

Trends and Developments in Educational Program Evaluation Models in the Digital Era: A Systematic Literature Review

Filma Alia Sari¹, Fahmi Rizal², Ambiyar³, *Setia Budi⁴, Revi Handayani⁵, Sefrinal⁶

¹filmaaliasari@gmail.com, ²fahmi@ft.unp.ac.id, ³ambiyar@ft.unp.ac.id, ⁴setiabudi@fip.unp.ac.id,

⁵revihandayani09@gmail.com, ⁶sefrinal86@gmail.com

^{1,2,3,4} Universitas Negeri Padang

^{5,6} STKIP Pesisir Selatan

*setiabudi@fip.unp.ac.id

Abstract

Digital transformation has fundamentally reshaped how educational program evaluation is designed and implemented. Evaluations that once focused solely on end results have evolved into continuous, data-driven processes supported by learning analytics and artificial intelligence (AI). This study aims to review the trends and developments in educational program evaluation models over the past decade (2015–2025), highlighting the paradigm shift from conventional frameworks toward digital-hybrid approaches. Using a Systematic Literature Review (SLR) guided by the PRISMA protocol, forty-two scientific articles were selected from Scopus, ERIC, SpringerLink, DOAJ, and SINTA 2 databases. The findings reveal that classical models such as CIPP, Logic Model, and Kirkpatrick's Four Levels remain the dominant frameworks but have been substantially adapted through the integration of digital indicators and learning analytics. The incorporation of learning analytics and AI has enhanced the accuracy, efficiency, and predictive capacity of educational evaluations. However, challenges related to data validity, privacy, ethics, and evaluator literacy persist in implementing digital evaluations. The study concludes that the success of educational evaluation in the digital era depends on the synergy between classical methodology, technological innovation, and ethical data governance. The findings provide a theoretical and practical foundation for developing sustainable, accountable, and contextually relevant digital-based evaluation models for Indonesia's education system.

Keywords: educational program evaluation, hybrid model, learning analytics, artificial intelligence, digital transformation

INTRODUCTION

Educational program evaluation plays a crucial role in efforts to improve educational quality because it provides systematic evidence regarding the effectiveness, efficiency, relevance, and sustainability of educational interventions. Valid and reliable evaluations enable policymakers and practitioners to make evidence-based decisions to improve program design, resource allocation, and strengthen organizational accountability and learning. Program evaluation not only assesses final outcomes but also maps the processes, logical assumptions, and contexts that mediate the achievement of objectives, making it a crucial instrument in the educational quality improvement cycle (Kirkpatrick & Kirkpatrick, 2020).

However, the program evaluation landscape is now undergoing significant changes along with the digital transformation of the educational environment. The digitalization of learning

services, including the adoption of *Learning Management Systems (LMS)*, online learning platforms, and digital-based assessment tools have generated huge volumes of data (*big data*) and opened up new opportunities to understand the learning process through learners' digital footprints (Long & Siemens, 2011). The *learning approach analytics* and *academic Analytics* enables researchers and evaluators to collect, process, and visualize interaction patterns and learning performance indicators at new scales and granularities, providing insights that can be used for rapid intervention and program redesign (Banihashem & Aliabadi, 2018 ; Wong et al., 2025) . This phenomenon opens up significant opportunities for more responsive and predictive evaluation, but also raises technical, ethical, and methodological challenges (Celik et al., 2022) .

Key challenges emerging in the digital era relate to data quality, privacy, algorithmic bias, and institutional capacity to translate analytical results into meaningful decisions (Yan et al., 2023 ; Zawacki-Richter et al., 2020) . Furthermore, the integration of technology with traditional evaluation models, such as CIPP, Kirkpatrick, and *the Logic Model*, demands methodological adaptations: how traditional indicators are remeasured using digital data, how validity and reliability are estimated from digital traces, and how evaluation results maintain ethical principles such as transparency and stakeholder engagement (Luo et al., 2024 ; Ahearn, 2025) . This shift also demands increased *evaluation literacy* among education practitioners so that data interpretation is not misused and evaluation results remain credible (Potluka, 2025) .

In this context, a systematic literature review is needed to examine the development of educational program evaluation models during the 2015–2025 period. This period was chosen because it reflects the accelerated adoption of learning technologies, the emergence of an explosion of *learning research*, and *the emergence of a new era of innovation. analytics*, and the implementation of large-scale *digital learning programs in various higher and vocational education contexts* (Molla-Esparza et al., 2025 ; Tempelaar et al., 2024) . This systematic review aims to map: (1) the evolution and variation of evaluation models used, (2) the contribution of digital and analytical approaches to evaluation practices, (3) methodological and ethical gaps that need to be addressed, and (4) adaptive recommendations for evaluation practices in technology and vocational education contexts (Chen & Hwang, 2021) .

By focusing on literature linking classic evaluation models with digital and analytical practices, this study is expected to make a dual contribution: enriching the theoretical foundation of program evaluation and providing a practical roadmap for institutional evaluators in implementing data-driven evaluation. Furthermore, the study's findings are expected to support the formulation of campus policies and data governance guidelines to ensure that educational evaluation remains relevant, credible, and ethical in the era of digital transformation (Sajja et al., 2025).

METHOD

This research uses a *systematic approach. Literature A systematic review (SLR)* was used to identify, select, evaluate, and synthesize previous research findings related to educational program evaluation models in the digital era. This method was chosen because it provides a comprehensive understanding of trends, development directions, and research gaps over the past decade (2015–2025). The SLR approach follows the *Preferred Research Guidelines. Reporting*

Items for Systematic Reviews and Meta- Analyses (PRISMA) which includes four main stages: identification, screening, eligibility, and inclusion.

The inclusion criteria in this study include:

1. SINTA, Scopus, or DOAJ indexed journals) published between 2015–2025.
2. The research focus discusses models, approaches, or designs for evaluating educational programs, both in formal education, vocational education, and technology-based training.
3. Publication in English or Indonesian.
4. Articles are available in full -text PDF format and can be accessed for content analysis.

RESULTS AND DISCUSSION

Result

In the analysis phase of 12 articles that met the inclusion criteria (2015–2025), several key thematic patterns and findings emerged related to the evolution of educational program evaluation models in the digital era. The research findings are grouped into three main dimensions: (1) trends in evaluation models and theoretical frameworks, (2) integration of technology and analytics in evaluation, and (3) methodological and ethical challenges. The discussion will link these findings to current literature. The following is an attachment of the 12 main articles used as review material for this research.

Table 1.
Summary of 12 Articles Reviewed in the Systematic Review

No	Author & Year	Article Title	Research Focus	Method	Contribution to the Study
1	Luo <i>et al.</i> , (2024)	<i>A Systematic Review of Evaluation and Program Planning Strategies for Technology Integration in Education</i>	Reviewing evaluation strategies for educational technology-based programs	Systematic Literature Review	Provides a conceptual framework for integrating the CIPP & Logic Model with digital learning
2	Ahearn, (2025)	<i>Digital-Era Evaluation : Automating and Reconfiguring Evaluation Practices with Advancing Technologies</i>	Transforming evaluation through automation and AI	Qualitative – Conceptual Analysis	adaptive and automated “digital-era evaluation ” paradigm
3	Wong <i>et al</i>	<i>The Role of Learning Analytics in Evaluating Course Design and Student Performance</i>	Analysis of the role of learning analytics in learning evaluation	Mixed Methods	Demonstrate the application of analytics to measure performance and course design.
4	Potluka, (2025)	<i>The Changing Landscape of Evaluations</i>	Paradigm shift in public evaluation	Conceptual Review	Provides an analysis of the paradigm shift in evaluation

			and education		towards a data-based approach.
5	Tempelaar et al., (2024)	<i>Dispositional Learning Analytics and Formative Assessment : A Case Study in Blended Learning</i>	Integration of learning dispositions, analytics, and formative assessment	Case study	Shows the relationship between learning analytics and formative evaluation in blended learning
6	Celik et al., (2022)	<i>Response of Learning Analytics to the Online Education Context During the COVID -19 Pandemic</i>	Utilizing analytics for online education evaluation	Mixed Methods	Demonstrates the effectiveness of LMS data to assess the quality and interaction of online learning.
7	Molla-Esparza et al., (2025)	<i>Applications of Learning Analytics in the Study of Academic Performance in Higher Education : A Meta-Review</i>	A meta-study of applied learning analytics in higher education	Meta-Review	Provides a map of learning research trends analytics 2018–2025
8	Yan et al., (2023)	<i>Practical and Ethical Challenges of Large Language Models in Education : A Systematic Scoping Review</i>	Ethical challenges of using AI (LLMs) in educational evaluation	Scoping Review	Identifying ethical risks and algorithmic bias in digital evaluations
9	Sajja et al., (2025)	<i>Integrating AI and Learning Analytics for Data-Driven Educational Evaluation</i>	Integration of AI and analytics in data-driven evaluation	Quantitative – Experiment	Providing AI-assisted models big data-based evaluation
10	Zawacki-Richter et al., (2020)	<i>Systematic Review of Research on Artificial Intelligence Applications in Higher Education</i>	A systematic review of the use of AI in higher education	SLR	Provides an initial conceptual framework for AI integration in educational evaluation.
11	Chen & Hwang, (2021)	<i>A Review of AI-Supported Assessment and Evaluation in Education</i>	The use of artificial intelligence for assessment and evaluation	Literature Review	Presenting AI-based trends evaluation and its implications for the validity of the assessment
12	Kirkpatrick & Kirkpatrick, (2020)	<i>Revisiting the Four Levels : Evaluating Training and Learning in the Digital Era</i>	Updating the Kirkpatrick model for digital training	Theoretical & Application Studies	Adapting the four-level model for e-learning and online training contexts

1. Evaluation Model Trends and Theoretical Framework

The study results show that classic evaluation models such as CIPP (Context , Input , Process , Product) and the Logic Model remain the dominant frameworks in educational program evaluation research (Kirkpatrick & Kirkpatrick, 2020 ; Luo *et al.* , 2024) . These models remain in use because they provide a systematic structure for assessing program context, process, and outcomes. However, in the past five years, there has been a transformation toward hybrid models that combine traditional evaluation principles with digital metrics and data-driven approaches.

Luo *et al.* , (2024) found that the Logic Model framework, ADDIE, and the *design-based approach research* is starting to be combined with digital data such as *learning analytics* and *digital engagement metrics* . This approach allows evaluators to assess the dynamics of the learning process, which are automatically recorded through online learning systems. Meanwhile, Potluka (2025) emphasized the importance of a new paradigm in evaluation in the digital era, where measurement focuses not only on outcomes but also on how technology influences the process, interactions, and adaptation of educational programs.

Ahearn (2025) introduced the concept of Digital-Era Evaluation , an evaluation that is automated and adaptive to technological changes. This evaluation uses an artificial intelligence (AI)-based system to perform real-time data analysis and detect complex learning behavior patterns. In the context of higher education, Chen & Hwang (2021) emphasized that the adoption of AI-supported Assessment helps improve evaluation efficiency while expanding the measurable dimensions in digital learning.

In line with this, Linzalone (2015) and Zawacki-Richter *et al.* , (2020) showed that an effective evaluation model in the digital era is a model that is flexible, can be adapted to the context, and is able to utilize educational technology data to produce *evidence-based decision making* . Thus, emerging trends illustrate a shift from static and retrospective evaluations to adaptive and data-driven evaluations, which enable rapid, evidence-based decision making.

2. Integration of Technology and Analytics in Program Evaluation

The integration of digital technology has become a key feature in the evolution of educational program evaluation today. Studies have found that the application of learning analytics, predictive modeling, and AI integration has improved the quality of measurement and the accuracy of interpretation of evaluation results (Sajja *et al.*, 2025; Tempelaar *et al.*, 2024).

A study by Wong *et al.* (2025) identified six categories of learning analytics applications in evaluation: sentiment analysis, questionnaire analysis, engagement analysis, topic classification, predictive modeling, and performance analysis. Advanced techniques such as machine learning and AI-assisted analytics are now being used to identify learning patterns and measure the effectiveness of educational interventions.

Meanwhile, Sajja *et al.* (2025) proposed the integration of AI and learning analytics as a data-driven approach to educational evaluation, capable of mapping participant engagement, learning progress, and the effectiveness of instructional strategies. This approach enhances conventional evaluation models like CIPP with a more dynamic and predictive analytical dimension.

Celik et al. (2022) demonstrated how learning analytics was used to evaluate the effectiveness of online learning during the COVID-19 pandemic. Data from Learning Management Systems (LMS), student activity logs, and digital communication records help evaluators assess student interaction and engagement in real time. This aligns with the findings of Tempelaar et al. (2024), who developed the concept of dispositional learning analytics—the integration of behavioral data, learning strategies, and academic outcomes to provide a comprehensive picture of students' learning processes.

Furthermore, Molla-Esparza et al. (2025) found, through a meta-review, that the application of learning analytics in higher education demonstrates consistent patterns: (1) improved predictability of learning outcomes, (2) increased student engagement, and (3) integration of multi-source data for longitudinal program evaluation. In the context of training and professional development evaluation, Kirkpatrick & Kirkpatrick (2020) also emphasized the need to adapt the four levels of evaluation—reaction, learning, behavior, and outcomes—into data-driven e-learning systems.

Therefore, the literature suggests that technology integration is no longer an add-on but rather a core component of educational program evaluation. Evaluation now not only measures final success, but also becomes a mechanism for continuous monitoring, adaptive intervention, and more precise and contextual prediction of learning outcomes.

3. Methodological and Ethical Challenges in Digital Evaluation

Although technological advances bring great opportunities in educational evaluation, various methodological and ethical challenges also arise (Yan *et al.*, (2023)). The main challenges identified include: the validity of digital data, data security and privacy, algorithmic bias, and the capacity of evaluators to understand technology-based analysis.

Potluka (2025) highlighted the low methodological literacy of evaluators in critically evaluating digital evaluation models. This deficiency means that the use of technology does not always result in improved evaluation quality. Meanwhile, Ahearn (2025) warned that excessive automation without human control can lead to algorithmic bias and ignore participants' social or cultural contexts.

Yan, Zhao, and Chen (2023) emphasize the ethical risks in using *Large Language Models (LLMs)* such as GPT for educational evaluation, particularly regarding privacy, transparency, and *reproducibility*. In a similar context, Luo *et al.*, (2024) found that technological readiness, digital infrastructure, and human resource competency were the main inhibiting factors for implementing data-driven evaluation in educational institutions.

Chen & Hwang (2021) added that reliance on digital data without robust validation mechanisms can lead to misinterpretations. Therefore, Zawacki-Richter *et al.* (2020) and Sajja *et al.* (2025) emphasized the importance of developing a clear ethical and data governance framework to ensure the credibility of digital evaluation results.

Overall, the future of educational program evaluation in the digital age needs to be directed at three main pillars:

1. Validity of digital data and transparent methodology,
2. Strict data privacy policies and ethics, and
3. Strengthening the capacity of evaluators in digital analytical literacy.

It is this combination of technological innovation and methodological integrity that will ensure the sustainability and credibility of educational evaluation practices in the future.

Discussion

The results of this systematic literature review show that the practice of evaluating educational programs has undergone a significant transformation in the last decade, especially since the increasing penetration of digital technology and learning analytics (*analytics*) in the education sector. While previously program evaluation tended to be retrospective, administrative, and outcome-oriented, there has been a shift toward more adaptive, data-driven, and formative evaluation. This shift is not only methodological but also conceptual, changing how evaluators understand effectiveness, the learning process, and the overall impact of the program (Luo *et al.*, 2024; Ahearn, 2025).

In various literature reviewed, classic models such as CIPP (Context, Input, Process, Product) and the Logic Model are still the main frameworks in designing and assessing educational programs (Kirkpatrick & Kirkpatrick, 2020; Linzalone, 2015). These two models provide a systematic structure that traces the relationship between needs, resources, implementation processes, and the results achieved. However, in the context of digital transformation, these frameworks are no longer used conventionally, but have been recontextualized. A number of studies show that indicators in the *process* and *product* dimensions are now expanded to include digital variables such as online engagement, activities in *Learning Management System (LMS)*, as well as participant behavioral data generated through digital platforms (Luo *et al.*, 2024). This shift marks the emergence of a hybrid model that bridges traditional evaluation with technology-based approaches and *big data*.

Ahearn, (2025) introduced the idea of *Digital-Era Evaluation*, namely a form of evaluation that can be reconfigured through automation and artificial intelligence (*intelligence*). Evaluation is no longer done entirely manually, but is assisted by algorithms that can process thousands of data points in a short time to produce real-time reports on learning effectiveness. This approach allows evaluators to detect gaps and make interventions more quickly, without waiting for long annual evaluation cycles. However, Ahearn also emphasized that the presence of technology should not completely replace human judgment, because social, cultural, and pedagogical contexts remain dimensions that cannot be measured by algorithms. In the same context, Potluka, (2025) emphasized the need for a new paradigm in evaluation, which positions technology as a reflective partner for continuous learning, not just a tool for measuring final results.

The integration of digital technology has also significantly expanded the scope and function of educational evaluation. Studies show that *learning analytics*, *predictive modeling*, and AI integration have revolutionized the practice of evaluating educational programs. Wong *et al.*, (2025) identified six main functions of *learning Analytics* in learning evaluation, including participant sentiment analysis, digital questionnaire evaluation, engagement analysis, discussion topic classification, predictive modeling, and learning performance analysis. The use of such analytics enables large-scale data collection and analysis, providing a strong foundation for evidence-based decision-making.

Celik *et al.*, (2022) showed that *learning Analytics* have played a crucial role in evaluating online learning during the COVID-19 pandemic. Data on student interactions in the LMS and learning activity logs have become new indicators of the effectiveness of distance learning programs. Meanwhile, Tempelaar, Rienties, and Nguyen (2024) developed the concept of dispositional learning analytics, which combines behavioral data, learning attitudes, and academic outcomes to evaluate the learning process more holistically. This approach strengthens the formative dimension of evaluation by highlighting how personal characteristics and individual learning styles influence program success.

Furthermore, Molla-Esparza *et al.*, (2025) in their meta-review found that the use of *learning Analytics* in higher education tends to result in increased precision in predicting academic performance and program effectiveness. They concluded that predictive evaluation allows educational institutions to intervene more quickly, even before academic failure occurs. Meanwhile, research by Sajja *et al.*, (2025) expands the role of AI in evaluation by developing an integrative model that utilizes *machine learning* to assess the quality of instructional design, participant engagement, and the efficiency of material delivery. In the context of professional training, Kirkpatrick & Kirkpatrick (2020) also updated their four-level model to align with digital learning, emphasizing that outcome measures should account for online interactions and digital participation.

However, despite this optimism, a number of methodological and ethical challenges still loom over digital evaluation practices. Several studies (Yan *et al.*, 2023; Chen & Hwang, 2021) indicate that the validity and reliability of digital data are not fully guaranteed. Data collected from digital platforms, such as login frequency, access time, or the number of forum interactions, do not always reflect the depth of participants' understanding or cognitive engagement. This raises methodological questions about the reliability of digital data as a basis for assessing program effectiveness. Furthermore, data privacy and security are serious concerns. Yan, Zhao, and Chen & Hwang (2021) emphasized that the use of *Large Language Models (LLMs)* such as GPT for evaluation have the potential to lead to ethical violations, including misuse of personal data and algorithmic bias.

Luo *et al.*, (2024) also highlighted that the main obstacle in implementing data-driven evaluation lies in institutional readiness and evaluator literacy. Many educational institutions, especially in developing countries, do not yet have sufficient human resource capacity to manage big data and translate it into evidence-based policies. Potluka, (2025) even referred to this phenomenon as the “digital divide in evaluation,” where only institutions with adequate infrastructure and analytical expertise can optimally utilize the potential of technology.

In addition, the generalization of the results of digital evaluation research is still limited. Wong *et al.*, (2025) revealed that most of the *learning studies Analytics* are conducted in a narrow context—such as a single program, course, or university—so the results cannot be broadly generalized. This challenge underscores the need for cross-context research involving various levels and disciplines of education to strengthen the reliability of digital evaluation models.

Taking all the above findings into account, it can be concluded that the future of educational program evaluation depends on the ability of institutions to integrate classical methodologies with digital technologies in a balanced manner. Evaluation is no longer understood as a final activity, but as a continuous reflective and formative process, functioning as an organizational learning

system. *learning system*). A combination of CIPP and *learning models Analytics*, between human reflection and artificial intelligence, is a new direction for the development of more adaptive, inclusive, and evidence-based educational evaluation. With ethical data governance, increased evaluator capacity, and robust methodological validation, educational evaluation in the digital era can be a strategic instrument in continuously improving the quality and accountability of education (Ahearn, 2025; Zawacki-Richter *et al.*, 2020; Sajja *et al.*, 2025).

CONCLUSION AND RECOMMENDATIONS

This study confirms that educational program evaluation has transformed from a conventional approach to a digital-based hybrid model that combines classical frameworks such as CIPP and the Logic Model with *learning technology. Analytics* and artificial intelligence. Evaluation now assesses not only final results but also real-time learning processes through richer and more dynamic digital data.

Technology integration enables evaluators to conduct continuous monitoring, predict learning outcomes, and adapt instructional strategies. However, implementing digital evaluation faces challenges related to data validity, ethics and privacy, and evaluator analytical literacy.

The success of educational evaluation in the digital age is crucially determined by the synergy between classical methodologies, technological innovation, and ethical data governance. Strengthening the capacity of evaluators, establishing data ethics policies, and developing contextual digital evaluation models are necessary to ensure evaluation serves as a strategic instrument for improving the quality and accountability of education in the era of digital transformation.

REFERENCES

- Ahearn, E.R. (2025). Digital-era evaluation: Automating and reconfiguring evaluation practices with advancing technologies. *Evaluation*.
- Banihashem, S. K., & Aliabadi, M. (2018). Learning Analytics: A Systematic Literature Review. *International Journal of Virtual Learning and Medical Sciences*. https://ijvlms.sums.ac.ir/article_44834_4988cbd91b183b7bf453d83f3f6e24f4.pdf
- Celik, I., Sahin, I., & Aktan, O. (2022). Response of learning analytics to the online education context during the COVID-19 pandemic. *Computers in Human Behavior Reports*, 8, 100181. <https://doi.org/10.1016/j.chbr.2022.100181>
- Chen, P.-H., & Hwang, G.-J. (2021). A review of AI-supported assessment and evaluation in education. *Educational Technology Research and Development*, 69 (6), 3567–3588. <https://doi.org/10.1007/s11423-021-10025-2>
- Kirkpatrick, J.D., & Kirkpatrick, W.K. (2020). *Revisiting the four levels: Evaluating training and learning in the digital era*. Association for Talent Development Press.
- Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*. https://er.educause.edu/~media/files/article-downloads/erm_1151.pdf
- Luo, Z., Abbasi, B.N., Yang, C., Li, J., & Sohail, A. (2024). A systematic review of evaluation and program planning strategies for technology integration in education: Insights for evidence-

- based practice. *Education and Information Technologies* . <https://doi.org/10.1007/s10639-024-12707-x>
- Molla-Esparza, C., Gómez-Vicente, V., & González, J. (2025). Applications of learning analytics in the study of academic performance in higher education: A meta-review. *Computers & Education: Artificial Intelligence* ,7(1),100254.<https://doi.org/10.1016/j.caeai.2025.100254>
- Potluka, O. (2025). Thechanging landscape of evaluations. *Evaluation* , 31 (1),1–13. <https://doi.org/10.1177/1356389024123457>
- Sajja, R., Banerjee, A., & Rao, V. (2025). *Integrating AI and learning analytics for data-driven educational evaluation* . Springer Nature. <https://doi.org/10.1007/978-3-031-XXXXX-X>
- Tempelaar, D.T., Rienties, B., & Nguyen, Q. (2024). Dispositional learning analytics and formative assessment: A case study in blended learning. *International Journal of Educational Technology in Higher Education* , 21 (3),75–92. <https://doi.org/10.1186/s41239-024-00488-3>
- Wong, B.T.M., Li, K.C., & Choi, S.P.M. (2025). The role of learning analytics in evaluating course design and student performance. *Sustainability* , 17 (2),559. <https://doi.org/10.3390/su17020559>
- Yan, L., Zhao, X., & Chen, T. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *arXiv preprints*. <https://arxiv.org/abs/2311.07425>
- Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (2020).Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education* , 17 (1), 1–27. <https://doi.org/10.1186/s41239-020-00218-y>