

# **Indonesian Journal of Computer Science**

ISSN 2549-7286 (*online*) Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

# Usability Evaluation Of Social Security For Workers Applications In Public Institution

#### Asra Afrizon<sup>1</sup>, Panca Hadi Putra<sup>2</sup>

<sup>1</sup>Asra.afrizon@ui.ac.id, <sup>2</sup>hadiputra@ui.ac.id Fakultas Ilmu Komputer, Magister Teknologi Informasi, Universitas Indonesia

Article Information	Abstract
Submitted : 18 Jul 2023 Reviewed: 24 Jul 2023 Accepted : 30 Aug 2023	Workforce social security application services known as JMO are owned by the Indonesian government and held by the Social Security Administering Body on Employment. Initially, the application only provided social security programs, membership, and benefits information. In September 2021, the application was transformed by including transactions for claim redemption.
Keywords	User reviews of the application gave ratings of 1.5 and 4.4, respectively
Usability, User Experience Questionnaire, Task Scenario, Post Test Interview	highest, accounting for 65% and 5%, respectively. Through analysis of user reviews and complaints, it has been determined that users encounter multiple constraints when utilizing applications concerning accessibility, information, authentication, procedures, and comprehension of usage. The research method involves a quantitative approach using a user experience questionnaire and usability testing methods through task scenarios. Followed by a qualitative approach using a post-test interview to conduct a deep dive into user problems and expectations. This study aims to evaluate the application's usability and provide recommendations for improvement. Based on the analysis from UEQ, the application gets an average value positive, and the benchmark result is above average. In contrast, the novelty gets neutral, and the benchmark result is below average, reviewed as conventional, slow, less valuable, challenging to learn, and monotonous. The test task scenario described the effectiveness of the application. Meanwhile, time-based efficiency is 0,04 goals/second, which can be interpreted as high speed. Task scenario finding explains 30 problems in which three tasks are prioritized. Post-test interviews explained user need for frequently accessed features, difficulty to use, and usage limitations. The recommendations aimed to adjust the interface and functionality to meet user needs, providing a satisfying experience and shortening the self-service process.

#### A. Introduction

In the present era, characterized by the prevalence of information and communication, individuals rely on computer technology to connect with others and their communities. Mobile technology is available in a lightweight and portable form, supported by software that provides users with reliable performance for specific tasks [2]. Mobile applications can be downloaded and utilized on mobile devices with compatible operating systems [1]. These applications are designed to run on smartphones or other portable devices, allowing users to easily connect to the internet [2].

Given the significance of mobile applications in enhancing the quality of public services, companies must be able to innovate and adapt to meet customer expectations, while for the Government it can overcome the digital divide through personalization that suits the needs of each community [22]. The implementation of digital claim services has improved the overall performance of public service applications. User feedback through application providers such as the App Store [16] with a rating of 1.5 out of 5 and the Google Play Store [17] with a rating of 4.4 in February 2022 and 4.7 in March 2023. However, some users still face difficulty in using social security services for workers, including issues related to application policies, data quality, application quality, environmental factors, devices, and user errors.

The number of users of the JMO application through Google Play is higher with 10 million downloads with 2.3 million reviews while the downloaded App Store has a smaller number of users with 7 thousand reviews. Reviews through the Google Play store have a higher rating because assessments through reviews with high ratings some of which explain the shortcomings and obstacles to using the application, and users can provide changes to reviews and ratings so that after being given a reliever service the user provides the best rating. Through user reviews that provide a very good rating assessment of 4.7 out of 5 on Google Play and 1.5 out of 5 on the App Store, but the description of the contents of the review still shows the obstacles faced by users, so there are reviews that are not consistently delivered with the numerical assessment given. In contrast to the App Store with users and reviews tending to be low, it has a lagging version with Android, so changes and obstacles that have been fixed still cannot be felt by iOS users. Based on the results of unfavorable reviews, it is necessary to evaluate and test users to get more detailed information to find out the details of the problem by looking at the user's pain points when interacting with the application, to find out the types and criteria of obstacles found, as well as user understanding of the features used.

Rating	Reviews	Category		
5	* * * * * 5 April 2022 19 Aplikasi tidak bisa dibuka lagi setelah diperbaharui, selalu tulisan nya kesalahan di akun sm password berbeda, dan akun tidak ditemukan. Padahal akun dan password nya tidak ada sy rubah msh tetap yg pertama. Kenapa skrg tidak bisa dibuka lagi???	Data not found, cannot be accessed		
4	<ul> <li>* * * * 11 April 2022</li> <li>Proses pengkinian data sangat sulit, saya sampai tahap biometrik dan tidak bisa lanjut lagi setelah gagal di 3x kesempatan, karna dikolom syarat dan ketentuan tidak ada tulisan</li> <li>"hanya 3x kesempatan" mohon diperbaiki aplikasinya, paling tidak beri alternatif yg lain, kalau memang harus lapor mohon be</li> </ul>	Failed verification		

 Table 1. Review on providers

1	Unknown error 401     Fri       ★☆☆☆☆     rika asmiati       Ini kenapa yak apk nya seperti ini , data kartu tidak ada , cek saldo juga ga bisa . Tulisannya selalu unknown error ???????       Tolong diperbaiki apk nya	Pelayanan jelek         Wed           ★☆☆☆☆         nabilafarhan           Mau log-in aja ko susah banget si, berulang2         masukin sandi aja salah 🥹           Tidak bisa login lagi padahal sudah daftar         Thu           ★☆☆☆☆         Moch Abil	Cannot be accessed the application, failed task
	<b>т</b> ие	Ini bagaimana yaa kan saya sudah daftar nah saya keluar lagi pas mau masuk lagi kok gk bisa	and data not
	Cek saldo gak bisa, padahal internet kenceng buka fitur lain bisa cek saldo doang yang gak bisa. Aplikasi aneh! Hmmmmmmmmmmmmmmmmm	Buruk Tue ★☆☆☆☆ dutaaah Buruk banget pengkinian data eror terus	found

The issues expressed through app provider reviews and information complaints can result in a negative user experience. However, the root causes of these obstacles may originate from other aspects, necessitating a more thorough analysis to identify the factors impacting these obstacles. Root causes into 5 domains: user, hardware, software, data, and network [25]. It is important to establish the link between causal factors and the effectiveness, productivity, security, and user satisfaction in user interface interactions within labor social security applications.

Based on the findings from problem identification, it is evident that there is a disparity between user expectations and the actual problems encountered. Therefore, research is needed to determine how to assess the usability evaluation of workers' social security services in public service applications?

#### **B.** Literature Review

The study aims to determine the effectiveness of the application through a mixed-method research design. This approach combines quantitative and qualitative methods to provide a comprehensive evaluation of the study. The quantitative method utilizes numerical and statistical data to explain the object under study, while the qualitative method focuses on an in-depth exploration of the object to uncover user opinions and facts through interviews. In qualitative research, the approach used is thematic analysis. Thematic analysis is a method that can identify, analyze and find patterns contained in information and facts [23]. The results of the analysis are expected to reveal constraints and expectations of the application's usage.

Based on previous research and literature, the research flow evaluation of the mixed method comprises five categories with ten subject research. The study begins with a quantitative method utilizing questionnaires and test task scenarios, followed by a qualitative method using post-test interview instruments to deepen the findings. The research concludes with practical implications and recommendations for improvement in institutions and user convenience.

#### 1. Scope

The objective of the study is to assess the efficacy from the perspective of the end-users. The software has the capability to function as a source of self-help information, digital compensation and redemption service, and grievance redressal mechanism. The end-user engagement will entail numerous phases of transactions that are contingent upon several variables such as data veracity, precision of biometric authentication, network connectivity, proper navigation, and comprehensibility.

# 2. Usability Testing

Assessing the user interface through usability testing involves directing users to complete tasks within the scope of the evaluation. The purpose of this evaluation, according to Wahyuningrum [3] is to gather information about the need for a user interface design that aligns with the user's natural environment. Jacobsen (1999) and Wahyuningrum [3] classify usability testing into four types: exploratory, predictive, formative, and summative, based on the evaluation's purpose. In the meantime, based on an alternative interpretation, usability refers to the capacity to be utilized or implicitly, specifically the capability of an object to be employed [21].

Usability testing can be divided into two categories based on the examiner: analytic and empirical. Usability analytics involves inspection methods that include heuristic evaluation, cognitive walkthrough, and guidelines, among others and utilizes experts. Empirical evaluation, on the other hand, is based on user experience and includes methods such as thinking aloud, user performance tests, remote usability testing, beta test, forum test, cooperative evaluation, and coaching methods. Methods that rely on subjective user statements include questionnaires, observations, focus group discussions, and interviews [3].

Remote usability testing is a method that separates testers and users at different times and locations. A/B testing compares two types of application views, while diary studies collect qualitative data on user behavior, activities, and experiences during a specific period. The heuristic evaluation identifies usability problems in interface design based on recognized usability principles. Cognitive walkthrough explores user convenience in a system design. Questionnaires collect data that has subjective user value.

# 3. Study Literature

The examined research was categorized according to the topics discussed in order to identify similarities and discrepancies in the issues addressed by diverse techniques employed for assessing the excellence of application utilization. The Concept Of Study Literature Is Explained In Figure 1.



# **Figure 1.** Usability evaluation research group **Usability Evaluation Research in e-Government**

Several studies by Situmorang [4], Rizaldy [5], Oloan Lubs [6], Lupita Dyayu [7], Karisma Sara [8], and Nugraha[9] have been conducted on Usability Evaluation Research in E-Government. These studies aim to analyze the usability of e-government identify user problems and provide recommendations for improvements based on the findings. A questionnaire-based research method is

used to provide a comprehensive evaluation of usability problems in e-government case studies. The evaluation focuses on aspects such as learnability, efficiency, interaction, satisfaction, and errors.

In the meantime, utilizing task scenarios [4], [5],[7], [8] as a research method can yield more specific and measurable evaluation outcomes, resulting in more concrete recommendations for resolving issues identified through task scenario testing. On the other hand, utilizing interviews [5], [4], [7] as a research method can reveal more in-depth issues and capture user expectations and proposed enhancements needed by users.

Research that employs gap analysis [8], [9] can expound on the gap analysis of performance levels and levels of importance, which can be mapped through quadrants as a basis for prioritizing improvement. The six studies discussing the evaluation of e-government applications for general users with different segmentations highlight the importance of determining their usefulness, particularly in terms of efficiency, application performance, satisfaction, and security, through questionnaires and gap analysis. Task scenarios [4], [5], [7], [8] and comparing the evaluation outcomes of general users with expert users, as conducted in Situmorang [4], can be utilized to identify user problems and provide recommendations for improvement. Interviews can be carried out to deepen usability, with the expected outcomes in the form of appropriate recommendations for improvement based on usability guidelines in Situmorang [4].

# Usability Evaluation Research in Financial and Banking Services

Studies on assessing usability in financial and banking services were conducted by Habib et al. [10], Nugroho [11], and Yuwono et al. [12]. The main objective was to evaluate usability to determine user experience and requirements. The research [10], [11], [12] generally evaluates the aspects of appeal and engagement, comprehensibility, efficiency, and functionality. The user activity criteria are the same for utilizing the application, specifically for conducting financial transactions. The research [10], [12] used a questionnaire instrument from UEQ which assessed six aspects and tested task scenarios. Jak one mobile has superior value through the parameters task completed, error during task, time completed task, and number of clicks while research [10] measures success rate, time-based efficiency, and user satisfaction variables.

The study carried out by [10] is an assessment before the launch using a prototype of an application. This is done in order to apply the findings of the improvement evaluation during the launch of the application. The methodology utilized is design thinking which prioritizes customer-centricity in crafting user experience, producing novel inventions, and comprehending user requirements. In contradistinction to the research conducted by [12], which compares the evaluations of two comparable applications by delivering optimistic and pessimistic evaluations of the characteristics employed.

Initially, these four investigations used techniques to assess the efficacy by instrument task scenarios, surveys, and interviews, yet they have not correlated the issues' discoveries with the formulated suggestions. Additionally, this exploration appraises a heavily regulated financial service, however, the evaluated facets overlook user safety.

#### **Usability Evaluation Research in Self Services**

Studies on the Assessment of Self-Service Usability Situmorang [4], Lupita Dyayu [7], Yuwono [12], and Ahnan [13] delve into the evaluation of the usability of self-service applications. These applications furnish information and services that users can access autonomously, at any time and from any place. The UEQ or SUS questionnaire is employed as a tool to collect data via a survey, with the aim of gauging the user's overall experience while using the application, including factors such as ease of use, comprehensibility, and interface. The study conducted by Situmorang [4] primarily emphasizes the importance of responsiveness, application efficiency, and user interface design as the fundamental criteria for a self-service platform with regard to the m-KantorPos application.

Dissimilar to the prior investigation [12] which centered on examining the user experience of mobile banking applications based on their attractiveness, clarity, efficiency, accuracy, stimulation, and novelty. This study reinforced the comprehensiveness of the evaluation by conducting tests to gauge application performance in terms of time, errors, and clicks. This allowed for quantitative metrics to elucidate the functionality of the applications. Additionally, an interview process was conducted to capture user expectations and issues beyond the outcomes of a controlled appraisal utilizing a questionnaire and testing tool. Nevertheless, the outcomes of this intricate study have not yet furnished suggestions for enhancing the evaluations performed.

The study conducted by Ahnan [13] solely relies on questionnaires and interviews as means of gathering data. As a result, the evaluations that can be analyzed only shed light on explicit aspects of user experience. On the other hand, [7] research employs task scenario testing instruments, questionnaires, and interviews. However, the findings only provide a cursory review of the user's efficiency and fail to delve into the root causes of errors that occurred. Additionally, the recommendations presented lack a basis in user experience guidelines.

#### **Usability Evaluation Research on Insurance Services**

The research carried out by Oloan Lubs [6], Nugraha [9], and Habib [10] focused on assessing the quality of insurance services based on specific user limitations. The objective was to evaluate service excellence in terms of user comprehension, ease of use, interface design, navigation, application functionality, and service operations.

The study [10] carried out an examination of the final desired software by employing comprehensive tools such as task simulations, interviews, and two categories of surveys (SUS and UEQ). The results effectively identified recommendations for enhancement, such as refining language usage, enhancing interface design, and incorporating business process flows.

A research [6] conducted to assess the JKN mobile application among its users, assessed system, user, and interaction aspects in a restricted manner. The findings of the research indicate that all three aspects were rated as very good and good. Nonetheless, the study did not cover any details regarding usage issues, user expectations, and improvements proposed were solely based on questionnaire

measurements, with the user aspect receiving the lowest value. The study lacks specificity on the areas that require improvement.

In their study, Nugraha and colleagues [9] employed a gap analysis strategy to evaluate the anticipations of application performance across three dimensions: system, user, and interaction. They subsequently charted the significance and level of performance of each aspect to determine their respective priorities. The outcomes revealed the users' requirements and preferences, which can be utilized to streamline the development process by concentrating on specific features and services.

The synopsis of former studies indicates that the use of surveys has effectively obtained appraisals grounded on user experience. However, to apprehend issues regarding usage, application performance, and efficiency, conducting direct testing via task scenarios that illustrate user activity is necessary. Moreover, to apprehend expectations and intangible aspects that cannot be perceived through surveys and testing, an exhaustive examination must be performed through user interviews.

# 4. Theoretical Framework

ISO adalah [18]. The Application Quality Standard by ISO 9126-4 [19] to assess the Quality used has four factors that measure the quality of application use, namely effectiveness, productivity, safety, and satisfaction. The efficiency and productivity factors as measuring tools in conducting evaluations use measuring tools through task test scenarios, while the safety and satisfaction factors use user experience questionnaires and post-test interview tools. Standards for evaluation, tools, and user feedback data serve as the means to assess usability. The results of the evaluation findings are presented in the form of usability metrics, problem detection, user satisfaction evaluation, and user anticipation.



# Figure 2. Design Thinking

# C. Research Methods

The investigation commences at the initial phase by means of scrutinizing the data and constructing the context of the issue in order to arrive at deductions from

the findings of the research analysis and the creation of reports. Each step and procedure in the progression of the investigation is elucidated below:

The initial step of scrutinizing information and issues is accomplished by performing discussions with the application administration group, executing desktop exploration on auxiliary information utilizing fundamental statistical techniques derived from grievances of application problems, and consumer feedback on application suppliers. The stage of reviewing the library involved an analysis of the literature by examining theories of usability evaluation and literature reviews of past research. The research was classified into four categories: egovernment applications, financial service applications, self-service applications, and insurance applications.

The exploration of literature was conducted by utilizing key phrases present in the research inquiries, pertinent concepts, and previously conducted research. The 3C2S approach (evaluate, differentiate, disapprove, amalgamate, condense) was employed to analyze the data. This phase resulted in the establishment of a theoretical foundation and research methodology, which will be illustrated in a framework to establish connections between variables in the study.

Utilizing the framework established during the literature review phase, which involves aligning various reference standards for usability evaluation, the optimal research instrument has been selected to achieve the greatest outcomes. Specifically, this study will adopt a sequential mixed method approach with a sequential transformative strategy, wherein a theoretical perspective is employed to determine specific procedures by selecting one of the methods in the initial stage [15]. The data collection and analysis phase of this research will prioritize quantitative analysis, followed by qualitative analysis.

This method of testing usability is conducted using a combination of quantitative and qualitative approaches to gathering comprehensive and detailed evaluation outcomes. The UEQ and Test Task Scenario can elucidate the application's usability through user experience and assessment during testing. Simultaneously, interviews can reveal the perceived limitations and expectations of application users.

The UEQ questionnaire [20] has six aspects which include efficiency, attractiveness, perspicuity, dependability, stimulation, and novelty, and consists of 26 question items.

The phases involved in carrying out the study encompass conducting research, primarily undertaking trials on the questionnaire tool to evaluate its accuracy and consistency. Following this, the data from the questionnaire, which has been completed by users who meet the criteria established by previous research instruments, is collected. The objective is to obtain responses from 100 participants at this stage. Once the data collection goal has been attained, the subsequent step is to scrutinize the questionnaire data using UEQ's data analysis tools, specifically version 10.

The number of participants will be determined using the Slovin method, considering the total population of 13 million users, in order to reach a sample size of 100 participants. Respondents will be selected based on their similarity to reviewers through service providers. Out of the 108 respondents, 84 are Android users, while the remaining 24 are not. This is due to the fact that Android users tend

to provide more reviews and there are inconsistencies between ratings and review outcomes. A portion of the respondents will represent users of iOS, who have fewer users and give lower ratings. These respondents will be selected based on their level of experience with the application.

The next phase of the numerical approach involves administering assessments utilizing evaluation exercises that have been formulated in the investigative tool to focus on eight participants. The findings of the assessments will be evaluated by applying the commonly accepted ISO 9126-4 quality-in-use formula for quality in terms of efficiency and efficacy. ISO 9126-4 [19] is used as a standard for assessing task scenarios carried out through aspects and variables that are assessed in quality in use in an application. Meanwhile, the benchmark taken through the Research and Development of the Ministry of Home Affairs is used as a reference for assessing effectiveness through the success of the task which is categorized as very ineffective to very effective. This assessment reference is referenced through research conducted by Desiana [14].

The third phase involves a qualitative approach, wherein a post-test interview is conducted to evaluate components of the UEQ survey and ISO 9126-4 quality inuse standards. The objective is to obtain detailed insights and data pertaining to the utilization of service applications, limitations, and user expectations.

The phases of analysis of the outcomes execute analyses on numeric and nonnumeric research by harmonizing user assessments with issue discoveries and user anticipations. The analysis outcomes will be utilized in composing notional and pragmatic consequences and suggestions for enhancement. The aim of this section is to furnish conclusions and recommendations for enhancing future research endeavors. This can be achieved by utilizing diverse methodologies, approaches, and tools.

#### D. Result and Discussion

The study outcomes were obtained through the application of the usability assessment technique, employing a quantitative methodology that involved questionnaires and task-based scenarios, complemented by a qualitative technique utilizing post-test interviews. These were communicated via an interview session.

The time frame for gathering user feedback from the Google Play Store and App Store was between January and April 2022. The data collection for the UEQ questionnaire took place from October to November 2022, resulting in a five-month disparity between the secondary data collection and questionnaire distribution.

In October 2022, when the distribution of the questionnaire, task scenario testing, and post-test interviews were conducted, the user ratings on Google Play remained at 4.5, and the ratings on the App Store remained at 1.6. This indicates that the review conditions and user experience were consistent. However, the users' reviews still displayed inconsistencies between the rating and the review description, despite the five-point rating system. The review descriptions revealed that users faced usage problems.

The findings of the examination that took into account the user characteristics, specifically gender, indicate that out of the 108 participants, the larger proportion was female, comprising of 66 participants (61.1%), whereas the male participants accounted for 42 respondents (38.9%). According to the age classification of the 108 participants, it is evident that the bulk of the respondents, comprising 74 individuals

(68.5%), fall within the age range of 20-30 years. On the other hand, the group of respondents aged between 41-50 years is the smallest, with only 7 individuals (6.5%). According to the educational background, it was revealed that among 108 participants, most of them possessed a bachelor's degree, which accounted for 65 participants (65.2%). On the other hand, a minimum of 9 participants (8.3%) had a master's degree. According to the duration of employment, it was discovered that among the 108 participants, most of them had a work experience of 1-5 years, which accounted for 61 respondents (56.5%). On the other hand, 13 respondents (12%) had an experience of more than 6 years. According to the level of usage, it is evident that among the 108 participants, most of them, that is 59 participants (54.6%), utilizing the application once every month, whereas almost on a daily basis, around 10 participants (9.3%) used it. According to the user's OS, it is evident that among 108 participants, most of them prefer the application on Android OS, which constitutes 84 respondents (77.8%). Conversely, the iOS OS has 24 respondents (22.2%).

#### 1. Questionnaire Analysis

Data from the UEQ questionnaire were gathered from 108 individuals who utilized the social security service application for employees. Next, the data was evaluated using the UEQ data tools application. The responses were sorted into negative values (1-3), neutral values (0), and positive values (4-7). This categorization helped to clarify the distribution of responses provided by the participants. The overall percentage of respondent answers revealed that 14.4% of the responses were negative, 15.6% were neutral, and 70% were positive.



#### Distribution of Answers per Item

Figure 3. Conclusion of the UEQ aspect variable assessment

The findings of the assessment of the six components of the UEQ clarify that 80% of the aspect variables (5 out of 6) account for the favorable user experience, while one aspect variable has a neutral impact. The items in the survey that received notably low ratings (over 20% negative) are traditional, sluggish, of limited utility, challenging to grasp, and monotonous.

Table 2. Conclusion of the UEQ aspect variable assessment					
Variabl	e <i>Mea</i>	n Assessment Evalu	lation		
Attractive	nes 1,4	Positive			
S					
Perspicuit	y 1,3	Positive			
Efficiency	1,3	Positive			
Dependab	ility 1,2	Positive			
Stimulatio	n 1,1	Positive			
Novelty	0,6	Neutral			

Once the mean value for each variable is calculated, a comparison is made with the benchmark dataset to assess the application's relative quality compared to other products. The UEQ data analysis tools were utilized to conduct the benchmark process, and the results are presented below.



Figure 4. The result of the UEQ benchmark

According to the data presented in Figure 4, it is evident that applications receive higher than average ratings in terms of their attractiveness, clarity, effectiveness, reliability, and engagement, but they score lower than average in terms of novelty.

# 2. Test Task Scenario Analysis

Table 3. Overview of the result test task scenarios				
Scenario	Feature	Avg time (second)		
Pogistor as a new participant (S1)	List of new participants	25%	876	
Register as a new participant (S1)	Dues payment	25%	74	

#### **T-11-0** 0 of the regult test test cooperi

Scenario	Feature	Success	Avg time (second)
Registration as a new user (S2)	New user list	88%	556
	Application login	81%	59
Enter the application and find the latest	View program information	100%	15
information (S3)	see the news	100%	8
	see promos	100%	7
	Add Participant Card (KPJ)	100%	80
View membership information (digital	See JHT balance	100%	18
card, balance and adding a new card) (S4)	View RSJHT	100%	22
	View digital card	100%	26
Make a complaint (T5)	Make a Complaint	38%	28
Changing amail and paceword (S6)	Change <i>email/cellphone</i>	44%	74
changing email and password (50)	Change password	38%	65
Undate data and claims (\$7)	Data update	44%	270
opuate data and claims (37)	Claim	31%	426
Changing profile photo and exiting	Logout	100%	13
application (S8)	Changing Profile Photo	50%	31

Test results for each task explain that the average success rate over all users is 70%. The success rate percentage of users can be calculated by adding the successful tasks and half of the partially successful tasks, then dividing the result by the total number of tasks and multiplying it by the number of participants [24]. The final value should be multiplied by 100 to obtain the percentage.

Success rate = $(92+(18 \times 0,5))/144 \times 100\% = 70.14\%$ .

The user has achieved a success rate of 70.14% in the tested task scenario. As per the research conducted by [14], following the reference standards of the Ministry of Home Affairs Research and Development in 1991, it can be inferred that the user's accomplishment in completing the task scenario falls under the fairly effective category, which ranges from 60% to 79.99%.

The tasks performed by users yield quite impressive scores, but a thorough discussion is required to evaluate the categorization of successful usage. The success of tasks executed on straightforward and uncomplicated activities, such as accessing informational features, is evident. However, when it comes to tasks with low scores, they are mainly transactional in nature. These include registering new participants, making complaints, changing emails and passwords, updating data, and making claims, as well as changing profile photos. Therefore, it can be concluded that the testing of viewing tasks runs smoothly, while the testing of transactional tasks still encounters issues, with an average value below 37%.

Upon examining the frequency of errors committed by each user, it becomes apparent that there are both failed and successful tasks. A portion of these errors led to tasks not being completed, resulting in a total of 24 unfinished tasks across all users. This test indicates that the error rate remains relatively high, particularly in regard to transactional tasks. Meanwhile, users experience an average error rate of 17% throughout the task, with P3 users reporting the highest inaccuracy of 28%.

Table 4. Error rate

Participant	Fail	Partially success	Amount	Error Rate
P1	2	2	4	22%
P2	1	2	3	17%
P3	1	4	5	28%
P4	0	3	3	17%
P5	1	3	4	22%
P6	1	0	1	6%
P7	0	2	2	11%
P8	0	2	2	11%
Aver	age		17%	

The outcomes of user evaluation concerning interaction duration, as presented in Table 4, reveal that among 8 users, the mean overall time taken is 1,454 seconds (equivalent to 24.2 minutes), and the mean time taken to accomplish one task is 92.88 seconds (1.5 minutes). In instances where the tasks failed, specifically tasks 4, 9, and 10, the users encountered an average failure time of 11 seconds and discontinued their attempts after 19 seconds upon receiving an error message from the application. Time-based efficiency (Total time) / (Total number of tasks x Number of participants).

nij = Total time

N = Number of tasks

R = Number of participants

Pt = 5,86/144

Pt = 0,04 goals/seconds

The efficiency calculation based on time is 0.04 objectives per second or can be understood as taking 24.54 seconds to accomplish a single task. Understanding the time intervals based on the conduct of time [14] required to deal with task situations is recognized to be extremely rapid, falling within the 60 – 300 seconds range.

The evaluation of the duration necessary to finish the entire assignment given to 8 participants was comparatively swift, but through further deliberation, the pace of the typical assignment processing time also encompassed both varieties of assignments, namely transactional and viewing. Consequently, access to viewing features merely entails the action of observing and locating the required information by the user, without necessitating any interaction with the interface. On transactional pages, users are mandated to input comprehensive data and adhere to the navigation of the application interface, thereby resulting in constraints and impediments that lead to a rather high assessment of the completion rate and error rate for transactional task categories. With regards to the time required, users who are unable to finalize the task will cease due to an error and subsequently become a partially successful task.

Overall relative efficiency=(Total time)/(Total task x Number of participants) Pt = 13.308/13.374

Pt = 99,51%

The computation of the total comparative effectiveness yields a temporal proportion of 99.51% among individuals who satisfactorily accomplish assignments.

Table 4. Problem finding

Problem	С	Ι	F	Р
Participant's card number is unknown	4	4	1	9
Biometric verification is blocked because it has taken	3	4	2	8,5
three pictures				
Filling in the mobile number verification code gets an	3	3	2	7,5
error message that the request was rejected without				
giving a reason and solution				

Information:

(C) Critical level is categorized as (1) Low, (2) Moderate, (3) High, (4) Very high.

(I) The Impact Score is categorized as 1 – suggestions from users, 2 – minor problems have little impact, 3 - major problems have frustrating effects and delays, 4 - obstacle blockers prevent users from completing tasks.

(F) The score for the frequency of occurrence of problems is categorized based on the percentage at 1 - below 25%, 2 between 25% - 50%, 3 between 50% - 75% and 4 above 75%.

As per the results obtained from the task scenario analysis, it is evident that there exist a number of impediments that need to be addressed on a priority basis for enhancing the application development process. Some of the major challenges identified include unidentified card numbers of participants (9), biometric verification getting blocked after three attempts (8.5), and encountering an error message while entering the cellphone number verification code, without any rationale or resolution explanation (7.5).

#### 3. Post-Test Interview Analysis

Individuals utilize the application for various purposes, including professional obligations, interest in membership details, and checking the account balance. These motives are associated with the informative capabilities offered by the application. Users hold diverse viewpoints when it comes to utilizing the application. The general consensus is that the application is beneficial, visually appealing, and user-friendly. However, users have varying opinions regarding the functionality of the application. Some find it challenging to navigate, while others find it effortless or encounter some obstacles. Therefore, it can be inferred that the application is relatively easy to learn, but there are still some limitations to its usage.

Certain challenges faced by users include discrepancies in displayed balances, complexities during registration and login procedures, limitations in biometric verification, and restraints when submitting claims. As per user feedback, some of these issues can be resolved independently, while others necessitate assistance from external sources such as visiting a branch office or reaching out to the contact center.

The application users have reported that their personal data is well-protected during the process of providing it, as the application is under the control of a governmental institution. As per the feedback from users, the security measures implemented by the application policies are quite stringent, which at times, can make it challenging for them to complete the process, leading to a few instances of failure. According to users, the application's level of responsiveness is inadequate, moderately satisfactory, and has room for further enhancement. In summary, the responsiveness component of the application requires improvement. The aspiration of users in the development of applications is to offer an interface that is not dull, ensure a prompt response, simplify the process of data authentication, and enable accessibility for users with disabilities.

The display of the outcomes arising from the consequences of the qualitative investigation indicates that there exist hindrances that users still encounter and are unable to overcome on their own, necessitating them to seek assistance from officers at the branch office or contact center. This highlights the significance of application development by giving attention to the findings of the limitations and user feedback to enhance the entire application concerning providing accurate information, furnishing more informative feedback on errors or failures that do not confuse the user and impede problem resolution, presenting a clear and userfriendly interface, providing easy navigation to frequently used features, furnishing user-friendly supportive features such as information on membership types and program benefits before becoming an application user, complaint services for nonusers to solve registration and login problems, and supportive features for people with disabilities.

# 4. Recommendation

On the user initiation page, the issue lies in the fact that users are unable to comprehend the vocabulary presented by the software. Hence, a well-designed user initiation page must be presented prior to the user's registration. Potential resolutions that can be executed include creating a landing page for non-registered users, and offering additional clarification about the software and the advantages it offers.

On the sign-in page, it is common for users to overlook their password due to the lengthy usage period of over 3 months. Users may find it challenging to comply with the password creation guidelines that require capital letters, lowercase letters, numbers, and special characters. To address these issues, the following enhancements are recommended: providing the option to sign in with a Google account or Apple ID to simplify the process and minimize the likelihood of forgetting login credentials; introducing a user-friendly help button on the help page that includes a detailed description of the problem in case of account or password retrieval issues; and incorporating a language translation feature.

On the sign-up page, the user is unclear about the distinction between registering as a new participant or a new user. They are also unsure about which type of membership to select. Additionally, users are unaware of the advantages offered by each social security program. To address this issue, it is recommended to provide more detailed instructions on the registration process and information about the different programs and their benefits. Clear guidelines should be provided to differentiate between registering as a new participant or a new user. Additional navigation buttons should be included, leading to program information pages and warranty benefits. Clear examples or information should be provided to assist with membership segmentation selection. Finally, a help center navigation option should be available to users who require assistance with using the application.

On the help center webpage, the user encounters an issue that cannot be rectified during login and registration for new users. Users seek brief and effective solutions to familiar difficulties. Users encounter difficulties in determining the status of their grievances. Recommendations for practical actions include introducing a help button that directs non-users to the support page, which incorporates the contact center's phone function, live chat, and complaint monitoring. Moreover, to assist users in resolving their concerns, the frequently asked questions (FAQ) function may be included.

When users access the homepage, they primarily seek superior features, particularly during their initial visit. They are unaware of their membership details and account balance. They also lack comprehension of the terms and conditions of each program. To address these issues, it is recommended to add navigation buttons to the menu for the most commonly used functions, namely viewing balances, claiming program benefits and tracking claims, accessing digital cards, and obtaining program information. Included a search bar in the service menu. Displays the complete balance held. Distinguish between active and inactive memberships with unique indicators. Incorporated navigation buttons to clarify program details and requirements, such as comprehension, advantages, contribution payments, and claim submissions.

The user encounters difficulties on the page for updating data and claims, which include inputting incorrect data, uncertainty about the relevant membership to be filled in, inability to locate the filled-in data, and being blocked due to verification failure. To address these issues, it is recommended to include data updates in the list flow to encourage users to complete them after the listing process. Additionally, an onboarding page can be added to explain the advantages and purposes of updating data, while guidance can be provided on updating data during biometric verification procedures. The reasons for biometric failures can be explained to users, and a help page button can be included for requesting a reset.

# 5. Implication

Inconsistency between the Android and iOS operating systems can cause confusion for users who switch devices, leading to the need for adjustment. Additionally, difficulty in filling out data can make the application challenging to use. Users expect ease and speed when utilizing the application, which should be capable of replacing physical service processes with digital services to minimize user effort in accessing social security services. Security concerns are especially important for applications that provide financial transaction services. Users trust the provider organization to protect their data and verify the information provided.

By implementing recommendations for application improvements, service quality can be improved to meet user needs and expectations. This can lead to an increase in the number of services offered, optimization of existing resources, the shift of operational burdens to digital services, a reduction in complaints, and an improved reputation for the organization.

# E. Conclusion And Limitation

The investigation effectively generated suggestions for enhancing the user experience on service applications by analyzing results obtained from UEQ,

interviews, and statistical tests. The recommended domains for correction and enhancement of user experience pertain to novelty as it still falls below the average benchmark. The applications necessitate development based on usability assessment, which is indicated by query items reflecting the application's sluggishness, complexity in learning, and monotonousness.

According to the assessment carried out on 8 participants, it has been established that the mean achievement rate for the task under scrutiny is 70.14%. As per the research and development benchmarks of the Research and Development of the Ministry of Home Affairs from 1991 [14] it can be inferred that the users' accomplishment in executing the task falls within the moderately efficient bracket, which spans from 60% to 79.99%.

Derived from the outcomes of user interviews, it is evident that users exhibit diverse reactions contingent on their individual experiences. Users have justifications for utilizing the services, primarily to acquire information regarding their labor union membership status and social security participation, such as accessing details about their benefits, balance amount, and reported wages. An additional intriguing requirement for users is the availability of digital services that streamline the process of filing claims.

According to the users, the application is an intriguing tool that proves to be beneficial. However, there are a few who contend that the application is not userfriendly due to inadequate comprehension, errors while entering data, and complexity in operating it. The limitations faced by users are predominantly related to logging in, biometric verification, and accessing membership details such as digital cards, balances, and wages. When encountering difficulties, users feel that the responses provided by the application are inconsistent. Some are lucid, while others are ambiguous about what the next step should be, resulting in prolonged response times. Users desire that the application is improved by focusing on user-friendliness, taking into account their needs and preferences, and addressing issues that surface through feedback channels and user reviews.

The research investigates usability concerns related to user experience during interface interactions. However, this study does not cover usability issues related to application and database services, which prevents us from identifying the root causes of problems and their connection to usability problems during user interface interactions.

The research is specifically focused on evaluating applications on two platforms, Android and iOS, by gathering data through questionnaires, task scenarios, and posttest interviews. Failure to differentiate between the two platforms may introduce evaluation bias.

The research mainly evaluates existing users who are participants in the workers' social security program, which means that some user experiences are already familiar to them. External factors such as network conditions, hardware, lighting conditions that impact photo-taking, and temperature conditions that affect interaction with the touchscreen display interface cannot be controlled. This study has not taken into account the aforementioned external factors in its evaluation.

#### F. Acknowledgment

The researcher is grateful to the University of Indonesia and the Supervisor for their encouragement and support throughout the research process.

# G. References

- [1] Islam, R., & Mazumder, T. (2010). Mobile application and its global impact. *International Journal of Engineering & ..., 06*, 72–78.
- [2] Turban, E., King, D., Lee, J. K., Liang, T.-P., & Turban, D. (2012). Electronic Commerce: a Managerial and social Networks Perspective 2012. In *Commerce A Managerial Perspective (5th ed.*
- [3] Wahyuningrum, T. (2021). Buku Referensi Mengukur Usability Perangkat Lunak.
- [4] Situmorang, T. K., Muslimah Az-Zahra, H., & Herlambang, A. D. (2019). *Evaluasi Usability Pada Aplikasi m-KantorPos dengan Menggunakan Metode Usability Testing* (Vol. 3, Issue 5). http://j-ptiik.ub.ac.id
- [5] Rizaldy, A. A., Nathaniel, V., Rifa'i, R. D., & Wardhana, A. C. (n.d.). Evaluasi Aplikasi Kai Access Menggunakan Metode Moderated Usability Testing. In *Jurnal ICTEE* (Vol. 3, Issue 2).
- [6] Oloan Lubs, B., Salim, A., Kamal Raya No, J., & Barat, R. (n.d.). *EVALUASI* USABILITY SISTEM APLIKASI MOBILE JKN MENGGUNAKAN USE QUESTIONNAIRE.
- [7] Lupita Dyayu, A., & Yani, H. (2023). Evaluasi Usability Aplikasi PeduliLindungi Menggunakan Metode Usability Testing dan System Usability Scale (SUS). Jurnal Manajemen Teknologi Dan Sistem Informasi (JMS), 3(1). http://ejournal.unama.ac.id/index.php/jms
- [8] Karisma Sara, F. D., Soedijono, B. W., & Rudyanto Arief, M. (2019). Evaluasi Website Pemerintah Kabupaten Rembang Dengan Metode Usability Testing. *Jurnal Teknologi Informasi*, *3*. https://rembangkab.go.id/.
- [9] Nugraha, R., Mazia, L., & Asri Utami, L. (2022). ANALISA KEGUNAAN DAN KEMUDAHAN LAYANAN MOBILE JKN PADA UPTD PUSKESMAS DEPOK DENGAN USE QUESTIONNAIRE DAN IPA (Vol. 16, Issue 2). https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index
- [10] Habib, M., Chatib, F., Tolle, H., & Fanani, L. (2022). Perancangan User experience Aplikasi Mobile Fitur Asuransi Jago Last Wish menggunakan Pendekatan Design Thinking (Studi Kasus Magang di PT. Impact Byte Teknologi Edukasi Challenge Partner Bank Jago) (Vol. 6, Issue 7). http://j-ptiik.ub.ac.id
- [11] Nugroho, W., & Disetujui, D. D. (2022). Evaluasi Kualitas Digital Payment OVO Berdasarkan Faktor Usability Standar ISO/IEC 9126. *Journal Computer Science*, 1(1).
- [12] Yuwono, R., Wibowo, A., Hadi Wijoyo, S., & Rokhmawati, R. I. (2019). Analisis Pengalaman Pengguna Pada Aplikasi Mobile Banking di Indonesia Dengan Menggunakan Usability Testing dan User Experience Questionnaire (UEQ) (Studi pada JakOne Mobile dan BCA Mobile) (Vol. 3, Issue 6). http://j-ptiik.ub.ac.id
- [13] Ahnan, M., Prasetya Wibawa, A., Ari, I., & Zaeni, E. (n.d.). Evaluasi User Experience pada Pengembangan Aplikasi Sistem ADI STETSA SMAN 4 Malang menggunakan User Experience Questionnaire. Jurnal Inovasi Teknik Dan

Edukasi	Teknologi,	1(7),	518-525.
https://doi.org	/10.17977/um068v1i720	21p518-525	

- [14] Desiana, P., & Ayu, W. (2017). Analisis Pengukuran Tingkat Efektivitas dan Efisiensi Sistem Informasi Manajemen Surat STIKOM Bali. *Jurnal Sistem Dan Informatika*, *11*(2).
- [15] Creswell, J.W. (2003). Chapter One, A Framework for Design. Research design: Qualitative, quantitative, and mixed methods approaches. Thousand Oaks, CA: Sage Publications.
- [16] Apple. JMO Jamsostek Mobile. Apple Store. Retrieved July 24, 2023, from https://apps.apple.com/id/app/jmo-jamsostek-mobile/id1444834757
- [17] Google. JMO (Jamsostek Mobile). Google Play Store. Retrieved July 24, 2023,from https://play.google.com/store/apps/details?id=com.bpjstku
- [18] International Organisation for Standardisation. (2018). Ergonomics of humansystem interaction - Part 11: Usability: Definitions and Concepts ISO 9241-11:2018. ISO 9241-11:2018(E), November.
- [19] ISO 9126. (2001). ISO / IEC 9126-4 Software engineering Product quality Part 4: Quality in use metrics. *Quality*, 1(514).
- [20] Kushendriawan, M. A., Santoso, H. B., Putra, P. O. H., & Schrepp, M. (2021). Evaluating User Experience of a Mobile Health Application Halodoc using User Experience Questionnaire and Usability Testing. In *Journal of Information System*) (Vol. 17, Issue 1).
- [21] Bevan, N., Carter, J., & Harker, S. (2015). Iso 9241-11 revised: What have we learnt about usability since 1998? Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9169. https://doi.org/10.1007/978-3-319-20901-2\_13
- [22] Ganapati, S. (2015). Using Mobile Apps in Government. *Washington DC: IBM Center for The Business of Government*.
- [23] Braun, V., & Clarke, V. (2022). Thematic Analysis: A Practical Guide. QMiP Bulletin, 1(33). https://doi.org/10.53841/bpsqmip.2022.1.33.46
- [24] Sabandar, V. P., & Santoso, H. B. (2018). Evaluasi Aplikasi Media Pembelajaran Statistika Dasar Menggunakan Metode Usability Testing. *50 TEKNIKA*, 7(1).
- [25] James A. O'Brien, & George M. Marakas. (2011). Management Information System Tenth Edition. In *Mc.Graw-Hill Companies*.