasizes algorithmic explainability and clear outcomes.

Transparency ensures clarity in institutional policies and AI algorithms. Ungerer and Slade (2022) advocate for transparent student data practices and auditability, aligned with GDPR (Regulation, 2018). Pagallo (2017) stresses accessible data for fair processing. Kong et al. (2023) emphasize clear communication of AI's benefits and limitations. Gorur et al. (2020) propose ethical frameworks to enhance transparency. Beerkens (2022) notes challenges with transparency in large datasets. Pereira et al. (2021) prioritize interpretable predictive models. Luger and Sellen (2016) highlight user concerns about transparency in AI interactions, relevant to education. Transparency is achieved through designs that clarify black-box models. Litman et al. (2021) propose hybrid models combining feature-based and neural network approaches.

Ethics

Ethics, discussed in three studies, encompasses fairness, accountability, and transparency. Ethics in AI for higher education encompasses moral philosophy, data security, and privacy. Williams et al. (2020) explore computational and psychological ethical decision-making. Beerkens (2022) examines data ownership. Ungerer and Slade (2022) advocate for governance and clear policies. Kong et al. (2023) outline ethical principles: autonomy, beneficence/nonmaleficence, and fairness. Michel-Villarreal et al. (2023) propose a "thing ethnography" approach, viewing AI as a co-participant to enrich ethical understanding. Consequentialist perspectives prioritize equitable outcomes, while deontological approaches emphasize duties like transparency and consent. Segal-Halevi et al. (2020) use stochastic dominance to select ethical theories, aligning with consequentialist frameworks by optimizing outcomes.

The Evaluated Publications Examined FATE

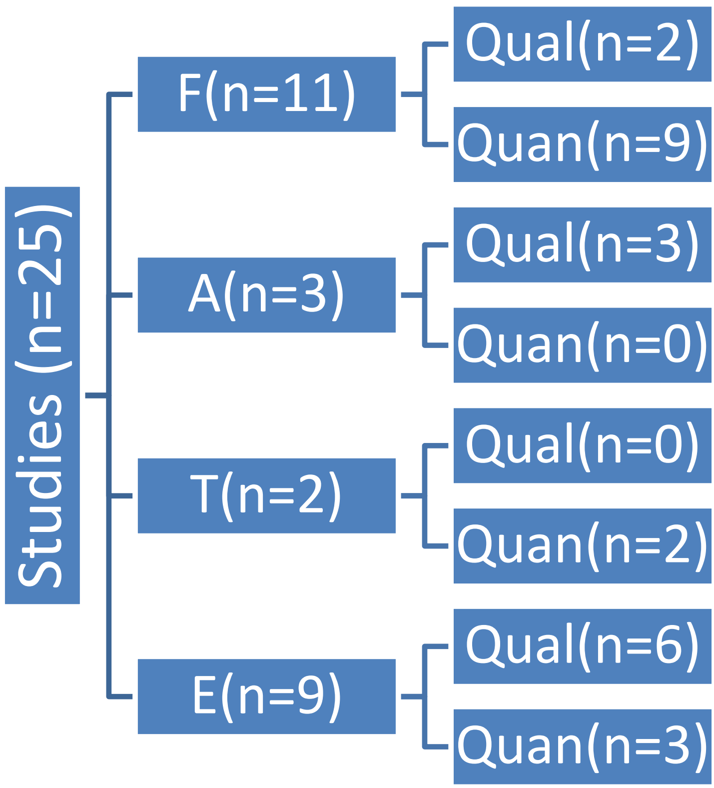
Figure 5 summarizes the analysis of FATE using qualitative (e.g., literature reviews, focus groups) or quantitative (e.g., surveys, empirical analyses) methods. Fairness and ethics are most prevalent, with accountability and transparency less studied.

Fairness

Fairness, addressed in 10 studies, is the most explored FATE concept (Figures 4 and 5). Mashhadi et al. (2022) use focus groups to examine algorithmic fairness, emphasizing transparency and interactivity. Baker and Hawn (2022) identify known and emerging biases, recommending diverse datasets. A consequentialist lens highlights mitigating harmful outcomes, while a deontological approach stresses equitable processes. Barbierato et al. (2022) generate synthetic datasets to test bias mitigation. Shulner-Tal et al. (2023) study fairness perceptions in AI recruitment systems. Afrin et al. (2022) analysed student perceptions of success predictors. Pereira et al. (2021) develop an explainable pipeline using SHAP. Litman et al. (2021) detect biases in automated scoring. Jiang and Pardos (2021) improve group fairness in grade prediction. Fang et al. (2020) ensure fair predictions via differential impact groupings. Segal-Halevi et al. (2020) explore utility differences. Lavidas et al. (2024) identify performance expectancy as a fairness factor in AI adoption.

Figure 5

Keywords (Fairness, Accountability, Transparency, Ethics) in reviewed studies, categorized by analysis method: qualitative (literature reviews, reflective analyses, focus groups, interviews) or quantitative (surveys, empirical analyses).



Accountability

Three qualitative studies address accountability. Beerns (2022) examines ethical implications of big data, advocating quality assurance. Pagallo (2017) links accountability to trust and proposes legal frameworks. Bearman et al. (2022)

identify research gaps in AI's impact on accountability. A deontological perspective emphasizes clear responsibility frameworks.

Transparency

Two studies quantitatively explore transparency. Sosnovsky and Brusilovsky (2015) enhance transparency via QuizGuide. Shin et al. (2022) examine algorithmic literacy's role in trust, advocating service trust initiatives. A consequentialist lens underscores transparency's role in equitable outcomes.

Ethics

Nine studies—six qualitative, three quantitative—address ethics. Tenorio et al. (2023) note limited collaboration in AI ethics literacy. Kobis and Mehner (2021) discuss ethical challenges in AI mentorship. Ungerer and Slade (2022) advocate for auditability. Prinsloo and Slade (2017) emphasize student risk profiles. Gorur et al. (2020) critique the lack of macro-ethical focus. Casacuberta and Guersenzvaig (2019) propose pre-reflective ethics approaches. The European Parliament (2024) stresses regulatory frameworks for ethical AI. A deontological lens prioritizes student privacy and consent. Williams et al. (2020) evaluate an ethics curriculum, identifying learning opportunities. Kong et al. (2023) survey students, emphasizing privacy and transparency. Lucic et al. (2022) use reproducibility to teach ethical AI challenges. A consequentialist perspective assesses ethics based on educational outcomes.

DISCUSSION

This systematic review of 25 Scopus-indexed studies examines the application of FATE in AI within higher education. The analysis reveals an uneven focus—**fairness** receives significant attention due to concerns about algorithmic bias, while **accountability** and **transparency** are underexplored, and **ethics** lacks precise definition.

By using **consequentialist** (outcome-oriented) and **deontological** (duty-oriented) ethical frameworks, this section synthesizes key themes, clarifies definitional boundaries, and identifies opportunities for future research. Foundational insights from Holmes et al. (2021), Inuwa-Dutse (2023), Raji et al. (2021), and Su and Yang (2023) guide this thematic analysis and underscore FATE's importance in shaping equitable and responsible AI use in education.

Fairness: Navigating Bias and Equity

The literature's focus on **fairness** reflects a growing awareness of algorithmic bias in educational AI systems. From a **consequentialist** lens, fairness emphasizes equitable outcomes, often operationalized through metrics such as group fairness or differential impact modeling (Jiang & Pardos, 2021). From a **deontological** perspective, fairness is a moral imperative to uphold justice, echoing frameworks like those of Prinsloo and Slade (2017).

Yet, defining fairness remains context-sensitive and complex, particularly concerning socio-demographic diversity. Baker and Hawn (2022) advocate for tailored fairness criteria that respond to specific educational settings and stakeholder groups.

To advance this field, future research should:

- Develop causal models of bias and test interventions using controlled components (Barbierato et al., 2022).
- Create **accessible fairness metrics** for broader public and educator understanding (Shulner-Tal et al., 2023).
- Explore **hybrid human-AI evaluation systems** that balance fairness with predictive accuracy (Litman et al., 2021).

Accountability: Defining Responsibility

Accountability, discussed in six studies, concerns assigning responsibility for AI-related decisions and outcomes in educational environments. A **deontological** framing emphasizes the duty to establish transparent responsibility among AI developers, data curators, and end-users (Pagallo, 2017). Beerkens (2022) discusses accountability gaps arising in big data-driven education, while Bearman et al. (2022) caution against blurring responsibilities between educators and AI systems.

The **IDEE framework** (Su & Yang, 2023), which integrates accountability into ethical AI use, supports **consequentialist goals** of equitable and traceable decision-making. However, most existing studies are qualitative, limiting the development of standardized accountability tools or metrics. Pagallo's (2017) call for meta-regulatory frameworks—echoed by the European Parliament (2024)—highlights the need for international harmonization.

Future research directions include:

- Developing **quantitative accountability metrics**, such as decision audit trails for AI systems.
- Investigating models of **shared accountability** between educators and AI systems to prevent miscommunication (Bearman et al., 2022).

- Exploring **global regulatory frameworks** to address disparities in AI accountability implementation across regions (Pagallo, 2017).

Transparency: Fostering Trust and Clarity

Transparency, though less studied, is critical for fostering trust, interpretability, and compliance. A **consequentialist** view emphasizes its role in achieving equitable outcomes, as demonstrated by systems like QuizGuide (Sosnovsky & Brusilovsky, 2015). From a **deontological** perspective, transparency is a right tied to informed consent, supporting policies such as GDPR (Regulation, 2018; Ungerer & Slade, 2022).

Multiple studies cite transparency as essential for trust, yet the literature remains limited. Shin et al. (2022) connect algorithmic literacy with perceived transparency, while Luger and Sellen (2016) underscore users' discomfort with opaque systems. Pereira et al. (2021) and Beerkens (2022) call for interpretable models and data clarity, while Michel-Villarreal et al. (2023) propose viewing AI as a co-participant to foster mutual understanding.

Most studies in this domain are **quantitative**, indicating a gap in qualitative research on lived user experiences with transparency.
To address this, future research should:

- Establish **multidisciplinary transparency standards** for adaptive AI systems.
- Design **user-centered interfaces** that improve explainability and trust (Luger & Sellen, 2016).
- Integrate **technical and descriptive approaches** to make transparency more accessible to non-expert stakeholders (Shin et al., 2022).

Ethics: Bridging Conceptual and Practical Gaps

Ethics is the **least represented** FATE concept in the dataset, though it inherently underpins the others. A **consequentialist** perspective emphasizes ethics as a framework to assess AI's long-term educational outcomes (Williams et al., 2020). Conversely, **deontological ethics** stress core principles such as privacy, autonomy, and justice (Kong et al., 2023).

Several studies address ethical concerns indirectly. For example, Holmes et al. (2021) and Inuwa-Dutse (2023) explore ethical complexity in AI design and use. Casacuberta and Guersenzvaig (2019) argue for **pre-reflective ethics**, while Gorur et al. (2020) critique the dominance of **micro-ethics** over systemic ethical thinking. Kobis and Mehner (2021) highlight challenges in implementing AI mentorship ethically.

Educational efforts like Kong et al.'s (2023) AI literacy program and Lucic et al.'s (2022) reproducibility-centered teaching model offer **practical pathways** to enhance ethical understanding.

However, the **descriptive orientation** of ethics research risks reducing ethics to surface-level norms. Scholars like Raji et al. (2021) call for **macro-ethical frameworks** that systematize ethics across institutions and cultures. Future work should:

- Foster **global collaboration** to standardize ethical AI literacy (Tenorio et al., 2023).
- Develop **micro-ethical documentation protocols**, adapted from fields like medicine (Lucic et al., 2022).
- Use **practical, scenario-based teaching tools** to make ethical reasoning tangible (Kong et al., 2023).

Ethical Statement

This study is a secondary analysis based solely on published academic literature and does not involve the collection of primary data or engagement with human participants. All reviewed sources are publicly available and have been appropriately cited in accordance with academic integrity standards.

Although no formal ethical approval was required, the research involved analysis of literature that discusses sensitive topics such as student data, algorithmic decision-making, and educational surveillance. Accordingly, the authors maintained a critical ethical stance in synthesizing these materials and adhered to institutional and disciplinary standards for responsible conduct of research.

CONCLUSIONS

This review synthesizes insights from 25 studies, identifying research gaps and offering practical recommendations for applying FATE in AI-driven higher education. It emphasizes actionable strategies for institutions, educators, and policymakers to ensure equitable and ethical AI integration, while highlighting the need for further research to refine FATE principles.

Institutional Strategies

The prominence of fairness research, driven by concerns about algorithmic bias (Adiguzel et al., 2023; Baker & Hawn, 2022), underscores the need for bias-mitigation measures. Consequentially, institutions should adopt tools like SHAP for explainable AI in grading and admissions to promote equitable outcomes

(Pereira et al., 2021). Deontologically, there is a duty to enhance faculty AI literacy, as outlined in Su and Yang's (2023) IDEE framework. Recommended actions include:

- Forming interdisciplinary AI ethics committees to oversee tool implementation and address risks of educational inequities (Adiguzel et al., 2023).
- Integrating fairness-focused training into faculty development, prioritizing accessible metrics for broader understanding (Shulner-Tal et al., 2023).
- Collaborating with tech developers to design context-sensitive fairness models that mitigate cultural biases (Afrin et al., 2022).

Curriculum Development

Curriculum enhancements are critical to embedding FATE principles, particularly given the broad scope of ethics (Gorur et al., 2020). Kong et al.'s (2023) AI literacy program illustrates how ethical awareness equips students to navigate AI's societal impacts. Institutions should:

- Incorporate macro-ethical perspectives into computer science and education curricula to address Gorur et al.'s (2020) critique of micro-ethical focus.
- Develop interdisciplinary courses blending technical AI skills with ethical decision-making, inspired by Lucic et al.'s (2022) reproducibility exercises.
- Use practical scenarios to teach pre-reflective ethical reasoning in AI contexts, as advocated by Casacuberta and Guersenzvaig (2019).

Policy and Governance

Accountability gaps (Bearman et al., 2022) and transparency challenges (Luger & Sellen, 2016) necessitate robust regulatory frameworks, as emphasized by Pagallo (2017) and the European Parliament (2024). Policymakers and practitioners should:

- Mandate transparent AI systems aligned with GDPR principles (Regulation, 2018) and Michel-Villarreal et al.'s (2023) co-participant approach.
- Fund longitudinal studies to evaluate FATE's long-term impacts, addressing Liu et al.'s (2023) concerns about AI's cognitive effects.
- Develop guidelines for educator-AI collaboration to clarify responsibility models (Beerkens, 2022).

LIMITATIONS OF THE STUDY

The review's exclusive reliance on Scopus may have excluded relevant literature from complementary databases such as Web of Science, IEEE Xplore, ERIC, or ACM Digital Library. These platforms may contain technical, interdisciplinary, or region-specific contributions that could enrich the findings.

Additionally, the focus on FATE specifically within the higher education context narrows the generalizability of insights to broader AI ethics debates. Several of the included studies provide only superficial treatments of FATE concepts, limiting the depth of analysis.

Finally, with a sample size of 25 studies, this review captures emerging patterns but may not fully represent the global diversity or methodological range of FATE research. Future reviews should consider expanded inclusion criteria, multilingual searches, and triangulation across academic and grey literature to build a more holistic evidence base.

IMPLICATIONS OF THE STUDY

This review underscores the need for clarity, balance, and rigor in FATE-related research. By applying both consequentialist and deontological lenses, it highlights how interdisciplinary and ethically informed approaches can improve AI use in higher education.

Future work should:

- Standardize FATE definitions across disciplines to reduce conceptual ambiguity.
- Conduct longitudinal studies that measure FATE's long-term educational impact, especially on cognition and learner equity (Liu et al., 2023).
- Expand qualitative inquiry into how students and faculty experience AI systems, promoting inclusive and adaptive design (Michel-Villarreal et al., 2023).
- Employ inter-rater reliability and transparency audits to strengthen methodological robustness in future reviews.

These efforts can directly inform **curriculum development**, **faculty training**, and **institutional governance policies**, fostering a more ethical, transparent, and accountable AI ecosystem. Aligning with international efforts—such as UNESCO's AI in education guidelines—can further position higher education institutions as global leaders in responsible AI integration.

FUTURE RESEARCH

To strengthen FATE's application, future research should integrate qualitative and quantitative methods to bridge descriptive and technical definitions. Key priorities include:

- In-depth studies on transparency and accountability, developing metrics such as AI decision audit trails (Pagallo, 2017).
- Longitudinal, reproducible, and open-access research to enhance FATE's rigor, as recommended by Holmes et al. (2021) and Raji et al. (2021).
- Global collaboration to standardize ethical AI literacy, responding to Tenorio et al.'s (2023) call for cooperation.

By adopting these strategies, higher education institutions, educators, and policymakers can leverage AI's potential while upholding FATE principles, fostering equitable, transparent, and ethically sound educational environments.

REFERENCES

Abbasi, B. N., Wu, Y., & Luo, Z. (2025). Exploring the impact of artificial intelligence on curriculum development in global higher education institutions. *Education and Information Technologies*, *30*(1), 547-581. https://doi.org/10.1007/s10639-024-13113-z.

Adams-Becker, S., Cummins, M., Davis, A., Freeman, A., Hall-Giesinger, C., & Ananthanarayanan, V. (2017). *NMC Horizon Report: 2017 Higher Education Edition.* The New Media Consortium.

Afrin, F., Hamilton, M., Thevathyan, C., & Majrashi, K. (2022). Investigating perceptions of AI-based decision making in student success prediction. *International conference on Artificial Intelligence in Education: AIED 2022,* 23, 315–319. https://doi.org/10.1007/978-3-031-11647-6_60

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

Barbierato, E., Della Vedova, M. L., Tessera, D., Toti, D., & Vanoli, N. (2022). A methodology for controlling bias and fairness in synthetic data generation. *Applied Sciences, 12*(9), 1-15. https://doi.org/10.3390/app12094436

Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education, 86*(2), 369–385. https://doi.org/10.1007/s10734-022-00937-2

Beerkens, M. (2022). An evolution of performance data in higher education governance: A path towards a 'big data' era? *Quality in Higher Education, 28*(1), 29–49. https://doi.org/10.1080/13538322.2021.1951451

Bengesi, S., El-Sayed, H., Sarker, M. K., Houkpati, Y., Irungu, J., & Oladunni, T. (2024). Advancements in generative AI: A comprehensive review of GANs, GPT, autoencoders, diffusion model, and transformers. *IEEE Access, 12*, 69812–69837. https://doi.org/10.1109/ACCESS.2024.3397775

Casacuberta, D., & Guersenzvaig, A. (2019). Using Dreyfus' legacy to understand justice in algorithm-based processes. *AI & Society, 34*(2), 313–319. https://doi.org/10.1007/s00146-018-0803-2

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. *Communications of the ACM, 67*(2), 56–67. https://doi.org/10.1145/3624720

European Parliament. (2024). *Artificial intelligence and democracy: The impact on human rights, privacy and ethics.* European Parliament Publications. https://www.europarl.europa.eu

Fang, B. L., Jiang, M., Cheng, P. Y., Shen, J., & Fang, Y. (2020). Achieving outcome fairness in machine learning models for social decision problems. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 444–450. https://doi.org/10.24963/ijcai.2020/62

Gorur, R., Hoon, L., & Kowal, E. (2020). Computer science ethics education in Australia – A work in progress. *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering,* 945–947. https://doi.org/10.1109/TALE48869.2020.9368375

Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J., & Wu, Y. (2023). How close is ChatGPT to human experts? Comparison corpus, evaluation, and detection. *arXiv Preprint, 1-20*. https://arxiv.org/abs/2301.07597

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education, 32*(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

Inuwa-Dutse, I. (2023). FATE in AI: Towards algorithmic inclusivity and accessibility. *arXiv Preprint, 1-12*. https://doi.org/10.48550/arXiv.2301.01590

Jiang, W., & Pardos, Z. A. (2021). Towards equity and algorithmic fairness in student grade prediction. *Proceedings of the 2021 AAAI/ACM Conference* *on AI, Ethics, and Society*, 4, 608–617. https://doi.org/10.1145/3461702.3462623

Kobis, L., & Mehner, C. (2021). Ethical questions raised by AI-supported mentoring in higher education. *Frontiers in Artificial Intelligence, 4*, 1-9. https://doi.org/10.3389/frai.2021.624050

Kong, S.-C., Cheung, W. M.-Y., & Zhang, G. (2023). Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. *Educational Technology & Society, 26*(1), 16–30. https://doi.org/10.30191/ETS.202301_26(1).0002

Krouska, A., Troussas, C., Sgouropoulou, C., & Virvou, M. (2023). A hybrid AI-based recommender system in e-learning environments. *Electronics, 13*(3). https://doi.org/10.3390/electronics13030498

Lavidas, K., Voulgari, I., Papadakis, S., Athanassopoulos, S., Anastasiou, A., Filippidi, A., Komis, V., & Karacapilidis, N. (2024). Determinants of humanities and social sciences students' intentions to use artificial intelligence applications for academic purposes. *Information, 15*(6), 1-17. https://doi.org/10.3390/info15060314

Li, C., Xing, W., & Leite, W. (2022). Using fair AI to predict students' math learning outcomes in an online platform. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2022.2115076

Litman, D., Zhang, H. R., Correnti, R., Matsumura, L. C., & Wang, E. (2021). A fairness evaluation of automated methods for scoring text evidence usage in writing. *International Conference on Artificial Intelligence in Education: AIED 2021,* 22, 255–267. https://doi.org/10.1007/978-3-030-78292-4_21

Liu, B., Li, W., Du, Y., & Zhu, J. (2023). Cognitive effects of generative AI in education: A framework for understanding ChatGPT's role in learning. *arXiv Preprint*, 1-18. https://arxiv.org/abs/2304.01487

Lucic, A., Bleeker, M., Jullien, S., Bhargav, S., & de Rijke, M. (2022). Reproducibility as a mechanism for teaching fairness, accountability, confidentiality, and transparency in artificial intelligence. *AAAI Conference on Artificial Intelligence,* 36, 12792-12800. https://doi.org/10.1609/aaai.v36i11.21533

Luger, E., & Sellen, A. (2016). "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1-12. https://doi.org/10.1145/2858036.2858288

Mai, D. T. T., Da, C. V., & Hanh, N. V. (2024). The use of ChatGPT in teaching and learning: A systematic review through SWOT analysis approach. *Frontiers in Education,* 9, 1-13. https://doi.org/10.3389/feduc.2024.1328769

Mashhadi, A., Zolyomi, A., & Quedado, J. (2022). A case study of integrating fairness visualization tools in machine learning education. *CHI Conference on Human Factors in Computing Systems,* 1-12. https://doi.org/10.1145/3491101.3503568

Merriam-Webster. (2023). Merriam-Webster dictionary. Merriam-Webster, 1-100. https://www.merriam-webster.com/dictionary.

McDonald, N., Johri, A., Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, 100121. https://doi.org/10.1016/j.chbah.2025.100121.

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education* Sciences, 13(9), 1-15. https://doi.org/10.3390/educsci13090856.

Overono, A. L., & Ditta, A. S. (2025). The rise of artificial intelligence: A clarion call for higher education to redefine learning and reimagine assessment. *College Teaching*, *73*(2), 123-126. https://doi.org/10.1080/87567555.2023.2233653.

Pagallo, U. (2017). From automation to autonomous systems: A legal phenomenology with problems of accountability. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence,* 17-23. https://doi.org/10.24963/ijcai.2017/3

Pereira, F. D., Fonseca, S. C., Oliveira, E. H. T., Cristea, A. I., Bellhauser, H., Rodrigues, L., Oliveira, D. B. F., Isotani, S., & Carvalho, L. S. G. (2021). Explaining individual and collective programming students' behavior by interpreting a black-box predictive model. *IEEE Access,* 9, 117097-117119. https://doi.org/10.1109/ACCESS.2021.3105956

Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room - the obligation to act. *International Conference on Learning Analytics and Knowledge,* 7, 46-55. https://doi.org/10.1145/3027385.3027406

Raji, I. Dgeq., Scheuerman, M. K., & Amironesei, R. (2021). You can't sit with us: Exclusionary pedagogy in AI ethics education. *ACM Conference on Fairness, Accountability, and Transparency,* 515-525. https://doi.org/10.1145/3442188.3445914

Regulation (EU) 2016/679 (General Data Protection Regulation). (2018). General data protection regulation (GDPR). *Official Journal of the European Union*, 24(1), 1-88. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26(2), 582-599. https://doi.org/10.1007/s40593-016-0110-3

Segal-Halevi, E., Hassidim, A., & Aziz, H. (2020). Fair allocation with diminishing differences. *Journal of Artificial Intelligence Research,* 67, 471-507. https://doi.org/10.1613/jair.1.11928.

Shahzad, M. F., Xu, S., & Asif, M. (2025). Factors affecting generative artificial intelligence, such as ChatGPT, use in higher education: An application of technology acceptance model. *British Educational Research Journal*, *51*(2), 489-513. https://doi.org/10.1002/berj.4084.

Shin, D., Rasul, A., & Fotiadis, A. (2022). Why am I seeing this? Deconstructing algorithm literacy through the lens of users. *Internet Research,* 32(4), 1214-1234. https://doi.org/10.1108/INTR-02-2021-0087

Shulner-Tal, A., Kuflik, T., & Kliger, D. (2023). Enhancing fairness perception - towards human-centred AI and personalized explanations understanding the factors influencing laypeople's fairness perceptions of algorithmic decisions. *International Journal of Human-Computer Interaction,* 39(7), 1455-1482. https://doi.org/10.1080/10447318.2022.2095705

Soni, H., Yadav, R. K., & Patra, S. K. (2025). Global impact of urbanization on ecosystems: A comprehensive bibliometric analysis. *Natural Hazards Research*, *5*(1), 21-35.

Sosnovsky, S., & Brusilovsky, P. (2015). Evaluation of topic-based adaptation and student modeling in QuizGuide. *User Modeling and User-Adapted Interaction,* 25(4), 371-424. https://doi.org/10.1007/s11257-015-9164-4

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355-366. https://doi.org/10.1177/20965311231168423

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2021). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments,* 1-19. https://doi.org/10.1080/10494820.2021.1875000.

Tenorio, K., Olari, V., Chikobava, M., & Romeike, R. (2023). Artificial intelligence literacy research field: A bibliometric analysis from 1989 to 2021. *ACM Symposium on Computer Science Education,* 54, 1-7. https://doi.org/10.1145/3545945.3569874

Ungerer, L., & Slade, S. (2022). Ethical considerations of artificial intelligence in learning analytics in distance education contexts. *Springer Briefs in Open and Distance Education,* 105-120. https://doi.org/10.1007/978-981-19-0786-9_8

Williams, T., Zhu, Q., & Grollman, D. (2020). An experimental ethics approach to robot ethics education. *AAAI Conference on Artificial Intelligence*, 34, 13485-13492. https://doi.org/10.1609/aaai.v34i09.7067

Woolf, B. P. (2015). AI and education: Celebrating 30 years of marriage. *AIED Workshops,* 4, 38-47.

Woolf, B. P., Lane, H. C., Chaudhri, V. K., & Kolodner, J. L. (2013). AI grand challenges for education. *AI Magazine,* 34(4), 66-84. https://doi.org/10.1609/aimag.v34i4.2490

ARAKHITA BEHERA, PhD, Assistant Professor in Economics of the Department of Social Science and Humanities at Regional Institute of Education, Ajmer. His major research interests include sustainable tourism and hospitality, agricultural economics, women's studies, and higher education.
Email id: beheraarakhita996@gmail.com

PRANZAL TRIVEDI is pursuing Master in Economics in the Department of Economics at School of Social Science at Central University of Rajasthan. His major research interest includes macroeconomics and growth and development.
Email: pranzal03@gmail.com.

SURESH KUMAR PATRA, PhD, is an Assistant Professor in Economics in the Department of Economics at School of Social Science at Central University of Rajasthan. His major research interest includes macroeconomics, and growth and development, applied econometrics, climate change and disaster management.
Email: suresh.patra@curaj.ac.in.

CHANDAYYA MAKENI, PhD, is an Assistant Professor in Economics in the Department of Economics at Nagaland University. His major research interest includes Agriculture Economics, Labour Economics, and Development Economics Email: makenichandu@gmail.com

Acknowledgement: We express our gratitude to the Editor and the anonymous reviewers for their perceptive criticism, which significantly improved the quality of the paper.

Note: The authors did not use OpenAI's ChatGPT or any other AI tools in the drafting, editing, or refining of this manuscript. All content was generated, reviewed, and refined solely by the authors.
