INTERNATIONAL SOCIAL SCIENCES AND EDUCATION JOURNAL (ISSEJ)
ISSN: 3005-3463 (Online)
DOI: 10.61424/issej
Journal Homepage: https://bluemarkpublishers.com/index.php/ISSEJ

ISSEJ
BLUEMARK PUBLISHERS

Ethical and Responsible Use of AI in Education: Insights from Bibliometrics

Stephen Naul[1], **John Kenneth Amolata**[2], **Lavelle Compendio**[3], **Rio Katrina Avila**[4], **Cyrene Napoles**[5] and **Jomar Monte De Ramos**[6]
[123456]*Southern Leyte State University – Main Campus*

Corresponding Author: Stephen Naul **E-mail:** snaul@southernleytestateu.edu.ph

ARTICLE INFO

Received: October 19th, 2025
Accepted: November 28th 2025
Published: December, 12th 2025
Volume: 3
Issue: 4
DOI: 10.61424/issej.v3i4.613

KEYWORDS

ChatGPT, generative AI, academic integrity, bibliometric analysis, machine learning, educational technology, AI ethics

ABSTRACT

This study provides a comprehensive bibliometric analysis of global research on the ethical and responsible use of artificial intelligence (AI) in education. Using 699 peer-reviewed journal articles indexed in Scopus (1998–2025), the review examines publication trends, leading contributors, influential sources, collaboration patterns, and the conceptual structure of the field. Findings show a rapid rise in research output beginning in 2019 and peaking in 2024, driven largely by scholars in the United States, United Kingdom, China, Spain, and Australia. Co-citation and keyword analyses reveal three central knowledge domains: technological foundations, educational, and ethical frameworks emphasizing transparency, privacy, fairness, and responsibility. Despite this strong conceptual base, practice-oriented areas such as teacher training, professional development, and classroom-level implementation remain underexplored. The study highlights the need for broader global participation, stronger interdisciplinary collaboration, and more empirical work examining how ethical principles operate in real educational contexts. These insights offer direction for advancing responsible and inclusive AI integration in education.

1. Introduction

Artificial intelligence (AI) has moved from a largely experimental concept into an everyday tool in education. It now shapes classroom instruction, learning-management systems, assessment practices, and even institutional planning. These technologies promise meaningful gains—more tailored learning experiences, timely feedback, streamlined administrative work, and better use of resources (Usher & Barak, 2024). This expansion, however, has brought forward a set of urgent ethical questions. Concerns about algorithmic bias, opaque decision processes, threats to privacy, and uneven access to digital tools serve as reminders that technological progress carries real risks (Mouta et al., 2024).

Given the speed and scope of this growth across fields, researchers have begun to take stock of how knowledge in this area has developed. Bibliometric analysis offers a useful way to do this, providing a systematic lens for tracing publication trends and spotting shifts in research interests (Swindell et al., 2024). By mapping research output, citation patterns, collaborative networks, and thematic groupings, bibliometrics helps reveal both the foundations of current

understanding and the directions in which new work is moving (Kamali et al., 2024). It complements narrative and systematic reviews by offering a broad, data-driven view of the field.

Ethical considerations appear consistently in recent discussions on AI in education. Researchers emphasize the need for fairness, openness, accountability, and governance models that reflect not only instructional aims but also the core values of education (An et al., 2024). Such principles are essential; neglecting them risks deepening social inequalities, limiting learner autonomy, and creating new forms of dependence on automated systems (Swindell et al., 2024). Even with this growing attention, questions remain about whether current scholarship fully addresses the ethical, pedagogical, and practical challenges that surface when AI is used in everyday teaching environments (Fu & Weng, 2024).

This bibliometric study examines how scholars across the globe have engaged with the ethical and responsible use of AI in educational settings. Using a curated Scopus dataset that brings together research at the intersection of AI, education, and ethics, the study maps both the structure and the evolution of this expanding field of inquiry.

To frame this analysis, six guiding research questions are posed:

1. How has scholarship on the ethical and responsible use of AI in education developed across time?
2. Which countries, institutions, and authors have played leading roles in advancing this conversation?
3. Which journals, publications, and citation networks exert the greatest influence on how AI ethics in education is discussed and understood?
4. What thematic clusters and emerging research fronts can be traced within this literature?
5. How fully does existing scholarship engage with the practical, ethical, and pedagogical challenges that come with integrating AI into educational settings?
6. What gaps remain, and what directions should future research explore at the intersection of AI, ethics, and education?

2. Literature Review

Early writing on AI in education often presents a split narrative. Innovations such as adaptive assessments, intelligent tutoring systems, and data-informed guidance are portrayed as important milestones in teaching and learning. At the same time, these developments raise difficult ethical questions—among them algorithmic bias, heightened surveillance, and the potential narrowing of learners' independence (Kamali et al., 2024).

Across the wider field of responsible AI and digital ethics, researchers generally agree on a core set of values: fairness, accountability, transparency, privacy safeguards, human oversight, and mechanisms for contesting automated decisions. How consistently these principles are translated into educational tools and classroom practice, however, varies greatly (An et al., 2024).

Earlier reviews highlight several unresolved issues, including limited empirical testing of ethical AI frameworks, insufficient attention to teacher-centered design, and gaps between technological innovation and existing governance structures. Many scholars argue for approaches that are sensitive to local educational contexts, support curricular aims, and protect student rights (Mouta et al., 2024).

Building on these concerns, more recent studies stress that responsible AI cannot remain an abstract ideal. Ethical design needs to be connected to the day-to-day realities of teaching, with technology reinforcing teacher judgment rather than sidelining it. Work in this area calls for ethics to be integrated directly into curriculum planning and to strengthen, rather than diminish, teacher agency (Usher & Barak, 2024).

A complementary line of inquiry urges that digital ethics be embedded in professional development so teachers are equipped to engage critically with AI and use it responsibly in their practice (An et al., 2024; Mouta et al., 2024).

Scholarship on governance further emphasizes the need for clear accountability structures and participatory oversight when AI is adopted in educational systems. Without such measures, institutions risk deepening existing inequities and undermining public confidence (Castelló-Sirvent et al., 2024; Fu & Weng, 2024; Kurian, 2025). Several studies also point out that ethical concerns often arise in specific local contexts—such as under-resourced schools or

culturally diverse learning environments—highlighting the importance of flexible governance models that can adapt to varying needs (Kamali et al., 2024).

Recent work has begun mapping the thematic contours of AI ethics in education. One cluster focuses on student data governance and privacy, especially within large-scale learning analytics systems (Fu & Weng, 2024). Another investigates bias in algorithmic decision-making and its consequences for assessment and admissions (Kamali et al., 2024; Usher & Barak, 2024). A third examines broader sociocultural dynamics, including how AI reshapes teacher–student relationships and influences notions of authority and autonomy (Mouta et al., 2024; Swindell et al., 2024).

Collectively, this literature forms a strong conceptual base, yet it also reveals notable shortcomings. Empirical insights remain limited in comparison with theoretical discussions. Still missing are detailed examinations of how ethical frameworks function inside classrooms, how educators and learners experience AI in everyday practice, and how policy can support innovation while upholding equity and justice. These gaps reinforce the value of bibliometric approaches, which help clarify where scholarly attention has been concentrated and where further inquiry is urgently needed.

3. Methodology

This study used a bibliometric approach to examine how scholars have investigated the ethical and responsible use of artificial intelligence (AI) in education. Bibliometric analysis offers a structured way to detect and visualize research patterns, making it possible to follow publication trajectories, identify influential contributors, and map the intellectual contours of a rapidly developing field. Through this lens, the study captures emerging trends, key thematic areas, and the works that have shaped ongoing conversations about AI ethics in educational settings.

The dataset was drawn from Scopus using a targeted search strategy that combined terms related to artificial intelligence, education, and ethics or responsibility. To keep the focus precise, only peer-reviewed journal articles that directly addressed the ethical or responsible use of AI in educational contexts were included; publications outside the education domain or lacking substantive ethical engagement were removed. This screening process resulted in a curated body of work that provides a reliable basis for tracing the evolution of research in this area (Hasan & Lateef, 2023).

The analysis incorporated several core bibliometric techniques, including assessments of publication trends, citation and co-citation patterns, co-authorship networks, and keyword co-occurrence. Visualization tools were also used to reveal structural dynamics such as collaborative linkages, citation clusters, and emerging thematic groupings. Together, these methods offer a comprehensive view of the field's development, the organization of its knowledge base, and the research fronts that continue to influence discussions on responsible AI in education. This approach also brings existing gaps into clearer focus, positioning the field within wider debates on educational technology and digital ethics while highlighting avenues for future inquiry (Hasan & Lateef, 2023).

Step 1. Data sourcing and specification

Table 1 outlines the methodological standards used in this bibliometric analysis of the ethical and responsible use of artificial intelligence in education. To ensure rigor, the dataset was sourced exclusively from Scopus using a deliberately crafted search query. This query combined terms such as AI, Artificial Intelligence, and Generative AI with ethical descriptors including ethics and responsible, and was applied across article titles, abstracts, and keywords. The search covered all publications available up to September 2025, allowing the study to capture both early foundational work and the most recent developments in the field.

The breadth of the dataset reflects the interdisciplinary character of AI ethics in education, spanning computer science, engineering, social sciences, psychology, arts and humanities, business, and multidisciplinary research. To maintain consistency and quality, only peer-reviewed journal articles written in English and classified as final publications were retained. The selection was globally inclusive, drawing from all countries and territories without

restricting open-access status. The final dataset was exported in CSV format, providing a structured and replicable foundation for the bibliometric analysis.

Category	**Specific standard requirements**
Database	Scopus
Search query	(“AI" OR “Artificial Intelligence” OR “Generative AI” AND “ethics” OR “responsible”
Search category	Article, title, abstract, keywords
Time frame	Articles published until September 2025
Subject area	Computer science, Social Sciences, Arts and Humanities, Engineering, Business Management, and Accounting, Multidisciplinary, Psychology, Mathematics, Material`s Science, and Environmental Science.
Document type	Articles
Source Type	Journal
Language	English
Country/Territory	All
Publication Stage	Final
Open Access	All Open Access
Data export	CSV

Table 1. Data requirements of the bibliometric study

Step 2. Data screening

Figure 1 illustrates the PRISMA-based screening process used in this study. The initial Scopus search produced 24,675 records, with no exclusions applied at the outset. During screening, 23,976 records were removed for not meeting the inclusion criteria. Of these, 3,001 fell outside the designated subject areas, 2,999 were not the correct article type, and 3,005 were not published in journals. A further 2,997 records were excluded because they were not written in English, 3,000 were still in press, 3,002 were not final versions, 3,980 did not meet the open-access requirement, and 3,992 were classified as irrelevant to the study’s focus. After completing this process, 699 articles met all criteria and formed the final dataset for the bibliometric analysis.

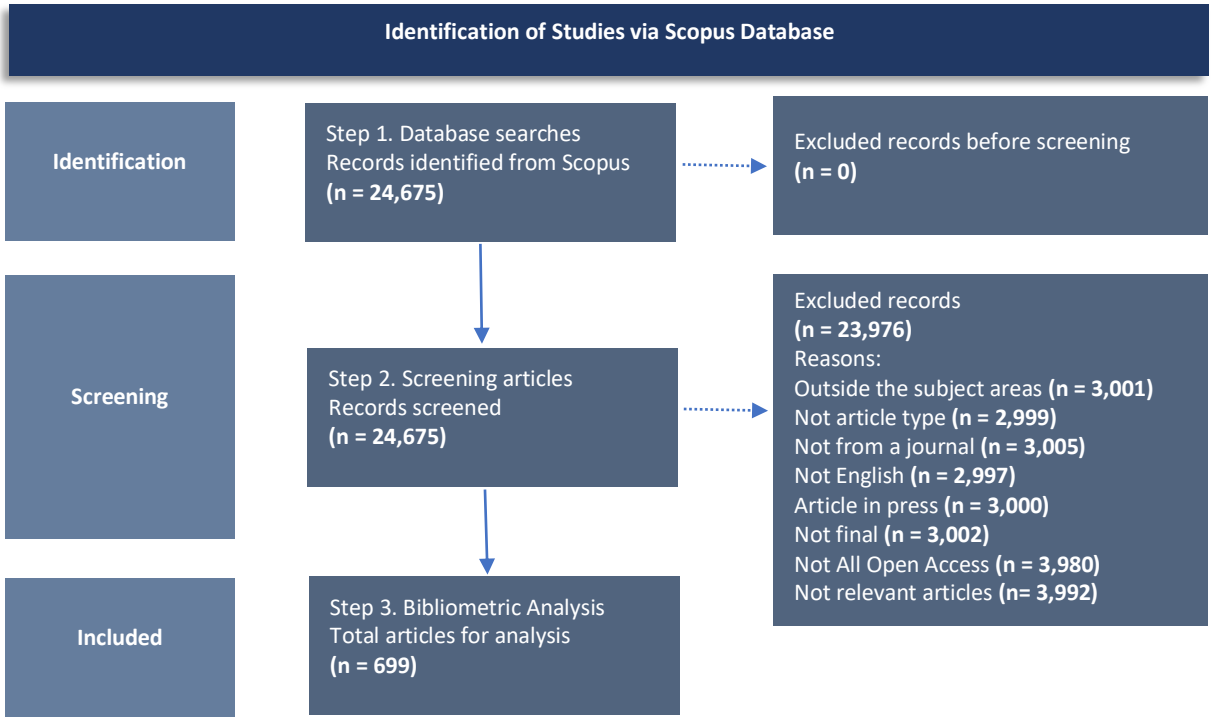


Figure 1. The bibliometric review process

Step 3. Data analysis

This study used a combination of Biblioshiny, VOSviewer, and Microsoft Excel to conduct and visualize the bibliometric analysis. Biblioshiny, the web-based interface of the Bibliometrix package in RStudio, was used to import the Scopus dataset and carry out trend mapping, co-citation analysis, keyword co-occurrence, and thematic evolution. These procedures helped identify influential authors, highly cited works, and the major thematic clusters shaping discussions on AI ethics in education. VOSviewer (version 1.6.20) was then employed to produce network-based visualizations, including co-authorship networks, institutional collaborations, citation structures, and keyword maps. These outputs revealed the field's key scholarly groupings and the intellectual connections among them. Microsoft Excel complemented these tools by supporting customized analyses of publication growth, geographic distribution, and citation patterns. Taken together, these platforms strengthened the precision, interpretability, and visual clarity of the study's bibliometric findings

4. Results

4.1 Bibliometric overview of studies

Table 2 provides an overview of the bibliometric dataset used in this study. Spanning 1998 to 2025, the dataset includes 699 journal articles published across 392 sources, reflecting the steady growth and interdisciplinary reach of research on the ethical and responsible use of AI in education. The field shows a strong annual growth rate of 18.73%, and with documents averaging 2.09 years in age, the literature is both rapidly expanding and highly current. The dataset contains 5,802 references and an average of 39.82 citations per article, indicating a growing influence within the broader research community.

The breadth of scholarship is further demonstrated by 2,107 indexed terms (Keywords Plus) and 3,787 author keywords, revealing wide thematic coverage across pedagogy, ethics, governance, and technology. Authorship patterns point to a notably collaborative environment: 4,308 authors contributed to the dataset, with an average of 9.78 co-authors per article and an international co-authorship rate of 27.36%. All included documents are peer-reviewed journal articles, ensuring consistency in academic rigor and publication quality.

Description	Results
Main information about data	
Timespan	1998:2025
Sources	392
Documents	699
Annual Growth Rate	18.73%
Document Average Age	2.09
Average Citations per Doc	39.82
References	5,802
Document contents	
Keywords Plus (ID)	2,107
Authors Keyword (DE)	3,787
Authors	
Authors	4,308
Authors collaboration	
International Co-Authorship	27.36%
Co-Authors per Doc	9.78%
Document types	
Article	699

Table 2. Description of the bibliometric data

4.2 Annual Scientific Production.

Figure 2 presents the annual publication output from 1998 to 2025. The trend shows a long period of minimal activity, with only a small number of studies appearing before 2015. A gradual rise begins around 2017, followed by a sharp acceleration from 2019 onward. The years 2023 and 2024 are particularly notable, with 2024 marking the highest output at 266 publications. In contrast, 2025 shows a drop to 103 documents. The trend line illustrates an overall exponential growth pattern in recent years, underscoring the rapidly increasing scholarly attention to the ethical and responsible use of AI in education.

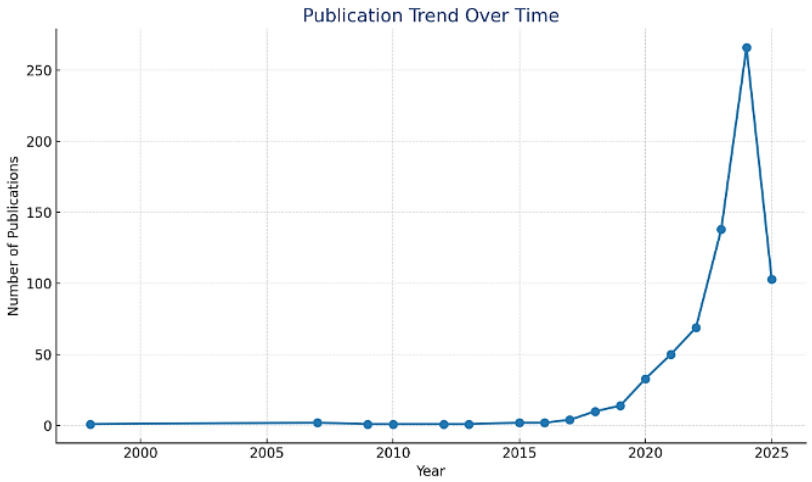


Figure 2. Annual scientific productions

4.3 Country Production Over Time.

Figure 3 shows the annual publication output of five leading countries—Australia, China, Spain, the United Kingdom, and the United States—from 1998 to 2025. Across all five countries, research activity remained minimal until about 2015. A gradual rise begins around 2018, followed by sharper increases after 2020. The United States stands out as the most prolific contributor, with output climbing steeply and exceeding 300 publications by 2025. The United Kingdom, Australia, China, and Spain also display clear upward trends, especially between 2021 and 2024, though

at lower overall volumes than the U.S. Taken together, these patterns suggest accelerating global engagement and expanding international collaboration in research on the ethical and responsible use of AI in education.

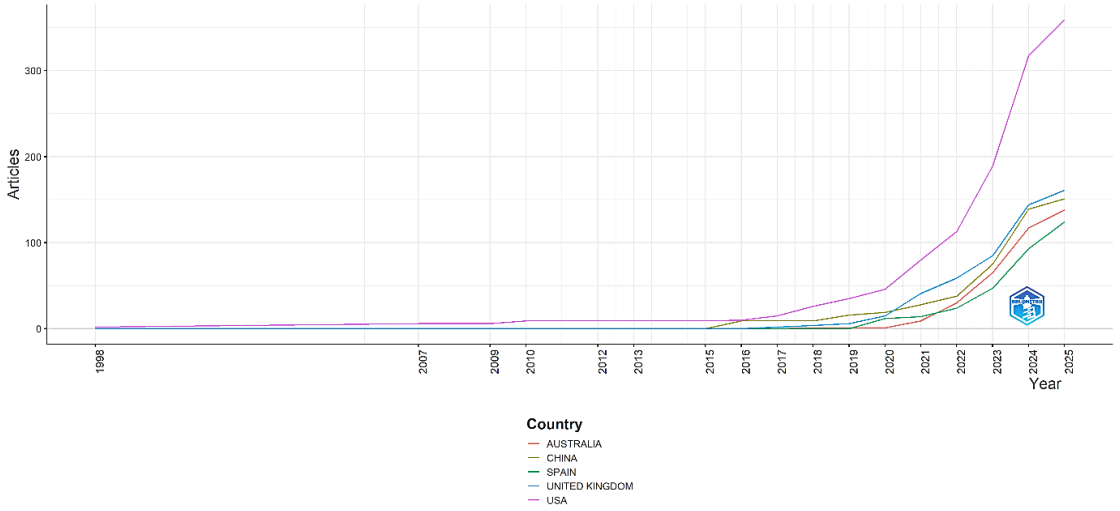


Figure 3. Country production

4.4 Most Relevant Affiliations.

Figure 4 displays the institutions with the highest publication output in this field. Monash University leads with 32 articles, underscoring its prominent research role. The category labeled "Not Reported" follows with 18 publications, indicating cases where affiliation data were missing or unclassified. The University of Hong Kong and Beijing Normal University appear next, with 14 and 13 publications, respectively. Other active contributors include Technische Universität München (12), the College of Engineering (11), Universidad de Murcia (11), Universidad de Salamanca (11), the Chinese University of Hong Kong (10), and King Faisal University (10). This distribution shows that although research is conducted across a wide range of institutions worldwide, a select group of universities contributes a particularly notable share of the scholarly output in this area.

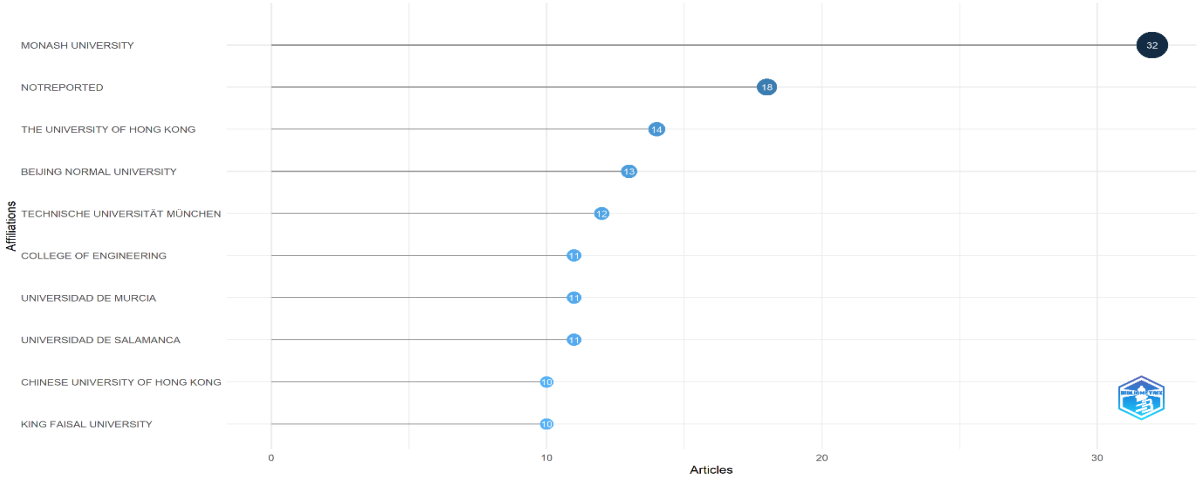


Figure 4. Most relevant affiliations

4.5 Most relevant authors.

Figure 5 depicts the publication activity of leading authors from 2020 to 2025, showing both their output and citation influence over time. Prominent contributors include Thomas K. F. Chiu, Javiera Atenas, and Sarah K. Burriss, each producing multiple publications in 2024, signaling their strong engagement with the field during this period. Natalia Díaz-Rodríguez appears among the earliest active authors, with work published as early as 2020. Other scholars—such as Simon J. Buckingham Shum, Roberto Martinez Maldonado, Thilo Hagendorff, and Leo Havemann—demonstrate steady contributions between 2022 and 2024. Overall, the figure indicates a widening distribution of scholarly activity, reflecting growing collaboration and participation among authors, with 2024 standing out as an especially productive year.

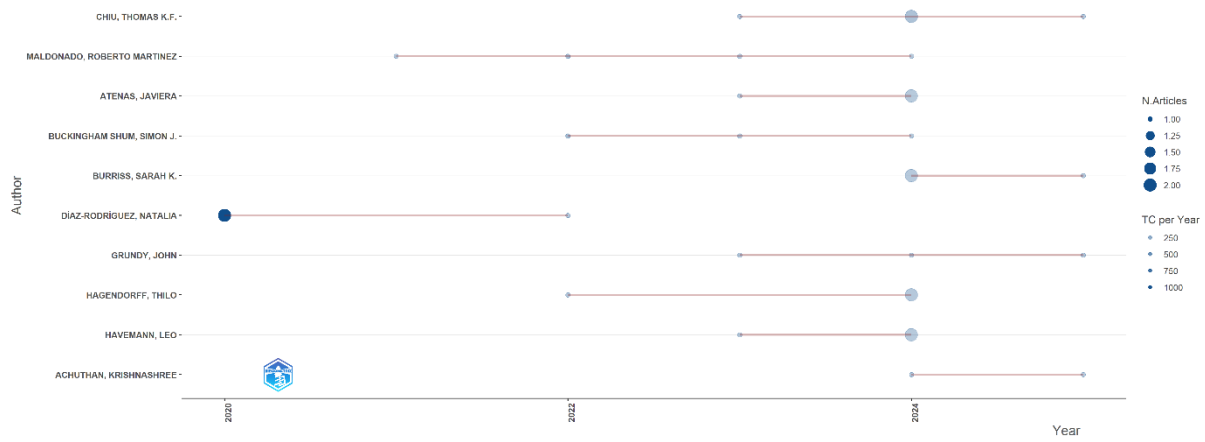


Figure 5. Most relevant authors

4.6 Co-Citation by Cited Authors.

Figure 6 presents the co-citation network of authors most frequently cited together in the literature on ethical and responsible AI in education. Generated using VOSviewer, the map reveals clusters of scholars who form distinct intellectual communities within the field. One prominent cluster includes Abd-Alrazaq, Adiguzel, and Bender, whose scholarship is closely linked to discussions on AI applications and ethical considerations. Another cluster centers on influential publication outlets—such as *Communications of the ACM*, *Lecture Notes in Computer Science*, and *Proceedings of the IEEE*—which function as key venues for knowledge exchange. Peripheral contributors, including Diakopoulos and Caulfield, introduce emerging perspectives that broaden the scope of the field. The density of connections across clusters reflects the interdisciplinary nature of this research area, drawing together insights from computer science, education, ethics, and policy. Overall, the co-citation pattern underscores that the intellectual foundations of responsible AI in education rest on an interplay of technical and sociocultural scholarship.

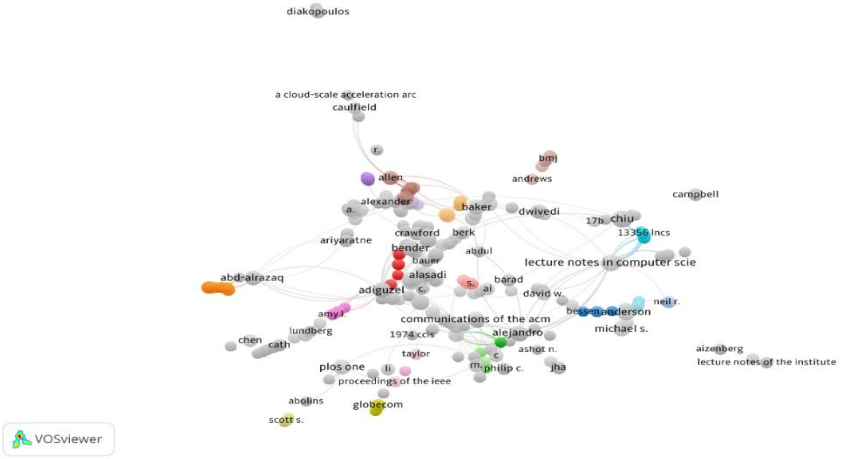


Figure 6. Co-citation by cited authors

4.7 Co-Citation by Cited Journals.

Figure 7 shows the network of journals most frequently cited together in research on the ethical and responsible use of AI in education. Distinct clusters emerge, reflecting the field's interdisciplinary foundations. *Computers & Education*, *Interactive Learning Environments*, and the *British Journal of Educational Technology* form a central educational technology cluster, highlighting the pedagogical grounding of much of this work. A second cluster—featuring *AI and Society*, *Philosophy and Technology*, and *Science and Engineering Ethics*—captures the ethical and philosophical dimensions of AI adoption. Meanwhile, journals such as *Information Fusion*, *Nature*, and *IEEE Intelligent Systems* represent the technical and computer science domain, linking AI advancements to their broader societal implications.

The presence of high-impact interdisciplinary outlets, including *Science* and *MIS Quarterly*, demonstrates how discussions of AI ethics in education extend beyond education itself, intersecting with governance, management, and the social sciences. Overall, the co-citation network shows that scholarship in this area is built on an integration of educational, ethical, and technical literatures.

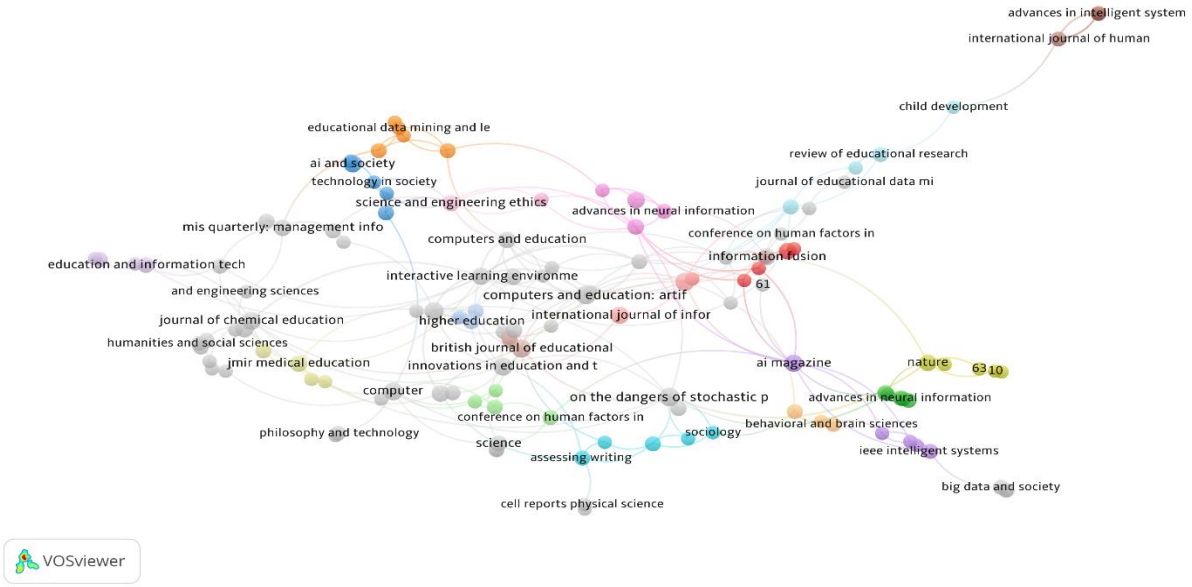


Figure 7. Co-Citation by Cited Journals

4.8 Co-Citation by Cited References.

Figure 8 presents the network of highly co-cited references that underpin scholarship on the ethical and responsible use of AI in education. Central to the map are influential works such as Crawford's *Atlas of AI*, Ajzen's Theory of Planned Behavior, and Burrell's analysis of algorithmic opacity, which ground the field in ethical and sociotechnical perspectives. Technical contributions—including Barredo-Arrieta's scholarship on explainable AI and Bolukbasi's work on bias in word embeddings—highlight the computational foundations of responsible AI debates. On the applied education side, studies published in the *Asian Journal of Distance Education* and research on generative AI in learning environments illustrate how theoretical insights are being translated into educational practice. Peripheral yet significant references, such as Bostrom's *Superintelligence* and Hagendorff's *Ethics of AI*, broaden the discussion by addressing long-term risks and governance challenges.

Collectively, the range of co-cited sources underscores that responsible AI in education is shaped by the intersection of ethical theory, technical development, and practical considerations within teaching and learning environments.

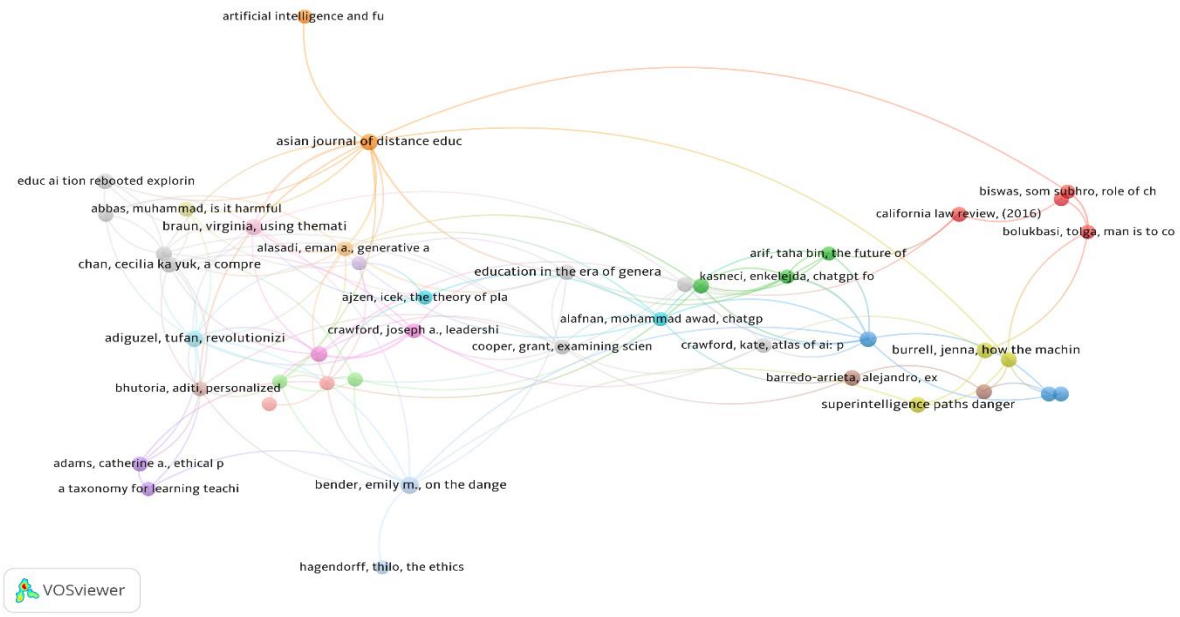


Figure 8. Co-Citation by Cited References

4.9 Conceptual Structure of the Field.

Figure 9 presents the factorial map of keywords generated through Multiple Correspondence Analysis (MCA), offering a view of the conceptual structure of research on ethical and responsible AI in education. Three prominent clusters emerge. On the left, keywords such as *large language models*, *language model*, and *generative adversarial networks* reflect the technical foundation of the field, highlighting growing interest in tools like ChatGPT and other generative AI systems. At the top, terms including *plagiarism*, *academic integrity*, *higher education*, and *students* point to pressing educational issues situated in classroom and institutional contexts. On the right, a dense cluster brings together *AI ethics*, *responsible AI*, *transparency*, *privacy*, *explainable AI*, and *decision-making*, underscoring ongoing debates around governance, accountability, and ethical design. Collectively, these clusters demonstrate the field's interdisciplinary character, where technological innovation intersects with pedagogical challenges and ethical considerations.

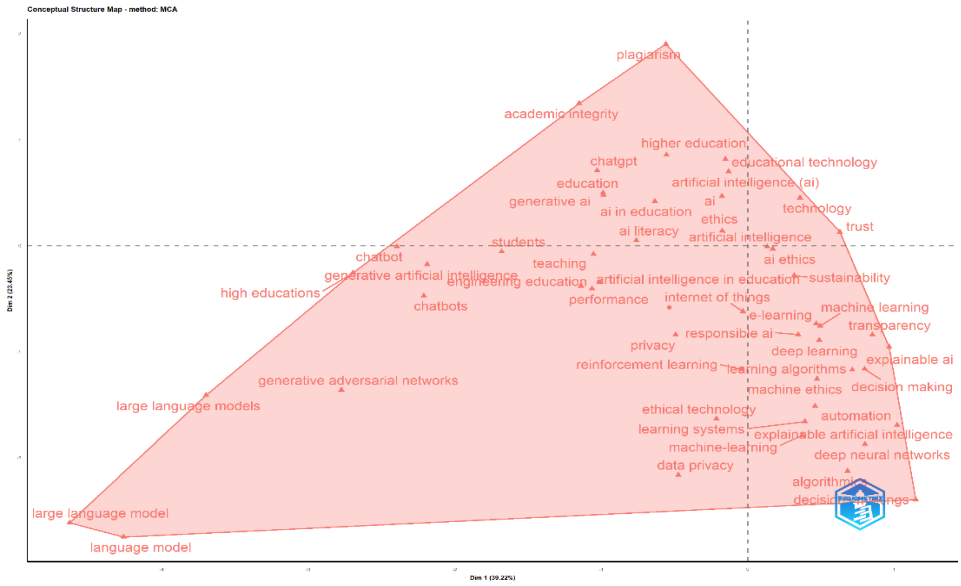


Figure 9. Conceptual structure map

4.10 Keyword Distribution.

Figure 10 presents a word cloud of the most frequently used keywords in the literature on ethical and responsible AI in education. *Artificial intelligence* and *machine learning* appear most prominently, reaffirming their central role in shaping the field. Closely associated terms such as *ethics*, *higher education*, *ChatGPT*, *deep learning*, and *learning systems* highlight the close connection between emerging AI technologies and their educational applications. Keywords like *decision-making*, *academic integrity*, and *responsible AI* signal the strong attention given to ethical considerations and governance frameworks. The inclusion of terms such as *generative AI* and *ethical technology* points to newer research frontiers driven by rapid advances in AI capabilities and their associated moral questions. Overall, the keyword distribution underscores that this body of scholarship is rooted in both technological development and ethical–pedagogical inquiry, reflecting the field's deeply interdisciplinary nature.

Figure 10. Keyword distribution

4.11 Thematic Evolution

Figure 11 presents the thematic map, which organizes research themes according to their centrality (importance within the field) and density (degree of development). In the Motor Themes quadrant, concepts such as *machine learning* and *ethical technology* occupy prominent positions, indicating their strong relevance and advanced development as key drivers of current scholarship. The Basic Themes quadrant features *artificial intelligence*, *ethics*, and *ChatGPT*, suggesting that these topics are highly relevant but continue to function as broad, foundational areas rather than specialized niches. In the Niche Themes quadrant, topics such as *convolutional neural networks* and *feature extraction* appear, reflecting specialized lines of inquiry that are well developed but less central to the field as a whole. The Emerging or Declining Themes quadrant highlights areas such as *teacher training*, *professional development*, and *generative artificial intelligence (GenAI)*, pointing to themes that are either gaining momentum or remain insufficiently explored in existing research.

Taken together, the thematic structure suggests that although technical and ethical debates continue to dominate the field, pedagogical concerns—especially those related to teacher readiness and professional development—are increasingly recognized as essential areas for future investigation.

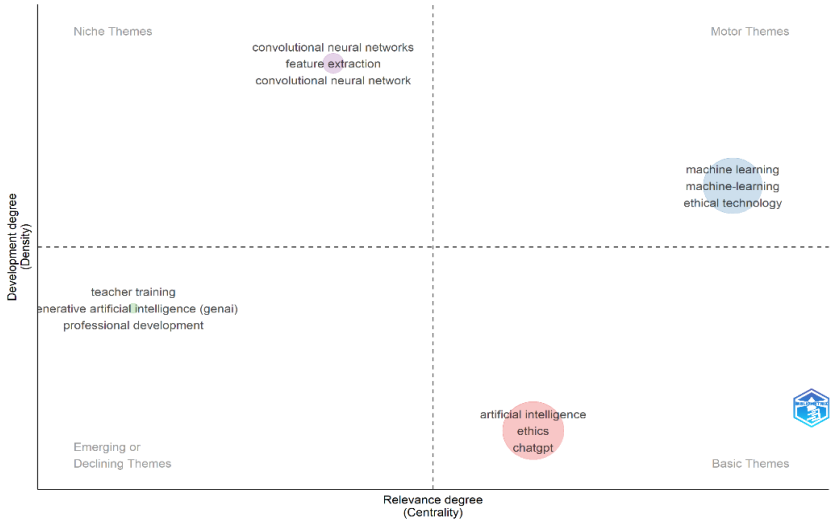


Figure 11. Thematic evolution

4.12 Co-Authorship by Authors.

Figure 12 illustrates the co-authorship network of leading researchers in the field of ethical and responsible AI in education. The visualization highlights several distinct collaborative clusters, each representing groups of scholars who frequently publish together. Notable clusters include the group comprising Herrera, Del-Ser-Lorente, Molina, and Poyatos, as well as another centered around Díaz-Rodríguez, Othmani, Berry, and Cully—both showing strong internal collaboration. Central figures such as Díaz-Rodríguez and Bennetot act as bridging nodes, connecting different clusters and enabling knowledge exchange across research communities. Although the network reflects active collaboration, many partnerships appear geographically or thematically concentrated, with limited interaction across clusters. This pattern points to opportunities for expanding global and interdisciplinary collaborations, particularly those that integrate technical, ethical, and pedagogical perspectives within the study of AI in education.

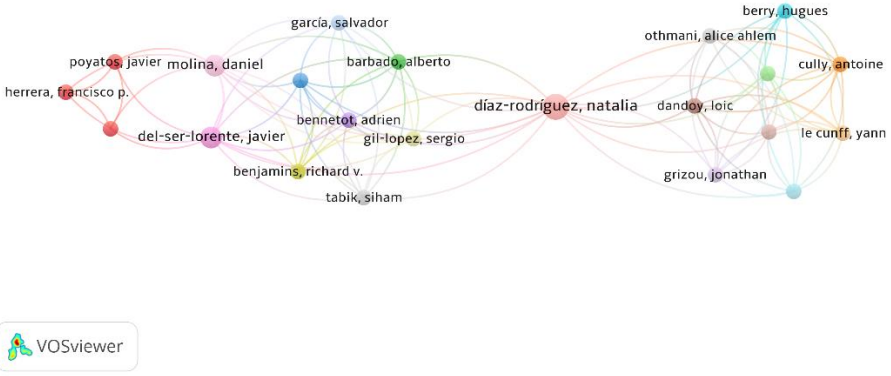


Figure 12. Co-authorship by authors

4.13 Co-Authorship by Countries.

Figure 13 illustrates the global collaboration network among countries contributing to research on the ethical and responsible use of AI in education. The United States, United Kingdom, Spain, Germany, Australia, and China emerge as central hubs, reflecting both their high publication output and strong international partnerships. Notable collaborative links connect the United States with several European countries—including the United Kingdom, Germany, Spain, and France—while China shows close ties with regional partners such as Hong Kong, Malaysia, and Indonesia. Countries including Saudi Arabia, India, and South Africa represent growing contributors, participating in collaborations with both Western and regional networks. In contrast, nations such as Ecuador, Israel, and Armenia appear more peripheral, indicating limited integration into global research exchanges. Overall, the map

reveals a highly connected international research landscape led by a small number of dominant countries, while highlighting opportunities to broaden collaboration with underrepresented regions to enhance inclusivity and diversify global perspectives.

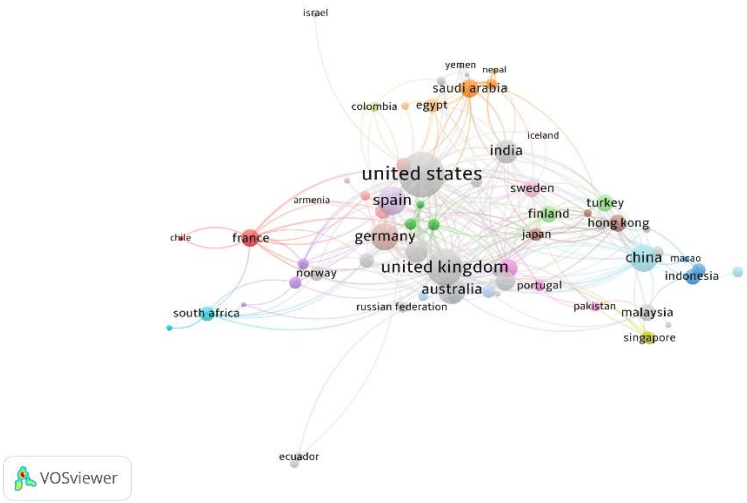


Figure 13. Co-authorship by countries

5. Discussion

Drawing on the insights from this bibliometric review, the paper identifies several key areas that contribute to a more comprehensive understanding of global scholarship on the ethical and responsible use of artificial intelligence (AI) in education.

First, the bibliometric evidence shows a marked surge in research output over the past decade, with publication activity accelerating after 2018 and reaching its peak in 2024. The United States, the United Kingdom, China, Spain, and Australia have emerged as the most prolific contributors, shaping much of the current research agenda. Their dominance reflects the advantages of strong institutional infrastructures, substantial research investment, and established interdisciplinary collaborations across education, computer science, and ethics (Fu & Weng, 2024; Kurian, 2025). In contrast, participation from developing countries remains limited, reducing the diversity of perspectives represented in global debates. Broadening contributions from underrepresented regions—through increased international collaboration and greater access to high-visibility publication outlets—is essential for ensuring that ethical frameworks for AI in education are informed by a wider range of contexts and experiences.

Second, patterns of institutional and author-level collaboration indicate a growing yet uneven network. Institutions such as Monash University, the University of Hong Kong, Beijing Normal University, and Technische Universität München stand out for their productivity, while scholars including Thomas K. F. Chiu, Javiera Atenas, and Natalia Díaz-Rodríguez have played influential roles in shaping recent discourse. However, co-authorship networks remain relatively concentrated, with collaboration occurring primarily within regional or institutional clusters. Expanding cross-border and cross-disciplinary partnerships—especially those linking technical, ethical, and pedagogical expertise—would help diversify the field and prevent scholarly influence from being concentrated primarily within technologically advanced nations.

Third, the co-citation and conceptual analyses reveal three core pillars underpinning the field: (a) technological foundations, including machine learning, deep learning, and generative AI; (b) educational applications, such as learning systems, academic integrity, and student-centered uses of AI; and (c) ethical and governance frameworks

focused on transparency, privacy, fairness, and responsibility. Although these clusters point to a strong conceptual base, they also underscore notable gaps. Practical, classroom-oriented concerns—such as teacher preparation, professional development, and context-aware governance—remain comparatively underdeveloped. This signals the need for more empirical research examining how ethical frameworks operate in real educational settings and how teachers and students navigate AI tools in practice.

Finally, the thematic evolution highlights new areas of inquiry, particularly around generative AI and ChatGPT. Much of the existing work approaches these technologies from ethical or governance perspectives, with less attention to their pedagogical implications or their impact on instructional practice. Themes related to inclusivity, equity, and cross-cultural dimensions of responsible AI also remain insufficiently examined, despite their relevance to global education systems. Just as research on technology business incubators revealed gaps in understanding gender and inclusivity dynamics, AI ethics in education must broaden its scope to address questions of marginalization, accessibility, and learner diversity.

Overall, although the field has established a solid conceptual and ethical foundation, the next phase of research must focus on bridging theoretical frameworks with practical realities. Strengthening international collaboration, increasing representation from underserved regions, and addressing classroom-level challenges will be crucial for ensuring that AI is integrated into education in ways that are technologically robust, ethically grounded, and socially inclusive.

6. Conclusion

This bibliometric analysis demonstrates that research on the ethical and responsible use of artificial intelligence in education has expanded rapidly in recent years, reaching its peak in 2024. This trajectory reflects growing global interest in ensuring that technological innovation is developed and applied in alignment with ethical principles. The field is driven predominantly by contributions from countries such as the United States, the United Kingdom, China, Spain, and Australia—nations supported by strong institutional infrastructures and active research networks. Although influential universities and authors have shaped much of the discourse, participation from developing regions remains limited, highlighting the need for broader inclusion to enrich global perspectives.

The intellectual structure of the field is anchored in three interconnected pillars: (1) technological foundations, including machine learning, deep learning, and generative AI; (2) educational applications, such as learning systems, academic integrity, and student engagement; and (3) ethical frameworks that emphasize transparency, privacy, and responsibility. While these pillars form a solid conceptual base, gaps persist in practice-oriented domains, particularly teacher training, professional development, and classroom-level implementation.

Moving forward, the field must focus on bridging theory and practice through more empirical studies that evaluate how ethical frameworks operate in real educational settings. Strengthening teacher preparation and ongoing professional development is essential to ensure educators are equipped to engage critically and confidently with AI tools. Issues of inclusivity, equity, and cultural sensitivity must also take priority, so that AI supports diverse learners rather than reinforcing existing disparities. In addition, enhanced international and cross-disciplinary collaborations, supported by governance models that adapt to varied educational contexts, will be critical for making responsible AI both globally representative and locally meaningful.

In sum, the field has established a strong foundation, but its progress now depends on translating ethical commitments into everyday educational practice, expanding global participation, and promoting systems that are both inclusive and accountable. By pursuing these directions, future research can help ensure that AI in education evolves not only as a technological advancement but also as a force for equity, integrity, and meaningful learning.

References

An, Q., Yang, J., Xu, X., Zhang, Y., & Zhang, H. (2024). Decoding AI ethics from Users' lens in education: A systematic review. *Heliyon*, *10*(20), e39357. https://doi.org/10.1016/j.heliyon.2024.e39357

Atenas, J., Havemann, L., Rodés, V., & Podetti, M. (2023). Critical data literacy in praxis: An open education approach for academic development. *Edutec. Revista Electrónica de Tecnología Educativa*, *85*, 49-67. https://doi.org/10.21556/edutec.2023.85.2851

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Burriss, S. K., Hutchins, N., Conley, Z., Deweese, M. M., Doe, Y. J., Eeds, A., Villanueva, A., Ziegler, H., & Oliver, K. (2024). Redesigning an AI bill of rights with/for young people: Principles for exploring AI ethics with middle and high school students. *Computers and Education: Artificial Intelligence*, *7*, 100317. https://doi.org/10.1016/j.caeai.2024.100317

Castelló-Sirvent, F., Roger-Monzó, V., & Gouveia-Rodrigues, R. (2024). Quo Vadis, University? A Roadmap for AI and Ethics in Higher Education. *Electronic Journal of E-Learning*, *22*(6), 34-51. https://doi.org/10.34190/ejel.22.6.3267

Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Computers and Education Open*, *6*, 100171. https://doi.org/10.1016/j.caeo.2024.100171

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, *7*, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Jin, Y., Echeverria, V., Yan, L., Zhao, L., Alfredo, R., Tsai, Y.-S., Gašević, D., & Martinez-Maldonado, R. (2024). FATE in MMLA: A Student-Centred Exploration of Fairness, Accountability, Transparency, and Ethics in Multimodal Learning Analytics. *Journal of Learning Analytics*. https://doi.org/10.18608/jla.2024.8351

Kamali, J., Alpat, M. F., & Bozkurt, A. (2024). AI ethics as a complex and multifaceted challenge: Decoding educators' AI ethics alignment through the lens of activity theory. *International Journal of Educational Technology in Higher Education*, *21*(1), 62. https://doi.org/10.1186/s41239-024-00496-9

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, *3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kurian, N. (2025). AI's empathy gap: The risks of conversational Artificial Intelligence for young children's well-being and key ethical considerations for early childhood education and care. *Contemporary Issues in Early Childhood*, *26*(1), 132-139. https://doi.org/10.1177/14639491231206004

Kusters, R., Misevic, D., Berry, H., Cully, A., Le Cunff, Y., Dandoy, L., Díaz-Rodríguez, N., Ficher, M., Grizou, J., Othmani, A., Palpanas, T., Komorowski, M., Loiseau, P., Moulin Frier, C., Nanini, S., Quercia, D., Sebag, M., Soulié Fogelman, F., Taleb, S., … Wehbi, F. (2020). Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities. *Frontiers in Big Data*, *3*, 577974. https://doi.org/10.3389/fdata.2020.577974

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2024). Uncovering Blind Spots in Education Ethics: Insights from a Systematic Literature Review on Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, *34*(3), 1166-1205. https://doi.org/10.1007/s40593-023-00384-9

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2025). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. *Education and Information Technologies*, *30*(3), 3343-3387. https://doi.org/10.1007/s10639-024-12946-y

Swindell, A., Greeley, L., Farag, A., & Verdone, B. (2024). Against Artificial Education: Towards an Ethical Framework for Generative Artificial Intelligence (AI) Use in Education. *Online Learning*, *28*(2). https://doi.org/10.24059/olj.v28i2.4438

Usher, M., & Barak, M. (2024). Unpacking the role of AI ethics online education for science and engineering students. *International Journal of STEM Education*, *11*(1), 35. https://doi.org/10.1186/s40594-024-00493-4

Uslu, S., Kaur, D., Rivera, S. J., Durresi, A., Babbar-Sebens, M., & Tilt, J. H. (2025). A Trustworthy and Responsible Decision-Making Framework for Resource Management in Food–Energy–Water Nexus: A Control-Theoretical Approach. *ACM Transactions on Intelligent Systems and Technology*, *16*(1), 1-29. https://doi.org/10.1145/3660640

Author's biography

Authors Name: Stephen V. Naul, LPT

Stephen V. Naul is a graduate of Bachelor of Elementary Education majoring in General Education and a Licensed Professional Teacher in the Philippines. He currently serves as a Research Assistant at Southern Leyte State University Main Campus and is pursuing a Master's in Education, majoring in Educational Management. With a strong academic background and a passion for improving the educational system, Stephen is dedicated to advancing his knowledge and contributing to the development of effective educational practices through research and publication.

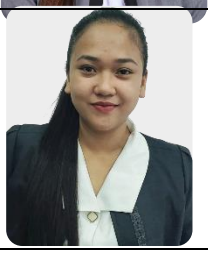
Authors Name: John Kenneth Amolata

John Kenneth L. Amolata is a graduate of the Bachelor of Science in Information Technology, majoring in Programming. He has an academic background in computing technologies, information systems, and problem-solving methodologies, with an interest in the application of information technology to research and innovation. His professional and academic experiences reflect a commitment to advancing knowledge in the field of information technology through research, system development, and practical technological solutions.

Authors Name: Lavelle Compendio

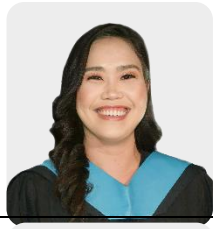
Lavelle Compendio is a graduate of Bachelor of Science in Civil Engineering and has completed 18 units of the Bachelor of Science in Secondary Education, major in Mathematics. She is currently employed at Southern Leyte State University as an Administrative Staff and is pursuing a Master's degree in Management. She is strongly committed to continuous professional development and actively seeks to enhance her expertise through research, innovation, and academic publication.

Authors Name: Rio Katrina Avila



Rio Katrina Avila is a Science Research Assistant and a Bachelor of Science in Biotechnology graduate. She is actively involved in research support and academic initiatives, contributing to scientific studies, research coordination, and capacity-building activities within the university. With a strong background in biotechnology and experience in intellectual property (IP) awareness and research-related IP support, she helps promote responsible research practices, protection of research outputs, and scientific innovation with meaningful societal impact.

Authors Name: Cyrene Napoles, LPT



Cyrene A. Napoles holds a degree in Bachelor of Technology and Livelihood Education and is both a Licensed Professional Teacher and a Civil Service passer in the Philippines. She is currently working as a Research Assistant while pursuing her Master of Technology Education. With her technical background and growing research experience, Cyrene is committed to strengthening her professional capabilities and contributing to the improvement of educational systems through innovative studies, scholarly engagement, and active participation in academic development.

Authors Name: Jomar Monte De Ramos, LPT

Jomar S. Monte De Ramos is a graduate of the Bachelor of Technology and Livelihood Education, major in Home Economics, and is a Licensed Professional Teacher. He has been practicing his profession for three years in a private educational institution. He is currently serving as a Research Assistant at Southern Leyte State University–Main Campus. His dedication and passion for teaching drive him to continually elevate his profession by navigating the field of education and applying insights gained from research findings.