https://research.erukasyah.com/index.php/emaet/submissions

E - ISSN: XXXX-XXXX

Development of an Ethical AI-Based LMS and Its Impact on the Learning Outcomes of Teacher Education Students

Maharani[1], **Hamzidar**[2], **Hasniar** [3]

[1] Universitas Muhammadiyah Enrekang, Indonesia
[2] Universitas Muhammadiyah Enrekang, Indonesia
[3] Universitas Muhammadiyah Enrekang, Indonesia

ABSTRACT

The development of artificial intelligence (AI) in higher education demands attention to ethical aspects such as transparency, fairness, and privacy, particularly for Education and Teacher Training students as future educators. This study aims to develop a Learning Management System (LMS) that develops ethical AI principles and measures differences in student learning outcomes before and after its use. The study used a Research and Development approach with the Dick and Carey model, combined with a One-Group Pretest-Posttest design on 32 Teacher Training and Education students in South Sulawesi in 2026. Validation by media and materials experts resulted in an average Aiken's V coefficient of 0.83 (valid category). Field trials showed an increase in the average score from 60.77 to 68.28, with a significant paired t-test result ($p < 0.001$). These findings indicate an increase in learning outcomes after LMS use, but should be interpreted cautiously given the limitations of the design without a control group and the limited sample size. This study is expected to contribute to the development of LMSs based on ethical AI principles in higher education.

KEYWORDS : Dick and Carey, Ethical AI, Learning Outcomes, Teacher Education

Citation: Maharani, Maharani., Hamzidar, Hamzidar., & Hasniar, Hasniar. (2026). Development of an Ethical AI-Based LMS and Its Impact on the Learning Outcomes of Teacher Education Students. Ethical Media Architecture and Educational Technology, *1*(1). 24-30
https://doi.org/10.65224/emaet.75

Correspondence:

Maharani,
Maharani.unimen@studen.ic.id

Received: March 02 2026

Accepted: June 07 2026

Published: July 29 2026

INTRODUCTION

Artificial intelligence (AI) has evolved from an experimental stage to an integral part of higher education operations globally. Ellucian's annual survey of higher education institutions recorded an increase in institutional-level AI adoption from 49% in 2024 to 66% in 2025. A 17-point increase that marks the shift from mere experimentation to an operational integration strategy. This acceleration raises ethical concerns of equal magnitude. Since 2021, UNESCO has established the Recommendation on the Ethics of Artificial Intelligence as a global framework, which has been further popularized through the Guidance for Generative AI in Education and Research in 2023, which emphasizes human-centered privacy, data privacy, ethical standards, and teacher training, and the AI Competency Framework for Students in 2024, which focuses on four competency areas: human-centered thinking, AI ethics, technical applications, and design systems. This phenomenon confirms that the integration of AI in higher education cannot be separated from questions of algorithm fairness, transparency, and system accountability.

In Indonesia, similar trends are occurring at the policy and practice levels. The Ministry of Education, Culture, Research, and Technology has issued official guidelines for the use of generative AI in higher education, covering four ethical areas: academic integrity, security and safety of use, equity and transparency, and environmental impact (Kemendikbudristek, 2024). However, these policies may not necessarily align with student understanding. Salsabila and Sohidin found that students implicitly have an understanding of academic ethics in the use of AI, but this understanding has not been fully implemented—practices of academic ethics violations are still found (Salsabila & Sohidin, 2024). This finding is particularly important for students in Education and Teacher Training programs in Indonesia, as they are not only users of AI during their studies but also future educators who will shape ethical awareness of AI in their own students—thus, developing ethical understanding at the student level has the potential to be replicated in future classrooms.

On the other hand, the development of AI-based Learning Management Systems (LMS) has so far been largely directed at personalized learning and administrative efficiency, rather than at how the system is designed ethically from the outset (ethics-by-design). A study on a five-dimensional trust framework for educational AI systems identified recurring risks such as data breaches, bias, opaque decision-making, and ambiguous responsibilities in educational AI systems, as well as incentives for governance strategies based on ethics-by-design principles (Innovation and New Technology, 2025). Meanwhile, studies on AI integration in the Indonesian higher education ecosystem still focus on personalized learning, administrative efficiency, and data-driven research, with the main challenges being infrastructure fragmentation and technological literacy (Jurnal Teknologi Sistem Informasi, 2025) — without explicitly discussing system design that integrates ethical AI principles as a functional component of the LMS itself, while empirically measuring its impact on student learning outcomes for prospective teachers.

Based on these similarities, this study aims to develop a Learning Management System (LMS) that integrates ethical AI principles (transparency, fairness, and data privacy) using the Dick and Carey development model, and to measure the differences in learning outcomes of students in the Teacher Education and Training study program before and after using the LMS.

RESEARCH METHODOLOGY

Research Design

This study uses a Research and Development (R&D) approach, a research method used to produce a specific product and test its effectiveness or differences in results after use (Sugiyono, 2019). This approach is combined with a pre-experimental One-Group Pretest-Posttest Design to measure differences in learning outcomes before and after product use. This combination was chosen because the study has two distinct objectives: to produce a product in the form of a Learning Management System (LMS) that applies ethical AI principles, and to measure differences in student learning outcomes before and after use of the product. These two objectives are approached with separate methodological instruments so that the validity of each claim (product validity vs. differences in learning outcomes) does not interfere with each other.

Research Location and Subjects

The population in this study was all 32 students of the Teacher Education Study Program in the academic year of 2026. Considering the population size was under 100, the sample was taken as a whole (total sampling), following the guideline that if the research subjects number less than 100 people, it is best to take all of them (Arikunto, 2012). Thus, all 32 students participated as subjects of the field trial as well as subjects of learning outcome measurements (pretest-posttest).

Development Procedure

The product development phase adapts the Dick and Carey model, first introduced in the book, The Systematic Design of Instruction, and continually updated in its latest edition (Dick, Carey, & Carey, 2015). This model consists of ten interrelated components that function as a system, rather than as separate stages: (1) identification of instructional objectives, (2) instructional analysis, (3) learner and context analysis, (4) formulation of performance objectives, (5) development of assessment instruments, (6) development of instructional strategies, (7) development and selection of instructional materials, (8) formative evaluation, (9) revision, and (10) summative evaluation.

In this study, the seventh component (instructional materials development) was adapted into the LMS development process, incorporating features that represent ethical AI principles—transparency, fairness, and privacy—core dimensions consistently emerging in the framework of trustworthy AI in educational contexts (Innovation and Emerging Technologies, 2025). This adaptation is necessary because Dick and Carey was primarily designed for developing strategies and instructional materials, not for developing software systems. Therefore, the products strictly "developed" using the steps of this model are the instructional design and functional system specifications, while the technical implementation (coding) follows these specifications as a derivative.

Summative evaluation of the steps involved the implementation of a One-Group Pretest-Posttest design to measure differences in learning outcomes (Sugiyono, 2019), with the notation:

O_1 X O_2

O_1 is the measurement of learning outcomes before the treatment (pretest), X is the treatment in the form of using an LMS based on ethical AI principles in the learning process, and O_2 is the measurement of learning outcomes after the treatment (posttest). Differences between O_1 and O_2 were analyzed using paired-difference tests to determine the significance of the changes.

This design was chosen due to its practical implementation within the context of a single class/study program without a comparison group. Specifically, this design did not control for internal validity threats such as maturation effects, history, or testing effects. Consequently, the findings of this study should be interpreted as indicating differences in learning outcomes following product use, rather than as evidence of a causal relationship between the LMS and improved learning outcomes.

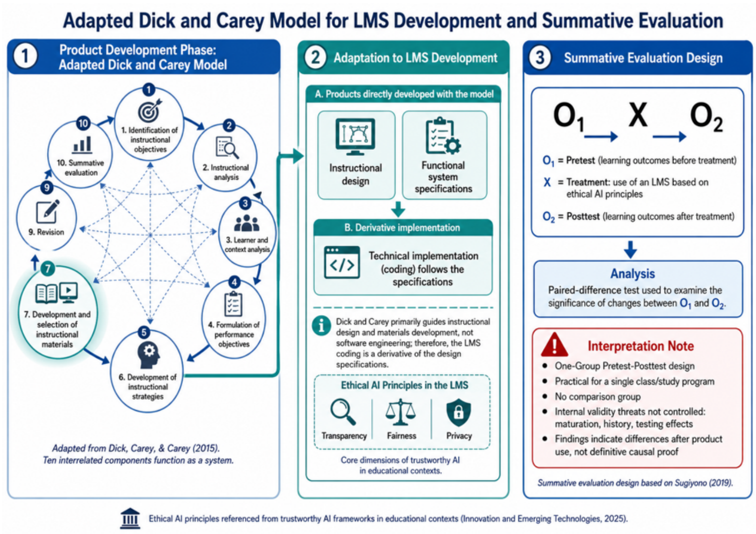

Figure 1. Adapted Dick and Carey Model for Ethical AI-Based LMS Development and Summative Evaluation

Instrumen

This study used three instruments. First, a Likert-scale expert validity questionnaire (1–4) to assess the feasibility of the LMS from a media/system perspective (technical feasibility and standards of transparency, fairness, and privacy) and a content perspective (content alignment with instructional objectives). Data were analyzed using Aiken's V coefficient, $V = S / [n(c-1)]$, where S is the sum of the rater scores from the lowest score, n is the number of validators, and c is the number of assessment categories; V values range from 0–1, with the closer to 1 the more valid (Aiken, 1985). With only two validators, these results are susceptible to extreme bias, and the alternative Content Validity Ratio method even requires a minimum of five validators for its critical value table to be valid (Lawshe, 1975) — it would be better to increase the number of validators to three.

Second, a Likert-scale practicality questionnaire to measure students' perceptions of the system's ease of navigation and transparency was completed after the field trial.

Third, the learning outcome instrument consisted of a combined cognitive test (pretest-posttest multiple-choice/essay) to measure understanding of the material, and an AI ethical awareness scale (non-test, Likert-type) to measure changes in students' awareness of AI transparency, fairness, and privacy. The reliability of both instruments was tested using Cronbach's Alpha, with a minimum threshold of 0.70 as the generally accepted criterion.

RESULT AND DISCUSSION

The expert validation results showed that the Ethical AI-Based Learning Management System (LMS) achieved an average Aiken’s V coefficient of 0.83, indicating that the product was valid and appropriate for implementation. The highest validation score was obtained for usability (V = 1.00), followed by technical system feasibility and content appropriateness (V = 0.83). Meanwhile, AI transparency obtained the lowest score (V = 0.67), indicating that this aspect required further improvement. Aiken’s V is commonly used to assess content validity based on expert judgments, with coefficients ranging from 0 to 1 (Aiken, 1985).

Table 1. Expert Validation Results of the Ethical AI-Based LMS

Assessed Aspect	Media Expert	Content Expert	Aiken's V	Category
Technical system feasibility	4	3	0.83	Valid
Implementation of AI transparency	3	3	0.67	Moderately valid
Content appropriateness	3	4	0.83	Valid
Usability	4	4	1.00	Highly valid
Average			**0.83**	**Valid**

The findings indicate that the developed LMS was generally feasible for use in the learning process. However, the relatively lower score for AI transparency suggests the need to provide clearer information regarding the role of AI, the purposes for which AI is used, the limitations of AI-generated outputs, and the management of user data. This is consistent with UNESCO (2023), which emphasizes that the use of AI in education should adopt a human-centered approach and ensure transparency, privacy protection, data security, and ethical responsibility. Therefore, transparency should not merely indicate the presence of AI features but should also enable users to understand how and why AI contributes to the learning process.

The impact assessment involved 32 participants using a One-Group Pretest–Posttest design. The mean pretest score was 60.77**,** which increased to 68.28 in the posttest, representing a mean increase of 7.51 points**.** Descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of Pretest and Posttest Scores

Statistic	Pretest	Posttest
N	32	32
Mean	60.77	68.28
Standard deviation	8.51	8.74
Minimum	44.80	49.50
Maximum	78.70	81.80

The Shapiro–Wilk normality test performed on the difference scores yielded a statistic of 0.965 with $p = 0.383$**.** Since $p > 0.05$, the difference scores did not significantly deviate from normality, allowing the use of a parametric paired-samples t-test (Shapiro & Wilk, 1965).

Table 3. Paired-Samples t-Test Results

Analysis	Value
Mean difference	7.51
t-value	8.25
df	31
p-value	< 0.001
Cohen's d	1.46

The paired-samples t-test revealed a statistically significant difference between pretest and posttest scores, **$t(31) = 8.25$, $p < 0.001$**. The effect size was also substantial, with **Cohen's $d = 1.46$**, indicating a large magnitude of change based on Cohen's (1988) conventional interpretation. These findings suggest that the implementation of the Ethical AI-Based LMS was followed by a substantial improvement in students' learning outcomes.

The improvement may be associated with the integration of structured instructional design, system usability, and ethical AI principles such as transparency, fairness, and privacy. A well-designed LMS can facilitate more systematic access to learning materials and activities, while ethical AI principles can support more transparent, responsible, and human-centered technology use. In particular, the high usability validation score suggests that students were able to interact with the system with relatively minimal technical barriers, allowing greater attention to be directed toward learning activities.

Nevertheless, these findings should be interpreted cautiously. Because the study employed a **One-Group Pretest–Posttest design without a control group**, the observed improvement cannot be attributed exclusively to the LMS. Potential threats to internal validity, including maturation, history, and testing effects, may also have influenced the posttest scores. Therefore, the findings provide evidence of a **significant improvement in learning outcomes following the use of the Ethical AI-Based LMS**, rather than definitive evidence of a causal effect. Future studies should employ quasi-experimental or experimental designs with comparison groups to establish stronger evidence regarding the effectiveness of the developed LMS.

CONCLUSION

This study produced a Learning Management System (LMS) that applies ethical AI principles through an adaptation of the Dick and Carey model, with an average validity level of 0.83 based on the assessment of media and material experts (valid category). The results of a field trial sample on 32 students of the Teacher Education Study Program indicated an increase in learning outcomes after using the LMS, from an average of 60.77 to 68.28 ($p < 0.001$), although this finding needs to be interpreted with caution considering the study design without a control group and limited numbers so it cannot be generalized as evidence of causal effectiveness. Further research is recommended to use a design with a control group and a larger number of expert validators to strengthen the validity of the findings.

ACKNOWLEDGEMENT

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

AUTHORS' CONTRIBUTION

Look this example below:

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

Author 2: Conceptualization; Data curation; In-vestigation.

Author 3: Data curation; Investigation.

.

REFERENCES

Banzon, A. M. E., Beever, J., & Taub, M. (2024). Facial expression recognition in classrooms: Ethical considerations and proposed guidelines for affect detection in educational settings. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2023.3275624

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet, 327(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), Usability evaluation in industry (pp. 189–194). Taylor & Francis.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6(4), 284–290.

Ekman, P. (1992). Are there basic emotions? Psychological Review, 99(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Guerrero-Roldán, A. E., Rodríguez-González, M. E., Karadeniz, A., Kocdar, S., Aleksieva, L., & Peytcheva-Forsyth, R. (2020). Students' experiences on using an authentication and authorship checking system in e-assessment. Hacettepe Eğitim Dergisi. https://doi.org/10.16986/HUJE.2020063670

Gülen, E., & Göktaş, Y. (2026). Facial expression recognition for investigating attention and affective states in synchronous online higher education. Educational Technology Research and Development. https://doi.org/10.1007/s11423-026-10631-0

Hakim, G., Simangunsong, G., Ningrat, R., Rabika, J., Rusafni, M., Giri, E., & Mindara, G. (2024). Real-time facial emotion detection application with image processing based on convolutional neural network (CNN). International Journal of Electrical Engineering, Mathematics and Computer Science, 1, 27–36. https://doi.org/10.62951/ijeemcs.v1i4.123

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology, 52, 139–183.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75–105. https://doi.org/10.2307/25148625

Hidayat, Y., & Ansar, A. (2026). Kebijakan, kesenjangan dan tantangan terhadap transformasi digital dalam pengembangan pendidik di Sulawesi Selatan. Jurnal Manajemen Pendidikan, 11(2), 2289–2305. https://doi.org/10.34125/jmp.v11i2.2093

Lickona, T. (1991). Educating for character: How our schools can teach respect and responsibility. Bantam Books.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Picard, R. W. (2000). Affective computing. MIT Press.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health, 29(5), 489–497.

Samala, A. D., Rawas, S., Rahmadika, S., Criollo-C, S., Fikri, R., & Sandra, R. P. (2025). Virtual reality in education: Global trends, challenges, and impacts—game changer or passing trend? Discover Education, 4, Article 229. https://doi.org/10.1007/s44217-025-00650-z

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Whitehill, J., Serpell, Z., & Lin, Y.-C. (2014). The faces of engagement: Automatic recognition of student engagement from facial expressions. IEEE Transactions on Affective Computing, 5(1), 86–98. https://doi.org/10.1109/TAFFC.2014.2316163