Engineering and Technology Journal e-ISSN: 2456-3358

Volume 11 Issue 04 April-2026, Page No.-9600-9614

DOI: 10.47191/etj/v11i04.34, I.F. – 8.595

Student Affairs Merit System: A Digital Performance Assessment System Using Rule-Based Algorithm

Nathaniel Andrei Gonzales[1], Xyrus Lex Pajarillaga[2], Jhan Robin Sy[3], Rhugene Villegas[4], Eliza B. Ayo[5]

[1,2,3,4,5]School of Science and Technology, Centro Escolar University – Manila, Philippines

ABSTRACT: The digitalization of student affairs processes has become essential to promoting transparency, efficiency, and fairness in higher education. This study introduces and evaluates the Student Affairs Merit System (SAMS), a web-based platform developed to modernize the evaluation of co-curricular and extracurricular activities at Centro Escolar University (CEU) Manila. The system integrates an automated merit point computation engine, an artificial intelligence (AI)–powered chatbot for user guidance, a notification module, and a role-based authentication framework within a multi-tiered validation workflow. A descriptive-developmental research design was employed. System quality was assessed using the ISO/IEC 25010 framework across seven domains: Functionality Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, and Maintainability. Survey data were gathered from 52 end-users who rated the system on a four-point Likert scale, and responses were analyzed using descriptive statistics. All seven domains yielded mean scores within the Strongly Agree range (M = 3.73–3.81). Compatibility recorded the highest mean (M = 3.81), while Usability recorded the lowest (M = 3.73), identifying interface navigation as the primary area for refinement. The findings indicate that SAMS is a reliable, efficient, and secure digital solution that is ready for broader institutional deployment, with targeted enhancements recommended to optimize the overall user experience. The study contributes to the growing literature on digital performance assessment by demonstrating how standardized evaluation frameworks and AI-assisted guidance can be combined to support equitable merit recognition in higher education.

KEYWORDS: digital merit system, student affairs, ISO/IEC 25010, web-based evaluation, AI chatbot, extracurricular assessment

1. INTRODUCTION

The recognition of student achievement beyond the classroom is a fundamental component of holistic higher education, contributing to the development of leadership, personal growth, and professional readiness. Across universities worldwide, merit-based recognition systems serve as formal mechanisms for acknowledging student contributions through co-curricular and extracurricular engagement. Despite their importance, many institutions continue to rely on manual, paper-based processes that struggle to keep pace with the demands of increasingly diverse and growing student populations. At Centro Escolar University (CEU) Manila, the existing merit assessment workflow depends heavily on physical documentation and manual computation, resulting in inefficiencies, inconsistencies, and limited transparency.

Traditional evaluation processes are characterized by well-documented weaknesses, including subjectivity in scoring, inconsistent documentation practices, and administrative delays that postpone the timely recognition of student contributions (Kalischko & Riedl, 2021). Research on electronic performance monitoring (EPM) has demonstrated that digital platforms reduce such inefficiencies by minimizing human error, accelerating reporting cycles, and producing auditable records that paper systems cannot reliably provide. Consequently, the migration from paper-based to digital assessment represents not merely a technological upgrade but a structural improvement in institutional fairness and accountability.

1.1 Research Problem

Although digital transformation in student affairs has gained momentum, limited research has examined the implementation of integrated merit assessment platforms in Philippine higher education institutions. At CEU Manila, the reliance on paper-based submissions and manual calculations exemplifies the documented weaknesses of traditional evaluation systems. Such processes introduce variability into merit recognition, delay administrative workflows, and erode student confidence in the fairness of outcomes. Addressing these limitations requires a purpose-built digital solution that combines automated computation, transparent validation, and accessible guidance for all user roles.

1.2 Gap in the Literature

Prior studies have explored various components of digital assessment, including rule-based performance evaluation systems (Alfonso et al., 2022), electronic performance monitoring (Kalischko & Riedl, 2021), and the application of AI chatbots in educational contexts (Dawood, 2024; Rahim et al., 2022). However, few studies have examined the integration of these components into a unified, institutionally deployed merit

system evaluated against internationally recognized software quality standards. Furthermore, limited empirical evidence exists regarding the end-user acceptance of such platforms in Southeast Asian higher education settings. This study addresses these gaps by presenting the development and evaluation of a web-based merit assessment system tailored to the institutional context of CEU Manila and evaluated systematically against the ISO/IEC 25010 framework.

1.3 Objectives of the Study

The overarching aim of this study is to transform CEU Manila's manual merit assessment process into a structured, transparent, and technologically supported digital system. Specifically, the study seeks to: (a) identify the challenges associated with implementing a digital merit assessment system for co-curricular and extracurricular activities at CEU Manila; (b) determine the technologies most suitable for developing the Student Affairs Merit System (SAMS); (c) examine how SAMS improves the efficiency and accessibility of merit evaluation; (d) assess the system's performance in acceptance, usability, security, compatibility, performance, and stress testing; and (e) evaluate end-user perceptions of SAMS in accordance with the ISO/IEC 25010 quality standards.

1.4 Significance of the Study

This study holds significance for students, faculty, and administrators at CEU Manila and for higher education institutions considering comparable digital transformation initiatives. By replacing manual evaluation with an automated, rule-based workflow, SAMS enhances the fairness, transparency, and efficiency of merit recognition. The integrated AI chatbot reduces administrative workload by providing students with instant, accurate guidance on eligibility, documentation, and scoring criteria (Velázquez-García et al., 2024), while the notification module supports timely task completion by alerting users to pending actions. Taken together, these features establish a more reliable foundation for recognizing student contributions and promote institutional commitment to progressive, technology-enabled student affairs management (Betsurmath et al., 2024).

1.5 Scope and Delimitations

SAMS is designed as a web-based platform for tracking and assessing the co-curricular and extracurricular activities of CEU Manila students who are eligible for leadership awards as defined by the Student Affairs Office (SAO) scoring framework. The system covers the entirety of a student's enrollment period and processes only SAO-approved activities supported by appropriate documentation. The platform does not integrate with external systems such as Google Drive, institutional learning management systems, or school portals, and all submissions must occur within the SAMS environment. The AI chatbot is limited to responding to predefined query categories and does not interpret ambiguous documentation or override human evaluation decisions. Violations and demerits fall outside the functional scope of the system. Furthermore, SAMS does not guarantee the subjective fairness of individual rater or validator judgments; it ensures only that input is processed and recorded consistently.

LITERATURE REVIEW

This section reviews the theoretical foundations and empirical evidence relevant to the development and evaluation of digital merit systems in higher education. The discussion is organized around five thematic areas: (a) digital performance evaluation frameworks, (b) electronic performance monitoring and merit systems, (c) AI chatbots in education, (d) digital badges and recognition systems, and (e) persistent challenges in digital merit assessment. Each subsection situates SAMS within the broader scholarly conversation and highlights the gaps that this study seeks to address.

2.1 Digital Performance Evaluation Frameworks

Objective and scalable assessment requires structured frameworks that translate participation into measurable outcomes. Alfonso et al. (2022) demonstrated that a web-based performance evaluation system employing a rule-based algorithm ensures that contributions are assessed against consistent, predefined criteria rather than evaluator intuition. Arboleda-Velasquez (2024) extended this argument through the MERIT framework for manuscript peer review, which emphasizes transparency and equitable criteria design. The structural parallels between manuscript evaluation and student merit evaluation are instructive: both require explicit criteria, multi-reviewer checks, and mechanisms that limit subjective drift.

Complementing these frameworks, Idiyatova et al. (2024) and Tkachenko et al. (2024) have shown that digital assessment tools improve evaluation effectiveness only when users possess sufficient digital literacy to engage with them meaningfully. This finding underscores the importance of intuitive interface design and supplementary guidance mechanisms—considerations that inform the AI chatbot component of SAMS.

2.2 Electronic Performance Monitoring and Merit Systems

Electronic Performance Monitoring (EPM) provides the theoretical anchor for SAMS's core tracking function. Kalischko and Riedl (2021) conducted an extensive review of EPM in digital workplaces, documenting its effects on productivity, engagement, and perceived fairness. Their findings indicate that transparent monitoring with clearly communicated criteria increases engagement, whereas opaque criteria diminish it. Geshkov (2021) similarly reported that automated tracking in performance appraisal improves efficiency and accuracy, while cautioning against implementations that fail to communicate evaluation logic to those being evaluated. Both studies reinforce a design principle embedded in SAMS: users should always be able to see exactly how their merit points are calculated.

Merit systems in professional organizations offer additional insight. Curzi and Ferrarini (2023) found that high-performance work systems combining digital monitoring tools with structured employee participation produced meaningful innovation gains, while simultaneously surfacing the inherently subjective dimensions of meritocracy. Busso et al. (2024) examined a teacher selection reform in Colombia and reported that merit-based criteria improve selection fairness only when evaluation transparency is high and stakeholders trust the process.

A similar tension is present in student merit evaluation, where the credibility of outcomes depends on whether students perceive the scoring criteria as genuinely fair and consistently enforced.

At the institutional level, Alvaran et al. (2022) and Betsurmath et al. (2024) developed and validated digital performance appraisal systems that demonstrate the value of rule-based, automated evaluation in reducing evaluator variability, accelerating processing, and producing more defensible scoring decisions. SAMS adopts a comparable hybrid architecture, using text extraction and rule-based algorithms to assign preliminary merit scores before human reviewers apply contextual judgment.

2.3 AI Chatbots in Higher Education

AI chatbots in educational settings have matured from keyword-matching utilities into conversational agents capable of context-sensitive guidance. Dawood (2024) traced this evolution, noting that modern chatbots can effectively address a broad range of student inquiries, including academic advising tasks that formerly required dedicated staff time. Rahim et al. (2022) surveyed AI chatbot adoption in higher education and reported consistent efficiency gains, including faster response times, reduced administrative workload, and improved student satisfaction with information access.

Several studies have examined chatbot integration in merit-tracking and student support contexts. Dehankar et al. (2022) (2024) further established a direct correlation between badge relevance and sustained engagement.

Velázquez-García et al. (2024) found that badges serve as effective extrinsic motivators in higher education, though their long-term effectiveness depends on perceived credibility and meaningful differentiation between award levels. These findings carry direct implications for the design of SAMS's Gold, Silver, and Bronze award tiers, which function as an achievement ladder calibrated to give students visible and meaningful milestones. As with other motivational mechanisms, the critical design variable is perceived credibility: awards motivate participation only when students believe that the evaluation process is fair and that the criteria are genuinely attainable through effort.

2.5 Challenges and Research Gaps

Despite rapid advancement in digital assessment tools, several persistent challenges moderate enthusiasm for purely automated solutions. Fairness and transparency remain the most significant

demonstrated how chatbot systems deliver comprehensive information through institutional platforms, while Ndunagu et al. (2025) reported that AI-driven student support systems improved engagement and academic accessibility in distance learning environments. Chang et al. (2023) further emphasized that chatbots are most effective when they support self-regulated learning—helping students set goals, understand feedback, and monitor their own progress. Aditi et al. (2024) examined chatbot applications for accessibility, demonstrating how automated conversational tools reduce information asymmetries for users with varying levels of digital literacy. Collectively, these findings validate the design choices underlying the SAMS chatbot, which is deliberately scoped to platform navigation and merit policy queries to improve reliability and limit the risk of misleading guidance.

Nevertheless, chatbots are not a universal solution. Dawood (2024) observed that while chatbots manage routine queries effectively, they struggle with nuanced advising scenarios requiring human judgment. Manasa (2024) identified predefined dataset limitations as a constraint on chatbot adaptability, and Chang et al. (2023) stressed the need for personalization strategies to maintain effectiveness across diverse user populations. SAMS addresses these limitations by narrowing the chatbot's scope to merit-related queries and system navigation, while preserving human oversight for all evaluation decisions.

2.4 Digital Badges and Recognition Systems

Digital badges have emerged as a mechanism for recognizing achievement in ways that complement conventional grading. Sousa-Vieira et al. (2021) designed a badge system for higher education that recognizes specific competencies beyond aggregate grades, while Goulding et al. (2023) experimented with badges as alternatives to traditional grading and reported improved student engagement. Hartnett (2021) documented similar motivational gains in New Zealand's higher education sector, and Huber et al. (2022) demonstrated badge efficacy in experiential learning contexts. Funa

concerns. Patel et al. (2023) documented participation disparities in a merit-based incentive payment system for oncologists, a finding that echoes concerns about equity in student merit evaluation. Sun and Shi (2023) identified data inconsistencies and subjective criteria as recurring problems in digital government performance evaluation, issues that demand robust validation mechanisms rather than technological fixes alone. Vinith and Pinto (2023) raised ethical concerns regarding performance surveillance in workplace IT settings—concerns that translate directly to student data privacy within SAMS. Reliability and validity are equally critical. Saltos-Rivas et al. (2022) conducted a systematic mapping study on digital competence evaluations in higher education and concluded that many existing instruments fall short on one or both dimensions. Their findings reinforce the decision to evaluate SAMS against the ISO/IEC 25010 quality framework, which provides structured, measurable benchmarks for reliability and other quality attributes. Wellberg and Evans (2022) similarly

cautioned that performance assessment reforms often fail when they do not account for implementation contexts—a consideration that informed the decision to involve students, raters, and SAO staff in defining evaluation criteria prior to development.

Synthesizing across the literature, a coherent picture emerges: effective digital merit systems are not simply automated versions of their paper predecessors but carefully designed sociotechnical systems that balance algorithmic consistency with human judgment, efficiency with ethical accountability, and user engagement with institutional integrity. The most successful implementations share a commitment to transparent criteria, multi-level review processes, and meaningful feedback mechanisms. SAMS is designed with these lessons in mind. Its rule-based scoring algorithm, multi-tiered validation workflow, AI chatbot guidance, and notification system represent a coherent response to the documented weaknesses of manual merit evaluation. The present study contributes empirical evidence regarding how such an integrated system performs when implemented and evaluated in a Philippine higher education context

3. METHODOLOGY

3.1 Research Design

This study employed a descriptive-developmental research design to develop, evaluate, and iteratively refine the Student Affairs Merit System (SAMS). The descriptive component documented the development process in detail, recording design decisions, implementation challenges, and stakeholder responses as they occurred. The developmental component structured the core activity of the research: constructing a functional system, subjecting it to systematic testing, and refining it based on empirical feedback. Together, these complementary approaches supported the assessment of SAMS not merely as a technical artifact but as a sociotechnical system embedded in a real institutional context.

For end-user evaluation, a quantitative survey design was adopted. Participants rated structured statements about system quality using a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree) covering all seven ISO/IEC 25010 quality domains. Descriptive statistics—including mean scores, standard deviations, and frequency distributions—were computed for each item and domain to identify performance patterns and overall user sentiment. A pilot test was conducted prior to full survey deployment to verify item clarity and to refine question wording based on preliminary feedback.

3.2 Participants

A total of 52 respondents participated in the end-user evaluation, representing students, raters, and validators who interacted with SAMS during the beta testing phase. The sample was drawn from multiple departments within CEU Manila to ensure a diverse and representative respondent pool. Stakeholder interviews were also conducted with faculty members, SAO staff, and student organization leaders to provide qualitative context for interpreting the quantitative survey findings.

3.3 Instruments

The primary evaluation instrument was a structured questionnaire aligned with the seven quality domains of the ISO/IEC 25010 software quality framework: Functionality Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, and Maintainability. Each domain was operationalized through four to six indicator items, yielding a total of 32 items. Responses were collected on a four-point Likert scale, and verbal interpretation categories were assigned as follows: Strongly Agree (3.26–4.00), Agree (2.51–3.25), Disagree (1.76–2.50), and Strongly Disagree (1.00–1.75).

Complementary system testing was conducted against the ISO/IEC 25010 framework to generate objective performance data across all seven quality domains. This combination of quantitative survey responses and technical testing supported the triangulation of findings and strengthened the validity of the evaluation.

3.4 System Development

SAMS was developed using the Agile methodology, organized into six iterative phases: Planning, Design, Development, Review, Deployment, and Launch. Each phase produced specific deliverables that informed the next, while sprint retrospectives created feedback loops for continuous improvement. During the Planning phase, stakeholder consultations identified requirements and documented gaps in the existing manual process. The Design phase produced system architecture diagrams, database schemas, user role definitions, user interface and user experience (UI/UX) wireframes, and chatbot conversation flow maps. The Development phase implemented core functionality using a React.js frontend, a Node.js (Express.js) backend, and a MongoDB database. The Review phase encompassed comprehensive ISO/IEC 25010 testing, pilot user feedback collection, and chatbot accuracy evaluation. The Deployment phase configured the production server environment, implemented security measures, and conducted user training. The

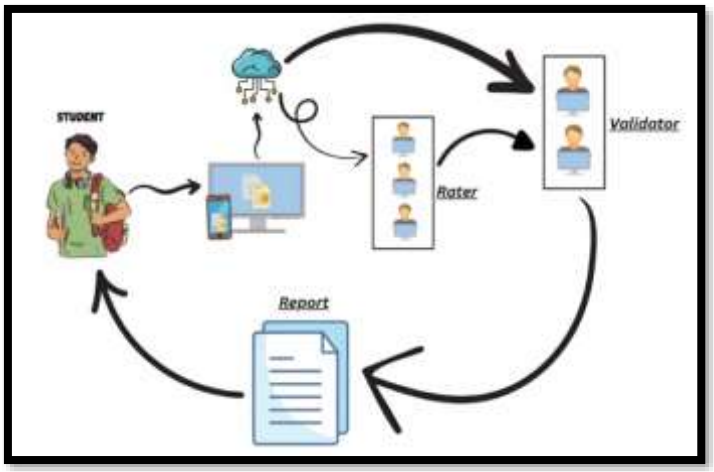

Fig. 1. Conceptual Framework

Launch phase introduced SAMS through a phased rollout supported by dedicated support channels and performance monitoring.

A notable practical challenge emerged during chatbot calibration. During early testing, a small but non-trivial proportion of student queries contained informal phrasing or department-specific terminology that the initial intent-recognition model did not handle reliably. Rather than expanding the training dataset indiscriminately, a targeted approach was adopted: misclassified queries from the first two testing sprints were collected, categorized by failure type, and used to retrain only the affected intent categories. This approach reduced the misclassification rate from approximately 18% to under 6% prior to beta deployment. The experience reinforced an important design principle: chatbot reliability depends not only on the volume of initial training data but also on the ongoing curation of domain-specific edge cases.

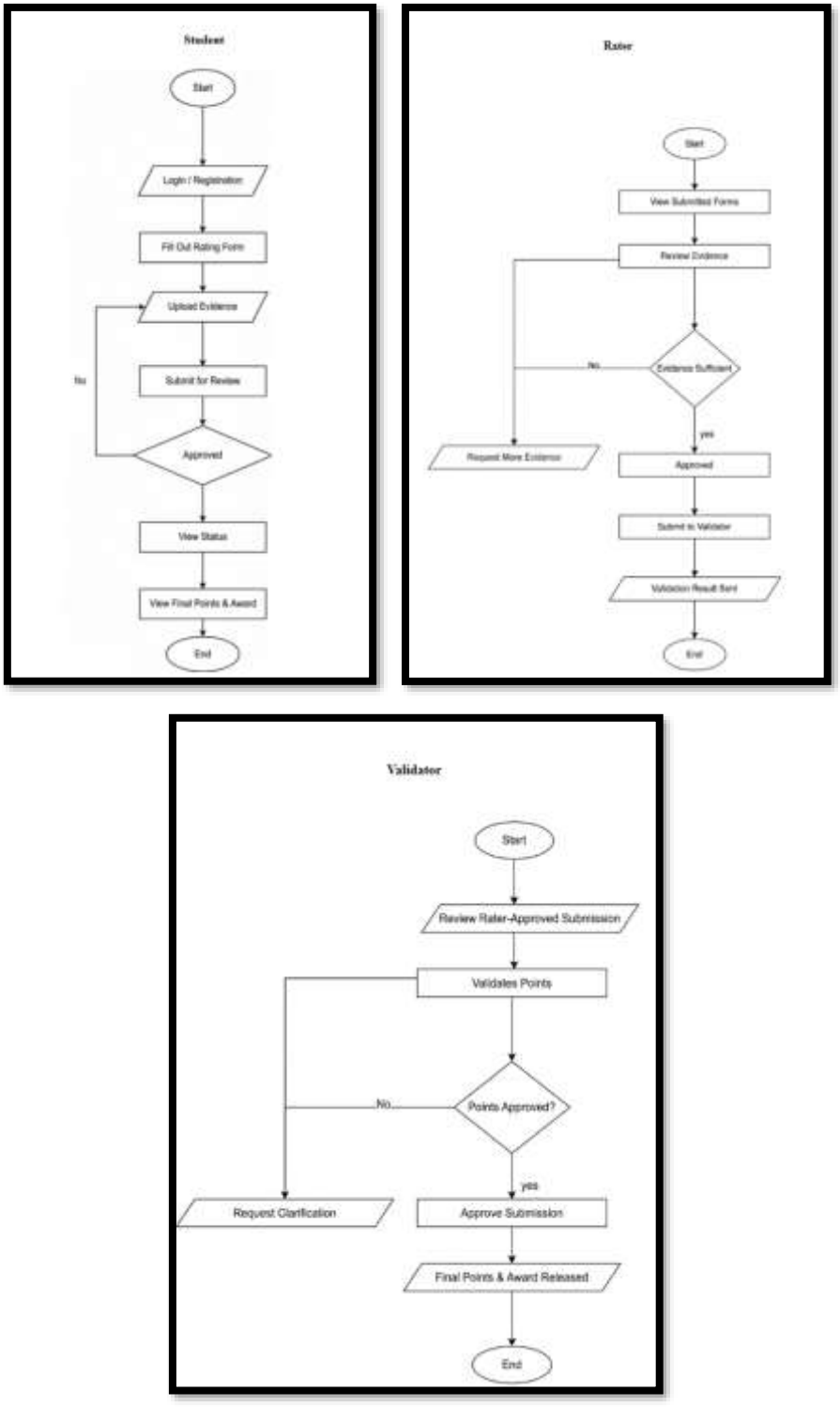

Fig. 2. ***Flowcharts***

3.5 Merit Point Computation

SAMS assigns merit points through a rule-based algorithm that employs text extraction to analyze uploaded documents. When a student submits a certificate or supporting document, the system retrieves key text fields and matches them against predefined criteria to determine the activity type and associated point value. If the document meets the required standards, points are automatically assigned; otherwise, the submission is flagged for manual review or returned to the student with a request for additional documentation. This approach ensures consistent automated preliminary scoring while preserving the role of human raters in resolving edge cases.

Total merit points are computed using the following summation formula:

$$S = \Sigma (x_i + y_i)$$

Where S represents the total accumulated points, x_i denotes points from co-curricular activities, y_i denotes points from extracurricular activities, and Σ represents the summation across all approved submissions. Medal thresholds for Gold, Silver, and Bronze awards are determined by the Vice President for Student Affairs and administered by the SAO in accordance with institutional policy.

Rule-Based Algorithm

Our algorithm validates documents using three rule-based checks: a File Extension Rule that only accepts .pdf, .doc, .docx, and .png files, a Text Extraction Rule that scans for the keyword "Certificate", and a Date Extraction Rule that retrieves the issuer date. A document is only valid if

```
it passes all three rules.
const fs = require('fs');
const path = require('path');

const RULES = {
allowedExtensions: ['.pdf', '.doc', '.docx', '.png'],
requiredKeyword: 'Certificate',
datePatterns: [
/issued\s*[:\-]?\s*(\d{1,2}[\/\-]\d{1,2}[\/\-]\d{2,4})/i,
/date\s*[:\-]?\s*(\d{1,2}[\/\-]\d{1,2}[\/\-]\d{2,4})/i,
/(\w+ \d{1,2},? \d{4})/,
/(\d{1,2}[\/\-]\d{1,2}[\/\-]\d{2,4})/,
]
};

function validateExtension(filename) {
const ext = path.extname(filename).toLowerCase();
const isValid = RULES.allowedExtensions.includes(ext);
return {
rule: 'File Extension',
extension: ext,
passed: isValid,
message: isValid
? `Valid file type: ${ext}`
: Invalid file type: "${ext}". Allowed: ${RULES.allowedExtensions.join(', ')}`
};
}
```

```
function extractKeyword(text) {
const keyword = RULES.requiredKeyword;
const found =
text.toLowerCase().includes(keyword.toLowerCase());
const lines = text.split('\n');
const matchLine = lines.find(line =>
line.toLowerCase().includes(keyword.toLowerCase())
);
return {
rule: 'Keyword Extraction',
keyword,
passed: found,
foundIn: matchLine ? matchLine.trim() : null,
message: found
? `✅ Keyword "${keyword}" found: "${matchLine?.trim()}"`
: `❌ Keyword "${keyword}" not found in document`
};
}

function extractIssuerDate(text) {
let extractedDate = null;
for (const pattern of RULES.datePatterns) {
const match = text.match(pattern);
if (match) {
extractedDate = match[1] || match[0];
break;
}
}
return {
rule: 'Issuer Date Extraction',
extractedDate,
passed: extractedDate !== null,
message: extractedDate
? `✅ Issuer date found: "${extractedDate}"`
: `❌ No issuer date found in document`
};
}

function validateDocument(filename, fileContent) {
const results = [];
const extResult = validateExtension(filename);
results.push(extResult);
console.log(extResult.message);

if (!extResult.passed) {
```

```
return { filename, passed: false, results };
}

const keywordResult = extractKeyword(fileContent);
results.push(keywordResult);
console.log(keywordResult.message);

const dateResult = extractIssuerDate(fileContent);
results.push(dateResult);
console.log(dateResult.message);

const allPassed = results.every(r => r.passed);
console.log(allPassed ? '✅ RESULT: Document is VALID' :
'❌ RESULT: Document FAILED validation');

return { filename, passed: allPassed, results };
}

const content = fs.readFileSync('./your-file.pdf', 'utf-8');
validateDocument('your-file.pdf', content);
```

3.6 Data Collection and Analysis

Primary data were collected through three complementary mechanisms: structured system testing, user feedback surveys, and stakeholder interviews. Students evaluated the system's ease of use, clarity of merit point presentation, and overall experience relative to the previous manual process. Faculty members and student organization leaders assessed the efficiency of document verification and evaluation reporting. Administrators contributed insights on system reliability, security, and integration with existing institutional workflows. Observational data gathered during beta testing sessions further informed interface and workflow refinements. Survey responses were analyzed using descriptive statistics to compute mean scores and standard deviations at both the item and domain levels, and findings were interpreted in relation to the verbal interpretation ranges defined above.

3.7 Hardware and Software Requirements

The hardware and software configurations used during the development and implementation of SAMS are summarized in Table 1. A server-centric architecture was adopted so that the system could perform most computations on the server side, ensuring accessibility for students using modest hardware. This design choice aligns with CEU Manila's commitment to equitable access across its diverse student population

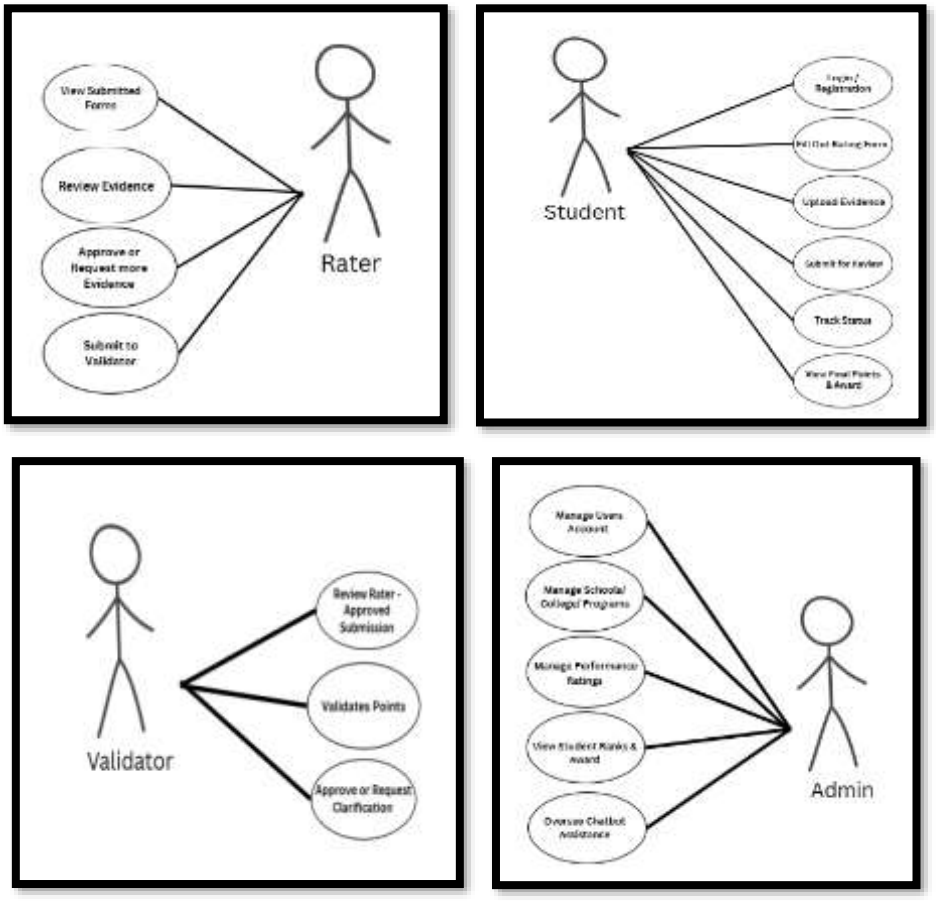

Fig. 3. Use Cases

Table 1. Hardware and Software Requirements by Development Phase

Category	Development Phase	Implementation Phase
Hardware	CPU: Intel Core i5 or higher; RAM: 8 GB or higher; Storage: 256 GB SSD; GPU: Integrated or dedicated (for UI/UX testing); Network: Gigabit Ethernet or higher.	Server: Intel Xeon or AMD Ryzen (quad-core or higher); RAM: 16 GB or higher; Storage: 1 TB SSD; Network: Gigabit Ethernet. End-user devices: Intel Core i3, 4 GB RAM, 128 GB or higher storage.
Software	Frontend: HTML, CSS, JavaScript, React.js; Backend: Node.js (Express.js); Database: MongoDB; Chatbot: AI-powered (API-integrated); Tools: VS Code, Git, Postman, GitHub.	Frontend: React.js; Backend: Node.js (Express.js); Database: MongoDB; Hosting: AWS (EC2, RDS, S3); Security: SSL certificates and firewalls.

Note. Configuration specifications reflect minimum requirements for reliable development and deployment.

4. RESULTS

End-user evaluation of SAMS was based on responses from 52 participants across the seven quality domains of the ISO/IEC 25010 standard. Mean scores, standard deviations (SDs), and verbal interpretations are reported for each indicator and domain. Standard deviation values are reported alongside mean scores to indicate the degree of consensus among respondents.

4.1 Functionality Suitability

Table 2. Functionality Suitability

Indicator	Mean	SD	Verbal Interpretation
1.1. The system correctly calculates and displays total merit points.	3.79	0.41	Strongly Agree
1.2. Submitting activity documents (e.g., certificates, forms) through SAMS is easy.	3.83	0.38	Strongly Agree
1.3. The chatbot helps users find answers about navigating and using the system.	3.75	0.52	Strongly Agree
1.4. Each module (submission, evaluation, chatbot, notifications) performs its purpose effectively.	3.77	0.43	Strongly Agree
Overall Mean	3.78	0.44	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

As presented in Table 2, Functionality Suitability yielded an overall mean of 3.78 (SD = 0.44), with all four items falling within the Strongly Agree range. Document submission ease (Item 1.2) received the highest rating (M = 3.83, SD = 0.38), whereas chatbot guidance (Item 1.3) received the lowest score within the domain (M = 3.75) but also recorded the highest standard deviation (SD = 0.52), indicating greater variation in how respondents experienced chatbot support.

Table 3. Performance Efficiency

Indicator	Mean	SD	Verbal Interpretation
2.1. The system loads pages and features quickly.	3.75	0.56	Strongly Agree
2.2. Button clicks and form submissions respond without long delays.	3.75	0.48	Strongly Agree
2.3. The chatbot responds promptly to queries.	3.75	0.44	Strongly Agree
2.4. Checking merit points or submission status is fast.	3.81	0.40	Strongly Agree
2.5. Uploading documents or activity files is fast and efficient.	3.83	0.38	Strongly Agree

Indicator	Mean	SD	Verbal Interpretation
2.6. The system performs well even with multiple simultaneous users.	3.85	0.36	Strongly Agree
Overall Mean	3.79	0.44	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

As shown in Table 3, Performance Efficiency averaged 3.79 across the six indicators. Concurrent-user performance (Item 2.6) earned the highest score within the domain (M = 3.85, SD = 0.36), with the low standard deviation indicating near-unanimous confidence. Page loading speed (Item 2.1) registered the highest variability within the domain (SD = 0.56), suggesting that device and network conditions contributed to differing experiences.

Table 4. Compatibility

Indicator	Mean	SD	Verbal Interpretation
3.1. The system works properly on the user's laptop or desktop computer.	3.81	0.40	Strongly Agree
3.2. It functions smoothly on the user's preferred web browser (e.g., Chrome, Firefox, Edge).	3.85	0.41	Strongly Agree
3.3. Users do not experience unexpected errors related to their device or browser.	3.79	0.46	Strongly Agree
3.4. The system functions well across various operating systems.	3.81	0.44	Strongly Agree
Overall Mean	3.81	0.43	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

Compatibility, as summarized in Table 4, recorded the highest overall domain mean in the evaluation (M = 3.81, SD = 0.43). Browser compatibility (Item 3.2) led the domain (M = 3.85, SD = 0.41), indicating consistent behavior across Chrome, Firefox, and Edge. Item 3.3, which tracked unexpected device- or browser-related errors, scored the lowest within the domain (M = 3.79, SD = 0.46). The low standard deviations across all four items (0.40–0.46) reflect strong respondent consensus.

Table 5. **Usability**

Indicator	Mean	SD	Verbal Interpretation
4.1. The system is easy to use and navigate.	3.71	0.46	Strongly Agree
4.2. Buttons and menus are clear and easy to understand.	3.75	0.44	Strongly Agree
4.3. It was easy to learn how to use SAMS upon first use.	3.73	0.49	Strongly Agree
4.4. Users can find what they need (e.g., submission history, points) without difficulty.	3.75	0.48	Strongly Agree
Overall Mean	3.73	0.46	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

As reported in Table 5, Usability received the lowest overall mean in the evaluation (M = 3.73, SD = 0.46), although it remained firmly within the Strongly Agree range. Overall ease of use and navigation (Item 4.1) scored the lowest within the domain (M = 3.71), while menu clarity (Item 4.2) and ease of locating information (Item 4.4) tied for the highest scores in the domain (M = 3.75).

4.5 Reliability

Table 6. Reliability

Indicator	Mean	SD	Verbal Interpretation
5.1. The system is almost always available when needed.	3.81	0.40	Strongly Agree
5.2. The system rarely freezes, crashes, or logs users out unexpectedly.	3.75	0.44	Strongly Agree
5.3. Uploaded documents and entered information are saved correctly.	3.77	0.47	Strongly Agree
5.4. Users experience minimal to no errors while using SAMS.	3.81	0.40	Strongly Agree
5.5. The system displays accurate, current submission status.	3.77	0.43	Strongly Agree
Overall Mean	3.78	0.42	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

Table 6 shows that Reliability averaged 3.78 across the five indicators, with consistently low standard deviations (0.40–0.47). System availability (Item 5.1) and minimal errors (Item 5.4) tied for the highest ratings in the domain (M = 3.81, SD = 0.40). Item 5.2, concerning crashes, freezes, and unexpected logouts, received the lowest rating within the domain (M = 3.75, SD = 0.44).

Table 7. Security

Indicator	Mean	SD	Verbal Interpretation
6.1. The login process ensures that only authorized users can access the system.	3.75	0.44	Strongly Agree
6.2. The system consistently ensures that only valid users can access and use its features.	3.77	0.55	Strongly Agree
6.3. The system prevents unauthorized viewing or modification of submitted documents.	3.75	0.44	Strongly Agree
6.4. The system prevents users from accessing restricted sections.	3.77	0.43	Strongly Agree
Overall Mean	3.76	0.46	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

As shown in Table 7, Security averaged 3.76 across the four indicators. Items 6.2 and 6.4, which addressed valid-user enforcement and restricted-section prevention, tied for the highest scores (M = 3.77). Notably, Item 6.2 recorded a standard deviation of 0.55—the highest observed for any single item across the entire evaluation—indicating greater variability in respondent experience for this specific aspect of access control.

Table 8. Maintainability

Indicator	Mean	SD	Verbal Interpretation
7.1. System updates are introduced clearly and do not cause confusion.	3.83	0.38	Strongly Agree
7.2. System updates rarely introduce new problems or bugs.	3.73	0.45	Strongly Agree
7.3. It is easy to find where to report a problem or request help.	3.81	0.40	Strongly Agree
7.4. Reported problems are typically resolved within a reasonable time.	3.81	0.40	Strongly Agree
7.5. Help resources (e.g., the chatbot) contain up-to-date information.	3.79	0.41	Strongly Agree
Overall Mean	3.79	0.41	Strongly Agree

Note. N = 52. Ratings on a four-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree).

Table 8 summarizes Maintainability, which averaged 3.79 across the five indicators. Update clarity (Item 7.1) received the highest score within the domain (M = 3.83, SD = 0.38). Items 7.3 and 7.4, concerning access to support channels and

timeliness of issue resolution, both scored 3.81. Item 7.2, concerning whether updates introduce new bugs, received the lowest score within the domain (M = 3.73, SD = 0.45).

4.8 Summary of Evaluation Results

A consolidated view of the seven ISO/IEC 25010 domain mean scores is presented in Table 9. All seven domains received Strongly Agree ratings, with domain means ranging from 3.73 to 3.81. Compatibility recorded the highest overall mean (M = 3.81), while Usability recorded the lowest (M = 3.73).

Table 9. Summary of Domain Mean Scores and Verbal Interpretations

Quality Domain	Mean	SD	Interpretation
Functionality Suitability	3.78	0.44	Strongly Agree
Performance Efficiency	3.79	0.44	Strongly Agree
Compatibility	3.81	0.43	Strongly Agree
Usability	3.73	0.46	Strongly Agree
Reliability	3.78	0.42	Strongly Agree
Security	3.76	0.46	Strongly Agree
Maintainability	3.79	0.41	Strongly Agree

Note. N = 52. Domain means are the averages of their constituent indicator items.

Across all items, standard deviations ranged from 0.36 to 0.56, indicating generally strong agreement among respondents. The two highest standard deviations were observed for page loading speed (Item 2.1, SD = 0.56) and access validation (Item 6.2, SD = 0.55), representing the most salient areas of variation in user experience.

5. DISCUSSION

The evaluation findings support the conclusion that SAMS effectively addresses the principal weaknesses of the manual merit assessment process at CEU Manila. With all seven ISO/IEC 25010 domains receiving Strongly Agree ratings, the data indicate that the system meets accepted software quality benchmarks and is perceived positively by its intended users. The following discussion interprets each domain-level finding in relation to the existing literature and highlights implications for subsequent development.

5.1 Functional Adequacy and the Role of the Chatbot

The strong ratings for functionality (M = 3.78) affirm that SAMS's core features—particularly document submission and merit computation—operate reliably. These results are consistent with prior research on rule-based performance evaluation systems, which has shown that predefined criteria and automated scoring reduce evaluator variability and accelerate processing (Alfonso et al., 2022; Betsurmath et al., 2024). The chatbot component, while also receiving a positive rating, exhibited greater variation in user experience (Item 1.3, SD = 0.52), which aligns with findings reported by Dawood (2024) and Manasa (2024) that chatbots often perform unevenly when users pose informal or domain-specific queries outside the predefined intent set. This supports the need for ongoing refinement of the chatbot's knowledge base and the eventual introduction of hybrid escalation pathways in which complex queries are routed to human advisors—a strategy consistent with the recommendations of Chang et al. (2023).

5.2 Performance and Compatibility as System Strengths

Performance Efficiency (M = 3.79) and Compatibility (M = 3.81) emerged as particular strengths. Near-unanimous confidence in concurrent-user performance (Item 2.6, SD = 0.36) suggests that the server architecture is well calibrated for the anticipated load at CEU Manila. Similarly, high ratings for cross-browser and cross-device compatibility indicate that SAMS performs reliably across the heterogeneous hardware typically used by students. These outcomes are consistent with the emphasis in the digital assessment literature on accessibility and infrastructure considerations (Saltos-Rivas et al., 2022; Tkachenko et al., 2024). The higher variability observed for page loading speed (Item 2.1, SD = 0.56) is likely attributable to external factors such as network conditions and device capability, rather than to system-level deficiencies; nevertheless, this finding points toward performance optimization, particularly for initial page rendering, as a worthwhile target for future iterations.

5.3 Usability as the Primary Area for Refinement

Although Usability remained within the Strongly Agree range (M = 3.73), it recorded the lowest domain mean and therefore represents the most salient opportunity for improvement. Item 4.1, which addressed overall ease of navigation, was the lowest-rated item in this domain (M = 3.71), suggesting that initial interactions with the interface may still present friction for some

users. This finding aligns with prior research emphasizing the central role of intuitive design in digital assessment platforms (Idiyatova et al., 2024; Osabutey et al., 2022). Targeted improvements—such as streamlining the primary navigation structure, strengthening button affordance, and reducing the number of steps required to complete common tasks—are likely to yield the most substantial usability gains. Importantly, the observed variability in usability experience may reflect differences in prior digital tool familiarity among users rather than fundamental design flaws, underscoring the complementary role of the chatbot in scaffolding digital literacy.

5.4 Reliability and Data Integrity

Reliability ratings (M = 3.78) demonstrate that respondents generally perceive SAMS as stable and dependable. The strong ratings for system availability (Item 5.1, M = 3.81) and minimal errors (Item 5.4, M = 3.81) indicate that the system's backend data-handling mechanisms are functioning as intended. Minor disruptions reflected in Item 5.2 are consistent with the intermittent, network-dependent behaviors common to browser-based applications. These findings echo the observations of Ober et al. (2023) and Faber et al. (2022), who noted that digital monitoring systems deliver their greatest value when accompanied by robust server monitoring and proactive maintenance protocols.

5.5 Security and the Need for Enhanced Access Controls

The Security domain yielded an overall mean of 3.76, with the role-based access controls operating effectively in the majority of user interactions. However, the elevated standard deviation observed for Item 6.2 (SD = 0.55) merits particular attention: it suggests that while most users are confident in SAMS's access control mechanisms, a meaningful minority harbors uncertainty or has encountered experiences that reduce that confidence. Strengthening access validation through session timeout enforcement, detailed access logging, and an eventual pathway toward multi-factor authentication would reduce this variability and enhance overall security confidence. These recommendations are consistent with the broader literature on performance monitoring ethics and data protection (Vinith & Pinto, 2023; Widayanto et al., 2021).

5.6 Maintainability and Post-Deployment Sustainability

Maintainability ratings (M = 3.79) indicate that users are generally satisfied with update communication, support accessibility, and the timeliness of issue resolution. The relatively lower rating for Item 7.2 (M = 3.73)—concerning whether updates introduce new bugs—highlights a challenge common to iterative software development, in which even well-tested updates can produce unexpected behaviors in production environments. Strengthening pre-release testing protocols and introducing more structured regression testing would mitigate this risk and support the long-term sustainability of the system. The broader literature on digital assessment platforms emphasizes that post-deployment support quality significantly shapes long-term user trust (Rahim et al., 2022; Sowmya et al., 2024).

5.7 Synthesis and Implications

Taken together, the findings indicate that SAMS is a well-developed digital platform that delivers consistent, reliable, and secure service across diverse user contexts. The combination of rule-based automation, multi-tiered validation, AI-assisted guidance, and role-based authentication provides a coherent response to the documented weaknesses of manual merit evaluation. The results contribute empirical support to the broader scholarly argument that effective digital merit systems are sociotechnical in nature: their success depends not only on technical sophistication but also on transparency, accessibility, and institutional alignment (Franţuzan et al., 2023; Kalischko & Riedl, 2021). For higher education institutions considering comparable initiatives, the present study provides a concrete example of how structured quality evaluation can inform iterative refinement and support evidence-based deployment decisions.

6. CONCLUSION

This study developed and evaluated the Student Affairs Merit System (SAMS), a web-based platform designed to digitalize merit documentation, submission, validation, and evaluation for co-curricular and extracurricular activities at CEU Manila. Using a descriptive-developmental research design and the ISO/IEC 25010 software quality framework, the study assessed the system across seven quality domains through a survey of 52 respondents.

The findings indicate that SAMS performs strongly across all evaluated dimensions. Domain mean scores ranged from 3.73 to 3.81 on a four-point Likert scale, with every domain receiving a Strongly Agree interpretation. Compatibility achieved the highest rating (M = 3.81), reflecting consistent performance across devices, browsers, and operating systems. Usability received the lowest rating (M = 3.73) but remained within the Strongly Agree range, identifying interface navigation as the primary target for refinement. These results confirm that SAMS represents a meaningful improvement over manual merit evaluation, providing a more reliable, efficient, and transparent alternative that aligns with CEU Manila's institutional commitment to technology-enabled student affairs management.

6.1 Limitations

Several limitations should be acknowledged. First, the sample size of 52 respondents, while sufficient for descriptive analysis, limits the generalizability of the findings to the broader CEU Manila student population and to other higher education institutions. Second, the study was conducted over a single evaluation cycle during beta deployment; longitudinal evidence of system performance and user satisfaction is not yet available. Third, although the ISO/IEC 25010 framework provides a robust structure for evaluating software quality, it does not fully capture all dimensions of student engagement or institutional

adoption dynamics. Finally, the chatbot component was evaluated as part of the overall system; a dedicated, standalone evaluation of chatbot accuracy and user satisfaction remains an area for future research.

6.2 Recommendations

Based on the evaluation findings, several priorities are recommended for subsequent development cycles. Interface refinement should be treated as the most urgent improvement target, with emphasis on simplifying the primary navigation structure, improving button affordance and visual hierarchy, and reducing the number of steps required to complete core tasks. Performance optimization—particularly for initial page loading under diverse network conditions—would address the variability observed in the Performance Efficiency domain. Security enhancements, including session timeout enforcement, detailed access logging, and a pathway toward multi-factor authentication, would reduce the variability reflected in access validation ratings.

The AI chatbot's knowledge base should be expanded and regularly updated, with misclassification rates monitored systematically to guide retraining cycles. Hybrid escalation pathways—in whichcomplex queries that the chatbot cannot resolve confidently are routed to human advisors—would significantly improve chatbot reliability in edge cases. At the institutional level, structured training sessions for raters and validators would further improve scoring consistency. Students should be consistently reminded, through the notification system and chatbot guidance, to submit documentation promptly and accurately. Formalizing SAMS's integration into CEU Manila's student affairs workflow, with clearly defined submission windows and validation deadlines, would maximize the system's utility and support smoother adoption across all departments. Finally, a longitudinal evaluation study tracking user satisfaction, merit computation accuracy, and student engagement over one or two academic years following full deployment would provide valuable evidence for further refinement and for institutional decision-making.

In summary, SAMS demonstrates the value of combining structured digital workflows, automated computation, and AI-assisted guidance to modernize student affairs processes. The system is ready for broader institutional deployment, and, with the targeted enhancements identified in this study, offers a robust foundation for equitable and efficient merit recognition at CEU Manila

Link to software demonstration:
https://tinyurl.com/ycxv8wya

REFERENCES

1. Aditi, A., Saxena, H., Prof, G., & Malik, A. (2024). AI-enhanced accessibility solutions using chatbot. International Journal of Scientific Research in Engineering and Management. https://doi.org/10.55041/ijsrem34701
2. Alfonso, L., Maaliw, R., Adao, R., Lagman, A., Juanatas, I., & Fernando-Raguro, M. (2022). Web-based performance evaluation system platform using rule-based algorithm. 2022 10th International Conference on Information and Education Technology (ICIET), 124–129. https://doi.org/10.1109/ICIET55102.2022.9778996
3. Arboleda-Velasquez, J. (2024). MERIT (Manuscript Evaluation Reflecting Intrinsic Tenets): A framework for efficient and equitable academic peer review. ScienceBank. https://doi.org/10.61340/rwng4653
4. Betsurmath, C., Mamatha, H., Chidambaram, S., & Prashant, V. (2024). Design, development and validation of a digital performance-based appraisal system. Indian Journal of Science and Technology. https://doi.org/10.17485/ijst/v17i6.2711
5. Busso, M., Montaño, S., Muñoz-Morales, J., & Pope, N. (2024). The unintended consequences of merit-based teacher selection: Evidence from a large-scale reform in Colombia. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4958544
6. Chang, D., Lin, M., Hajian, S., & Wang, Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning: Goal setting, feedback, and personalization. Sustainability, 15(17), 12921. https://doi.org/10.3390/su151712921
7. Curzi, Y., & Ferrarini, F. (2023). High-performance work systems and firm innovation: The moderating role of digital technology and employee participation. Management Research Review. https://doi.org/10.1108/mrr-11-2022-0751
8. Dawood, M. (2024). Assessing the effectiveness of chatbots in providing personalized academic advising and support in higher education. Studies in Technology Enhanced Learning, 4(1), 1–12. https://doi.org/10.21428/8c225f6e.7140f8f4
9. Dehankar, A., Kirpan, S., Jha, A., Thosare, G., Bhaisare, A., & Mankar, S. (2022). AI chatbot using Dialog Flow. International Journal of Computer Science and Mobile Computing. https://doi.org/10.47760/ijcsmc.2022.v11i05.006
10. Dhina, M., Hadisoebroto, G., Setiawati, I., Undang, G., Masluh, M., & Mubaroq, S. (2021). Digital performance assessment: Measuring a pharmacy physics laboratory's skill. Journal of Physics: Conference Series, 1869, 012099. https://doi.org/10.1088/1742-6596/1869/1/012099
11. Duarte-Vidal, L., Herrera, R., Atencio, E., & Rivera, M. (2021). Interoperability of digital tools for the monitoring and control of construction projects. Applied Sciences, 11(21), 10370. https://doi.org/10.3390/app112110370

12. Faber, J., Feskens, R., & Visscher, A. (2022). A best-evidence meta-analysis of the effects of digital monitoring tools for teachers on student achievement. School Effectiveness and School Improvement, 34, 169–188. https://doi.org/10.1080/09243453.2022.2142247
13. Firoiu, D., Pîrvu, R., Jianu, E., Cismaș, L., Tudor, S., & Lățea, G. (2022). Digital performance in EU member states in the context of the transition to a climate-neutral economy. Sustainability, 14(6), 3343. https://doi.org/10.3390/su14063343
14. Franţuzan, L., Callo, T., & Bâlici, V. (2023). The strategic concept of meritology in learning. Revista Romaneasca pentru Educatie Multidimensionala. https://doi.org/10.18662/rrem/15.1/698
15. Funa, A. (2024). Digital badges as rewards in science education: Students' perceptions and experiences. Dalat University Journal of Science, 14(2). https://doi.org/10.37569/dalatuniversity.14.2.1212(2024)
16. Geshkov, M. (2021). Application of digital technologies in performance appraisal. Trakia Journal of Sciences. https://doi.org/10.15547/tjs.2021.s.01.016
17. Goulding, J., Twining, P., & Sharp, H. (2023). Awarding digital badges instead of grades. ASCILITE Publications. https://doi.org/10.14742/apubs.2023.599
18. Hartnett, M. (2021). How and why are digital badges being used in higher education in New Zealand? Australasian Journal of Educational Technology, 37(3), 104–118. https://doi.org/10.14742/AJET.6098
19. Huber, M., Heath, C., Baxter, C., & Reed, A. (2022). Using digital badges to design a comprehensive model for high-impact experiential learning. In Handbook of Research on Credential Innovations for Inclusive Pathways to Professions. IGI Global. https://doi.org/10.4018/978-1-7998-3820-3.ch021
20. Idiyatova, Y., Sadvakassova, A., & Mukhatayev, A. (2024). Modern approaches to the assessment of academic achievements in higher education institutions: A systematic review. Higher Education in Kazakhstan, 46(2), 25–37. https://doi.org/10.59787/2413-5488-2024-46-2-25-37
21. Kalischko, T., & Riedl, R. (2021). Electronic performance monitoring in the digital workplace: Conceptualization, review of effects and moderators, and future research opportunities. Frontiers in Psychology, 12, 633031. https://doi.org/10.3389/fpsyg.2021.633031
22. Ma'rup, M., Tobirin, & Rokhman, A. (2024). Utilization of artificial intelligence chatbots in improving public services: A meta-analysis study. Open Access Indonesia Journal of Social Sciences, 7(4), 1610–1617. https://doi.org/10.37275/oaijss.v7i4.255
23. Márquez, J., Lazcano, L., Bada, C., & Arroyo-Barrigüete, J. (2023). Class participation and feedback as enablers of student academic performance. SAGE Open, 13. https://doi.org/10.1177/21582440231177298
24. Mashilo, M., & Shekgola, M. (2024). Utilising artificial intelligence chatbots for student support at comprehensive open distance e-learning institutions in the Fifth Industrial Revolution. Journal of Education Society & Multiculturalism, 26–47. https://doi.org/10.2478/jesm-2024-0003
25. Mohd Rahim, N. I., Iahad, N. A., Yusof, A. F., & Al-Sharafi, M. A. (2022). AI-based chatbots adoption model for higher-education institutions: A hybrid PLS-SEM-neural network modelling approach. Sustainability, 14(19), 12726. https://doi.org/10.3390/su141912726
26. Naquita, M. P., & Balagtas, M. U. (2022). Use of interdisciplinary approach in performance-based assessment in the new normal. European Journal of Educational Research, 11(4), 2475–2486. https://doi.org/10.12973/eu-jer.11.4.2475
27. Ndunagu, J. N., Ezeanya, C. U., Onuorah, B. O., Onyeakazi, J. C., & Ukwandu, E. (2025). A chatbot student support system in open and distance learning institutions. Computers, 14, 96. https://doi.org/10.3390/computers14030096
28. Nikolova, N. (2024). Development of a system for assessment of digital professional suitability in the company environment. Environment. Technologies. Resources. Proceedings of the International Scientific and Practical Conference. https://doi.org/10.17770/etr2024vol3.8136
29. Nugroho, A., Rahayu, N., & Yusuf, R. (2023). The role of e-government to improve the implementation of merit system in Indonesian local governments. KnE Social Sciences. https://doi.org/10.18502/kss.v8i11.13570
30. Ober, T., Cheng, Y., Carter, M., & Liu, C. (2023). Leveraging performance and feedback-seeking indicators from a digital learning platform for early prediction of students' learning outcomes. Journal of Computer Assisted Learning, 40, 219–240. https://doi.org/10.1111/jcal.12870
31. Osabutey, E., Senyo, P., & Bempong, B. (2022). Evaluating the potential impact of online assessment on students' academic performance. Information Technology & People, 37, 152–170. https://doi.org/10.1108/itp-05-2021-0377
32. Patel, V., Cwalina, T., Gupta, A., Nortjé, N., Mullangi, S., Parikh, R., Shih, Y., & Hussaini, S. (2023). Oncologist participation and performance in the merit-based incentive payment system. The Oncologist, 28, e228–e232. https://doi.org/10.1093/oncolo/oyad033

33. Prange, A., & Sonntag, D. (2021). Assessing cognitive test performance using automatic digital pen features analysis. Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization. https://doi.org/10.1145/3450613.3456812
34. Saltos-Rivas, R., Novoa-Hernández, P., & Rodríguez, R. (2022). How reliable and valid are the evaluations of digital competence in higher education: A systematic mapping study. SAGE Open, 12. https://doi.org/10.1177/21582440211068492
35. Sousa-Vieira, M., Ferrero-Castro, D., & López-Ardao, J. (2021). Design, development and use of a digital badges system in higher education. Applied Sciences, 12(1), 220. https://doi.org/10.3390/app12010220
36. Sowmya, A. M., Sarkar, R. I., Singh, S., Sinha, S., & Manjunath, R. V. (2024). Smart virtual assistant academic model for a university website. International Journal of Innovative Science and Research Technology, 9(4). https://doi.org/10.38124/ijisrt/IJISRT24APR1711
37. Sun, Y., & Shi, Y. (2023). Problems and countermeasures in digital government performance evaluation. Social Security and Administration Management. https://doi.org/10.23977/socsam.2023.040806
38. Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. Journal of Intelligence, 11, 54. https://doi.org/10.3390/jintelligence11030054
39. Tkachenko, I., Tarasova, K., & Gracheva, D. (2024). Exploring the relationship between performance and response process data in digital literacy assessment. Journal of Modern Foreign Psychology. https://doi.org/10.17759/jmfp.2024130105
40. Topping, K. (2021). Digital peer assessment in school teacher education and development: A systematic review. Research Papers in Education, 38, 472–498. https://doi.org/10.1080/02671522.2021.1961301
41. Truskowska, E., Emmett, Y., & Guerandel, A. (2023). Digital badges: An evaluation of their use in a psychiatry module. The Asia Pacific Scholar. https://doi.org/10.29060/taps.2023-8-2/oa2869
42. Uanhoro, J., & Young, S. (2022). Investigation of the effect of badges in the online homework system for undergraduate general physics course. Education Sciences, 12(3), 217. https://doi.org/10.3390/educsci12030217
43. Velázquez-García, L., Longar-Blanco, M., Cedillo-Hernandez, A., & Bustos-Farías, E. (2024). Gamification in the classroom: Motivating higher education students using digital badges. International Journal of Learning and Teaching. https://doi.org/10.18178/ijlt.10.4.532-538
44. Vinith, K., & Pinto, D. (2023). Information technology in surveillance of employee performance. EPRA International Journal of Environmental Economics, Commerce and Educational Management. https://doi.org/10.36713/epra14255
45. Viphanphong, W., Limna, P., Kraiwanit, T., & Jangjarat, K. (2023). Merit piggy bank in the digital economy. Shanti Journal, 2(1). https://doi.org/10.3126/shantij.v2i1.53727
46. Wang, Y., Li, K., Luo, Y., Li, G., Guo, Y., & Wang, Z. (2024). Fast, robust and interpretable participant contribution estimation for federated learning. 2024 IEEE 40th International Conference on Data Engineering (ICDE), 2298–2311. https://doi.org/10.1109/ICDE60146.2024.00182
47. Wellberg, S., & Evans, C. (2022). Assumptions underlying performance assessment reforms intended to improve instructional practices: A research-based framework. Practical Assessment, Research, and Evaluation, 27(1), 23. https://doi.org/10.7275/pare.1334
48. Widayanto, M. T., Yuliana, I., Ismawati, & Natan, N. (2021). Implementation of performance assessment to determine employee performance. International Journal of Science, Technology & Management, 2(3), 81–88. https://doi.org/10.46729/ijstm.v2i5.302
49. Wu, W., Berestova, A., Lobuteva, A., & Stroiteleva, N. (2021). An intelligent computer system for assessing student performance. International Journal of Emerging Technologies in Learning, 16. https://doi.org/10.3991/IJET.V16I02.18739
50. Zhang, Q. (2024). An empirical study of Mythware platform on improving academic performance of college students. International Journal of Sociologies and Anthropologies Science Reviews. https://doi.org/10.60027/ijsasr.2024.4452
51. Zhu, S., Guo, Q., & Yang, H. (2023). Beyond the traditional: A systematic review of digital game-based assessment for students' knowledge, skills, and affections. Sustainability, 15(5), 4693. https://doi.org/10.3390/su1505469