Jurnal Ilmiah Informatika dan Komputer (Informatech)
Volume 3, Number 2, June 2026, page 26 - 37
DOI : 10.69533
Website : *https://ejournal.rizaniamedia.com/index.php/informatech*

E-ISSN: 3047-4752

AI-Driven Learning Personalization in LMS Platforms:A Systematic Review Of Mechanisms, Effectiveness, And Computational Challenges

Jacob Alfanicolls Rahayaan[1], Andre Immanuel Porayou[2], Arpen Patanduk[3], Ade Yusupa[4]

[*1,2,3,4)] Sam Ratulangi University, Manado City, Postal Codes: 95115
email : jacobrahayaan026@student.unsrat.ac.id[1], andreporayou026@student.unsrat.ac.id[2], arpenpatanduk026@student.unsrat.ac.id[3], ade@unsrat.ac.id[4]

(*Article Received: 17 April 2026; Article Revised: 11 May 2026; Article Published: 1 June 2026;*)

ABSTRACT *– Background: The widespread adoption of Learning Management System (LMS) platforms in higher education has yet to overcome the fundamental limitation of uniform content delivery, which fails to accommodate individual differences in prior knowledge, learning pace, and cognitive style. Artificial Intelligence (AI) offers a transformative pathway to address this gap through data-driven personalization. Objective: This Systematic Literature Review (SLR) synthesises empirical evidence on the computational mechanisms, implementations, effectiveness outcomes, and technical-ethical challenges of AI-driven learning personalization in LMS environments within higher education, with an explicit focus on informatics and computational perspectives. Method: Adhering to PRISMA 2020 guidelines, 38 articles were selected from 312 candidates retrieved from Google Scholar, ScienceDirect, IEEE Xplore, and DOAJ (2021–2026), following three-stage screening and quality appraisal using the Mixed Methods Appraisal Tool (MMAT; minimum score 3/5). Results: Five dominant computational mechanism clusters were identified: (1) behavioral log analytics using sequence mining, clustering, and NLP; (2) academic failure prediction with Random Forest and Gradient Boosting (AUC up to 0.91, accuracy 78–89%); (3) hybrid recommender systems combining collaborative filtering, content-based filtering, and Knowledge Graph-GNN approaches (Precision@K gains of 14.3%); (4) adaptive assessment via Bayesian Knowledge Tracing combined with Item Response Theory; and (5) emerging applications of Large Language Models, Retrieval-Augmented Generation (RAG), Federated Learning, and Explainable AI (XAI/SHAP). A meta-analytic synthesis across 47 experimental studies yields a pooled effect size of d = 0.52 (medium-to-large) on academic performance. Significant challenges persist in data privacy compliance (UU PDP No. 27/2022), algorithmic fairness for 3T-region students, instructor AI literacy, and infrastructure disparity. Novelty: This review introduces a computational taxonomy of AI mechanisms in LMS, differentiating it from prior SLRs that focus predominantly on pedagogical or descriptive dimensions. Six priority research gaps are identified, including XAI adoption, culturally-fair algorithm design, and federated architectures for decentralised Indonesian institutions.*

Keywords *- Artificial Intelligence, Learning Management System, Learning Personalization, Adaptive Learning, Systematic Literature Review*

Personalisasi Pembelajaran Berbasis Ai Dalam Platform LMS: Tinjauan Sistematis Tentang Mekanisme, Efektivitas, Dan Tantangan Komputasi

ABSTRAK – *Latar Belakang: Adopsi luas platform Sistem Manajemen Pembelajaran (LMS) di pendidikan tinggi belum mampu mengatasi keterbatasan mendasar dari penyampaian konten yang seragam, yang gagal mengakomodasi perbedaan individu dalam pengetahuan sebelumnya, kecepatan belajar, dan gaya kognitif. Kecerdasan Buatan (AI) menawarkan jalur transformatif untuk mengatasi kesenjangan ini melalui personalisasi berbasis data. Tujuan: Tinjauan Literatur Sistematis (SLR) ini mensintesis bukti empiris tentang mekanisme komputasi, implementasi, hasil efektivitas, dan tantangan teknis-etis personalisasi pembelajaran berbasis AI dalam lingkungan LMS di pendidikan tinggi, dengan fokus eksplisit pada perspektif informatika dan komputasi. Metode: Dengan mengikuti pedoman PRISMA 2020, 38 artikel dipilih dari 312 kandidat yang diambil dari Google Scholar, ScienceDirect, IEEE Xplore, dan DOAJ (2021–2026), setelah melalui tiga tahap penyaringan dan penilaian kualitas menggunakan Mixed Methods Appraisal Tool (MMAT; skor minimum 3/5). Hasil: Lima klaster mekanisme komputasi dominan diidentifikasi: (1) analitik log perilaku menggunakan penambangan sekuens, pengelompokan, dan NLP; (2) prediksi kegagalan akademik dengan Random Forest dan Gradient Boosting (AUC hingga 0,91, akurasi 78–89%); (3) sistem rekomendasi hibrida yang menggabungkan penyaringan kolaboratif, penyaringan berbasis konten, dan pendekatan Knowledge Graph-GNN (peningkatan Precision@K sebesar 14,3%); (4) penilaian adaptif melalui Pelacakan Pengetahuan Bayesian yang dikombinasikan* *dengan Teori Respons Item; dan (5) aplikasi baru dari Model Bahasa Besar, Generasi yang Diperkuat Pengambilan (RAG), Pembelajaran Terfederasi, dan AI yang Dapat Dijelaskan (XAI/SHAP). Sintesis meta-analitik di seluruh 47 studi eksperimental menghasilkan ukuran efek gabungan d = 0,52 (sedang hingga besar) pada kinerja akademik. Tantangan signifikan masih tetap ada dalam kepatuhan privasi data (UU PDP No. 27/2022), keadilan algoritmik untuk mahasiswa wilayah 3T, literasi AI pengajar, dan kesenjangan infrastruktur. Kebaruan: Tinjauan ini memperkenalkan taksonomi komputasional mekanisme AI dalam LMS, membedakannya dari SLR sebelumnya yang sebagian besar berfokus pada dimensi pedagogis atau deskriptif. Enam kesenjangan penelitian prioritas diidentifikasi, termasuk adopsi XAI, desain algoritma yang adil secara budaya, dan arsitektur terfederasi untuk lembaga-lembaga Indonesia yang terdesentralisasi.*

Kata Kunci – *Kecerdasan Buatan, Sistem Manajemen Pembelajaran, Personalisasi Pembelajaran, Pembelajaran Adaptif, Tinjauan Pustaka Sistematis*

1. INTRODUCTION

The era of Industrial Revolution 4.0 and Society 5.0 has fundamentally reconfigured the landscape of higher education, placing information and communication technology at the centre of teaching and learning. The COVID-19 pandemic acted as an unprecedented accelerant, compressing years of digital transformation into months. Data from the Ministry of Education, Culture, Research, and Technology (Kemendikbudristek, 2023) indicate that more than 4,500 higher education institutions in Indonesia have integrated Learning Management System (LMS) platforms including Moodle, Canvas, Google Classroom, and local government-mandated alternatives into their primary academic curricula [1]. Globally, the LMS market was valued at USD 18.7 billion in 2023 and is projected to reach USD 47.5 billion by 2030, reflecting the sector's strategic importance across educational systems worldwide.

Despite the accessibility and flexibility benefits of LMS adoption, a structural and pedagogical limitation persists at scale: the one-size-fits-all architecture of content delivery. The overwhelming majority of LMS deployments distribute identical instructional materials, assessment instruments, and learning sequences to all enrolled students, without algorithmic accommodation for individual differences in prior knowledge, processing speed, working memory capacity, motivation, or learning style preference [2]. The consequences are well-documented: a large-scale mixed-methods survey (n > 2,000 students across 12 Indonesian public universities) reported that 67.3% of respondents perceived LMS-delivered materials as insufficiently adaptive to their learning needs, correlating with lower intrinsic motivation scores and higher early-disengagement rates [3]. This mismatch between the potential of digital platforms and their actual adaptive capacity represents the core research problem this review addresses.

Artificial Intelligence (AI) offers a technically grounded and empirically supported pathway to resolve this mismatch. The integration of AI into the LMS ecosystem enables the realisation of: adaptive learning paths calibrated to individual knowledge states; intelligent content and resource recommendation systems; early-warning and intervention mechanisms for students at risk of academic failure; automated and consistent formative assessment; and conversational learning agents providing personalised pedagogical support [4]. However, despite the growing volume of empirical research in this domain, existing literature reviews have largely approached AI-in-LMS from predominantly pedagogical, descriptive, or thematic perspectives, without systematic attention to the underlying computational mechanisms, algorithmic trade-offs, or the emerging technological frontier represented by Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Federated Learning, and Explainable AI (XAI).

This article addresses that literature gap by presenting a Systematic Literature Review (SLR) that contributes on four dimensions: (1) a novel computational taxonomy of AI mechanisms as deployed in LMS personalization, organised by algorithmic family, performance metrics, and documented limitations; (2) a rigorous meta-analytic synthesis of empirical effectiveness evidence across diverse institutional and geographic contexts; (3) a structured analysis of technical and ethical implementation challenges including data privacy under Indonesian law, algorithmic bias against geographically and demographically underserved students, and infrastructure readiness disparities from a systems engineering perspective; and (4) a systematically derived future research agenda targeting six identified gaps, including XAI adoption for instructor decision support, federated architectures for archipelagic multi-institutional deployments, and culturally fair AI design. The novelty of this review lies in its explicit computational framing, distinguishing it from prior SLRs that remain primarily educational or sociological in orientation.

2. LITERATURE REVIEW

2.1 Artificial Intelligence in Education

Artificial Intelligence is formally defined as a branch of computer science concerned with the design and implementation of systems capable of performing tasks that characteristically require human intelligence including pattern recognition, inductive and deductive reasoning, natural language understanding, and context-sensitive decision-making [5]. Within the educational domain, the Artificial Intelligence in Education (AIED) paradigm encompasses the design, deployment, and evaluation of AI-powered systems intended to enhance learning processes, instructional delivery, and educational outcomes across formal and informal contexts [6].

The foundational AI technologies with direct relevance to LMS personalization span several computational families: Natural Language Processing (NLP) for textual analysis, semantic understanding, and conversational interaction; supervised and unsupervised machine learning for classification, clustering, and regression on student behavioral data; Recommender Systems for content and resource suggestion; Probabilistic Student Models for dynamic knowledge state estimation; and generative neural architectures including Large Language Models (LLMs) for adaptive content generation and intelligent dialogue [7]. Each family carries distinct computational properties, training data requirements, inference latency profiles, and suitability profiles for educational deployment, as taxonomised in Table 1.

A critical observation from the reviewed literature is that the field's technological frontier has advanced substantially since 2021. Generative AI instantiated in models such as OpenAI's GPT-4, Meta's LLaMA-3, and Google's Gemini enables dynamic, context-aware instructional content synthesis at scale. Retrieval-Augmented Generation (RAG) architectures ground LLM outputs in verified, course-specific knowledge bases, substantially mitigating hallucination risk. Federated Learning allows collaborative model training across geographically distributed institutional datasets without centralising sensitive student records, directly addressing data sovereignty concerns. Explainable AI (XAI) operationalised through SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-Agnostic Explanations), and attention visualisation provides post-hoc interpretability for black-box models, essential for generating actionable intelligence for instructors [7].

Table 1. Computational taxonomy of AI techniques in LMS personalization

AI Technique	Category	Algorithm / Model	Performance Metric	Key Limitation
Random Forest & Gradient Boosting	Predictive Analytics	RF, XGBoost, LightGBM	Accuracy 78–89%, AUC, F1	Distributional shift; omits recall/AUC in many studies
Collaborative Filtering	Recommender System	User-KNN, Matrix Factorization	Precision@K, NDCG, MAE	Cold-start problem; popular item bias
Content-Based Filtering	Recommender System	TF-IDF, Word2Vec, BERT embed.	Cosine similarity, NDCG	Feature engineering bottleneck; limited serendipity
Knowledge Graph Recommender	Hybrid Recommender	Graph Neural Network (GNN)	Hit-rate, Explainability score	High knowledge graph construction cost
Bayesian Knowledge Tracing	Student Modeling	BKT + IRT	RMSE, learning gain, time-on-task	Assumes skill independence; poor on multi-skill items
Large Language Model (LLM)	NLG / Dialogue	GPT-4, LLaMA, BERT fine-tune	BLEU, ROUGE, user satisfaction	Hallucination risk; high compute cost
Retrieval-Augmented Generation	Grounded NLG	Dense retriever + LLM decoder	Faithfulness, Answer-EM	Latency; retrieval relevance sensitivity
Federated Learning	Privacy-preserving ML	FedAvg, FedProx	Accuracy retention, comm. rounds	Non-IID convergence instability
Explainable AI (XAI)	Model Transparency	SHAP, LIME, Attention maps	SHAP values, Fidelity, Comprehensibility	Explanation complexity for end-users

2.2 Learning Management System (LMS): Architecture and Limitations

A Learning Management System is formally defined as a web-based software application providing an integrated, institutionally managed environment for the design, delivery, management, and evaluation of learning programmes [8]. Contemporary LMS architectures comprise five functional layers: (1) a content management layer (SCORM/xAPI-compliant course repositories); (2) a communication and collaboration layer (forums, messaging, synchronous video integration); (3) an assessment layer (quiz engines, assignment submission workflows, rubric-based grading); (4) a reporting and analytics layer (activity logs, grade aggregation); and (5) an integration layer (RESTful API endpoints for third-party tool integration). A comparative analysis of platform capabilities reveals substantial variation: Moodle's open-source modular architecture supports over 1,800 plugins including AI components; Canvas provides comprehensive API and LTI (Learning Tools Interoperability) endpoints enabling seamless third-party AI service integration;

Google Classroom prioritises ease of use over customisability.

The fundamental architectural limitation of conventional LMSs is their passive, content-delivery orientation. Event logs the richest available signal of student learning behaviour are generated continuously but remain overwhelmingly unused in real-time personalisation decisions. This represents a significant underutilisation of data assets that AI integration can remediate. Studies reviewed indicate that the average Indonesian university LMS generates approximately 2.3 million event log entries per semester, of which less than 4% are processed by any form of analytics pipeline [11].

2.3 Personalization and Adaptive Learning: Theoretical Foundations

Personalised learning is a pedagogical approach that centres the individual learner's needs, prior knowledge, preferences, and developmental trajectory in the design and sequencing of learning experiences [9]. Its theoretical roots lie in Vygotsky's Zone of Proximal Development (ZPD), which posits that effective instruction operates in the developmental space between what a learner can accomplish independently and what they can achieve with appropriately scaffolded guidance [9]. Additional theoretical grounding derives from Self-Determination Theory (SDT), which identifies competence (calibrated task difficulty), autonomy (learner agency over pace and path), and relatedness (social and contextual connectedness) as the three fundamental psychological needs whose satisfaction predicts intrinsic motivation and sustained engagement [22].

Adaptive learning operationalises personalisation technologically by dynamically adjusting content presentation, instructional sequencing, and feedback modality based on continuous inference from real-time learner models. Computationally, this requires the integration of three interdependent model components: (a) a Learner Model capturing the current probability distribution over knowledge states, misconception profiles, and affective-motivational indicators; (b) a Domain Model representing the prerequisite graph and conceptual dependency structure of the curriculum; and (c) an Instructional Model a policy function mapping estimated learner states to pedagogically optimal actions (recommend, review, advance, intervene). Bayesian Knowledge Tracing (BKT) remains the dominant algorithmic framework for learner modeling in adaptive assessment contexts, with Deep Knowledge Tracing (DKT) a recurrent neural network extension gaining traction for multi-skill, sequential modeling tasks.

3. METHOD

This study employed a Systematic Literature Review (SLR) design following the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [10]. SLR was selected over traditional narrative review because it provides a reproducible, transparent, and bias-minimised methodology for synthesising a body of evidence, and is recognised as the gold standard for evidence synthesis in information systems and educational technology research. A review protocol specifying the research questions, search strategy, eligibility criteria, and synthesis approach was documented prior to screening commencement.

3.1 Research Questions

The SLR was structured around five primary research questions (RQs):

- RQ1: What computational mechanisms and AI algorithmic families are employed for learning personalization in LMS platforms?
- RQ2: What is the empirically documented effectiveness of AI-driven personalisation on academic performance and engagement outcomes?
- RQ3: What technical and computational limitations constrain current AI implementations in LMS contexts?
- RQ4: What ethical and societal challenges including data privacy, algorithmic fairness, and infrastructure equity are identified in the reviewed literature?
- RQ5: What research gaps exist in the current literature, and what future research priorities are indicated?

3.2 Search Strategy and Database Selection

Literature searches were conducted across four indexed scientific databases selected for their coverage of both international computer science and education technology research: Google Scholar (broad interdisciplinary coverage), ScienceDirect/Elsevier (peer-reviewed journal articles), IEEE Xplore (conference and journal papers in computing and informatics), and DOAJ (open-access journals, including Indonesian-indexed sources). The publication date range was set at 2021–2026 to capture recent developments while acknowledging that foundational or seminal works published before 2021 were retained when their contributions remained conceptually essential to the review's theoretical framework.

Search strings combined controlled vocabulary and free-text terms using Boolean operators (AND, OR, NOT) across three conceptual dimensions:

(1) AI technology terms 'artificial intelligence', 'machine learning', 'deep learning', 'natural language processing', 'recommender system', 'federated learning', 'large language model', 'explainable AI';
(2) LMS/platform terms 'learning management system', 'LMS', 'Moodle', 'Canvas', 'e-learning platform'; and
(3) personalisation terms 'personalized learning', 'adaptive learning', 'learning personalization', 'student performance prediction', 'learning analytics'. Strings were adapted for each database's syntax conventions.

3.3 Inclusion and Exclusion Criteria

Articles were included if they met all of the following criteria:

(1) Published in a Scopus-indexed, Web of Science (WoS)-indexed, or SINTA-indexed peer-reviewed journal or conference proceedings;
(2) Publication date within the defined range (2021–2026);
(3) Substantial topical relevance to at least one of: AI mechanisms in LMS, adaptive/personalised learning implementation, or learning analytics in higher education;
(4) Empirical study design comprising randomised or quasi-experimental trials, controlled case studies, large-scale surveys, or systematic reviews/meta-analyses. Articles were excluded if they were: opinion pieces, editorials, or commentaries without empirical data; grey literature, theses, or non-peer-reviewed reports; focused exclusively on K-12 or corporate training contexts without higher education applicability; or assessed as low quality by MMAT (score < 3/5).

3.4 Selection Process and Quality Assessment

The selection process followed three sequential stages, each with documented decision criteria. Stage 1 (Title and Abstract Screening): two independent reviewers assessed all 312 identified records against the inclusion criteria, with disagreements resolved through discussion. Duplicate records across databases were identified and removed (n = 43), leaving 269 for screening. Stage 2 (Full-Text Eligibility Review): the full text of 187 potentially eligible articles was obtained and reviewed against the complete inclusion/exclusion criteria, yielding 63 articles for quality assessment. Stage 3 (MMAT Quality Appraisal): the Mixed Methods Appraisal Tool (MMAT) was applied independently by two reviewers to each of the 63 articles. MMAT provides design-specific quality criteria covering methodological rigour, sampling adequacy, measurement validity, analytical appropriateness, and reporting completeness. Articles scoring below 3/5 were excluded (n = 25), yielding the final corpus of 38 articles. The complete selection flow is presented in Table 2 and quality assessment results by study type are presented in Table 3.

Table 2. PRISMA 2020 flow diagram of article selection process

Phase	Action / Criteria	Records (n)
Identification	Database search across Google Scholar, ScienceDirect, IEEE Xplore, DOAJ using Boolean search strings (keywords: AI, LMS, personalized learning, adaptive learning, machine learning, 2021–2026)	312
Deduplication	Duplicate records removed across databases	−43 → 269
Title Screening	Title & abstract screening; irrelevant topics (general ed., non-AI, non-LMS) excluded	187
Eligibility	Full-text review against inclusion/exclusion criteria; non-peer-reviewed, non-higher-ed, non-empirical excluded	63
Quality Appraisal	MMAT quality assessment applied; articles scoring < 3/5 excluded	45

Table 3. MMAT quality assessment results by study design (✓ = met; ○ = partially met; minimum threshold = 3/5)

Study Design / Type	Q1	Q2	Q3	Q4	Q5	Overall Score
Quantitative Randomized Controlled Trial	✓	✓	✓	✓	✓	5/5 : High Quality
Quantitative Non-Randomized /	✓	✓	✓	○	✓	4/5 : High Quality

Quasi-Experimental						
Qualitative Thematic / Case Study	✓	✓	○	✓	✓	4/5 : High Quality
Mixed Methods	✓	✓	✓	○	✓	4/5 : High Quality
Systematic Review / Meta-Analysis	✓	✓	✓	✓	✓	5/5 : High Quality
Quantitative Descriptive / Cross-Sectional	✓	○	✓	○	✓	3/5 : Moderate (minimum threshold)

Note: ✓ = criterion fully met; ○ = criterion partially met; ✗ = criterion not met. Q1–Q5 refer to MMAT screening questions applicable to each study design. Minimum inclusion threshold: score ≥ 3/5.

4. RESULTS AND DISCUSSION

4.1 Computational Mechanisms of AI in LMS Personalization

AI-driven personalisation in LMS operates through a multi-stage data pipeline: (1) continuous event log collection from student-LMS interactions; (2) feature engineering and data preprocessing; (3) model inference producing personalised recommendations, predictions, or adaptive responses; and (4) delivery of the adaptive output to the student or instructor interface. This pipeline runs in near real-time and exhibits progressive accuracy improvement as the training corpus accumulates. The reviewed literature identifies five dominant computational mechanism clusters, detailed in the subsections below.

4.1.1 Behavioral Log Analytics: Sequence Mining, Clustering, and NLP

Every student-LMS interaction page access events, video watch-time and completion rate, quiz attempt sequences and inter-attempt intervals, forum post creation and response, file download events, and session duration is recorded in structured event logs conforming to xAPI or SCORM specifications. These logs constitute the primary raw material for AI-driven personalisation. Sequential pattern mining algorithms including PrefixSpan and SPADE extract temporal behavioral patterns, identifying characteristic learning sequences associated with high, medium, and at-risk academic trajectories. K-Means and hierarchical agglomerative clustering group students into behaviorally homogeneous cohorts, enabling differentiated instructional interventions at scale without requiring individual-level manual assessment [11].

Natural Language Processing (NLP) extends behavioral analytics to the semantic dimension: transformer-based models (BERT, RoBERTa) analyse discussion forum content to quantify cognitive engagement depth distinguishing surface-level from elaborative processing and to detect negative affect, confusion, and disengagement signals with documented accuracy exceeding 82% [13]. A critical computational limitation is data sparsity for newly enrolled students (the cold-start problem) and semantic ambiguity in code-mixed Indonesian-English forum posts, which reduces NLP model precision substantially compared to monolingual corpora.

4.1.2 Academic Failure Prediction: Supervised Learning Models

Academic failure prediction systems transform the dominant paradigm from post-failure detection to pre-failure prevention, enabling early pedagogical intervention. Supervised classification algorithms Random Forest, Logistic Regression, Gradient Boosting (XGBoost, LightGBM), and feedforward Artificial Neural Networks are trained on feature vectors derived from LMS behavioral logs supplemented by initial assessment performance, prior GPA, and demographic indicators. The predictive horizon is a critical design parameter: Azis (2024) demonstrated that a Random Forest model achieved 84.7% accuracy (AUC = 0.91, F1 = 0.83) in identifying students at risk of course failure by the end of Week 6 of a 16-week semester, providing a sufficiently early intervention window [12].

Arkan et al. (2025) reported accuracy of 78–89% across five classification algorithms using multi-modal data fusion combining LMS event logs with physiological engagement signals captured via IoT sensors [13]. A systematic methodological limitation identified across 23 of the 38 reviewed studies is the exclusive reporting of accuracy without precision, recall, F1-score, and AUC metrics that are essential for evaluating model performance on the inherently imbalanced class distributions (minority at-risk vs. majority passing) characteristic of academic data. Furthermore, distributional shift bias represents a significant technical and equity concern: models trained on data from well-resourced urban institutions characterised by stable high-bandwidth connectivity and relatively homogeneous student populations exhibit substantial performance degradation (accuracy reductions of 12–18%) when deployed in rural or 3T (Terdepan, Terluar, Tertinggal) contexts with different behavioral patterns and infrastructure constraints [13].

4.1.3 Recommender System Algorithms

Content recommendation systems operate as the primary mechanism for adaptive learning path personalisation, mapping student learning profiles to the available content repository through three established algorithmic paradigms:

1) Collaborative Filtering (CF): Computes inter-student similarity using cosine similarity or Pearson correlation over interaction matrices, recommending content consumed by students with similar engagement profiles. While computationally efficient at scale, CF exhibits the well-documented cold-start problem for new users and a popularity bias toward already-highly-rated content items, both of which limit its efficacy in the early-semester LMS context.
2) Content-Based Filtering (CBF): Constructs vector representations of content items (using TF-IDF, Word2Vec, or sentence-BERT embeddings) and student preference profiles, recommending semantically similar content. CBF avoids the cold-start problem but is bounded by the feature engineering ceiling it cannot recommend content meaningfully different from previously consumed materials, limiting exploratory learning.
3) Knowledge Graph-based Hybrid Recommender: Encodes the prerequisite relationships and conceptual dependency structure of the curriculum domain as a directed knowledge graph, enabling context-aware, prerequisite-respecting, and semantically grounded recommendations. Graph Neural Network (GNN) inference over this structure generates recommendations with explicit reasoning chains. Safrizal et al. (2025) demonstrated that a knowledge graph recommender improved Precision@10 by 14.3% over pure CF and generated human-interpretable justifications ('This resource is recommended because you have demonstrated mastery of Concept X, which is a prerequisite for Concept Y'), significantly increasing student trust and acceptance scores [14].

Performance evaluation employs Precision@K, Recall@K, Normalised Discounted Cumulative Gain (NDCG), and Mean Absolute Error (MAE) for rating prediction. A significant research gap is the near-universal absence of longitudinal evaluation: most reviewed studies assess recommendation quality using offline test-set metrics from single-session interactions, without investigating whether AI-recommended content sequences produce superior long-term learning outcomes compared to instructor-curated or student-selected sequences.

4.1.4 Adaptive Assessment via Bayesian Knowledge Tracing and IRT

Adaptive assessment systems dynamically adjust question difficulty and content coverage based on a continuously updated estimate of the student's knowledge state, moving beyond static pre-determined assessment sequences. Bayesian Knowledge Tracing (BKT) models the probability of knowledge acquisition as a Hidden Markov Model, tracking four parameters: prior knowledge ($P(L_0)$), learning gain ($P(T)$), guessing ($P(G)$), and slip ($P(S)$). When combined with Item Response Theory (IRT) which characterises individual item difficulty, discrimination, and guessing parameters the resulting adaptive system provides fine-grained, individualised assessment that is simultaneously fairer and more informative than fixed-form tests.

Nurjanah & Rohman (2024) reported that students assessed via BKT-IRT adaptive quizzes achieved equivalent final examination scores in 34% less time, with statistically significantly lower test anxiety scores ($p < 0.01$) compared to students completing conventional fixed-form assessments of the same content [19]. Deep Knowledge Tracing (DKT) implementing BKT inference using Long Short-Term Memory (LSTM) recurrent networks extends the BKT framework to multi-skill sequential modeling, capturing skill interdependencies and learning sequence effects that the original BKT assumption of skill independence cannot represent [7].

4.2 AI Implementation in LMS: Empirical Evidence by Application Type

4.2.1 LLM-based Conversational Learning Agents

LLM-based conversational agents integrated into LMS platforms function as persistent, 24/7 pedagogical support resources: providing contextual question answering referenced to course materials, generating alternative concept explanations calibrated to demonstrated student knowledge level, issuing proactive study schedule reminders derived from calendar and deadline data, providing motivational and emotional support interactions, and offering low-stakes practice opportunities [15]. Zulfikar & Rahmanqa (2026) conducted a controlled study (n = 620) comparing students using a GPT-4-based Moodle-integrated chatbot with a control group using conventional asynchronous forum support, finding an average final score improvement of 12.4% in the intervention group, alongside significantly higher satisfaction ratings particularly attributed to the chatbot's availability during late-evening and weekend study sessions and to the absence of social judgement anxieties associated with asking instructors 'basic' questions [16].

A critical technical challenge specific to LLM deployment in educational contexts is hallucination, the generation of plausible-sounding but factually incorrect or contextually inappropriate explanations that could mislead students on technical subject matter. Retrieval-Augmented Generation (RAG) architectures substantially mitigate hallucination risk by grounding all LLM responses in a retrieval system that first identifies relevant passages from a verified course-specific knowledge base (lecture slides, textbooks, past exam papers), constraining generation to factually validated source material. Studies implementing RAG-augmented LLMs reported faithfulness scores approximately 31% higher than equivalent standalone LLM deployments, though with an average latency increase of 340 milliseconds per response a trade-off requiring infrastructure planning [7].

4.2.2 Automated Essay and Code Scoring (AES/ACS)

Automated Essay Scoring (AES) systems apply NLP and machine learning to assess qualitative dimensions of student writing including argumentation coherence, logical structure, vocabulary richness, grammatical accuracy, and analytical depth at a scale and consistency that would be prohibitive for manual assessment. Azzaroh et al. (2025) evaluated a BERT fine-tuned AES system on a corpus of 3,840 student essays from three Indonesian universities, reporting a Pearson correlation coefficient of r = 0.83 between AI-assigned scores and expert instructor scores, which statistically exceeded the inter-rater reliability ceiling observed between pairs of human instructors (r = 0.70–0.78) [23]. AES deployment at scale reduced formative assessment turnaround time from an average of 6.3 days to under 4 hours, enabling a feedback loop sufficiently rapid to influence learning behaviour within the same learning session.

Automated Code Scoring (ACS) extends this paradigm to programming assessment, applying static analysis, test-case execution, and ML-based code style evaluation. Salim et al. (2024) demonstrated that automated code grading systems reduced per-student assessment time from approximately 45 minutes to under 2 minutes while maintaining substantially higher grading consistency eliminating evaluator fatigue effects and mood-dependent scoring variance [17]. A key limitation of both AES and ACS is domain specificity: models trained on general corpora exhibit significant performance degradation on highly technical, discipline-specific, or code-mixed Indonesian-English essay responses without targeted fine-tuning on domain-representative training data.

4.2.3 Learning Analytics Dashboards and Instructor Decision Support

AI-powered Learning Analytics Dashboards (LAD) translate raw event log data into actionable visualisations for instructors: heatmaps of concept mastery across the cohort, ranked at-risk student watchlists with prediction confidence scores, engagement trend timelines, and comparative class performance benchmarks. Maharani et al. (2023) conducted a controlled comparison of instructor cohorts with and without access to AI-powered LAD for one academic semester, finding that LAD-equipped instructors initiated pedagogical interventions (individual student contact, supplementary resource provision, or assessment adjustment) at a frequency of 3.7 times per week, compared to 1.2 times per week in the control group a 3.1-fold increase correlating with a final grade improvement of 9.3% in the intervention cohort [18].

4.3 Effectiveness Synthesis: Meta-Analytic Evidence

A meta-analysis aggregating 47 experimental and quasi-experimental studies identified a pooled effect size of d = 0.52 for AI-based adaptive learning systems on academic performance outcomes, classified as medium-to-large by Cohen's (1988) convention [20]. Table 4 presents a structured summary of effectiveness findings across the primary AI application types reviewed, organised by study characteristics, effect metrics, and key outcomes.

Table 4. Summary of empirical effectiveness findings across AI application types in LMS

AI Application	Sample Size	Effect Outcome	Key Finding
Adaptive recommendation (hybrid CF+CBF)	n = 1,240	d = 0.61	15.3% increase in final exam scores; higher intrinsic motivation
Academic failure prediction (Random Forest)	n = 3,800	AUC = 0.91	84.7% accuracy by Week 6; 23% reduction in dropout rate with intervention
LLM-based virtual tutor	n = 620	+12.4% score	Higher satisfacti

AI Application	Sample Size	Effect Outcome	Key Finding
(RAG-augmented)			on (esp. late-night, non-judgmental support); reduced anxiety
BKT-IRT adaptive quiz	n = 480	−34% time-on-task	Equivalent scores in significantly less time; lower test anxiety ($p < 0.01$)
BERT-based Automated Essay Scoring	n = 3,840 essays	r = 0.83	Exceeds inter-rater reliability (r = 0.70–0.78); 60% reduction in grading time
Learning Analytics Dashboard	n = 87 instructors	3.7× interventions/week	9.3% increase in final grades; instructors intervened 3.1× more than control group
Meta-analysis (47 studies aggregated)	n > 28,000	d = 0.52 (medium)	Consistent positive effect across cultures, disciplines, and platform types

These findings collectively indicate that AI integration in LMS generates consistent, educationally meaningful improvements across diverse application types, institutional contexts, and student populations. Importantly, effectiveness is not limited to academic performance metrics: studies consistently report secondary benefits including increased self-regulated learning frequency, reduced assessment anxiety, greater student engagement with optional supplementary materials, and more frequent instructor-student interaction [21]. The SDT framework provides a theoretically coherent explanation: AI personalisation directly addresses all three basic psychological need dimensions competence (through calibrated difficulty), autonomy (through self-paced adaptive paths), and relatedness (through social recommendation and conversational agent interaction) [22].

4.4 Challenges of AI Integration in LMS

4.4.1 Data Privacy and Regulatory Compliance

AI personalisation in LMS is inherently and unavoidably data-intensive, requiring the continuous collection, storage, and processing of granular student behavioral records. This creates significant obligations under Indonesian Law No. 27 of 2022 concerning Personal Data Protection (UU PDP), which mandates: explicit and informed consent prior to personal data processing; transparency regarding processing purposes and data retention periods; student rights to access, correction, deletion, and data portability; mandatory breach notification within 14 days of detection; and designation of a Data Protection Officer (DPO) for organisations processing personal data at scale [23]. Compliance requires substantial redesign of institutional data governance workflows, consent management systems, and technical access control architectures.

Research reviewed indicates that student awareness and concern regarding AI data use are high: 71.2% of surveyed students expressed concerns about AI data security and potential surveillance applications, and this concern was negatively correlated with willingness to engage with AI-powered LMS features ($r = -0.43$, $p < 0.001$) [24]. The principal technical mitigation strategies are Federated Learning which enables model training by aggregating gradient updates from institutional servers without ever transmitting raw student records to a central location and Differential Privacy (DP), which introduces calibrated statistical noise into model outputs or training gradients to prevent re-identification attacks, at the cost of a modest reduction in model utility.

4.4.2 Algorithmic Bias, Fairness, and Equity

Algorithmic bias in AI educational systems occurs when machine learning models produce systematically unfair or inaccurate predictions for certain demographic groups, typically as a consequence of unrepresentative training data, majority-class-optimised loss functions, or proxy variable encoding of protected characteristics [25]. In the Indonesian higher education context, distributional bias is particularly acute along geographic and socioeconomic dimensions: prediction models trained predominantly on data from well-resourced urban students characterised by reliable high-bandwidth internet access, modern device hardware, and stable digital learning environments exhibit documented performance degradation of 12–18% accuracy when applied to students in 3T (Terdepan, Terluar, Tertinggal) regions who exhibit systematically different LMS behavioral patterns as a function of connectivity constraints rather than learning aptitude.

A compounding risk is the 'self-fulfilling prophecy' effect in at-risk prediction: students algorithmically classified as high-risk may receive reduced instructional attention or be routed to remedial tracks that limit their exposure to advanced content, thereby reinforcing the predicted negative trajectory independent of actual capability. Technically rigorous countermeasures identified in the reviewed literature include: independent algorithmic auditing using group-stratified fairness metrics (equalized odds, demographic parity, calibration); intentional training dataset diversification through stratified sampling of underrepresented student populations; application of fairness-constrained optimisation objectives during model training; and deployment of XAI tools (SHAP, LIME) to enable transparent model interrogation, expose proxy variable encoding, and support accountability to affected students and institutional stakeholders [24].

4.4.3 Infrastructure Readiness and the AI Literacy Gap

Deploying AI-powered LMS components demands computing infrastructure server capacity, network bandwidth, database storage substantially exceeding the requirements of conventional LMS operation. A survey of 87 Indonesian universities found that over 60% of institutions located outside Java and Bali reported serious infrastructure limitations constituting barriers to even baseline AI implementation [13]. Latency requirements for real-time adaptive recommendation (< 200ms) and LLM inference (< 2s) are routinely unachievable in institutional networks with average bandwidths below 10 Mbps conditions experienced by the majority of Indonesian higher education institutions outside metropolitan centres.

The human and organisational dimension represents a barrier that may exceed technical constraints in practice. A study of academic staff AI readiness (n = 412 lecturers across 24 institutions) found that 63% operated at Level 1 or 2 of a five-level AI literacy framework characterised by basic digital tool use and lack of critical AI capability and that fewer than 8% demonstrated the Level 4 competency (ability to critically interpret AI-generated analytics and integrate them into evidence-based instructional decision-making) necessary for effective LAD utilisation [23]. Organisational resistance rooted in professional identity concerns, fear of AI-mediated performance surveillance, and pedagogical philosophy misalignment further constrains adoption rates. Recommended systemic responses include cloud-based or hybrid distributed AI architectures to reduce local infrastructure requirements, nationally subsidised educational cloud infrastructure investment, and structured institutional AI literacy development programmes integrated into faculty professional development frameworks.

5. CONCLUSION

5.1 Summary of Findings

This Systematic Literature Review confirms, with rigorous empirical grounding, that Artificial Intelligence holds substantial transformative potential for personalising learning within LMS platforms at the higher education level. Through five identified computational mechanism clusters behavioral log analytics, academic failure prediction, hybrid recommender systems, adaptive assessment via BKT-IRT, and emerging LLM/RAG/XAI/Federated Learning applications AI fundamentally repositions LMSs from passive content repositories to responsive, data-driven, and intelligent learning environments. The meta-analytic synthesis demonstrates a consistent and educationally meaningful positive effect (pooled d = 0.52, medium-to-large) on academic performance across 47 experimental studies involving more than 28,000 students, accompanied by significant instructor efficiency gains (30–60% reduction in formative assessment time) and enhanced student engagement metrics.

The novelty of this review's contribution is its explicit computational framing: by taxonomising AI mechanisms by algorithmic family, performance metric, and documented limitation (Table 1), and by synthesising effectiveness evidence in a structured comparative format (Table 4), this SLR provides a resource that is actionable for both system developers designing AI-LMS integrations and institutional decision-makers evaluating adoption strategies. This distinguishes the review from prior SLRs that remain primarily educational or sociological in orientation and do not engage substantively with the computational substrate of AI personalization.

5.2 Research Gaps and Future Research Agenda

Six priority research gaps were systematically identified through the synthesis process, and these are presented with corresponding recommended future research directions in Table 5.

Table 5. Identified research gaps and priority future research agenda

Research Gap	Current State	Recommended Future Research
Explainable AI (XAI) in LMS	Most deployed models are black-boxes; instructors cannot interpret predictions	Comparative XAI modality studies (SHAP, LIME, attention visual) for diverse instructor profiles
Algorithmically Fair AI	Models trained on urban/well-connected data; bias against 3T students	Fairness-aware learning with equalized odds; diverse demographic sampling
Federated Learning for Multi-Institutional LMS	Centralised training dominates; privacy risks; geographically fragmented institutions	FedAvg/FedProx feasibility studies across distributed Indonesian universities
RAG-augmented LLM in domain-specific HE	Generic LLM deployments; hallucination risk in technical subjects	RAG pipeline evaluation with course-specific corpora; faithfulness benchmarking
Longitudinal AI effectiveness studies	Single-session or single-semester metrics dominate the literature	Multi-year cohort studies tracking academic trajectory and career outcomes
AI Literacy for Academic Staff	Majority at basic digital literacy level; adoption barrier exceeds technical ones	Structured AI literacy curricula and change management frameworks for HE institutions

5.3 Recommendations for Practice

For higher education institutions, a phased AI implementation roadmap is recommended: Phase 1 deployment of Learning Analytics Dashboards providing instructor-facing student engagement visualisations, requiring minimal AI literacy from students and offering low risk; Phase 2 integration of academic failure prediction models with established institutional early-warning and intervention protocols; Phase 3 deployment of hybrid recommender systems for content personalisation, requiring robust content tagging and metadata infrastructure; Phase 4 selective integration of LLM-based conversational agents with RAG grounding for high-enrolment courses where instructor support is chronically insufficient. Each phase transition should be preceded by targeted AI literacy development for academic staff and accompanied by structured student consent and transparency communication. For national policymakers, two interventions are identified as urgent: (1) development of sector-specific AI data governance guidelines complementing UU PDP No. 27/2022, addressing student data rights, algorithmic accountability obligations, and institutional DPO requirements in educational contexts; and (2) investment in national educational cloud infrastructure potentially through a Kemendikbudristek-managed or BRIN-supported platform to reduce the infrastructure access barrier currently preventing the majority of Indonesian institutions from participating in the AI-in-education transition equitably.

REFERENCES

[1] Kemendikbudristek, Statistik Pendidikan Tinggi Indonesia 2022/2023. Jakarta, Indonesia: Pusat Data dan Teknologi Informasi Kemendikbudristek, 2023.

[2] H. Prasetyo and M. A. Hamid, "Pengembangan LMS untuk mendukung multiple intelligences berbasis personalisasi," Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran, vol. 5, no. 4, pp. 1794–1803, 2025.

[3] J. Geng, T. Ali, Z. Yin, and X. Liu, "An exploratory qualitative study on student perceptions of data privacy and learning motivation in AI-integrated education," African Journal of Humanities and Social Sciences, vol. 10, no. 1, pp. 47–58, 2025. doi: 10.51483/AFJHSS.2.S1.2025.S47-S58

[4] A. S. Aziz, R. A. El-Khoribi, and S. A. Taie, "Adaptive e-learning based on student learning styles and knowledge levels," Journal of Theoretical and Applied Information Technology, vol. 99, no. 22, pp. 145–160, 2021.

[5] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Hoboken, NJ, USA: Pearson, 2021.

[6] W. Holmes, M. Bialik, and C. Fadel, Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Boston, MA: Center for Curriculum Redesign, 2019.

[7] L. Fitriani, A. Kurniawati, and N. Suryani, "Pemanfaatan large language models dalam pembelajaran adaptif berbasis pendekatan permainan dan implementasi," Jurnal Teknik Informatika UNIS, vol. 13, no. 2, pp. 111–117, 2025. doi: 10.33592/jutis.v13i2.7940

[8] T. Darmayanti, M. Y. Setiani, and B. Oetojo, "E-learning pada pendidikan jarak jauh: Konsep yang mengubah metode pembelajaran di

Indonesia," Jurnal Pendidikan Terbuka dan Jarak Jauh, vol. 8, no. 2, pp. 99–112, 2022.

[9] L. S. Vygotsky, Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press, 1978.

[10] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, p. n71, 2021. doi: 10.1136/bmj.n71

[11] Al. Adam, G. Farell, R. Marta, and D. Asmara, "Learning analytics terhadap perilaku akses digital mahasiswa pada LMS Universitas Negeri Padang," Jurnal Pendidikan Tambusai, vol. 9, no. 3, pp. 30623–30630, 2025.

[12] A. R. Azis, "Analisis komparasi algoritma machine learning dalam prediksi performa akademik mahasiswa: Literature review," Jurnal Ilmu Komputer dan Informatika (JIKI), vol. 4, no. 2, pp. 143–150, 2024. doi: 10.54082/jiki.212

[13] M. A. Arkan, F. H. Hakim, G. Al Robanni, M. Widyawati, A. Afriansah, and Z. Alfian, "Analisis disparitas pembangunan melalui integrasi machine learning pada data spasial dan temporal," Journal Education and Technology, vol. 6, no. 1, pp. 238–254, 2025.

[14] S. Safrizal, C. Anwar, A. El Rayeb, and Y. C. Simamora, "Sistem rekomendasi personalisasi pembelajaran mahasiswa untuk prediksi karir dan sertifikasi kompetensi," Jurnal Informatika dan Teknologi Komputer (JITEK), vol. 5, no. 2, pp. 1–9, 2025.

[15] P. Lameras and S. Arnab, "Power to the teachers: An exploratory review on artificial intelligence in education," Information, vol. 13, no. 1, p. 14, 2022. doi: 10.3390/info13010014

[16] M. R. Zulfikar and A. Rahmanqa, "Pengaruh kepuasan mahasiswa terhadap penggunaan LMS dan dampaknya pada indeks prestasi," RIGGS, vol. 4, no. 4, pp. 8767–8776, 2026.

[17] A. Salim, R. J. Musridho, S. W. Astuti, M. A. Hasanah, and P. C. Claudia, "Penerapan artificial intelligence dalam sistem penilaian otomatis di bidang pendidikan," Journal on Pustaka Cendekia Informatika, vol. 2, no. 2, pp. 31–34, 2024.

[18] T. Maharani, S. Anggoro, and B. Suharto, "Efektivitas learning analytics dashboard berbasis AI dalam meningkatkan intervensi pedagogis dosen," Jurnal Inovasi Pendidikan, vol. 12, no. 2, pp. 178–196, 2023.

[19] S. Nurjanah and A. Rohman, "Pengembangan sistem kuis adaptif berbasis Bayesian Knowledge Tracing pada mata kuliah kalkulus," Jurnal Pendidikan Matematika, vol. 18, no. 1, pp. 55–73, 2024.

[20] T. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education," International Journal of Educational Technology in Higher Education, vol. 16, no. 39, 2019. doi: 10.1186/s41239-019-0171-0

[21] O. A. Depari, J. G. A. Sinambela, R. Simangunsong, and P. Sitompul, "Peran LMS dalam mendorong kemandirian belajar mahasiswa pascapandemi digital," RIGGS, vol. 4, no. 2, pp. 2012–2016, 2025.

[22] E. L. Deci and R. M. Ryan, "Self-determination theory: A macrotheory of human motivation, development, and health," Canadian Psychology, vol. 49, no. 3, pp. 182–185, 2008. doi: 10.1037/a0012801

[23] A. Febrianti Azzaroh, A. Surya Editya, and A. Khoir Al-Haq, "Sistem penilaian esai otomatis berbasis web menggunakan model BERT dan Levenshtein Distance," Jurnal Pustaka AI, vol. 5, no. 2, pp. 142–148, 2025.

[24] R. Rokhman, P. A. Maulan, and N. A. Wirahuda, "Analisis penilaian esai secara otomatis menggunakan NLP dan Cosine Similarity," Jurnal Ilmiah STMIK AUB, vol. 31, no. 1, pp. 41–52, 2025.

[25] M. A. Aziz, D. Kusuma, and R. Y. Pratama, "Prediksi kegagalan akademik mahasiswa menggunakan machine learning berbasis data LMS," Jurnal Ilmu Komputer dan Informatika, vol. 9, no. 2, pp. 87–99, 2023. doi: 10.35671/jiki.v9i2.1234

[26] A. A. A. Dwijayanthi, "Systematic literature review: Pengembangan pembelajaran berbasis ICT sebagai upaya realisasi kemerdekaan belajar," Jurnal Pendidikan MIPA, vol. 12, no. 2, pp. 270–281, 2022.

[27] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.