

Mapping the Evolution of Educational Evaluation and E-Learning Research in Indonesia Post-COVID-19: A Bibliometric Analysis (2020–2024)

Asma Khoirunnisa ^{a,1,*}, Heri Retnawati ^{a,b,2}, Risky Setiawan ^{a,3}, Samsul Hadi ^{c,4},
Auliasari Marjan ^{a,5}, Sitti Aminah ^{a,6}, Hanifah Mar'atush Shalihah ^{a,7}, Israel Amatepey ^{d,8},
Aisha Juma Mussa ^{e,9}, Alaa Abdelkader ^{f,10}

^a Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Yogyakarta, 55281, Indonesia

^b Department of Educational Mathematics, Universitas Negeri Yogyakarta, Yogyakarta, 55281, Indonesia

^c Department of Electrical Engineering, Universitas Negeri Yogyakarta, Yogyakarta, 55281, Indonesia

^d Department of Chemistry Education, University of Education Winneba, Ghana

^e Department of Monitoring and Evaluation, The Open University of Tanzania, Dar es Salaam, 23409, Tanzania

^f Department of Art Education, Helwan University, Cairo, 4034572, Egypt

¹ asmakhoirunnisa.2025@student.uny.ac.id; ² heri_retnawati@uny.ac.id; ³ riskysetiawan@uny.ac.id;

⁴ samsul_hd@uny.ac.id; ⁵ auliasarimarjanp@gmail.com; ⁶ inawahid03@gmail.com;

⁷ hanifahmaratush.2025@student.uny.ac.id; ⁸ amatepeyisrael2123@gmail.com;

⁹ ayshamussa22@gmail.com; ¹⁰ alaaeltlwany@gmail.com

* Corresponding Author

ARTICLE INFO

Article history

Received May 05, 2026

Revised June 02, 2026

Accepted July 01, 2026

Keywords

E-Learning;

Evaluation Model;

Learning Assessment;

Educational Technology;

Post-COVID Education

ABSTRACT

The rapid shift to online and technology-mediated learning during the COVID-19 pandemic significantly transformed Indonesia's educational practices and intensified attention to how learning processes and outcomes are evaluated. This study maps the evolution of educational evaluation and e-learning research in Indonesia after the pandemic by examining publication trends, thematic patterns, and intellectual structures from 2020 to 2024. A total of 133 publications published between 2020 and 2024 were retrieved from Scopus using a structured search query construction procedure and analyzed through Biblioshiny. The analysis included annual scientific production, average citation trends, cumulative growth modelling, keyword co-occurrence mapping, thematic mapping, and conceptual structure analysis. Data quality was strengthened through duplicate checking, keyword standardization, metadata consistency checks, and verification of Indonesian institutional affiliations. The findings reveal five major clusters: e-learning as a central theme connecting evaluation and digital pedagogy; learning assessment and student performance; educational technology and AI-supported learning tools; learning systems and platform-based instructional models; and emerging topics such as digital literacy, local wisdom, metaverse-related learning, and instrument development. These results indicate a post-pandemic shift toward technology-supported evaluation practices while also exposing gaps in validated digital assessment instruments, longitudinal evaluation, AI-integrated assessment frameworks, and equity-sensitive studies across diverse Indonesian educational contexts.

©2026 The Author. This is an open-access article under the CC-BY license.



1. Introduction

The emergence of the COVID-19 pandemic precipitated an unparalleled disruption, leading to an accelerated transformation within Indonesia's educational sector, necessitating an immediate transition towards online and hybrid instructional approaches. Educational institutions, ranging from universities to schools, confronted substantial hurdles, primarily due to insufficient preparedness for digital instruction and pervasive inequities in technological infrastructure and access among Indonesian students [1]. The implementation of emergency remote teaching, characterized as a provisional adaptation to distance education during a crisis, illuminated considerable deficiencies in educator readiness, technological capabilities, and impartial access to learning resources [2][3][4]. Although policy initiatives sought to ensure educational continuity via remote platforms, the practical implementation in Indonesia underscored varying levels of preparedness across different regions, particularly beyond Java and Sumatra, thereby exacerbating pre-existing educational disparities. This situation underscored the critical need for robust digital transformation initiatives in higher education to enhance competitiveness and mitigate the impact of such crises in the future [5].

In response to these unprecedented challenges, national policy frameworks critically envisioned an education system oriented toward unwavering resilience, equitable inclusivity, and dynamic digital adaptability. This forward-looking vision was concretized through initiatives such as Indonesia's "Merdeka Belajar" and reinforced by broader commitments via the UN Sustainable Development Goal 4, which collectively underscored the critical importance of not only quality of learning but also the urgent necessity for evaluation models to adapt effectively to emerging digital and hybrid modalities [6]. Consequently, policymakers explicitly articulated a pressing need for sophisticated evaluation models capable of seamlessly accommodating digital learning environments, effectively monitoring potential learning loss, and crucially providing real-time, actionable data for educators and administrators. Within this progressive ideal scenario, evaluation fundamentally transcends the limitations of traditional summative testing, evolving to encompass comprehensive process monitoring, iterative feedback loops, and advanced, technology-enabled analytics.

Nevertheless, the practical implementation of educational evaluation in Indonesia has lagged behind this vision in several key respects. First, while some institutions have adopted classical frameworks such as the CIPP model to assess Emergency Remote Teaching, the widespread adoption of adaptive, technology-based models remains limited [2][7][8][9][10]. Second, a scarcity of longitudinal studies tracking learning loss and recovery leaves policymakers with insufficient empirical evidence for targeted interventions [11][12]. Third, disparities in digital infrastructure, along with varying levels of teacher digital literacy and platform readiness, have constrained the effective implementation of robust evaluation systems [6][13][14]. Finally, specialized evaluation instruments tailored to the digital context, such as Learning Management System analytics, AI-driven assessment, and remote feedback mechanisms are still nascent and lack sufficient validation.

Although several studies have examined the implementation of online learning, teacher readiness, and the effectiveness of evaluation models during the pandemic, comprehensive scholarly research that systematically maps the development of educational evaluation in the Indonesian context remains scarce [15][16][17][18]. Global research indicates that bibliometrics mapping is an essential approach to understand research patterns, identify intellectual structures, and trace the evolution of scientific fields [19][20][21]. However, no study has applied a bibliometric approach to systematically evaluate publication trends, thematic clusters, and the evolution of research agendas in Indonesian educational evaluation after the pandemic until now. International studies also emphasize that the digital transformation of education requires more adaptive, data-driven, and technologically integrated evaluation models. Nevertheless, national literature has not provided a comprehensive picture of how the field of educational evaluation has responded to these challenges through scientific output. This gap constitutes a critical research gap that must be addressed to ensure that the development of policies and practices in educational evaluation is grounded in up-to-date and systematically mapped scientific evidence.

The present bibliometric analysis addresses this gap by mapping the development of educational evaluation and e-learning research in Indonesia during the post-COVID period. Using Scopus-indexed publications from 2020 to 2024 and Biblioshiny as the analytical platform, this study examines publication growth, citation patterns, keyword co-occurrence, thematic clusters, and conceptual structures without reporting empirical findings in the introductory section. The novelty of this study lies in its systematic mapping of Indonesian educational evaluation research after the pandemic through a structured search query construction procedure and bibliometric science mapping. Rather than reviewing individual studies descriptively, the article identifies the intellectual structure of the field, the dominant research themes, and the areas that require further methodological development. The contribution of this study is threefold. First, it provides an evidence-based overview of how educational evaluation research in Indonesia has responded to digital learning transformation. Second, it identifies dominant and emerging themes related to e-learning, learning assessment, educational technology, learning systems, and digital evaluation instruments. Third, it offers research and policy directions for strengthening technology-supported, context-sensitive, and equity-oriented educational evaluation in Indonesia.

2. Method

This investigation employs a bibliometric research design, grounded in quantitative statistical methods to examine patterns and structures within scholarly publication data. This approach systematically analyzes academic literature to identify trends, measure research output, and map intellectual landscapes. Bibliometric analyses enable researchers to investigate trajectories of scientific output, collaboration networks, thematic evolutions, and growth dynamics in established or emerging fields. Specifically, these methods allow for the quantification of research activity through metrics such as annual publication counts and cumulative growth curves, as well as the identification of key research areas and their interconnections through thematic clustering and network analysis [22].

In the current study, the focus is on mapping how educational evaluation models in Indonesia have been studied in the aftermath of COVID-19. By applying these quantitative techniques, this research aims to provide a comprehensive overview of the scholarly discourse, revealing prominent themes such as e-learning, learning assessment, and digital literacy, as well as identifying gaps like the relatively modest density of instrument-development research [23]. This methodological choice facilitates an objective assessment of the field's development, including its expansion, specialization, and the emergence of new research frontiers within the specific Indonesian educational context post-pandemic.

2.1. Data Source and Study Scope

The research encompasses publications from 2020 to 2024, a temporal span strategically chosen to capture the immediate aftermath of the COVID-19 pandemic, a period that instigated profound and considerable transformation within the Indonesian education sector. This specific timeframe allows for an in-depth examination of how educational evaluation models have evolved in response to the emergent challenges and adaptations during this critical phase, transitioning from crisis-driven responses to more structured evaluation paradigms, particularly those aligned with digital learning contexts [24]. The geographical scope of this inquiry is intrinsically linked to the national context of Indonesia, with publications meticulously selected based on author affiliations demonstrably situated within the country. This approach ensures that the analysis precisely reflects domestic scholarly contributions and perspectives on educational evaluation. While data acquisition leveraged global bibliographic databases, such as Scopus, thereby ensuring comprehensive coverage and rigorous data collection [22], the analytical emphasis was deliberately constrained to Indonesian scholarship. This focused approach aimed to provide a granular understanding of the unique trajectory of research and development in educational evaluation models within Indonesia's specific socio-educational landscape post-pandemic.

2.2. Inclusion Criteria

The inclusion criteria covered publications related to evaluation models, learning assessment, e-learning, digital learning, and associated thematic areas pertinent to the Indonesian educational landscape [8][9][23][25]. To realize this objective, bibliographic records that satisfied pre-defined keyword criteria and affiliation parameters were compiled and subsequently analyzed [21][26][27]. This analysis aimed to elucidate patterns in model application, research foci, thematic groupings, and the epistemological development within the field [28][29][30].

2.3. Unit of Analysis

The unit of analysis was the bibliographic record. Each record consisted of metadata such as title, author names, affiliations, year of publication, abstract, keywords, source title, and citation information. Because the study used publication metadata rather than human participants, no direct participant recruitment or primary survey procedure was involved.

2.3.1. Procedure

The procedure unfolds in five major steps, meticulously designed to ensure the rigor and reliability of the bibliometric analysis, as illustrated in Fig. 1.

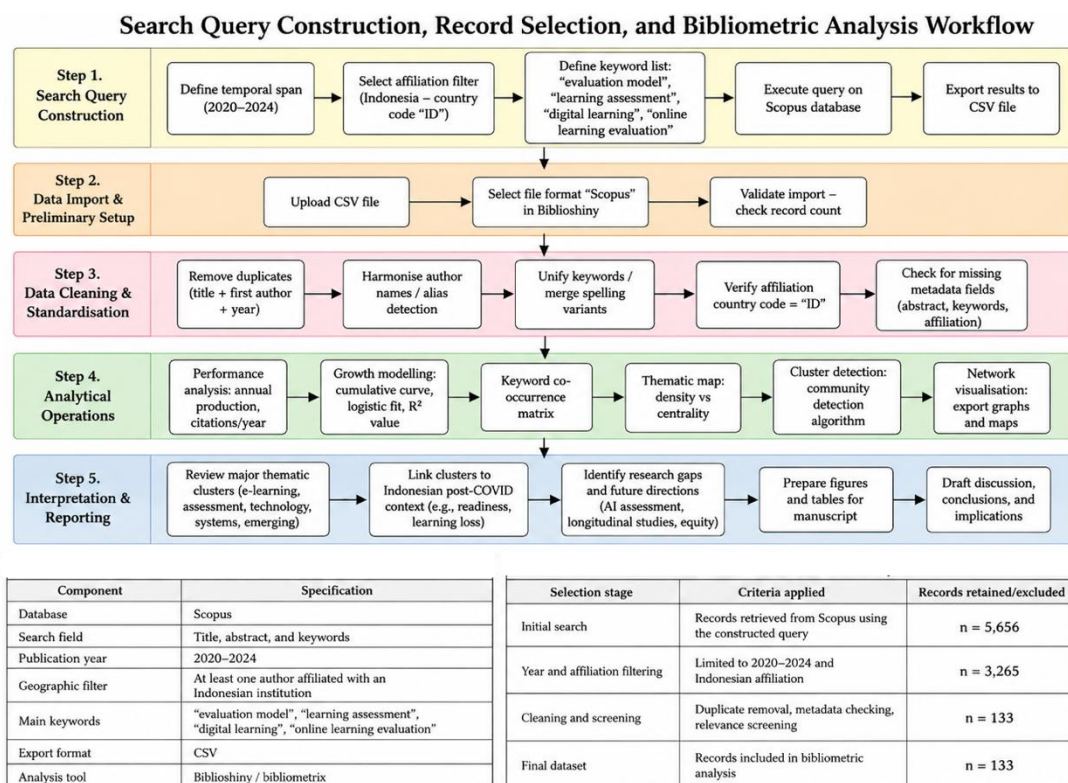


Fig. 1. Reserch Methodology

First, the initial phase involved the meticulous construction and execution of a precise search query within the Scopus database, a prominent global bibliographic platform known for its comprehensive coverage of academic literature [22]. This query was specifically designed to capture scholarly records published between 2020 and 2024, a temporal span strategically chosen to focus on the immediate aftermath of the COVID-19 pandemic and its transformative impact on Indonesian education [24]. The search string combined relevant keyword terms, including "evaluation model," "learning assessment," "digital learning," or "online learning evaluation," to identify pertinent research [21]. Crucially, an affiliation filter was applied to ensure that each retrieved record had at least one author demonstrably affiliated with an institution in Indonesia, thereby localizing the scope of the study to domestic scholarly contributions.

Second, following the successful execution of the defined search query, the retrieved bibliographic records were exported from Scopus in a Comma Separated Values file format. This raw dataset was then seamlessly uploaded into the Biblioshiny web interface, an integral component of the Bibliometrix R-package (Fig. 2). Biblioshiny served as the primary analytic instrument, providing an intuitive graphical user interface for subsequent data handling, analysis, and visualization, facilitating an efficient workflow for bibliometric research without requiring extensive programming knowledge.

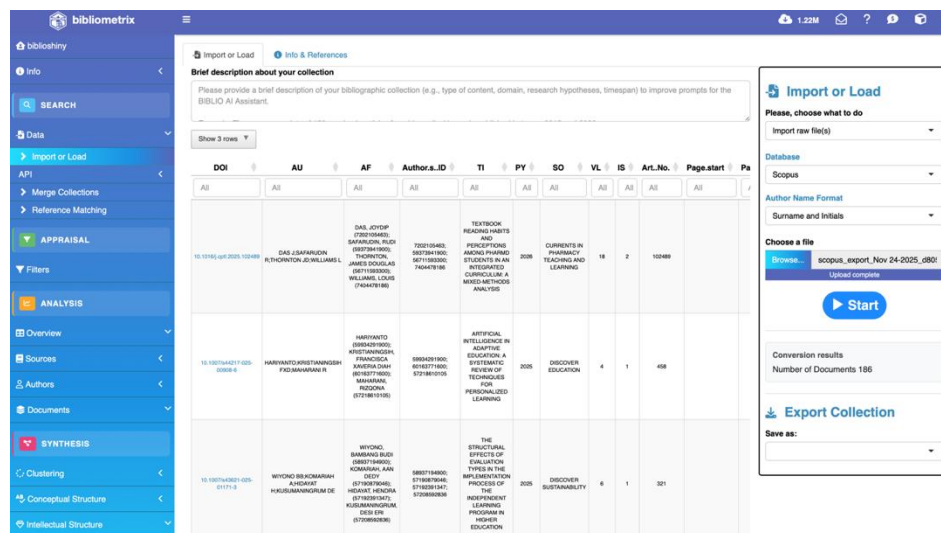


Fig. 2. Data Upload in the Biblioshiny Interface

Third, a crucial step in ensuring data integrity and consistency was the comprehensive data cleaning and standardisation process. This involved several sub-steps to refine the raw bibliographic metadata. Duplicate records were systematically identified and removed to prevent redundancy and skewed results [26]. Author name variants were carefully harmonised to consolidate publications by the same author under a single identity. Keywords, which are essential for thematic analysis, underwent a thorough standardization process, where author-supplied keywords and index keywords were unified, and variations representing the same concept (e.g., singular and plural forms) were merged to enhance analytical accuracy [26]. Additionally, affiliation country codes were rigorously verified to confirm the inclusion of "ID" for Indonesia, reaffirming the geographical focus of the study.

Fourth, a series of analytical operations were executed via Biblioshiny, leveraging its robust functionalities for bibliometric mapping and analysis. These operations encompassed performance analysis, which included quantifying annual scientific production and cumulative growth curves; growth modelling, often involving logistic regression to assess field maturity; science-mapping techniques, such as keyword co-occurrence matrix creation and thematic map generation; and sophisticated cluster detection using community-detection algorithms to identify prominent research areas and their interconnections [22][28][29][30].

Finally, the generated results from the analytical operations were meticulously interpreted. This interpretative phase involved drawing meaningful connections between the identified thematic clusters, the prevailing models and approaches to educational evaluation, and the observed patterns of publication growth within the specific context of Indonesia's post-COVID educational landscape. This comprehensive interpretation aimed to provide a nuanced understanding of the field's development, emergent trends, and areas ripe for future investigation.

2.4. Analytical Instruments

While direct human participant interaction was not a component of this research, the Biblioshiny application functioned as the principal analytical instrument [31][32]. This specialized software facilitated the importation, systematic cleaning, in-depth analysis, and comprehensive visualization of bibliographic data, thereby eliminating the necessity for extensive programming knowledge

[33][34][35]. The established reliability of Biblioshiny as a robust bibliometric tool has been consistently acknowledged within contemporary methodological discourse [36]. Furthermore, rigorous manual verification and standardization of the data were undertaken as crucial supplementary validation procedures, aimed at minimizing measurement inaccuracies and enhancing the consistency and validity of the thematic clusters identified [37][38][39].

2.5. Data Collection & Metadata Processing

Bibliographic data collection was performed through a structured search query construction procedure in Scopus. The final dataset consisted of 133 publications published between 2020 and 2024. Exported records included all available metadata fields. Upon upload to Biblioshiny, data selection criteria were applied (year ≥ 2020 , affiliation country code “ID”, document type journal article/proceedings). Metadata fields of interest included authors, title, abstract, keywords, affiliations, publication year, and times cited. Additional manual cleaning steps addressed incomplete fields and inconsistent keyword formatting.

2.6. Data Analysis Techniques

Two tiers of analysis were conducted. The first tier, descriptive performance analysis, quantified annual scientific production, average citations per year, and cumulative publication counts. A logistic regression model was fit to the cumulative curve to assess whether the field is approaching maturity, with R^2 reported as a measure of fit. The second tier, science-mapping analysis, used Biblioshiny’s functions to perform keyword co-occurrence matrix creation, network visualisation, thematic map generation, and cluster detection via community-detection algorithms. Interpretations drew on network metrics and thematic coherence to identify research fronts and gaps. Fig. 3 visually presents this comprehensive analytical framework.

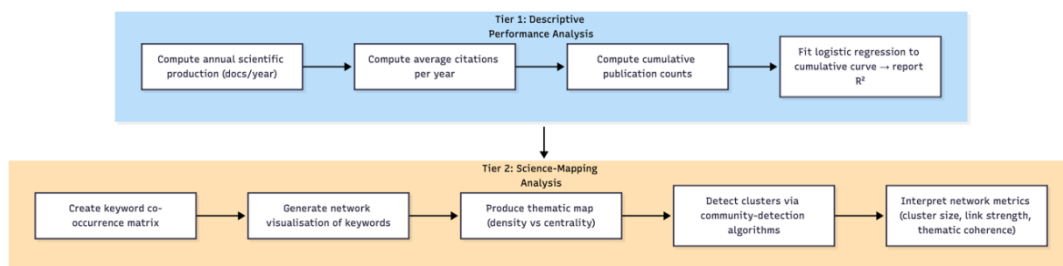


Fig. 3. Analytical Framework

2.7. Ethical Considerations

Because this study relies entirely on publicly available bibliographic data, it does not involve direct interaction with human participants and therefore does not require formal ethical approval. Nevertheless, the authors guarantee that data are used solely for research purposes and that all metadata are handled in accordance with database terms of service. Furthermore, the authors commit to anonymising any author-level information when reporting (e.g., using aggregated counts rather than individual identifiers) to protect academic privacy and avoid misinterpretation of performance metrics.

3. Results and Discussion

3.1. Annual Scientific Production

The bibliometric investigation into Indonesian scholarly research on educational evaluation models in the post-COVID-19 era offers multifaceted insights into both the trajectory of scholarly output and the thematic architecture of the field. These insights are derived from a comprehensive analysis encompassing performance metrics, growth modelling, science-mapping techniques, and sophisticated cluster detection, which collectively illuminate the evolution and structure of research in this domain. As depicted in Fig. 4, the volume of publications has undergone a marked acceleration.

Prior to 2020, annual counts were minimal; however, commencing in 2020, the trend line ascends steeply, reaching approximately fifty-plus articles by 2024. This surge underscores a concerted scholarly response to the profound educational disruptions precipitated by the global COVID-19 pandemic. These disruptions included widespread school closures, hindered academic progress, rescheduling of examinations, and the disruption of co-curricular activities and social skill development [40]. Consequently, there was a rapid and often forced pivot to e-learning and hybrid instructional formats, necessitating an urgent re-evaluation of established educational practices [40]. The transition exposed significant challenges, such as digital inequality, lack of preparedness among educators and learners for virtual platforms, and concerns regarding data privacy and the efficacy of remote instruction. It signifies that Indonesian researcher rapidly mobilised to examine models of teaching, assessment, and evaluation within this transformed educational environment, responding to the critical need for understanding and adapting to the emergent pedagogical landscape.

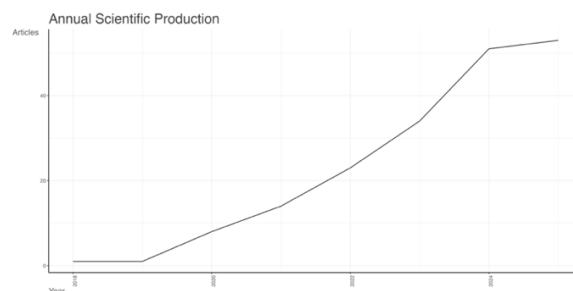


Fig. 4. Annual Scientific Production (2020 – 2024)

3.2. Average Citations per Year

A closer examination of this section provides crucial impact data, complementing the insights from scientific output analysis (see Fig. 5). While the average citations per article exhibit a modest peak, hovering around 4.5 to 5 citations per year, a noticeable disparity exists with earlier years showing significantly lower values. This pattern suggests that despite a recent surge in publication quantity within the Indonesian educational evaluation landscape, the field's maturity in terms of broader influence, cross-citation, and theoretical consolidation is still in its developmental stages.

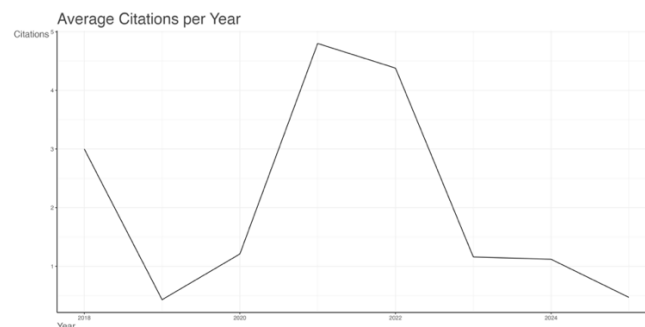


Fig. 5. Average Citations per Year

Several factors could contribute to this observed trend. The early phase of a research field often sees fewer citations as it builds its foundational literature. As the field evolves, foundational texts tend to attract higher impact. The slight decline in average citations after the peak observed in the figure might reflect a saturation effect, where a rapid influx of newer publications requires more time to accumulate citations. Additionally, non-scientific factors such as article length, author collaboration networks, journal impact factors, and even online accessibility can influence citation counts [41]. Disciplinary differences in citation practices also play a role, with some fields inherently having lower citation rates than others [42]. Therefore, while citation counts are a common measure of academic impact, they can be influenced by various parameters beyond the intrinsic quality of individual papers [43]. The relatively modest citation impact further underscores the need for deeper engagement,

perhaps through interdisciplinary collaborations, more rigorous methodological approaches, and broader dissemination strategies to enhance the visibility and influence of Indonesian research in educational evaluation.

3.3. Cumulative Growth Curve

The cumulative perspective on the field's scholarly output is vividly presented in Fig. 6(a), where the observed cumulative number of publications is juxtaposed with a fitted logistic regression model. This widely adopted “S-shaped” curve is a characteristic representation of growth in many scientific and technological domains, illustrating an initial slow period, followed by rapid, exponential expansion, and culminating in a saturation phase. The curve distinctly illustrates that, by 2024, the cumulative number of publications in Indonesian educational evaluation research approaches a level where modelled saturation is projected.

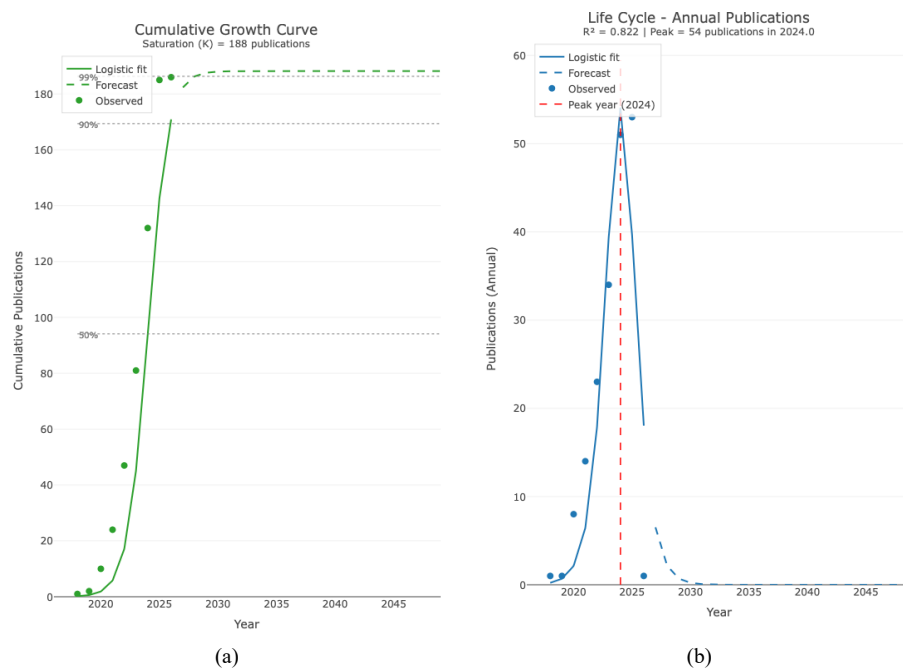


Fig. 6. Cumulative Growth Curve

The predictive capability of this logistic model is supported by an R^2 value of approximately 0.822, which indicates a strong fit between the model and the observed data. This robust fit suggests that the field may indeed be entering a plateau phase, a stage often characterized by a deceleration in the rate of new publications after a period of rapid growth. Consequently, this implies a potential shift in the field's developmental trajectory: future growth will likely move away from a simple increase in publication quantity towards greater diversification and depth. In essence, the “explosive growth” phase, marked by a rapid surge in scholarly contributions, is now transitioning into a more stable and mature stage of academic activity. This maturation may entail a focus on refining existing knowledge, exploring interdisciplinary connections, or deepening theoretical frameworks, rather than merely expanding the volume of published works. This transition can offer opportunities for innovation and specialization within the established framework of the field.

Complementing other temporal metrics, Fig. 6(b) offers a categorization of the publication trend into established life-cycle phases: emerging, growth, maturity, and decline. This graphical representation specifically indicates that the peak publication year is 2024, with 54 articles. This observed peak aligns precisely with the maximum point projected by the logistic growth model, which is frequently employed in bibliometrics to describe the evolution and eventual saturation of scientific fields, mirroring natural growth processes.

The depiction within Fig. 6(b) strongly suggests that the field of Indonesian educational evaluation research has now reached a stage of maturity. This implies an impending transition into phases characterized more by refinement or consolidation rather than rapid quantitative expansion. As fields approach saturation, the focus often shifts from increasing the sheer volume of output to enhancing depth and specialization. Consequently, for Indonesian educational evaluation research, this mature phase indicates that while the rate of generative growth in publications may naturally slow, new and significant opportunities for innovation and specialization within existing research areas will likely emerge and become more prominent. This shift from an "explosive growth" phase to a more stable, mature stage offers a chance to deepen scholarly contributions.

3.4. Thematic map of Keyword Clusters

Turning to the thematic architecture of the field, Fig. 7 provides insights into how key topics are distributed based on their centrality and density within the research landscape. The map distinctly highlights clusters positioned within the "motor themes" quadrant, characterized by both high centrality and high density. A prime example of such a motor theme is "students-learning assessment," which demonstrates a well-developed and central role in the research field. This indicates its foundational importance and established research activity. Other clusters, such as "e-learning/online learning evaluation," also appear strongly, reflecting significant engagement.

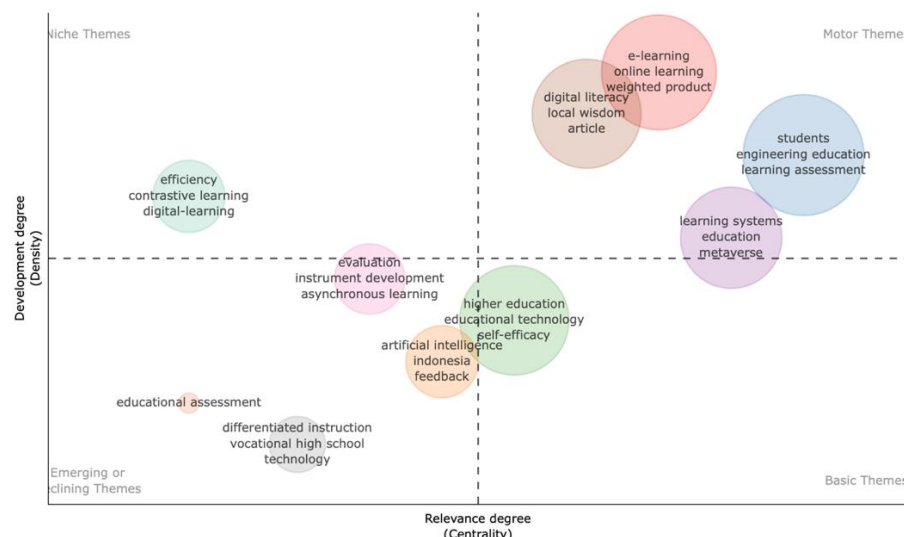


Fig. 7. Thematic Map of Keyword Clusters

In contrast, other themes, including "instrument development" and "digital literacy/local wisdom," are located in less developed or emerging quadrants. This mapping is crucial as it not only pinpoints where research efforts are currently concentrated but also identifies potential areas where future scholarly work might be directed, highlighting gaps or nascent fields of inquiry.

3.5. Keyword Network Map

A more granular understanding of keyword interrelationships is presented in Fig. 8, which visually illustrates the sizes of nodes and the linkages between them. In this visualization, the size of each node typically represents the frequency of occurrence for that particular theme or keyword, reflecting its prominence within the dataset. The linkages, or edges, between these nodes indicate thematic overlap and the integrated nature of the concepts, with denser connections signifying stronger interrelationships. This detailed co-occurrence network reveals several highly central nodes, notably "e-learning," "learning systems," "students," and "educational technology."

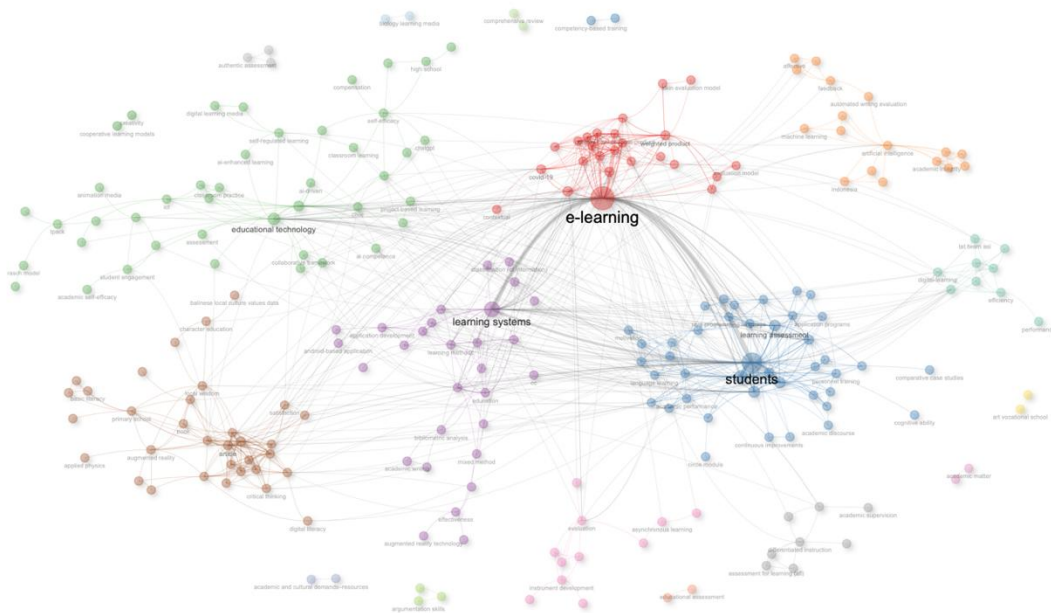


Fig. 8. Keyword Network Map

The dense interconnections observed among these central nodes underscore the inherent thematic overlap and integrated nature of the field. Nodes with high centrality in such networks often act as crucial connectors, facilitating information flow and integrating diverse ideas across different thematic clusters. This is particularly evident in how e-learning research now converges with and influences various evaluation models, student outcomes, learning systems, and technology adoption strategies, forming a cohesive research front. The strong presence and interconnectedness of these themes highlight their foundational importance and established research activity within Indonesian educational evaluation.

Furthermore, the map identifies smaller clusters situated on the periphery of the main network, such as "augmented reality," "metaverse," and "local wisdom." These less integrated clusters signal nascent or emergent trajectories within the research landscape, suggesting areas that are currently undergoing initial exploration and hold significant potential for future development and expanded scholarly engagement. For instance, the inclusion of "augmented reality" and "metaverse" indicates a growing interest in advanced technological applications within education, aligning with the evolution of e-learning towards more interactive 3D environments incorporating artificial intelligence and virtual reality [40]. The presence of "local wisdom" as a peripheral cluster suggests an emerging focus on culturally relevant or indigenous approaches to educational evaluation, potentially offering unique contributions to the field. This mapping is crucial as it not only pinpoints where research efforts are currently concentrated but also identifies potential areas where future scholarly work might be directed, highlighting gaps or nascent fields of inquiry.

The implications of these findings for Indonesia's educational landscape are substantial. First, the accelerated publication growth reflects a responsive scholarly community but also underscores urgency: the pandemic created immediate evaluative needs (e.g., measurement of learning loss, adaptation of evaluation frameworks, deployment of digital assessment tools). The fact that the students-learning assessment cluster is central indicates that researchers are aligning with national priorities around student outcomes and accountability. Yet, the moderate citation impact and modelled saturation suggest that while volume has grown, the depth of citations, collaboration, and theoretical sophistication is still in development.

Second, the presence of clusters in "emerging" quadrants, such as digital-literacy, AI-enabled evaluation, metaverse in education signals future frontiers. For Indonesian research, this translates into an imperative: to move beyond descriptive studies and into advanced methodological and

technological development, longitudinal designs, context-adapted evaluation models (e.g., for remote areas or culturally diverse regions), and evidence of real-world impact on policy and practice. The relative under-development of these themes (according to thematic map density) suggests research gaps worthy of targeted investment.

Third, from a policy and practice standpoint, these findings suggest that Indonesian educational authorities should support not only quantitative growth (i.e., more papers) but also structural quality and integration. For instance, investing in national dashboards for e-learning evaluation, developing validated instruments for digital learning assessment, and forging international collaboration to boost citation and impact profiles. Researchers would benefit from interdisciplinary approaches linking educational evaluation with data science, AI, and regional equity studies.

Finally, the logistic growth modelling and publication life-cycle analysis caution that the field may soon transition from growth to consolidation. This means that future research may yield diminishing returns if it simply replicates established themes. Instead, innovation will likely come from niche and emerging themes. This maturation phase offers an opportunity to deepen, refine, and broaden the field: better validated instruments, mixed-method longitudinal studies, impact evaluations of interventions, and greater attention to under-researched regions of Indonesia. In conclusion, this study illustrates a field that is dynamically growing, thematically coherent, and entering a stage of maturation. For Indonesian educational evaluation research, the challenge is no longer merely "produce more" but rather "produce better, more impactful, more contextually relevant" work. The scholarly community and policy stakeholders alike should thus turn attention toward depth, integration, and strategic focus in upcoming years.

Indonesian educational evaluation research exhibits significant growth, yet bibliometric analysis consistently shows a pronounced reliance on the CIPP evaluation model, mirroring previous national assessments. This sustained preference highlights researchers' continued use of a familiar, decision-oriented framework that systematically examines the context, input, process, and product of programs [44]. Nevertheless, thematic mapping and network analysis reveal underexplored areas and indicate a maturing field, signaling the need for a strategic redirection. While the CIPP model retains its foundational value for program accountability and decision-making [7][44][45][46], it often falls short in providing the depth of contextual information and understanding "how and why" an intervention works [47]. Future inquiry should therefore broaden its methodological scope beyond CIPP's general applicability. To enhance impact and relevance, scholars ought to investigate alternative and complementary evaluation frameworks. For instance, realist evaluation can provide deeper insights into causal mechanisms and the interplay of context and outcomes, addressing the "how and why" questions often missed by traditional models [47]. Similarly, developmental evaluation and logic-model approaches are crucial for explicating implementation pathways and making meaningful allowances for the complexity and context of interventions [48] (Coldwell & Maxwell, 2018). Furthermore, given Indonesia's diverse cultural landscape, adopting culturally responsive evaluation methods, which are rooted in indigenous values and knowledge systems, can ensure evaluations are more contextually grounded and relevant to culturally distinct educational environments [49]. These diverse models could more effectively address contemporary challenges like digital transformation, emergent technologies, and culturally distinct educational environments by offering increased theoretical depth and analytical sophistication into Indonesian scholarship [7][45][46]. Such diversification would align with a new phase of research focused on producing deeper, methodologically robust, and contextually grounded evaluations, rather than merely increasing publication volume.

4. Conclusion

This study concludes that educational evaluation research in Indonesia has undergone a substantial paradigm shift in the post-COVID era, transitioning from traditional approaches to more adaptive, technology-driven, and data-oriented evaluation models. The bibliometrics analysis revealed a sharp rise in publication output from 2020 to 2024, indicating a strong scholarly response to the

challenges posed by digital learning transitions. The thematic mapping identified five major clusters: e-learning and digital pedagogy, learning assessment and student performance, educational technology, platform-based learning systems, and emerging themes such as digital literacy, metaverse, and instrument development. These findings suggest that e-learning and assessment serve as motor themes driving the research landscape, whereas topics such as digital literacy and instrument development remain underexplored, highlighting research gaps and opportunities for future studies.

Despite this growth and thematic diversification, an important limitation emerges from the bibliometric evidence: the overwhelming reliance on a single evaluation framework, namely the CIPP evaluation model (Context–Input–Process–Product), underscores a lack of diversity in evaluation model types employed. While CIPP remains a robust and comprehensive model widely used in Indonesian research (assessing context, inputs, processes, and products), the minimal use of alternative or complementary models (such as outcome-based, responsive, goal-free, or theory-driven frameworks) may constrain the capacity of research to address complex, emergent evaluation needs in digital and hybrid learning environments. As the field matures and opportunities for thematic expansion grow, future studies should deliberately diversify methodological and theoretical frameworks, integrating different evaluation models, including contextually adapted or mixed-method approaches to ensure educational evaluation research in Indonesia remains relevant, rigorous, and responsive to evolving educational realities.

The cumulative growth trend demonstrates that the field is entering a maturity phase, implying that the future trajectory of research should shift from quantitative expansion toward methodological refinement, validation of contextualised assessment tools, AI-integrated evaluation frameworks, and longitudinal studies to assess the sustained impact of educational interventions. Accordingly, this study not only provides a comprehensive scientific overview of the domain but also offers an evidence-based foundation to inform educational policy, technological implementation in evaluation practices, and the development of more responsive and socially grounded assessment models. Thus, future research must prioritise methodological rigor, interdisciplinary collaboration, and the inclusion of underserved educational contexts to ensure that evaluation practices contribute meaningfully to equitable and innovation-driven educational transformation in Indonesia.

Supplementary Materials: Not applicable. No supplementary materials are provided for this article.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding.

Acknowledgment: The authors acknowledge the use of R software (R 4.5.2) and the Biblioshiny package for data analysis and visualization in this study [19].

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] M. Fitrah *et al.*, “Are Teachers Ready to Adopt Deep Learning Pedagogy? The Role of Technology and 21st-Century Competencies Amid Educational Policy Reform,” *Education Sciences*, vol. 15, no. 10, Oct. 2025, <https://doi.org/10.3390/educsci15101344>.
- [2] A. Cahyadi, Hendryadi, S. Widyastuti, and Suryani, “COVID-19, emergency remote teaching evaluation: the case of Indonesia,” *Education and Information Technologies*, vol. 27, no. 2, pp. 2165–2179, Aug. 2021, <https://doi.org/10.1007/s10639-021-10680-3>.
- [3] D. Mulyaman and C. Catherine, “Comparative Analysis on South Korea and Indonesia’s Outcome-Based Education (OBE) Resilience Amidst The COVID-19 Pandemic,” *Humaniora: Journal of Indonesia Culture and Society*, vol. 13, no. 3, pp. 265–275, Nov. 2022, <https://doi.org/10.21512/humaniora.v13i3.7734>.

-
- [4] M. D. H. Rahiem, "The emergency remote learning experience of university students in Indonesia amidst the COVID-19 crisis," *International Journal of Learning, Teaching and Educational Research*, vol. 19, no. 6, pp. 1–26, Jun. 2020, <https://doi.org/10.26803/ijlter.19.6.1>.
- [5] P. Priyanto and N. P. M. Suhandi, "Indonesian Higher Education Institutions Competitiveness and Digital Transformation Initiatives," *JPI (Jurnal Pendidikan Indonesia)*, vol. 11, no. 1, pp. 86–95, Feb. 2022, <https://doi.org/10.23887/jpi-undiksha.v11i1.34263>.
- [6] OECD. *Shaping Digital Education: Enabling Factors for Quality, Equity and Efficiency*. OECD Publishing, 2023, <https://doi.org/10.1787/bac4dc9f-en>.
- [7] M. F. Alquraan, S. Alazzam, and D. Farhat, "Assessing the structural relationships between the CIPP model components in teacher education program," *Asian Education and Development Studies*, vol. 14, no. 1, pp. 103–114, Jan. 2025, <https://doi.org/10.1108/AEDS-09-2024-0217>.
- [8] A. Bingölbali, A. Aslan, V. Batdi, and E. Cinkara, "Mixed-meta Method Concerning the Effect of Blended Learning Practices on Students' Academic Achievement in Higher Education Settings," *Sage Open*, vol. 15, no. 2, Apr. 2025, <https://doi.org/10.1177/21582440251336646>.
- [9] T. T. Wu, H. Y. Lee, P. H. Li, C. N. Huang, and Y. M. Huang, "Promoting Self-Regulation Progress and Knowledge Construction in Blended Learning via ChatGPT-Based Learning Aid," *Journal of Educational Computing Research*, vol. 61, no. 8, pp. 3–31, Jan. 2024, <https://doi.org/10.1177/07356331231191125>.
- [10] M. Bond, S. Bedenlier, V. I. Marín, and M. Händel, "Emergency remote teaching in higher education: mapping the first global online semester," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, p. 50, Dec. 2021, <https://doi.org/10.1186/s41239-021-00282-x>.
- [11] M. V. Carbonari *et al.*, "The Impact and Implementation of Academic Interventions During COVID-19: Evidence from the Road to Recovery Project," *AERA Open*, vol. 10, Jan. 2024, <https://doi.org/10.1177/23328584241281286>.
- [12] N. A. Dela Cruz, A. J. Adona, R. Molato-Gayares, and A. Park, "Learning loss and recovery from the COVID-19 pandemic: A systematic review of evidence," *Int. J. Educ. Dev.*, vol. 115, May 2025, <https://doi.org/10.1016/j.ijedudev.2025.103271>.
- [13] L. Boeskens and K. Meyer, "Policies for the digital transformation of school education: Evidence from the Policy Survey on School Education in the Digital Age," *OECD Education Working Papers*, Mar. 2025, <https://doi.org/10.1787/464dab4d-en>.
- [14] D. T. Figueroa, C. Rawkins, E. Qing, and H. M. de Sousa. *Education Policy Outlook 2024: Reshaping Teaching into a Thriving Profession from ABCs to AI*. OECD Publishing, 2024, <https://doi.org/10.1787/dd5140e4-en>.
- [15] A. Santoso, H. Retnawati, M. Nikma Rosyada, E. Apino, I. Rafi, and A. Dauletkulova, "Factors Influencing the Success of Students' Learning through Online Learning/Distance Education: A Bibliometric Analysis of Scopus Database," *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran, dan Pembelajaran*, vol. 10, no. 1, pp. 286–299, Mar. 2024, <https://doi.org/10.33394/jk.v10i1.10856>.
- [16] K. Aviory, H. Retnawati, and S. Sudiyatno, "Factors affecting mathematics achievement: Online learning, self-efficacy, and self-regulated learning," *Journal of Pedagogical Research*, vol. 9, no. 3, pp. 276–290, Jun. 2025, <https://doi.org/10.33902/JPR.202534085>.
- [17] D. M. Zulqadri, A. Mustadi, and H. Retnawati, "Digital Safety During Online Learning: What We Do to Protect Our Student?," *Jurnal Iqra': Kajian Ilmu Pendidikan*, vol. 7, no. 1, pp. 178–191, Jun. 2022, <https://doi.org/10.25217/ji.v7i1.1746>.
- [18] R. A. Fauzani, A. Senen, and H. Retnawati, "Challenges for Elementary School Teachers in Attitude Assessment During Online Learning," *Journal of Education Research and Evaluation*, vol. 5, no. 3, pp. 362–372, Jun. 2021, <https://ejournal.undiksha.ac.id/index.php/JERE/article/view/33226>.
- [19] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *J. Informetr.*, vol. 11, no. 4, pp. 959–975, Nov. 2017, <https://doi.org/10.1016/j.joi.2017.08.007>.
-

- [20] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *J. Bus. Res.*, vol. 133, pp. 285–296, Sep. 2021, <https://doi.org/10.1016/j.jbusres.2021.04.070>.
- [21] L. Lundberg and L. L. Se, "Bibliometric mining of research directions and trends for big data," *Journal of Big Data*, vol. 10, no. 1, pp. 112–, Jul. 2023, <https://doi.org/10.1186/s40537-023-00793-6>.
- [22] A. D. Oktarina, S. Hamdi, A. Wijaya, Z. Zafrullah, and S. Rashid, "Research Trends on the Implementation of Digital Literacy in Education: A Bibliometric Analysis of the Scopus Database," *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, vol. 7, no. 3, pp. 481–502, Sep. 2025, <https://doi.org/10.23917/ijolae.v7i3.11087>.
- [23] N. E. Sudakova, T. N. Savina, A. R. Masalimova, M. N. Mikhaylovsky, L. G. Karandeeva, and S. P. Zhdanov, "Online Formative Assessment in Higher Education: Bibliometric Analysis," *Educ. Sci. (Basel)*, vol. 12, no. 3, Mar. 2022, <https://doi.org/10.3390/educsci12030209>.
- [24] F. Siddiq, A. D. Olofsson, J. O. Lindberg, and L. Tomczyk, "Special issue: What will be the new normal? Digital competence and 21st-century skills: critical and emergent issues in education," *Education and Information Technologies*, vol. 29, no. 6, pp. 7697–7705, Aug. 2023, <https://doi.org/10.1007/s10639-023-12067-y>.
- [25] D. Permadi and E. H. Tiarto, "Bibliometric Analysis: Physics Online Learning in Indonesia (2020-2021)," *Online Learning In Educational Research (OLER)*, vol. 2, no. 1, pp. 11–18, Jun. 2022, <https://doi.org/10.58524/oler.v2i1.94>.
- [26] D. A. Parry, "Computing Research in South Africa: A Scientometric Investigation," *South African Computer Journal*, vol. 31, no. 1, pp. 51–79, Jul. 2019, <https://doi.org/10.18489/sacj.v31i1.674>.
- [27] G. Solorzano and V. Plevris, "Computational intelligence methods in simulation and modeling of structures: A state-of-the-art review using bibliometric maps," *Front. Built Environ.*, vol. 8, p. 1049616, Oct. 2022, <https://doi.org/10.3389/fbuil.2022.1049616>.
- [28] P. Chatterjee, A. Saha, S. Kadry, and G. Demir, *Bibliometric Analyses in Data-Driven Decision-Making*. Wiley, 2025, <https://doi.org/10.1002/9781394302581>.
- [29] M. Budler, I. Župič, and P. Trkman, "The development of business model research: A bibliometric review," *J. Bus. Res.*, vol. 135, pp. 480–495, Oct. 2021, <https://doi.org/10.1016/j.jbusres.2021.06.045>.
- [30] R. Maucuer and A. Renaud, "Business Model Research: A Bibliometric Analysis of Origins and Trends," *Management*, vol. 22, no. 2, pp. 176–215, 2019, <https://doi.org/10.3917/mana.222.0176>.
- [31] S. Ding, "Insights into bibliometric analyses," *JBI Evid. Synth.*, vol. 23, no. 4, pp. 564–565, Apr. 2025, <https://doi.org/10.11124/JBIES-24-00522>.
- [32] J. A. Moral-Muñoz, E. Herrera-Viedma, A. Santisteban-Espejo, and M. J. Cobo, "Software tools for conducting bibliometric analysis in science: An up-to-date review," *Profesional de la Informacion*, vol. 29, no. 1, pp. 1699–2407, Jan. 2020, <https://doi.org/10.3145/epi.2020.ene.03>.
- [33] B. D. Ghorbani, *Bibliometrix: Science Mapping Analysis with R Biblioshiny Based on Web of Science in Applied Linguistics*. Springer Nature, 2024, https://doi.org/10.1007/978-3-031-51726-6_8.
- [34] L. Karadağ and T. M. İ. Özberk, "Academic Mapping of Digital Marketing and Tourism Marketing: A Bibliometric Analysis," *Romaya Journal*, vol. 5, no. 1, pp. 21–36, Apr. 2025, <https://doi.org/10.5281/zenodo.15258006>.
- [35] H. T. Luong, "Foundations and trends in the darknet-related criminals in the last 10 years: a systematic literature review and bibliometric analysis," *Security Journal*, vol. 37, no. 3, pp. 535–574, May 2023, <https://doi.org/10.1057/s41284-023-00383-4>.
- [36] A. P. Costa, E. D. Moresi, I. Pinho, and M. Halaweh, "Integrating Bibliometrics and Qualitative Content Analysis for Conducting a Literature Review," *2023 24th International Arab Conference on Information Technology, ACIT 2023*, 2023, <https://doi.org/10.1109/ACIT58888.2023.10453680>.

-
- [37] B. Borsi, Z. Vida, and S. Soós, “Keyword standardization and restructuring: the impact on analysing network-based science maps in innovation management research,” *Scientometrics*, vol. 130, no. 2, pp. 593–617, Feb. 2025, <https://doi.org/10.1007/s11192-025-05232-2>.
- [38] L. Bredahl, “The Current and Evolving Landscape of Bibliometric Tools and Technologies,” *ALA TechSource*, vol. 58, no. 8, 2022, <https://doi.org/10.5860/ltr.58n8>.
- [39] G. Miller and E. Spiegel, “Guidelines for Research Data Integrity (GRDI),” *Scientific Data* 2025 12:1, vol. 12, no. 1, pp. 95–, Jan. 2025, <https://doi.org/10.1038/s41597-024-04312-x>.
- [40] R. L. Ravn, “The Effects of the COVID-19 Pandemic on School-to-School Transitions. A Study of Lower Secondary School Graduates,” *Soc. Policy Adm.*, vol. 59, no. 5, pp. 762–778, Sep. 2025, <https://doi.org/10.1111/spol.13092>.
- [41] S. Mammola, E. Piano, A. Doretto, E. Caprio, and D. Chamberlain, “Measuring the influence of non-scientific features on citations,” *Scientometrics*, vol. 127, no. 7, pp. 4123–4137, Jun. 2022, <https://doi.org/10.1007/s11192-022-04421-7>.
- [42] P. Dorta-González and E. Gómez-Déniz, “A Two-Stage Model for Factors Influencing Citation Counts,” *Publications*, vol. 13, no. 2, Jun. 2025, <https://doi.org/10.3390/publications13020029>.
- [43] T. Liskiewicz, G. Liskiewicz, and J. Paczesny, “Factors affecting the citations of papers in tribology journals,” *Scientometrics*, vol. 126, no. 4, pp. 3321–3336, Feb. 2021, <https://doi.org/10.1007/s11192-021-03870-w>.
- [44] J. Simangunsong and H. Mustikaningsih, “An evaluation of the implementation of study-from-home policy in senior high school during Covid-19 pandemic,” *REID (Research and Evaluation in Education)*, vol. 8, no. 1, pp. 67–77, Jun. 2022, <https://doi.org/10.21831/reid.v8i1.36734>.
- [45] R. Meylani and A. Prof, “A Comparative Analysis of Traditional and Modern Approaches to Assessment and Evaluation in Education,” *The Western Anatolia Journal of Educational Sciences (WAJES)*, vol. 15, no. 1, pp. 520–555, Apr. 2024, <https://doi.org/10.51460/baebd.1386737>.
- [46] K. Kabassi, “Comparing Multi-Criteria Decision Making Models for Evaluating Environmental Education Programs,” *Sustainability*, vol. 13, no. 20, Oct. 2021, <https://doi.org/10.3390/su132011220>.
- [47] A. C. Graham and S. McAleer, “An overview of realist evaluation for simulation-based education,” *Adv. Simul. (Lond.)*, vol. 3, no. 1, Dec. 2018, <https://doi.org/10.1186/s41077-018-0073-6>.
- [48] M. Coldwell and B. Maxwell, “Using evidence-informed logic models to bridge methods in educational evaluation,” *Review of Education*, vol. 6, no. 3, pp. 267–300, Oct. 2018, <https://doi.org/10.1002/rev3.3151>.
- [49] E. S. Boadu and I. Ile, “Evaluation innovation in Africa: Towards indigenously responsive evaluation (IRE) philosophies, methods and practices in Ghana,” *African Journal of Science, Technology, Innovation and Development*, vol. 15, no. 5, pp. 624–636, Jul. 2023, <https://doi.org/10.1080/20421338.2023.2173864>.