

The Eurasia Proceedings of Educational and Social Sciences (EPESS), 2026

Volume 49, Pages 87-98

ICRET 2026: International Conference on Research in Education and Technology

Application of Artificial Intelligence in Educational Evaluation: A Bibliometric Analysis

Wahyu Muhammad Ramdhani
Yogyakarta State University

Aman Aman
Yogyakarta State University

Endang Wati
Yogyakarta State University

Abstract: The purpose of this study's analysis is to map trends in research publications on the application of artificial intelligence to educational evaluation, using bibliometric analysis of data from Scopus. After entering keywords related to artificial intelligence, evaluation, and education at the data collection stage in Scopus, the author analysed 159 articles selected using the PRISMA method. The next step was to conduct the analysis using R and VOSviewer software. The results of the study show that research on the application of artificial intelligence in educational evaluation began to emerge in 2021 and increased significantly through 2025, with the highest number of publications recorded in 2025. The United States, Australia, and China occupy the top three positions in terms of the number of publications and citations, confirming their significant influence in research on the application of artificial intelligence in educational evaluation. Meanwhile, the most active institutions in this research are the University of the Health Sciences and Rensselaer Polytechnic Institute. In the study of artificial intelligence applications in educational evaluation, new trends have emerged, including 'AI literacy' and 'artificial intelligence education,' which still require further exploration.

Keywords: Artificial Intelligence, Evaluation, Education

Introduction

The development of artificial intelligence (AI) technology has had a significant impact on the field of education, including the evaluation of the learning process. Evaluation activities are a crucial element in the education system because they serve not only to evaluate student learning achievements, but also to provide useful feedback to teachers and institutions to improve the quality of teaching. However, conventional evaluation methods often face various obstacles, such as subjective assessments, time constraints, and difficulties in comprehensively assessing cognitive and non-cognitive aspects. In this context, artificial intelligence (AI) has emerged as an innovative alternative to improve objectivity, efficiency, and accuracy in educational evaluation (Husain, Mohd Matore, and Mohd Hamzah, 2024). The application of AI in educational evaluation has evolved in various ways. For example, automated assessment systems based on Natural Language Processing (NLP) can assess student essays quickly and consistently, while machine learning can be applied to predict student performance based on existing learning data. In addition, technologies such as Intelligent Tutoring Systems (ITS) enable personalised learning while assessing student progress in real time (Phạm and Hoàn, 2024). Therefore, AI not only functions as a tool in evaluation but also as a diagnostic tool that can provide deep insights into student progress.

Along with the increasing interest in this issue, many bibliometric studies have been conducted using the Scopus database to map the development of research related to AI in education. For example, a study conducted by Phạm and Hoàn (2024) analysed 686 articles listed in Scopus on the use of AI in teacher education during the period 1991-2023. Their findings show a significant growth trend in publications, with the journal *Computers and Education* as the main channel, and the topic of Intelligent Tutoring Systems as the most influential. On the other hand, Husain and colleagues (2024) researched teacher performance assessment in higher education, which has shifted from traditional methods to AI-based competency frameworks and real-time analytics.

Similar research has also been conducted by Irwanto (2025), who studied the application of artificial intelligence in K-12 education in Asia, covering publications from 1996 to 2025. The results of this study show that although there has been a significant increase in the number of publications, research is still concentrated in a few countries, and its distribution is not yet evenly spread across Asia. This indicates a lack of research that needs to be addressed, especially in developing countries. In addition, Nusantara et al. (2023) sought to explore the integration of AI into pedagogy through a bibliometric analysis of 94 articles indexed in Scopus during the period 2018 to 2023. The results of the study emphasise that the collaboration between AI and innovative pedagogical methods can open up new opportunities in how student learning is assessed. Although some of these studies have had a significant impact in describing the development of AI research in education, there is still a lack of studies that specifically discuss the application of AI in educational evaluation. Many studies focus more on pedagogical aspects, teacher education, or the general use of AI. Therefore, bibliometric research emphasising the topic of educational evaluation is crucial to conduct and represents an innovation.

From this explanation, it can be concluded that research on the application of artificial intelligence in education has developed rapidly, but there are still few studies that focus on the role of AI in educational evaluation. This raises a number of questions that must be answered in this study. First, what are the trends in publications on the application of artificial intelligence in educational evaluation in the Scopus database over the past 5 years? Second, which journals, authors, and countries have contributed most to publications on this topic? Third, what are the patterns of collaboration between authors and institutions involved in this research? Fourth, which articles and authors are most frequently cited in this field? Fifth, what keywords frequently appear in publications on this topic? Sixth, what are the main themes of the research that emerge, and how are these themes related to each other in bibliometric analysis? These questions form the basis for formulating the focus of this study.

In line with these questions, this study has several main objectives. First, this study aims to describe the latest developments in publications related to the application of artificial intelligence in educational evaluation based on data found in Scopus. Second, this study seeks to identify the journals, authors, and countries with the most significant contributions in this field. Third, this study aims to examine patterns of collaboration between authors and institutions in related research. Fourth, this study aims to determine the most frequently cited articles and authors in this area. Fifth, this study aims to identify the most commonly used keywords in publications on this topic. Sixth, this study also aims to explore emerging research themes and how they are interrelated in bibliometric analysis. Thus, it is hoped that the results of this study can provide a comprehensive understanding of publication trends, scientific collaboration, major themes, and future research directions related to the application of artificial intelligence in educational evaluation.

Method

This study utilises a bibliometric design based on Scopus data, given its multidisciplinary coverage and extensive content curation, making it suitable for mapping AI literature in cross-disciplinary and cross-national education evaluation (Thelwall, 2022; Prancutè, 2021). The search strategy was developed in accordance with the PRISMA 2020 guidelines and the PRISMA-S extension for search reporting, ensuring that the process of identifying, screening, assessing eligibility, and including articles was documented and replicable (Page et al., 2021; Rethlefsen et al., 2021). Keywords were combined in the query TITLE ('Artificial Intelligence') AND TITLE ('Evaluation') OR TITLE ('Assessment') AND TITLE (Education*) OR TITLE-ABS-KEY (Learning*), with restrictions on the document type of indexed articles and social studies for the subject area. The search results were exported (metadata: author, affiliation, title, abstract, keywords, journal, year, citations, DOI), then deduplicated and cleaned (standardisation of author/affiliation names, keyword synonyms; e.g., 'AI', 'artificial intelligence'; 'assessment', 'evaluation') before analysis.

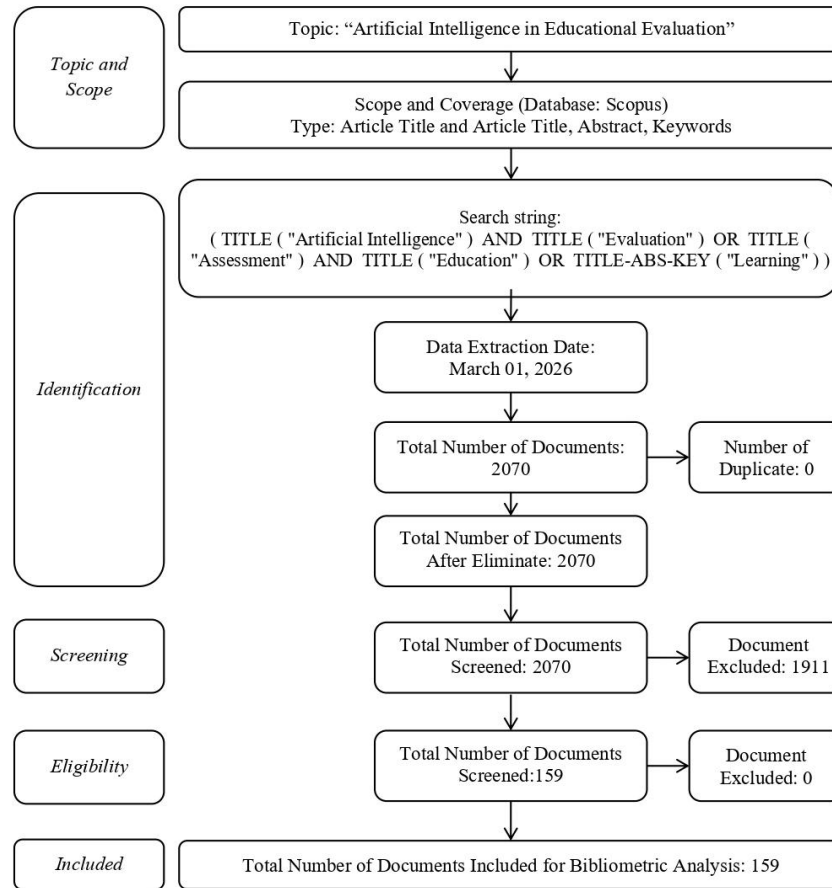


Figure 1. PRISMA flow diagram for bibliometric analysis

The analysis was conducted on three layers of indicators. First, performance indicators: annual publication growth, most productive journals, top authors and countries, as well as local and global citation impact to map the dynamics and contributions of key actors (Aria & Cuccurullo, 2017). Second, network analysis: (a) co-authorship (author/country/institution) to examine collaboration; (b) co-citation (author/source) to trace intellectual foundations; and (c) bibliographic coupling to identify thematic proximity between documents. Network mapping was conducted using VOSviewer with association strength normalisation and (fractional/full) counting in accordance with methodological recommendations, as well as a minimum frequency threshold (e.g., ≥ 5 keyword occurrences; ≥ 20 citations for co-citation) to maintain map readability (van Eck & Waltman, 2010). Third, thematic mapping (science mapping): co-word analysis of Author Keywords and Index Keywords was processed in Bibliometrix/Biblioshiny to produce a thematic map based on centrality (relevance) and density (internal development) metrics so that themes were grouped into motor, basic, niche, and emerging/declining; this procedure follows standard science mapping practices (Aria & Cuccurullo, 2017; Cobo et al., 2012). To assess the evolution of themes, the analysis was divided by period (e.g., pre-2020 vs. 2020–present), and thematic evolution was created to observe the shift from conventional AI to generative AI/LLM in the context of evaluation.

All quantitative analyses were performed in R with the bibliometrix package (including performance analysis, conceptual/intellectual/social structure) and the Biblioshiny interface, as well as in VOSviewer for large-scale network visualisation (Aria & Cuccurullo, 2017; van Eck & Waltman, 2010). Technical decisions (thresholds, normalisation methods, label/node sizes) were recorded for replication. Validation/robustness checks included: (i) varying frequency thresholds to ensure cluster stability; (ii) comparing Author co-word results with Index Keywords; and (iii) manual inspection of high-frequency keyword clusters. The limitations of the study are acknowledged, including database coverage bias (Scopus is not identical to the entire literature), the potential dominance of English-language publications, and keyword metadata variability; the implications are discussed in the discussion section (Thelwall, 2022; Prancutè, 2021). With this design, the study aims to comprehensively map the landscape, actors, and evolution of AI themes for educational evaluation, while identifying research gaps.

Results and Discussion

After collecting data from Scopus, researchers analyzed publications using a bibliometric approach related to the use of artificial intelligence in educational evaluation, using the Bibliometrix application available in the R program. The following is the main information obtained from the analysis:



Figure 2. Main information

The figure shows the results of a bibliometric analysis, a summary of key indicators from the topic of artificial intelligence in educational evaluation for the period 2021–2025. During this timeframe, a total of 159 documents were identified from 111 publication sources, with a very high annual growth rate of 81.42%. These findings indicate that the application of artificial intelligence in educational evaluation is a rapidly growing field that has garnered increasing attention within the global academic community in recent years. In terms of scientific collaboration, 915 authors contributed to these publications, with only 19 documents authored by a single author. This indicates that research in this field is highly collaborative, reinforced by an average of 5.95 authors per document. Additionally, the level of international collaboration reached 26.42%, reflecting the involvement of researchers across countries in developing studies related to artificial intelligence in educational evaluation. In terms of content characteristics, the authors used 570 keywords, reflecting the diversity of themes and approaches in the research, such as the use of machine learning, learning analytics, and AI-based evaluation systems. The total number of references used reached 21,990, indicating that research in this field is supported by a comprehensive literature base. The average document age of 2.04 years indicates that the analyzed publications are relatively recent, while the average citations per document of 32.95 suggest that works in this field have a fairly high scientific impact. Overall, these indicators demonstrate that the topic of applying artificial intelligence in educational evaluation is a dynamic, collaborative field of research with significant influence on the advancement of scientific knowledge.

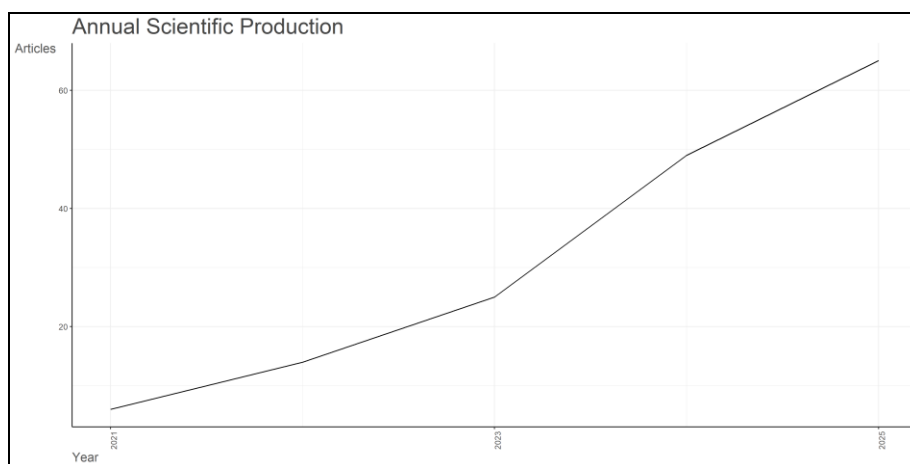


Figure 3. Annual scientific production

The figure shows the results of a bibliometric analysis of the annual scientific production in publications related to the topic of using artificial intelligence in educational evaluation. In 2021, the number of articles was still relatively low, at around 6 publications, which then gradually increased to approximately 14 articles in 2022 and 25 articles in 2023. A more significant surge is observed in 2024 with approximately 50 publications, peaking in

2025 with around 65 articles. This pattern indicates a strong acceleration in research interest over the past few years, particularly after 2023. This sharp increase can be attributed to the rapid development of artificial intelligence technologies, such as machine learning and large language models, as well as the growing need for more adaptive, automated, and data-driven educational evaluation systems. Overall, this trend confirms that the application of artificial intelligence in educational evaluation is an emerging and rapidly growing field with highly promising research prospects for the future.

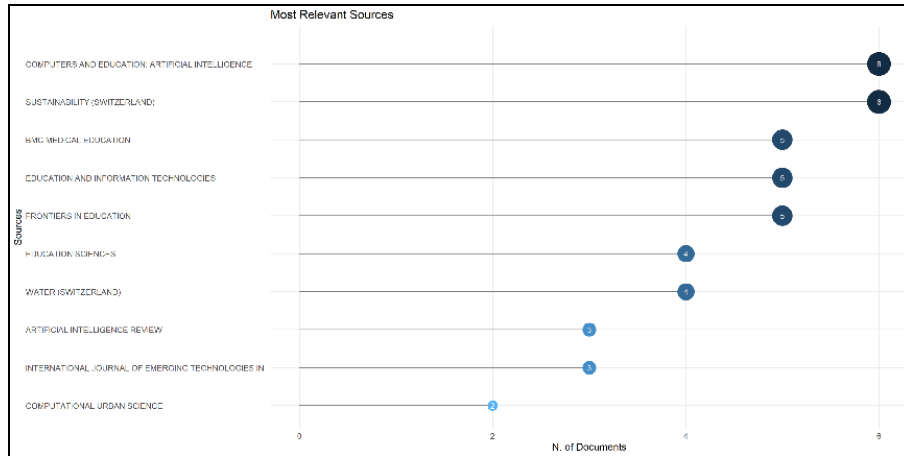


Figure 4. Most relevant source

Based on the figure above, the analysis results show that the journals Computers and Education: Artificial Intelligence and Sustainability (Switzerland) are the most dominant sources, each with 6 publications. This indicates that research on artificial intelligence in educational evaluation is not limited to educational technology but is also closely linked to sustainability and the development of adaptive educational systems. Furthermore, several other journals, such as BMC Medical Education, Education and Information Technologies, and Frontiers in Education, each contributed 5 publications, indicating significant contributions from various disciplines, including medical education and educational technology. Meanwhile, Education Sciences and Water (Switzerland) each have 4 publications, followed by Artificial Intelligence Review and the International Journal of Emerging Technologies in Learning with 3 publications each. Computational Urban Science ranks last with 2 publications. Overall, this distribution indicates that research on the application of artificial intelligence in educational evaluation is multidisciplinary and published across reputable journals focused on education, technology, and other applied fields. The dominance of certain journals also suggests the existence of key publication hubs (core sources) that serve as important references in the development of the literature in this field.

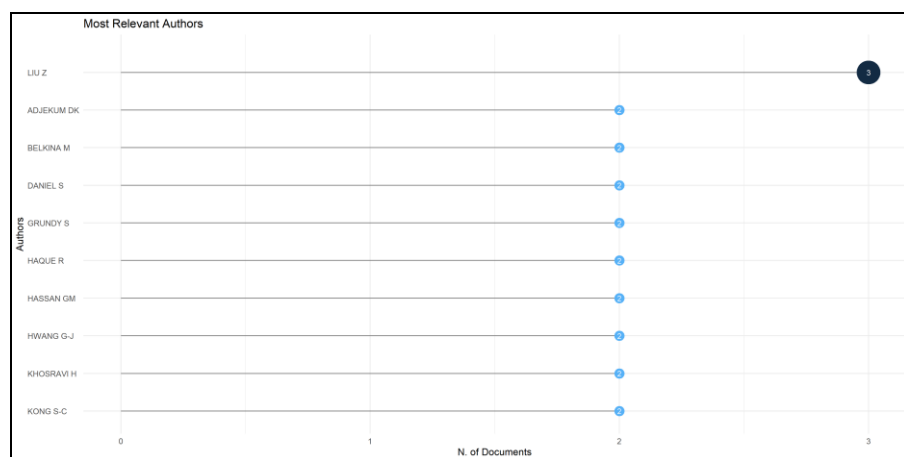


Figure 5. Most relevant authors

Based on the figure above, the analysis results show that LIU Z is the author with the highest contribution, with a total of 3 publications, and can therefore be considered one of the key contributors to the development of this topic. Meanwhile, several other authors, such as ADJEKUM DK, BELKINA M, DANIEL S, GRUNDY S, HAQUE R, HASSAN GM, HWANG G-J, KHOSRAVI H, and KONG S-C, each have 2 publications. This distribution indicates that there is no strong dominance by just one or two authors; rather, research contributions

are relatively evenly distributed among several researchers. This pattern suggests that the field of applying artificial intelligence in educational evaluation remains an emerging field, with broad participation from various researchers globally. Furthermore, the involvement of many authors with a relatively balanced number of publications also reflects the collaborative and dynamic nature of this field, where innovation and knowledge development are not centered on specific individuals but are distributed across a broader research network.

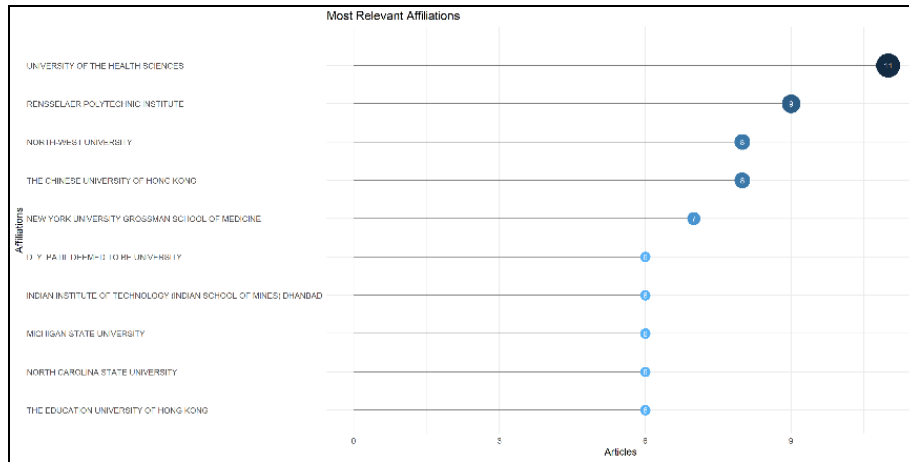


Figure 6. Most relevant affiliations

Based on the figure above, the analysis results show that the University of the Health Sciences ranks first with a total of 11 publications, highlighting the institution's dominant role in advancing research in this field. Next, Rensselaer Polytechnic Institute follows with 9 publications, followed by North-West University and The Chinese University of Hong Kong, each contributing 8 publications. Other institutions, such as New York University Grossman School of Medicine, also demonstrate significant contributions with 7 publications. Meanwhile, several institutions such as D. Y. Patil Deemed to be University, the Indian Institute of Technology (IIT) Dhanbad, Michigan State University, North Carolina State University, and the Education University of Hong Kong each have 6 publications. This distribution indicates that research contributions are not concentrated in a single geographic region but are spread globally, encompassing institutions from the United States, Asia, and other regions. Overall, these findings indicate that research on the application of artificial intelligence in educational evaluation is driven by strong cross-institutional collaboration involving various leading universities worldwide. The dominance of certain institutions also reflects the existence of centers of excellence that play a key role in driving innovation and scientific advancement in this field.

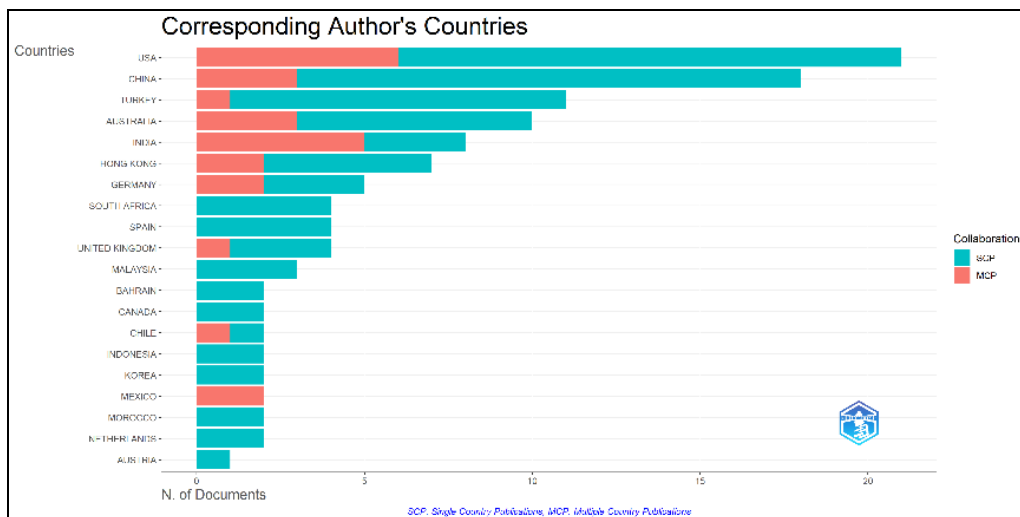


Figure 7. Corresponding author's countries

The figure shows the distribution of the countries of origin of corresponding authors in publications on the topic of the application of artificial intelligence in educational evaluation, distinguishing between Single Country Publications (SCP) and Multiple Country Publications (MCP). It is evident that the United States ranks first

with the highest number of publications, which are dominated by SCP but also include a significant proportion of MCP, indicating both high productivity and active engagement in international collaboration. China ranks second with a substantial contribution that tends to be dominated by SCP, indicating that research is conducted primarily domestically. Other countries such as Turkey, Australia, and India also exhibit similar patterns, with a dominance of national publications despite the presence of cross-border collaboration. Meanwhile, Hong Kong, Germany, and the United Kingdom show a relatively more balanced proportion of MCP, indicating more active involvement in international research networks. On the other hand, several countries, such as Spain, South Africa, Malaysia, Indonesia, Korea, and the Netherlands, have publications that are almost entirely SCP, indicating that research remains focused on the domestic sphere. Overall, these findings suggest that research on the application of artificial intelligence in educational evaluation is still dominated by major countries, with varying levels of international collaboration, thereby opening opportunities for enhanced global cooperation to strengthen the quality and impact of research in this field.

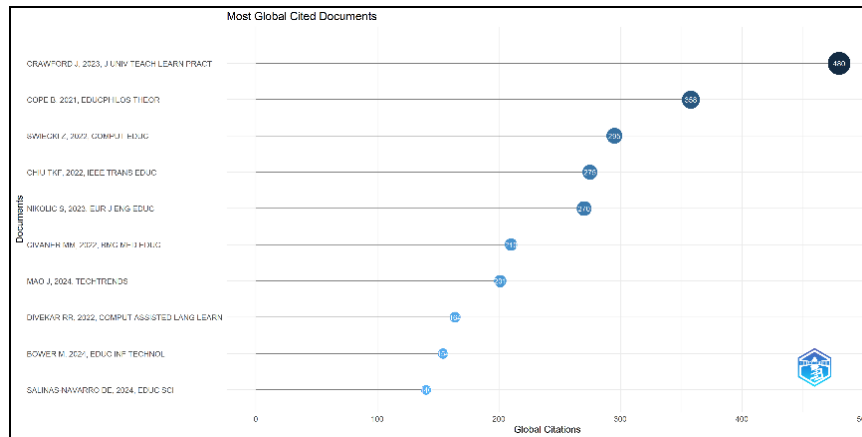


Figure 8. Most global cited documents

The figure shows the results of a bibliometric analysis of the most globally cited documents on the topic of artificial intelligence in educational evaluation. The article by Crawford (2023) ranks first with the highest number of citations, indicating a very strong influence in shaping the direction of research in this field. It is followed by Cope (2021) and Swiecki (2022), which also have significant citation counts, indicating that these works serve as primary references in the academic discourse regarding the application of AI in educational evaluation. Furthermore, several other articles, such as Chiu (2022), Nikolic (2023), and Civaner (2022) demonstrate significant contributions with relatively high citation counts, reflecting the relevance of the topics addressed to the development of this field. Interestingly, there are also relatively recent articles such as Mao (2024) and Bower (2024) that have already garnered a substantial number of citations, indicating accelerated growth and high interest in AI research within educational evaluation. Overall, this citation distribution indicates that the field is growing rapidly, with some key works serving as theoretical and empirical foundations, and demonstrates the evolving dynamics of research with significant contributions from recent publications.



Figure 9. Authors' local impact by H-index

The figure shows the results of a bibliometric analysis of authors' local impact based on the H-Index (Authors' Local Impact by H Index) in research related to artificial intelligence in educational evaluation. Liu Z ranks

highest with an H-index of 3, indicating that this author has made the most influential contributions to the collection of documents analyzed, with several publications that are consistently cited within the scope of this research dataset. Meanwhile, other authors such as Adjekum DK, Belkina M, Daniel S, Grundy S, Haque R, Hassan GM, Khosravi H, Kong S-C, and Kumar S have an H-index of 2, indicating that they also make a significant contribution, although their impact level remains below that of Liu Z. Overall, this relatively low and not highly varied distribution of H-index values suggests that the field of applying artificial intelligence in educational evaluation is still developing, with no strong dominance by a handful of authors, leaving ample opportunities for new contributions and research development in this field.

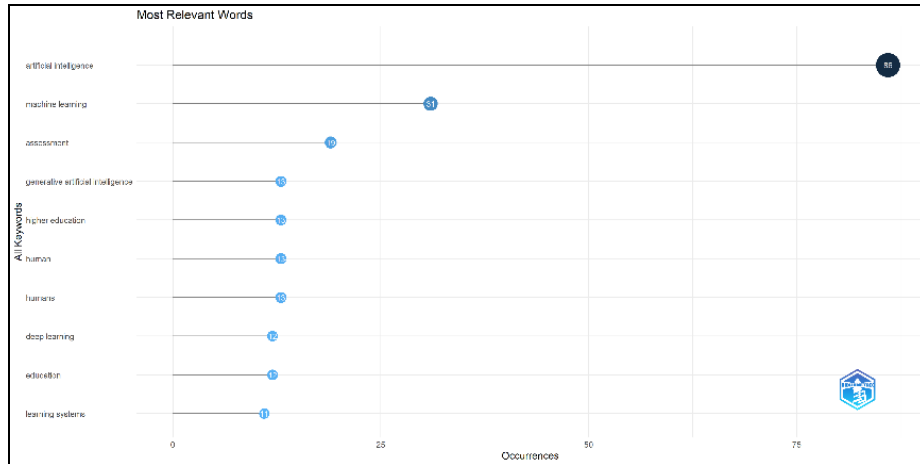


Figure 10. Most relevant words

The figure shows the results of a bibliometric analysis of the most relevant words in publications related to the topic of using artificial intelligence in educational evaluation. It is evident that the keyword “artificial intelligence” dominates significantly with the highest number of occurrences, confirming that the primary focus of the research is heavily centered on the use of AI in the context of educational evaluation. Furthermore, “machine learning” ranks second, indicating that machine learning-based approaches are the most widely used methods in the implementation of AI in this field. Other keywords, such as “assessment” also appear with fairly high frequency, indicating a research focus on the assessment or evaluation process in education. Additionally, the emergence of terms like “generative artificial intelligence,” “deep learning,” and “learning systems” reflects more specific and cutting-edge AI technological developments that are increasingly being adopted in research. Meanwhile, keywords such as higher education, education, and human/humans indicate that the context of AI application remains dominated by the higher education environment and continues to prioritize human aspects as the primary subjects in the evaluation process. Overall, this keyword distribution illustrates that research in this field focuses not only on AI technology itself but also on its integration into complex, human-centered educational evaluation systems.

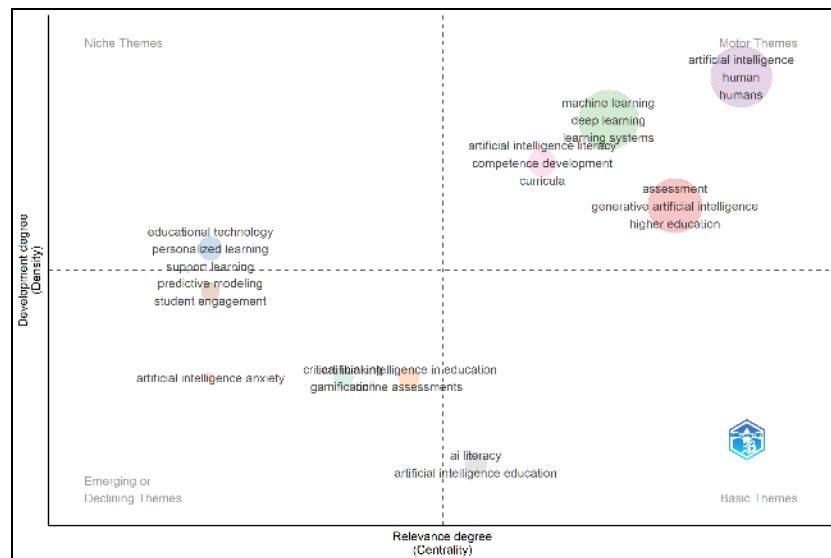


Figure 11. Thematic map

The figure is a thematic map illustrating the structure and development of research topics in the study titled “The Application of Artificial Intelligence in Educational Evaluation” based on two dimensions: centrality and density. In the “motor themes” quadrant (top right), topics such as artificial intelligence, human, humans, machine learning, deep learning, learning systems, as well as assessment and generative artificial intelligence are observed to have high centrality and density. This indicates that these themes constitute the core of mature research and are highly influential in this field. Furthermore, in the “basic themes” quadrant (bottom right), topics such as AI literacy and artificial intelligence education exhibit high centrality but low density, indicating that while these themes are important, they are still in the developmental stage. In the “niche themes” quadrant (top left), topics like educational technology, personalized learning, and predictive modeling emerge; these have high development levels but are less connected to core themes, making them specific and limited in scope. Meanwhile, in the emerging or declining themes quadrant (bottom left), there are topics such as artificial intelligence anxiety, gamification, and online assessments, indicating that these themes are still emerging or beginning to see a decline in research attention. Overall, this map shows that research on AI in educational evaluation is dominated by core technological and assessment themes, but is also expanding into AI literacy, pedagogical aspects, and emerging issues with the potential to develop in the future.

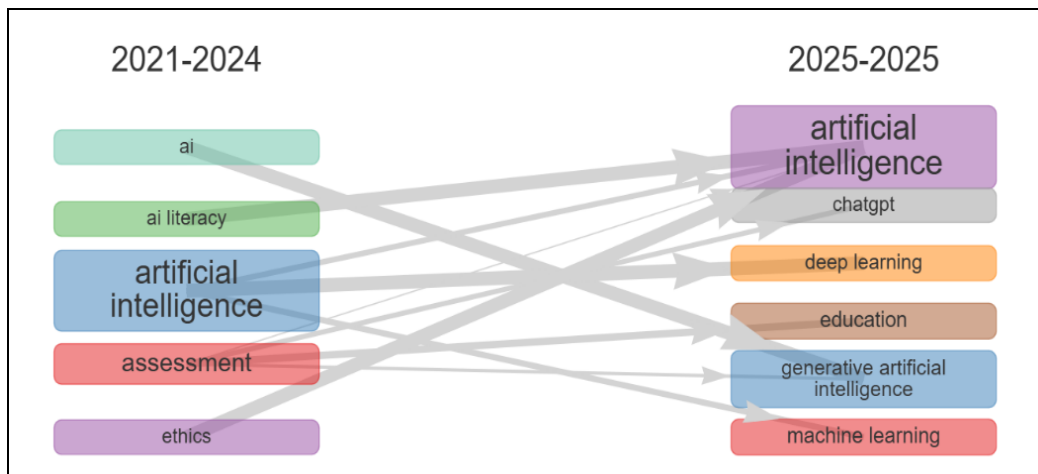


Figure 12. Evolution map

The figure represents a map of the evolution of research themes from 2021 to 2024, leading up to 2025, illustrating shifts and a deepening focus in the field of artificial intelligence (AI) within educational assessment. In the initial period (2021–2024), themes such as artificial intelligence, AI literacy, assessment, and ethics emerged as the primary foundations, with artificial intelligence serving as the central hub connecting various topics. Entering the 2025 period, there is an expansion and specialization of themes that are more applied and technical, marked by the emergence of generative artificial intelligence, machine learning, deep learning, and direct integration with educational contexts and tools such as ChatGPT. The arrows connecting the two periods indicate both continuity and transformation, where foundational AI concepts evolve into more complex and implementable approaches in educational assessment. Overall, this visualization indicates that research is moving from the conceptual and literacy stages toward the utilization of cutting-edge AI technologies focused on automation, personalization, and innovation in the learning evaluation process.

Conclusion

The results of a bibliometric study on the application of artificial intelligence in educational evaluation show that this topic is developing rapidly and has become one of the main areas of interest in academic literature worldwide. Analysis of publication trends shows a significant increase in the number of studies over the past few years, reflecting the importance and urgency of this issue in line with the acceleration of digital transformation in education. The most prominent focus is on artificial intelligence, machine learning, and deep learning, which serve as the main drivers of AI-based assessment technology development. In addition, terms such as assessment, ChatGPT, and generative artificial intelligence have emerged as new issues, indicating a shift in research towards the use of generative technology to improve the quality of evaluation, both in online and face-to-face learning. On the other hand, the mapping results show that despite the emphasis on technological innovation, there are still very few in-depth studies on the integration of AI with pedagogical approaches, formative assessment models, and learning theories. This opens up great opportunities for further research that focuses more on the integration of technology and educational strategies so that the use of AI truly

supports the improvement of learning quality, not just technical efficiency. In addition, research dissemination is still focused on developed countries, while publications from developing countries are still very limited. This indicates a gap in research and a need for context-based studies that can produce findings relevant to the educational challenges in developing countries.

Recommendations

Based on a bibliometric analysis of the application of artificial intelligence in educational evaluation, there are a number of important suggestions that can be used as a basis for future research. First, future studies should explore specific themes, such as the use of natural language processing in essay assessment, utilizing learning analytics to monitor student performance, and developing adaptive assessment systems tailored to the needs of each student. Second, the integration of artificial intelligence with teaching methods has not been sufficiently researched. Therefore, further research needs to focus on the relationship between AI and learning theories and pedagogical strategies, such as constructivism, project-based learning, and formative assessment, so that technology plays a role not only as a tool but also in strengthening the learning process. Third, publications listed in the Scopus database still mostly originate from developed countries, while studies in developing countries, including Indonesia, are still minimal. Therefore, it is important to conduct studies that are more relevant to the local context regarding the application of AI in educational evaluation so that the research results can contribute to the development of national education policy.

In addition, future research should encourage collaboration between different disciplines, including computer science, educational psychology, learning technology, and public policy, as well as strengthen international research networks to broaden perspectives. Ethical and privacy issues are also crucial topics that have not been widely researched, so future research should pay more attention to algorithmic bias, student data protection, and regulations on the use of AI in the education sector. Finally, future research should not only focus on theoretical studies, but also produce applicable AI-based evaluation models, such as prototypes of intelligent assessment systems that can be tested in schools or universities. In this way, the results of the research will not only be academic but also practical in supporting improvements in the quality of educational evaluation.

Scientific Ethics Declaration

* The authors declare that the scientific ethical and legal responsibility of this article published in EPES journal belongs to the authors.

* The author states that the study was conducted in accordance with the principles of scientific integrity and applicable research ethics standards. This study does not involve human subjects, animals, or confidential personal data, as all data used are secondary data legally obtained from the Scopus database and analyzed using bibliometric methods without any data manipulation or distortion. The author also ensures that all sources used have been properly cited and acknowledged in accordance with academic standards, and that there are no elements of plagiarism in this study. Furthermore, this manuscript is not currently under consideration or has not been published in any journal or other media in any form. The author declares no conflicts of interest that could influence the research results, and all contributing parties have met the criteria for authorship and approved the final version of the manuscript. The author is committed to maintaining the transparency, accuracy, and reproducibility of the research results and takes full responsibility for the content of this article.

Conflict of Interest

* The author declares that there are no conflicts of interest in the conduct and publication of the research. This research was conducted independently without influence from any party, whether financial, institutional, or personal, that could affect the results, analysis, or interpretation of the data. The entire research process, from data collection and bibliometric analysis to report preparation, was conducted objectively and in accordance with established scientific principles. Thus, the author affirms that the results of this study are free from any specific interests that could introduce bias.

Funding

* This study, titled “The Application of Artificial Intelligence in Educational Evaluation: A Bibliometric Analysis,” was supported by funding from the scholarship program of Lembaga Pengelola Dana Pendidikan (LPDP) under the Ministry of Finance of the Republic of Indonesia. This funding support was provided in the form of study funding that indirectly facilitated the conduct of this research, including access to academic resources such as the Scopus database, bibliometric analysis software, and other necessities that support the data collection, processing, and analysis process. Funding from LPDP also contributed to supporting the process of drafting the scientific manuscript through to the publication stage, both through the provision of academic resources and a conducive research environment.

Acknowledgements or Notes

* This article was presented as an oral presentation at the International Conference on Research in Education and Technology (www.icret.net) held in Konya/Türkiye on April 02-05, 2026.

* The author would like to express my deepest gratitude to my supervisor, colleagues, and all those who have provided support, guidance, input, and motivation in completing this research entitled “The Application of Artificial Intelligence in Educational Evaluation: A Bibliometric Analysis”. The author also expresses my appreciation to the Scopus database administrators for providing access to scientific information, which was the main source for this bibliometric study. The support in the form of credible data and literature was instrumental in the smooth running of the analysis. In particular, the author would like to express m deep gratitude and appreciation to Lembaga Pengelola Dana Pendidikan (LPDP) of the Ministry of Finance of the Republic of Indonesia for providing educational funding support. This assistance not only enabled the author to complete our studies but also provided us with the opportunity to develop this research as an academic contribution in the field of educational technology. It is hoped that the results of this research will benefit the development of science, particularly in the application of artificial intelligence to improve the quality of educational evaluation and will be a tangible contribution to education in Indonesia and internationally.

References

- Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2012). SciMAT: A new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology*, 63(8), 1609–1630. <https://doi.org/10.1002/asi.22688>
- Husain, N. A. S., Mohd Matore, M. E., & Mohd Hamzah, M. I. (2024). Performance evaluation for educators in higher education from bibliometric analysis views. *International Journal of Learning, Teaching and Educational Research*, 23(6), 124–142. <https://doi.org/10.26803/ijlter.23.6.7>
- Irwanto. (2025). Research trends on artificial intelligence in K-12 education in Asia: A bibliometric analysis using the Scopus database (1996–2025). *Discover Artificial Intelligence*, 5(1), 155. <https://doi.org/10.1007/s44163-025-00389-4>
- Nusantara, A. F., Sutrisno, S., & Puspitasari, D. (2023). Bibliometric analysis of artificial intelligent studies in education and pedagogy. *Nusantara: Jurnal Pendidikan Indonesia*, 9(3), 451–464. <https://doi.org/10.14421/njpi.2023.93.325>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(n71). <https://doi.org/10.1136/bmj.n71>
- Phạm, V. Q., & Hoàn, N. T. K. (2024). A bibliometric analysis of the use of artificial intelligence in teacher education based on the Scopus database. *Education Sciences*, 12(8), 569. <https://doi.org/10.18173/2354-1075.2024-0690>
- Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12. <https://doi.org/10.3390/publications9010012>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., & Koffel, J. B. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Journal of the Medical Library Association*, 109(2), 174–200. <https://doi.org/10.5195/jmla.2021.962>
- Thelwall, M. (2022). Scopus 1900–2020: Growth in articles, abstracts, countries, fields and journals. *Quantitative Science Studies*, 3(1), 37–50. https://doi.org/10.1162/qss_a_00155

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>

Author(s) Information

Wahyu Muhammad Ramdhani

Yogyakarta State University (Universitas Negeri
Yogyakarta)
Jln. Colombo No.1, Karangmalang, Sleman, Indonesia
ORCID iD: <https://orcid.org/0009-0000-3816-6100>

*Contact e-mail: wahyumuhammad.2024@student.uny.ac.id

Aman Aman

Yogyakarta State University (Universitas Negeri
Yogyakarta)
Jln. Colombo No.1, Karangmalang, Sleman, Indonesia
ORCID iD: <https://orcid.org/0000-0002-2395-446X>

Endang Wati

Yogyakarta State University (Universitas Negeri
Yogyakarta)
Jln. Colombo No.1, Karangmalang, Sleman, Indonesia
ORCID iD: <https://orcid.org/0009-0008-4264-7984>

*Corresponding author(s) contact email

To cite this article:

Ramdhani, W.M., Aman, A. & Wati, E. (2026). Application of artificial intelligence in educational evaluation: A bibliometric analysis. *The Eurasia Proceedings of Educational and Social Sciences (EPESS)*, 49, 87-98.