



The Global Landscape of Learning Analytics Research in Higher Education: A Bibliometric Analysis

Reni Kurnia^{1*}, Ambiyar², Fahmi Rizal³

¹ Department of Curriculum and Educational Technology, Faculty of Education, Universitas Negeri Padang, Padang, Indonesia

² Department of Electronics Engineering, Faculty of Engineering, Universitas Negeri Padang, Padang, Indonesia

³ Department of Civil Engineering, Faculty of Education, Universitas Negeri Padang, Padang, Indonesia

ARTICLE INFO

Article history:

Received: July, 2026

Received in revised from: July, 2026

Accepted: September, 2026

Available online: September, 28, 2026

Keywords: Learning analytics, Bibliometrics analysis, educational data mining, Self-regulated learning

ABSTRACT

The growing adoption of digital learning environments has intensified interest in learning analytics as a means of understanding student behaviour and improving educational decision-making in higher education. Despite the increasing volume of studies in this field, a clear picture of its global research development remains limited. This study examines the research landscape of learning analytics in higher education through a bibliometric analysis of publications indexed in the Scopus database from 2015 to 2025. Following a multi-stage screening process based on publication year, document type, language, and accessibility, 479 open-access journal articles were selected for analysis. Bibliometric mapping was conducted using Scopus metadata and VOSviewer to identify publication trends, influential contributors, collaboration networks, and thematic patterns. The results indicate a consistent rise in publication output over the past decade, reflecting expanding scholarly attention to data-driven learning practices. Research activity is largely concentrated in several productive countries and leading academic journals, with collaboration networks centred around a few dominant institutions. Keyword analysis reveals a shift from basic descriptive analytics toward predictive modelling, artificial intelligence, student engagement, and ethical issues related to data use. These findings offer a clearer understanding of the field's intellectual structure and may serve as a reference for future research and policy development in learning analytics.

1. Introduction

Higher education has undergone substantial digital transformation over the last decade. Learning management systems, virtual classrooms, online assessments, and educational platforms now generate vast amounts of learner data at a scale that was difficult to imagine only a few years ago. Every click, submission, discussion post, and interaction leaves a digital trace that can be captured and analysed. Universities no longer operate solely as knowledge-delivery institutions; increasingly, they function within data-rich ecosystems where teaching and learning processes are continuously mediated by digital technologies. This shift accelerated during and after the COVID-19 period, when online and hybrid learning became normalized across many countries (Bond et al.,

* Corresponding author.

E-mail address: renikurnia@fip.unp.ac.id

<https://doi.org/10.56806/jh.v7i3.455>

2021; Crawford et al., 2020). As digital infrastructures expanded, so did interest in extracting meaningful patterns from educational data to support instructional improvement, student success, and institutional planning. Within this context, Learning Analytics has gained considerable scholarly attention as a research domain concerned with understanding learning processes through data-driven approaches (Long & Siemens, 2011.; Viberg et al., 2018).

Learning analytics is commonly defined as the measurement, collection, analysis, and reporting of data about learners and their contexts for the purpose of understanding and optimizing learning and the environments in which it occurs (Long & Siemens, 2011). While the definition appears straightforward, the field itself is highly interdisciplinary, drawing from education, computer science, data science, psychology, and human-computer interaction. Its applications in higher education are wide-ranging. Institutions use analytics to monitor engagement, identify at-risk students, improve retention, personalize feedback, and support evidence-based academic decision-making (Alonso-Fernández et al., 2019; Ifenthaler & Yau, 2020; Lim et al., 2021). The appeal of learning analytics lies not merely in prediction but in intervention—using data to act before academic difficulties become irreversible. For universities facing growing pressure to improve completion rates and student outcomes, such capabilities are increasingly attractive. At the same time, the adoption of analytics introduces complex challenges, including algorithmic bias, data governance, privacy, and questions about how data should inform pedagogical decisions (Tsai et al., 2019). These tensions make learning analytics not only a technical domain but also a deeply educational and ethical one.

Research in learning analytics has diversified rapidly. Early studies focused largely on descriptive analytics, often examining learner behavior through dashboards or LMS logs. Over time, the field moved toward predictive modeling and increasingly sophisticated analytical techniques. Machine learning models are now widely used to predict student performance, dropout risk, and learning trajectories (Aldowah et al., 2019; Viberg et al., 2020). Another strong research stream examines student engagement, particularly behavioral, emotional, and cognitive engagement in online environments (Matcha et al., 2020). Early warning systems have also become more prevalent, enabling institutions to identify students needing academic support before failure occurs. More recently, the rise of generative AI and advanced predictive systems has further expanded the field, opening possibilities for adaptive learning and personalized educational pathways. Yet technological progress has also intensified debate around ethics. Questions surrounding transparency, fairness, consent, and surveillance have moved from the margins to the center of scholarly discussion (Alalawi et al., 2025; Holmes et al., 2022; Shibani et al., 2020). This growing complexity reflects the global expansion of learning analytics research, with contributions increasingly coming from North America, Europe, Asia, and Australia.

Given the rapid expansion of publications, bibliometric approaches have become useful for mapping the intellectual structure of this field. Several review studies have attempted to synthesize learning analytics research through systematic reviews and bibliometric analyses. Existing studies have examined publication growth, influential authors, keyword trends, and citation structures to understand how the field has evolved (Engrini et al., 2026; Leitner et al., 2017; Pardo & Teasley, 2014). Some reviews focused on methodological trends, particularly the increasing use of machine learning and predictive analytics, while others explored collaboration networks and thematic clusters. These studies have contributed to a broader understanding of research development, but they also reveal certain limitations. Many earlier bibliometric studies relied on datasets collected before the post-pandemic acceleration of digital learning. Several included learning analytics across all educational levels, making it difficult to isolate patterns specific to higher education. In some cases, analyses emphasized publication counts and citations while offering limited exploration of

thematic evolution over time. As a result, although existing reviews provide useful snapshots, they do not fully capture the most recent structural shifts within the field.

This limitation becomes more pronounced when considering the research landscape between 2023 and 2025. The emergence of generative AI, large language models, and increasingly sophisticated adaptive learning systems has reshaped how educational data can be interpreted and used. Research published before this period could not account for these developments. A second issue concerns geographical representation. Global collaboration in learning analytics is often discussed, yet detailed mapping of inter-country collaboration patterns remains underexplored, particularly in identifying dominant research hubs and peripheral contributors. A third gap relates to thematic evolution. Earlier studies frequently identified static keyword clusters but rarely examined how research themes have shifted from descriptive analytics toward predictive, adaptive, and AI-driven paradigms. Finally, despite higher education being one of the primary application domains of learning analytics, many bibliometric studies treat education as a broad category rather than investigating higher education as a distinct ecosystem with unique institutional, pedagogical, and governance characteristics. These gaps suggest the need for a more updated and focused bibliometric analysis that captures both structural and thematic changes in the field.

Against this background, the present study offers a focused bibliometric mapping of global research on learning analytics in higher education using Scopus-indexed publications from 2015 to 2025. Unlike earlier studies, this work concentrates specifically on higher education and incorporates recent literature produced during the rapid expansion of AI-driven educational technologies. The analysis maps publication growth, influential countries, leading journals, citation structures, collaboration networks, and keyword co-occurrence patterns to identify emerging research trajectories. By combining bibliometric indicators with network visualization, this study seeks to provide a clearer picture of the field's intellectual structure and evolving thematic directions. The findings are expected to contribute not only to scholarly understanding of learning analytics but also to future research agendas, institutional strategy, and policy discussions concerning the responsible use of educational data in higher education.

2. Methodology

This study employed a bibliometric research design to examine the global development of research on learning analytics in higher education. Bibliometric analysis was selected because it enables systematic mapping of scientific publications, allowing researchers to identify publication trends, citation patterns, collaboration networks, and thematic developments within a research field (Donthu et al., 2021; Zupic & Čater, 2015). To support knowledge mapping and visualization, this study combined bibliometric performance analysis with science mapping techniques using specialized bibliometric software (Aria & Cuccurullo, 2017). Bibliographic data were retrieved from the Scopus database due to its broad coverage of peer-reviewed journals and reliable metadata, making it suitable for bibliometric studies in interdisciplinary fields such as Learning Analytics (Baas et al., 2020). The search was conducted using the TITLE-ABS-KEY field with the query TITLE-ABS-KEY ("learning analytics" AND "higher education"). The keyword *learning analytics* represented the core research domain, while *higher education* was used to focus the search on universities and tertiary institutions.

A PRISMA-inspired screening procedure was adopted to ensure transparency and reproducibility in the article selection process. Although the PRISMA 2020 framework was originally developed for systematic reviews and meta-analyses, its structured stages—identification, screening, eligibility, and inclusion—provide a useful framework for documenting article selection in bibliometric studies (Page et al., 2021). The screening process is illustrated in Figure 1. The initial search identified 2,028 records.

After limiting publications to the period 2015–2025, the dataset decreased to 1,734 records. Only journal articles were retained, reducing the dataset to 736 records. After filtering for English-language publications, 714 records remained. The open-access filter further reduced the dataset to 479 records. a result, 479 articles were included in the final analysis.

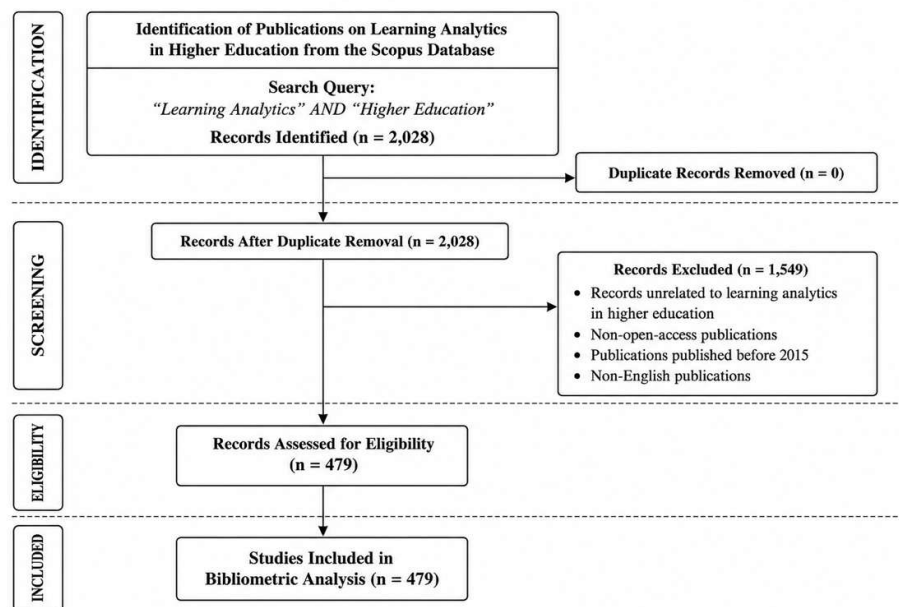


Fig. 1. PRISMA flow

The selected records were exported in CSV format, including bibliographic metadata, abstracts, author keywords, affiliations, citations, and references. The dataset was cleaned to reduce inconsistencies in author names, keywords, and institutional affiliations. Data analysis focused on annual publication trends, citation analysis, co-authorship patterns, country collaboration, source analysis, and keyword co-occurrence to identify research productivity, intellectual influence, and thematic developments within the field. Bibliometric mapping and visualization were conducted using Vosviewer (Van Eck & Waltman, 2010). The software was used to generate co-authorship, citation, and keyword co-occurrence networks. The full counting method was applied, with a minimum keyword occurrence threshold of 5. In the visualization maps, larger nodes indicate greater productivity or influence, while stronger links represent closer relationships among bibliometric units.

3. Results

3.1 Publication Characteristics and Growth Trends

The final dataset consisted of 479 journal articles related to learning analytics in higher education published between 2015 and 2025. These publications were retrieved from the Scopus database through a structured search strategy and screened using a PRISMA-inspired procedure to ensure consistency and relevance, as presented in Figure 1. Only peer-reviewed, English-language, open-access journal articles specifically related to learning analytics in higher education were included in the final dataset. The subject area distribution reflects the interdisciplinary nature of learning analytics research. The largest share of publications was classified under Social Sciences (38.7%), followed by Computer Science (27.0%). This distribution suggests that learning analytics is shaped by both pedagogical and computational perspectives. Additional contributions from Engineering (7.0%),

Psychology (5.6%), and Business, Management, and Accounting (3.0%) indicate that the field increasingly draws from multiple disciplinary traditions, including behavioral sciences, technical modeling, and institutional management. Smaller contributions from mathematics, decision sciences, health-related disciplines, and humanities further reinforce the broad intellectual reach of this research area.

Publication output showed a strong upward trajectory over the eleven-year period, although growth was not strictly linear. Three distinct development phases can be identified. The early stage (2015–2018) was marked by relatively limited publication activity, growing gradually from 6 articles in 2015 to 18 articles in 2018. During this phase, learning analytics was still consolidating its conceptual and methodological foundations within higher education research. A more pronounced expansion occurred during the growth stage (2019–2022). Publications increased from 32 articles in 2019 to 52 in 2020, indicating rapidly growing scholarly interest. This surge coincided with the large-scale digital transformation of higher education, particularly during the COVID-19 pandemic, when institutions worldwide adopted online and hybrid learning environments. The widespread use of learning management systems and digital platforms generated substantial educational data, creating favorable conditions for analytics-driven research. Although publication output declined slightly to 45 articles in 2021, the upward trend resumed in 2022 with 57 publications.

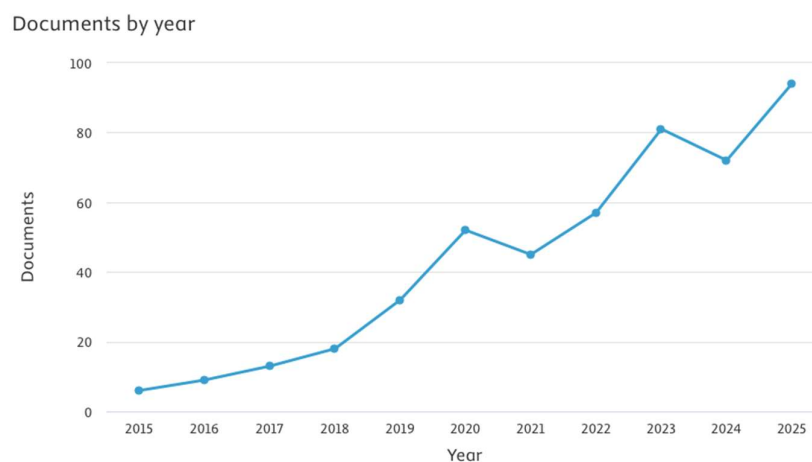


Fig. 2. Annual Distribution of Scopus-Indexed Publications on Learning Analytics in Higher Education (2015–2025)

The acceleration stage (2023–2025) displayed the most substantial expansion. Research output rose sharply to 81 publications in 2023, decreased modestly to 72 in 2024, and reached its highest level in 2025 with 94 publications. This rapid increase suggests that learning analytics has evolved from an emerging niche into a mature and fast-growing research domain. The acceleration observed after 2022 appears closely linked to the increasing integration of artificial intelligence, predictive modeling, and adaptive learning technologies in higher education.

3.2 Geographics and Source Distribution

Research productivity in learning analytics within higher education was concentrated in a relatively small number of countries, with strong dominance from English-speaking and technologically advanced nations. Australia emerged as the most productive country with 82 publications, followed closely by the United Kingdom (81 publications). The United States ranked third with 61 publications, while Spain (56 publications) also demonstrated strong research activity.

A second tier of contributors included Germany and the Netherlands, each with 31 publications, followed by Saudi Arabia (22), Finland (21), Chile (18), and China (17). The geographic distribution reveals clear concentration in Europe, North America, and Oceania, regions characterized by strong digital infrastructure and sustained investment in educational technology research. The prominence of Australia and the United Kingdom is particularly notable, reflecting their active institutional adoption of learning analytics and well-established academic research networks. At the same time, contributions from countries such as Saudi Arabia, Chile, and China suggest that learning analytics research is gradually expanding beyond traditional Western research hubs.

The distribution of publications across journal sources indicates that knowledge production is similarly concentrated in a relatively small group of specialized journals. The *Journal of Learning Analytics* emerged as the most dominant publication outlet, maintaining consistent output from 2019 onward and reaching a peak of 10 articles in 2024. Its central position reflects its role as a primary venue for advancing theoretical, methodological, and applied research in learning analytics. Other influential publication sources included the *International Journal of Educational Technology in Higher Education*, the *British Journal of Educational Technology*, *Education Sciences*, and *Sustainability*. The increasing contribution of *Education Sciences*, particularly in 2023 and 2025, suggests growing interest from broader educational research communities. Overall, journal concentration indicates that learning analytics research is largely disseminated through publication venues positioned at the intersection of education, digital learning, and computational analysis, creating a relatively cohesive scholarly communication ecosystem.

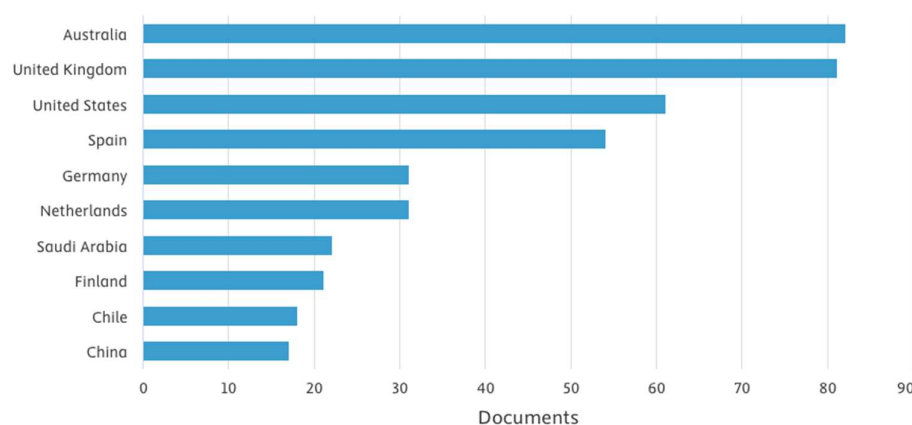


Fig. 3. Distribution of Publications by Countries in Learning Analytics Research in Higher Education (2015–2025)

3.3 Research Themes and Emerging Trends

Keyword co-occurrence analysis using VOSviewer revealed a dense conceptual network, indicating strong thematic interconnectedness across the literature. Among all keywords, machine learning emerged as the most dominant and central term, represented by the largest node in the network. Its centrality suggests that machine learning has become a major methodological foundation in contemporary learning analytics research. Strong relationships were observed between machine learning and other frequently occurring keywords such as students, predictive models, education computing, computer-aided instruction, and e-learning. This pattern reflects the growing reliance on computational techniques to analyze student behavior and improve data-informed educational decision-making.

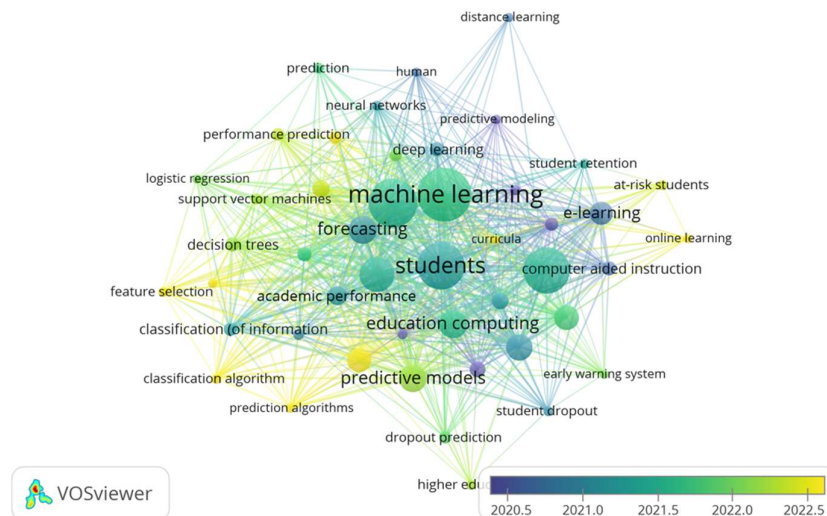


Fig. 4. VOSviewer Overlay Map Showing Temporal Keyword Evolution in Learning Analytics Research in Higher Education

Based on the figure above, four major thematic clusters were identified. The first cluster, Predictive Analytics and Performance Modeling, centered on keywords such as forecasting, prediction, performance prediction, and predictive modeling. This cluster indicates strong research emphasis on anticipating academic outcomes through data-driven models. The second cluster, Student Engagement, Retention, and Dropout Detection, included keywords such as student retention, student dropout, early warning system, and at-risk students. Research in this cluster focuses on identifying students at risk of academic failure and enabling timely interventions.

The third cluster, Machine Learning and AI Methods, contained advanced analytical techniques such as deep learning, neural networks, support vector machines, decision trees, logistic regression, and classification algorithms. The presence of these methods suggests a clear methodological shift from traditional descriptive analytics toward more sophisticated predictive and classification approaches. The fourth cluster, Digital Learning Environments, was associated with e-learning, distance learning, and computer-aided instruction, reflecting the technological environments in which educational data are generated and analysed.

Overlay visualization revealed clear temporal evolution in research focus. Earlier studies (2020–2021) were strongly associated with foundational themes such as e-learning, distance learning, and computer-aided instruction, reflecting early efforts to understand digital learning environments. More recent topics (2022–2023) shifted toward performance prediction, classification algorithms, feature selection, and at-risk students. This evolution suggests that the field is moving beyond descriptive monitoring toward increasingly predictive, intervention-oriented, and AI-driven analytics. Rather than merely analysing past learning behaviour, recent research increasingly aims to forecast academic risks and support personalized educational interventions.

4. Conclusions

This bibliometric analysis mapped the global development of research on learning analytics in higher education using 479 Scopus-indexed journal articles published between 2015 and 2025. The publication trajectory shows a clear expansion of the field, moving from a relatively modest research area into a rapidly growing domain of educational inquiry. The sharp rise in publications during the last few years reflects the accelerating digital transformation of higher education, particularly after

the widespread adoption of online learning environments and the growing integration of artificial intelligence in educational systems. At the same time, the disciplinary distribution confirms that learning analytics is inherently interdisciplinary, shaped primarily by the convergence of social sciences and computer science.

The geographic distribution reveals that scholarly production remains concentrated in a small number of developed countries, with Australia, the United Kingdom, and the United States contributing the largest share of publications. This concentration suggests that institutional readiness, digital infrastructure, and research investment continue to shape global participation in learning analytics research. Publication sources also show a relatively cohesive scholarly ecosystem, where a limited number of specialized journals function as central hubs for knowledge dissemination and academic exchange.

Based on a thematic perspective, the field has undergone a noticeable shift. Early research largely focused on digital learning environments such as e-learning and computer-assisted instruction. More recent work increasingly centers on predictive modeling, machine learning, and early identification of at-risk students. This transition suggests that learning analytics is no longer confined to descriptive monitoring of learner behavior. Instead, the field is moving toward predictive, intervention-oriented, and increasingly intelligent systems capable of supporting personalized learning and institutional decision-making.

Several limitations should be acknowledged. The analysis relied solely on the Scopus database and included only open-access journal articles published between 2015 and 2025, which may have excluded relevant studies indexed in other databases or available through restricted access. Future research could broaden the analytical scope by incorporating multiple databases such as Web of Science or Google Scholar. Further investigation is also needed to address emerging concerns surrounding AI-driven learning analytics, particularly issues related to explainability, fairness, privacy, and ethical data use. As learning analytics continues to evolve, future research should not only pursue greater predictive accuracy but also ensure that technological advancement remains aligned with pedagogical values and equitable educational practice.

References

- Alalawi, K., Athauda, R., Chiong, R., & Renner, I. (2025). Evaluating the student performance prediction and action framework through a learning analytics intervention study. *Education and Information Technologies*, 30(3), 2887–2916. <https://doi.org/10.1007/s10639-024-12923-5>
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- Alonso-Fernández, C., Pérez-Colado, I. J., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). *Using Simva to evaluate serious games and collect game learning analytics data*.
- 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>
- Bond, M., Bedenlier, S., Marín, V. I., & Händel, M. (2021). Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, 18(1), 50. <https://doi.org/10.1186/s41239-021-00282-x>
- Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P. A., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy

- responses. *Journal of Applied Learning and Teaching*, 3(1), 09–28. <https://doi.org/10.37074/jalt.2020.3.1.7>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Engrini, F., Asrizal, Mufit, F., & Novitra, F. (2026). Mapping Research Trends of Project Based Learning for Students' Digital Literacy: A Bibliometric Analysis. *Jurnal Penelitian Pendidikan IPA*, 12(5), 23–35. <https://doi.org/10.29303/jppipa.v12i5.14058>
- 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. (2022). 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>
- Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>
- Leitner, P., Khalil, M., & Ebner, M. (2017). Learning Analytics in Higher Education—A Literature Review. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundamentals, Applications, and Trends* (Vol. 94, pp. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-52977-6_1
- Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2021). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*, 72, 101202. <https://doi.org/10.1016/j.learninstruc.2019.04.003>
- Long, P., & Siemens, G. (2011). *Penetrating the Fog: Analytics in Learning and Education*.
- Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Pérez-Sanagustín, M., & Tsai, Y.-S. (2020). Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics*, 7(2), 45–71. <https://doi.org/10.18608/jla.2020.72.3>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Pardo, A., & Teasley, S. (2014). Learning Analytics Research, Theory and Practice: Widening the Discipline. *Journal of Learning Analytics*, 1(3), 4–6. <https://doi.org/10.18608/jla.2014.13.2>
- Shibani, A., Knight, S., & Buckingham Shum, S. (2020). Educator perspectives on learning analytics in classroom practice. *The Internet and Higher Education*, 46, 100730. <https://doi.org/10.1016/j.iheduc.2020.100730>
- Tsai, Y., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2019). Complexity leadership in learning analytics: Drivers, challenges and opportunities. *British Journal of Educational Technology*, 50(6), 2839–2854. <https://doi.org/10.1111/bjet.12846>
- 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>
- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89, 98–110. <https://doi.org/10.1016/j.chb.2018.07.027>

- Viberg, O., Wasson, B., & Kukulska-Hulme, A. (2020). Mobile-assisted language learning through learning analytics for self-regulated learning (MALLAS): A conceptual framework. *Australasian Journal of Educational Technology*, 36(6), 34–52. <https://doi.org/10.14742/ajet.6494>
- Zupic, I., & Čater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, 18(3), 429–472. <https://doi.org/10.1177/1094428114562629>