ISSN: 2088-7043

JURNAL SISTEM INFORMASI

JOURNAL OF INFORMATION SYSTEM

Volume 22, Issue 2, October 2026

ISSN: 2088-7043

JOURNAL OF INFORMATION SYSTEM

JURNAL SISTEM INFORMASI

Volume 22, Issue 2, October 2026

Director:
Wisnu Jatmiko, Universitas Indonesia, Indonesia

Chief Editor:
Dana Indra Sensuse, Faculty of Computer Science, Universitas Indonesia, Indonesia

Board of Editor:
Achmad Nizar Hidayanto, Faculty of Computer Science, Universitas Indonesia
Betty Purwandari, Faculty of Computer Science, Universitas Indonesia
Harry Budi Santoso, Faculty of Computer Science, Universitas Indonesia
Indra Budi, Faculty of Computer Science, Universitas Indonesia
Juliana Sutanto, Faculty of Information Technology, Monash University, Australia
Kongkiti Phusavat, Department of Industrial Engineering. Faculty of Engineering, Kasetsart University, Thailand
Wade Goodridge, Utah State University, United States
Yu-Qian Zhu, Department of Information Management, National Taiwan University, Taiwan

Assistant Editor:
Fatimah Azzahro, Faculty of Computer Science, Universitas Indonesia, Indonesia
Nabila Clydea Harahap, Faculty of Computer Science, Universitas Indonesia, Indonesia
Panca O. Hadi Putra, Faculty of Computer Science, Universitas Indonesia, Indonesia

Address:
Faculty of Computer Science
Universitas Indonesia
Kampus Baru UI Depok, 16424, Indonesia
Telp. +62-21-7863419, Fax. +62-21-7863415
Email: jsi@cs.ui.ac.id
Website: http://jsi.cs.ui.ac.id/

Published by:
Faculty of Computer Science
Universitas Indonesia

ISSN: 2088-7043

JOURNAL OF INFORMATION SYSTEM

JURNAL SISTEM INFORMASI

Volume 22, Issue 2, October 2026

TABLE OF CONTENTS

1. Generative AI in Software Engineering Education: A Study on Competency and Perception
Ghozy Ghulamul Afif, Panca Oktavia Hadi Putra 1–17

2. User Experience of a Local Government Website: A Design Thinking and UEQ Approach
Muhammad Galih Wonoseto, Ghina Dersa, Ihwanikha Fazli Pratama, Aqib Khoiruzaman, Zahra Zakila Anindha Rahmanti, Adriel Devara Sandji, Eko Hadi Gunawan 18–31

3. Acceptance Analysis of the Eco-Routing Feature in Google Maps from the Perspective of TAM and Environmental Concern
Siti Nur Kholifah, Intan Sartika Eris Maghfiroh, Mochamad Chandra Saputra, Almira Syawli 32–42

4. Examining Digital Egalitarianism through the Moderating Effect of Job Level on G2E Systems Adoption
Hanif Kridho Pangantyo, Mardiana Purwaningsih 43–63

5. Hybrid Cybersecurity Compliance Audit Framework for an On-Premise Hotel Property Management System: Mapping NIST CSF 2.0 Controls to Indonesia's Personal Data Protection Law
I Made Sudana, Bana Handaga, Jan Wantoro 64–80

Generative AI in Software Engineering Education: A Study on Competency and Perception

Ghozy Ghulamul Afif*
Faculty of Computer Science
Universitas Indonesia
Depok, Indonesia
ghozy.ghulamul@ui.ac.id

Panca Oktavia Hadi Putra
Faculty of Computer Science
Universitas Indonesia
Depok, Indonesia
hadiputra@cs.ui.ac.id

Abstract

Software Engineering (SE) education struggles to align graduate competencies with industry needs, a challenge compounded by Generative AI (GenAI). This study's purpose is to explore how structured GenAI utilization can enhance student competencies and learning perceptions. This exploratory case study used a mixed-method design involving four students in a controlled experiment. Data was collected via quantitative (SWEBOK-based artifact assessments, questionnaires) and qualitative narrative reflections. Results show the group using AI produced higher-quality artifacts and reported a more efficient, positive experience. A key finding is the transformation of work processes, shifting the primary challenge from manual execution to the critical evaluation of AI outputs. Due to its small scale, the findings are preliminary but suggest GenAI can enhance competencies. This requires an educational shift toward human-AI collaboration skills like critical evaluation and prompt engineering. Further validation with larger, more comprehensive studies is essential to generalize these initial insights.

Keywords: Generative AI, software engineering education, skills gap, SWEBOK, case study, human-AI collaboration, prompt engineering

Introduction

The global technology sector is navigating a period of notable adjustment. Recent trends, including widespread layoffs during the "tech winter," have created a complex job market (Afifa, 2023). Paradoxically, this volatility has only intensified a persistent skills gap, with a majority of business executives now predicting a significant talent shortage (Cantrell, 2022). This misalignment between the competencies of university graduates and the dynamic needs of the software industry has been a documented challenge in Software Engineering (SE) education for decades. Educators continuously struggle to develop relevant competencies—such as requirements analysis, architectural design, and complex problem-solving—that extend beyond mere coding proficiency (Garousi et al., 2020; Oguz & Oguz, 2019; Zönnchen et al., 2024).

This foundational challenge is now being reshaped by the emergence of Generative Artificial Intelligence (GenAI). The rapid adoption of GenAI tools in the industry is transforming software development workflows, with studies reporting significant productivity gains and a shift in the nature of developer tasks (Coutinho et al., 2024; Gao, 2024). This technological wave presents both a challenge

* Corresponding Author

and an opportunity for SE education. It presents challenges to traditional pedagogical approaches while also offering a powerful means to enhance learning, for instance, by helping students brainstorm requirements (Pereira et al., 2024), generate initial code (Zönnchen et al., 2024), or understand complex design concepts (Chen, 2024). It is important to move beyond speculation and empirically investigate how these tools can be leveraged effectively.

While the potential of GenAI in education is a topic of widespread discussion, there remains a notable gap in systematic, evidence-based research on its structured application within the SE curriculum. Current educational approaches face hurdles in keeping curricula relevant to rapid technological change (Kozov et al., 2024) and in providing scalable, personalized feedback to students (Sengul et al., 2024). Many existing studies on GenAI in this context are exploratory, focus on user perceptions, or are limited to basic programming tasks (Sengul et al., 2024). It is still largely unclear how the integration of GenAI as a core component of the development process—rather than just an auxiliary tool—affects specific student competencies as defined by established frameworks like the Software Engineering Body of Knowledge (SWEBOK).

This discrepancy between the rapid adoption of GenAI in industry and its cautious, often unstructured, application in higher education creates a notable tension. While industry professionals are leveraging these tools to accelerate development, academic institutions are left without clear, evidence-based guidance on how to prepare the next generation of software engineers for this new reality. This study addresses this research gap by examining the direct impact of structured GenAI use on student performance and learning experiences in a controlled educational setting. Therefore, the primary objective of this research is to explore how the structured integration of GenAI into software engineering coursework influences student competency development and their perceptions of the learning process.

To achieve this objective, this study is guided by the following central research question: *"How can the utilization of Generative AI in the Software Engineering learning process enhance the competencies of IT students to meet industry needs?"* By focusing on this question, the research aims not to provide a comprehensive solution to the entire skills gap, but to investigate the potential of GenAI as a strategic tool to address specific facets of it, such as accelerating the mastery of fundamental SE skills and exposing students to the latest technologies being adopted by the industry.

This study is expected to provide contributions to both theory and practice. Theoretically, it will offer a preliminary framework for understanding the role of human-AI collaboration in competency development within SE education. This aligns with a growing body of research indicating an evolving skill shift away from manual coding and toward strategic competencies like critical output evaluation and prompt engineering (Fan et al., 2023; Ozkaya & Schmidt, 2024; Zönnchen et al., 2024). Practically, the findings will provide educators and curriculum designers with initial, evidence-based insights into how they can effectively adapt their teaching methodologies for the GenAI era. By illuminating the benefits, challenges, and the "learning paradox"—where effective AI use requires strong foundational knowledge—this research aims to help shape a pedagogical approach that prepares students for the future of software engineering.

Literature Review

Definition and Fundamental Principles of SE

Software Engineering (SE) is formally defined as the application of a systematic, disciplined, and measurable approach to the development, operation, and maintenance of software (International Organization for Standardization, 2011). This discipline, as articulated by pioneers like Fritz Bauer and Ian Sommerville, emphasizes producing economically efficient and technically reliable software that encompasses all aspects of its production lifecycle (Bauer, 1972; Sommerville, 2011). Despite its status, SE is an evolving field, and unlike traditional engineering disciplines, it has not yet reached a universal consensus on its fundamental principles, reflecting an ongoing refinement of its theoretical foundations (Al-Sarayreh et al., 2021).

SWEBOK: The Map of Software Engineering Knowledge

A significant step toward standardizing competencies and knowledge in SE is the development of the Software Engineering Body of Knowledge (SWEBOK). Compiled as an official guide by the IEEE Computer Society and adopted as ISO/IEC TR 19759, SWEBOK establishes a generally accepted knowledge framework for the SE community. Its primary goal is to foster a uniform understanding of the discipline globally and to serve as a basis for educational curricula and professional certification (Washizaki et al., 2025).

Table 1. SWEBOK v4.0 KAs (Washizaki, 2024)

No	Knowledge Areas
1	Software Requirements
2	Software Architecture
3	Software Design
4	Software Construction
5	Software Testing
6	Software Engineering Operations
7	Software Maintenance
8	Software Configuration Management
9	Software Engineering Management
10	Software Engineering Process
11	Software Engineering Models and Methods
12	Software Quality
13	Software Security
14	Software Engineering Professional Practice
15	Software Engineering Economics
16	Computing Foundations
17	Mathematical Foundations
18	Engineering Foundations

The latest edition, SWEBOK v4.0 released in 2024, organizes the SE discipline into 18 main Knowledge Areas (KAs) that cover the entire software lifecycle and relevant engineering practices. These KAs range from requirements analysis to software maintenance, including managerial and quality assurance aspects, underscoring that SE is a complex and comprehensive discipline that extends far beyond coding. Table 1 lists these 18 Knowledge Areas.

Competency Gap

SE education aims to prepare students for industry careers by delivering the conceptual foundations of the software lifecycle and developing professional competencies through practical experience (Kirova et al., 2024). However, a persistent gap exists between the competencies of IT graduates and the actual needs of the software industry (Garousi et al., 2020; Oguz & Oguz, 2019). This gap is often attributed to factors such as curricula that are not aligned with industry dynamics, insufficient student experience in real-world projects, and suboptimal development of soft skills during the educational process (Garousi et al., 2020; Ozkaya & Schmidt, 2024).

SE competency itself is a crucial combination of knowledge, skills, and personal attributes required by the industry for effective software development (Assyne et al., 2022b, 2022a; Rivera-Ibarra et al., 2010). This includes technical mastery aligned with SWEBOK knowledge areas and essential soft skills like

problem-solving, communication, and teamwork (Assyne et al., 2022a; Garousi et al., 2020; Kirova et al., 2024). The rise of GenAI is predicted to shift competency needs, demanding that professionals be able to analyze, validate, and integrate AI-generated work meticulously (Pereira et al., 2024; Zönnchen et al., 2024). The focus may shift from manual coding to higher-level tasks such as design, requirements engineering, and testing of AI outputs (Zönnchen et al., 2024). This shift adds another layer of complexity to the challenge of bridging the existing skills gap, underscoring the urgency for curriculum adaptation in SE education (Ozkaya & Schmidt, 2024).

Generative Artificial Intelligence (GenAI)

Generative Artificial Intelligence (GenAI) is technology capable of producing new and original content such as code, designs, or documentation in Software Engineering (SE) (Pereira et al., 2024). This technology is being adopted to increase productivity and simplify the development process (Coutinho et al., 2024), with applications spanning various stages from requirements elicitation, diagram creation, and API simulation to code writing, explanation, testing, and documentation (Pereira et al., 2024).

While tools like GitHub Copilot and ChatGPT are often used to support these activities (Coutinho et al., 2024), the effectiveness of their use heavily depends on user interaction, where careful prompt design (Pereira et al., 2024) and the ability to formulate precise instructions are key to obtaining desired results (Coutinho et al., 2024). The generated outputs often require manual refinement and critical validation by the user (Coutinho et al., 2024; Kuhail et al., 2024; Pereira et al., 2024) due to significant challenges. These challenges include issues of reliability and accuracy (Coutinho et al., 2024; Pereira et al., 2024), the potential for hallucination or generating misleading information (Pereira et al., 2024), the risk of over-reliance that can hinder skill development (Pereira et al., 2024), potential biases in the models (Pereira et al., 2024), and concerns related to security and privacy when handling sensitive data (Coutinho et al., 2024). Therefore, understanding both the potential and the challenges, along with the importance of skilled interaction, is essential for the effective use of GenAI in SE.

Methodology

Research Design

This study employs an explanatory mixed-method design, which is well-suited for research that aims to first identify quantitative outcomes and then explain those results with in-depth qualitative data (Creswell, 2014). The research was structured as an evaluative case study involving an experimental design. Four third-year (6th semester) IT students were selected through purposive sampling and divided into two groups: a control group (manual work) and an experimental group (using Generative AI).

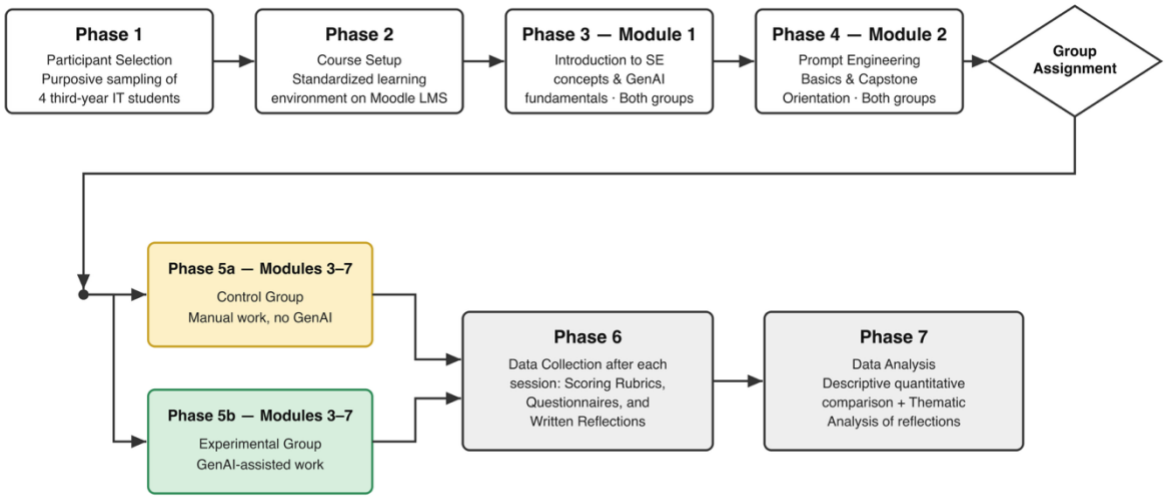


Figure 1. Research Process Flowchart

Figure 1 illustrates the overall research process from beginning to end. The study was organized into seven phases. It began with participant selection and the assignment of students to control and experimental groups, followed by the preparation of the course on the Moodle LMS. In Module 1, both groups were introduced to fundamental concepts in Software Engineering (SE) and Generative AI (GenAI). Module 2 focused on prompt engineering and provided guidance for the capstone project.

The main project activities were conducted across five sessions (Modules 3–7). During these sessions, the control group completed the tasks manually, while the experimental group was allowed to use GenAI tools. Each session was aligned with a specific SWEBOK Knowledge Area. After every session, data were collected using scoring rubrics, questionnaires, and students' written reflections. The final stage consisted of a two-phase analysis that combined descriptive quantitative comparisons with thematic analysis of the qualitative data. The figure also indicates the stage where the two groups began to differ, starting from Module 3, and the point where their data were brought together again for the final analysis.

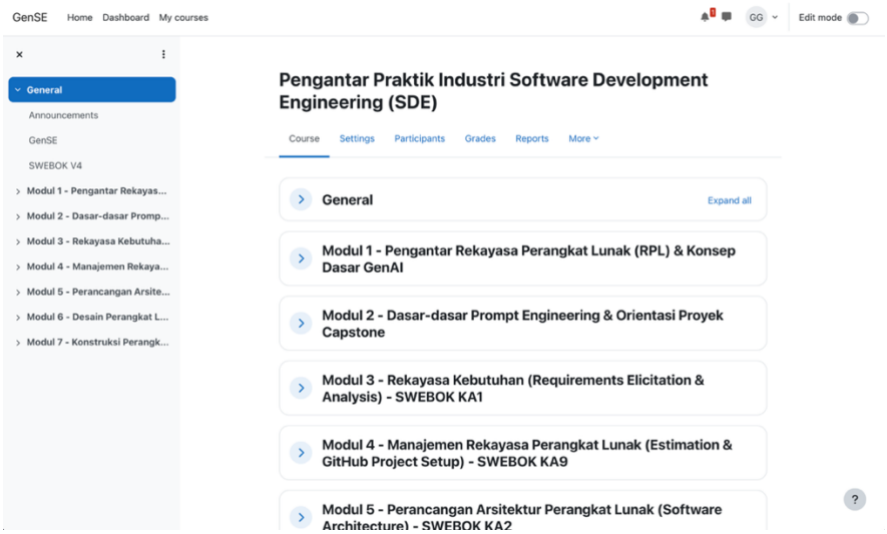


Figure 2. Structured Learning Environment on Moodle

The independent variable was the use of GenAI, while the dependent variables were the quality of the software engineering artifacts produced and the students' perceptions of the learning experience. The entire experiment was facilitated through a dedicated course on the Moodle Learning Management System (LMS) to ensure standardized delivery of materials, instructions, and data collection across both groups. The research process followed a sequential structure, beginning with two preparatory modules and followed by five core project sessions aligned with key SWEBOK Knowledge Areas: Software Requirements, SE Management, Software Architecture, Software Design, and Software Construction. Figure 2 shows the resulting structured learning environment on Moodle.

This study is positioned as an exploratory, education-oriented case study that aims to identify emerging patterns in students' learning processes and competency development when Generative AI is integrated into Software Engineering coursework. Given the small sample size and the pedagogical focus of the research, the study does not seek statistical generalization. Instead, it emphasizes in-depth analysis of observed tendencies, supported by descriptive quantitative results and qualitative reflections, in line with established practices in computing education research.

Research Instruments

To align with the mixed-method research design, both quantitative and qualitative data collection instruments were employed as follows:

1. **Scoring Rubrics**

   Structured rubrics were developed before data collection began for each of the five project sessions. Each rubric translated the relevant SWEBOK competency into a set of specific criteria

rated on a five-point scale, with distinct behavioral descriptors defined for every score level. The criteria addressed aspects such as the completeness and testability of acceptance criteria, the quality of architectural justification, and the functional correctness of the implemented code.

All artifacts were scored by the primary researcher. The researcher holds a graduate degree in Computer Science with a focus on Software Engineering and has practical experience in software development across the knowledge areas being assessed. Only one scorer was involved, which is acknowledged as a limitation later in this paper. To maintain consistency in scoring, three measures were applied. First, the rubric descriptors were finalized before any student work was reviewed so that the evaluation criteria remained fixed throughout the study. Second, a written justification was recorded for each assigned score at the time of assessment, providing a clear rationale for every evaluation decision. Third, each artifact was assessed against the rubric criteria independently rather than through comparison with other participants' work. Because the study used a single scorer, no inter-rater reliability coefficient was calculated. This limitation is discussed further in a later section.

2. **Learning Experience Questionnaire**

   A Likert-scale questionnaire was administered after each session to gather quantitative data on students' subjective learning experiences. The instrument captured responses related to perceived task difficulty, confidence in the quality of their outputs, satisfaction with the development process, perceived efficiency, and comprehension of key software engineering concepts.

3. **Written Reflections**

   To obtain qualitative insights, participants were invited to complete a guided reflection consisting of open-ended questions after each session. These reflections solicited narrative accounts regarding the challenges encountered, problem-solving strategies adopted, and critical takeaways from the learning experience. The qualitative data served to contextualize and deepen the interpretation of the quantitative results. Figure 3 presents the written reflection prompt used for this data collection.



Figure 3. Written Reflection Prompt for Data Collection

Data Collection

Data collection was conducted systematically through the Moodle platform following the completion of each of the five core project sessions (Session 3 to Session 7). This centralized approach ensured that all data was collected in a standardized manner.

The quantitative data collection involved two primary sources. First, the software engineering artifacts (e.g., Product Requirements Document, architecture diagrams, source code) produced by each student

were uploaded to designated submission slots on Moodle. Second, responses to the learning experience questionnaire were collected directly through the platform's questionnaire feature at the end of each session.

The qualitative data were gathered from the written narrative reflections submitted by participants through Moodle. This method allowed for the consistent collection of in-depth insights into each student's thought process and experience throughout the project.

Data Analysis

The data analysis was conducted in two phases that together reflect the explanatory mixed-method design of this study. In the first phase, the quantitative data, consisting of artifact scores and questionnaire responses, were examined. An important consideration is that each group included only two participants. Because of this small sample, applying standard inferential statistical tests would be inappropriate. For this reason, the analysis relies on descriptive comparison rather than hypothesis testing. Each participant's scores were averaged across the five sessions to produce a single aggregate value per person, and these values were then compared between groups. To provide a standardized indication of the magnitude of the observed difference, Hedges' g was calculated using participant-level mean scores ($g \approx 3.48$), indicating a very large descriptive separation between groups. However, given the extremely small sample size ($n = 2$ per group), this value should be interpreted only as an exploratory magnitude indicator rather than a stable estimate of population effect size. Accordingly, no p-values are reported, because inferential testing would require treating repeated session scores as independent observations, which would violate the assumption of independence and risk pseudoreplication.

In the second phase, the written reflections were analyzed using Thematic Analysis. The responses were reviewed, coded, and grouped into recurring patterns that captured how students in each group experienced the project, including the challenges they encountered, the strategies they used, and the lessons they drew from each session. The resulting themes were then used to help interpret the quantitative findings from the first phase, linking the numerical patterns in the scores to the underlying learning experiences that may have produced them.

Result

Descriptive Overview of Group Performance

Before discussing the results by knowledge area, a brief note on interpretation is necessary. With only two participants per group, none of the comparisons here should be read as statistically conclusive. The numbers are reported because they reveal a consistent directional pattern worth examining, but readers should treat them as preliminary observations from a small exploratory study rather than evidence that generalizes beyond these four students.

When each participant's scores are averaged across the five sessions, P3 achieved a mean of 93.0, and P4 achieved 89.0, while P1 and P2 scored 74.0 and 69.0, respectively. Both experimental group members thus outperformed both control group members on this aggregate measure. Hedges' g computed from these four participant means was 3.48, suggesting a large observed difference (Hedges, 1981). However, with $n = 2$ per group, this value should be interpreted as descriptive only. Table 2 breaks down the scores session by session.

Table 3 presents each participant's average questionnaire score for each dimension across the five sessions (scale 1–5). A consistent pattern is evident. On every dimension, both P3 and P4 rated their experience higher than P1 and P2. The gap in completion time was also notable. Because P1's time for the Software Construction session was not recorded, time comparisons are based on the four sessions where complete data exists for all participants. On this basis, the control group averaged approximately 260 minutes per session (P1: 290 min; P2: 230 min), compared to around 76 minutes for the experimental group (P3: 76 min; P4: 76 min).

Table 2. Artifact Scores by Knowledge Area and Participant

Knowledge Area	P1 (Control)	P2 (Control)	Control Avg	P3 (Exp.)	P4 (Exp.)	Exp. Avg	Difference
Software Requirements	100	80	90.0	100	100	100.0	+10.0
SE Management	70	90	80.0	100	70	85.0	+5.0
Software Architecture	80	30	55.0	80	80	80.0	+25.0
Software Design	90	65	77.5	95	95	95.0	+17.5
Software Construction	30	80	55.0	90	100	95.0	+40.0
Overall Mean	**74.0**	**69.0**	**71.5**	**93.0**	**89.0**	**91.0**	**+19.5**

Table 3. Per-Participant Average Questionnaire Scores (Scale 1–5)

Dimension	P1 (Control)	P2 (Control)	Control Avg	P3 (Exp.)	P4 (Exp.)	Exp. Avg
Perceived Task Difficulty	1.4	1.8	1.6	3.4	3.8	3.6
Confidence in Output Quality	1.0	1.8	1.4	4.4	4.2	4.3
Process Satisfaction	2.4	1.8	2.1	4.4	4.2	4.3
Process Efficiency	1.6	1.8	1.7	4.2	4.6	4.4
Core Concept Understanding	2.2	1.8	2.0	3.8	4.4	4.1

Taken together, the descriptive data points in the same direction across both instruments. The experimental group produced higher-quality artifacts and reported a more positive learning experience throughout the sessions. The difference was modest in the earlier, more structured activities and became more pronounced as the tasks grew more technically demanding, particularly in Software Architecture, Software Design, and Software Construction. The sections below examine each knowledge area in detail, drawing on the qualitative reflections to explain what drove these patterns.

Artifact Quality and Performance

The primary quantitative findings are derived from the descriptive comparison of software engineering artifact quality across five SWEBOK Knowledge Areas (KAs). An aggregate review of rubric scores reveals a consistent tendency for participants in the GenAI-supported group to produce artifacts with higher average scores across most assessed competencies compared to the control group. These results reflect observable patterns rather than statistically significant differences and are interpreted within the exploratory scope of the study. Figure 4 compares the competency profiles of the individual participants.

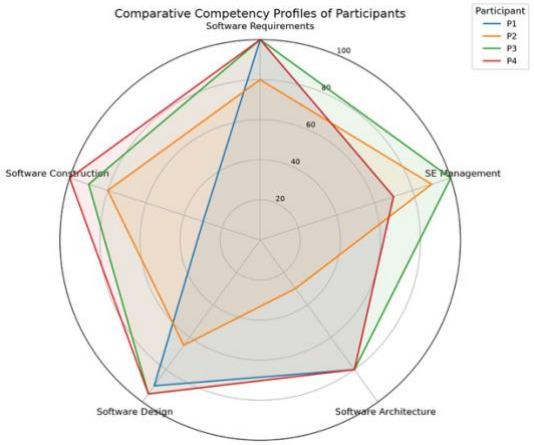


Figure 4. Comparative Competency Profiles of Participants

A breakdown of performance by each SWEBOK Knowledge Area provides more specific insights, integrating qualitative reflections to explain the quantitative scores. Figure 5 summarizes the average artifact quality scores by knowledge area.

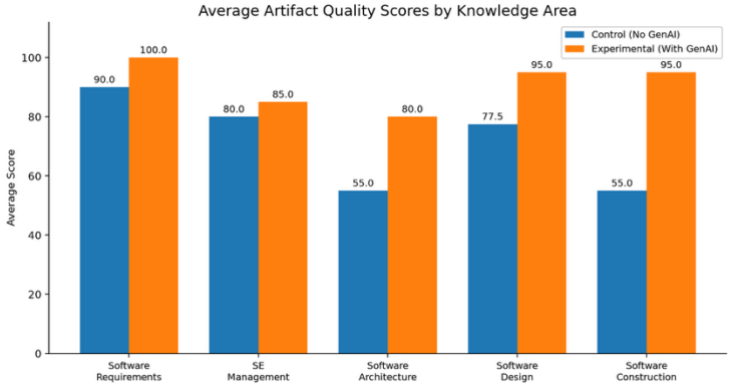


Figure 5. Average Artifact Quality Scores by Knowledge Area

1. **Software Requirements (KA1)**

   The experimental group achieved consistently optimal scores, while the control group's performance was varied. This discrepancy is explained by the different strategies employed. The control group relied on manual cognitive effort, with one participant noting the primary challenge was *"...composing clear and unambiguous Acceptance Criteria and Non-Functional Requirements"* (P2, Control). In contrast, the experimental group leveraged GenAI as a core strategy. One participant stated their process was simply *"Using GenAI"* (P4, Experimental), which standardized their output at a high level of quality by assisting with tasks like brainstorming and identifying edge cases.

2. **Software Engineering Management (KA9)**

   Performance in this area was highly heterogeneous, with the GenAI intervention not guaranteeing superior results. The key challenge for both groups was justifying effort estimation, a task requiring experience. The experimental group's reliance on AI for this task proved to be a "black box." One participant (P3, Experimental) admitted, *"...even though GPT gave an estimate, I couldn't justify it,"* while another (P4, Experimental) stated, *"The determining factors: GenAI decided them here."* This lack of personal rationale explains why their estimation scores were lower than a high-performing control group participant who relied on experience.

3. **Software Architecture (KA2)**

   The results were divergent. The experimental group excelled in providing detailed justifications for technology choices, using AI as a research assistant to compare options. However, a control group participant (P1) produced the highest quality architectural diagram. The experimental group's reflections reveal the limitations of AI in this visual task, with one noting, *"...using ChatGPT to help make the diagram—the format was often inconsistent..."* (P3, Experimental). This suggests that for tasks requiring formal visual abstraction, individual capability remains dominant.

4. **Software Design (KA3)**

   The experimental group showed a slight overall superiority, driven by notably higher scores in creating Sequence Diagrams. The qualitative data reveal a fundamental difference like the challenge. The control group experienced a high cognitive load, which one participant described as *"traumatizing"* (P1, Control). The experimental group, however, faced a different challenge: evaluation. Their advantage in creating Sequence Diagrams stemmed from an AI-driven reverse-engineering strategy, where they prompted the AI to analyze existing code to generate the diagram automatically.

5. **Software Construction (KA4)**

   A clear performance hierarchy emerged, with the experimental group demonstrating markedly superior outcomes. GenAI acted as a powerful accelerator for participants who already possessed foundational skills. This was articulated by one participant (P3, Experimental) who noted, *"...the development process using LLM was faster, especially... the Copilot agent could read the code structure."* This allowed them to focus on business logic rather than boilerplate code. Conversely, for participants with weaker foundational skills, the time saved on writing code was lost in the effort required to integrate and revise the AI's output, diminishing the tool's effectiveness.

Student Perceptions and Learning Experience

Data from the post-session questionnaires revealed notable differences in the subjective learning experience between the two groups across several dimensions.

1. **Time and Efficiency**

   The experimental group consistently reported requiring substantially less time to complete tasks. For instance, in the SE Management session, the control group spent an average of 245 minutes, whereas the experimental group spent only 50 minutes. This trend was consistent across all sessions. This pattern is illustrated in Figure 6 and Figure 7.

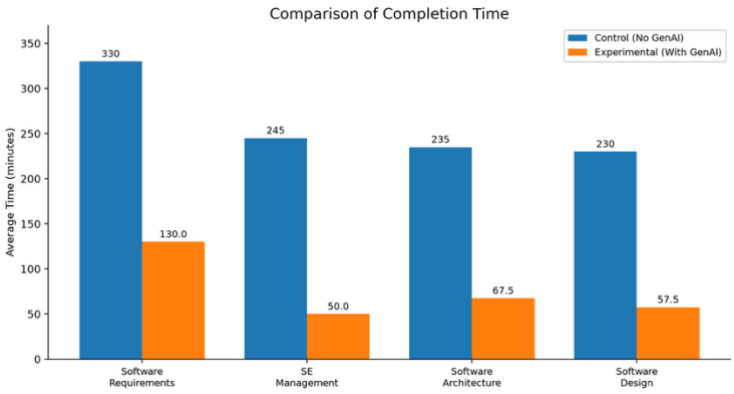


Figure 6. Comparison of Completion Time

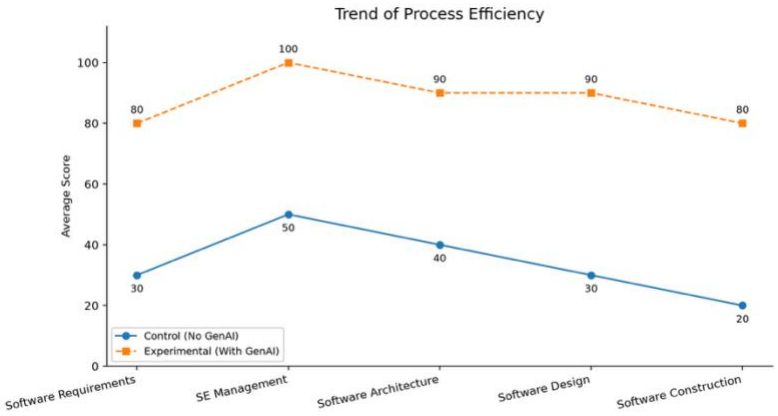


Figure 7. Trend of Process Efficiency

2. Confidence and Difficulty

Participants using GenAI reported substantially higher confidence in the quality of their work and perceived the tasks as less difficult compared to the control group. These trends are shown in Figure 8 and Figure 9.

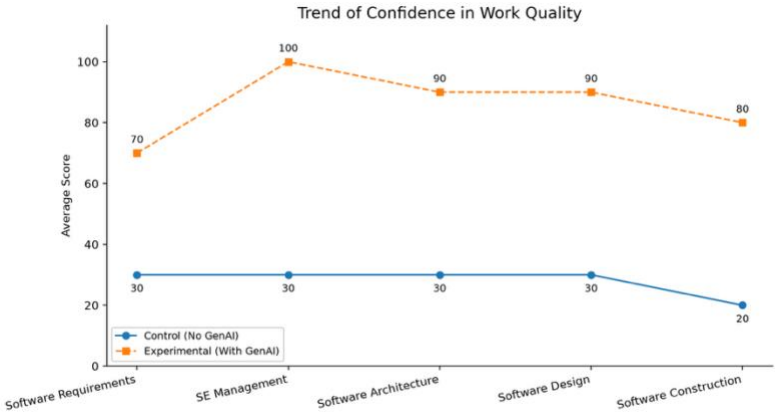


Figure 8. Trend of Confidence in Work Quality

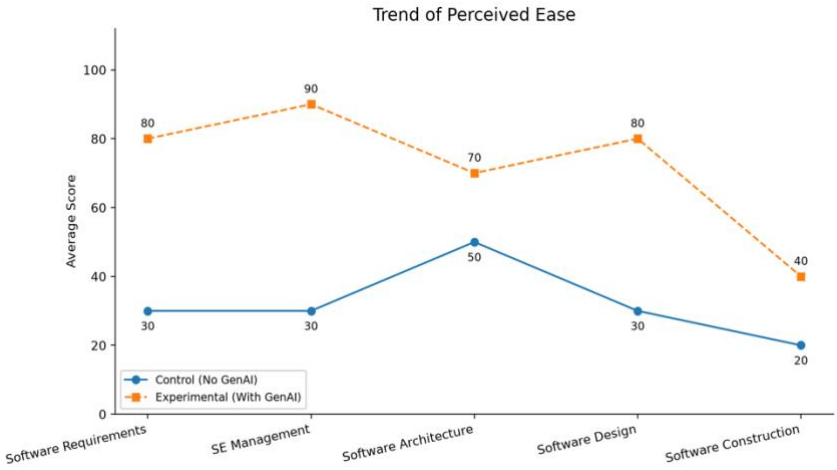


Figure 9. Trend of Perceived Ease

3. Concept Comprehension

The experimental group consistently reported a high level of understanding of the core concepts for each session. The control group's self-reported understanding was lower and more variable, particularly in the more complex design and construction phases. Figure 10 depicts this trend in core concept understanding.

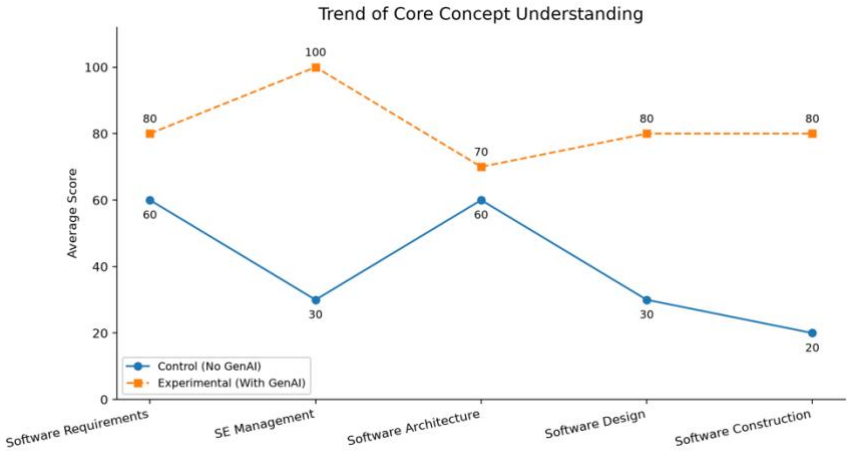


Figure 10. Trend of Core Concept Understanding

Theme 1: Paradigm Shift in Work Strategy: Manual vs. AI-Orchestrated

A fundamental difference in work strategy was evident between the groups. The control group relied on manual cognitive processes, involving independent research and analogical reasoning to complete tasks. This placed the entire analytical burden on the individual.

"...looking at examples... searching for precedents from similar applications... [doing] additional research..." - P1 (Control)

In contrast, the experimental group integrated GenAI as a core component of their workflow, transforming the process from manual execution to AI-assisted orchestration. This advanced strategy included using AI for reverse engineering, which helps explain the clear descriptive performance difference observed in the quantitative scores for complex tasks like Software Design and Construction.

"...asked the agent copilot to access the source code and create an ERD, Sequence Diagram... based on the source code..." - P4 (Experimental)

Theme 2: Transformation of Challenges: From Burden of Execution to Burden of Evaluation

The nature of the primary challenge faced by the participants was fundamentally different. The control group struggled with the burden of execution—the high cognitive load required to create complex artifacts from scratch. This was acutely felt in the design phase.

"[The design task was] traumatizing because of the high complexity and precision required in the manual process." - P1 (Control)

Conversely, the experimental group's main challenge shifted to the burden of evaluation. Their difficulty was not in generating content, but in critically assessing, contextualizing, and validating the output provided by the AI. This explains why their performance was not uniformly perfect, particularly in the SE Management task, where analytical justification was paramount.

"...creating the diagram... wasn't hard... the hard part was having to evaluate the results/answers from GenAI." - P4 (Experimental)

Theme 3: The Learning Paradox: Efficiency vs. Depth of Understanding

A recurring theme was the perceived trade-off between the efficiency gained from using AI and the depth of understanding acquired. Participants acknowledged that a manual process, while slower, often felt like it led to a more profound comprehension. However, they also realized that using AI effectively was not a shortcut that bypassed the need for knowledge. Instead, a strong foundational understanding was a prerequisite to guide the AI and validate its output, preventing it from becoming a passive user.

"...if we can't understand the code, we will surely be easily replaced by GenAI in the future." - P4 (Experimental)

Theme 4: Redefinition of Engineer Competencies

Across their reflections, participants described a shift in how they understood their role as engineers, moving away from the view of a developer as someone who manually produces outputs toward a more evaluative and coordinating role. Figure 11 presents the four competencies that emerged from this theme and illustrates how they relate to one another. The arrows are not merely decorative. They represent a working cycle. When working with AI, an engineer first provides context, since without it the AI tends to generate generic and often unusable outputs. This context then informs the prompt, introducing the second competency, prompt design. The AI subsequently generates an output, after which the third competency, critical validation, becomes necessary to review, correct, and determine what should be retained. Overarching these steps is the fourth role, ethical guardianship, which functions less as a separate stage and more as an ongoing responsibility that influences decisions throughout the cycle. The arrows between the components are bidirectional to indicate that the process is iterative. If validation fails, the engineer may return to refine the prompt or reconsider the context until the output meets the required standard.

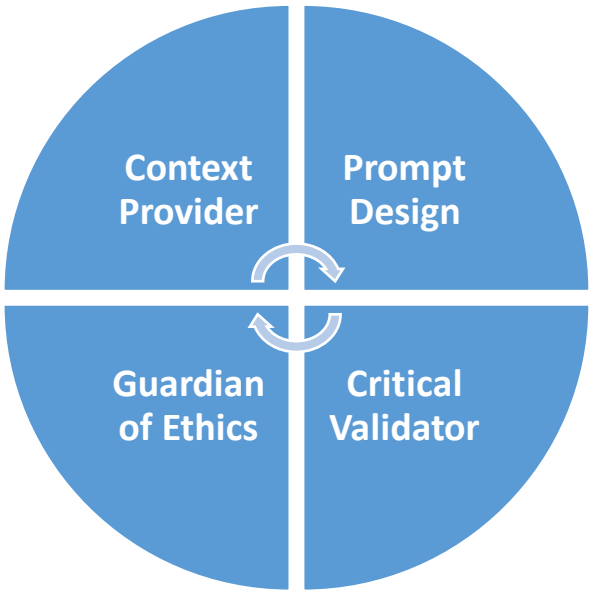


Figure 11. Four Key Competencies of Software Engineers in the AI Era

Each competency is elaborated below, supported by participants' own reflections:

1. **Context Provider**

   Successful interaction with AI was found to be highly dependent on providing precise and complete problem context. Participants realized that without it, AI tends to produce generic or

unsuitable output. As one participant concluded, *"...the thinking must still come from the programmer, because the [AI's] context is limited"* (P4, Experimental).

2. **Prompt Design**

   The ability to design effective and clear commands, or prompts, was recognized as key to obtaining desired results from the AI. One participant explicitly mentioned the challenge of *"...composing a clear prompt..."* as a primary difficulty when interacting with the LLM.

3. **Critical Validator**

   The first and most fundamental competency is the ability to not accept AI output at face value, but to actively validate, test, and refine it. This was described as the most difficult part of the process, with one participant stating, *"...the hard part is having to evaluate the results/answers from GenAI..."* (P4, Experimental). This competency shifts the engineer's role from "creator" to "curator" and "quality assurer".

4. **Guardian of Ethics**

   Finally, participants recognized a broader role that includes the responsibility to reshape *"...ethical standards..."* in using AI technology responsibly. These four pillars collectively redefine the profile of a software engineer who is relevant and ready to face the challenges of the modern industry.

Discussion and Implications

This study demonstrates that integrating Generative AI (GenAI) into software engineering coursework fundamentally shifts students' focus from routine artifact creation toward critical evaluation and higher-order design activities. Rather than simply automating coding or modeling tasks, GenAI tools serve as catalysts for deeper engagement: students allocate saved time to validate AI outputs, refine architectural decisions, and articulate rationale for their design choices (Fan et al., 2023).

A key novel insight is the "learning paradox": while GenAI accelerates production of deliverables, it simultaneously demands stronger conceptual understanding to detect and correct AI-introduced inaccuracies. In educational settings, this paradox implies that GenAI adoption without scaffolded support may inadvertently mask learning gaps (Zönnchen et al., 2024). Accordingly, curricula should be redesigned to balance tool-driven efficiency with structured opportunities for manual problem solving and error analysis (Pereira et al., 2024).

Practically, educators can leverage these findings by:

1. Embedding prompt engineering exercises that guide students in formulating effective queries and interpreting AI responses (Coutinho et al., 2024).
2. Incorporating reflection checkpoints where learners must critique AI outputs, document identified errors, and propose corrective actions.
3. Adopting hybrid assignments that alternate between fully manual implementations and AI-augmented workflows, ensuring proficiency in both modalities (Ozkaya & Schmidt, 2024).
4. Updating assessment rubrics to evaluate not only artifact quality but also the rigor of students' AI interaction strategies and validation processes.

Unexpectedly, a subset of students became over-reliant on AI suggestions, producing superficially correct artifacts that concealed underlying misconceptions. This finding highlights the importance of continuous monitoring and targeted intervention to prevent surface-level learning. Furthermore, the small sample size and single-institution context limit broad generalizability; future research should replicate this approach across diverse cohorts and course formats.

In sum, GenAI holds substantial promise for enhancing software engineering education, but only when it is supported by pedagogical frameworks that emphasize critical oversight and conceptual rigor. By aligning learning objectives with a new profile of competencies, namely the Context Provider, Prompt

Design, Critical Validator, and Ethical Guardian, educators can better prepare students for a professional landscape where human–AI collaboration is increasingly the norm.

Conclusion

This exploratory case study provides preliminary evidence that the structured integration of Generative AI (GenAI) may support the development of software engineering competencies among students. The findings suggest that GenAI not only accelerates artifact production but was also associated with more positive learning experiences, including higher self-reported confidence, efficiency, and conceptual understanding. The research identified a fundamental transformation in the engineering workflow, shifting the primary challenge from manual execution to the critical evaluation of AI-generated outputs. This shift underscores the need to redefine SE competencies, emphasizing skills like critical validation, context provision, and prompt engineering.

Limitations and Future Works

Despite these contributions, this study has several limitations that must be acknowledged. First, the case study approach with a very small sample size (four students) and reliance on subjective perceptual data significantly limits the statistical generalizability of the findings to a broader population. Second, the research was conducted in a controlled academic environment, which may not fully reflect the complexities and dynamics of real-world industry contexts. Third, the study did not involve direct validation from industry practitioners; therefore, the definition of "industry needs" was based on theoretical frameworks and literature rather than direct market assessment.

Based on these findings and limitations, we propose several directions for future research. First, there is a clear need to validate and expand upon these initial findings through larger-scale studies that include direct involvement from industry practitioners in the evaluation process. This would test the transferability of the identified human-AI collaboration dynamics to contexts aligned with actual industry standards. Second, the "learning paradox" identified in this study—where effective AI use requires strong foundational knowledge—highlights an urgent need for research focused on developing new pedagogical models. Future studies could investigate how to best leverage GenAI not just as a productivity tool but as an instrument to reinforce fundamental concepts, potentially leading to strategic frameworks for its effective and responsible integration into SE curricula.

A further limitation concerns the scoring of student artifacts. Because all assessments were conducted by a single researcher, no inter-rater reliability measure, such as Cohen's Kappa or an intraclass correlation coefficient, could be calculated. Steps were taken to support consistency, including predefined rubric anchors and documented justifications for each score. However, these measures do not substitute for independent verification by a second qualified rater. Future studies in this area should involve at least two assessors with relevant domain expertise and report agreement statistics to establish the reliability of the scoring process.

References

Afifa, L. (2023). *Layoffs Caused Almost One Million People to Lose Jobs Last Year, Apindo Says - News En.tempo.co*. https://en.tempo.co/read/1675803/layoffs-caused-almost-one-million-people-to-lose-jobs-last-year-apindo-says

Al-Sarayreh, K. T., Meridji, K., & Abran, A. (2021). Software engineering principles: A systematic mapping study and a quantitative literature review. *Engineering Science and Technology, an International Journal*, *24*(3), 768–781. https://doi.org/10.1016/j.jestch.2020.11.005

Assyne, N., Ghanbari, H., & Pulkkinen, M. (2022a). The essential competencies of software professionals: A unified competence framework. *Information and Software Technology*, *151*. https://doi.org/10.1016/j.infsof.2022.107020

Assyne, N., Ghanbari, H., & Pulkkinen, M. (2022b). The state of research on software engineering competencies: A systematic mapping study. *Journal of Systems and Software*, *185*. https://doi.org/10.1016/j.jss.2021.111183

Bauer, F. L. (1972). Software engineering, Information processing. *Proceedings of the Software Engineering*, 1899–2227.

Cantrell, S. G. M. J. R. H. J. (2022). *The skills-based organization: A new operating model for work and the workforce*. https://www.deloitte.com/us/en/insights/topics/talent/organizational-skill-based-hiring.html

Chen, T. (2024). The Impact of AI-Pair Programmers on Code Quality and Developer Satisfaction: Evidence from TiMi studio. *ACM International Conference Proceeding Series*, 201–205. https://doi.org/10.1145/3665348.3665383

Coutinho, M., Marques, L., Santos, A., Dahia, M., França, C., & De Souza Santos, R. (2024). The Role of Generative AI in Software Development Productivity: A Pilot Case Study. *AIware 2024 - Proceedings of the 1st ACM International Conference on AI-Powered Software, Co-Located with: ESEC/FSE 2024*, 131–138. https://doi.org/10.1145/3664646.3664773

Creswell, J. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*.

Fan, A., Gokkaya, B., Harman, M., Lyubarskiy, M., Sengupta, S., Yoo, S., & Zhang, J. M. (2023). Large Language Models for Software Engineering: Survey and Open Problems. *Proceedings - 2023 IEEE/ACM International Conference on Software Engineering: Future of Software Engineering, ICSE-FoSE 2023*, 31–53. https://doi.org/10.1109/ICSE-FoSE59343.2023.00008

Gao, Y. (2024). *Research: Quantifying GitHub Copilot's impact in the enterprise with Accenture*. https://github.blog/news-insights/research/research-quantifying-github-copilots-impact-in-the-enterprise-with-accenture/

Garousi, V., Giray, G., Tuzun, E., Catal, C., & Felderer, M. (2020). Closing the Gap between Software Engineering Education and Industrial Needs. *IEEE Software*, *37*(2), 68–77. https://doi.org/10.1109/MS.2018.2880823

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect size and Related Estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

International Organization for Standardization. (2011). *ISO/IEC 25010:2011 – Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*.

Kirova, V. D., Ku, C. S., Laracy, J. R., & Marlowe, T. J. (2024). Software Engineering Education Must Adapt and Evolve for an LLM Environment. *SIGCSE 2024 - Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, *1*, 666–672. https://doi.org/10.1145/3626252.3630927

Kozov, V., Ivanova, G., & Atanasova, D. (2024). Practical Application of AI and Large Language Models in Software Engineering Education. In *IJACSA) International Journal of Advanced Computer Science and Applications*. *15*(1). https://doi.org/10.14569/IJACSA.2024.0150168

Kuhail, M. A., Mathew, S. S., Khalil, A., Berengueres, J., & Shah, S. J. H. (2024). "Will I be replaced?" Assessing ChatGPT's effect on software development and programmer perceptions of AI tools. *Science of Computer Programming*, *235*. https://doi.org/10.1016/j.scico.2024.103111

Oguz, D., & Oguz, K. (2019). Perspectives on the Gap between the Software Industry and the Software Engineering Education. *IEEE Access*, *7*, 117527–117543. https://doi.org/10.1109/ACCESS.2019.2936660

Ozkaya, I., & Schmidt, D. (2024). *Generative AI and Software Engineering Education*. https://insights.sei.cmu.edu/blog/generative-ai-and-software-engineering-education/

Pereira, J., López, J.-M., Garmendia, X., & Azanza, M. (2024). *Leveraging Open Source LLMs for Software Engineering Education and Training*. 1–10. https://doi.org/10.1109/cseet62301.2024.10663055

Rivera-Ibarra, J. G., Rodríguez-Jacobo, J., Fernández-Zepeda, J. A., & Serrano-Vargas, M. A. (2010). Competency framework for software engineers. *Software Engineering Education Conference, Proceedings*, 33–40. https://doi.org/10.1109/CSEET.2010.21

Sengul, C., Neykova, R., & Destefanis, G. (2024). Software engineering education in the era of conversational AI: current trends and future directions. In *Frontiers in Artificial Intelligence* (Vol. 7). Frontiers Media SA. https://doi.org/10.3389/frai.2024.1436350

Sommerville, Ian. (2011). *Software engineering*. Pearson.

Washizaki, H. (2024). *SWEBOK v4*.

Washizaki, H., Sanchez-Segura, M.-I., Garbajosa, J., Tockey, S., Nidiffer, K. E., & Reilly, A. D. (2025). *An Overview of the SWEBOK Guide - SEBoK*. https://sebokwiki.org/wiki/An_Overview_of_the_SWEBOK_Guide

Zönnchen, B., Thurner, V., & Böttcher, A. (2024). On the Impact of ChatGPT on Teaching and Studying Software Engineering. *IEEE Global Engineering Education Conference, EDUCON*. https://doi.org/10.1109/EDUCON60312.2024.10578680

How to cite:

Afif, G. G., & Putra, P. O. H. (2026). Generative AI in Software Engineering Education: A Study on Competency and Perception. *Jurnal Sistem Informasi (Journal of Information System), 22*(2), 1–17.

User Experience of a Local Government Website: A Design Thinking and UEQ Approach

Muhammad Galih Wonoseto*
Faculty of Science and Technology
UIN Sunan Kalijaga Yogyakarta
Sleman, Indonesia
muhammad.wonoseto@uin-suka.ac.id

Ghina Dersa
Faculty of Science and Technology
UIN Sunan Kalijaga Yogyakarta
Sleman, Indonesia
ghinadersa2603@gmail.com

Ihwanikha Fazli Pratama
Faculty of Science and Technology
UIN Sunan Kalijaga Yogyakarta
Sleman, Indonesia
ifazlipratama@gmail.com

Aqib Khoiruzaman
Faculty of Science and Technology
UIN Sunan Kalijaga Yogyakarta
Sleman, Indonesia
aqbkhrzmn@gmail.com

Zahra Zakila Anindha Rahmanti
Faculty of Science and Technology
UIN Sunan Kalijaga Yogyakarta
Sleman, Indonesia
23106050019@student.uin-suka.ac.id

Adriel Devara Sandji
MSc Technopreneurship & Innovation
Nanyang Technological University
Singapore, Singapore
adrielde001@e.ntu.edu.sg

Eko Hadi Gunawan
Faculty of Science and Technology
UIN Sunan Kalijaga Yogyakarta
Sleman, Indonesia
eko.gunawan@uin-suka.ac.id

Abstract

This study evaluates the user experience (UX) of a local government website using an integrated Design Thinking and User Experience Questionnaire (UEQ) approach. A mixed-methods design was applied to the Pesisir Barat Regency Government website, combining qualitative usability testing with quantitative UEQ measurement involving 42 respondents. The Design Thinking framework—empathize, define, ideate, prototype, and test—was employed to identify usability issues and guide interface redesign. Pre-redesign results showed that all UEQ dimensions were below the global benchmark average, with notable weaknesses in novelty, navigation clarity, and information structure. Following redesign, all six UEQ dimensions demonstrated substantial improvement, reaching the "excellent" category and ranking within the top 10% of global benchmark results. The findings confirm that user-centered redesign significantly enhances both pragmatic and hedonic quality of e-government websites. This study provides empirical evidence supporting the effectiveness of integrating standardized UX evaluation with Design Thinking to improve digital public service quality.

Keywords: user experience evaluation, design thinking, UEQ, e-government, usability testing

* Corresponding Author

Introduction

The rapid advancement of digital technologies has significantly transformed the way governments deliver public services, shifting from conventional administrative processes toward digital-based service models commonly referred to as e-government. In this context, local government websites serve as a primary digital interface between public institutions and citizens, facilitating information dissemination, public communication, and access to administrative services. The effectiveness of these platforms is therefore not solely determined by their functional availability but increasingly by the quality of user experience (UX) they provide (Janowski, 2015; United Nations, 2022).

User experience has been widely recognized as a critical factor influencing user satisfaction, trust, and continued use of digital government services. Previous studies indicate that inadequate UX can reduce public trust, limit service adoption, and negatively affect perceptions of government credibility (Carter & Bélanger, 2005; Verkijika & De Wet, 2018). This challenge is particularly relevant for local government websites, which must accommodate diverse user groups with varying levels of digital literacy, information needs, and service expectations.

The Pesisir Barat Regency Government website serves multiple stakeholders, including local residents, tourists, business actors, and public institutions. Such diversity increases the complexity of website design, particularly in terms of navigation structure, information architecture, content readability, and interface responsiveness. Studies examining local government websites in developing regions consistently report recurring usability issues, including unclear navigation paths, inconsistent layouts, and limited mobile optimization, all of which hinder effective access to public information and services (Anthopoulos et al., 2016).

Regional government websites play a strategic role as the primary representation of government digital services in providing public information, promoting regional potential, and providing public access to administrative services. In the context of e-government, government websites play a critical role in promoting transparency, accountability, citizen engagement, and the delivery of high-quality digital public services (Anthopoulos et al., 2016; Janowski, 2015). The Pesisir Barat Regency Government website is utilized by various user groups, such as local residents, students, government officials, business actors, and local and international tourists. This diverse user base requires regional government websites to provide a high-quality user experience, as UX quality plays a critical role in facilitating access to information, improving user satisfaction, and encouraging the use of digital public services (Aladwani, 2013; United Nations, 2022). However, various studies in Indonesia over the past five years have shown that the quality of regional government websites still faces issues with usability and user experience. Ernungtyas, Boer, & Qadrifa (2023) found that less intuitive navigation and inconsistent interfaces result in low user satisfaction. Similar findings have been reported in previous studies, which suggest that poorly structured content, ineffective information architecture, and inadequate navigation design make it difficult for users to locate relevant information and negatively affect the overall user experience (Aladwani, 2013; ISO, 2019).

The quality of local government websites significantly influences citizens' trust, satisfaction, and willingness to engage with digital public services. Poor user experience may discourage citizens from adopting e-government services, whereas positive user experiences can strengthen trust and continued usage intentions (Aladwani, 2013; Verkijika & De Wet, 2018). In the context of regions with tourism potential, the quality of local government websites also influences the region's image and tourist interest, as official websites are often the initial source of information accessed by users (Wijayanti, Y., Suyoto, S., & Hidayat, A. T., 2023).

There is a relationship between website quality and user interaction. According to Shi and Yuan (2019), the usability of web content, personalized services, and emotional states have the greatest impact on users' satisfaction, which provides the relevant theoretical guides on usability construction of websites. To assess and improve the quality of user interaction with digital systems, various UX evaluation methods have been proposed. Among these, the User Experience Questionnaire (UEQ) has gained prominence due to its ability to quantitatively measure both pragmatic and hedonic dimensions of UX using a standardized and benchmark-based approach (Laugwitz et al., 2008; Schrepp et al., 2014;

Schrepp et al., 2017). The UEQ has been widely applied in e-government and public service contexts, demonstrating its effectiveness in capturing user perceptions related to efficiency, clarity, stimulation, and novelty (Schrepp et al., 2014). However, while UEQ is effective in identifying UX strengths and weaknesses, it provides limited guidance on how evaluation results should be translated into concrete design improvements.

Addressing this limitation requires the integration of UX evaluation with a user-centered design methodology capable of generating actionable solutions. Design Thinking offers a human-centered and iterative framework that emphasizes understanding user needs, defining core problems, generating ideas, developing prototypes, and testing solutions in real contexts (Brown, 2019; Dam & Siang, 2020). Prior research has shown that Design Thinking can effectively enhance usability and innovation outcomes in digital public services when applied systematically (Liedtka, 2020; Micheli et al., 2019).

Despite increasing interest in UX evaluation and user-centered design approaches, empirical studies that integrate UEQ-based evaluation with Design Thinking in the context of local government websites remain limited, particularly in developing-country settings. The analysis reveals that effective visual communication relies on the synergy between UI aesthetics and UX functionality, where UI serves as a tangible manifestation of broader UX strategy rather than an isolated discipline (Kholik et al., 2024; Masveta & Manyangara, 2025). Existing research often treats UX measurement and interface redesign as separate activities, resulting in evaluation outcomes that are not fully operationalized into design solutions. To address this gap, this study integrates UEQ-based user experience evaluation with a Design Thinking-driven redesign process to improve the usability and overall UX of a local government website. The findings are expected to contribute empirical evidence and methodological insights for enhancing user-centered digital public services and strengthening e-government implementation. Table 1 shows the research gaps filled by this study.

Table 1. Research Gap

Study	UEQ	Design Thinking	E-Government	Before-After Evaluation
Putra & Santiyasa (2023)	✓	✗	✓	✗
Micheli et al. (2019)	✗	✓	✗	✗
This Study	✓	✓	✓	✓

Literature Review

User Experience in E-Government Websites

User experience (UX) refers to users' perceptions and responses resulting from the use or anticipated use of a system, product, or service, encompassing emotions, beliefs, preferences, and behaviors (ISO, 2019). In the context of e-government, UX has emerged as a critical determinant of citizen satisfaction, trust, and continued use of digital public services. Government websites with poor UX may hinder access to information, reduce transparency, and negatively affect public perceptions of institutional credibility (Carter & Bélanger, 2005; Janowski, 2015).

Previous studies highlight that UX in e-government websites is influenced by both pragmatic qualities, such as usability and efficiency, and hedonic qualities, such as stimulation and attractiveness (Aladwani, 2013). Particularly in local government contexts, users often represent heterogeneous groups with diverse goals, technological skills, and expectations, making UX design and evaluation more complex (Anthopoulos et al., 2016). As a result, systematic UX evaluation is essential to ensure that government websites effectively support public service delivery and citizen engagement.

User Experience Questionnaire (UEQ)

The User Experience Questionnaire (UEQ) is a standardized instrument developed to measure user experience across six dimensions: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty (Laugwitz et al., 2008; Schrepp et al., 2014). These dimensions collectively capture both pragmatic aspects of usability and hedonic aspects related to user engagement and emotional response. UEQ has been widely adopted due to its reliability, ease of use, and availability of benchmark data, enabling comparative evaluation across systems and domains.

Several empirical studies demonstrate the effectiveness of UEQ in evaluating e-government platforms and public digital services. Schrepp et al. (2017) emphasize that UEQ provides reliable results even with relatively small sample sizes, making it suitable for applied UX research. Furthermore, Aladwani (2013) showed that citizens' views of e-government interface quality differ across user groups, underlining the need for instruments that capture users' perceptions directly. However, while UEQ effectively identifies UX strengths and weaknesses, it does not prescribe specific design solutions, highlighting the need for complementary design-oriented approaches.

Usability Testing in UX Evaluation

Usability testing is a qualitative evaluation method that examines how users interact with a system to accomplish specific tasks, focusing on effectiveness, efficiency, and satisfaction (Nielsen, 1994). In UX research, usability testing is commonly used to uncover interaction-level issues such as navigation errors, unclear information structure, and interface inconsistencies that may not be fully captured through questionnaire-based evaluation.

In e-government research, usability testing has been shown to provide valuable insights into real user behavior and contextual challenges faced by citizens when accessing digital services (Verkijika & De Wet, 2018). By observing task performance and collecting user feedback, researchers can identify critical usability barriers that affect service accessibility. When combined with quantitative UX instruments such as UEQ, usability testing enhances the depth and validity of evaluation results by linking subjective perceptions to observable interaction problems.

Design Thinking as a User-Centered Design Approach

Design Thinking is a human-centered design methodology that emphasizes empathy with users, iterative problem-solving, and solution prototyping (Brown, 2019). The Design Thinking process typically consists of five stages: empathize, define, ideate, prototype, and test (Dam & Siang, 2020). This approach encourages designers and researchers to deeply understand user needs and contextual constraints before developing and evaluating design solutions.

In the public sector, Design Thinking has gained increasing attention as an effective approach for improving digital public services and fostering innovation (Liedtka, 2020; Micheli et al., 2019). Studies indicate that Design Thinking helps bridge the gap between UX evaluation and practical system improvement by transforming user insights into actionable design interventions. When integrated with UX evaluation methods, Design Thinking supports iterative refinement and continuous improvement of digital systems.

Integration of UEQ and Design Thinking in E-Government Research

Although UEQ and Design Thinking have been widely applied independently, research integrating both approaches within a single evaluation–redesign framework remains limited. Most e-government studies focus either on measuring UX outcomes or on redesigning interfaces without systematic evaluation. Integrating UEQ with Design Thinking enables researchers to quantitatively assess UX, identify underlying usability problems, and iteratively develop design solutions based on empirical evidence.

Recent studies suggest that combining standardized UX measurement with user-centered design methodologies results in more effective and sustainable improvements in digital service quality

(Janowski, 2015). Therefore, adopting an integrated UEQ and Design Thinking approach provides a comprehensive framework for evaluating and enhancing user experience in local government websites.

Methodology

Research Design

This study employed a mixed-methods research design to comprehensively evaluate and improve the user experience (UX) of a local government website. The mixed-methods approach enables the integration of quantitative and qualitative data to provide a more holistic understanding of user perceptions and interaction behaviors than a single-method approach (Creswell & Plano Clark, 2018). Quantitative data were collected using the User Experience Questionnaire (UEQ), while qualitative insights were obtained through usability testing and observations. The overall research process was guided by the Design Thinking framework to ensure a user-centered evaluation and redesign process.

Research Object

The object of this study was the official website of the Pesisir Barat Regency Government, which functions as a digital platform for public information dissemination, administrative communication, and regional promotion. The website serves a heterogeneous user population, including residents, tourists, business actors, and public institutions, making it a suitable case for UX evaluation in an e-government context.

Participants

A total of 42 participants were involved in this study. Participants were selected using purposive sampling to represent diverse user groups who frequently access local government websites. These groups included local residents, civil servants and tourists. Table 2 and Table 3 shows the demographic data of the 42 respondents.

Table 2. Distribution of Respondents by Gender

Category	Number of Respondents	Percentage
Male	20	47.6%
Female	22	52.4%

Table 3. Distribution of Respondents by Age

Category	Number of Respondents	Percentage
< 20 years old	9	21.4%
20-30 years old	30	71.4%
More than 31 years old	3	7.2%

According to Schrepp et al. (2017), a sample size above 30 participants is generally sufficient to produce stable UEQ results. Participants had varying levels of digital literacy to reflect real-world user diversity.

Instruments

User Experience Questionnaire (UEQ)

The User Experience Questionnaire (UEQ) was used to quantitatively measure user experience across six dimensions: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty (Laugwitz et al., 2008). The UEQ employs a semantic differential scale ranging from −3 to +3 and includes benchmark data that allow comparison with results from other systems across different domains (Schrepp et al., 2014). The UEQ was administered twice: before and after the website redesign.

Usability Testing

Usability testing was conducted to identify interaction-level usability issues that may not be fully captured through questionnaire-based evaluation. Participants were asked to complete predefined tasks, such as locating specific information, navigating between sections, and accessing public service content. Some of the tasks performed included searching for public service information, searching for tourism information, and accessing news. User behavior, task completion success, errors, and verbal feedback were recorded. Usability testing is widely recognized as an effective method for uncovering usability problems in digital systems (Nielsen, 1994).

Procedure

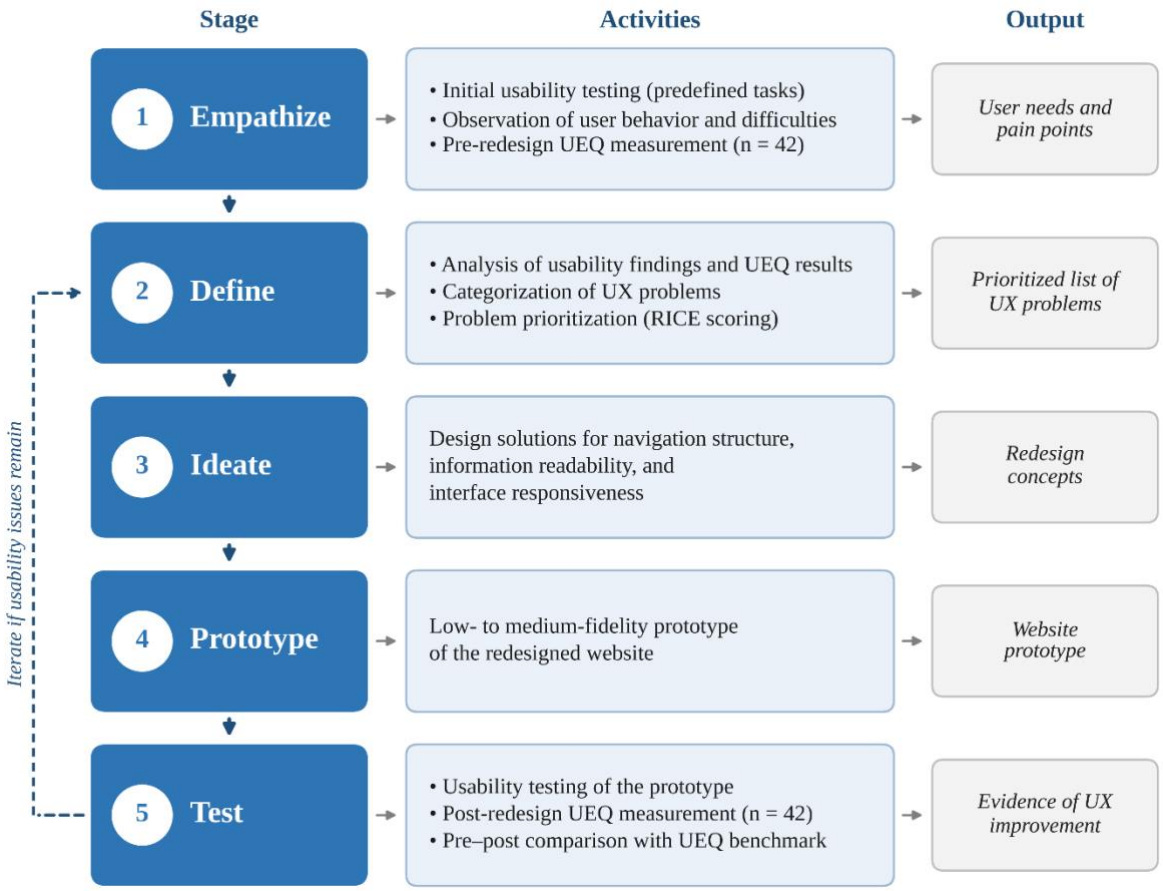


Figure 1. Research Procedure

The research procedure consisted of five main stages aligned with the Design Thinking framework:

1. **Empathize:** Initial usability testing and user observation were conducted to understand user needs, expectations, and difficulties when interacting with the website.
2. **Define:** Identified usability issues and UX problems were analyzed and categorized based on UEQ results and usability findings.
3. **Ideate:** Design solutions were generated to address the identified problems, focusing on navigation structure, information readability, and interface responsiveness.
4. **Prototype:** A low- to medium-fidelity prototype of the redesigned website was developed to represent the proposed improvements.
5. **Test:** The redesigned prototype was evaluated through usability testing and post-redesign UEQ measurement.

This iterative process ensured that evaluation results were directly translated into design improvements. Figure 1 or the chart below shows the stages or research procedures carried out.

Data Analysis

Quantitative data from the UEQ were analyzed using the official UEQ analysis tool to calculate mean scores for each UX dimension. Results were interpreted using the UEQ benchmark categories, including "below average," "above average," "good," and "excellent" (Schrepp et al., 2014). A descriptive comparison between pre- and post-redesign results was conducted to assess UX improvement.

Qualitative data from usability testing were analyzed using thematic analysis, focusing on recurring usability issues related to navigation, information structure, and interface design. Combining quantitative and qualitative findings enhanced the validity of the results by linking user perceptions with observed interaction behavior (Creswell & Plano Clark, 2018).

Ethical Considerations

All participants were informed about the purpose of the study and provided consent prior to participation. Data were collected anonymously and used solely for research purposes to ensure participant confidentiality and ethical compliance.

Results and Discussion

Results of User Experience Evaluation Using UEQ

The initial user experience evaluation was conducted using the User Experience Questionnaire (UEQ) to measure users' perceptions of the Pesisir Barat Regency Government website prior to redesign. The results indicate that all six UEQ dimensions—attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty—were rated below the global benchmark average. Table 4 is a list of questionnaires along with descriptions of the dimensions asked to respondents.

Based on the survey results, the lowest scores were observed in the novelty and perspicuity dimensions (Figure 2), indicating that users perceived the website as insufficiently innovative and lacking clarity, particularly in its navigation structure and information organization. These findings indicate that although the website was functionally accessible, it did not sufficiently support intuitive interaction or positive emotional engagement. According to Schrepp et al. (2014), low scores in perspicuity and novelty often reflect problems related to unclear information architecture and outdated interface design, which aligns with the observed conditions of the website prior to redesign. Figure 2 shows the UEQ measurements before the redesign.

Figure 2 shows that the UEQ level before redesign was below average. All UEQ dimensions—Attractiveness (0.94), Perspicuity (0.88), Efficiency (0.81), Dependability (0.80), Stimulation (0.85), and Novelty (0.40)—are in the Below Average category. Benchmark interpretation shows that approximately 50% of other evaluation results have better scores, while only 25% have.

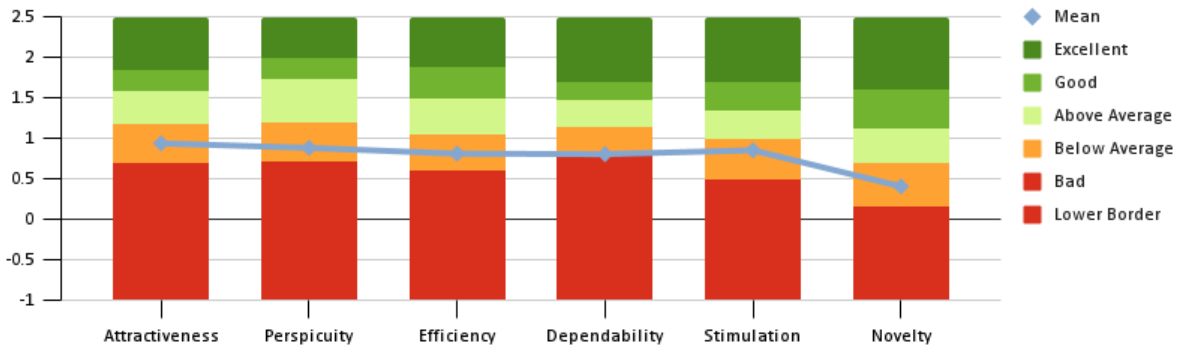


Figure 2. UEQ Before Redesign

Table 4. List of User Experience Questionnaire (Ambarwati & Mustikasari, 2021; Schrepp, 2023)

UEQ Dimensions		1	2	3	4	5	6	7	
Attractiveness	Annoying								Enjoyable
Perspicuity	Not Understandable								Understandable
Novelty	Creative								Dull
Perspicuity	Easy to learn								Difficult to learn
Stimulation	Valuable								Inferior
Stimulation	Boring								Exciting
Stimulation	Not Interesting								Interesting
Dependability	Unpredictable								predictable
Efficiency	Fast								Slow
Novelty	Inventive								conventional
Dependability	Obstructive								supportive
Attractiveness	Good								Bad
Perspicuity	Complicated								Easy
Attractiveness	Unlikable								Pleasing
Novelty	Usual								Leading edge
Attractiveness	Unpleasant								Pleasant
Dependability	Secure								Does not meet expectations
Stimulation	Motivating								Demotivating
Dependability	Meets expectations								Does not meet expectations
Efficiency	Inefficient								Efficient
Perspicuity	Clear								Confusing
Efficiency	Impractical								Practical
Efficiency	Organized								Cluttered
Attractiveness	Attractive								Unattractive
Attractiveness	Friendly								Unfriendly
Novelty	Conservative								Innovative

Table 5. Confidence Intervals Values Before Redesign

Confidence intervals (p=0.05) per scale						
Scale	**Mean**	**Std. Dev.**	**N**	**Confidence**	**Confidence interval**	
Attractiveness	0.937	1.419	42	0.429	0.507	1.366
Perspicuity	0.881	1.274	42	0.385	0.496	1.266
Efficiency	0.810	1.389	42	0.420	0.389	1.230
Dependability	0.804	1.392	42	0.421	0.383	1.225
Stimulation	0.851	1.303	42	0.394	0.457	1.245
Novelty	0.405	1.024	42	0.310	0.095	0.714

Table 5 shows the 95% confidence intervals for each UEQ dimension before the redesign based on 42 respondents. All dimensions have relatively wide confidence intervals, reflecting a fairly high variation in user perceptions. The Attractiveness dimension has a mean of 0.937 with a confidence interval between 0.507 and 1.366, while Perspicuity (0.496–1.266), Efficiency (0.389–1.230), Dependability (0.383–1.225), and Stimulation (0.457–1.245) also show lower and upper limits above zero, indicating a statistically positive evaluation trend. However, the Novelty dimension has the narrowest and lowest confidence interval (0.095–0.714), indicating that although still positive, the novelty and innovation aspects are perceived the weakest and least convincing by users. Overall, the results of these confidence intervals indicate that the pre-redesign user experience tends to be positive, but is not yet strong and

stable, especially in the innovation aspect, thus reinforcing the need for a redesign to improve the overall consistency and quality of the UX.

After redesigning the Pesisir Barat district government website, we conducted another user experience questionnaire with the same 42 respondents. Figure 3 and Table 6 show the comparison of the UEQ scores after the redesign against the global UEQ benchmark. All UEQ dimensions—Attractiveness (2.08), Perspicuity (2.04), Efficiency (2.06), Dependability (2.00), Stimulation (2.01), and Novelty (1.74)—are in the Excellent category, meaning they are all within the top 10% of results based on the UEQ benchmark. This achievement represents a very significant improvement compared to the condition before the redesign, where all dimensions were still in the Below Average category.

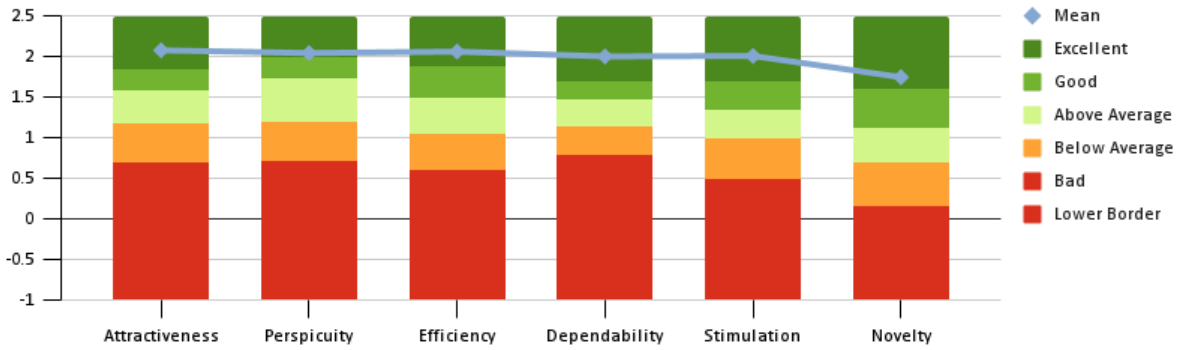


Figure 3. UEQ After Redesign

The visualization in Figure 3 shows that all mean values are well above the Above Average limit and consistently in the Excellent zone, indicating a highly competitive user experience quality globally. Overall, these benchmark results confirm that the redesign process has effectively improved interface quality, workflow efficiency, system reliability, as well as aspects of innovation and usability, enabling the system to deliver a superior user experience and meet international UX standards.

Table 6. Confidence Interval Values After Redesign

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	2.08	Excellent	In the range of the 10% best results
Perspicuity	2.04	Excellent	In the range of the 10% best results
Efficiency	2.06	Excellent	In the range of the 10% best results
Dependability	2.00	Excellent	In the range of the 10% best results
Stimulation	2.01	Excellent	In the range of the 10% best results
Novelty	1.74	Excellent	In the range of the 10% best results

Table 7. Confidence Intervals Values After Redesign

Confidence intervals (p=0.05) per scale						
Scale	**Mean**	**Std. Dev.**	**N**	**Confidence**	**Confidence interval**	
Attractiveness	2.075	0.936	42	0.283	1.792	2.359
Perspicuity	2.042	0.924	42	0.279	1.762	2.321
Efficiency	2.060	0.885	42	0.268	1.792	2.327
Dependability	2.000	1.014	42	0.307	1.693	2.307
Stimulation	2.006	1.056	42	0.319	1.687	2.325
Novelty	1.744	1.074	42	0.325	1.419	2.069

Table 7 shows the 95% confidence intervals for each UEQ dimension after the redesign based on 42 respondents. All dimensions have high mean values, with confidence intervals completely above 1.0, indicating a very positive and statistically significant user experience evaluation. The Attractiveness dimension has a mean of 2.075 with a confidence interval between 1.792 and 2.359, while Perspicuity (1.762–2.321), Efficiency (1.792–2.327), Dependability (1.693–2.307), and Stimulation (1.687–2.325) also show relatively narrow and consistent interval ranges. The Novelty dimension has the lowest mean of 1.744 with a confidence interval of 1.419–2.069, but remains in the very positive category, indicating that the novelty and innovation aspects have experienced a significant increase compared to the condition before the redesign. Overall, the results of these confidence intervals indicate that the user experience after the redesign is not only substantially improved, but also more stable and convincing, thus confirming the success of the redesign in improving the overall UX quality.

Usability Testing Findings

Qualitative usability testing revealed several recurring usability issues that explain the low UEQ scores. First, users experienced difficulties in locating specific information due to inconsistent menu structures and unclear labeling. Second, information readability was compromised by dense text, insufficient visual hierarchy, and inconsistent font usage. Third, users reported responsiveness issues when accessing the website on mobile devices, which negatively affected interaction efficiency.

These findings support previous research indicating that navigation complexity and poor content presentation are major barriers to effective use of e-government websites (Nielsen, 1994). The consistency between usability testing results and UEQ outcomes reinforces the validity of the evaluation and highlights the importance of combining quantitative and qualitative methods in UX research.

Design Thinking–Driven Redesign Outcomes

Based on the problems identified during the empathize and define stages, the ideate phase focused on improving navigation clarity, simplifying information structure, and enhancing visual consistency. The redesigned prototype introduced a clearer menu hierarchy, improved typography and spacing, and responsive layout adjustments to support mobile access.

Post-redesign evaluation using UEQ demonstrated a substantial improvement across all UX dimensions. All six UEQ scales reached the “excellent” category and ranked within the top 10% of the global benchmark dataset. The most significant improvements were observed in perspicuity, efficiency, and novelty, indicating that the redesign successfully enhanced both usability and user engagement.

These results confirm that applying Design Thinking as a guiding framework effectively translates UX evaluation findings into actionable design improvements. Similar outcomes have been reported in previous studies, where iterative, user-centered redesign significantly improved system usability and user satisfaction in public digital services (Liedtka, 2020; Micheli et al., 2019). Figure 4 presents a side-by-side comparison of the website appearance before and after the redesign.

As shown in Figure 4, the primary distinction in this redesign lies in the presentation of the news catalog, which has been carefully developed to improve usability, visual appeal, and informational clarity. The interface adopts a color palette inspired by the natural landscape of Pesisir Barat Regency, creating a strong regional identity while enhancing user engagement. A consistent typographic strategy is implemented through the use of one to three complementary fonts with hierarchical size variations, ensuring readability and structured information flow. Modern iconography is incorporated to reinforce visual aesthetics, while a more dynamic layout supports intuitive navigation. The homepage is designed to capture user attention through an attractive arrangement that includes a promotional video, complemented by high-quality hero images and videos—particularly drone footage of Bukit Barisan Selatan National Park and the region’s beaches—to establish an immersive first impression. A dark mode option is also provided to improve viewing comfort during nighttime browsing.

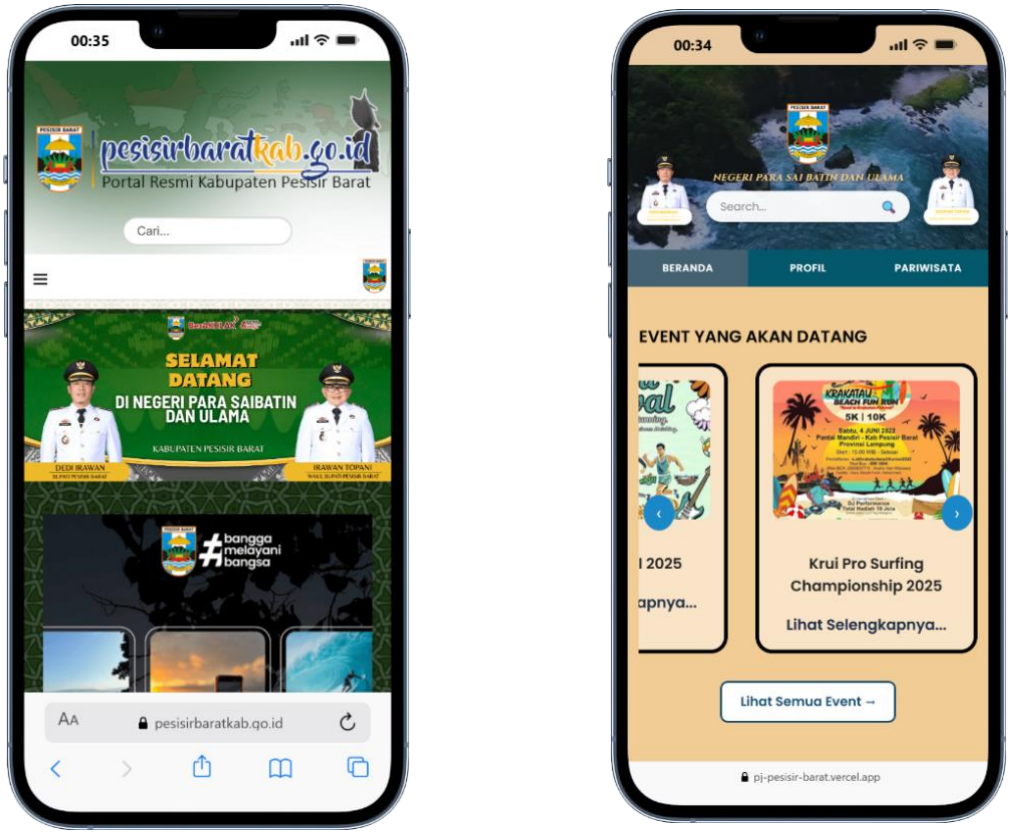


Figure 4. Comparison of Website Appearance Before and After Redesign

From a structural perspective, the platform features clearly defined primary menus—Home, Profile, and Tourism—allowing users to efficiently access essential information. Local government contact details are prominently displayed to support transparency and public communication. The system organizes content into meaningful categories, including regency history and regional symbols, as well as tourism destination classifications such as beaches (e.g., Tanjung Setia), mountains and forests (notably TNBBS), waterfalls, and cultural attractions. An event calendar is integrated to inform visitors about cultural festivals, surfing competitions, coffee harvest seasons, and other local events, thereby promoting regional activities.

To further enhance functionality, the platform integrates external services such as Google Maps and Instagram feeds, enabling users to access geospatial information and real-time social media updates. Automatic language detection based on browser settings is implemented to improve accessibility for international audiences. Additionally, a dedicated policy menu provides access to official government documents of Pesisir Barat Regency, reinforcing institutional credibility. The redesign also emphasizes optimized desktop responsiveness, ensuring that diverse content is presented effectively across larger screen environments. Collectively, these improvements position the platform as a modern, user-centered digital gateway for regional information and tourism promotion.

Discussion

The findings of this study demonstrate that integrating UEQ with usability testing and Design Thinking provides a comprehensive and effective approach to improving user experience in local government websites. The pre-redesign evaluation highlights that technical functionality alone is insufficient to ensure positive user experience, particularly in environments with diverse user groups. Instead, intuitive navigation, clear information presentation, and responsive design are critical determinants of UX quality. The substantial improvement observed in the post-redesign UEQ results suggests that user-centered design interventions can significantly enhance both pragmatic and hedonic aspects of UX. This

finding supports previous research emphasizing that UX quality directly influences citizens' trust and willingness to engage with e-government services (Aladwani, 2013; Carter & Bélanger, 2005). The substantial increase in perspicuity suggests that improvements in information architecture and navigation hierarchy play a critical role in enhancing usability in e-government websites.

From a methodological perspective, this study contributes empirical evidence that UEQ benchmark analysis is a reliable tool for measuring UX improvements when combined with qualitative usability insights. Unlike studies that rely solely on questionnaire data, the integration of Design Thinking ensures that evaluation results are operationalized into concrete design solutions. This approach addresses a common limitation in UX evaluation research, where findings are reported without leading to practical system improvements. However, this study also has limitations. The evaluation was conducted within a single local government website context, which may limit the generalizability of the findings. Future research could apply the proposed evaluation–redesign framework across multiple government websites or incorporate longitudinal analysis to examine long-term user satisfaction and adoption behavior.

Conclusion and Implications

Conclusion

This study evaluated the user experience (UX) of a local government website by integrating the User Experience Questionnaire (UEQ), usability testing, and the Design Thinking approach within a mixed-methods research framework. The initial evaluation revealed that, despite functional availability, the website exhibited low UX quality across all UEQ dimensions, particularly in terms of perspicuity, novelty, and navigation clarity. These findings indicate that technical functionality alone is insufficient to ensure effective digital public service delivery.

Following a user-centered redesign guided by the Design Thinking process, substantial improvements were observed across all UX dimensions. Post-redesign UEQ results reached the "excellent" category and ranked within the top 10% of the global benchmark dataset, demonstrating significant enhancement in both pragmatic and hedonic aspects of user experience. The results confirm that integrating standardized UX evaluation with iterative, empathy-driven design processes effectively translates evaluation insights into measurable UX improvements. Overall, this study provides empirical evidence supporting the value of user-centered design approaches in strengthening the quality of e-government websites.

Implications

Theoretical Implications

This study contributes to the UX and e-government literature by demonstrating the effectiveness of integrating UEQ-based quantitative evaluation with Design Thinking as a user-centered design methodology. The findings extend existing research by illustrating how standardized UX measurement can be operationalized into actionable design interventions, addressing a common limitation in prior UX evaluation studies that focus solely on measurement without redesign.

Practical Implications

From a practical perspective, the results offer actionable insights for local government institutions seeking to improve digital public services. The integrated evaluation–redesign framework can be adopted as a systematic approach to identify usability issues, prioritize design improvements, and enhance user satisfaction. Improving UX quality is expected to increase public trust, accessibility, and engagement with e-government services, particularly for users with diverse digital literacy levels.

Policy Implications

At the policy level, this study highlights the importance of incorporating user-centered design principles into digital government development strategies. Government agencies are encouraged to move beyond compliance-driven website development toward continuous UX evaluation and iterative improvement processes. Embedding UX evaluation frameworks such as UEQ and Design Thinking into digital governance policies can support sustainable e-government transformation and improve the overall quality of public digital services.

References

Aladwani, A. M. (2013). A cross-cultural comparison of Kuwaiti and British citizens' views of e-government interface quality. *Government Information Quarterly, 30*(1), 74–86. https://doi.org/10.1016/j.giq.2012.08.003

Ambarwati, P., & Mustikasari, M. (2021). Usability evaluation of the restaurant finder application using inspection and inquiry methods. *Jurnal Sistem Informasi, 17*(2), 1–17. https://doi.org/10.21609/jsi.v17i2.1049

Anthopoulos, L., Reddick, C. G., Giannakidou, I., & Mavridis, N. (2016). Why e-government projects fail? An analysis of the Healthcare.gov website. *Government Information Quarterly, 33*(1), 161–173. https://doi.org/10.1016/j.giq.2015.07.003

Brown, T. (2019). Change by Design, Revised and Updated: How Design Thinking Transforms Organizations and Inspires Innovation. HarperCollins. https://books.google.co.id/books?id=FPFduAEACAAJ

Carter, L., & Bélanger, F. (2005). The utilization of e-government services: Citizen trust, innovation and acceptance factors. *Information Systems Journal, 15*(1), 5–25. https://doi.org/10.1111/j.1365-2575.2005.00183.x

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications. https://collegepublishing.sagepub.com/products/designing-and-conducting-mixed-methods-research-3-241842

Dam, R. F., & Siang, T. Y. (2020). *Design thinking: Get started with Prototyping*. Interaction Design Foundation. https://ixdf.org/literature/article/design-thinking-get-started-with-prototyping

Ernungtyas, N. F., Boer, R. F., & Qadrifa, S. S. (2023). The government website as user's information source: A model of user satisfaction, information, and system quality. *Informasi, 53*(2), 197–214. https://doi.org/10.21831/informasi.v53i2.60579

ISO. (2019). *ISO 9241-210:2019 Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems*. International Organization for Standardization. https://cdn.standards.iteh.ai/samples/77520/33cc3cf17e5e4f1894ee872ed4f54fb6/ISO-9241-210-2019.pdf

Janowski, T. (2015). Digital government evolution: From transformation to contextualization. *Government Information Quarterly, 32*(3), 221–236. https://doi.org/10.1016/j.giq.2015.07.001

Kholik, A., Soegiarto, A., & Sari, W. P. (2024). Strategi komunikasi visual dalam user interface (UI) dan user experience (UX) untuk membangun kepuasan pengguna. *TUTURAN: Jurnal Ilmu Komunikasi, Sosial dan Humaniora, 2*(4), 335–344. https://doi.org/10.47861/tuturan.v2i4.1358

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger (Ed.), *HCI and usability for education and work* (Lecture Notes in Computer Science, Vol. 5298, pp. 63–76). Springer. https://doi.org/10.1007/978-3-540-89350-9_6

Liedtka, J. (2020). Putting technology in its place: Design thinking's social technology at work. *California Management Review, 62*(2), 53–83. https://doi.org/10.1177/0008125619897391

Masveta, D., & Manyangara, M. E. (2025). The UX-UI continuum: Exploring the interplay between user experience and user interface in e-learning platforms. *Cogent Education, 12*(1), Article 2536531. https://doi.org/10.1080/2331186X.2025.2536531

Micheli, P., Wilner, S. J. S., Bhatti, S. H., Mura, M., & Beverland, M. B. (2019). Doing design thinking: Conceptual review, synthesis, and research agenda. *Journal of Product Innovation Management, 36*(2), 124–148. https://doi.org/10.1111/jpim.12466

Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann. https://books.google.co.id/books?id=95As2OF67f0C&printsec=frontcover&hl=id - v=onepage&q&f=false

Schrepp, M. (2023). *User Experience Questionnaire handbook* (Version 11). https://www.ueq-online.org/Material/Handbook.pdf

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. In A. Marcus (Ed.), *Design, user experience, and usability: Theories, methods, and tools for designing the user experience* (Lecture Notes in Computer Science, Vol. 8517, pp. 383–392). Springer. https://doi.org/10.1007/978-3-319-07668-3_37

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*(6), 103–108. https://doi.org/10.9781/ijimai.2017.09.001

Shi, M., & Yuan, H. (2019). Impact of e-commerce website usability on user satisfaction. *Journal of Advanced Computational Intelligence and Intelligent Informatics, 23*(1), 91–96. https://doi.org/10.20965/jaciii.2019.p0091

United Nations. (2022). *United Nations e-government survey 2022: The future of digital government*. United Nations Department of Economic and Social Affairs. https://desapublications.un.org/sites/default/files/publications/2022-09/Web%20version%20E-Government%202022.pdf

Verkijika, S. F., & De Wet, L. (2018). E-government adoption in sub-Saharan Africa. *Electronic Commerce Research and Applications, 30*, 83–93. https://doi.org/10.1016/j.elerap.2018.05.012

Wijayanti, Y., Suyoto, S., & Hidayat, A. T. (2023). Evaluasi Pengalaman Pengguna Pada Aplikasi Seluler Visiting Jogja Menggunakan Metode User Experience Questionnaire (UEQ). Jurnal Janitra Informatika Dan Sistem Informasi, 3(1), 10-17. https://doi.org/10.25008/janitra.v3i1.169

How to cite:

Wonoseto, M. G., Dersa, G., Pratama, I. F., Khoiruzaman, A., Rahmanti, Z. Z. A., Sandji, A. D., & Gunawan, E. H. (2026). User Experience of a Local Government Website: A Design Thinking and UEQ Approach. *Jurnal Sistem Informasi (Journal of Information System)*, *22*(2), 18–31.

Acceptance Analysis of the Eco-Routing Feature in Google Maps from the Perspective of TAM and Environmental Concern

Siti Nur Kholifah
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
kholifahsitinur@student.ub.ac.id

Intan Sartika Eris Maghfiroh*
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
intansartika@ub.ac.id

Mochamad Chandra Saputra
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
andra@ub.ac.id

Almira Syawli
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
a.syawli@ub.ac.id

Abstract

The transportation sector is one of the main contributors to global greenhouse gas emissions. Google Maps' Eco-Routing feature is a technological innovation that recommends environmentally friendly routes. This study aims to examine the factors influencing students' intention to use Eco-Routing by integrating the Technology Acceptance Model (TAM) with Environmental Concern and Trust. A quantitative approach used questionnaire of college students who actively use Google Maps through Likert and open-end questions, and analyzed using PLS-SEM. The results show that Perceived Ease of Use and Trust have a significant positive effect on Behavioral Intention to Use Eco-Routing, while Perceived Usefulness and Environmental Concern do not. These findings indicate a value–action gap between environmental awareness and technology usage. This study provides improvements in ease of use, system reliability, route accuracy, and contextual presentation of environmental impact, contributing to literature and practical insights into sustainable information systems for the development of eco-friendly navigation.

Keywords: eco-routing, Technology Acceptance Model, Environmental Concern, Trust, sustainable transportation, Google Maps

Introduction

The transportation sector is one of the main contributors to greenhouse gas emissions globally, which has a significant impact on climate change and environmental degradation. Research shows that air pollution generated by transportation, particularly PM2.5 and ozone, contributes to 315,000 to 385,000 premature deaths worldwide, with emissions growing by 105% between 1990 and 2019 (Han, 2025).

In Indonesia, the dominance of private vehicles is clearly evident from data compiled by the Central Statistics Agency, which shows that motorcycles account for more than 84% of all motorized vehicles (Badan Pusat Statistik, 2023). The high use of these vehicles has triggered an increase in carbon dioxide

* Corresponding Author

(CO_2), nitrogen oxide (NO_x), and particulate matter (PM2.5) emissions, with the transportation sector in Jakarta alone contributing 93% of CO emissions (Lestari et al., 2020). With the serious impacts on health and the environment, sustainable transportation practices are now a crucial need, including implementing energy efficiency.

Amidst the rapid development of digital technology, navigation applications such as Google Maps have introduced a strategic innovation in the form of an eco-routing feature. This feature is designed to recommend the most fuel-efficient travel routes by taking into account variables such as elevation and traffic density (Google Sustainability, 2021). The implementation of this feature has been proven to reduce fuel consumption by 10% to 25% and help reduce more than 2.9 million metric tons of carbon emissions by the end of 2023. Implementation of this feature has been shown to reduce fuel consumption by 10% to 25% and help reduce more than 2.9 million metric tons of carbon emissions by the end of 2023, which is relevant to sustainable transportation goals to reduce carbon footprints (Huang et al., 2018).

This research is very important to be conducted at this time, given the global momentum in reducing the rate of global warming requires active participation from all levels of society through the adoption of green technology. Students were chosen as the main research subjects because they are a generation of digital natives who have a very high level of technological literacy and mobility. As the group that most frequently uses navigation applications in their daily activities, students have great potential to become early adopters who will influence future environmentally friendly behavior trends. Understanding students' perspectives on sustainable technology is key to predicting the effectiveness of future emission reduction policies.

To analyze this phenomenon, the study uses the Technology Acceptance Model (TAM) framework to identify the factors that drive the intention to use these features. Based on this model, behavioral intention is influenced by perceived usefulness, such as fuel savings, and perceived ease of use in accessing the feature (Sakulajo et al., 2023).

In addition, psychological factors such as environmental concern and trust in the accuracy of AI-based routes also influence the acceptance of eco-routing features. While this feature offers various benefits, there are still challenges such as uncertainty about travel time estimates and limited user understanding, which pose challenges to its adoption (Shao & Xu, 2023). Therefore, this study aims to evaluate these driving and inhibiting factors to optimize the use of environmentally friendly features among the younger generation.

Literature Review

Sustainable transportation involves developing transport systems that reduce energy use and environmental pollution while ensuring equitable access to mobility. Its implementation requires transportation policies and planning strategies that balance environmental protection, social well-being, and economic considerations (Heidari et al., 2023). One form of sustainable transportation is eco-routing, which recommends routes with lower energy consumption and emissions compared to the fastest route, taking into account traffic, road conditions, and vehicle type (Huang et al., 2018; Jovanovic et al., 2024). The effectiveness of eco-routing depends on route accuracy and suitability to user needs.

TAM can be used in sustainable technology studies to explain technology acceptance through Perceived Usefulness and Perceived Ease of Use, which directly influence the intention to use the technology. Perceived usefulness is the belief that the technology provides tangible benefits to users. Perceived Ease of Use describes the ease of using the technology without excessive effort (Davis, 1989). In eco-routing, PU related to fuel efficiency and emission reduction influences the intention to use sustainable transportation technology (Huang et al., 2018). while PEOU relates to the ease of understanding and accessing features, thereby increasing the acceptance of everyday navigation technology (Jovanovic et al., 2024).

Environmental Concern reflects an individual's level of concern for environmental issues, which is often associated with the adoption of environmentally friendly technologies. However, concern does

not always lead to actual behavior due to a gap between values and actions (Barbett & Landmann, 2024). In this case, trust in eco-routing is also necessary to follow routes that are not always the fastest, which can strengthen intentions to use digital technology. Behavioral intentions also reflect the likelihood of using technology in the future. BIU is directly influenced by perceived usefulness and perceived ease of use, and trust plays a significant role in increasing intentions to use recommendation-based systems (Davis, 1989).

Methodology

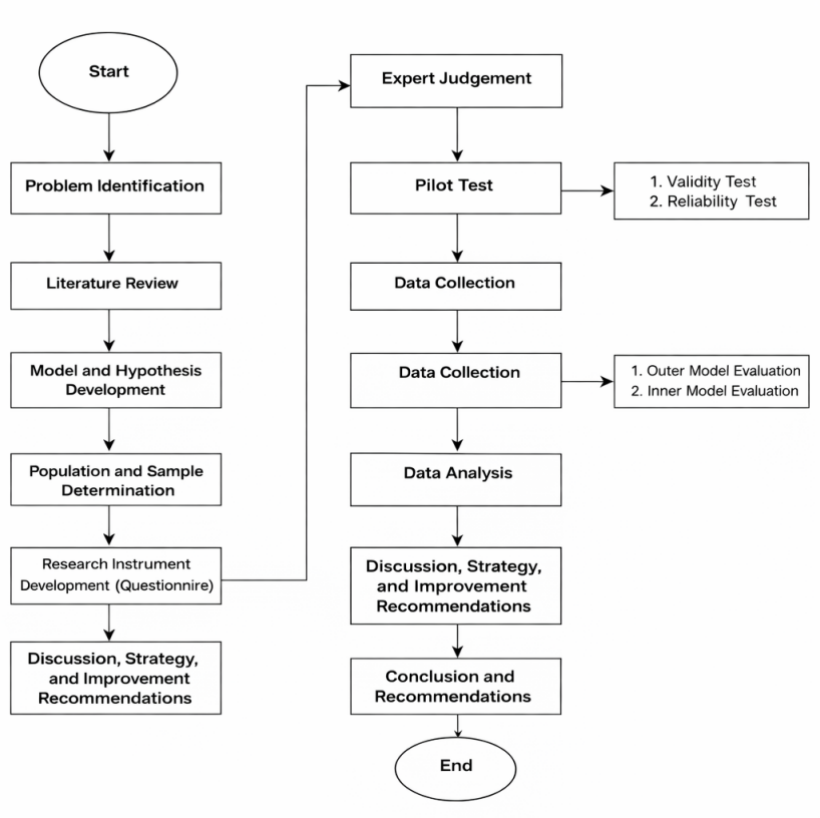


Figure 1. Method Flow

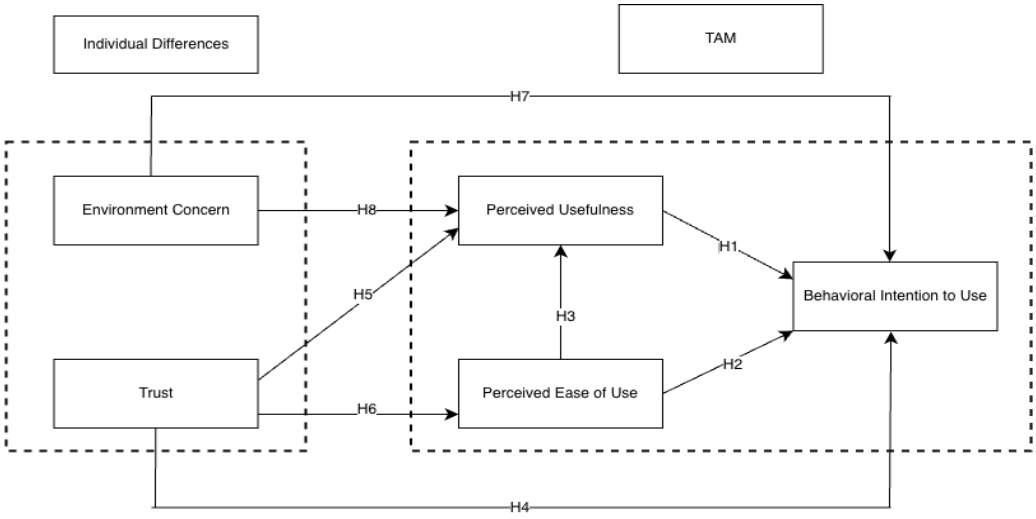


Figure 2. Research Model

This study is classified as non-implementation analytical research using a quantitative approach with quantitative data collection methods, in which a questionnaire survey was conducted to obtain feedback from respondents on both applications offering this efficient route feature. The research methodology can be seen in Figure 1.

The research model uses the Technology Acceptance Model (TAM) expanded with external variables, namely environmental concern and trust, to influence students' behavioral intention to use. This research model can be seen in Figure 2.

The researchers formulated eight main hypotheses based on the research model in Figure 2. The explanation related to the formulation of these eight hypotheses is as follows:

H1: Perceived Usefulness (PU) has a positive and significant effect on the Behavioral Intention to Use the Eco-Routing feature on Google Maps among students at Brawijaya University.

H2: Perceived Ease of Use (PEOU) has a positive and significant effect on Behavioral Intention to Use the Eco-Routing feature on Google Maps among students at Brawijaya University.

H3: Perceived Ease of Use (PEOU) has a positive and significant effect on Perceived Usefulness (PU) in the use of the Eco-Routing feature on Google Maps among students at Brawijaya University.

H4: Trust (T) has a positive and significant effect on Behavioral Intention to Use the Eco-Routing feature on Google Maps among students at Brawijaya University.

H5: Trust (T) has a positive and significant effect on Perceived Usefulness (PU) of the Eco-Routing feature on Google Maps among students at Brawijaya University.

H6: Trust (T) has a positive and significant effect on Perceived Ease of Use (PEOU)of the Eco-Routing feature on Google Maps among students at Brawijaya University.

H7: Environment Concern (EC) has a positive and significant effect on Behavioral Intention to Use the Eco-Routing feature on Google Maps among students at Brawijaya University.

H8: Environment Concern (EC) has a positive and significant effect on the Perceived Usefulness (PU) of the Eco-Routing feature on Google Maps among students at Brawijaya University.

The population in this study consisted of 62,415 active undergraduate students at Brawijaya University. The sampling technique employed was purposive sampling, which involved selecting respondents based on specific criteria relevant to the research objectives. Respondents were required to be familiar with and have used the eco-routing feature. Demographic information, including gender, age, and faculty, was collected through a questionnaire distributed via Google Forms.

The minimum sample size was determined using Slovin's formula, as presented in Equation 1,

$$n = \frac{N}{1 + Ne^2} \qquad \text{Equation (1)}$$

where n represents the minimum sample size, N represents the population size, and e represents the margin of error. Using a population size of 62,415 students and a margin of error of 10% ((e = 0.10)), the minimum sample size was calculated to be approximately 100 respondents.

The questionnaire consisted of Likert-scale items measuring the Technology Acceptance Model (TAM) variables and open-ended questions to collect suggestions regarding the feature under study. The questionnaire was evaluated by three experts, and Aiken's V was used to assess the content validity of the instrument. A pilot test involving 30 respondents was conducted to assess item validity using Pearson's correlation and reliability using Cronbach's alpha in SPSS. The collected data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM), involving outer and inner model evaluations performed with SmartPLS software. Responses to the open-ended questions were analyzed using thematic analysis.

Result

Respondent Demographic Results

The demographics of the respondents provide an overview of the characteristics of the research sample, which consisted of 101 people, all of whom met the specified criteria. Based on these results, the majority of respondents were aged 21–23 years (78.9%) and 18-20 years (21.1%), with the majority coming from the Faculty of FILKOM (29.4%). The complete results regarding the demographics of the respondents are presented in Table 1.

Table 1. Demographic Results of Respondents

Characteristic	Category	N	%
Age	18–20	21	21.1%
	21–23	80	78.9%
Faculty	FH	3	2.8%
	FEB	4	3.7%
	FIA	7	6.4%
	FP	4	3.7%
	FAPET	7	6.4%
	FT	6	5.5%
	FK	3	2.8%
	FPIK	4	3.7%
	FMIPA	5	4.6%
	FTP	12	11%
	FISIP	3	2.8%
	FIB	2	1.8%
	FKH	2	1.8%
	FILKOM	32	29.4%
	FKG	8	7.3%
	FIKES	7	6.4%

Results of Descriptive Statistical Analysis

Descriptive statistical analysis aims to present the distribution of data for each indicator representing a variable. The researchers present the results of the descriptive analysis in the form of scores (1 to 5), mode, median, mean, standard deviation, and variance. The results of the analysis provide an overview

of the positive trends of respondents for each variable in supporting the intention to use sustainable transportation through the Google Maps application.

Outer Model Test Results

Indicator reliability testing is used to assess how well the variance of each indicator can be explained by the relevant latent variable. The ideal outer loading value should be > 0.7 (Hair et al., 2021). Indicator reliability testing in this study was conducted twice because there were several indicators that had outer loading values < 0.7. The final test results showed that all indicators had outer loadings > 0.7, so they were declared valid and suitable for use in subsequent tests.

Internal consistency reliability test to assess the extent to which indicators in a variable have consistent relationships (Hair et al., 2021). This test was conducted using Cronbach's Alpha (CA) and Composite Reliability (CR), each of which must have a value > 0.7 for the variable to be declared reliable (Urbach & Ahlemann, 2010). The results of the internal consistency reliability test show that all variables have a value > 0.7, so they are declared reliable. Table 2 shows the results of the internal consistency reliability test in detail.

Table 2. Internal Consistency Reliability Results

Variable	CA	CR	Desc
Perceived Usefulness (PU)	0.789	0.878	Reliable
Perceived Ease of Use (PEOU)	0.857	0.903	Reliable
Environment Concern (EC)	0.782	0.902	Reliable
Trust (T)	0.855	0.902	Reliable
Behavioral Intention to Use (BIU)	0.847	0.908	Reliable

The convergent validity test aims to ensure that the indicators of a variable converge with each other compared to the indicators of other variables. This test uses the Average Variance Extracted (AVE) method with an AVE value that must be > 0.50 (Urbach & Ahlemann, 2010). The test results show that all variables obtained an AVE value > 0.5, thus meeting the validity criteria. Table 3 shows the detailed results of the convergent validity test.

Table 3. Convergent Validity Test Results

Variable	AVE	Desc
Perceived Usefulness (PU)	0.706	Valid
Perceived Ease of Use (PEOU)	0.699	Valid
Environment Concern (EC)	0.821	Valid
Trust (T)	0.697	Valid
Behavioral Intention to Use (BIU)	0.766	Valid

Discriminant validity testing is used to ensure that each variable is empirically different from other variables in the structural model (Hair et al., 2021). There are three methods for testing discriminant validity, namely Cross-loadings, Fornell-Larcker criterion, and Heterotrait-Monotrait Ratio (HTMT). The test results show that all indicators have the highest loading on their own variables, the AVE square root value of each variable is > its correlation with other variables, and all HTMT values are < 0.90. Thus, all variables meet discriminant validity and are eligible to proceed to inner model testing.

Inner Model Test Result

The coefficient of determination test uses the R-square value to measure the extent to which independent variables explain the variance in dependent variables (Tanlim & Ruslim, 2024). The interpretation of the R-square value consists of three criteria, namely a value of 0.670 is considered substantial, a value of 0.333 is considered moderate, and a value of 0.190 is considered weak (Urbach & Ahlemann, 2010). The test results show that the independent variables have a fairly good ability to explain the variance in each dependent variable. The Perceived Ease of Use (PEOU) variable can be explained in the moderate category (44.5%), and Perceived Usefulness (PU) can also be explained in the moderate category (49.6%) by the exogenous variables. Meanwhile, Behavioral Intention to Use (BIU) can be explained in the moderate category (57.2%) by the variables included in the research model. Thus, this research model is able to describe the influence of factors affecting users' intention to use the eco-routing feature. Table 4 presents the detailed results of the coefficient of determination test.

Table 4. Results of the Coefficient of Determination Test

Variable	R-Square	Desc
Behavioral Intention to Use (BIU)	0.572	Moderat
Perceived Ease of Use (PEOU)	0.445	Moderat
Perceived Usefulness (PU)	0.496	Moderat

Test path coefficients to assess the strength and direction of influence between variables in the structural model. The testing was conducted using the bootstrapping method with Smart-PLS. The hypothesis was considered valid if the path coefficient value was between -1 and +1. The relationship between variables was considered significant if the t-statistic value was > t-table (1.96) and the p-value was < 0.05 (5%) (Hair et al., 2021). The test results show that 8 research relationships are positive and significant ($t > 1.96$, $p < 0.05$), while 2 research relationships are insignificant because the t-statistic and p-value values do not meet the specified criteria ($t < 1.96$, $p > 0.05$). These results also prove that there is sufficient evidence to accept 8 research hypotheses and insufficient evidence to accept 4 research hypotheses based on the hypotheses made in the research model. Table 5 shows the detailed results of the path coefficients test.

Table 5. Path Coefficient Test Result

Path	Path Coef.	T-Statistic	P-values	Hypothesis
PU→BIU	0.096	1.092	0.275	H1 Rejected
PEOU→BIU	0.291	3.235	0.001	H2 Accepted
PEOU→PU	0.605	5.639	0.000	H3 Accepted
T→ BIU	0.534	4.865	0.000	H4 Accepted
T→PU	0.116	0.1038	0.299	H5 Rejected
T→ PEOU	0.667	10.849	0.000	H6 Accepted
EC→BIU	-0.106	1.256	0.209	H7 Rejected
EC→PU	0.027	1.255	0.798	H8 Rejected

Effect size testing uses the f-square value to assess the magnitude of the substantial influence of independent variables on dependent variables in structural models. The interpretation of the f-square value is divided into three categories, namely 0.02 for a low effect, 0.15 for a moderate effect, and 0.35

for a large effect (Urbach & Ahlemann, 2010). The test results show that 8 relationships meet the criteria and 2 relationships have no effect on the model ($f^2 < 0.02$), namely the influence of attitude and subjective norm on the intention to sell used items online. Overall, 4 relationships have a low effect, 2 relationships have a moderate effect, and 2 other relationships show a large effect. Thus, the model is able to provide a clear picture of the strength of each variable in influencing structural relationships. Table 6 presents the effect size test results in detail.

Table 6. Effect Size Test Result

Path	F-square	Desc
PU→BIU	0.011	No Effect
PEOU→BIU	0.077	Low
PEOU→PU	0.396	High
T→ BIU	0.261	Mid
T→PU	0.011	No Effect
T→ PEOU	0.803	High
EC→BIU	0.013	No Effect
EC→PU	0.001	No Effect

The predictive relevance test uses the Q-square value to assess the extent to which the research model is able to accurately predict latent dependent variables. This test is conducted through PLS predict in SmartPLS software. A model is said to have good predictive relevance if the Q-square value is > 0 (Urbach & Ahlemann, 2010). The results of the predictive relevance test show that all dependent variables have a Q-square value > 0.

Discussion

Factors that Encourage Users to Use the Eco-Routing Feature on the Google Maps Application

The results of the study indicate that Perceived Ease of Use (PEOU) has a positive and significant effect on students' intention to use the Eco-Routing feature ($t=3.235$; $p=0.001$; $f^2=0.077$). This finding supports the study by Li & in (2022), but differs from Maghfiroh et al. (2023), who found a negative effect on other platforms. Analysis of open-ended questions reinforces this, with 69.4% of respondents feeling that this feature is easy to use because it is directly integrated with the familiar Google Maps interface, thereby minimizing cognitive load in daily navigation.

The Trust (T) variable was identified as the most dominant predictor of usage intention ($t=4.865$; $p=0.000$; $f^2=0.261$). Trust in the accuracy of the system and the credibility of Google Maps were the main foundations for students to adopt environmentally friendly routes. These findings are in line with Sakulajo et al. (2023) and supported by qualitative data showing that 74.1% of respondents felt they could contribute to reducing emissions without sacrificing comfort, confirming that system reliability is key to behavioral change.

Factors that Prevent Users from Using the Eco-Routing Feature on the Google Maps Application

Conversely, Perceived Usefulness (PU) was not found to have a significant effect on usage intention ($t=1.092$; $p=0.275$; $f^2=0.011$), thus rejecting H1. In line with (Phuangsuwan et al., 2024), the ecological benefits of this feature are not yet perceived as a major functional advantage over the fastest route.

Qualitative data showed that 43.5% of respondents considered the Eco-Routing route to be less than optimal in terms of time, so based on Task-Technology Fit Theory (Goodhue & Thompson, 1995), this technology does not yet fully meet users' practical needs for efficiency.

The Environmental Concern (EC) factor also did not show a significant effect on usage intention or perceived usefulness ($t=1.256$; $p=0.209$; $f^2=0.013$). This phenomenon indicates the existence of a value–action gap, where students' environmental concerns do not automatically translate into behavior to use the feature of Barbett & Landmann (2024) that pro-environmental behavior is multidimensional; without the support of convenience and direct practical benefits, environmental awareness alone is not strong enough to encourage the sustainable adoption of Eco-Routing.

Recommendations and Improvements to Make the Eco-Routing Feature on Google Maps More Acceptable to Users

The strategy for improving the perceived usefulness of the Eco-Routing feature focuses on improving functional performance, referring to Task–Technology Fit (Goodhue & Thompson, 1995), which emphasizes the alignment of technology with the user's main task, namely travel efficiency. The perception of more circuitous routes indicates the need for vehicle profile personalization to improve route accuracy, in line with the principle of personalization in User Experience Design. Furthermore, based on the Value-Based Adoption Model (Kim et al., 2007), the benefits of Eco-Routing need to be communicated in utilitarian terms, such as estimated fuel cost savings, to reinforce its usefulness and reduce the perceived trade-off in travel time.

The Environmental Concern handling strategy focuses on reducing the value–action gap, considering that environmental concerns do not directly encourage the use of Eco-Routing, in line with Barbett & Landmann (2024). A digital green nudging approach based on Nudge Theory (Thaler & Sunstein, 2009) is recommended through emissions feedback and social reference, as well as increased discoverability of features through contextual orientation. This strategy aims to transform abstract environmental concerns into concrete and repeated actions in the use of Eco-Routing.

Conclusion

This study concludes that Perceived Ease of Use (PEOU) and Trust (T) are the main drivers of students' intention to use the Eco-Routing feature. Low cognitive load ease of access and high trust in system reliability are the main foundations for adopting this technology. This shows that the integration of familiar features and route accuracy are more decisive in user interest than mere functionality.

Conversely, Perceived Usefulness (PU) and Environmental Concern (EC) were not found to be significant. A value–action gap was found, whereby environmental concern does not automatically trigger usage if the practical benefits (time efficiency) do not exceed those of the fastest route. Users still prioritize travel comfort and speed over abstract ecological benefits.

As a development strategy, Eco-Routing must transform into behavior-shaping technology by optimizing route algorithms to remain competitive in terms of time. Improvements need to focus on transparency of real environmental impacts and strengthening intuitive interfaces. By balancing functional efficiency and system trust, this feature can shift from being a mere add-on to an integral part of sustainable transportation habits. Further research needs to expand the model by integrating behavioral and system quality variables such as habit, system quality, and perceived risk. In addition, the generalization of results can be improved by involving more heterogeneous respondents, such as the general public or workers with high mobility. The use of mixed methods and comparative studies between environmentally friendly navigation platforms is highly recommended to explore user experiences more comprehensively and evaluate the effectiveness of cross-application features.

References

Badan Pusat Statistik. (2023). Statistik Transportasi Darat 2023. https://www.bps.go.id/assets/publication/2024/11/25/cdcf9b5e74dd2e9bb3458ee4/statistik-transportasi-darat-2023.html

Barbett, L., & Landmann, H. (2024). What Is Pro-Environmental Behaviour? Nuanced Maps of Multi-Dimensional Pro-Environmental Behaviours That Highlight the Relevance of Impact and Diversity. *OSF*, (1), 4–8. https://doi.org/10.31219/osf.io/h75fw

Davis, F. D. (1989). Technology acceptance model: TAM. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption*, *205*(219), 5.

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, (19) 213–236. https://doi.org/10.2307/249689

Google Sustainability. (2021). Google Maps Eco-Friendly Routing Accelerating the journey to sustainability for one billion people. *Sustainability Report*. https://sustainability.google/reports/google-maps-eco-friendly-routing/

Hair, J. F. J., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial Least Squares Structural Equation Modeling (Pls-Sem) Using R. Classroom Companion: Business. *Cham: Springer International Publishing*, (1), 100-131. https://doi.org/10.1007/978-3-030-80519-7

Han, H. (2025). Increases in Global Transportation-Induced Air Pollution Mortality and Radiative Forcing During 1990–2019. *Frontiers in Environmental Science*, (13), 1-3. https://doi.org/10.3389/fenvs.2025.1545924

Heidari, I., Toloie Eshlaghy, A., & Seyyed Hoseini, S. M. (2023). Sustainable transportation: Definitions, dimensions, and indicators—Case study of importance-performance analysis for the city of Tehran. *Heliyon, 9*(10), 2-4. https://doi.org/10.1016/j.heliyon.2023.e20457

Huang, Y., N.g, E. C. Y., Zhou, J. L., Surawskia, N. C., Chan, E. F. C., & Hong, H. (2018). Eco-Driving Technology For Sustainable Road Transport: A Review. *Renewable and Sustainable Energy Reviews*, 93, 596–609. https://doi.org/10.1016/j.rser.2018.05.030

Jovanovic, A., Gavric, S., & Stevanovic, A. (2024). Evaluating Google Maps' Eco-Routes: A Metaheuristic-Driven Microsimulation Approach. *Geographies*, *4*(4), 732–752. https://doi.org/10.3390/geographies4040040

Kim, H.-W., Chan, H. C., & Gupta, S. (2007). Value-Based Adoption of Mobile Internet: An Empirical Investigation. *Decision Support Systems*, *43*(1), 111–126. https://doi.org/10.1016/j.dss.2005.05.009

Lestari, P., Damayanti, S., & Arrohman, M. K. (2020). April. Emission Inventory of Pollutants (CO, SO2, PM2. 5, and NOX) in Jakarta Indonesia. *In IOP Conference Series: Earth and Environmental Science*, *489*(1), 012014. https://doi.org/10.1088/1755-1315/489/1/012014

Li, X., & Lin, H. (2022). Using the Extended Acceptance Model to Understand Continuance Intention of Dockless Bike-Sharing. *Frontiers In Psychology*, *13*, 786693. https://doi.org/10.3389/fpsyg.2022.786693

Maghfiroh, I. S. E., Setiawan, N. Y., Syawli, A., Wibowo, V. M., & Nandika, B. J. P. (2023). Analisis Upaya Pengurangan Jejak Karbon dalam Aplikasi Perangkat Bergerak Grab dan Gojek Menggunakan Technology Acceptance Model. *Jurnal Sistem Informasi, Teknologi Informasi, Dan Edukasi Sistem Informasi*, *4*(2), 85–93. https://doi.org/10.25126/gafbf803

Phuangsuwan, P., Siripipatthanakul, S., Chaiprakarn, S., & Puangsang, M. (2024). Investigating Intentions to Use Google Maps for Travelling Among Users. *RMUTT Global Business and Economics Review*, *19*(1), 91–110. https://doi.org/10.60101/rmuttgber.2024.270317

Sakulajo, P., Satiennam, W., & Jaensirisak, S. (2023). Understanding Intention to Use Conditionally Automated Vehicles in Thailand, Based on An Extended Technology Acceptance Model. *Sustainability*, *15*(3), 1868. https://doi.org/10.3390/su15031868

Shao, Z., & Xu, Y. (2023). Moving Towards Carbon Neutral Lifestyle Through Fintech Social Media Platform: A Case Study of Ant Forest. *Frontiers In Environmental Science*, *11*, 1160986.

Tanlim, V., & Ruslim, T. S. (2024). Measuring customers loyalty through satisfaction in e-commerce: An empirical study on Tokopedia. *Jurnal Manajemen*, *28*(3), 567–587. https://doi.org/10.24912/jm.v28i3.1967

Thaler, R. H., Sunstein, C. R., & Milligan, S. (2009). Nudge: Improving decisions about health, wealth, and happiness. *London: Penguin books*. (47), 9.

Urbach, N., & Ahlemann, F. (2010). Structural Equation Modeling in Information Systems Research Using Partial Least Squares. *Journal of Information Technology Theory and Application (JITTA)*, *11*(2), 19-21. https://aisel.aisnet.org/jitta/vol11/iss2/2

How to cite:

Kholifah, S. N., Maghfiroh, I. S. E., Saputra, M. C., & Syawli, A. (2026). Acceptance Analysis of the Eco-Routing Feature in Google Maps from the Perspective of TAM and Environmental Concern. *Jurnal Sistem Informasi (Journal of Information System)*, *22*(2), 32–42.

Examining Digital Egalitarianism through the Moderating Effect of Job Level on G2E Systems Adoption

Hanif Kridho Pangantyo*
Faculty of Information and Technology
Perbanas Institute
Jakarta, Indonesia
hanif.kridho30@perbanas.id

Mardiana Purwaningsih
Faculty of Information and Technology
Perbanas Institute
Jakarta, Indonesia
mardiana@perbanas.id

Abstract

To support budget efficiency and bureaucratic reform, while embodying the ASN "Adaptive" Core Value regarding technology, it is essential to evaluate the existing Government-to-Employee (G2E) systems in the Deputy of Economic Cooperation and Investment, Coordinating Ministry of Economic Affairs, Republic of Indonesia. In this context, the Sipentakel platform was developed to digitally integrate program and governance services. This study aims to evaluate the performance of Sipentakel by employing a hybrid framework that integrates the DeLone & McLean IS Success Model, the Unified Theory of Acceptance and Use of Technology (UTAUT), and the Human-Organization-Technology (HOT) Fit model. Additionally, this study assesses the moderating role of Job Level on the relationship between Human Factors and Behavioral Intention. The analysis was performed using Partial Least Squares Structural Equation Modeling (PLS-SEM) on data from 52 system users. The findings indicate that Behavioral Intention to use the system is significantly driven by Human and Technology factors, whereas Organization factors and User Satisfaction do not demonstrate a significant effect. Furthermore, Net Benefits are primarily determined by Behavioral Intention. The Technology factor emerged as the most critical driver, suggesting that enhancing system and information quality is the most effective strategy for optimizing Sipentakel implementation. Notably, the assessment reveals that Job Level does not significantly moderate the influence of Human Factors on Behavioral Intention. This finding highlights a phenomenon of digital egalitarianism, indicating that user acceptance is driven by system utility rather than hierarchical position.

Keywords: DeLone and McLean, UTAUT, HOT-Fi, job level

Introduction

Bureaucratic reform, as a national priority agenda aimed at creating good governance, has accelerated the digitalization of government administrative services. As a main subject in the body of the Minister of Administrative and Bureaucratic Reform (PANRB) Regulation No. 9 of 2023, bureaucratic reform includes the strengthening of an integrated and comprehensive Electronic-Based Government System (SPBE) (Kementerian PANRB, 2023). The implementation of this SPBE is manifested in the achievement of the Bureaucratic Reform (RB) index through the SPBE index and the implementation level of the SPBE architecture policy on the strategic target of "The Realization of an Effective, Agile,

* Corresponding Author

and Collaborative Digital Government Governance," as it becomes an element of evaluation through the annual measurement of the bureaucratic reform index.

However, the Coordinating Ministry for Economic Affairs, the institution responsible for coordinating economic issues (Presiden Republik Indonesia, 2024), has not yet achieved an optimal bureaucratic reform score. Based on the results of the bureaucratic reform evaluation by the Ministry of Administrative and Bureaucratic Reform, which serves as the Coordinating Ministry's internal data, the achievement score of the Coordinating Ministry of Economic Affairs is 88.58, with a breakdown of a general RB score of 81.82 and a thematic RB of 6.76. The breakdown results from the general RB score related to SPBE are the implementation level of the SPBE architecture policy, valued at 3.00, and the SPBE index, valued at 4.00 out of a 5-point scale. This score is still not optimal to reach the aggregate RB target value of 90.77 in 2025. As a note, referring to the Minister of PANRB Regulation No. 8 of 2024, an RB index above 90.01 is required to propose a performance allowance adjustment to 100% (Kementerian PANRB, 2024). The suboptimality of the RB score could lead to a decrease in employee welfare due to the impossibility of proposing a performance allowance adjustment.

The implementation of bureaucratic reform through the implementation of the SPBE architecture policy and the SPBE index is the practical responsibility of all work units (Iswanto & Putra, 2023), which is then orchestrated by the Bureau of Performance Management and Cooperation, including the Deputy of Economic Cooperation and Investment at the Coordinating Ministry for Economic Affairs. As stipulated in Presidential Regulation No. 143 of 2024, the Coordinating Ministry of Economic Affairs coordinates issues in the economic sector, one of which relates to economic cooperation and investment, through the Deputy of Economic Cooperation and Investment (Presiden Republik Indonesia, 2024). Referring to the Regulation of the Coordinating Minister for Economic Affairs No. 15 of 2024 concerning the Organization and Work Procedure of the Coordinating Ministry for Economic Affairs, the provision of services and program support, including the implementation of bureaucratic reform and the management of information systems at the Deputy for Economic Cooperation and Investment, falls under the responsibility of the Deputy's Secretary. In order to strengthen SPBE implementation, which is a component of bureaucratic reform implementation (Iswanto & Putra, 2023), the Secretary of the Deputy Minister for Economic Cooperation and Investment developed a Sistem Pelayanan Program dan Tata Kelola (Sipentakel) that integrates program and governance services into a unified and integrated information system.

The Sipentakel information system contains program, governance, and other services. Program services include important matters in the deputy's program administration, covering budget planning and execution, including budget monitoring, completeness of financial accountability documents (SPJ), and performance agreements. Governance services include monitoring of key performance indicators (IKU), official correspondence management, to the documentation of performance appraisals and employee work targets. Other services in Sipentakel include add-on services that fall within the program and governance categories, such as scientific papers for functional policy analysts, proposals for study publications, and performance satisfaction surveys.

Until the time this research was conducted, there had been no effort to gather constructive feedback through an adequate performance evaluation of the Sipentakel information system, even though continuous and user-centric evaluation is key to the success of digital public service platforms (Weigl et al., 2024). The absence of this information system performance evaluation effort has the potential to cause budget inefficiencies and hinder efforts to improve the effectiveness of government system digitalization. The procurement of an information system that does not provide good net benefit value is certainly an inefficiency that contradicts Presidential Instruction (Inpres) No. 1 of 2025 concerning Expenditure Efficiency in the Implementation of the State Budget (APBN) and Regional Budget (APBD) for the 2025 Fiscal Year. This has the potential to become a finding by government internal control apparatus (the Financial and Development Supervisory Agency) and independent government auditors (the Supreme Audit Agency). The absence of an information system performance evaluation also has the potential to impact the performance assessment of the Deputy's Secretariat in the context of Key Performance Indicator (IKU) and bureaucratic reform index evaluations. Both of these are important indicators of an organization's performance and will certainly affect all employees at the Deputy for Economic Cooperation and Investment.

As aligned with the effort for continuous improvement as an implementation of the core value of State Civil Apparatus (ASN) Berakhlak in the value of service orientation (Syaputra et al., 2023) and the strengthening of bureaucratic reform, the Sipentakel information system needs to undergo a performance evaluation. The practice of information system evaluation in the public sector has become a standard to ensure service quality. In Indonesia, various studies have been conducted to measure the performance and user satisfaction of government information systems. A study conducted by Hidayatullah et al. in 2022 measured the success rate of the PeduliLindungi application using the DeLone & McLean IS Success Model evaluation method. This method assesses system success by measuring success dimensions and their relationship with system quality, information quality, user satisfaction, and net benefit (Hidayatullah et al., 2022). The performance of an information system can also be evaluated from the perspective of user acceptance using the Unified Theory of Acceptance and Use of Technology (UTAUT) method, as in the research conducted by Marsintauli et al. (2023) which evaluated the acceptance of the government's E-Invoicing Web system. The UTAUT method is used to measure factors such as performance expectancy and social influence that affect users' behavioral intentions towards a technology (Marsintauli & Sari, 2023; Omol & Ondiek, 2021). The alignment between users, organization, and technology can also be an evaluation perspective through the Human Organization Technology (HOT) Fit method. The HOT-Fit method measures the success of an information system by evaluating three main factors: humans as users, the organization, and technology, as well as the fit of the relationships among the three in generating net benefits (Arsyad & Sfenrianto, 2023). Research with this approach was conducted by Rifial et al. (2025) in the evaluation of the Hospital Management Information System (SIMRS) at Madani Regional General Hospital in Palu. The study concluded that the system usage factor is the most dominant factor influencing the utilization and net benefits of the SIMRS (Rifial et al., 2025).

In the context of evaluating the performance of the Sipentakel information system, the required evaluation perspectives may include the performance success of the information system, the acceptance of target users towards the use of technology within the information system, and the alignment between human resources at the Deputy of Economic Cooperation and Investment work unit with the organization and the applied technology. The performance success dimension is necessary to directly evaluate the system's performance so that technical improvements can be made. The acceptance of target users, in this case, the employees at the Deputy's office, needs to be evaluated to observe the behavioral shift of employees in accepting technology that is in line with the adaptive value in the ASN BerAKHLAK core values. The dimension of alignment between employees, organization, and technology needs to be evaluated to assess the support of the organization (Secretariat of Deputy of Economic Cooperation and Investment) and its relationship with employees and technology in viewing the Sipentakel information system.

To comprehensively study these three dimensions, the most appropriate approach is to integrate several theoretical frameworks with complementary advantages in success, acceptance, and alignment models through a combination of the DeLone & McLean, UTAUT, and HOT-Fit evaluation models. This integration method has been successfully adopted by Dewangga et al. in 2025 in evaluating the LPD Mobile application in Bali. The DeLone & McLean, UTAUT, and HOT-Fit evaluation model is an integrated model that can provide a better representation for evaluating from the aspects of success, acceptance, and alignment (Dewangga et al., 2025). DeLone & McLean is specifically designed to evaluate the success of a system. This model measures various variables such as system quality, information quality, and service quality, which then impact the intention to use, user satisfaction, and ultimately net benefit (Purwati et al., 2021). From the perspective of acceptance, the UTAUT method has high explanatory power to explain technology acceptance (Almaiah et al., 2019; Xue et al., 2024). UTAUT identifies four main factors that influence behavioral intention and use, namely performance expectancy, effort expectancy, social influence, and facilitating conditions (Nugroho et al., 2023). Meanwhile, HOT-Fit is used to evaluate the alignment or fit of the relationship between three crucial components: human, organization, and technology (Kumorowani & Mulyanti, 2023).

Previous evaluation studies utilizing these frameworks have predominantly focused on Business-to-Consumer (B2C) or Government-to-Citizen (G2C) applications, where technology adoption is largely voluntary and users operate outside of a formal hierarchical structure (Dewangga et al., 2025). A

significant research gap exists in evaluating internal, mandatory Government-to-Employee (G2E) systems within highly bureaucratic environments. In such settings, users are bound by structural hierarchies and job levels, which may fundamentally alter the dynamics of technology acceptance (Kuciapski, 2019). To address this gap, this study pioneers a comprehensive internal evaluation of a mandatory G2E system (Sipentakel) by integrating the DeLone & McLean, UTAUT, and HOT-Fit models. Furthermore, it specifically investigates the moderating role of hierarchical job levels to determine if structural rank influences digital acceptance in a modern bureaucracy (Kuciapski, 2019).

Based on the aforementioned problem background, to provide an adequate performance evaluation, especially in supporting bureaucratic reform and budget efficiency, this study will evaluate the performance of the Sipentakel information system of the Deputy of Economic Cooperation and Investment from the aspects of success, acceptance, and alignment through the integrated method of DeLone & McLean, UTAUT, and HOT-Fit to identify areas that require improvement and provide strategic recommendations for the relevant work unit. Therefore, this study aims to answer the following research questions: (1) How does the Sipentakel information system perform based on the integrated DeLone & McLean, UTAUT, and HOT-Fit framework? (2) What strategic recommendations can be formulated for the improvement and development of the system based on these findings?

Literature Review

Information System Performance Evaluation Models

The evaluation of information system performance is a fundamental process used to assess the condition of an implemented system, justify IT investments, identify areas requiring improvement, and provide data-driven foundations for managerial decision-making and organizational learning (Khairani et al., 2022). A widely accepted and foundational framework in this domain is the DeLone and McLean (D&M) Information Systems Success Model. Originally grounded in Shannon and Weaver's communication theory, the model was expanded to provide a comprehensive structure for measuring system success (Meilani et al., 2020). The updated 2003 D&M model operates on six primary dimensions: system quality, information quality, service quality, intention to use/use, user satisfaction, and net benefits (Rahmatullah et al., 2025). System quality assesses the technical characteristics of the system, such as ease of use and reliability, which are assumed to directly enhance user satisfaction and usage intention (Purwati et al., 2021). Information quality focuses on the accuracy, relevance, and completeness of the system's output, while service quality measures the responsiveness and assurance provided by IT support (Meilani et al., 2020). Although the D&M model allows researchers to systematically diagnose system strengths and weaknesses (Putri et al., 2021), it is limited in its ability to explain the psychological antecedents and individual perceptions that drive initial user acceptance.

To address the psychological and behavioral aspects of technology adoption, the Unified Theory of Acceptance and Use of Technology (UTAUT) model offers a robust framework designed to explain individuals' intentions and actual use behaviors (Purwanto & Loisa, 2020; Xue et al., 2024). Synthesized from eight prior technology acceptance theories by Venkatesh et al., the UTAUT model evaluates adoption through four main determinants: performance expectancy, effort expectancy, social influence, and facilitating conditions (Nugroho et al., 2023). While performance expectancy, effort expectancy, and social influence theoretically shape the behavioral intention to use a technology, facilitating conditions directly affect actual use behavior (Purwanto & Loisa, 2020). Despite its comprehensive approach to behavioral antecedents, the UTAUT model's primary limitation lies in its lack of explicit evaluation of intrinsic system qualities—such as technical reliability and information accuracy—which are critical for long-term satisfaction and sustainable success.

Offering an alternative perspective that emphasizes alignment, the Human-Organization-Technology Fit (HOT-Fit) model posits that the success of an information system depends heavily on the harmonization of human, organizational, and technological components. Introduced by Yusof et al., the HOT-Fit framework was developed to assess systems by evaluating not only technical aspects but also user contexts and organizational environments (Saputro et al., 2020). The technology component mirrors the D&M quality metrics (system, information, and service quality) (Ayuardini & Ridwan, 2019), while the human component evaluates system use and user satisfaction. Finally, the

organizational component accounts for internal structures, such as management support, and external environmental factors (Sari et al., 2016). Together, these components are evaluated based on their collective influence on achieving optimal net benefits.

Integration of DeLone & McLean, UTAUT, and HOT-Fit

In information systems research, relying on a single theoretical model often presents inherent limitations when attempting to capture the multifaceted phenomenon of system success (Omol & Ondiek, 2021). For instance, while the D&M model is highly reliable for post-implementation evaluation, it lacks a strong focus on the contextual factors driving initial user acceptance. Conversely, the UTAUT model excels at predicting behavioral intentions but does not deeply measure intrinsic system quality or long-term organizational impacts. Meanwhile, the HOT-Fit model offers a holistic perspective on alignment, though its concept of "fit" can sometimes remain abstract in practical application. Consequently, integrating the D&M, UTAUT, and HOT-Fit models provides a highly comprehensive framework capable of simultaneously assessing user acceptance, implementation success, and contextual alignment (Dewangga et al., 2025; Pamugar et al., 2014).

Structurally, this integrated approach synthesizes the dimensions of the three models into a cohesive, parsimonious framework. Adopting the architectural foundation of the HOT-Fit model, independent variables are categorized into human, technology, and organization factors. The human factor consistently integrates the core constructs from the UTAUT model, specifically performance expectancy, effort expectancy, and social influence. The technology factor adopts the quality dimensions from the D&M model, which include information quality, system quality, and service quality. Furthermore, the organization factor merges the facilitating conditions construct from UTAUT with contextual organizational variables, such as organizational support, which have proven crucial in shaping system success (Dewangga et al., 2025; Pamugar et al., 2014). Together, these three primary factors are hypothesized to influence the mediating variables—behavioral intention (derived from UTAUT) and user satisfaction (derived from D&M)—which ultimately culminate in determining the overall net benefits of the information system.

Conceptual Model

The research is built upon an integrated conceptual framework that synthesizes dimensions from three perspectives of information system evaluation models: the DeLone & McLean IS Success Model, the Unified Theory of Acceptance and Use of Technology (UTAUT), and the Human-Organization-Technology (HOT) Fit model. This integrated model, adapted from Dewangga et al. (2025), categorizes antecedent variables into three core factors: the Human Factor (comprising Performance Expectancy, Effort Expectancy, and Social Influence), the Technology Factor (including Information Quality, System Quality, and Service Quality), and the Organization Factor (integrating Facilitating Conditions and Organization Support) (Dewangga et al., 2025). These factors are hypothesized to influence the central variables of Behavioral Intention and User Satisfaction, which in turn are expected to affect the Net Benefits.

While the foundational integration of these three models is adapted from Dewangga et al. (2025), this study introduces a theoretical modification by incorporating Job Level (JL) as a moderating variable. This modification aims to investigate the phenomenon of digital egalitarianism within a hierarchical bureaucratic context (Kuciapski, 2019). The Job Level stated as a moderating variable on the relationship between the Human Factor and Behavioral Intention. Job Level is proposed as a moderating variable because hierarchical positions in government organizations often correlate with different administrative tasks, experiences, and access to resources. As suggested by Kuciapski (2019), the type of job position can significantly influence technology acceptance (Kuciapski, 2019). Therefore, it is hypothesized that a hierarchical job level may moderate the influence of the Human Factor on Behavioral Intention (Kuciapski, 2019). The conceptual workflow illustrated as in Figure 1.

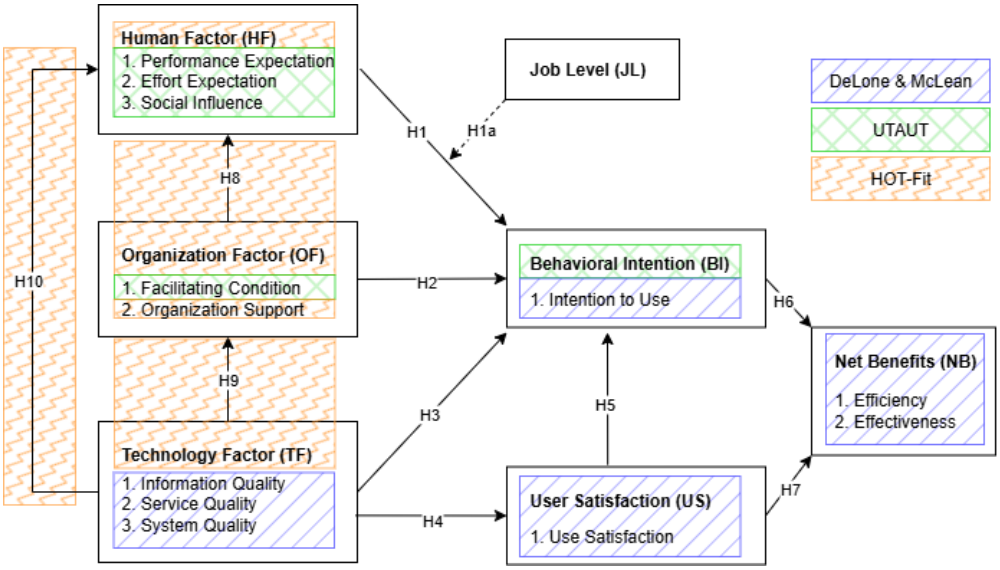


Figure 1. Proposed Conceptual Model

Based on the structural equation model depicted in Figure 1, this study incorporates six latent variables and one moderating variable to assess the causal relationships among constructs. The Technology Factor serves as the primary exogenous variable, hypothesized to influence the Human Factor, Organization Factor, Behavioral Intention, and User Satisfaction. Conversely, the endogenous variables include the Human Factor, Organization Factor, User Satisfaction, Behavioral Intention, and Net Benefits (Apsari et al., 2022; Dewangga et al., 2025; Pamugar et al., 2014). Additionally, Job Level is proposed as a moderating variable to examine whether a higher hierarchical position strengthens the positive influence of the Human Factor on Behavioral Intention. Finally, Net Benefits functions as the ultimate dependent variable, aggregating the comprehensive system assessment derived from the integrated DeLone & McLean, UTAUT, and HOT-Fit frameworks. The structural model was subsequently analyzed through a two-stage assessment process.

Based on the integrated theoretical framework of DeLone & McLean, UTAUT, and HOT-Fit as illustrated in Figure 1, this study proposes eleven hypotheses to evaluate the Sipentakel system, as described in Table 1.

Table 1. Hypothesis and Its Description

Hypothesis		**Description**
H_1	Human Factor has a positive and significant effect on Behavioral Intention.	The Human Factor encompasses performance expectancy, effort expectancy, and social influence, which act as direct determinants of behavioral intention. Consistent with Marsintauli et al. (2023) and Dewangga et al. (2025), these elements, particularly the expectation of improved job performance, significantly drive users' intent to adopt a system (Dewangga et al., 2025; Marsintauli et al., 2023).
H_{1a}	Job Level moderates the relationship between Human Factor and Behavioral Intention.	Hierarchical job positions correlate with varying administrative responsibilities, experiences, and resource access. As posited by Kuciapski (2019), the type of job position significantly influences technology acceptance (Kuciapski, 2019). Therefore, Job Level is proposed to moderate the impact of the Human Factor on Behavioral Intention to observe digital egalitarianism within bureaucracies.
H_2	Organization Factor has a positive and significant effect on Behavioral Intention.	The Organization Factor reflects the availability of technical resources, helpdesks, and management backing through facilitating conditions and organizational support. Adequate facilities and leadership support enhance users' confidence and their subsequent

Hypothesis		Description
		intention to utilize the technology (Marsintauli et al., 2023; Pamugar et al., 2014).
H_3	Technology Factor has a positive and significant effect on Behavioral Intention.	The Technology Factor represents intrinsic system qualities, including information, system, and service quality. High technical reliability and accurate outputs create positive user perceptions. Akbar & Fajar (2024) demonstrate that superior system and information quality significantly enhances the behavioral intention to use the system (Akbar & Fajar, 2024).
H_4	Technology Factor has a positive and significant effect on User Satisfaction.	System reliability, ease of access, and comprehensive information are prerequisites for user gratification. Studies confirm that high technology factor qualities directly and positively affect overall user satisfaction (Akbar & Fajar, 2024; Dewangga et al., 2025).
H_5	User Satisfaction has a positive and significant effect on Behavioral Intention.	User Satisfaction reflects the affective feedback following system interaction. Research by Dewangga et al. (2025) and Kurnianto et al. (2019) establishes that users who are satisfied with a system's interface and performance exhibit a stronger behavioral intention for continued and sustainable use (Dewangga et al., 2025; S. Kurnianto et al., 2019).
H_6	Behavioral Intention has a positive and significant effect on Net Benefits.	Net Benefits represent the ultimate positive impacts, such as improved operational efficiency, derived from system implementation. High usage intention drives actual adoption, which in turn realizes these benefits. Studies by Akbar & Fajar (2024) and Dewangga et al. (2025) validate that behavioral intention is a significant predictor of net benefits (Akbar & Fajar, 2024; Dewangga et al., 2025).
H_7	User Satisfaction has a positive and significant effect on Net Benefits.	According to the DeLone & McLean model, the realization of organizational and individual impacts is heavily dependent on user satisfaction. Consistent with Akbar & Fajar (2024) and Dewangga et al. (2025), satisfied users are more likely to achieve and report operational net benefits (Akbar & Fajar, 2024; Dewangga et al., 2025).
H_8	Organization Factor has a positive and significant effect on Human Factor.	The HOT-Fit framework emphasizes the necessary alignment between human and organizational components. As stated by Apsari et al. (2022) and Arsyad & Sfenrianto (2023), robust organizational support and strategic implementation directly influence human readiness and performance expectations (Apsari et al., 2022; Arsyad & Sfenrianto, 2023).
H_9	Technology Factor has a positive and significant effect on Organization Factor.	Technological quality shapes organizational capabilities. Arsyad & Sfenrianto (2023) and Apsari et al. (2022) highlight that the implementation of high-quality technology significantly strengthens organizational support systems and facilitating conditions (Apsari et al., 2022; Arsyad & Sfenrianto, 2023).
H_{10}	Technology Factor has a positive and significant effect on Human Factor.	User productivity and readiness are heavily influenced by their fit with the provided technology. Apsari et al. (2022) and Arsyad & Sfenrianto (2023) assert that superior system and information quality positively influences human expectations regarding performance and effort (Apsari et al., 2022; Arsyad & Sfenrianto, 2023).

Methodology

Research Design

This study employs a quantitative approach with a causal explanatory design (Richter et al., 2022). While primary data was collected via questionnaires, unstructured interviews and direct observations were utilized exclusively as supplementary methods to contextualize and enrich the interpretation of the quantitative findings, rather than constituting a formal mixed-methods design (Dewangga et al., 2025). The research evaluates the impact of independent variables—specifically human, organization, and technology factors—on dependent variables, including behavioral intention, user satisfaction, and net benefits, based on the integrated DeLone & McLean, UTAUT, and HOT-Fit models. The population for this study consists of 100 government employees at the Deputy for International Economic Cooperation and Investment who interact with the Sipentakel information system across various work units. A purposive sampling technique was utilized to select respondents based on specific criteria aligned with the research objectives, ensuring that the participants chosen are highly relevant and capable of providing accurate diagnostic data regarding the system's performance (Ahmed, 2024; Dewangga et al., 2025; Winarno, 2013).

Data Collection

Primary data was collected through an online questionnaire distributed via Google Forms during September 2025, utilizing a combination of closed and open-ended questions. A 4-point Likert scale (ranging from Strongly Disagree to Strongly Agree) was applied to the closed questions to explicitly avoid the central tendency effect, preventing respondents from selecting a neutral middle option (Dewangga et al., 2025). Open-ended questions were included to gather in-depth qualitative data regarding specific issues, critiques, and suggestions. To enrich and validate the quantitative findings, interviews were conducted with personnel involved in Sipentakel's design and management, alongside direct observations of user interactions with the system's interface. Furthermore, secondary data was gathered through a systematic literature review of relevant scientific journals, books, and regulations to build a strong theoretical foundation. Prior to widespread distribution, an Aiken's V test was conducted to ensure the readability and validity of the research instrument (B. Kurnianto & Sarwono, 2023). Following the "10-times rule" for partial least squares structural equation modeling (PLS-SEM), the minimum sample size required was determined to be 50 respondents, calculated as ten times the maximum number of structural paths (five) directed at the Behavioral Intention construct (J. F. Hair et al., 2021).

The population for this study consisted of 100 government employees at the Deputy of Economic Cooperation and Investment who use the Sipentakel information system. A sample of 52 respondents was selected using a purposive sampling technique to ensure that the participants were directly relevant to the research objectives (Ahmed, 2024; Winarno, 2013). The sample size adheres to the 10-times rule recommended for PLS-SEM (J. F. Hair et al., 2021). The study involved a sample of 52 respondents representing all work units within the Deputy of Economic Cooperation and Investment, with data collected via an online questionnaire. The demographic analysis focused on work unit, tenure, and specifically job level, as the latter is hypothesized and analyzed as a crucial moderating variable in the relationship between the human factor and behavioral intention.

Prior to participation, respondents were provided with an informed consent form integrated into the first page of the online questionnaire. This form explicitly stated the academic purpose of the study, guaranteed the confidentiality and anonymity of the respondents' data, and emphasized that participation was entirely voluntary.

The following in Table 2 are the respondents' overview based on work unit, job level, and work period.

Table 2. Respondents' Overview

	Respondent	Percentage (%)
Work Units		
Secretariat of the Deputy	13	25.0
Assistant Deputy of Bilateral Economic Cooperation	8	15.4
Assistant Deputy of Regional Economic Cooperation	7	13.5
Assistant Deputy of Multilateral Economic Cooperation	8	15.4
Assistant Deputy of Investment and Downstreaming Acceleration	10	19.2
Assistant Deputy of Macro and Fiscal Economy	6	11.5
Job Levels		
Echelon II	1	1.9
Echelon III or equivalent	3	5.8
Echelon IV or equivalent	9	17.3
Staffs	39	75.0
Work Period		
< 5 years	13	25.0
5–10 years	27	51.9
11–20 years	7	13.5
> 20 years	5	9.6

Data Analysis

The data analysis in this study was conducted using Structural Equation Modeling (SEM) with a Partial Least Squares (PLS) approach, known as PLS-SEM. This method was selected due to its capability to analyze complex research models and its primary focus on prediction (Richter et al., 2022). The conceptual model built in WarpPLS 8.0 as illustrated on Figure 2.

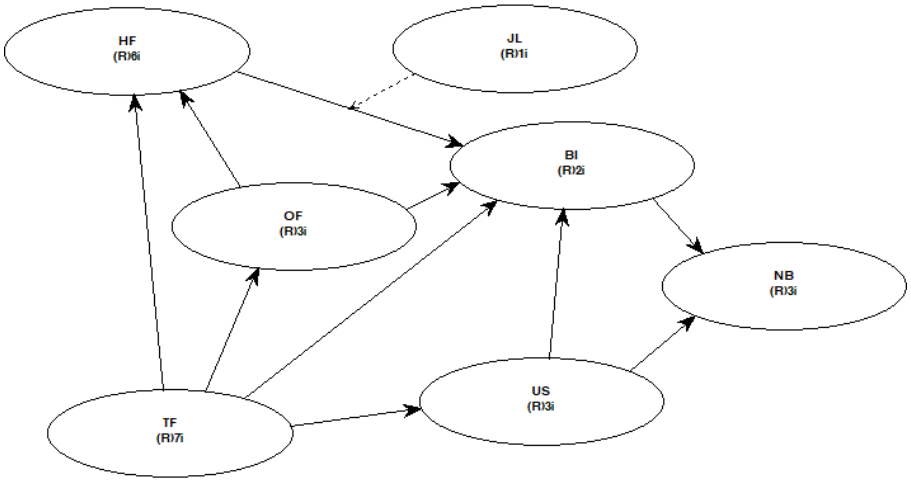


Figure 2. Conceptual Model Built in WarpPLS 8.0

The analysis process was executed using WarpPLS 8.0 software with Stable3 p-value calculation algorithm (Kock, 2024), encompassing two main stages: the evaluation of the measurement model (outer model) and the structural model (inner model).

The analysis of the measurement model (outer model) was conducted to ensure the validity and reliability of the constructs. The evaluation was based on the following criteria:

1. **Convergent Validity**

   Convergent validity assesses the extent to which a measure correlates positively with alternative measures of the same construct. It is evaluated using outer loadings and Average Variance Extracted (AVE). According to Hair et al. (2021), an indicator is considered valid if its outer loading value exceeds 0.708, and the construct's AVE value is greater than 0.50 (J. F. Hair et al., 2021).

2. **Discriminant Validity**

   Discriminant validity ensures that a construct is empirically distinct from other constructs in the structural model. This study employs the Heterotrait-Monotrait (HTMT) ratio of correlations, which is considered superior to the Fornell-Larcker criterion. To establish discriminant validity, the HTMT value should be below the threshold of 0.90 (J. F. Hair et al., 2021).

3. **Composite Reliability and Cronbach's Alpha**

   Reliability measures the consistency of the instrument. This study uses Composite Reliability and Cronbach's Alpha. A construct is deemed reliable if both values exceed 0.70 (J. F. Hair et al., 2021).

After confirming the validity and reliability of the measurement model, the structural model was evaluated to assess the causal relationships between constructs and the model's predictive power (J. F. Hair et al., 2021). This evaluation focuses on the Coefficient of Determination (R^2), Predictive Relevance (Q^2), Effect Size (f^2), and the significance of path coefficients.

Research Instruments

The structural model comprises six latent variables and one moderating variable, Job Level, which assesses the impact of job rank on the relationship between Human Factor and Behavioral Intention. The sole exogenous variable, Technology Factor, consists of 7 indicators measuring information quality, system quality, and service quality. The endogenous variables are operationalized as follows: Human Factor includes 6 indicators measuring performance expectancy, effort expectancy, and social influence; Organization Factor includes 3 indicators assessing facilitating conditions and organizational support; Behavioral Intention is measured by 2 indicators focused on intention to use; User Satisfaction utilizes 3 indicators ; and Net Benefits acts as the peak aggregating variable, measured by 4 indicators evaluating efficiency and effectiveness.

The measurement instruments were adapted from validated scales in prior studies, including Dewangga et al. (2025), Pamugar et al. (2014), and Apsari et al. (2022). The complete list of indicators and questionnaire items is provided in Appendix A.

Results

Measurement Model Evaluation (Outer Model)

The outer model evaluation aims to assess the validity and reliability of each construct used in the research. This process involves three key assessments: convergent validity, discriminant validity, and composite reliability (J. F. Hair et al., 2021).

Convergent Validity

Convergent validity measures the extent to which a construct explains the variance of its indicators (J. F. Hair et al., 2021). The assessment of convergent validity for reflective measurement models is based on the correlation between item scores and construct scores. The primary criteria used are the outer loading value of each indicator and the Average Variance Extracted (AVE). An indicator is considered reliable if it has an AVE value above 0.50 and an outer loading above 0.708, indicating that the construct is able to explain more than 50% of the indicator's variance (J. Hair & Alamer, 2022).

In the initial model analysis, the study included all hypothesized indicators. However, the evaluation revealed that several indicators did not meet the validity criteria. Specifically, they had outer loadings below 0.708, AVE values lower than other indicators in the construct, or contributed to high HTMT values and low internal consistency. Consequently, a stepwise removal of indicators was performed (J. F. Hair et al., 2021). The indicators excluded from the model were HF1, HF3, OF3, TF2, TF4, TF5, TF6, US1, US2, and NB3.

The convergent validity test result after a stepwise removal of indicator is as ilustrated on the Table 3.

Table 3. Outer Loading Test Result

	JF	HF	OF	TF	BI	US	NB	JL*HF
JL	-1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HF2	-0.336	(0.800)	-0.076	0.003	0.057	-0.516	0.062	0.329
HF4	0.146	(0.828)	-0.232	0.132	0.191	0.049	-0.132	-0.108
HF5	0.065	(0.857)	0.053	0.100	0.059	-0.045	0.046	-0.039
HF6	0.118	(0.785)	0.264	-0.251	-0.324	0.523	0.026	-0.179
OF1	-0.007	-0.109	(0.909)	-0.113	0.446	0.096	-0.368	-0.033
OF2	0.007	0.109	(0.909)	0.113	-0.446	-0.096	0.368	0.033
TF1	-0.067	-0.212	-0.055	(0.933)	-0.189	0.139	0.326	0.034
TF3	0.079	-0.032	0.038	(0.879)	0.275	0.032	-0.171	-0.129
TF7	-0.008	0.256	0.020	(0.882)	-0.074	-0.179	-0.174	0.093
BI1	-0.039	0.081	0.049	0.179	(0.942)	-0.261	-0.048	0.022
BI2	0.039	-0.081	-0.049	-0.179	(0.942)	0.261	0.048	-0.022
US3	0.000	0.000	0.000	0.000	0.000	-1.000	0.000	0.000
NB1	0.046	0.155	0.030	-0.074	-0.163	0.087	(0.962)	-0.049
NB2	-0.046	-0.155	-0.030	0.074	0.163	-0.087	(0.962)	0.049
JL*HF	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-1.000

Based on the analysis of the outer loadings presented in Table 3, all values have met the threshold of being above 0.708. The exceptionally high outer loading value (-1.000) is attributed to the fact that the construct consists of only one indicator (single-item construct) as the User Satisfaction consists only one indicator, justified by Hair that satisfaction-related construct oftenly being a single-item construct (J. F. Hair et al., 2021). Following the indicator removal process, the User Satisfaction (US) construct was reduced to a single-item construct. This is methodologically acceptable in PLS-SEM, as Hair et al. (2021) justify that constructs related to general satisfaction are often robustly and appropriately measured as single-item constructs without compromising validity (J. F. Hair et al., 2021).

Therefore, the outer loading test results demonstrate that the model fulfills the convergent validity requirements, allowing for the subsequent assessment of the AVE values.

Based on the AVE analysis results presented in Table 4, all AVE values have exceeded the 0.50 threshold. Therefore, the model testing results based on primary data are concluded to fulfill the convergent validity requirements.

Table 4. AVE Test Result

	JL	HF	OF	TF	BI	US	NB	JL*HF
JL	-1.000	-0.016	0.312	0.095	0.015	-0.158	0.163	0.241
HF	-0.016	(0.818)	0.393	0.761	0.738	0.667	0.688	-0.010
OF	0.312	0.393	(0.909)	0.459	0.288	0.281	0.398	-0.012
TF	0.095	0.761	0.459	(0.898)	0.728	0.723	0.687	0.067
BI	0.015	0.738	0.288	0.728	(0.942)	0.645	0.788	0.030
US	-0.158	0.667	0.281	0.723	0.645	-1.000	0.501	0.125
NB	0.163	0.688	0.398	0.687	0.788	0.501	(0.962)	0.008
JL*HF	0.241	-0.010	-0.012	0.067	0.030	0.125	0.008	-1.000

Discriminant Validity

Discriminant validity measures the extent to which a construct is empirically distinct from other constructs in the structural model. The recommended method for this assessment is the Heterotrait-Monotrait (HTMT) ratio of correlations (J. F. Hair et al., 2021). The HTMT criterion is considered more reliable than the Fornell-Larcker criterion and cross-loadings (J. F. Hair et al., 2021). The suggested threshold for HTMT is below 0.90 to ensure distinctiveness (J. F. Hair et al., 2021). The HTMT test results are presented in the Table 5.

Table 5. HTMT Test Result

	JF	HF	OF	TF	BI	US	NB	JL*HF
JF								
HF								
OF		0.485						
TF		0.888	0.551					
BI		0.863	0.347	0.830				
US								
NB		0.783	0.466	0.762	0.878			
JL*HF								

Based on the discriminant validity test results on Table 5, the HTMT indicated that all values in the adjusted model were within the acceptable range (< 0.90), confirming that the constructs are valid and distinct.

Composite Reliability and Cronbach's Alpha

Composite reliability measures the internal consistency of the indicators forming a construct. This test ensures that all indicators consistently measure the same construct. The metrics used are Composite Reliability and Cronbach's Alpha. A construct is considered to have good reliability if the values of Composite Reliability and Cronbach's Alpha are greater than 0.70 (J. F. Hair et al., 2021). These two indicators are used complementarily because Cronbach's Alpha tends to be conservative, producing lower reliability values, while Composite Reliability can be optimistic (J. F. Hair et al., 2021). Furthermore, the composite reliability and Cronbach's Alpha test result is as presented in Table 6.

Table 6. Composite Reliability and Cronbach's Alpha Test Result

	JF	HF	OF	TF	BI	US	NB	JL*HF
Composite Reliability	1.000	0.890	0.905	0.926	0.941	1.000	0.962	1.000
Cronbach's Alpha	1.000	0.835	0.791	0.880	0.874	1.000	0.920	1.000

Based on the assessment of Composite Reliability and Cronbach's Alpha presented in Table 6, all constructs generally fall within the acceptable thresholds. Notably, the Net Benefits (NB) construct exhibited a Composite Reliability value exceeding 0.95 (with a standard deviation of 0.012), while its corresponding Cronbach's Alpha remained within the excellent category.

It is important to acknowledge that reliability metrics can sometimes be overstated; therefore, a robust assessment requires a simultaneous examination of both Composite Reliability and Cronbach's Alpha (J. F. Hair et al., 2021). However, in practical PLS-SEM applications, a Composite Reliability score exceeding 0.95 remains valid and acceptable, as supported by recent empirical findings from (Rahmatullah et al., 2025). Consequently, based on the primary data analysis, it can be concluded that the measurement model satisfies the requirements for internal consistency reliability.

Structural Model Evaluation (Inner Model)

The structural model evaluation was conducted to assess the predictive capabilities of the model and the relationships between constructs. This assessment encompasses the coefficient of determination (R^2), predictive relevance (Q^2), effect size (f^2), and the significance of path coefficients.

Coefficient of Determination (R^2)

The R^2 value measures the model's predictive power, indicating the proportion of variance in the endogenous variable explained by the exogenous variables (J. F. Hair et al., 2021). According to Hair et al. (2021), R^2 values of 0.75, 0.50, and 0.25 are interpreted as strong, medium, and weak, respectively (J. F. Hair et al., 2021). Futhermore, the R^2 values are as presented in Table 7.

Table 7. Coefficient of Determination (R^2) Value

	JL	HF	OF	TF	BI	US	NB	JL*HF
R-squar	-	0.609	0.244	-	0.882	0.522	0.635	-
Adj, R-	-	0.593	0.229	-	0.869	0.513	0.620	-
Interpretation	-	Medium	Weak	-	Strong	Medium	Medium	-

Based on the coefficient of determination value presented in the Table 7, the analysis results demonstrate that the model possesses substantial predictive accuracy for Behavioral Intention (R^2= 0.882), indicating that 88.2% of the variance in intention is explained by the predictors. Meanwhile, the constructs of Net Benefits (R^2= 0.635) and User Satisfaction (R^2= 0.522) fall into the moderate category.

Predictive Relevance (Q^2)

To further validate the model's predictive accuracy, the Stone-Geisser Q^2 test was employed. This metric evaluates the model's capability to reconstruct observed values and parameter estimates (Ratnasari et al., 2022). A Q^2 value greater than zero indicates that the model has predictive relevance for a specific endogenous construct (J. F. Hair et al., 2021). Futhermore, the Q^2 values are as presented in Table 8.

Table 8. Predictive Relevance (Q^2) Value

	JL	HF	OF	TF	BI	US	NB	JL*HF
Q-squar	-	0.609	0.233	-	0.745	0.534	0.631	-

Based on the predictive relevance value presented in Table 8, the results confirm that all endogenous variables exhibit values above zero, thereby confirming the model's predictive relevance.

Effect Size (f^2)

The effect size (f^2) assesses the relative impact of a predictor construct on an endogenous construct regarding the f^2 value. Hair and Alamer (2022) classify f^2 values of 0.02, 0.15, and 0.35 as representing

small, medium, and large effects, respectively (J. Hair & Alamer, 2022). Futhermore, the f^2 values are as presented in Table 9.

Table 9. Effect Size (f^2) Value

No	Constructs Relation	f^2	Interpretation
1	HF → BI	0.302	Medium
1_a	JL.HF → BI	0.038	Small
2	OF → BI	0.025	Small
3	TF → BI	0.474	Large
4	TF → US	0.522	Large
5	US → BI	0.044	Small
6	BI → NB	0.631	Large
7	US → NB	0.005	Very Small
8	OF → HF	0.073	Small
9	TF → OF	0.244	Medium
10	TF → HF	0.536	Large

Based on the effect size value presented in Table 9, the analysis reveals varying degrees of impact, for instance, the Technology Factor demonstrates a large effect on both Behavioral Intention (f^2 = 0.474) and User Satisfaction (f^2 = 0.522). In contrast, the Organization Factor shows a small effect on Behavioral Intention (f^2 = 0.025).

Path Coefficients and Hypothesis Testing

Hypothesis testing was performed by examining the path coefficients and their corresponding significance levels. The significance was determined using the Stable3 algorithm in WarpPLS 8.0, a quasi-parametric approach based on exponential smoothing for p-value calculation (Kock, 2024). Hypothesis testing was conducted applying a 5% significance level (α = 0.05) using a two-tailed approach, with a critical t-value of 1.960 (J. F. Hair et al., 2021). A relationship between constructs is considered statistically significant if the *p-value* is less than the significance level of 0.05 (J. F. Hair et al., 2021). The specific results of the path analysis, which serve as the basis for accepting or rejecting the proposed hypotheses, are detailed in Table 10.

Table 10. Inner Model Calculation

No	Construct Relation	*p-value*	*Path Coefficient*	*t-value*	Hypothesis Status
1	HF → BI	<0.001	0.407	3.423	Supported
1_a	JL.HF → BI	0.123	0.153	1.172	Rejected
2	OF → BI	0.266	0.085	0.630	Rejected
3	TF → BI	<0.001	0.571	5.170	Supported
4	TF → US	<0.001	0.723	6.844	Supported
5	US → BI	0.309	0.068	0.500	Rejected
6	BI → NB	<0.001	0.791	7.687	Supported
7	US → NB	0.473	0.009	0.068	Rejected
8	OF → HF	0.122	0.154	1.177	Rejected
9	TF → OF	<0.001	0.494	4.295	Supported
10	TF → HF	<0.001	0.687	6.542	Supported

In summary, the structural model analysis indicates that hypotheses H_1, H_3, H_4, H_6, H_9, and H_{10} were supported, whereas H_{1a}, H_2, H_5, H_7, and H_8 were rejected. The path coefficients across the hypothesized relationships exhibited considerable variation. The finalized structural relationships following the analysis are illustrated in Figure 3. Note: Solid green lines indicate accepted hypotheses; thin red lines indicate rejected hypotheses.

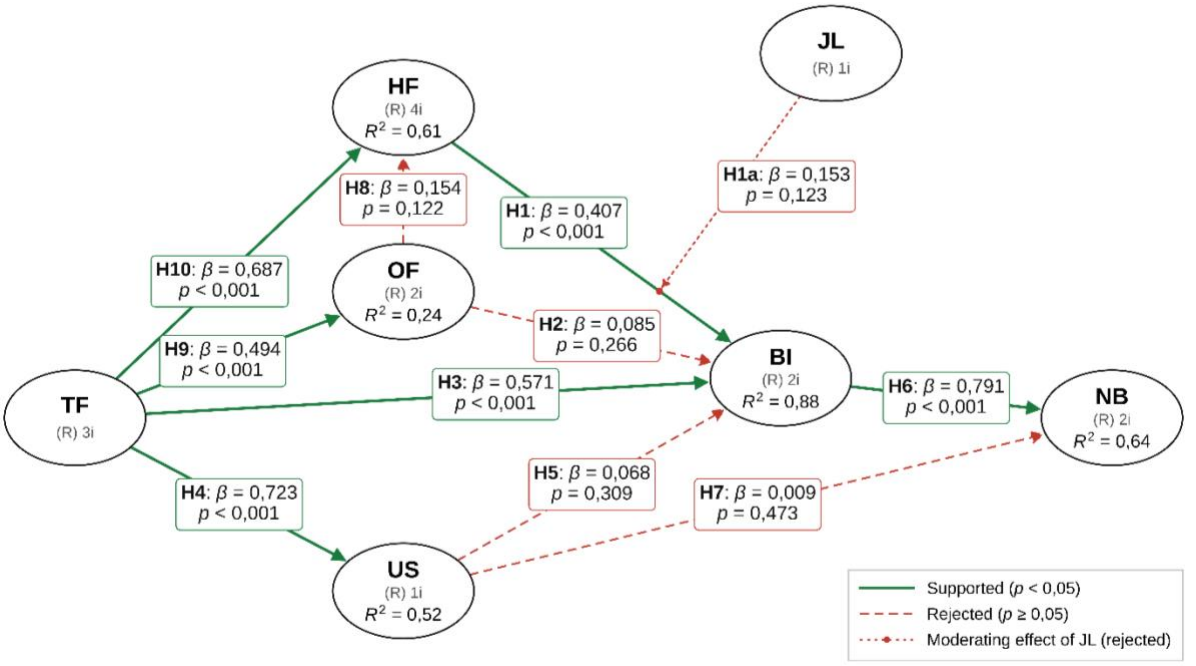


Figure 3. Structural Model Result

Discussions

The study highlights that Human Factors and Technology Factors significantly influence Behavioral Intention. This finding aligns with previous studies, such as Dewangga et al. (2025), which found that user intention is strongly driven by performance expectancy and system quality. In the context of Sipentakel, qualitative observations of user interactions support this outcome; when employees perceive the system as facilitating administrative tasks—such as efficiently accessing key performance indicators or managing travel orders—and see it being embraced by their peers, their motivation to use it increases. Direct observations confirmed that seamless navigation through these specific modules visibly reduces administrative bottlenecks, reinforcing why enhancing intrinsic system functionality and information completeness is more critical than formal persuasion.

Additionally, this study confirms that Technology Factors significantly boost User Satisfaction. This result is in line with Apsari et al. (2022), indicating that accurate information and system reliability are prerequisites for user gratification. However, an interesting insight from this study is that User Satisfaction is not a significant predictor of Behavioral Intention or Net Benefits, which contrasts with the findings of Dewangga et al. (2025). One logical reason for this is the mandatory nature of the system within the Deputy of Economic Cooperation and Investment. This statistical paradox is strongly validated by qualitative insights gathered from unstructured interviews. Interviewees expressed that because their primary daily workload consists of substantive policy tasks, administrative tools like Sipentakel are viewed strictly through a functional lens. Since the system is integral to daily workflows, employees are compelled to use it to fulfill bureaucratic obligations quickly, regardless of their satisfaction levels. This reflects a utilitarian culture where users focus on Net Benefits (efficiency and effectiveness) achieved through the act of using the system (Behavioral Intention), rather than emotional satisfaction.

Furthermore, regarding organizational dynamics, the study reveals that Job Level does not moderate the relationship between Human Factors and Behavioral Intention. This contrasts with Kuciapski (2019)

and suggests a form of digital egalitarianism where the system's utility is perceived equally across hierarchies, from high-ranking officials to execution staff. Similarly, Organization Factors were found to have no significant impact on Behavioral Intention or Human Factors. Qualitative interviews further enriched this finding; users explicitly confirmed that formal technical guidance is currently absent and organizational support is limited primarily to domain and server hosting. This lack of impact can be attributed to the users' self-sufficiency. Despite the limited formal support (e.g., lack of a dedicated helpdesk), users reported relying on the system's intuitive design. Thus, Technology Factors (quality and ease of use) effectively bridge the gap, as evidenced by the significant influence of Technology on both Organization and Human Factors. This implies that in this specific government setting, investing in system quality yields a higher return on adoption than investing in traditional organizational support mechanisms.

Implications

Theoretical & Managerial Implications

The theoretical implication of this study is that it extends the integrated evaluation framework of DeLone & McLean, UTAUT, and HOT-Fit by applying it to the Government-to-Employee (G2E) context in Indonesia, specifically to explore the determinants of Sipentakel system adoption. While previous research often examines these models in isolation or focuses heavily on voluntary user settings, this study fills a significant gap by investigating mandatory digital adoption within a high-stakes bureaucratic environment. The findings suggest that Technology Factors and Human Factors play a more dominant role than User Satisfaction and Organization Factors in the Indonesian government context. This contribution challenges the conventional DeLone & McLean theory by providing empirical evidence of a satisfaction paradox, where satisfaction is not a significant predictor of continuance intention or net benefits in mandatory settings. This highlights that in a rigid administrative culture, usage is driven primarily by functional utility and obligation rather than affective gratification. Furthermore, the study introduces the concept of digital egalitarianism by demonstrating that Job Level does not moderate adoption intention, extending the understanding that digital acceptance in modern bureaucracies is uniform across hierarchical lines.

Managerial implications of this study are that it provides actionable insights for the Secretariat of Deputy of Economic Cooperation and Investment to optimize the implementation of Sipentakel and future e-government systems. To ensure sustained engagement and budget efficiency, the Deputy Secretariat should pivot its strategy from satisfaction-centric to utility-centric. Since Organization Factors (such as formal training and helpdesks) were found to have no significant impact on intention, the organization can implement a self-service support strategy. This involves reducing budget allocation for formal technical guidance aligning with Presidential Instruction No. 1 of 2025 regarding expenditure efficiency and redirecting resources towards enhancing System Quality and Information Quality. By improving the system's intrinsic reliability and data completeness, the system becomes a facilitator that drives adoption without the need for extensive extrinsic support. Additionally, given that Job Level is not a barrier, policymakers should deploy a unified internal communication strategy that emphasizes practical efficiency gains for all employees, regardless of their rank, ensuring that the system is perceived as a critical tool for simplifying their daily administrative duties.

Conclusions

This study concludes that the success of the mandatory Sipentakel G2E system is fundamentally driven by Technology Factors and Human Factors. Specifically, the intention to use the system is significantly influenced by intrinsic system quality, information accuracy, and user expectations of performance, rather than by formal Organization Factors or overall User Satisfaction. A critical finding is that Net Benefits, primarily defined by operational efficiency, are realized solely through active usage (Behavioral Intention) independent of the user's emotional gratification. Furthermore, the assessment theoretically contributes by confirming a phenomenon of digital egalitarianism within a bureaucratic setting, where hierarchical Job Level does not moderate system acceptance, contrasting previous

assumptions regarding hierarchical influence on adoption (Kuciapski, 2019). The system's utility is perceived uniformly across hierarchical lines, from high-ranking officials to execution staff.

This research is subject to several methodological limitations. The relatively small sample size (n=52) evaluated against a highly complex conceptual model comprising 11 hypotheses introduces a potential risk of overfitting. Although the sample size satisfies the minimum threshold of the "10-times rule" commonly accepted in PLS-SEM literature (J. F. Hair et al., 2021), the high ratio of estimated structural paths to observations implies that model parsimony is stretched, which may limit the broader generalizability of these findings (Richter et al., 2022).

The rigorous measurement model evaluation necessitated the removal of numerous indicators to fulfill the strict criteria for convergent and discriminant validity (J. F. Hair et al., 2021). While this step ensured statistical reliability, it inherently compromised content validity by narrowing the conceptual meaning of several constructs. For instance, the necessary removal of the Service Quality and Organization Support dimensions restricted the original theoretical breadth of the Technology and Organization Factors (Pamugar et al., 2014). Furthermore, reducing Net Benefits solely to the Efficiency dimension constrained the holistic measurement of the system's impact (Rahmatullah et al., 2025).

To address these methodological constraints, future research should expand the population scope to secure much larger sample sizes. A larger sample would not only mitigate the risk of overfitting but also allow for the application of Covariance-Based SEM (CB-SEM) to perform stricter model fit evaluations. Additionally, utilizing a wider rating scale (e.g., 5-point or 7-point Likert scales) is strongly recommended to capture more nuanced user response variance (Alan & Atalay Kabasakal, 2020; Menon et al., 2025). This adjustment may prevent the excessive dropping of indicators, thereby preserving the comprehensive theoretical meaning of the integrated constructs in future evaluations.

References

Ahmed, S. K. (2024). How to choose a sampling technique and determine sample size for research: A simplified guide for researchers. Oral Oncology Reports, 12, Article 100662. https://doi.org/10.1016/j.oor.2024.100662

Akbar, M. K., & Fajar, A. N. (2024). Evaluation of the warehouse management system application using DeLone and McLean model in North Jakarta Indonesia. Ingénierie des Systèmes d'Information, 29(6), 2515–2524. https://doi.org/10.18280/isi.290636

Alan, Ü., & Atalay Kabasakal, K. (2020). Effect of number of response options on the psychometric properties of Likert-type scales used with children. Studies in Educational Evaluation, 66, Article 100895. https://doi.org/10.1016/j.stueduc.2020.100895

Almaiah, M. A., Alamri, M. M., & Al-Rahmi, W. (2019). Applying the UTAUT model to explain the students' acceptance of mobile learning system in higher education. IEEE Access, 7, 174673–174686. https://doi.org/10.1109/ACCESS.2019.2957206

Apsari, N. W. N., Suaryana, I. G. N. A., Wirajaya, I. G. A., & Ratna Sari, M. M. (2022). Evaluasi keberhasilan dan penerimaan sistem keuangan desa menggunakan model terintegrasi di Kabupaten Tabanan. E-Jurnal Ekonomi dan Bisnis Universitas Udayana, 11(01), 107–116. https://doi.org/10.24843/EEB.2022.v11.i01.p10

Arsyad, M. D., & Sfenrianto. (2023). Evaluation of automated reconciliation application at Bank XYZ using HOT-Fit model. Journal of System and Management Sciences, 13(1), 117–132. https://doi.org/10.33168/JSMS.2023.0107

Ayuardini, M., & Ridwan, A. (2019). Implementasi metode HOT FIT pada evaluasi tingkat kesuksesan sistem pengisian KRS terkomputerisasi. Faktor Exacta, 12(2), 122–131. https://doi.org/10.30998/faktorexacta.v12i2.3639

Dewangga, A. A. B. D. P., Ariyanto, D., Wirakusuma, M. G., & Sujana, I. K. (2025). Integrasi model DeLone & McLean, UTAUT, dan HOT-Fit untuk menilai keberhasilan dan keberterimaan aplikasi LPD Mobile di Bali. Jurnal Sistem Informasi Bisnis, 15(1), 34–44. https://doi.org/10.14710/vol15iss1pp34-44

Hair, J., & Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second

language and education research: Guidelines using an applied example. Research Methods in Applied Linguistics, 1(3), Article 100027. https://doi.org/10.1016/j.rmal.2022.100027

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). A primer on partial least squares structural equation modeling (PLS-SEM) (3rd ed.). SAGE Publications.

Hidayatullah, S., Alvianna, S., Sugeha, A. Z., & Astuti, W. (2022). Model of information systems success DeLone and McLean in using PeduliLindungi application in the tourism sector of Malang City. Jurnal Pariwisata Pesona, 7(1), 49–57. https://doi.org/10.26905/jpp.v7i1.7505

Iswanto, D., & Putra, R. A. (2023). Evaluation of bureaucratic reforms in realizing world class government in Tuban. Spirit Publik: Jurnal Administrasi Publik, 18(2), 139–156. https://doi.org/10.20961/sp.v18i2.72811

Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi. (2023). Peraturan Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 9 Tahun 2023 tentang Evaluasi Reformasi Birokrasi. https://peraturan.bpk.go.id/Details/266063/permen-panrb-no-9-tahun-2023

Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi. (2024). Peraturan Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 8 Tahun 2024 tentang Tata Cara Penyesuaian Tunjangan Kinerja. https://peraturan.bpk.go.id/Details/297886

Khairani, P., Hamzah, M. L., Megawati, M., & Jazman, M. (2022). Evaluasi kinerja Sistem Informasi Elektronik Kinerja ASN (SI-EKA) menggunakan framework COBIT 5 pada Kementerian Agama Kota Pekanbaru. INTECOMS: Journal of Information Technology and Computer Science, 5(1), 8–18. https://doi.org/10.31539/intecoms.v5i1.3693

Kock, N. (2024). WarpPLS user manual: Version 8.0. ScriptWarp Systems. https://www.scriptwarp.com/warppls/UserManual_v_8_0.pdf

Kuciapski, M. (2019). How the type of job position influences technology acceptance: A study of employees' intention to use mobile technologies for knowledge transfer. IEEE Access, 7, 177397–177413. https://doi.org/10.1109/ACCESS.2019.2957205

Kumorowani, R. P., & Mulyanti, D. (2023). Analisis penerapan sistem informasi manajemen rumah sakit (SIMRS) dengan pendekatan HOT FIT: Systematic literature review. DIAGNOSA: Jurnal Ilmu Kesehatan dan Keperawatan, 1(2), 25–33. https://doi.org/10.59581/diagnosa-widyakarya.v1i2.174

Kurnianto, B., & Sarwono, R. (2023). Pengembangan perangkat pembelajaran berbasis TPACK dalam meningkatkan aktivitas belajar dan kemampuan pemecahan masalah siswa. Scholaria: Jurnal Pendidikan dan Kebudayaan, 13(3), 210–221. https://doi.org/10.24246/j.js.2023.v13.i3.p210-221

Kurnianto, S., Kurniawansyah, D., & Ekasari, W. F. (2019). Menilai keberhasilan sistem keuangan desa (Siskeudes): Validasi model keberhasilan sistem informasi DeLone dan McLean. Jurnal Riset Akuntansi dan Bisnis Airlangga, 4(2), 687–706. https://doi.org/10.31093/jraba.v4i2.177

Marsintauli, F., Chosmalianto, D. Y., & Zaldin, A. (2023). Evaluation of e-government systems to realize sustainable development: E-invoicing web systems as a case study. E3S Web of Conferences, 426, Article 02041. https://doi.org/10.1051/e3sconf/202342602041

Marsintauli, F., & Sari, M. (2023). Evaluation of ERP Oracle NetSuite implementation using the unified theory of acceptance and use of technology (UTAUT) model to create a sustainable business. E3S Web of Conferences, 426, Article 02037. https://doi.org/10.1051/e3sconf/202342602037

Meilani, L., Suroso, A. I., & Yuliati, L. N. (2020). Evaluasi keberhasilan sistem informasi akademik dengan pendekatan model DeLone dan McLean. Jurnal Sistem Informasi Bisnis, 10(2), 137–144. https://doi.org/10.21456/vol10iss2pp137-144

Menon, V., Grover, S., Gupta, S., Indu, P. V., Chacko, D., & Vidhukumar, K. (2025). A primer on reliability testing of a rating scale. Indian Journal of Psychiatry, 67(7), 725–729. https://doi.org/10.4103/indianjpsychiatry_584_25

Nugroho, A. P., Zulni, D. A., & Andriansyah, Y. (2023). Exploring the adoption of digital wallets among Islamic millennials in Yogyakarta, Indonesia using an extended UTAUT model: The role of Islamic consumption ethics. Journal of Theoretical and Applied Information Technology, 101(14), 5528–5540. http://www.jatit.org/volumes/Vol101No14/9Vol101No14.pdf

Omol, E., & Ondiek, C. (2021). Technological innovations utilization framework: The complementary powers of UTAUT, HOT-Fit framework and DeLone and McLean IS model. International Journal of Scientific and Research Publications, 11(9), 146–151. https://doi.org/10.29322/IJSRP.11.09.2021.p11720

Pamugar, H., Winarno, W. W., & Najib, W. (2014). Model evaluasi kesuksesan dan penerimaan sistem informasi e-learning pada lembaga diklat pemerintah. Scientific Journal of Informatics, 1(1), 13–27. https://doi.org/10.15294/sji.v1i1.3638

Presiden Republik Indonesia. (2024). Peraturan Presiden Nomor 143 Tahun 2024 tentang Kementerian Koordinator Bidang Perekonomian. https://peraturan.bpk.go.id/Details/306481

Purwanto, E., & Loisa, J. (2020). The intention and use behaviour of the mobile banking system in Indonesia: UTAUT model. Technology Reports of Kansai University, 62(6), 2757–2767.

Purwati, A. A., Mustafa, Z., & Deli, M. M. (2021). Management information system in evaluation of BCA mobile banking using DeLone and McLean model. Journal of Applied Engineering and Technological Science, 2(2), 70–77. https://doi.org/10.37385/jaets.v2i2.217

Putri, A. A., Jayadianti, H., & Yuwono, B. (2021). Evaluation of Jogja application success from user's perspective using development of DeLone and McLean models to support the realization of the smart province. Telematika, 18(2), 181–193. https://doi.org/10.31315/telematika.v18i2.5316

Rahmatullah, R., Habibi, A., Khaeruddin, K., Yaqin, L. N., Alharmali, T. M., Fauzee, M. S. O., & Mahat, J. (2025). A study of user satisfaction and net benefits in Indonesia through the DeLone and McLean model for e-government success. Discover Sustainability, 6, Article 710. https://doi.org/10.1007/s43621-025-01645-4

Ratnasari, D. T., Saud, U. S., Agustin, M., & Permana, J. (2022). The effect of principal leadership on the teaching performance of elementary school teachers and their impact on school quality in Lebak-Banten District. Proceedings of the International Conference on Elementary Education, 4(1), 152–162. https://proceedings2.upi.edu/index.php/icee/article/view/1982

Richter, N. F., Hauff, S., Ringle, C. M., & Gudergan, S. P. (2022). The use of partial least squares structural equation modeling and complementary methods in international management research. Management International Review, 62(4), 449–470. https://doi.org/10.1007/s11575-022-00475-0

Rifial, M., Razak, A., Darmawansyah, D., Indar, I., & Rahman, A. (2025). Evaluation of the utilization of the hospital management information system (SIMRS) at Madani Regional General Hospital, Palu. Journal of Public Health and Pharmacy, 5(2), 274–286. https://doi.org/10.56338/jphp.v5i2.6169

Saputro, P. H., Faizah, A., & J.F, R. A. (2020). Evaluasi revolusi industri 4.0 pada bidang pertanian menggunakan model integrasi DeLone and McLean, UTAUT dan HOT Fit. Jurnal Ilmiah Informatika, 8(1), 53–58. https://doi.org/10.33884/jif.v8i01.1845

Sari, M. M., Sanjaya, G. Y., & Meliala, A. (2016). Evaluasi sistem informasi manajemen rumah sakit (SIMRS) dengan kerangka HOT-FIT. In Seminar Nasional Sistem Informasi Indonesia (SESINDO 2016). http://is.its.ac.id/pubs/oajis/index.php/home/detail/1665/Evaluasi-Sistem-Informasi-Manajemen-Rumah-Sakit-SIMRS-Dengan-Kerangka-HOT-FIT

Syaputra, H., Abdullah, A., & Sjafirah, N. A. (2023). Transformation, internalization, and media socialization of core values for state civil apparatus (ASN). Jurnal Lensa Mutiara Komunikasi, 7(1), 11–22. https://doi.org/10.51544/jlmk.v7i1.3812

Weigl, L., Roth, T., Amard, A., & Zavolokina, L. (2024). When public values and user-centricity in e-government collide – A systematic review. Government Information Quarterly, 41(3), Article 101956. https://doi.org/10.1016/j.giq.2024.101956

Winarno, M. E. (2013). Metodologi penelitian dalam pendidikan jasmani. UM Press.

Xue, L., Rashid, A. M., & Ouyang, S. (2024). The unified theory of acceptance and use of technology (UTAUT) in higher education: A systematic review. SAGE Open, 14(1). https://doi.org/10.1177/21582440241229570

How to cite:

Pangantyo, H. K., & Purwaningsih, M. (2026). Examining Digital Egalitarianism through the Moderating Effect of Job Level on G2E Systems Adoption. *Jurnal Sistem Informasi (Journal of Information System), 22*(2), 43–63.

Appendices

Appendix A
Research Instrument

Construct	Dimension	Questionnaire Item	Adapted From
Human Factor (HF)	Performance Expectancy	1. Sipentakel increases my work productivity.	Dewangga et al. (2025); Pamugar et al. (2014)
		2. Sipentakel makes program administration management easier.	
	Effort Expectancy	1. Sipentakel is easy to use.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. Sipentakel makes it easier to obtain information.	
	Social Influence	1. My work environment considers the use of Sipentakel important.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. My work environment supports the use of Sipentakel.	
Organization Factor (OF)	Facilitating Condition	1. The Deputy Secretariat provides adequate resources and technical support.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. The Deputy Secretariat provides the necessary guidance for users.	
	Organization Support	1. The Deputy Secretariat encourages the use of Sipentakel in daily work.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
Technology Factor (TF)	Information Quality	1. Sipentakel provides complete information.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. Sipentakel provides accurate information.	
		3. Sipentakel provides information relevant to my work.	
	System Quality	1. Sipentakel has an attractive interface and features.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. Sipentakel is easily accessible.	
	Service Quality	1. Sipentakel administrators provide quick responses to complaints.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. Sipentakel provides easy-to-understand usage instructions.	
Behavioral Intention (BI)	Intention to Use	1. I will continue to use Sipentakel in my work.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. I recommend Sipentakel to my colleagues.	
User Satisfaction (US)	Use Satisfaction	1. I am satisfied with the Sipentakel interface.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. I am satisfied with the information provided by Sipentakel.	
		3. Overall, I am satisfied using Sipentakel.	

Construct	Dimension	Questionnaire Item	Adapted From
Net Benefits (NB)	Efficiency	1. Sipentakel speeds up the completion of my work.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)
		2. Sipentakel makes my work easier.	
	Effectiveness	1. Sipentakel helps reduce errors in my work.	Apsari et al. (2022); Dewangga et al. (2025); Pamugar et al. (2014)

Hybrid Cybersecurity Compliance Audit Framework for an On-Premise Hotel Property Management System: Mapping NIST CSF 2.0 Controls to Indonesia's Personal Data Protection Law

I Made Sudana*
Master of Informatics
Universitas Muhammadiyah Surakarta
Surakarta, Indonesia
l208250008@student.ums.ac.id

Bana Handaga
Master of Informatics
Universitas Muhammadiyah Surakarta
Surakarta, Indonesia
bana@ums.ac.id

Jan Wantoro
Master of Informatics
Universitas Muhammadiyah Surakarta
Surakarta, Indonesia
jan@ums.ac.id

Abstract

The accelerating digitalization of Property Management Systems (PMS) in the hospitality sector introduces substantial cybersecurity exposure, particularly concerning the protection of guests' personal data. This study audits the cybersecurity posture of an on-premises e-VHP PMS operated by an international chain hotel in Surakarta, Indonesia, whose local PMS core is integrated with cloud-based third-party Channel Manager and Online Travel Agency (OTA) platforms managed at the corporate level. A hybrid audit instrument was developed to bridge technical controls under the NIST Cybersecurity Framework (CSF) 2.0 with legal obligations under Indonesia's Personal Data Protection Law (UU PDP) No. 27/2022. Primary data were collected through structured interviews with six informants representing three organizational layers: managerial (General Manager), technical (two IT Officers), and operational/front-line staff (Front Office, General Cashier/Accounts Payable, and Income Audit/Accounts Receivable), using a five-point maturity rubric (1 = Not Implemented to 5 = Optimized and Continuously Reviewed). Aggregated maturity scores showed that Detect (3.33) and Govern (3.00) achieved the highest relative maturity, whereas Respond (1.50) and Recover (1.89) showed the most severe capability deficits; Protect (2.50) and Identify (2.44) fell in between. Gap analysis against the optimal target (5.00) placed Respond at the highest remediation priority (Gap = 3.50), followed by Recover (Gap = 3.11). A structured risk register identified critical exposure in endpoint backup failure and unmanaged credential sharing among front-line staff, alongside high power-management and integration-channel risks. Structural mapping shows these technical and behavioral gaps correspond to potential non-compliance with Articles 35, 39, and 46 of the UU PDP, exposing the organization to administrative fines of up to 2% of annual revenue. The study proposes a tiered Cybersecurity Governance Blueprint emphasizing local incident-response

* Corresponding Author

capability, standardized power-backup procurement, encrypted endpoint backup automation, and a structured security-awareness program for front-line personnel.

Keywords: cybersecurity audit, hotel PMS, NIST CSF 2.0, Personal Data Protection Law (UU PDP), multi-stakeholder risk assessment

Introduction

The rapid digitalization of hotel operations through Property Management Systems (PMS) has expanded the cyber-attack surface of the hospitality industry, which manages highly sensitive guest data, including Personally Identifiable Information (PII), passport records, and payment credentials (Karadayi-Usta, 2024; Pipyros & Liasidou, 2025). In Indonesia, this exposure is compounded by the enactment of Law No. 27 of 2022 on Personal Data Protection (UU PDP), which imposes substantial administrative and criminal liability on data controllers that fail to safeguard consumer data (Sari et al., 2023; Wibowo et al., 2023).

Within hotel operations, security responsibility is frequently distributed across a Shared Responsibility Model that is more fragmented than the conventional cloud-vendor/customer dichotomy typically discussed in the literature. Field evidence collected for this study indicates that the audited property's PMS core operates on-premises, within a closed local network, while cloud-based connectivity is confined to a centrally managed Channel Manager integration linking the property to third-party Online Travel Agencies (OTAs). This hybrid on-premise/cloud-integration architecture means that local management retains direct operational responsibility for physical infrastructure resilience (power continuity, endpoint backup, local access control), while encryption and integration-channel security for external data flows are governed at the corporate level. This nuance is rarely addressed in prior literature, which tends to assume either a fully cloud-hosted PMS or a fully isolated on-premises system (Demir & Demir, 2025; Kim et al., 2013).

Prior research has generally bifurcated into two strands: technical cybersecurity maturity assessments that do not engage legal compliance implications (Bernardo et al., 2025; Casaril & Galletta, 2025), and legal-compliance studies that address regulatory obligations without operational or infrastructural specificity (Voigt & Von dem Bussche, 2017). Few studies integrate multi-stakeholder field evidence spanning managerial, technical, and front-line operational perspectives into a single audit instrument that links infrastructural weaknesses to specific statutory provisions.

This study addresses that gap by implementing a hybrid audit instrument that integrates: (1) standardized technical controls based on NIST CSF 2.0, (2) a structured risk quantification approach informed by triangulated interview data from six respondents across three organizational layers, and (3) structural mapping of audit findings to specific UU PDP provisions. The resulting Cybersecurity Governance Blueprint is intended to support continuous, evidence-based compliance rather than a one-time, point-in-time assessment.

To address this gap, this study proposes a novel Hybrid Cybersecurity Compliance Framework that systematically integrates the technical sub-categories of the NIST Cybersecurity Framework (NIST CSF 2.0), specifically the *Protect* and *Detect* functions, with the statutory data governance mandates of the UU PDP Law. Unlike conventional case-based audits that merely report post-hoc compliance failures, the theoretical contribution of this research lies in establishing a structured, risk-prioritized mapping mechanism. This framework provides a replicable conceptual bridge that translates technical host and network hardening metrics directly into legal compliance maturity levels for hospitality operations.

- A traceable audit artifact systematically mapping NIST CSF 2.0 sub-categories to specific UU PDP statutory obligations addressing a methodological gap in the Indonesian regulatory landscape.
- The first multi-layer empirical evidence detailing the security posture of hybrid on-premise/cloud-integrated hotel PMS architectures.

- A conceptual bridge demonstrating how physical infrastructure gaps (e.g., power continuity, local backup failures) map directly to statutory legal exposures under data protection laws.

Literature Review

Research on information-system audits and cybersecurity governance across industries commonly adopts international standard frameworks such as the NIST Cybersecurity Framework (CSF) and ISO/IEC 27001 as primary references for evaluating information-technology security posture (Culot et al., 2021). Much of this literature concentrates on banking and critical-network environments, given their high material-loss exposure (Abbas & Alazawi, 2026), and demonstrates that NIST CSF and ISO/IEC 27001 are effective for mapping security gaps, identifying network-penetration threats, and periodically measuring cyber-capability maturity (Abbas & Alazawi, 2026; Bernardo et al., 2025; Brezavšček & Baggia, 2025). However, these studies typically evaluate rigid information-system architectures supported by dedicated internal IT teams (Büyüközkan & Güler, 2025), so their resulting recommendations are often difficult to apply directly to organizations with more dynamic operations and limited on-site IT staffing.

A second strand of literature addresses the legal dimension of data security and privacy regulation, both global instruments such as the General Data Protection Regulation (GDPR) in Europe and domestic instruments such as Indonesia's UU PDP, which has grown substantially in recent years (Republik Indonesia, 2022; Voigt & Von dem Bussche, 2017). This strand generally focuses on organizational compliance readiness from a legal, data-governance, and macro risk-management perspective (Badawi et al., 2026), examining how UU PDP provisions shape data-controller obligations and legal certainty (Sari et al., 2023; Wibowo et al., 2023). While this work contributes significantly to the legal dimension, it generally lacks applicable operational or technical guidance, remaining largely theoretical-juridical and offering little direct intervention for site-level hardware or infrastructural failures. These studies have not yet bridged how physical anomalies, such as the absence of backup power at a core router or failure of endpoint backup systems, can directly satisfy the elements of a data-protection-regulation violation.

Given the limitations of both research strands, studies specifically auditing cybersecurity within hotel Property Management Systems (PMS) remain underdeveloped. Most conventional hospitality-technology research evaluates information-technology adoption from the standpoint of investment strategy, operational efficiency, or macro-level technology utilization (Demir & Demir, 2025; Cobanoglu et al., 2021), often overlooking the fact that integrating hotel information systems with third-party platforms creates an attack surface highly susceptible to data-security incidents (Karadayi-Usta, 2024; Kim et al., 2013). Where hospitality-sector cybersecurity research does exist, its scope is typically limited to generic risk-assessment techniques or macro-level threat awareness rather than in-depth audits of the underlying local infrastructure (Campoverde-Molina & Luján-Mora, 2025; Pipyros & Liasidou, 2025). Consequently, the hospitality industry may develop a false sense of security from adopting internationally branded information systems while remaining unaware of latent vulnerabilities in local infrastructure.

This study addresses this research gap by integrating the technical defenses of NIST CSF 2.0 with the legal-compliance obligations of UU PDP No. 27/2022 through a hybrid, multi-stakeholder audit approach (National Institute of Standards and Technology, 2024; Republik Indonesia, 2022). In contrast to prior single-method, single-informant studies, this research translates field-level infrastructural findings such as inconsistent UPS coverage and unrealized endpoint backup routines into concrete legal-compliance risk implications, triangulated across managerial, technical, and front-line respondents. Rather than assuming a one-time, absolute compliance state, this study adopts a feedback-loop perspective that treats legal-compliance assurance as a continuous, adaptive governance process (Badawi et al., 2026).

Theoretical Framework

To move beyond a descriptive audit report and establish robust explanatory power, this study anchors its analysis on two foundational theories: Socio-Technical Systems (STS) theory and Institutional Theory (Coercive Isomorphism).

- Socio-Technical Systems (STS) Theory: Originating from Trist and Bamford, STS posits that organizational performance depends on the optimization of both technical subsystems (hardware, software, networks) and social subsystems (people, policies, behaviors). In the context of a hybrid hotel PMS, technical security controls alone are insufficient if the social subsystem, such as front-line staff sharing passwords or handling unverified credentials, fails. This theory justifies our multi-layer assessment approach (managerial, technical, and operational).
- Institutional Theory (Coercive Isomorphism): Drawing on DiMaggio and Powell, coercive isomorphism explains how formal external pressures such as Indonesia's Personal Data Protection Law (UU PDP No. 27 of 2022) force organizations to adopt structural changes. However, organizations often exhibit *decoupling*, wherein formal corporate policies (*Govern* functions) are ceremonially adopted, while operational implementation and incident response capabilities (*Respond* and *Recover* functions) lag behind due to a lack of local technical capacity.

From these theories, two core propositions are formulated and examined through our empirical data:

- P_1 In organizations with centralized corporate IT, maturity scores for high-level governance (*Govern*) and detection (*Detect*) will be significantly higher than local execution capabilities (*Respond* and *Recover*), reflecting weak coupling across socio-technical subsystems.
- P_2 Regulatory pressures primarily induce documentary compliance rather than operational behavioral shifts, leading to structural decoupling between policy mandates and front-line practices.

Table 1. Summary of Related Literature and Positioning of This Study

Study	Sector	Framework / Standard	Informants / Layers	Legal Mapping	Instrument Validation	System Architecture
Bernardo et al. (2025)	General	NIST CSF	–	No	–	–
Pipyros & Liasidou (2025)	Hospitality	Risk Assessment	–	No	–	–
Karadayi-Usta (2024)	Hospitality	Stakeholder Model	–	No	–	–
Voigt & Von dem Bussche (2017)	Cross-sector	GDPR	–	Yes (legal only)	–	–
This Study	**Hospitality (PMS)**	**NIST CSF 2.0 × UU PDP**	**6 respondents, 3 layers**	**Yes, traceable per article**	**Yes (CVI + ICC)**	**Hybrid on-premise / cloud-integrated**

Table 1 presents a comparison of previous studies based on sector, framework or standard, informants and organizational layers, legal mapping, instrument validation, and system architecture. The comparison indicates that previous studies have addressed cybersecurity and data protection from different perspectives, but these dimensions have generally been examined separately.

Bernardo et al. (2025) applied the NIST Cybersecurity Framework (NIST CSF) to address cybersecurity assessment in a general context. Although the study provides a relevant cybersecurity framework, it does not specifically address the hospitality sector, legal mapping to data protection regulations, or the characteristics of hotel information system architecture. Pipyros & Liasidou (2025) focused on cybersecurity risk assessment within the hospitality sector, demonstrating the importance of risk-based approaches in hotel environments. However, the study does not explicitly integrate cybersecurity controls with applicable data protection legislation or provide a validated assessment instrument. Similarly, Karadayi-Usta (2024) emphasized stakeholder perspectives in the hospitality context, but did not combine stakeholder assessment with a formal cybersecurity framework, legal requirements, and technical system architecture.

Voigt & Von dem Bussche (2017), in contrast, provide a legal perspective through the General Data Protection Regulation (GDPR). Their work is important for understanding data protection obligations; however, the legal perspective is not directly integrated with a technical cybersecurity control framework, multi-layer organizational assessment, or empirical validation of an assessment instrument. Thus, while these studies provide valuable foundations, the comparison reveals that cybersecurity frameworks, hospitality risk assessment, stakeholder perspectives, and legal compliance have largely been addressed as separate components rather than as an integrated assessment approach.

The gap becomes particularly relevant in modern hotel Property Management Systems (PMS), where information processing occurs across organizational and technological boundaries. A hotel PMS may operate through on-premises infrastructure while simultaneously connecting to cloud-based Channel Managers and multiple Online Travel Agencies (OTAs). Consequently, cybersecurity governance cannot be evaluated solely from the perspective of internal technical controls or organizational policies. It also requires consideration of data protection obligations, stakeholder responsibilities, operational practices, and the shared responsibility between on-premises and cloud-based environments.

In response to this gap, the present study proposes an integrated assessment approach that combines NIST CSF 2.0 and Indonesia's Personal Data Protection Law (UU PDP) within the context of a hotel PMS. The study further incorporates six respondents across three organizational layers managerial, technical, and operational/front-line to support multi-perspective assessment and data triangulation. In addition, the assessment instrument is validated using the Content Validity Index (CVI) and Intraclass Correlation Coefficient (ICC), strengthening the methodological reliability of the proposed assessment.

Another distinguishing feature of this study is its explicit consideration of a hybrid on-premise/cloud-integrated system architecture. This allows the assessment to account for cybersecurity risks and responsibilities associated not only with the hotel's internal PMS infrastructure but also with external cloud-based integrations and OTA connectivity. Furthermore, the legal mapping is designed to be traceable to relevant provisions of the UU PDP, enabling cybersecurity controls to be evaluated alongside their corresponding legal obligations.

Therefore, compared with the previous studies presented in Table 1, this study contributes an integrated approach that connects cybersecurity controls, legal compliance, organizational perspectives, validated assessment instruments, and hybrid system architecture. The assessment results are subsequently translated into risk prioritization and a Hybrid Cybersecurity Governance Blueprint, providing a structured pathway from system diagnosis and control assessment to practical cybersecurity governance recommendations for hotel PMS environments.

Methodology

Research Design and Site

This study used a descriptive quantitative audit design combined with qualitative triangulation (mixed-methods) and was conducted at an international chain hotel property located in Surakarta, Central Java, Indonesia (hereafter referred to as "Hotel X" to preserve organizational confidentiality, consistent with the consent terms granted by site management). Data collection took place in June 2026. Hotel X operates an on-premise e-VHP PMS integrated, via a centrally managed Channel Manager, with cloud-based OTA distribution channels. Before data collection, the audit protocol, including the interview

instrument and the confidentiality terms governing the disclosure of site-level findings, was reviewed and approved by site management, and verbal informed consent was obtained from each respondent.

This study adopts a systematic, multi-phase evaluation methodology structured into four sequential phases to bridge technical cybersecurity metrics with regulatory compliance obligations within a hybrid hotel Property Management System (PMS) environment. The overall research flow, illustrated in Figure 1, is designed to ensure methodological rigor, moving from diagnostic scoping to framework formulation.

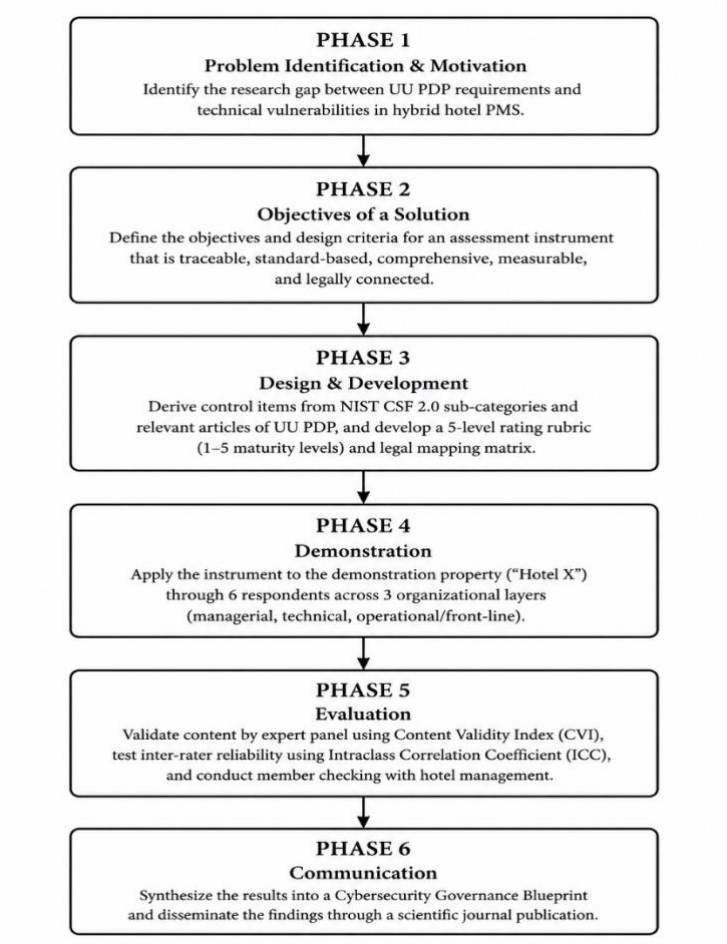


Figure 1. Six-Phase Research Methodological Flow for the Development and Evaluation of a Hybrid Hotel PMS Cybersecurity Governance Assessment

Phase 1: Problem Identification and Motivation

The first phase establishes the research problem and its underlying motivation. The study begins by identifying the gap between the requirements of Indonesia's Personal Data Protection Law (UU PDP) and the technical and organizational cybersecurity conditions of a hybrid hotel PMS environment. The problem is particularly relevant because contemporary hotel PMS environments increasingly operate across interconnected infrastructures, including on-premise servers and databases, cloud-based services, Channel Managers, and Online Travel Agencies (OTAs). Such interconnectedness creates multiple points at which personal and operational data may be processed, transmitted, stored, or accessed.

At this stage, the study examines the cybersecurity and data protection context of the hotel PMS and identifies potential discrepancies between regulatory expectations and actual technical or organizational practices. The analysis provides the basis for defining the research problem, establishing the research scope, and formulating the research gap. The output of this phase is therefore a clearly defined problem

statement and research motivation that justify the need for a context-specific cybersecurity assessment approach.

Phase 2: Objectives of a Solution

Following the identification of the problem, Phase 2 defines the objectives and design requirements of the proposed solution. The primary objective is to develop an assessment instrument capable of evaluating cybersecurity controls within a hybrid hotel PMS environment while maintaining clear relationships between technical controls, recognized cybersecurity standards, and applicable legal requirements.

Several design criteria are established in this phase. The instrument must be traceable, standard-based, comprehensive, measurable, and legally connected. Traceability requires that assessment items can be linked to their underlying cybersecurity control and corresponding legal provisions. The standard-based requirement is addressed through the use of the NIST Cybersecurity Framework (CSF) 2.0, while the legal dimension is incorporated through the mapping of relevant provisions of the UU PDP. The instrument must also be sufficiently comprehensive to address relevant cybersecurity practices, measurable through a consistent maturity-rating mechanism, and applicable to the organizational and technical characteristics of a hybrid hotel PMS. The output of this phase is a set of explicit instrument-design objectives and a conceptual framework that determines how NIST CSF 2.0 and UU PDP requirements will subsequently be integrated.

Phase 3: Design and Development

Phase 3 translates the objectives established in Phase 2 into the actual assessment instrument. In this stage, relevant control dimensions are derived from the selected NIST CSF 2.0 sub-categories, particularly those associated with the Protect and Detect functions. These control dimensions are then mapped to relevant provisions of the UU PDP to establish a legally traceable relationship between cybersecurity practices and personal data protection obligations.

The design process involves the development of assessment items, indicators, and a structured legal mapping matrix. Each assessment item is formulated to represent an observable cybersecurity practice or control that can subsequently be assessed within the hotel PMS environment. The mapping process is designed to preserve traceability, allowing each relevant control or assessment item to be associated with the corresponding legal provision.

A five-level maturity-rating rubric is also developed to support consistent assessment. The maturity scale ranges from Level 1 (Initial) to Level 5 (Optimized), representing progressive levels of cybersecurity capability and control maturity. The resulting draft instrument therefore consists of assessment items, control indicators, legal mappings, and a standardized rating mechanism. This phase produces the initial methodological artifact that will be applied and subsequently evaluated.

Phase 4: Demonstration

Phase 4 demonstrates the practical application of the developed instrument within an actual hotel PMS environment, represented in the study as “Hotel X.” The demonstration provides an empirical context in which the proposed assessment approach can be applied to a hybrid PMS architecture and its associated organizational processes. Data are collected from six respondents representing three organizational layers, namely the managerial layer, technical layer, and operational/front-line layer. The managerial layer provides evidence concerning governance, policies, and organizational oversight. The technical layer provides evidence concerning system configurations, technical controls, logs, backups, and infrastructure-related practices. The operational/front-line layer provides evidence concerning the actual day-to-day use of the PMS and operational security practices.

The use of multiple organizational layers enables the study to capture different dimensions of cybersecurity practice rather than relying on a single source of evidence. Data obtained through interviews, questionnaires, documentary evidence, technical walkthroughs, and relevant observations are subsequently compared and triangulated. This process enables the study to identify similarities,

inconsistencies, and gaps between documented policies, technical implementation, and operational practices.

The output of the demonstration phase is the empirical dataset and preliminary assessment results obtained from the application of the proposed instrument in the hotel PMS environment.

Phase 5: Evaluation

Phase 5 focuses on establishing the validity, reliability, and contextual credibility of the developed instrument and the resulting assessment. The evaluation consists of three complementary components: content validation, inter-rater reliability assessment, and member checking.

- First, the content of the instrument is evaluated by an expert panel using the Content Validity Index (CVI). This process determines whether the assessment items adequately represent the intended cybersecurity and data protection constructs and whether the items are relevant to the research objectives.
- Second, the consistency of assessments among raters is examined using the Intraclass Correlation Coefficient (ICC). The ICC provides an empirical basis for evaluating inter-rater reliability and determining whether the assessment produces sufficiently consistent ratings when applied by different evaluators.
- Third, member checking is conducted with hotel management to confirm the interpretation of findings and ensure that the assessment results appropriately reflect the actual organizational and operational context. Feedback obtained through this process is used to clarify or revise relevant elements where necessary. The combination of CVI, ICC, and member checking strengthens the methodological rigor of the study by addressing different dimensions of quality: content validity, measurement reliability, and contextual credibility. The output of this phase is a validated and reliable assessment instrument together with refined assessment findings.

Phase 6: Communication

The final phase translates the research findings into practical and scholarly outputs. The validated assessment results are analyzed to identify cybersecurity weaknesses, maturity gaps, and priority risks within the hybrid hotel PMS environment. These findings are then synthesized into a Cybersecurity Governance Blueprint that provides a structured basis for improving cybersecurity governance.

The blueprint is intended to connect assessment findings with practical governance responses, including cybersecurity priorities, control recommendations, governance mechanisms, and potential implementation directions. In this way, the research does not terminate at the identification of cybersecurity deficiencies but translates empirical findings into an actionable governance-oriented output.

At the scholarly level, the findings, methodological approach, instrument development, validation results, and governance implications are consolidated into the present scientific manuscript. Thus, communication serves two purposes: disseminating the research contribution to the academic community and providing a practical governance reference for organizations managing hybrid hotel PMS environments.

Instrument and Respondents

The audit instrument consisted of structured interview checklists derived from NIST CSF 2.0 functions (Govern, Identify, Protect, Detect, Respond, Recover), supplemented by two targeted sub-instruments: a Human-Factor checklist (front-line security awareness) and an Access-Control checklist (least-privilege and credential-handling practices). Each item was scored using a five-point maturity rubric, summarized in Table 2.

Table 2. Five-point Maturity Rubric Applied in the Field Instrument

Score	Maturity Level	Description
1	Not Implemented	Control is absent.
2	Inconsistent	Control exists but is applied irregularly or is undocumented.
3	Implemented	Control is applied as a standard practice.
4	Implemented & Monitored	Control is applied and actively monitored.
5	Optimized & Reviewed	Control is optimized and periodically evaluated.

Data were collected through structured interviews with six respondents purposively selected to represent three organizational layers directly responsible for, or affected by, cybersecurity controls: (1) managerial the General Manager; (2) technical two IT Officers; and (3) operational/front-line a Front Office Leader, a General Cashier/Accounts Payable officer, and an Income Audit/Accounts Receivable officer (Table 3). This multi-layer sampling strategy was adopted to triangulate self-reported managerial policy claims against technical implementation evidence and front-line behavioral reality, addressing a key methodological limitation: reliance on single-informant data identified in prior single-respondent audit studies.

Table 3. Respondent Profile

Respondent	Role	Instrument Scope
R1	IT Officer	Govern, Identify, Protect, Detect, Respond, Recover
R2	IT Officer	Govern, Identify, Protect, Detect, Respond, Recover
R3	General Manager	Govern, Identify, Protect, Detect, Respond, Recover
R4	Front Office Leader	Human Factor (security awareness)
R5	General Cashier / Accounts Payable	Access Control
R6	Income Audit / Accounts Receivable	Access Control

For items assessed by the three managerial/technical respondents (R1–R3), the domain-level maturity score for each NIST CSF function was calculated as the unweighted arithmetic mean of item scores across respondents and across items within the domain, following Equation 1. Unweighted averaging was used, rather than the weighted-average approach proposed in earlier drafts of this instrument, because no validated weighting scheme for item criticality had yet been established for this audit context; this is acknowledged as a limitation in Section 5,

$$SM = \frac{(\Sigma Si)}{n} \qquad \text{Equation (1)}$$

where SM is the domain maturity score, Si is the score of item i across all respondents who completed that item, and n is the total number of completed item-respondent observations within the domain.

Likelihood and Impact ratings used in the risk register (Section 3.4) were derived qualitatively from interview evidence and triangulated across respondents, using a 1–5 ordinal scale where 1 indicates rare/negligible, and 5 indicates near-certain/severe, consistent with common qualitative risk-matrix conventions (Abbas & Alazawi, 2026). Likelihood and Impact definitions are summarized in Table 4.

Table 4. Likelihood and Impact Rating Scale

Scale	Likelihood Descriptor	Impact Descriptor
1	Rare, no recorded occurrence	Negligible, no operational disruption
2	Unlikely, isolated occurrence	Minor limited, quickly recoverable disruption
3	Possible, occurs under specific conditions	Moderate partial operational disruption
4	Likely, recurring or structurally probable	Major significant disruption or data exposure
5	Almost certain, actively occurring or unmitigated	Severe system-wide failure or irreversible data loss

Limitations of the Legal Mapping Approach

The mapping of technical findings to specific UU PDP articles in Section 3.5 is interpretive and indicative; it identifies the statutory provisions a finding is most likely to implicate if exploited, rather than constituting a definitive legal determination of fault, which would require formal evidentiary and procedural processes beyond the scope of this audit.

The systematic alignment between NIST CSF 2.0 sub-categories and Indonesia's UU PDP (Law No. 27 of 2022) was established to bridge technical controls with legal mandates. The complete, granular cross-walk mapping, along with the required audit evidence and verbatim statutory references, is provided in Appendix A.

Result

Cybersecurity Maturity Profile

Aggregated maturity scores across the six NIST CSF 2.0 functions are presented in Figure 2 and summarized in Table 5.

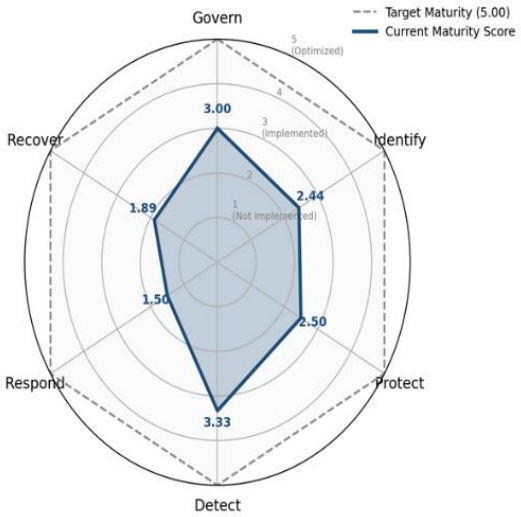


Figure 2. Cybersecurity Maturity Profile of the Six NIST CSF 2.0 Functions (Spider Chart)

As depicted in Figure 2, the solid blue line shows the hotel's current maturity score for each function on a 1–5 scale (1 = Not Implemented, 5 = Optimized and Continuously Reviewed; see Table 1 for full rubric); the dashed grey line marks the optimal target score (5.00) applied to every function. The larger the gap between the blue line and the dashed line on a given axis, the more urgent that function is for remediation. The chart highlights that Detect (3.33) and Govern (3.00) achieved scores closest to the target, while Respond (1.50) and Recover (1.89) exhibited the widest capability gaps.

Table 5 presents the same scores in tabular form, together with a qualitative maturity-status label for each function, for readers who prefer a numeric summary. Current Score is the average of all item-level interview scores within each function, rated on the 1–5 maturity rubric in Table 2 (1 = Not Implemented; 5 = Optimized and Continuously Reviewed). Maturity Status summarizes, in plain language, what the score implies about the current state of controls in that function. Domains are listed from highest to lowest maturity.

Table 5. Aggregated Maturity Scores by NIST CSF 2.0 Function

Domain	Current Score (1–5)	Maturity Status
Detect	3.33	Implemented, partially monitored
Govern	3.00	Implemented
Protect	2.50	Inconsistent to implemented
Identify	2.44	Inconsistent
Recover	1.89	Inconsistent
Respond	1.50	Largely not implemented

Detect achieved the highest relative maturity (3.33), driven primarily by consistent system-level activity logging (DE1 = 4.00), although periodic log review remains reactive rather than scheduled (DE2 = 2.67). Govern (3.00) reflects the existence of centrally issued IT security SOPs and a designated IT Officer role, though localized review at the property level remains limited. In contrast, Respond (1.50) and Recover (1.89) show the most severe capability deficits: no respondent reported prior incident-response simulation (RS2 averaged 1.00), and restore testing for backup data had never been performed (RC2 = 1.00), with site-level endpoint backup execution rated as largely unrealized (RC3 = 1.33) despite an annual capital-expenditure plan.

Gap Analysis

To prioritize domain remediation, the gap score for each NIST CSF 2.0 domain was calculated using Equation 2,

$$Gap = Target\ Score - Current\ Score \quad \text{Equation (2)}$$

where $Target\ Score$ represents the optimal score of 5.00 consistent with the maturity rubric in Table 2, and $Current\ Score$ denotes the domain average derived from Equation 1. The detailed gap values and corresponding priority rankings from P1 (Critical) to P6 (Low) are detailed in Table 6.

Respond ranks as the highest-priority remediation domain, reflecting the absence of a documented incident-response SOP at the property level and the complete absence of incident-simulation exercises. Notably, the existing organizational incident-reporting channel (escalation to the General Manager, Regional Director of Finance, Vice President, and ownership) was designed for guest- and facility-safety incidents and has not been adapted for IT/data-breach scenarios, leaving a structural gap relative to the UU PDP's incident-notification requirements.

Table 6. Gap Analysis and Remediation Priority

Domain	Current	Target	Gap	Priority
Respond	1.50	5.00	3.50	P1 (Critical)
Recover	1.89	5.00	3.11	P2 (High)
Identify	2.44	5.00	2.56	P3 (High)
Protect	2.50	5.00	2.50	P4 (Moderate)
Govern	3.00	5.00	2.00	P5 (Moderate)
Detect	3.33	5.00	1.67	P6 (Low)

Human Factor and Access Control Findings

Front-line security awareness, assessed through a ten-item Human Factor checklist administered to the Front Office Leader, yielded an average score of 2.10, the lowest aggregate score among all instrument categories. Notable findings include a lack of phishing-concept awareness (FO2 = 1.00), routine password-sharing across Front Office staff, daily workers, and on-the-job trainees (FO6 = 1.00; FO7 = 1.00), and recurring operational disruption from power loss affecting Front Office terminals when automatic backup generator activation failed (FO9 = 3.00).

Access-control findings were more heterogeneous: the Income Audit/Accounts Receivable respondent reported consistently enforced least-privilege access (average 3.25), whereas the General Cashier/Accounts Payable respondent reported broader-than-necessary access spanning Receiving, Inventory, and Front Office cashiering functions, alongside the absence of incident-reporting awareness (average 2.25). Both respondents confirmed the absence of Uninterruptible Power Supply (UPS) protection for their workstations, with backup limited to data already captured within the centralized PMS interface; locally held spreadsheet-based financial data was not subject to any backup routine.

Risk Register

Likelihood and Impact are both rated on the 1–5 ordinal scale defined in Table 4 (1 = rare/negligible, 5 = almost certain/severe). The overall Risk Score is calculated as $Score = Likelihood \times Impact$ ($maximum\ possible = 25$) and grouped into four risk categories: Score 1–6 = Low, 7–11 = Moderate, 12–15 = High, and 16–25 = Critical. The quantitative risk evaluation for the property is synthesized in Table 7, ordered from highest to lowest risk score.

Table 7. Risk Register

ID	Risk Description	Likelihood	Impact	Score	Category
R01	Inconsistent UPS coverage is causing PMS/power-dependent terminal failure during outages	4	4	16	High
R02	Endpoint and ancillary (spreadsheet) data backup failure; no restore testing performed	5	4	20	Critical
R03	Reliance on a centrally managed encryption/integration channel (Channel Manager–OTA) with no local TLS verification capability	3	4	12	High
R04	Unmanaged credential sharing and low phishing awareness among front-line staff	4	3	12	High

R02 represents the most critical exposure (Score = 20), driven by the combination of unrealized scheduled backups (RC3) and the absence of any restore-testing precedent (RC2), compounded by uncontrolled local storage of financial spreadsheet data outside the PMS backup scope (AC7). This condition, if exploited by ransomware or hardware failure, presents a high likelihood of unrecoverable

data loss. R01 (power management) and R04 (human factor/credential hygiene) both register as High, reflecting tangible, previously experienced operational disruptions (FO9) and routine, multi-party credential sharing that undermines individual accountability for system actions.

Structural Mapping to UU PDP Provisions

Table 8 maps the audit findings to the UU PDP provisions they are most likely to implicate. As noted in Section 2.3, this mapping is indicative rather than a formal legal determination. Each row links one technical finding (NIST Domain / Audit Finding) to the specific UU PDP article it could potentially trigger if left unaddressed, and explains in plain language what that would mean for the organization's legal exposure (Indicative Implication). This mapping identifies risk, not a finding of legal fault.

Table 8. Indicative Mapping of Audit Findings to UU PDP Provisions

NIST Domain	Audit Finding	Related UU PDP Article	Indicative Implication
Protect	Inconsistent UPS coverage / power-dependent system failure	Article 35	Potential failure to maintain a reliable data-processing system.
Recover	Unrealized endpoint backup; no restore testing	Article 39	Potential failure to ensure continuity of data protection.
Respond	No property-level incident-response SOP or simulation	Article 46	Potential failure to meet incident-handling and notification obligations.

If these gaps culminate in an actual data-breach event, Hotel X could face administrative sanctions under Article 57 of the UU PDP, including fines of up to 2% of annual revenue, in addition to potential civil claims and reputational damage. These exposures are presented as risk projections to inform proactive remediation, not as predictions of any pending regulatory action.

Discussion

The empirical findings detailed in Section 4 reveal deep operational vulnerabilities that extend far beyond technical glitches, providing strong support for our theoretical framework.

- **Analysis of P_1 (Subsystem Coupling):** Grounded in Socio-Technical Systems (STS) theory, the severe gap between high-level governance and localized execution illustrates a weak coupling between the technical subsystem (e.g., hardware terminals, unbacked spreadsheets, power dependencies) and the social subsystem (front-line staff and trainees). While corporate oversight ensures high maturity in structural policies (*Govern*), local operational capabilities such as disaster recovery and endpoint backup collapse because front-line personnel lack proper training and independent technical authority (evidenced by R01 and R02). When power failures or backup anomalies occur, the socio-technical loop breaks because the human actors cannot independently remediate the technical failure.
- **Analysis of P_2 (Institutional Decoupling):** Grounded in Institutional Theory (Coercive Isomorphism), the coexistence of formal compliance structures and widespread operational non-compliance (such as routine multi-party credential sharing, FO6/FO7 = 1.00) confirms a state of structural decoupling. External pressures stemming from Indonesia's Personal Data Protection Law (UU PDP No. 27/2022) successfully compel the organization to adopt ceremonial policies. However, daily operational routines at the front-line layer continue to prioritize workflow convenience over data security, meaning regulatory coercion drives documentary compliance rather than true behavioral changes.

When evaluated against existing literature, these findings refine our understanding of cybersecurity postures in the hospitality sector. While prior studies by Karadayi-Usta (2024) and Pipyros and Liasidou (2025) largely focused on high-level managerial risk perceptions or general risk assessments, our multi-

layer approach exposes the exact friction points in hybrid cloud-local PMS architectures, extending the insights of Bernardo et al. (2025). Furthermore, unlike standard GDPR cross-sector analyses (e.g., Voigt & Von dem Bussche, 2017), our findings demonstrate how physical infrastructure deficiencies such as unbacked local financial spreadsheets and inconsistent UPS coverage translate directly into statutory violations of data continuity and protection mandates under Articles 35 and 39 of the UU PDP Law.

To bridge the gap between identified risks (such as R02's critical score of 20 and R01/R04's high scores) and regulatory compliance under Articles 35, 39, and 46, organizations managing hybrid hotel properties must implement tiered, structural interventions rather than temporary fixes:

- **Respond / Recover Interventions (P1–P2):** Establishing property-level incident response capabilities aligned with UU PDP 3x24-hour notification timelines.
- **Power Resilience Upgrades (R01):** Standardizing UPS coverage with Automated Voltage Regulation across all front-line terminals, moving beyond server-room-only protections.
- **Endpoint & Backup Security (R02):** Deploying EDR agents, automated AES-256 encrypted NAS backups, and periodic restore testing that explicitly encompasses local spreadsheet-based financial records.
- **Human Factor & Access Hygiene (R04):** Eradicating unmanaged credential sharing through role-based access control (RBAC) and conducting phishing awareness simulations for front-line staff and trainees.

Conclusion

This study successfully developed and demonstrated a hybrid cybersecurity compliance audit framework tailored for on-premises hotel Property Management Systems (PMS) integrated with cloud-based channels, bridging NIST CSF 2.0 technical controls with Indonesia's Personal Data Protection (UU PDP) Law No. 27 of 2022.

Utilizing DSR methodology and multi-layer assessment at Hotel X, findings confirm that while high-level governance achieved moderate maturity (3.00), local technical execution and incident response lagged substantially (1.50). Furthermore, the study verified two core theoretical propositions: subsystem coupling (P_1) failures in power/backups and institutional decoupling (P_2) through ceremonial compliance versus front-line non-compliance.

Through a rigorous cross-walk mapping matrix and qualitative risk prioritization, the research translated physical infrastructure gaps and human factor vulnerabilities into explicit statutory legal exposures under Articles 35, 39, and 46 of the UU PDP Law. Future work should focus on scaling and validating this audit instrument across broader multi-property portfolios.

References

Abbas, A. S., & Alazawi, S. A. H. (2026). Risk evaluation for bank information security management in network environment considering ISO/IEC 27001 and ISO/IEC 27005. *AIP Conference Proceedings*, *3408*(1), 020012. https://doi.org/10.1063/5.0329787

Badawi, S. A., Takruri, M., Salameh, K., Guessoum, D., ElBadawi, I., & Al-Qaisi, A. (2026). Data governance and security assurance in ISO27001 programs: A context study on quality and compliance in Middle Eastern organizations. *Applied Mathematics & Information Sciences*, *20*(1), 253–265. https://doi.org/10.18576/amis/200117

Bernardo, L., Malta, S., & Magalhães, J. (2025). An evaluation framework for cybersecurity maturity aligned with the NIST CSF. Electronics, 14(7), 1364. https://doi.org/10.3390/electronics14071364

Brezavšček, A., & Baggia, A. (2025). Recent trends in information and cyber security maturity assessment: A systematic literature review. *Systems*, *13*(1), 52. https://doi.org/10.3390/systems13010052

Büyüközkan, G., & Güler, M. (2025). Cybersecurity maturity model: Systematic literature review and a proposed model. *Technological Forecasting and Social Change*, *213*, 123996. https://doi.org/10.1016/j.techfore.2025.123996

Campoverde-Molina, M., & Luján-Mora, S. (2025). Cybersecurity in smart agriculture: A systematic literature review. *Computers & Security*, *150*, 104284. https://doi.org/10.1016/j.cose.2024.104284

Casaril, F., & Galletta, L. (2025). Developing security metrics for space systems: A study considering the NIST Cybersecurity Framework 2.0 and the NIS2. *International Journal of Critical Infrastructure Protection*, *51*, 100805. https://doi.org/10.1016/j.ijcip.2025.100805

Culot, G., Nassimbeni, G., Podrecca, M., & Sartor, M. (2021). The ISO/IEC 27001 information security management standard: Literature review and theory-based research agenda. *The TQM Journal*, *33*(7), 76–105. https://doi.org/10.1108/TQM-09-2020-0202

Demir, M., & Demir, Ş. Ş. (2025). The relationship between technology investments, innovation strategies, and competitive performance in the hospitality industry: A mixed methods approach. International *Journal of Hospitality Management*, *128*, 104151. https://doi.org/10.1016/j.ijhm.2025.104151

Karadayi-Usta, S. (2024). Cybersecurity risks analysis in the hospitality industry: A stakeholder perspective on sustainable service systems. *Systems*, *12*(10), 397. https://doi.org/10.3390/systems12100397

Kim, H. B., Lee, D. S., & Ham, S. (2013). Impact of hotel information security on system reliability. *International Journal of Hospitality Management*, *35*, 369–379. https://doi.org/10.1016/j.ijhm.2012.06.002

National Institute of Standards and Technology. (2018). Framework for improving critical infrastructure cybersecurity (Version 1.1) (NIST Cybersecurity White Paper CSWP 4). U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.04162018

National Institute of Standards and Technology. (2024). The NIST cybersecurity framework (CSF) 2.0 (NIST Special Publication SP 1299). U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29

Pipyros, K., & Liasidou, S. (2025). A new cybersecurity risk assessment framework for the hospitality industry: Techniques and methods for enhanced data protection and threat mitigation. *Worldwide Hospitality and Tourism Themes*, *17*(1), 48–61. https://doi.org/10.1108/WHATT-12-2024-0296

Republik Indonesia. (2022). Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi. Lembaran Negara Republik Indonesia Tahun 2022 Nomor 196. *Sekretariat Negara*.

Sari, I. P., Ramadhan, A., & Hermawan, B. (2023). Analisis Pasal 56 dalam Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi dari perspektif kepastian hukum. *Jurnal Hukum Privasi*, *2*(1), 45–58.

Tarumingkeng, R. C. (2022). Analisis risiko siber dan tata kelola keamanan informasi pada sektor industri strategis. *Penerbit Buku Ilmiah*.

Cobanoglu, C., Dogan, S., Berezina, K., & Collins, G. (2021). Hospitality and tourism information technology. *USF M3 Publishing*, LLC. https://doi.org/10.5038/9781732127593

Voigt, P., & von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR): A practical guide. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-57959-7

Wibowo, A., Sofyan, M., & Budiman, A. (2023). Urgensi penguatan regulasi pelindungan data dan keamanan siber di Indonesia terhadap ancaman hacking dalam sistem pemerintahan berbasis elektronik. *Jurnal Keamanan Siber Nasional*, *5*(2), 112–128.

How to cite:

Sudana, I. M., Handaga, B., & Wantoro, J. (2026). Hybrid Cybersecurity Compliance Audit Framework for an On-Premise Hotel Property Management System: Mapping NIST CSF 2.0 Controls to Indonesia's Personal Data Protection Law. *Jurnal Sistem Informasi (Journal of Information System)*, *22*(2), 64–80.

Appendices

Appendix A
Detailed Cross-Walk Mapping Matrix (NIST CSF 2.0 Sub-Categories to UU PDP Law)

NIST CSF 2.0 Function	Sub-Category ID	NIST CSF 2.0 Sub-Category Description	UU PDP Law (No. 27/2022)	Statutory Mandate / Requirement Summary	Required Audit Evidence
PROTECT (PR)	PR.AC-01	Identities and credentials are authenticated and managed.	Article 35 (Paragraph 1 & 2)	Obligation to protect personal data from unauthorized access through secure authentication mechanisms.	Active user accounts list in PMS vs. active front-line staff roster, password complexity policy, multi-factor authentication (MFA) configuration logs.
PROTECT (PR)	PR.DS-01	Data-at-rest is protected against unauthorized access and disclosure.	Article 35 (Paragraph 1) & Article 39	Duty to implement technical safeguards (encryption/access restriction) for personal data storage.	Database encryption status reports, physical server room access logs, local NAS backup encryption configurations.
PROTECT (PR)	PR.DS-02	Data-in-transit is protected against unauthorized access and disclosure.	Article 35 (Paragraph 1) & Article 39	Ensuring data security during transmission across internal networks and external cloud integrations (OTAs).	Channel Manager API encryption protocols (TLS/SSL certificates), firewall rule sets, network topology diagrams.
PROTECT (PR)	PR.DS-11	Backups of data are created, protected, and tested.	Article 39 (Paragraph 1 & 2)	Obligation to implement security measures to prevent loss or unauthorized alteration of personal data, including backup management.	Backup execution logs, off-site backup schedules, hardware redundancy inventory (UPS/AVR status), and restoration test reports.
DETECT (DE)	DE.AE-02	Anomalous activity is detected across the information system.	Article 46 (Paragraph 1)	Obligation to monitor systems for security incidents and anomalies that may compromise personal data.	System audit trails, PMS access logs, error logs, and endpoint detection/monitoring configurations.
DETECT (DE)	DE.CM-01	Networks and physical environments	Article 35 & Article 46	Continuous monitoring of processing	Physical inspection of server rooms, CCTV logs covering

NIST CSF 2.0 Function	Sub-Category ID	NIST CSF 2.0 Sub-Category Description	UU PDP Law (No. 27/2022)	Statutory Mandate / Requirement Summary	Required Audit Evidence
		are monitored to find potentially unauthorized activities.		systems to detect vulnerabilities, unauthorized nodes, or breaches.	terminal access, and network monitoring tool configurations.
RESPOND / GOVERN	RC/GV	Incident response and organizational governance procedures are established.	Article 46 (Paragraph 2) & Article 57	Mandatory notification protocols for data breaches (within 3x24 hours) and administrative sanctions for non-compliance.	Incident response SOP documents, data breach notification templates, and records of past security incident handling.

ISSN: 2088-7043

JURNAL SISTEM INFORMASI

JOURNAL OF INFORMATION SYSTEM

http://jsi.cs.ui.ac.id