

SRIWIJAYA INTERNATIONAL JOURNAL OF DYNAMIC ECONOMICS AND BUSINESS

<http://ejournal.unsri.ac.id/index.php/sijdeb>

Supporting IFRS 17 Readiness: A Usability Case Study of Tugu's Core Insurance System in Indonesia

Asyhari Achmad Ichsan¹, and Harry Budi Santoso²

^{1,2}Faculty of Computer Science, Universitas Indonesia

Corresponding authors: asyhari.achmad@ui.ac.id

Abstract: To comply with IFRS 17, Tugu Insurance implemented the Tugu Insurance Core System (TICS) in 2024 to improve policy-data collection and accuracy. This study evaluates and improves the usability of TICS, focusing on workflows that are critical for IFRS 17 reporting. A mixed-methods user-centered design approach was applied, combining task-based usability testing with standardized questionnaires (System Usability Scale and Post-Study System Usability Questionnaire). Sixteen employees completed the questionnaires and ten participated in usability tests on the as-is system and the redesigned prototype. The SUS score increased from 50.8 (“poor”) to 77.8 (“good”), while PSSUQ scores improved across all dimensions. The case study demonstrates how systematic usability evaluation and redesign of an insurance core system can directly enhance IFRS 17 readiness by reducing data-entry problems, shortening task time, and increasing user satisfaction.

Keywords: IFRS 17 Readiness, Insurance Information Systems, Usability Evaluation, Usability Heuristics, User-Centered Design

Introduction

The global insurance sector is undergoing a major regulatory change with the implementation of IFRS 17. The standard requires more granular, data-driven reporting on insurance contracts and forces insurers to change long-established accounting practices (Dahiyat & Owais, 2021). In Indonesia, Tugu Insurance has responded by developing the Tugu Insurance Core System (TICS) to support IFRS 17-compliant policy-data collection and reporting. Internal user feedback in 2024 shows that TICS still has usability problems, slow data entry, complex navigation, and frequent errors, that can delay IFRS 17 reporting and increase the risk of misstatements (Eckert & Osterrieder, 2020; Meyeroff & Tremoulet, 2021).

Tugu Insurance was selected as the case-study site because it is a state-owned enterprise subsidiary with a large and diverse portfolio and has been an early mover in preparing for IFRS 17. Since 2023 it has invested in TICS specifically to support IFRS 17-compliant data collection and reporting. The complexity of its products and transaction volumes makes usability particularly critical, and the organization provided access to domain experts and end-users, enabling in-depth usability testing that would not have been feasible at many other insurers.

Usability, as defined in ISO-based frameworks, covers effectiveness, efficiency and satisfaction in a specified context of use (Almasi et al., 2023). Poor usability leads to data-entry errors, longer processing times and user frustration, which are all problematic under tight regulatory deadlines. Existing IFRS 17 research has focused mainly on accounting, actuarial modelling and disclosure issues, while a smaller body of work analyses data and systems architecture (EY, 2017, 2019; BCG, 2023; Scott, 2023; Parwitasari, 2022). Very few studies evaluate the usability of front-end core systems that collect IFRS 17-relevant data using structured methods such as task-based testing, SUS and PSSUQ (Hou & Mayernik, 2021; Mator et al., 2020). This gap is the starting point for the present study.

This study evaluates and improves the usability of TICS as an IFRS 17-oriented insurance core system. It uses a mixed-methods, user-centred design approach: heuristic analysis based on Nielsen, Marenkov and Shneiderman; task-based usability testing; and standardized questionnaires (System Usability Scale and Post-Study System Usability Questionnaire) administered to 16 internal users, complemented by think-aloud observations and open-ended questions (Clark et al., 2021; Subiyakto et al., 2021).

The study contributes in three ways. First, it provides an empirical usability evaluation of TICS as a core system that supports IFRS 17-relevant workflows such as policy creation, premium calculation and contract documentation. Second, it combines heuristic evaluation, task-based testing and SUS/PSSUQ scores to identify and prioritize usability problems and to quantify the impact of a redesigned prototype. Third, it derives concrete design recommendations that link usability improvements to IFRS 17 readiness and that may be useful for other Indonesian insurers facing similar implementation challenges.

Literature Review

IFRS 17 and Insurance Information Systems

IFRS 17 Insurance Contracts is the new global accounting standard for insurance contracts, issued by the IASB in 2017 and effective for annual reporting periods beginning on or after 1 January 2023. It replaces IFRS 4 and introduces a current-value measurement model, explicit risk adjustment, and the contractual service margin, with the aim of improving the transparency and comparability of insurers' financial statements (IASB, 2017). These changes are expected to have a significant positive effect on the quality of financial reporting, particularly in terms of comparability, faithful representation, and relevance of insurance company financial statements (Dahiyat & Owais, 2021).

From an information-systems perspective, IFRS 17 is highly data-intensive. The standard requires insurers to capture and process detailed cash-flow information, identify investment components, and group contracts into portfolios and cohorts based on similar risk characteristics, profitability, and issuance year. This leads to a substantial increase in the volume and granularity of data that must be stored, integrated, and reconciled across policy administration, actuarial, and general-ledger systems (EY, 2017; Chartis Research, 2021).

Industry analyses highlight that IFRS 17 implementation affects virtually every layer of insurers' data and IT architecture, demanding new data models, integration pipelines, and reporting platforms that can support more granular measurement, scenario analysis, and disclosures (BCG, 2023; Chartis Research, 2021). Empirical and practice-oriented studies on IFRS 17 implementation consistently emphasize technology and process challenges: upgrading legacy systems, redesigning data architectures, aligning actuarial models with accounting engines, and managing cross-functional collaboration between finance, actuarial, and IT teams (EY, 2019; Scott, 2023; Parwitasari, 2022).

However, this body of work typically focuses on data management, modelling, and control processes rather than on the usability of the front-end systems used by operational staff to enter and manage IFRS 17-relevant data. To the best of the authors' knowledge, there is still a scarcity of empirical studies that apply systematic usability evaluation, using task-based testing and standardized instruments such as SUS and PSSUQ, to core insurance systems specifically developed to support IFRS 17. This study addresses that gap by examining the usability of the Tugu Insurance Core System (TICS) as a case of an IFRS 17-oriented information system.

Usability Evaluation

Usability evaluation is a systematic process by which to assess how easily and effectively users can interact with a system or product to achieve specific goals (Sharfina & Santoso, 2016). The focus is not only on users' success in completing tasks but also on their efficiency, effectiveness, and satisfaction throughout the process (Kurniawan, 2003). To comprehensively evaluate the usability of the TICS application, this study employs a mixed-methods approach, combining quantitative data with qualitative insights.

Various means of evaluating usability exist, each with its strengths and limitations. These methods may be categorized into expert review-based methods (e.g., heuristic evaluation, cognitive walkthrough) and user testing-based methods (e.g., usability testing, think-aloud). Expert review-based methods are relatively fast and cost-effective but offer more limited insights due to their reliance on experts' perspectives (Nielsen, 2024). User testing-based methods are more comprehensive because they involve direct interaction with users, but they require more time and resources.

Usability Testing

Usability testing is an evaluation method that measures how effective a product or application is for its users. Usability testing also serves to provide an understanding of the extent to which users can use an application or product to achieve specific goals (Sharfina &

Santoso, 2016). This method entails the direct observation of users as they interact with the product with specific tasks in mind with the primary aim being to identify and analyze problems that may arise during the user experience.

By conducting usability testing, development teams can gain valuable insights into how users interact with the interface and identify areas in need of improvement to enhance the user experience (Nielsen, 2024). This method was chosen because it offers a deep and comprehensive understanding of how users interact with the system in real-world use contexts. Usability testing paired with the think-aloud approach facilitates the identification of unexpected usability issues and provides detailed insights into the ways in which users complete tasks. This method was chosen for its precision in uncovering complex and hidden usability problems that are difficult to identify through other methods.

System Usability Scale (SUS)

The System Usability Scale (SUS), originally developed by John Brooke in 1986, is a well-established instrument for evaluating system usability. Some key advantages of the SUS are its ease of administration and its proven reliability even with small sample sizes, which enable it to effectively distinguish between usable and unusable systems (Sharfina & Santoso, 2016). The SUS consists of ten questions with five response options ranging from "strongly disagree" to "strongly agree," which yields a final score between 0 and 100.

The average SUS score across various studies is 68. If a study's average SUS score falls below 68, the application is in need of improvement. A usability score of 68 represents a universally average value, equivalent to a letter grade of C. Although scores above 68 may be considered "above average," the general industry practice is to target an SUS score of 80, which is equivalent to a letter grade of A-.

Post-Study System Usability Questionnaire (PSSUQ)

The Post-Study System Usability Questionnaire (PSSUQ) is a standardized tool for quantitatively assessing users' satisfaction with software systems or applications. This measurement tool was developed by IBM's internal System Usability Metrics (SUMS) project with the goal of evaluating systems' ease of use and effectiveness in supporting users' needs (Kurniawan, 2003). The PSSUQ provides data with which developers and researchers can understand the overall user experience, serving as a strong basis for the development of improvements.

The first version of the PSSUQ was designed by Lewis in 1990 and contained 18 items aimed at assessing user satisfaction with a system. The second version was expanded to 19 items, now encompassing the system's support for productivity improvement (Lewis, 2002). The PSSUQ was chosen on account of it being more reliable than the SUS. The PSSUQ provides numerical scores that allow for objective comparison and statistical analysis, suiting the needs of quantitative research (Kurniawan, 2003). Furthermore, it was selected because it includes questions that specifically evaluate the presentation of information within applications; these questions align with this research's focus on the TICS application.

Three Core Principles of Usability and User Interface Design

This study is theoretically grounded in the ISO 9241-11 definition of usability, effectiveness, efficiency, and satisfaction in a specified context of use, as well as three established usability frameworks (Nielsen's heuristics, Marenkov's guidelines, and Shneiderman's golden rules). These frameworks define what counts as a usable interface, guide the identification and classification of problems in TICS, and provide a theoretical basis for interpreting changes in SUS and PSSUQ scores.

Marenkov's Nine Usability Guidelines offer a practical framework for user-centric interface design, focusing on principles that directly enhance user experience and efficiency (Lehnert, Doublet, & Sim, 2024). These guidelines prioritize design consistency, clear system feedback, minimal cognitive load, and robust error-prevention and recovery mechanisms. Adherence to Marenkov's principles fosters the development of systems that are intuitive, easy to navigate, and resilient, ultimately improving user satisfaction and success rates.

Similarly, Nielsen's Ten Usability Heuristics constitute a widely accepted method of systematic usability evaluation. This framework addresses several critical aspects, including system status visibility, consistency, user control, error prevention, and minimalist design, offering a comprehensive checklist for identifying common usability flaws (Marenkov, Robal, & Kalja, 2018). Applying these heuristics via expert review enables the early detection of potential issues in the design process, giving way to the development of more efficient and user-friendly products prior to extensive user testing.

Complementing these, Shneiderman's Eight Golden Rules establish a foundational guide for designing interactive systems that prioritize user productivity and confidence. These rules emphasize design consistency, informative feedback, reversible actions, and putting users in control, all while minimizing users' short-term memory load (Marcus, 2009). By following these principles, designers can create predictable and efficient interfaces that support human cognitive processes, leading to a greater sense of user mastery and broader system adoption.

Previous Usability Studies on Enterprise and Financial Systems

Prior research on complex information systems has shown that explicit usability mechanisms and heuristics can have a measurable impact on user performance. Ferreira et al. (2019), for example, conducted a controlled experiment to examine the effect of three usability mechanisms, Abort Operation, Progress Feedback, and Preferences, on efficiency, effectiveness, and satisfaction. Their findings indicate that certain mechanisms, particularly Abort Operation and Preferences, significantly improve task performance and perceived usability, and that these effects are robust across multiple replications and large samples.

This line of work empirically supports the idea that targeted interface changes grounded in usability principles can yield quantifiable gains in system use. Other studies have applied usability-testing frameworks to evaluate web-based information systems in organizational settings. Subiyakto et al. (2021) investigated the quality of an institutional repository website in a public university, combining query techniques, formal experiments, and think-aloud protocols with usability scales to assess effectiveness, efficiency, and user satisfaction.

Their study revealed that designs developed primarily from a developer perspective tended to underperform on usability metrics, and it produced concrete visual and interaction design recommendations for improving the site. Similarly, Almasi et al. (2023) systematically reviewed usability-evaluation tools for dashboards and found that most empirical work relies on generic questionnaires such as SUS and PSSUQ, while highlighting the need for more context-specific criteria in evaluating decision-support interfaces. These studies collectively demonstrate that heuristic-based evaluations, task-based usability testing, and standardized questionnaires are effective for identifying problems and guiding interface improvements in complex organizational systems. However, they largely focus on generic web applications, dashboards, or institutional repositories rather than on insurance core systems operating under a strict regulatory regime such as IFRS 17.

None of them explicitly link usability evaluation to regulatory data-quality requirements or to IFRS 17-specific workflows such as policy creation, premium calculation, and contract documentation. In contrast, the present study applies a similar toolkit, heuristic analysis, task-based testing, SUS, and PSSUQ, to an IFRS 17-oriented insurance core system (TICS) and uses the results to derive design recommendations directly tied to those critical insurance workflows.

Methods

To ensure a thorough evaluation of the TICS application, this research employed a mixed-method approach, integrating qualitative user-need discovery with quantitative usability assessment (Huyler & McGill, 2019). Initially, a qualitative investigation using task-based scenarios was conducted to inform the application's design. The resulting TICS prototype was then formally evaluated using quantitative instruments, namely the System Usability Scale (SUS) and the Post-Study System Usability Questionnaire (PSSUQ), to collect empirical user feedback. This study's methodology began with a problem-identification stage, gathering primary data through the examination of documentation and interviews. This process gave way to a deep understanding of the user-centric challenges posed by the current system. After defining the research questions, a systematic literature review was undertaken. This review focused on established theories and practices in usability evaluation broadly as well as the specific context of the TICS application to inform the methodological framework. The selected research instruments were multifaceted: task analysis was used to drive iterative design improvements during usability testing, while the SUS and PSSUQ quantitatively validated these enhancements by measuring user satisfaction and overall performance. Table 1 presents a summary of the employed methods.

Table 1. Research Approach and Methodology

Attributes	Information
Classifications	Integrated methodology (mixed-methods research)
Category	In-depth investigation via case study
Paradigm	Evaluative and assessment-focused
Purpose	To assess the usability of the TICS application
Results	Formulate enhancement recommendations for the TICS application
Data collection	Quantitative data (SUS & PSSUQ surveys); qualitative data (usability testing & open-ended inquiries)
Sampling approach	Purposive/selective sampling

In this study, participants were asked to engage with pre-established task scenarios, while subsequent interviews served to explore user feedback and identify issues in greater detail. Participants from the prior usability-testing phase also completed questionnaires. Following the initial problem analysis, identified issues were systematically classified using principles of heuristic evaluation. Solutions were then formulated and refined against the backdrop of three pertinent usability standards. The developed prototype was evaluated via user-centric task analysis, which was quantitatively substantiated by the administration of the SUS and PSSUQ instruments. The SUS facilitated a detailed evaluation of the TICS application.

This method was justified by its statistical reliability even with small sample sizes. Participant responses were collected using a five-point Likert scale (1–5). The PSSUQ fulfilled a dual role within the research design. Initially, it was deployed to establish a baseline assessment of the pre-existing information system. Later, it was used again to measure the effectiveness of the design enhancements in the final TICS application. This study employed the third version of the PSSUQ, which consists of 16 components measured along a seven-point Likert scale. The overall score was calculated by averaging the scores from each category.

Open-Ended Questions

To achieve an in-depth understanding of the system's usability, interface quality, and functionality from the user's perspective, open-ended questions were administered to respondents. The purpose of these questions (detailed in Table 2) was to gather nuanced insights into user experiences and identify specific areas for potential improvement. The resulting qualitative responses were instrumental in refining the system to better align with user needs.

Table 2. Set of Open-Ended/Free-Response Questions

Questions
What obstacles, if any, did you encounter while using the system's features?
Please describe any aspects of the interface that were not immediately clear or intuitive to you.
Did any features or functions behave in an unexpected way or fail to work correctly? Please describe any such instances.
How would you describe the system's responsiveness to your commands and actions?
Please describe your perception of the system's reliability. Did you feel it was dependable for completing your work?
If you could change any part of the system, what would be your top recommendations to enhance user satisfaction?
If you used the user manual, please describe its effectiveness and your experience with it.
Please provide any final recommendations you have for making the system more intuitive and efficient for your tasks.

Task Scenarios

To ensure a structured evaluation, task scenarios were designed to direct participants toward achieving specific objectives. (See Table 3 for a detailed breakdown of the task list.) This methodology was first applied to assess the usability of the current information system. Furthermore, the analysis of these user-performed tasks was instrumental to eliciting the key requirements for the TICS application.

Table 3. Task Scenarios

Task	Scenario
Search Policy	Please search for an existing policy number and view the detailed data.
Create New Policy	Create a new insurance policy based on the provided data.
Add Vehicle Details	Add the detailed data on the vehicle.
Calculate Premium	Calculate the premium based on the previously registered vehicle data.
Calculate R/I Placement	Calculate the R/I placement based on the detailed data.
Assign Account Holder	Assign an account holder to be responsible for the policy.
Assign Underwriter	Assign an underwriter to be responsible for the business class.
Validate Terms and Conditions	Ensure that the terms and conditions are correct.
Assign Document Signer	Assign an underwriter to sign the document.
Verify Quotation	Verify that the content in the quotation document is correct.
Submit Data for Policy Number	Submit all data to ensure the generation of a policy number.
Verify Insurance Document	Verify that the content in the insurance document is correct.

Sampling

The target population for this study comprised internal users of the Tugu Insurance Core System (TICS) at Tugu Insurance. These users are employees who rely on the TICS web application in their daily work for tasks such as policy creation, premium calculation, and related administrative processes. The unit of analysis is the individual TICS user. Data were collected in two stages.

In the first stage, an online usability questionnaire was distributed to TICS users between 5 and 16 May 2025. The questionnaire consisted of four parts: (1) nine general questions capturing respondent characteristics and usage patterns; (2) the ten items of the System Usability Scale (SUS); (3) sixteen items from the Post-Study System Usability Questionnaire (PSSUQ); and (4) eight open-ended questions designed to elicit more detailed feedback on usability problems and improvement suggestions. The initial target was 30 respondents across four user categories (end users, shared-services staff, programmers, and interns). In practice, 16 internal employees completed the questionnaire: 7 end users (43.75%), 4 shared-services staff (25%), 3 programmers (18.75%), and 2 interns (12.5%).

In the second stage, usability testing was conducted online between 21 and 28 May 2025 using predefined task scenarios that reflect key TICS workflows. Because the number of active TICS web-application users is limited, participants for this stage were selected using stratified sampling to ensure representation from each of the main user categories (end user, shared services, programmer, and intern).

Only respondents who had completed the Stage-1 questionnaire and explicitly agreed to take part in follow-up testing were eligible. A total of 10 participants (62.5% of questionnaire respondents) took part in the usability tests: 4 end users (40%), 2 shared-services staff (20%), 2 programmers (20%), and 2 interns (20%). These distributions ensure that both the survey and the usability testing reflect the diversity of TICS user roles within Tugu Insurance.

Data Collection

Empirical data for this study were gathered across two distinct stages. The initial stage was dedicated to a baseline assessment of the current TICS application. This entailed the acquisition of qualitative data through the direct observation of representative users as they engaged in task scenarios supplemented by interviews. Following these observational tasks, quantitative data were collected from the same participant cohort via an online survey comprising demographic inquiries, the PSSUQ and SUS instruments, and eight open-ended questions aimed at eliciting nuanced feedback on existing challenges.

The second data-collection stage centered on evaluating the implemented design enhancements. Employing a parallel methodology, this phase utilized the same approach (e.g., testing, interviews, SUS, PSSUQ, open-ended questions) to gather feedback on the new prototype. This two-phase approach was strategically designed to yield comparative usability metrics and enable a pre- and post-intervention analysis of the SUS and PSSUQ scores.

Data Processing

This study's data-processing strategy was separated into qualitative and quantitative components. For the qualitative data derived from usability tests and open-ended responses, a formal analytical approach was adopted. This approach commenced with data reduction, which entailed the systematic selection, coding, and categorization of raw data from session recordings and verbatim transcripts. In the subsequent data-presentation stage, transcribed qualitative data were meticulously analyzed through open coding. The resulting codes were then quantified into metrics and structured to reveal recurring patterns of user-reported usability deficits. In the final stage (drawing conclusions), the processed information was synthesized to form robust inferences regarding the challenges and feature requirements expressed by TICS users.

Findings

Data were collected through an online SUS/PSSUQ questionnaire and task-based usability testing. The questionnaire was distributed to TICS users between 5 and 16 May 2025 and was completed by 16 internal employees. Between 21 and 28 May 2025, 10 of these respondents also participated in task-based usability testing using predefined TICS scenarios. These participants represented the main user categories of the system, end users, shared-services staff, programmers, and interns. The respondent profile and recruitment across both stages are summarized in Table 4.

Table 4. Respondent Profile and Recruitment

User Category	Questionnaire Respondents n (%)	Usability-testing Participants n (%)
End user	7 (43.75%)	4 (40%)
Shared services	4 (25%)	2 (20%)
Programmer	3 (18.75%)	2 (20%)
Intern	2 (12.5%)	2 (20%)
Total	16 (100%)	10 (100%)

Initial TICS Assessment Results (As-Is)

The data from the usability questionnaires were analyzed in line with established guidelines for the SUS and PSSUQ instruments to ensure the integrity of the findings. The analysis encompassed responses from all 16 study participants. These instruments are particularly justified for studies with limited sample sizes. The SUS is known to produce dependable results with small cohorts; a sample of eight is sufficient for a 95% confidence level with ± 15 precision.

Likewise, the PSSUQ is sufficiently sensitive to consistently detect specific differences, making it appropriate for small sample sizes. As this study's primary instrument, the SUS featured ten questions measured along a five-point Likert scale (1 to 5). Each participant's responses were added together to calculate an overall SUS score. The detailed results from the 16 respondents are presented in Table 5.

Table 5. SUS Score of TICS (As-Is)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Score
R01	2	5	2	5	3	1	4	5	4	4	37.5
R02	3	5	3	5	2	4	2	1	3	4	35
R03	5	2	3	4	3	3	4	2	4	2	65
R04	2	2	4	1	2	3	2	3	3	5	47.5
R05	4	1	5	4	2	4	5	3	5	2	67.5
R06	3	2	4	1	3	3	2	2	5	1	70
R07	5	2	4	1	3	5	3	2	5	5	62.5
R08	2	4	4	5	4	5	4	2	4	2	50
R09	5	3	2	4	2	4	3	4	5	1	52.5
R10	3	4	2	5	3	1	2	4	2	2	40
R11	3	2	2	2	5	4	3	3	5	5	55
R12	2	3	2	2	5	2	3	2	2	2	57.5
R13	3	5	2	1	4	1	2	5	3	5	42.5
R14	3	4	2	4	5	5	4	5	5	2	47.5
R15	3	4	2	2	4	4	3	4	2	2	45
R16	5	4	3	4	2	5	3	5	2	2	37.5

The SUS score, ranging from 1 to 100, was employed to evaluate the TICS assessment. The overall SUS score derived from 16 usability questionnaire respondents was 50.8, below the average SUS research value of 68, suggesting the existence of usability issues within TICS that necessitate improvement. The individual scores ranged from a minimum of 35 to a maximum of 70. Given the overall SUS score of 50.8, TICS may be classified within the F or Poor category, indicating a clear need for enhancements to achieve a more favorable score. The second instrument used in this study, the PSSUQ, consisted of 16 questions measured along a scale of 1 to 7, with 1 signifying "strongly agree" and 7 signifying "strongly disagree." Table 6 presents the results of the PSSUQ questionnaire from the 16 usability questionnaire respondents, with each question falling under one of four categories.

Table 6. Comparative Analysis of PSSUQ Mean Scores (As-Is)

Category	Mean Value (As-Is)	Standard Value
System Usefulness (No. 1–6)	3.14	2.80
Information Quality (No. 7–12)	3.28	3.02
Interface Quality (No. 13–15)	3.06	2.49
Overall (No. 16)	3.17	2.82

This study compared its PSSUQ results against established benchmarks, which define standard values for overall satisfaction (2.82), system usefulness (2.80), information quality (3.02), and interface quality (2.49). Table 6 presents average values for each category among the 16 participants. Across all four categories, the average PSSUQ scores for the current TICS application are higher than the normative benchmark values. Because the PSSUQ uses a scale where lower scores indicate more favorable user experiences, these results show that TICS performs below the benchmark in terms of perceived system usefulness, information quality, interface quality, and overall satisfaction. In other words, users experience more friction and lower satisfaction than would be expected for a typical system in the normative dataset, confirming that substantial usability improvements are needed.

Figure 1. PSSUQ Evaluation (As-Is)

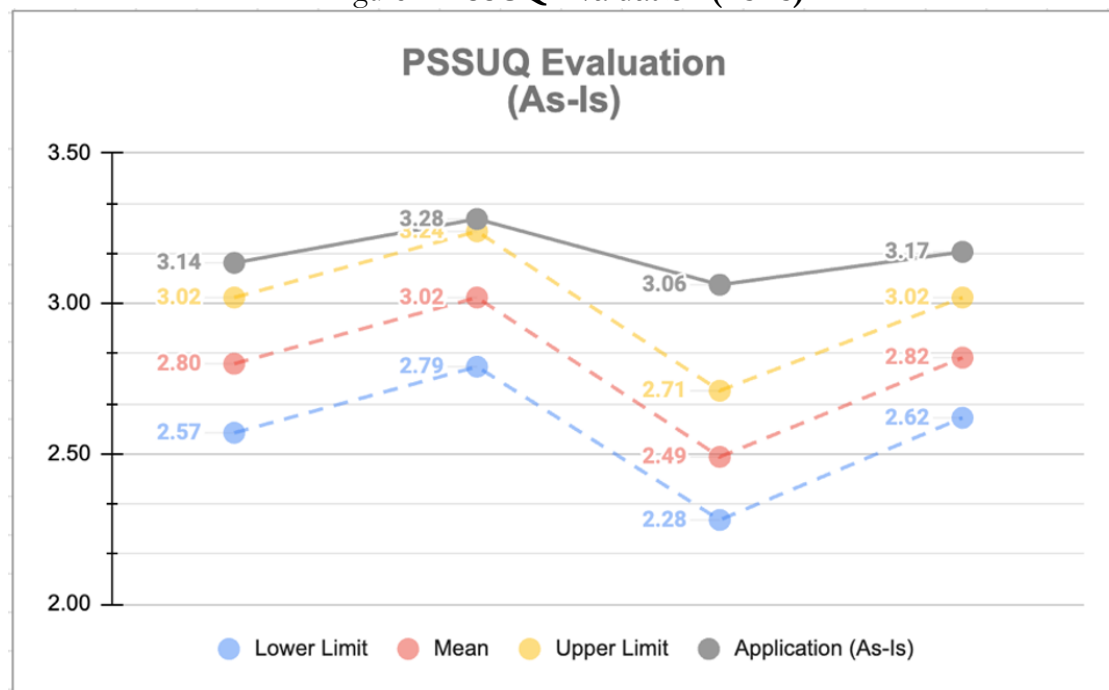


Figure 1 presents a comparative diagram that plots the study's average PSSUQ category scores against the established normative values. The figure shows that, for all four categories, the TICS As-Is scores lie above (i.e., worse than) the benchmark, visually reinforcing that the current system falls short of expected usability levels and that there is clear room for improvement. The scores obtained through these initial SUS and PSSUQ assessments form a baseline measurement that is crucial for the next phase of the study, in which it will be compared against new scores obtained following the evaluation of the proposed design enhancements. This subsequent analysis will ascertain the tangible impact of the modifications on the system's usability scores.

Open-Ended Analysis and Usability Testing Mapping

The open-ended questions yielded qualitative feedback from ten respondents, highlighting a range of user problems with the TICS application. Based on their responses, a total of 16 distinct usability issues were identified and collated. To structure the analysis, these issues were systematically coded and categorized by mapping them against three established usability frameworks, Nielsen's heuristics, Marenkov's rules, and Schneiderman's golden

rules, providing a formal basis for the development of targeted recommendations for interface design improvements.

Understanding these issues and the corresponding recommendations is crucial for enhancing the TICS interface. By addressing such user-reported problems, the system may be optimized to better meet user needs and improve overall usability, contributing to a more efficient and satisfactory user experience. By incorporating qualitative user feedback, this iterative-design-centered approach ensures that the TICS interface is continually refined to ensure alignment with user expectations and usability principles. It enhances user satisfaction as well as overall system effectiveness. Table 7 breaks down usability issues along the lines of Nielsen's Ten Heuristics. The most prominent area for improvement was "Aesthetic and Minimalist Design," which recorded the highest number of problems at four.

Table 7. Problem Results Mapped with Nielsen's Standards

Nielsen's Ten Usability Heuristics for Interface Design	Number of Problems
Visibility of System Status	-
Match Between the System and the Real World	1
User Control and Freedom	1
Consistency and Standards	3
Error Prevention	3
Recognition rather than Recall	1
Flexibility and Efficiency of Use	1
Aesthetic and Minimalist Design	4
Help Users Recognize, Diagnose, and Recover from Errors	2
Help and Documentation	-

A moderate number of issues were found in both "Consistency and Standards" and "Error Prevention," with three problems identified in each category. In contrast, the heuristics of "Match Between System and the Real World," "User Control and Freedom," "Recognition rather than Recall," and "Flexibility and Efficiency of Use" had only one problem each. Notably, no issues were reported for the remaining two heuristics. Table 8 investigates issues against Marenkov's Nine Usability Guidelines, indicating that usability issues mapped to only four of the nine principles. A plurality of the problems fell under "Organization of Information and Content" (6 issues), while fewer issues were identified in "Page Layout and Navigation" (5 issues) and "Screen-Based Controls" (4). Just one minor issue was noted when it came to "Home Page and Search", and the other five categories entailed no issues.

Table 8. Problem Results Mapped with Marenkov's Standards

Marenkov's Nine Usability Guidelines	Number of Problems
Accessibility and Compatibility	-
Home Page and Search	1
Page Layout and Navigation	5
Scrolling, Paging, and Links	-
Headings, Titles, and Labels	-
Text Appearance and Lists	-
Screen-Based Controls	4
Graphics, Images, and Multimedia	-
Organisation of Information and Content	6

Table 9 shows that only five of Schneiderman's Eight Golden Rules of Interface Design were violated. The "Prevent Errors" category had the highest number of problems at six, followed by "Cater to Universal Usability," "Support Internal Locus of Control," and "Reduce Short-Term Memory Load" with three problems each. The "Strive for Consistency" category had just one problem, while the remaining three categories had no problems.

Table 9. Problem Results Mapped with Schneiderman's Standards

Schneiderman's Eight Golden Rules of Interface Design	Number of Problems
Strive for Consistency	1
Cater to Universal Usability	3
Offer Informative Feedback	-
Design Dialogue to Yield Closure	-
Prevent Errors	6
Permit Easy Reversal of Actions	-
Support Internal Locus of Control	3
Reduce Short-Term Memory Load	3

Following the compilation of these issues, proposed improvement solutions were categorized based on the task scenarios employed during usability testing. This method served to directly connect each suggested enhancement to the specific user workflow where a problem was observed. This organizational structure, detailed in Table 10, clearly maps each recommendation to its relevant component within the prototype design.

Table 10. Task Name and Necessary Fixes

Task Name	Number of Fixing
Search Policy	4
Create New Policy	8
Add Vehicle Details	1
Calculate Premium	1
Calculate R/I Placement	No fix needed
Assign Account Holder	1
Assign Underwriter	No fix needed
Validate Terms and Conditions	No fix needed
Assign Document Signer	1
Verify Quotation	No fix needed
Submit Data for Policy Number	No fix needed
Verify Insurance Document	No fix needed

Second TICS Assessment Results (To-Be)

The refined prototype's design outcomes were presented as a clickable prototype developed using prototyping tools. Subsequently, the prototype underwent re-testing with a cohort of 10 usability-testing participants.

The evaluation of the enhanced prototype was conducted using contextual interviews aimed at assessing users' comprehension of and perspectives on the improved prototype via pre-established task scenarios.

In parallel, a second round of SUS and PSSUQ questionnaires was administered to the full group of 16 survey respondents (R01–R16) to obtain comparable pre- and post-intervention

usability scores. This iterative process enabled the direct observation of user interactions and feedback on specific design elements. Collected data facilitated the identification of areas in which the prototype excels and areas still in need of refinement to meet user expectations and usability standards. These methods gave way to a comprehensive understanding of the prototype's strengths and weaknesses, informing further design iterations and improvements. This thorough evaluation ensured that the final system would align closely with both user needs and usability principles.

Table 11. SUS Score of Prototype (To-Be)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Score
R01	4	4	2	4	2	2	3	2	3	4	75
R02	2	4	3	3	4	4	4	3	2	3	80
R03	4	4	2	3	2	4	2	4	3	2	75
R04	3	4	3	2	4	4	4	4	2	4	85
R05	4	3	2	3	2	3	3	2	3	2	67.5
R06	4	2	3	3	2	4	3	2	2	2	67.5
R07	3	4	3	2	4	4	4	4	3	2	82.5
R08	2	3	2	3	4	2	4	3	3	2	70
R09	3	4	3	4	2	3	4	4	2	2	77.5
R10	4	3	3	4	2	2	4	4	4	4	85
R11	2	4	3	2	3	4	4	3	3	2	75
R12	4	3	4	2	3	2	3	2	3	4	75
R13	2	3	3	4	3	3	2	2	4	2	70
R14	4	3	4	4	3	2	2	3	4	4	82.5
R15	4	3	4	4	3	4	2	3	4	4	87.5
R16	3	4	4	4	4	3	4	3	4	3	90

Table 11 presents the post-improvement SUS results, which showcase an average score of 77.8 from the 16 questionnaire respondents. This represents a substantial increase from the pre-intervention score of 50.8 (see Table 5) and places the TICS application well above the average usability benchmark of 68. The score of 77.8 corresponds to a grade of B+, indicating "Good" overall usability and suggesting that design enhancements effectively met user needs.

The PSSUQ, which was completed by the same 16 respondents, produced a mean score across all categories of 2.76, reinforcing the positive impact of the design changes. Table 12 presents the detailed PSSUQ results for the improved prototype.

Table 12. Comparison of Mean Values of PSSUQ

Category	Mean Value (As-Is)	Mean Value (To-Be)	Standard Value
System Usefulness (No. 1–6)	3.14	2.70	2.80
Information Quality (No. 7–12)	3.28	2.82	3.02
Interface Quality (No. 13–15)	3.06	2.63	2.49
Overall (No. 16)	3.17	2.76	2.82

The “Mean Value (To-Be)” column presents the PSSUQ scores obtained from the evaluation of the enhanced prototype. In all categories, the To-Be mean values are **lower** than the

corresponding As-Is means, indicating improved user perceptions across system usefulness, information quality, interface quality, and overall satisfaction.

Interpreting the PSSUQ scale, where lower scores represent more favorable evaluations, these reductions signal that the redesign reduced user frustration and made the system easier to use. The refined prototype therefore better aligns the interface and workflows with user expectations, leading to a more usable and efficient system overall.

Figure 2. PSSUQ Evaluation (As-Is and To-Be)

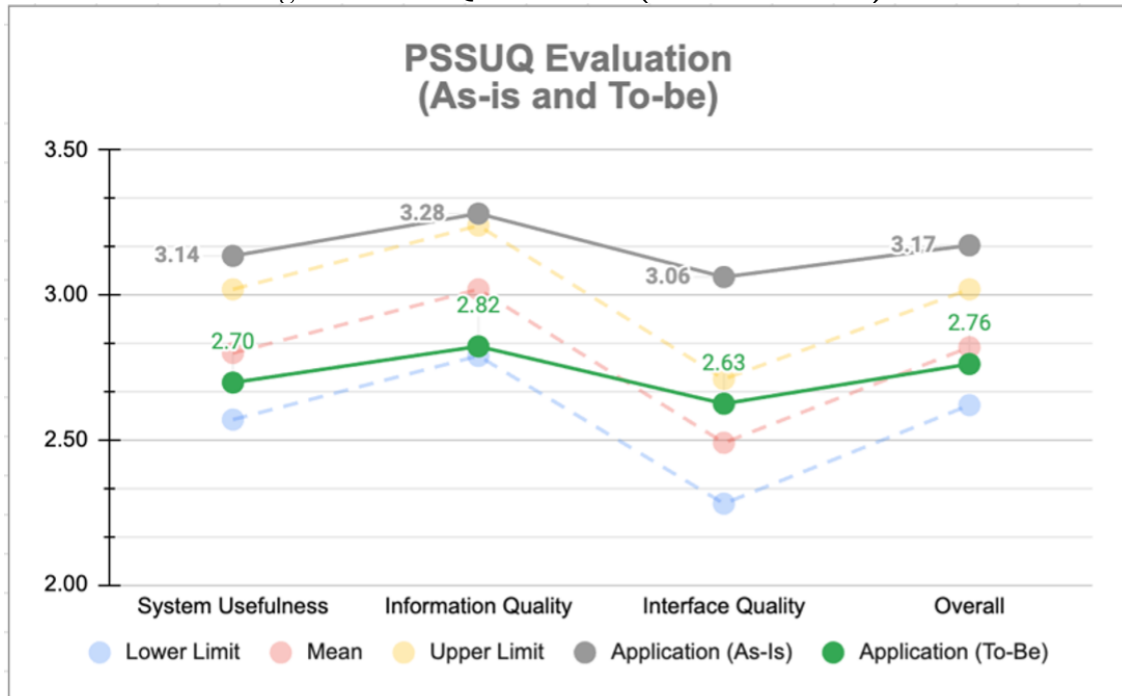


Figure 2 compares the As-Is and To-Be PSSUQ category means against the standard benchmark values. In the To-Be condition, three of the four categories—system usefulness, information quality, and overall satisfaction, now fall below (i.e., better than) the normative means, while interface quality is substantially closer to the benchmark than in the As-Is evaluation, although it remains slightly higher.

Taken together with the As-Is results, the figure shows a clear convergence toward or beyond the benchmark across all categories after the redesign. This pattern supports the conclusion that the implemented design changes had a tangible positive impact on the perceived usability of TICS. The pattern of improvement in SUS and PSSUQ scores after the redesign is consistent with prior research on usability mechanisms in complex information systems. Ferreira et al. (2019) found that targeted mechanisms such as Abort Operation, Progress Feedback, and Preferences can significantly improve efficiency, effectiveness, and user satisfaction when incorporated into system interfaces.

Similarly, Subiyakto et al. (2021) and Almasi et al. (2023) showed that combining heuristic analysis, task-based usability testing, and standardized instruments like SUS and PSSUQ is an effective strategy for diagnosing problems and guiding interface improvements in organizational and decision-support systems. The present case study extends these findings to an insurance core system that supports IFRS 17-relevant workflows. The reduction in

usability problems mapped to Nielsen’s heuristics, Marenkov’s guidelines, and Shneiderman’s rules, together with the substantial increase in the SUS score from 50.8 to 77.8 and the improvement of PSSUQ scores toward or beyond benchmark values, demonstrates that applying these established usability frameworks can materially enhance the usability of core transactional systems. In this way, the results provide empirical support for the transferability of usability-engineering practices from generic web and enterprise systems to the more specialized domain of insurance information systems.

Usability testing was also performed with ten participants to assess the enhanced prototype across 12 task scenarios. The entire testing procedure was recorded, and task-completion results were documented to pinpoint any challenges encountered during system use, as shown in Table 13.

Table 13. Problem Matrix for Usability Testing Respondents (To-Be)

Respondent	1	2	3	4	5	6	7	8	9	10	11	12
R01												
R03					X							
R05												
R06				X								
R10												
R11												
R13					X							
R14												
R15												
R16				X								
Total Problems	0	0	0	2	2	0	0	0	0	0	0	0
Frequency (%)	0%	0%	0%	20%	20%	0%	0%	0%	0%	0%	0%	0%

During this usability testing, all user-reported issues were documented alongside the frequency of their occurrence within each task scenario. To quantify this, a problem-frequency percentage was calculated for each issue by dividing the total number of times a problem occurred by the total number of test participants. As detailed in Table 14, these calculated frequencies were then classified into four distinct levels.

Table 14. Frequency of Problem Scenarios in Usability Testing (To-Be)

Level	Status	Frequency Range	Scenarios
1	Low	≤10%	1, 2, 3, 6, 7, 8, 9, 10, 11, 12
2	Medium	11–50%	4, 5
3	High	51–89%	-
4	Very high	≥90%	-

A comparison of the initial (“as-is”) and prototype (“to-be”) evaluations revealed a marked reduction in problem frequency. While the initial system had two tasks with level-3 problem frequencies and four tasks with level-2 problem frequencies, the improved prototype successfully eliminated all level-3 issues so that all task scenarios now fall within the low- or medium-frequency range (Levels 1 and 2).

From an IFRS 17 readiness perspective, the improved TICS design directly supports several system and data challenges highlighted in IFRS 17 implementation literature. More

streamlined flows for quotation creation, premium calculation, reinsurance placement, policy issuance, and policy-number generation help ensure that granular contract-level data are captured consistently and accurately at the point of entry, rather than being corrected downstream. Clearer screen layouts, reduced information clutter, and explicit validation messages lower the risk of omissions and inconsistencies in cash-flow inputs that later feed actuarial models and accounting engines.

Although this study does not evaluate the back-end IFRS 17 calculation engine, the evidence indicates that a user-centered front-end design can strengthen data quality and operational reliability in IFRS 17-relevant processes, thereby reducing the likelihood of manual rework and non-compliance in subsequent reporting cycles.

Conclusion

This study evaluated and improved the usability of the Tugu Insurance Core System (TICS), a core application used to support IFRS 17-related workflows at Tugu Insurance. Using a mixed-methods, user-centred design approach over ten stages, from problem identification and initial evaluation to redesign and re-testing, it combined heuristic analysis, task-based usability testing, and SUS and PSSUQ questionnaires.

The baseline evaluation with 16 respondents produced an overall SUS score of 50.8, placing TICS in the “poor” category, and PSSUQ scores above the benchmark in all dimensions, indicating lower perceived system usefulness, information quality, interface quality and overall satisfaction. After the redesign, a final evaluation using the same 16 respondents for SUS and PSSUQ and 10 participants for task-based usability testing showed a clear improvement: the SUS score rose to 77.8 (B+; “good”) and PSSUQ means moved toward or beyond the benchmark values. These results confirm that the implemented design changes substantially improved the system’s usability.

From an IFRS 17 readiness perspective, the improved flows for quotation creation, premium calculation, reinsurance placement, policy issuance and policy-number generation support more consistent and accurate capture of contract-level data at the point of entry. Clearer screen layouts and validation messages reduce data-entry errors and rework, helping to mitigate the risk of late or inaccurate IFRS 17 reporting. The case study therefore illustrates how usability engineering can be used not only to improve user satisfaction but also to strengthen the operational side of regulatory compliance in insurance information systems.

This research has several limitations. The sample size is modest (16 questionnaire respondents and 10 usability-testing participants), which constrains statistical generalization. The study also relies on the SUS, even though more recent alternatives such as ESUS have been proposed for enterprise applications (Schneider et al., 2024). In addition, the findings are based on a single organization. They should therefore be interpreted as analytical generalizations and design lessons for similar IFRS 17 implementations, rather than as estimates for the entire Indonesian insurance industry.

Future work could extend the analysis to other insurers, incorporate ESUS or other enterprise-focused instruments, and link front-end usability metrics more directly to back-end IFRS 17 calculation and reporting outcomes. At a practical level, the results suggest

that Indonesian insurers need to treat usability as an integral part of their IFRS 17 system programmes, rather than as a cosmetic layer. Coordinated effort between regulators, system developers and end-users will be required to develop information systems that both meet IFRS 17 requirements and support efficient day-to-day operations.

References

- Almasi, S., Bahaadinbeigy, K., Ahmadi, H., Sohrabei, S., & Rabiei, R. (2023). Review of Usability Evaluation of Dashboards: A Systematic Literature Review of Tools. *BioMed Research International*, 2023(1). Hindawi Publishing Corporation. <https://doi.org/10.1155/2023/9990933>
- Aziz, N. S., Sulaiman, N. S., Hassan, W. N. I. T. M., Zakaria, N. L., & Yaacob, A. (2021). Review of A Review of Website Measurement for Website Usability Evaluation. *Journal of Physics Conference Series*, 1874(1), 12045. IOP Publishing. <https://doi.org/10.1088/1742-6596/1874/1/012045>
- Boston Consulting Group. (2023). *IFRS 17 calls for a new take on data and technology*. Boston Consulting Group.
- Chartis Research. (2021). *The heart of IFRS 17 compliance: Why data and data architecture are vital in addressing insurers' technology challenges*. Chartis Research / Oracle.
- Clark, N., Dabkowski, M., Driscoll, P. J., Kennedy, D., Kloo, I., & Shi, H. (2021). Empirical Decision Rules for Improving the Uncertainty Reporting of Small Sample System Usability Scale Scores. *International Journal of Human-Computer Interaction*, 37(13), 1191. <https://doi.org/10.1080/10447318.2020.1870831>
- Dahiyat, A., & Owais, W. O. (2021). The expected impact of applying IFRS (17) insurance contracts on the quality of financial reports. *Accounting*, 581. <https://doi.org/10.5267/j.ac.2020.12.021>
- Eckert, C., & Osterrieder, K. (2020). How digitalization affects insurance companies: overview and use cases of digital technologies. *Zeitschrift Für Die Gesamte Versicherungswissenschaft*, 109(5), 333. <https://doi.org/10.1007/s12297-020-00475-9>
- EY. (2017). *Impacts of IFRS 17: Why data, systems and processes are important to IFRS 17*. EYGM Limited.
- EY. (2019). *IFRS 17 implementation: Practical challenges and lessons learned*. Ernst & Young Global Limited.
- Faiza, D., Farrell, G., Delianti, V. I., Thamrin, T., & Rahmadika, S. (2022). Eight Golden Rules Interface Analysis for Video Conference Information System. 131. <https://doi.org/10.1109/iceltics56128.2022.9932115>
- Ferreira, J. M., Acuña, S. T., Dieste, Ó., Vegas, S., Santos, A., Rodríguez, F. D., & Juristo, N. (2019). Impact of usability mechanisms: An experiment on efficiency, effectiveness and user satisfaction. *Information and Software Technology*, 117, 106195. <https://doi.org/10.1016/j.infsof.2019.106195>
- Hou, C., & Mayernik, M. S. (2021). Improving the Usability of Organizational Data Systems. *International Journal of Digital Curation*, 16(1), 21. <https://doi.org/10.2218/ijdc.v16i1.592>
- Huyler, D., & McGill, C. M. (2019). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, by John Creswell and J. David Creswell. Thousand Oaks, CA: Sage Publication, Inc. 275 pages, \$67.00 (Paperback). New Horizons in Adult Education and Human Resource Development, 31(3), 75. <https://doi.org/10.1002/nha3.20258>

- International Accounting Standards Board. (2017). *IFRS 17 Insurance Contracts*. London: IFRS Foundation.
- Kurniawan, S. (2003). Review of Interaction design. *ACM SIGCHI Bulletin*, 2003, 15. <https://doi.org/10.1145/967199.967218>
- Lehnert, F. K., Doublet, S., & Sim, G. (2024). Designing Usability/User Experience Heuristics to Evaluate E-Assessments Administered to Children. *Computer Standards & Interfaces*, 92, 103933. <https://doi.org/10.1016/j.csi.2024.103933>
- Lewis, J. R. (2002). Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, 14, 463. <https://doi.org/10.1080/10447318.2002.9669130>
- Marcus, A. (2009). Ben Shneiderman and Catherine Plaisant, with contributing authors Maxine S. Cohen and Steven M. Jacobs. Designing the User Interface. *Information Design Journal*, 17(2), 157. <https://doi.org/10.1075/idj.17.2.14mar>
- Marenkov, J., Robal, T., & Kalja, A. (2018). Design-Time Web Usability Evaluation with Guideliner. *Complex Systems Informatics and Modeling Quarterly*, (15), 90. <https://doi.org/10.7250/csimq.2018-15.05>
- Mator, J. D., Lehman, W., McManus, W., Powers, S. A., Tiller, L., Unverricht, J., & Still, J. D. (2020). Usability: Adoption, Measurement, Value. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 63(6), 956. <https://doi.org/10.1177/0018720819895098>
- Meyeroff, E., & Tremoulet, P. D. (2021). Etiometry's T3 Heuristic Evaluation. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 10(1), 37. <https://doi.org/10.1177/2327857921101015>
- Nielsen, J. (2024). 10 Usability Heuristics for User Interface Design. Retrieved from <https://principles.design/examples/10-usability-heuristics-for-user-interface-design>
- Parwitasari, L. (2022). *Analisis isu kritis akuntansi dalam rangka penerapan PSAK No. 74 (IFRS 17) pada program asuransi sosial X* [Analysis of critical accounting issues in the implementation of PSAK No. 74 (IFRS 17) in Social Insurance Program X] (Master's thesis). Universitas Indonesia, Depok, Indonesia.
- Schneider, S., Hillman, S., Bach, P., & Ma, G. (2024). ESUS: Aligning and simplifying SUS for enterprise applications. *Journal of User Experience*, 19, 60. Retrieved from <https://www.microsoft.com/en-us/research/publication/esus-aligning-and-simplifying-sus-for-enterprise-applications/>
- Sharfina, Z., & Santoso, H. B. (2016). An Indonesian adaptation of the System Usability Scale (SUS). 145. <https://doi.org/10.1109/icacsis.2016.7872776>
- Subiyakto, A., Rahmi, Y., Kumaladewi, N., Huda, M. Q., Hasanati, N., & Haryanto, T. (2021). Investigating quality of institutional repository website design using usability testing framework. *AIP Conference Proceedings*, 2331, 60016. <https://doi.org/10.1063/5.0041677>
- Scott, D. (2023). *Challenges applying IFRS 17: Insurance contracts*. World Bank, Centre for Financial Reporting Reform.

