Pattern Analysis of Global AI Adoption and Digital Literacy Development in Science Education: Basis for a Recommender System Framework

Honelly Mae S. Cascolan
College of Education-Pangasinan State University, Asingan Campus, Pangasinan, Philippines

***Abstract*—The rapid growth of artificial intelligence (AI) technologies has transformed educational environments and created new opportunities for enhancing digital literacy in science education. However, variations in technological readiness, infrastructure, educational development, and AI adoption continue to influence the effectiveness of digital transformation initiatives across regions. This study analyzed global AI adoption patterns and digital literacy development using the Global AI in Education Dataset (2015–2026) comprising 1,360 observations from multiple countries and regions. A data-driven Intelligent Recommender System Framework integrating Educational Data Mining, Machine Learning, and Explainable Artificial Intelligence techniques was developed to identify digital literacy profiles and generate evidence-based recommendations. Digital Literacy Scores were computed using AI adoption, educational readiness, and technology access indicators and were subsequently classified into low, moderate, and high digital literacy profiles. Random Forest, Naïve Bayes, and Hybrid Voting models were utilized for pattern discovery and profile classification. Results revealed a continuous increase in digital literacy development from 2015 to 2026, accompanied by substantial growth in student AI usage, teacher AI usage, and school AI adoption. Random Forest achieved the highest classification performance with an accuracy of 94.49%. Feature importance analysis identified student AI usage, urban AI usage, internet penetration, education index, school AI adoption, and teacher AI usage as the most influential determinants of digital literacy development. Cluster analysis further revealed distinct educational environments characterized by varying levels of infrastructure readiness and AI engagement. Based on these findings, the proposed Intelligent Recommender System Framework generated targeted recommendations related to infrastructure expansion, teacher professional development, curriculum enhancement, AI literacy integration, institutional AI adoption, and digital inclusion initiatives. The framework provides a scalable and explainable decision-support mechanism that can assist educators, administrators, curriculum developers, and policymakers in promoting digital literacy and advancing AI-enabled science education.**

Keywords*—Artificial intelligence adoption; digital literacy; Educational Data Mining; explainable artificial intelligence; Intelligent Recommender System

I. Introduction

The rapid advancement of artificial intelligence (AI) has transformed educational environments worldwide, creating new opportunities for teaching, learning, assessment, and educational decision-making. AI-powered tools such as intelligent tutoring systems, chatbots, learning analytics platforms, and generative AI applications have become increasingly integrated into educational processes, influencing how learners access information and develop digital competencies [1]. At the same time, the global education sector continues to emphasize the development of digital literacy and technology-related competencies as essential skills for sustainable development and lifelong learning in the twenty-first century [2].

Prior studies have highlighted the significant role of information and communication technologies (ICT) in enhancing student engagement, collaboration, communication, and learning outcomes across various educational settings [3], [4]. The integration of digital technologies has also expanded opportunities for personalized learning experiences and flexible instructional approaches [5], [6]. As educational technologies continue to evolve, teachers are increasingly expected to possess competencies that enable them to effectively integrate technological tools into their instructional practices while supporting meaningful student learning experiences [7], [8].

The theoretical foundation for technology integration in education is strongly influenced by the concepts of pedagogical content knowledge and technological pedagogical content knowledge. These frameworks emphasize the importance of combining subject matter expertise, pedagogical strategies, and technological knowledge to create effective learning environments [9]–[13]. Within science education, the ability of teachers and learners to utilize digital technologies effectively has become increasingly important as scientific inquiry, experimentation, and knowledge construction are increasingly supported by digital resources and AI-enabled tools.

Recent developments in generative AI and large language models have further accelerated the digital transformation of education. These technologies offer opportunities for personalized instruction, adaptive learning support, automated feedback, and intelligent content generation, potentially improving learning efficiency and accessibility [14], [15]. Higher education students have demonstrated growing acceptance and utilization of generative AI technologies for academic tasks, knowledge acquisition, and problem-solving activities [16]. However, the increasing reliance on AI tools also raises concerns regarding cognitive offloading, critical thinking development, and the need for sufficient digital literacy skills to ensure effective and responsible technology use [17], [18].

Understanding patterns of AI adoption and their relationship with digital literacy development has therefore become a critical area of educational research. While previous studies have examined technology acceptance and educational technology integration, limited research has explored how large-scale AI adoption trends across educational systems relate to variations in digital literacy outcomes within science education contexts. Identifying such patterns can provide valuable insights for educators, policymakers, and institutions seeking to maximize the benefits of AI while minimizing potential challenges.

The emergence of machine learning techniques has provided researchers with powerful analytical tools for identifying hidden patterns, relationships, and predictive factors within large educational datasets. Machine learning approaches have demonstrated strong performance in classification, prediction, and decision-support applications across various domains [19]–[21]. Among these techniques, Random Forest models have been widely recognized for their robustness, interpretability, and predictive accuracy, particularly when enhanced through systematic hyperparameter optimization methods such as Grid Search [22]–[24]. Ensemble learning approaches and voting-based classifiers further improve predictive performance by combining multiple algorithms to produce more reliable outcomes [25], [26]. In addition, probabilistic classification techniques such as Naïve Bayes continue to provide efficient and effective solutions for Educational Data Mining applications [27].

As machine learning models become increasingly utilized in educational analytics, explainable artificial intelligence has gained importance in supporting transparent and interpretable decision-making processes. Techniques such as SHAP and LIME enable researchers to identify influential variables and understand the factors contributing to model predictions, thereby improving the practical value of educational analytics findings [28]. These capabilities are particularly relevant when examining complex interactions among AI adoption indicators, educational characteristics, and digital literacy outcomes.

To investigate these relationships, this study utilizes the Global AI in Education Dataset (2015–2026), which provides comprehensive information on AI adoption, educational infrastructure, technology integration, and related educational indicators across different countries and educational contexts [29]. The dataset offers a unique opportunity to explore global trends and identify significant factors associated with digital literacy development in science education through data-driven analysis.

Beyond pattern identification and predictive modeling, the findings of educational analytics can be translated into actionable recommendations through recommender systems. Recommender systems are intelligent decision-support technologies designed to assist users by suggesting relevant resources, actions, or interventions based on observed patterns and user characteristics [30], [31]. Educational recommender systems have been increasingly applied to support personalized learning, resource allocation, learner guidance, and instructional decision-making [32]–[35]. These systems combine techniques from machine learning, knowledge representation, semantic technologies, and decision-support methodologies to provide context-aware recommendations that enhance educational effectiveness [36], [37].

The effectiveness of recommender systems is further strengthened through the incorporation of user profiling, preference modeling, and adaptive recommendation strategies [38], [39]. Fuzzy logic approaches have also demonstrated their ability to address uncertainty and complexity in recommendation processes, particularly in educational environments where learner needs and contextual factors vary significantly [40], [41]. Advances in collaborative filtering, representation learning, and preference elicitation techniques have further improved the capacity of recommender systems to generate meaningful and personalized recommendations [42], [43]. Similarly, semantic recommendation approaches have enabled the delivery of contextually relevant educational resources and interventions within dynamic learning ecosystems [44].

Despite the growing body of research on AI adoption, educational analytics, digital literacy, and recommender systems, limited empirical studies have specifically examined how global AI adoption patterns influence digital literacy development within science education. Science education increasingly depends on AI-supported simulations, virtual laboratories, scientific visualization tools, intelligent tutoring systems, and data-driven inquiry activities that require both teachers and students to possess adequate digital literacy competencies. Understanding the determinants of digital literacy therefore provides valuable evidence for designing targeted interventions that strengthen AI integration in science teaching, scientific inquiry, curriculum innovation, and technology-enhanced learning.

Therefore, this study aims to analyze global AI adoption patterns and their relationship with digital literacy development in science education using machine learning techniques. Specifically, it seeks to identify significant predictors of digital literacy, classify digital literacy levels based on AI adoption indicators, explain the contribution of influential variables, and develop a recommender system framework that can support data-driven educational planning and decision-making. The findings are expected to contribute to the advancement of Educational Data Mining, AI-enhanced science education, digital literacy research, and intelligent educational recommendation systems.

II. Methodology

A. Research Design

This study employed a data-driven recommender system development approach integrating Educational Data Mining (EDM), Machine Learning (ML), and Explainable Artificial Intelligence (XAI) techniques. The research was designed to identify global patterns of artificial intelligence (AI) adoption and digital literacy development in science education and subsequently utilize these patterns as the foundation for a recommendation framework that can support educational planning, curriculum enhancement, and digital literacy interventions. The study followed a six-stage framework consisting of data acquisition, data preprocessing, digital literacy profiling, pattern discovery, recommendation generation, and recommender system framework development.

Unlike conventional predictive studies that focus solely on classification accuracy, the primary objective of this research was to transform discovered patterns into actionable recommendations that may assist educators, institutions, and policymakers in improving science education in the digital age.

B. Research Framework

The proposed framework consists of six major phases:

- Dataset Acquisition
- Data Preparation and Transformation
- Digital Literacy Profiling
- Pattern Analysis and Knowledge Discovery
- Recommendation Generation
- Recommender System Framework Development

The framework transforms raw educational and AI adoption indicators into intelligent recommendations through a systematic knowledge discovery process. The output of the framework is a recommendation engine capable of identifying priority interventions based on the digital literacy profile and AI adoption characteristics of educational environments.

C. Data Source

The study utilized the Global AI in Education Dataset (2015–2026) developed by Hussain [29]. The dataset consists of 1,360 observations representing various indicators of artificial intelligence (AI) adoption and digital technology integration in educational environments across multiple countries and regions. It includes variables related to student AI usage, teacher AI usage, school AI adoption, average daily AI usage, internet penetration, education index, urban and rural AI usage, government AI policies, AI integration in the curriculum, gender gap in AI usage, geographic region, country, year, month, and the most frequently used AI tools. These variables collectively provide a comprehensive representation of educational technology adoption, digital infrastructure, technological readiness, educational development, and access to AI-enabled learning resources. As such, the dataset serves as a valuable source of information for examining global patterns of AI adoption and their relationship with digital literacy development in science education.

D. Data Preparation and Feature Engineering

Data preprocessing was conducted to ensure data quality and suitability for machine learning and recommendation analysis.

The following procedures were performed:

1) Data cleaning. Missing values, duplicate records, and inconsistent entries were identified and treated appropriately.

2) Data transformation. Categorical variables such as region, government AI policy, top AI tool, and AI curriculum integration were transformed using one-hot encoding.

3) Normalization. Numerical variables were normalized using Min-Max scaling to ensure comparable measurement ranges among indicators. To eliminate scale differences among variables, Min-Max normalization was applied using:

$$x'i = (xi - xmin)/(xmax - xmin) \quad (1)$$

where, xi denotes the original value, $xmin$ and $xmax$ represent the minimum and maximum values, respectively, and $x'I$ is the normalized value. Categorical attributes were transformed using One-Hot Encoding.

4) Feature construction. Additional indicators representing AI readiness and digital engagement were generated from existing variables to support recommendation generation.

E. Development of the Digital Literacy Profile

Since the dataset does not contain a direct measure of digital literacy, a Digital Literacy Profile (DLP) was developed using indicators associated with technological engagement, educational readiness, and artificial intelligence (AI) adoption. The profile was designed to represent the level of digital literacy development within educational environments by integrating variables that reflect access to technology, AI utilization, and educational capacity. Specifically, the profile was derived from student AI usage, teacher AI usage, school AI adoption, internet penetration, education index, urban AI usage, rural AI usage, and average daily AI usage. These indicators were selected because they collectively capture the technological, institutional, and infrastructural conditions that influence digital literacy development.

The selected indicators were chosen based on their representation of technological access, educational readiness, institutional AI adoption, and learner engagement, which are widely recognized as important dimensions influencing digital literacy in educational environments. Since the Global AI in Education Dataset does not contain a standardized or validated digital literacy instrument, the Digital Literacy Score (DLS) was operationalized as a composite proxy measure derived from these indicators. Equal weighting was adopted to ensure that each indicator contributed uniformly to the overall score and to avoid introducing subjective weighting schemes unsupported by empirical evidence. Consequently, the DLS should be interpreted as an analytical index for educational profiling rather than a psychometrically validated measure of individual digital literacy.

To ensure comparability among variables measured on different scales, all indicators were normalized using Min-Max normalization. A composite Digital Literacy Score (DLS) was then calculated by computing the average of the normalized indicator values. The Digital Literacy Score was determined using the following equation:

$$DLS = (\Sigma X_i) / n \quad (2)$$

where, DLS represents the Digital Literacy Score, X_i denotes the normalized value of each indicator, and n represents the total number of indicators included in the profile.

The resulting Digital Literacy Scores were subsequently categorized into three levels: Low Digital Literacy, Moderate Digital Literacy, and High Digital Literacy. These categories were used to establish digital literacy profiles that characterize the technological readiness and digital capability of different educational environments. The identified profiles served as the foundation for cluster analysis, pattern identification, and the development of the Intelligent Recommender System Framework. By classifying educational environments according to their digital literacy levels, the study was able to generate targeted recommendations aimed at enhancing AI adoption, technology integration, and digital literacy development in science education.

F. Pattern Discovery and Analytical Modeling

To identify significant relationships between AI adoption indicators and digital literacy development, analytical modeling techniques were employed to discover hidden patterns and influential factors within the dataset. The dataset was randomly partitioned using stratified sampling into 80% training data and 20% testing data to preserve the proportional distribution of digital literacy profiles across all classes. Model development was conducted exclusively on the training dataset, while predictive performance was evaluated using the unseen testing dataset. The objective of this phase was not solely to perform classification but to uncover meaningful relationships that could support digital literacy profiling and recommendation generation.

Random Forest analysis was utilized to examine nonlinear relationships among variables and determine the relative importance of AI adoption, educational readiness, and technological infrastructure indicators. Random Forest hyperparameters were optimized using Grid Search Cross-Validation. The final model employed 200 decision trees (n_estimators = 200), a maximum tree depth of 20, a minimum samples split of 2, a minimum samples leaf of 1, Gini impurity as the split criterion, bootstrap sampling enabled, and random_state = 42 to ensure reproducibility. The method is particularly effective for identifying influential predictors because it evaluates multiple decision trees and aggregates their results to produce robust estimates of variable importance. Through this process, the study identified the factors most strongly associated with digital literacy development across educational environments.

Five-fold cross-validation was employed during model training to assess performance stability and reduce the risk of overfitting. Mean Accuracy, Precision, Recall, and F1-score across validation folds were calculated before evaluating the final model on the testing dataset.

In addition, the Naïve Bayes algorithm was employed as a probabilistic analytical technique to examine the likelihood of digital literacy profile membership based on observed AI adoption characteristics. The algorithm assumes conditional independence among predictor variables and provides a computationally efficient approach for identifying relationships between educational indicators and digital literacy outcomes.

To enhance analytical reliability and support pattern validation, a hybrid ensemble approach combining Random Forest and Naïve Bayes techniques was implemented. The hybrid model integrated the strengths of both algorithms by aggregating their analytical outputs, thereby providing a more comprehensive assessment of relationships among variables. Rather than serving as a purely predictive model, the ensemble approach was used to validate discovered patterns and identify the most influential indicators contributing to variations in digital literacy development.

The results of these analytical procedures were subsequently integrated with clustering and association rule mining outputs to support the development of digital literacy profiles and the Intelligent Recommender System Framework. Consequently, machine learning techniques in this study functioned primarily as knowledge discovery tools that facilitated the identification of significant educational patterns and informed the generation of evidence-based recommendations for science education stakeholders.

G. Feature Importance and Explainable AI Analysis

A rule-based recommendation mechanism was developed using the outputs of machine learning classification and SHAP feature importance analysis. Recommendation rules were generated according to identified digital literacy profiles and influential AI adoption indicators.

Algorithm 1 presents the operational process of the proposed Intelligent Recommender System Framework. The algorithm begins by loading the educational dataset and initializing the digital literacy profiles, recommendation rules, and recommendation repository. After data preprocessing, including data cleaning, normalization, and transformation of categorical variables, Digital Literacy Scores (DLS) are computed and used to construct digital literacy profiles. For each educational environment, the framework analyzes AI adoption indicators and applies clustering, machine learning, SHAP analysis, and association rule mining to identify significant patterns and influential factors affecting digital literacy development. Based on the identified conditions, the system generates targeted recommendations related to internet infrastructure, teacher training, curriculum enhancement, institutional AI adoption, and rural digital literacy initiatives. These recommendations are then personalized according to the specific needs of each educational environment and stored within the recommendation repository. This algorithm transforms educational data into evidence-based recommendations that support digital literacy enhancement, AI integration, and informed decision-making among educators, administrators, curriculum developers, and policymakers.

Algorithm 1. Intelligent Recommender System Framework

```
Initialize Educational Dataset, Digital Literacy Profiles,
Recommendation Rules, and Recommendation Repository
Load AI Adoption Indicators, Educational Readiness Indicators,
and Digital Literacy Variables
Perform Data Preprocessing
    Clean Data
    Normalize Variables
    Transform Categorical Features
Generate Digital Literacy Scores (DLS)
Construct Digital Literacy Profiles
While (Educational Environment Exists) do
```

```
For (each Educational Environment) do
  Analyze AI Adoption Indicators
  Perform Cluster-Based Profile Identification
  Generate Machine Learning Outputs
  Compute SHAP Feature Importance Rankings
  Discover Patterns using Association Rule Mining
  If (Internet Penetration is Low AND Digital Literacy Profile
is Low) then
    Recommend:
      Expansion of Internet Infrastructure
      Community Connectivity Programs
      Technology Access Initiatives
  End
  If (Teacher AI Usage is Low) then
    Recommend:
      AI-Focused Teacher Training
      Professional Development Workshops
      AI Integration Certification Programs

  End
  If (AI Curriculum Integration is Absent) then
    Recommend:
      Curriculum Redesign Initiatives
      AI Literacy Modules
      Science-AI Integration Activities
  End
  If (School AI Adoption is Low) then
    Recommend:
      Institutional AI Adoption Programs
      Educational Technology Investments
      AI Implementation Roadmaps
  End
  If (Urban–Rural AI Usage Gap is High) then
    Recommend:
      Rural Digital Literacy Programs
      Technology Inclusion Projects
      Community-Based AI Learning Initiatives
  End
  Generate Personalized Educational Recommendations
  Store Recommendations in Recommendation Repository
 End
End
```

When multiple recommendation conditions were simultaneously satisfied, the recommendation engine ranked intervention priorities according to SHAP feature importance scores. Recommendations corresponding to higher-impact predictors were presented first, while additional applicable interventions were retained to provide comprehensive decision support.

H. Development of the Recommender System Framework

The proposed recommender system framework was developed using a hybrid knowledge-based recommendation approach. The framework consists of four interconnected layers:

1) Data layer. Responsible for collecting educational, technological, and AI adoption indicators.

2) Analytics layer. Performs preprocessing, classification, clustering, feature importance analysis, and pattern discovery.

3) Recommendation engine layer. Generates recommendations using:

- Digital Literacy Profiles
- Machine Learning Predictions
- SHAP Feature Importance Rankings
- Recommendation Rules

4) Decision support layer. Provides personalized recommendations for:

- Science Teachers
- School Administrators
- Curriculum Developers
- Educational Policymakers

The framework is designed to support evidence-based decision-making and strategic planning for AI-enhanced science education.

I. Evaluation of the Intelligent Recommender System Framework

The effectiveness of the proposed Intelligent Recommender System Framework was evaluated using three complementary dimensions: analytical performance, recommendation relevance, and explainability. These evaluation criteria were selected to ensure that the framework not only produced reliable analytical results but also generated meaningful and interpretable recommendations capable of supporting educational decision-making.

1) Analytical performance. The analytical performance of the framework was assessed using the outputs of the machine learning models employed in the study. Standard classification metrics, including Accuracy, Precision, Recall, and F1-Score, were utilized to evaluate the capability of the analytical models to correctly identify digital literacy profiles. These metrics provided quantitative measures of the reliability and consistency of the discovered patterns and profile classifications.

2) Recommendation relevance. Recommendation relevance was evaluated analytically by examining the consistency between generated recommendations and the influential predictors identified through SHAP feature importance analysis, clustering results, and association rule mining. Recommendations were considered appropriate when they directly addressed the dominant factors associated with each digital literacy profile. Since this study focused primarily on framework development, expert validation, user evaluation, and field deployment were beyond its scope. Future studies are recommended to validate the proposed recommendations through Delphi expert evaluation, practitioner assessment, and implementation in actual educational settings.

3) Explainability and transparency. The explainability of the framework was assessed through the interpretability of the recommendation rules and the transparency of the analytical procedures used to generate recommendations. SHAP feature importance analysis was employed to provide clear explanations regarding the contribution of each predictor variable to digital literacy development. In addition, the rule-based recommendation mechanism enabled stakeholders to trace generated recommendations back to specific educational conditions and influential factors. This transparency enhanced stakeholder trust and facilitated the practical application of recommendations in educational planning and decision-making.

The integration of explainable artificial intelligence within the framework ensured that generated recommendations remained understandable, transparent, and justifiable. By combining predictive analytics, pattern discovery, and explainable recommendation mechanisms, the proposed framework provides a robust decision-support tool capable of assisting policymakers, educational institutions, curriculum developers, and science educators in designing targeted interventions for improving digital literacy and AI adoption in science education.

J. Proposed Output

The final output of this study is an Intelligent Recommender System Framework designed to support digital literacy development and artificial intelligence integration in science education through data-driven and evidence-based recommendations. The framework serves as a strategic decision-support mechanism that transforms educational data into actionable insights by integrating digital literacy profiling, pattern analysis, machine learning, explainable artificial intelligence, and recommendation generation.

Specifically, the framework is capable of profiling educational environments according to their levels of digital literacy development, identifying significant patterns of AI adoption, and detecting critical factors that influence technological readiness and educational innovation. Through cluster analysis, association rule mining, and SHAP-based knowledge discovery, the framework determines priority intervention areas and generates targeted recommendations that address identified educational needs and challenges.

The optimal number of clusters was determined using the Silhouette Coefficient, Davies–Bouldin Index, and Calinski–Harabasz Index. Three clusters produced the highest clustering quality and were therefore selected for subsequent pattern analysis.

The proposed framework further supports the recommendation of science education enhancement strategies, including teacher professional development initiatives, curriculum improvement programs, digital infrastructure expansion, AI literacy integration, and institutional AI adoption plans. By linking identified digital literacy profiles with context-specific intervention strategies, the framework enables educational stakeholders to implement more effective and responsive digital transformation initiatives.

In addition, the framework assists educational institutions, curriculum developers, school administrators, and policymakers in planning and implementing AI integration strategies based on empirical evidence. The generated recommendations provide guidance for resource allocation, technology investments, instructional innovation, and policy formulation aimed at strengthening digital literacy and AI readiness.

Ultimately, the Intelligent Recommender System Framework serves as a strategic decision-support model for improving digital literacy development and advancing AI-enabled science education at institutional, regional, and national levels. By transforming analytical findings into practical recommendations, the framework contributes to evidence-based educational planning and supports the sustainable integration of artificial intelligence within science education systems.

III. Results and Discussion

A. Descriptive Results of AI Adoption and Digital Literacy Indicators

The descriptive results show that global AI adoption in education remains moderate but uneven. Student AI usage recorded a mean of 30.99%, while teacher AI usage recorded a mean of 29.91%. School AI adoption averaged 29.41%, indicating that institutional AI integration was still developing across the dataset. Internet penetration showed a relatively high mean of 74.70%, while the education index averaged 0.776.

TABLE I. Descriptive Statistics of Main Indicators

Indicator	Mean	SD	Min	Max
Student AI usage (%)	30.99	16.7	0	65
Teacher AI usage (%)	29.91	16.46	0	60
School AI adoption (%)	29.41	16.6	0	62
Internet penetration (%)	74.7	20.69	42	96
Education index	0.776	0.157	0.52	0.94
Urban AI usage (%)	33.56	16.76	0	70
Rural AI usage (%)	21.6	16.19	0	59
Average daily AI usage	32.82	15.56	5	60
Gender gap in AI usage (%)	5.41	2.88	1	10

The results presented in Table I indicate that AI use in education is expanding, but adoption remains uneven between urban and rural areas. Urban AI usage was higher than rural AI usage, suggesting a continuing digital access gap. This finding supports the need for a recommender system that does not only identify digital literacy levels but also recommends targeted interventions based on specific limiting factors.

B. Digital Literacy Profile Distribution

The computed Digital Literacy Score had a mean of 0.502, indicating a generally moderate level of digital literacy development. Using fixed classification ranges, 790 observations were categorized as moderate, 302 as high, and 268 as low. For model training and balanced classification, the profiles were also grouped into three balanced levels.

TABLE II. DIGITAL LITERACY PROFILE DISTRIBUTION

Profile	Fixed Range Count	Balanced Modeling Count
Low	268	453
Moderate	790	454
High	302	453

The dominance of the moderate category, as presented in Table II, implies that many educational environments have already developed some degree of digital readiness but still require strategic interventions to move toward higher AI-enabled digital literacy. This supports the importance of the proposed recommender system, since moderate-performing environments may benefit from differentiated recommendations rather than general technology adoption policies.

C. Regional Pattern of Digital Literacy Development

Regional analysis showed clear differences in digital literacy development. Oceania recorded the highest mean Digital Literacy Score, followed by Europe and North America. Africa and Asia recorded the lowest mean scores.

TABLE III. MEAN DIGITAL LITERACY SCORE BY REGION

Region	Mean DLS	SD	Count
Oceania	0.601	0.165	136
Europe	0.587	0.166	272
North America	0.584	0.17	272
South America	0.472	0.167	136
Asia	0.418	0.171	408
Africa	0.351	0.165	136

The results presented in Table III suggest that digital literacy development is strongly influenced by regional differences in infrastructure, educational readiness, and AI adoption capacity. The low mean scores in Africa and Asia indicate the need for stronger recommendations related to infrastructure development, teacher training, curriculum integration, and inclusive AI access.

D. Temporal Pattern of AI Adoption and Digital Literacy

The yearly trend showed a continuous increase in the Digital Literacy Score from 2015 to 2026. The mean score increased from 0.254 in 2015 to 0.799 in 2026. Student AI usage, teacher AI usage, and school AI adoption also increased steadily across the same period.

TABLE IV. DIGITAL LITERACY AND AI ADOPTION TREND

Year	Mean DLS	Student AI Usage	Teacher AI Usage	School AI Adoption
2015	0.254	4.98	4.04	3.38
2016	0.304	10.41	9.23	8.63
2017	0.347	15.4	14.01	13.47
2018	0.378	19.32	19.46	18.36
2019	0.444	25.47	23.74	23.63
2020	0.496	30.03	29.33	28.45
2021	0.544	35.3	34.26	33.52
2022	0.584	39.95	38.58	38.48
2023	0.639	44.84	44.11	43.54
2024	0.69	49.9	49.08	48.68
2025	0.741	55.33	53.45	53.43
2026	0.799	60.75	59	59.25

The increasing trend presented in Table IV confirms that AI adoption has become progressively embedded in educational environments. However, the variation across regions and indicators shows that growth is not uniform. Therefore, the proposed recommender system is necessary to provide context-sensitive interventions instead of a single universal strategy.

E. Machine Learning Performance for Digital Literacy Profiling

Random Forest, Naïve Bayes, and Hybrid Voting classifiers were used to support digital literacy profile classification and recommendation generation. Random Forest achieved the highest performance with an accuracy of 94.49% and an F1-score of 94.50%. The Hybrid Voting model achieved 90.81% accuracy, while Naïve Bayes achieved 82.35%.

TABLE V. MODEL PERFORMANCE

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.9449	0.9451	0.9449	0.945
Naïve Bayes	0.8235	0.824	0.8238	0.8232
Hybrid Voting	0.9081	0.9112	0.9081	0.909

Table V shows the model performance of selected machine learning algorithms. The Random Forest model performed best because it effectively captured nonlinear relationships among AI adoption indicators, infrastructure variables, and digital literacy profiles. The Hybrid Voting model also produced strong results, showing that combining probabilistic and ensemble learning methods can support reliable recommendation generation. Naïve Bayes performed lower but remained useful as a baseline model.

F. Confusion Matrix of the Best Model

The Random Forest model correctly classified most observations across the three digital literacy profiles. Misclassifications occurred mainly between adjacent categories, particularly between low and moderate, and between moderate and high.

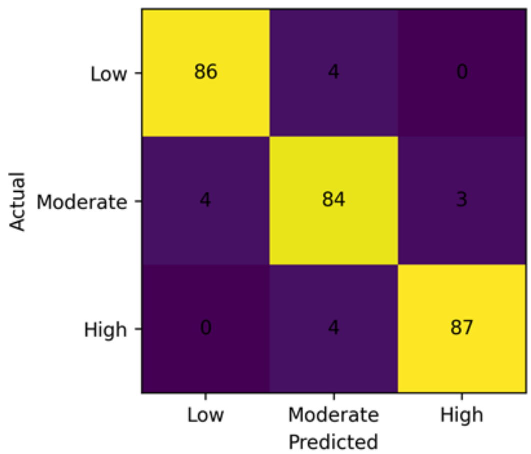

Fig. 1. Confusion matrix of the best-performing Random Forest model.

Fig. 1 shows the confusion matrix of the Random Forest. The confusion matrix indicates that the model was highly reliable in identifying low and high digital literacy profiles. Moderate cases were more difficult to classify because they share characteristics with both low and high profiles. This is important for the recommender system because moderate-profile environments require more refined recommendations based on their specific weaknesses.

G. Feature Importance and XAI-Based Interpretation

The feature importance results showed that student AI usage, urban AI usage, internet penetration, education index, school AI adoption, teacher AI usage, and rural AI usage were the strongest contributors to digital literacy classification.

TABLE VI. TOP PREDICTORS OF DIGITAL LITERACY PROFILE

Rank	Predictor	Importance Score
1	Student AI usage	0.1266
2	Urban AI usage	0.1237
3	Internet penetration	0.106
4	Education index	0.1026
5	School AI adoption	0.1008
6	Teacher AI usage	0.0958
7	Rural AI usage	0.0862
8	Year	0.069
9	Region	0.0631
10	Average daily AI usage	0.0589

The results presented in Table VI suggest that digital literacy development is influenced by both individual AI engagement and supportive educational conditions. Variables such as student AI usage and teacher AI usage represent the extent to which learners and educators actively utilize AI technologies in teaching and learning. Higher levels of AI engagement provide more opportunities to develop essential digital competencies, including information management, critical thinking, problem-solving, and the ethical use of digital technologies.

Meanwhile, internet penetration and the education index reflect the structural conditions that enable digital literacy. Reliable internet access provides users with opportunities to utilize digital platforms and AI applications, while a higher education index indicates a stronger educational environment that supports technology integration and digital skills development.

H. Cluster-Based Pattern Discovery

Cluster analysis identified three major patterns of AI adoption and digital literacy development.

TABLE VII. CLUSTER PROFILE

Cluster	Count	Mean DLS	Dominant Profile	Description
Cluster 0	453	0.383	Low	Low infrastructure and low readiness pattern
Cluster 1	453	0.414	Low	High infrastructure but low AI engagement pattern
Cluster 2	454	0.709	High	High readiness and high AI adoption pattern

Table VII shows that the respondents were grouped into three distinct clusters based on their Digital Literacy Score (DLS), infrastructure, and AI engagement.

Cluster 0 (n = 453) obtained a mean DLS of 0.383 and represents respondents with low infrastructure and low digital readiness. This group has limited access to digital resources and requires support in both infrastructure and foundational digital literacy.

Cluster 1 (n = 453) recorded a mean DLS of 0.414, also indicating low readiness. Although respondents have relatively better digital infrastructure, their AI engagement remains low, suggesting that access to technology alone does not ensure AI adoption. Targeted AI training and competency development are needed for this group.

Cluster 2 (n = 454) achieved the highest mean DLS of 0.709, reflecting high digital readiness and strong AI adoption. Respondents in this cluster are well-equipped to integrate AI into practice and can serve as early adopters or mentors.

I. Recommendation Rule Analysis

Table VIII summarizes the recommendation conditions generated by the proposed Intelligent Recommender System Framework based on identified digital literacy profiles and influential predictors.

TABLE VIII. RECOMMENDATION CONDITIONS AND INTERVENTION PRIORITIES.

Condition	Affected Records	Percentage	Main Recommendation
Low internet penetration	408	0.3	Internet infrastructure and connectivity programs
Low teacher AI usage	345	0.2537	AI-focused teacher training
No AI curriculum integration	673	0.4949	Curriculum redesign and AI literacy modules
Low school AI adoption	351	0.2581	Institutional AI adoption roadmap
High urban-rural AI gap	384	0.2824	Rural digital inclusion programs
High gender gap in AI usage	391	0.2875	Inclusive digital literacy interventions

The most frequent recommendation condition was the absence of AI curriculum integration, affecting 49.49% of the dataset. This indicates that curriculum enhancement is a major priority for strengthening AI-enabled science education. Low internet penetration, low teacher AI usage, low school AI adoption, and high urban-rural gaps also appeared as important intervention areas. The recommendation analysis confirms that the framework can translate analytical findings into actionable strategies. For low digital literacy environments, the system prioritizes access, infrastructure, and basic AI readiness. For moderate environments, the system recommends curriculum integration, teacher capacity building, and institutional adoption. For high-performing environments, the system may recommend advanced AI-enhanced science learning, innovation programs, and continuous improvement strategies.

J. Proposed Intelligent Recommender System Framework

Fig. 2 illustrates the proposed Intelligent Recommender System Framework developed to support digital literacy development and AI integration in science education.

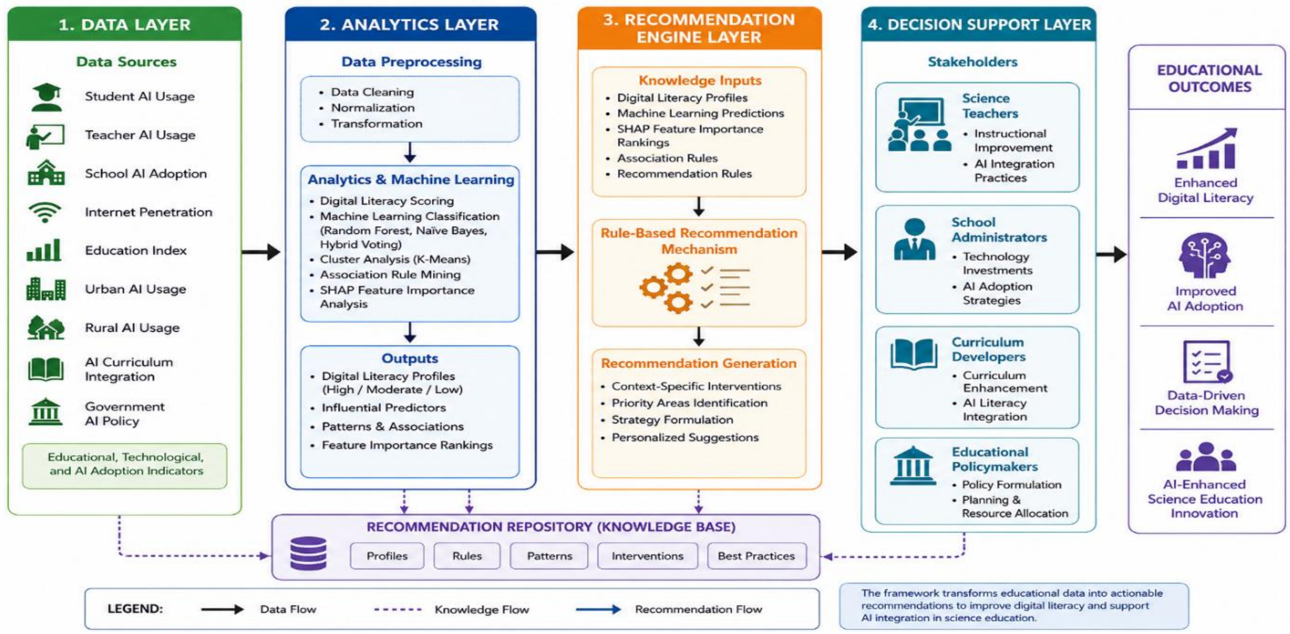

Fig. 2. The proposed Intelligent Recommender System Framework.

The framework consists of four interconnected layers that transform educational data into evidence-based recommendations for various educational stakeholders. The process begins with the Data Layer, which gathers educational, technological, and AI adoption indicators, including student and teacher AI usage, school AI adoption, internet penetration, education index, urban and rural AI usage, curriculum integration, and government AI policies. These variables provide the foundational information needed to assess digital literacy conditions and AI readiness across educational environments.

The collected data are then processed in the Analytics Layer, where data preprocessing, digital literacy profiling, machine learning analysis, cluster analysis, association rule mining, and SHAP-based explainable artificial intelligence are performed. This layer identifies significant patterns, influential predictors, and digital literacy profiles that characterize the technological readiness and AI adoption status of educational environments.

The extracted knowledge is subsequently transferred to the Recommendation Engine Layer, which serves as the core intelligence component of the framework. Using digital literacy profiles, machine learning outputs, SHAP feature importance rankings, and recommendation rules, the system generates targeted interventions and context-specific recommendations. These recommendations address identified gaps and priorities related to infrastructure, teacher competency development, curriculum enhancement, technology adoption, and digital inclusion.

The generated recommendations are delivered through the Decision Support Layer to key stakeholders, including science teachers, school administrators, curriculum developers, and educational policymakers. This layer enables stakeholders to utilize the recommendations for instructional improvement, technology planning, curriculum enhancement, resource allocation, and policy development.

Supporting all layers is the Recommendation Repository, which functions as a knowledge base that stores digital literacy profiles, discovered patterns, recommendation rules, intervention strategies, and best practices. This repository facilitates continuous knowledge accumulation and supports future recommendation generation.

This framework transforms educational data into actionable insights that promote enhanced digital literacy, improved AI adoption, evidence-based decision-making, and innovation in science education. By integrating Educational Data Mining, explainable artificial intelligence, and recommendation mechanisms, the framework provides a practical and scalable decision-support tool for advancing AI-enabled science education at institutional, regional, and national levels.

K. Discussion

The results demonstrate that global AI adoption and digital literacy development in science education are influenced by a combination of learner engagement, teacher readiness, institutional adoption, infrastructure, and regional conditions. The steady increase in Digital Literacy Score from 2015 to 2026 indicates that educational systems are becoming more digitally capable. However, regional differences and cluster patterns show that digital transformation remains uneven. These findings are particularly relevant to science education because effective integration of artificial intelligence supports inquiry-based learning, virtual experimentation, scientific problem-solving, and data-driven investigation. Educational environments demonstrating higher AI readiness are more capable of implementing AI-assisted science instruction, whereas institutions with lower digital literacy require targeted infrastructure, teacher development, and curriculum enhancement before AI can effectively support science learning.

The strong influence of student AI usage and teacher AI usage confirms that digital literacy development depends not only on access to technology but also on meaningful use of AI tools in teaching and learning. Teacher AI usage is especially important because teachers serve as facilitators of AI-enhanced science learning. When teacher AI usage is low, students may have limited opportunities to develop responsible and productive AI-related competencies.

Internet penetration and the education index were also among the most influential predictors. This finding suggests that digital literacy development requires a supportive ecosystem. Educational environments with strong infrastructure and higher educational development are more capable of integrating AI tools into science education. Conversely, low internet penetration limits access to AI-enabled learning resources and weakens the foundation for digital literacy development.

The cluster analysis provides further evidence that recommendation strategies must be context-specific. Some low-performing environments suffer from limited infrastructure, while others have adequate infrastructure but weak AI engagement. Therefore, a single intervention strategy is insufficient. The proposed Intelligent Recommender System Framework addresses this issue by matching each digital literacy profile with appropriate intervention priorities.

The results support the development of a recommender system framework that uses digital literacy profiling, machine learning classification, XAI-based feature interpretation, and rule-based recommendation generation. The framework can assist science teachers, administrators, curriculum developers, and policymakers in identifying priority areas for AI integration and digital literacy development. By converting data patterns into targeted recommendations, the framework provides a practical decision-support mechanism for improving science education in the digital age.

L. Limitations of the Study

One limitation of this study is that the Digital Literacy Score was developed as a composite indicator using normalized variables available in the dataset rather than a standardized digital literacy assessment instrument. Future studies should validate the proposed score against established digital literacy frameworks such as DigCompEdu [45] digital competency models.

IV. Conclusion and Recommendations

A. Conclusion

This study successfully developed an Intelligent Recommender System Framework for analyzing global AI adoption patterns and digital literacy development in science education. The findings demonstrated that digital literacy development is influenced by a combination of technological access, educational readiness, institutional AI adoption, and learner and teacher engagement with AI technologies. The results revealed a consistent increase in digital literacy levels and AI adoption indicators from 2015 to 2026, indicating the growing integration of artificial intelligence within educational environments.

Machine learning and explainable artificial intelligence analyses identified student AI usage, urban AI usage, internet penetration, education index, school AI adoption, and teacher AI usage as the most influential factors affecting digital literacy development. The Random Forest model achieved the highest classification performance, confirming its effectiveness in identifying digital literacy profiles and supporting recommendation generation. Cluster analysis further demonstrated that educational environments exhibit distinct patterns of technological readiness and AI adoption, highlighting the need for context-specific interventions rather than uniform policy approaches.

The proposed Intelligent Recommender System Framework transformed analytical findings into actionable recommendations through digital literacy profiling, pattern discovery, feature importance analysis, and rule-based recommendation generation. By linking educational conditions with targeted interventions, the framework provides a practical and explainable decision-support mechanism capable of supporting infrastructure development, teacher capacity building, curriculum enhancement, AI integration, and digital inclusion initiatives. Although the framework demonstrated high classification performance, its robustness under noisy, incomplete, or highly imbalanced educational datasets was not explicitly evaluated. Future investigations should examine the resilience of the proposed framework using datasets containing missing values, measurement uncertainty, and diverse educational contexts.

B. Recommendations

Based on the findings of the study, educational institutions should strengthen digital literacy development initiatives by increasing access to reliable internet infrastructure, particularly in regions and educational environments characterized by low technological readiness. Investments in connectivity and educational technology resources can provide the foundation necessary for effective AI integration and digital learning.

Teacher professional development programs should be expanded to enhance AI-related competencies, instructional innovation, and technology integration skills. Since teacher AI usage emerged as a significant predictor of digital literacy development, continuous training programs and certification initiatives should be prioritized to improve educators' capacity to utilize AI tools effectively in science education.

Curriculum developers are encouraged to incorporate AI literacy concepts, data literacy skills, and AI-supported learning activities within science curricula. Greater integration of AI into curriculum design can help learners develop the competencies necessary to engage responsibly and effectively with emerging technologies.

Educational policymakers should utilize data-driven approaches in planning AI adoption strategies and digital transformation initiatives. The implementation of targeted interventions based on digital literacy profiles can improve resource allocation and maximize the impact of educational investments.

Future researchers may extend the proposed framework by incorporating real-time educational data, longitudinal analyses, learning analytics indicators, and additional explainable AI techniques. Further validation of the framework using institutional, regional, or national datasets may also enhance its applicability and effectiveness in supporting AI-enabled educational innovation.

ACKNOWLEDGMENT

The researcher gratefully acknowledges the Pangasinan State University for its support in this study and Kaggle for providing the dataset used in the research. Artificial intelligence tools were utilized solely for language refinement, organization, and formatting. All research processes, analyses, findings, and conclusions were reviewed and validated by the researcher to ensure accuracy, originality, and academic integrity.

REFERENCES

[1] M. K. Rahman, M. A. Hossain, N. A. Ismail, M. S. Hossen, and M. Sultana, "Determinants of students' adoption of AI chatbots in higher education: The moderating role of tech readiness," Interactive Technology and Smart Education, vol. 22, no. 3, pp. 448–480, 2025.

[2] M. Rieckmann, Education for Sustainable Development Goals: Learning Objectives. Paris, France: UNESCO Publishing, 2017.

[3] N. Azmi, "The benefits of using ICT in the EFL classroom: From perceived utility to potential challenges," Journal of Educational and Social Research, vol. 7, no. 1, pp. 111–118, 2017.

[4] M. M. Yunus, N. Nordin, H. Salehi, M. A. Embi, and Z. Salehi, "The use of information and communication technology (ICT) in teaching ESL writing skills," English Language Teaching, vol. 6, no. 7, pp. 1–8, 2013.

[5] D. Fidaoui, R. Bahous, and N. N. Bacha, "CALL in Lebanese elementary ESL writing classrooms," Computer Assisted Language Learning, vol. 23, no. 2, pp. 151–168, 2010.

[6] P. Wang, "The effect of computer-assisted whole language instruction on Taiwanese university students' English learning," English Language Teaching, vol. 4, no. 4, pp. 10–27, 2011.

[7] A. Qoura, "EFL teacher competencies in the ICT age," Journal of Research in Curriculum, Instruction and Educational Technology, vol. 3, no. 2, pp. 127–159, 2017.

[8] T. Boudjadar, "ICT in the writing classroom: The pros and the cons," International Journal of Applied Linguistics and English Literature, vol. 4, no. 1, pp. 8–13, 2015.

[9] P. Mishra and M. J. Koehler, "Technological pedagogical content knowledge: A framework for teacher knowledge," Teachers College Record, vol. 108, no. 6, pp. 1017–1054, 2006.

[10] L. S. Shulman, "Those who understand: Knowledge growth in teaching," Educational Researcher, vol. 15, no. 2, pp. 4–14, 1986.

[11] L. S. Shulman, "Knowledge and teaching: Foundations of the new reform," Harvard Educational Review, vol. 57, no. 1, pp. 1–23, 1987.

[12] R. Rahmany, B. Sadeghi, and A. S. Chegini, "Normalization of CALL and TPACK: Discovering teachers' opportunities and challenges," Journal of Language Teaching and Research, vol. 5, no. 4, pp. 891–900, 2014.

[13] P. Redmond and J. Lock, "Secondary pre-service teachers' perceptions of technological pedagogical content knowledge (TPACK): What do they really think?" Australasian Journal of Educational Technology, vol. 35, no. 3, pp. 45–54, 2019.

[14] S. Sharma, P. Mittal, M. Kumar, and V. Bhardwaj, "The role of large language models in personalized learning: A systematic review of educational impact," Discover Sustainability, vol. 6, no. 1, pp. 1–24, 2025.

[15] M. A. Rodríguez-Ortiz, P. C. Santana-Mancilla, and L. E. Anido-Rifón, "Machine learning and generative AI in learning analytics for higher education: A systematic review of models, trends, and challenges," Applied Sciences, vol. 15, no. 15, Art. no. 8679, 2025.

[16] D. Tbaishat, G. Amoudi, and M. Elfadel, "Adapting teaching and learning with existing generative AI by higher education students," Computers and Education: Artificial Intelligence, vol. 8, Art. no. 100421, 2025.

[17] E. F. Risko and S. J. Gilbert, "Cognitive offloading," Trends in Cognitive Sciences, vol. 20, no. 9, pp. 676–688, 2016.

[18] J. Sweller, "Cognitive load during problem solving: Effects on learning," Cognitive Science, vol. 12, no. 2, pp. 257–285, 1988.

[19] A. Sharma and R. Rani, "A systematic review of applications of machine learning in cancer prediction and diagnosis," Archives of Computational Methods in Engineering, vol. 28, no. 7, pp. 4875–4896, 2021.

[20] M. Shamshuzzoha et al., "A novel framework for seasonal affective disorder detection: Comprehensive machine learning analysis using multimodal social media data and SMOTE," Acta Psychologica, vol. 256, Art. no. 105005, 2025.

[21] M. Schreiber et al., "A discourse on the use of machine learning in personality psychology," Journal of Research in Personality, vol. 119, Art. no. 104666, 2025.

[22] M. M. Ramadhan et al., "Parameter tuning in random forest based on grid search method," DEStech Transactions on Computer Science and Engineering, 2017.

[23] B. Sumathi, "Grid search tuning of hyperparameters in random forest classifier for customer feedback sentiment prediction," International Journal of Advanced Computer Science and Applications, vol. 11, no. 9, 2020.

[24] S. Subbiah et al., "Intrusion detection technique in wireless sensor network using grid search random forest with Boruta feature selection algorithm," Journal of Communications and Networks, vol. 24, no. 2, pp. 264–273, 2022.

[25] A. A. Salamai, E. S. M. El-Kenawy, and I. Abdelhameed, "Dynamic voting classifier for risk identification in supply chain 4.0," Computers, Materials & Continua, vol. 69, no. 3, 2021.

[26] N. A. Samee et al., "Metaheuristic optimization through deep learning classification of COVID-19 in chest X-ray images," Computers, Materials & Continua, vol. 73, no. 2, 2022.

[27] I. Rish, "An empirical study of the naive Bayes classifier," in Proc. IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, 2001, pp. 41–46.

[28] A. M. Salih et al., "A perspective on explainable artificial intelligence methods: SHAP and LIME," Advanced Intelligent Systems, vol. 7, no. 1, Art. no. 2400304, 2025.

[29] S. Hussain, "Global AI in Education Dataset (2015–2026)," Kaggle Dataset, 2026. [Online]. Available: https://doi.org/10.34740/KAGGLE/DSV/15685116. [Accessed: Jan. 7, 2026].

[30] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, Recommender Systems: An Introduction. New York, NY, USA: Cambridge University Press, 2011.

[31] W. Liu and L. Gao, "Recommendation system based on fuzzy cognitive map," Journal of Multimedia, vol. 9, no. 7, pp. 970–976, 2014.

[32] IEEE Learning Technology Standards Committee, "Learning Object Metadata (LOM) Standard." [Online]. Available: http://ieeeltsc.org/wg12LOM/

[33] B. Ojokoh, M. Omisore, O. Samuel, and T. Ogunniyi, "A fuzzy logic based personalized recommender system," International Journal of Computer Science Information Technology and Security, vol. 2, no. 5, pp. 1008–1015, 2012.

[34] K. Palanivel and R. Sivakumar, "Fuzzy multicriteria decision making approach for collaborative recommender systems," International Journal of Computer Theory and Engineering, vol. 2, no. 1, pp. 57–63, 2010.

[35] A. Rodriguez, J. Gago, L. Anido-Rifón, and R. Rodríguez, "A recommender system for non-traditional educational resources: A semantic approach," Journal of Universal Computer Science, vol. 21, pp. 306–325, 2015.

[36] F. Ricci, L. Rokach, B. Shapira, and P. Kantor, Eds., Recommender Systems Handbook. New York, NY, USA: Springer, 2011.

[37] L. Rifón, A. Cañas, V. Roris, J. Gago, and M. Iglesias, "A recommender system for educational resources in specific learning contexts," in Proc. 8th International Conference on Computer Science and Education (ICCSE), 2013, pp. 371–376.

[38] E. Stuart, N. Shadbolt, and D. De Roure, "Ontological user profiling in recommender systems," ACM Transactions on Information Systems, vol. 22, no. 1, pp. 54–88, 2004.

[39] R. Sikka, A. Dhankhar, and C. Rana, "A survey paper on e-learning recommender system," International Journal of Computer Applications, vol. 47, pp. 27–30, 2012.

[40] L. Terán, SmartParticipation: A Fuzzy-Based Recommender System for Political Community Building. Switzerland: Springer, 2014.

[41] C. Tsai and H. Chuang, "The role of cognitive decision effort in electronic commerce recommendation system," International Scholarly and Scientific Research & Innovation, vol. 5, no. 10, pp. 36–40, 2011.

[42] A. Almahairi, K. Kastner, K. Cho, and A. Courville, "Learning distributed representations from reviews for collaborative filtering," in Proc. 9th ACM Conference on Recommender Systems (RecSys '15), 2015, pp. 147–154.

[43] J. Neidhardt, R. Schuster, L. Seyfang, and H. Werthner, "Eliciting the users' unknown preferences," in Proc. 8th ACM Conference on Recommender Systems (RecSys '14), 2014, pp. 309–312.

[44] C. Mediani and M. H. Abel, "Semantic recommendation of pedagogical resources within learning ecosystems," in Proc. International Conference on Industrial Informatics and Computer Systems (CIICS), 2016,

[45] C. Redecker, European Framework for the Digital Competence of Educators: DigCompEdu, Y. Punie, Ed. Luxembourg: Publications Office of the European Union, 2017, EUR 28775 EN. doi: 10.2760/159770. [Online]. Available: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC107466/pdf_digcomedu_a4_final.pdf