Software Quality Analysis for Halodoc Application using ISO 25010:2011

Aditia Arga Pratama[1]

Faculty of Information System Management
Gunadarma Unversity, Depok, Indonesia

Achmad Benny Mutiara[2]

Faculty of Computer Science and Information Technology
Gunadarma University, Depok, Indonesia

Abstract—**The rapid spread of the Covid-19 disease and the ongoing social distancing have disrupted community activities. This consequently makes people use information and communication technology, especially in the fields of health, education, and business. The use of information and communication technology in the health sector plays an important role in preventing the spread of the Covid-19 disease. One of its uses is Telemedicine. One of the developing telemedicine in Indonesia is Halodoc. This study will test the quality of the Halodoc application to determine the quality of the Halodoc application using the model of the International Organization for Standardization (ISO), namely ISO 25010:2011. Software quality assurance can use this model as the basis for testing. Therefore the test will include 8 characteristics and 29 sub-characteristics. The assessment in this tests will use a weight where this weight is determined using Analytical Hierarchy Process (AHP) method. The testing method is carried out using Black-Box testing, stress testing, and distributing questionnaires to 100 respondents for Usability testing. Functional Suitability, Compatibility, Reliability, and Maintainability got a maximum score of 5. Performance Efficiency get a score of 4,886, Usability gets a score of 4, Security gets a score of 3,549, Portability gets a score of 3,718. The total results of the Halodoc application assessment get a score of 4,515 out of a maximum score of 5.**

Keywords—*Halodoc; ISO 25010:2011; software quality assurance; telemedicine*

I. INTRODUCTION

In 2020, the disease Covid-19 made the world difficult. The spread of this disease is so fast that the governments of each country are implementing isolation and social distancing. Indonesia is one of them. The Indonesian government has implemented social distancing to reduce the spread of the Covid-19 disease. The rapid spread of disease and social distancing rules that are still in effect disrupt community activities. Activities such as teaching and learning, working, doing business, and others cannot be carried out as usual. This effect makes people use information and communication technology, especially in the fields of health, education, and business [1]. The use of information and communication technology in the health sector plays an important role in preventing the spread of the Covid-19 disease. One of its uses is Telemedicine. Telemedicine is the service of remote health care services by all health care professionals using information and communication technologies. Another benefit of

telemedicine is that it can prevent the spread of Covid-19 in Indonesia.[2]. One of the developing telemedicine in Indonesia is Halodoc.

Halodoc provides features that make it easier to consult with doctors, including about the Covid-19 disease that many people suffer from. This feature provides convenience because patients can consult a doctor anytime and anywhere, doctors available for consultation are available from general practitioners to specialists, patients can ask doctors about their disease and symptoms, and will be given treatment prescriptions according to the patient's illness. Perceived usefulness, perceived behavioral control, trust, self-health awareness, system quality, and attitude are factors that influence users in using Halodoc [3]. Therefore if the Halodoc application wants to make users satisfied using the application, it must pay attention to these factors. The quality of the system or software factor can be assessed by testing the quality of the software. The quality of the system or software can be assessed by testing the quality of the software so that users remain satisfied.

There are various methods or models of software quality testing. McCall model, Boehm model, or model of the International Organization for Standardization (ISO) like 9126-1 and 25010:2011 model, and many other models. The ISO 25010:2011 model includes all the quality factors of the previous model like the McCall model, Boehm model, and 9126-1 model [4]. Therefore, this model can be used as a basis for testing software quality.

This journal will use the ISO 25010:2011 model as the basis for testing the quality of the Halodoc application. Furthermore, the test will use eight characteristics of the model. The total assessment of the quality of the Halodoc application will be calculated using the weight determined in the research using the Analytical Hierarchy Process (AHP) method multiplied by the score from each characteristic. In addition, this journal will also use all sub-characteristics except Modularity and Modifiability of the Maintainability characteristic in the test. The supporting software that will be used in this research is "TestProject.io" which is used for Time Behavior testing, "UI/Application Exerciser Monkey" for Maturity testing, "Apache JMeter" for Availability testing, and "SPSS" for testing Validity and Reliability of the questionnaire.

This journal will organize as follows. First, this journal will provide a brief overview of related work. After that, this journal will explain the research methods that will be used. Next, it will present the research result and discuss it. Finally, this journal will give a conclusion about this journal.

II. RELATED RESEARCH WORK

Previous research related to the topic of analyzing the quality of mobile applications already exists. First, there is a journal that uses the ISO 9126-1 model to measure electronic driving license application quality. The tool used for testing is a questionnaire, and the results from the journal are that the application meets the standard model 9126-1[5]. Nevertheless, this model is an old model. There is already a new model that develops from this model that is ISO 25010:2011 model [6].

Another research conducted a software quality analysis using the McCall model [7]. This model uses three characteristics, namely Product Revision, Product Transition, and Product Operation. This study aims to measure the quality of the financial system for cashiers. The results show that the performance of the system can still be improved. But the drawback of McCall's model is that the functionality of the application does not consider the functionality of the software [4].

Testing of quality analysis research uses the characteristics of the ISO 25010:2011 model. Those characteristics are Compatibility and Performance Efficiency. The test results from this study say that the Elektronika Analog application has a satisfactory quality. This research conducts the test using software cloud.bitbar.com. The test in this study has a weakness in which the Co-existence sub-characteristics are not tested [8].

Another research uses the ISO 25010:2011 model to ensure the quality of school information management applications. The test will use the characteristics like Functional Suitability, Usability, Performance Efficiency, Portability. Tests are carried out using tools such as gtmetrix for Performance Efficiency characteristics, System Usability Scale (SUS) for Usability characteristics, and Black-Box testing for Functional testing [9].

Other research uses the AHP method in calculating the final results of the test. The model used for testing is ISO 25010:2011. This research uses the characteristics like Functional Suitability, Reliability, Usability, and Performance Efficiency. The test method uses a questionnaire for all characteristics. The final result is calculated from the questionnaire data, which is then entered into the AHP. After that, the weights of the characteristics are compared to see which characteristics and sub-characteristics need improvement [10].

The difference between this journal and previous research is that it will use eight characteristics and all of their sub-characteristics except Modularity and Modifiability. Testing will also be carried out by determining the weight of each characteristic and sub-characteristic. The weight will be deterimined using AHP method. This weight is then multiplied by the value of each test performed.

III. RESEARCH METHODS

A. Software Quality Assurance

A set of activities that define and assess the adequacy of the software process to provide evidence that establishes confidence that the software process is appropriate and produces a software product of the appropriate quality for the intended process [11]. Software Quality Assurance (SQA) software is a fundamental activity for many businesses that produce products that will be used by both internal and external users of the company. Software assurance has several purposes such as ensure that the team has carefully reviewed the requirements model for high quality, The design model must be assessed by the software team to ensure the high-quality model and meets the requirements, source code and associated work products must conform to local coding standards and exhibit characteristics that will facilitate software maintenance, SQA analyzes the allocation of resources for review and testing to assess whether or not resources are allocated in the most effective manner [12].

B. Characteristics and Sub-characteristics of ISO 25010 Quality Model

This ISO 25010 quality model consists of eight characteristics that are related to the dynamic and static properties of a computer system. The characteristics and sub-characteristics provide a consistent term for defining, measuring, and evaluating the quality of systems and software [13]. Fig. 1 shows ISO 25010:2011 diagram.
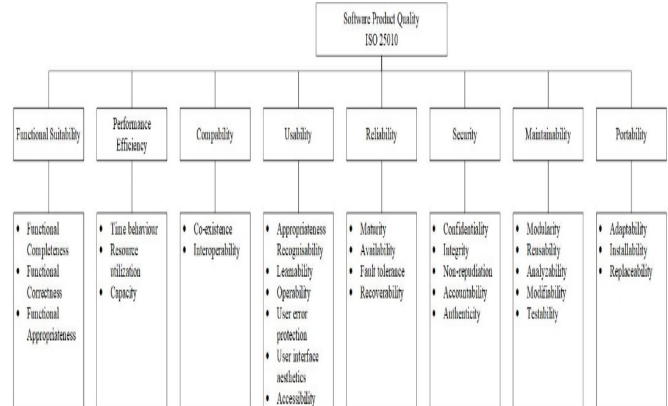

Fig. 1.   ISO 25010:2011 Model.

Eight characteristics of ISO 25010 is Functional Suitability which represents the extent to which a product or system provides a function that satisfies stated needs when used under certain conditions. This characteristic has three sub-characteristics, namely: Functional Completeness, Functional Correctness, and Functional Appropriateness [14]. Performance Efficiency represents performance relative to the resources used under the stated conditions. This characteristic consists of sub-characteristics, namely: Time Behavior, Resource Utilization, Capacity [14]. Compatibility represents the extent of the system or product can exchange information or perform required functions with other system or product while sharing the same environment. This characteristic consists of sub-characteristics, namely: Co-existence and

Interoperability [14]. Usability represents if a system or product can be used by certain users to achieve certain goals with effectiveness, efficiency, and satisfaction in a particular context of use. This characteristic consists of sub-characteristics, namely: Appropriateness Recognizability, Learnability, Operability, User Error Protection, User Interface Aesthetics, and Accessibility [14]. Reliability represents how well a system, product or component performs a particular function under certain conditions for a certain period of time. This characteristic consists of sub-characteristics, namely: Maturity, Availability, Fault Tolerance, and Recoverability [14]. Security is the characteristic to know the extent a system or product protects information and data with a level data access following type and level of authorization. This characteristic consists of sub-characteristics, namely: Confidentiality, Integrity, Non-repudiation, Accountability, and Authenticity [14]. Maintainability is used to know the level of effectiveness and efficiency with which a product or system can be modified to improve it, improve it or adapt it to changing environments, and requirements. This characteristic consists of sub-characteristics, namely: Modularity, Reusability, Analyzability, Modifiability, and Testability. In this research, the sub-characteristic that will be used only 3, namely: Reusability, Analyzability, and Testability [14]. Portability will be used to determine the degree of effectiveness and efficiency with which a system or product can be transferred from one hardware, software, or other operational or usage environment to another. This characteristic consists of sub-characteristics, namely: Adaptability, Installability, and Replaceability [14].

C. Black-box Testing

Black-Box Testing is testing based on requirements specification and there is no need to check the code in Black-Box testing. This is purely done from a user point of view, only the tester knows the predictable set of inputs and outputs [15].

Black-Box testing plays an important role in software testing; it helps in validating the overall functionality of the system. Black-Box testing is done based on customer requirements-so incomplete or unpredictable requirements can be easily identified and can be addressed later [16]. The main advantages of Black-Box testing are required relatively few resources compared to White-Box testing, resource effectiveness can be done by testing automatically hence contributing to a shorter testing period, and ability to perform almost all test case groups, such as availability (response time) reliability, load durability and test groups related to the operation, revision and transition factors [17].

D. Stress Testing

Stress testing is used to test the reliability and stability of the system. This test can determine the system's resilience and fault handling under very heavy load conditions. Tests outside the normal operating point and evaluating how the system performs under these extreme conditions are included in Stress testing. Stress Testing is carried out to ensure that the system will not crash under crisis situations. The existence of large amounts of data with very high frequency and volume will result in the performance of the software being disrupted or not [18].

E. Questionnaire

A questionnaire is a data collection instrument used to collect large amounts of data. The trick is to provide several written questions in a structured manner to respondents related to their responses to the various variables studied [19]. Questionnaires can also be referred to as written interviews because the contents of the questionnaire are a series of written questions addressed to the respondent and filled out by the respondent himself. Things that need to be considered in compiling questions in the questionnaire use words that are easy to understand, questions that are not too general, avoid suggestive questions, avoid questions that embarrass the respondent, avoid questions that rely on the respondent's memory [20]. Usability testing will be conducted using the J.R. Lewis Usability questionnaire. This questionnaire uses a Likert scale. The data that will be obtained must pass validity and reliability test [21] (see Appendix).

The measuring instrument must perform a validity test and then test its reliability so that the measuring instrument is reliable with a reliability test. The validity test is used to determine whether the questions in the research instrument are valid or can collect information according to the questions given, while the reliability test is carried out to see the level of reliability of the research instrument. Especially research that is a case study, collects research data using research instruments [22].

F. Analytical Hierarchy Process

The Analytic Hierarchy Process (AHP) was introduced by Thomas L. Saaty to solve problems that have many factors and many criteria. AHP is a structured method related to the decision-making process on complex problems, which consists of many alternatives such as projects, actions, and scenarios. AHP is developed according to the hierarchical structure of several alternative combinations in decision-making. For example, knowledge, experience, and intuition. This method can provide an opportunity for everyone to make decisions on various types of problems [23].

IV. RESEARCH RESULT AND DISCUSSION

A. The Weighting of Characteristics and Sub-characteristics of ISO 25010:2011

Determination of the characteristics and sub-characteristic weights of the ISO 25010:2011 quality model was carried out to measure the quality of the Halodoc software. Determination of this weight is done before evaluating the quality of the software. This Weighting method is based on previous research [24]. Weight determination will be carried out on 8 characteristics and 29 sub-characteristics of the ISO 25010:2011 quality model. Determination of the weights will use software assistance, namely Web-based AHP with the link address: https://bpmsg.com/ahp/ahp-calc.php.

B. The weighting of Characteristic of ISO 25010:2011

The results of determining the weight characteristics of the ISO 25010:2011 quality model obtained a weight of 8 characteristics as shown in Fig. 2.
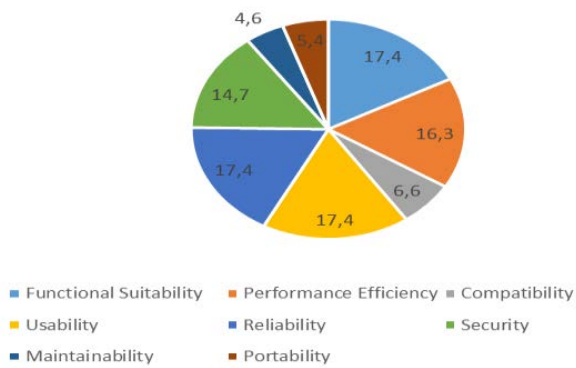
Fig. 2.   Result of Weighting Characteristics of ISO 25010:2011.

Characteristics of Functional Suitability, Usability, and Reliability each get a weight of 17.4%. Characteristic of Performance Efficiency gets a weight of 16.3%. Security characteristic gets a weight of 14.7%. Compatibility characteristic gets a weight of 6.6%. Portability characteristic gets a weight of 5.4% and Maintainability gets a weight of 4.6%.

After determining the weight of each characteristic, the determination of the weight of the sub-characteristics will be carried out on each characteristic of Functional Suitability, Performance Efficiency, Compatibility, Reliability, Security, Maintainability, and Portability. This journal will write 4 characteristics tests that will be used for weight determination and application testing. Those characteristics are Performance Efficiency, Reliability, Security, and Usability.

1) The weighting of performance efficiency: Three sub-characteristics that will be included are Resource Utilization, Time Behavior, and Capacity. The next step is to determine the priority of the three sub-characteristics.

The sub-characteristics of Resource Utilization and Time Behavior have the same important priority so that they get priority 1. It's because Time Behavior is used to examine the processing time when performing a function and Resource Utilization is used to determine the number of hardware resources used to carry out the function.

The Time Behavior sub-characteristic has a higher priority than Capacity with a priority of 5, which is of medium importance. It's because Time Behavior is used to examine the processing time when performing a function while Capacity is used to determine the maximum amount of hardware resources used to carry out the function.

The Resource Utilization sub-characteristic has a higher priority than Capacity and gets priority 3, which is slightly more important. This is because Resource Utilization is used to find out the number of hardware resources used to carry out functions while Capacity is used to find out the maximum number of hardware resources used by the software.

The result of determining the priority of sub-characteristics can be seen in Fig. 3. Time Behaviour gets a weight 48.1%, Resource Utilization gets 40.5%, and Capacity gets 11.4%.
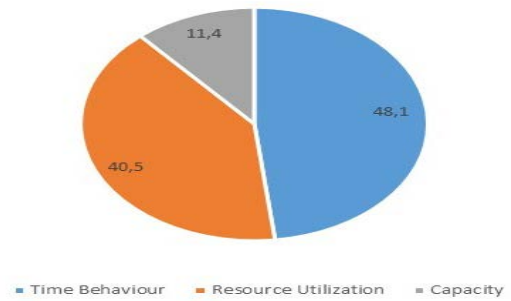

Fig. 3.   Performance Efficiency Weighting Result.

2) The weighting of reliability: The four sub-characteristics that have been included are Availability, Maturity, Recoverability, and Fault Tolerance. The next step is to determine the priority of the four sub-characteristics.

The Maturity and Fault Tolerance sub-characteristic is slightly more important than the Availability and Recoverability sub-characteristics. This gives Maturity and Fault Tolerance priority 3. This is because Maturity is used to determine software maturity in various conditions and Fault Tolerance is used to determine whether the software can run properly even if an error occurs while Availability is a sub-characteristic that serves to test software availability and Recoverability is used to determine whether the software can recover data and systems as desired if a failure occurs.

The sub-characteristics of Maturity and Fault Tolerance have the same level of importance so that they get priority 1, which is equally important. This is because Maturity is used to determine the maturity of the software under various conditions and Fault Tolerance is used to determine whether the software can run properly even if an error occurs.

The sub-characteristics of Availability and Recoverability have the same level of importance so that they get priority 1, which is equally important. This is because Availability is a sub-characteristic that serves to test the availability of software and Recoverability is used to find out whether the software can recover data and systems as desired if a failure occurs.

The result of determining the priority of sub-characteristics can be seen in Fig. 4. Maturity gets a weight 37.5%, Availability gets 12.5%, Fault Tolerance gets 37.5%, and Recoverability gets 12.5%.
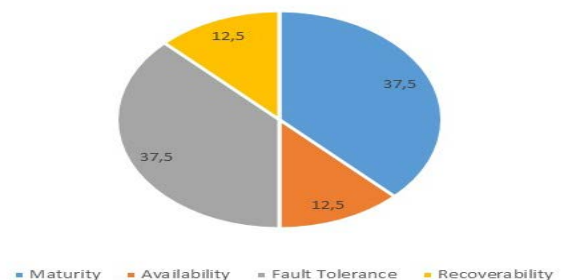

Fig. 4.   Reliability Weighting Result.

3) The weighting of security: The five sub-characteristics included are Integrity, Non-repudiation, Confidentiality Authenticity, and Accountability. The next step is to determine the priority of the five sub-characteristics.

The sub-characteristics of Confidentiality, Integrity, Authenticity have the same level of importance so each of them get priority 1, that means is equally important. This is because the Confidentiality sub-characteristic ensures that application data can only be modified and accessed by authorized parties. Integrity ensures the prevention of alteration or deletion of information by unauthorized parties, and Authenticity ensures that the identity of the subject can be proven as stated by the subject.

The Confidentiality sub-characteristic is slightly more important than the Non-repudiation and Accountability sub-characteristic so it get priority 3 that means it is slightly more important. This is because the Confidentiality sub-characteristic ensures that application data can only be modified and accessed by authorized parties. While non-repudiation is used to determine the extent to which events or actions can be proven to have occurred and Accountability is used to know that the actions of an entity can be traced to that entity in a unique way.

The Integrity sub-characteristic is slightly more important than the Non-repudiation and Accountability sub-characteristic with priority 3. This is because Integrity ensures the prevention of changes or deletion of information by unauthorized parties while Non-repudiation is used to determine the extent to which events or actions can be proven to have occurred and Accountability is used to know that the actions of an entity can be traced to that entity in a unique way.

The Authenticity sub-characteristic is slightly more important than the Non-repudiation and Accountability sub-characteristic with priority 3. This is because Authenticity ensures that the identity of the subject can be proven as stated by the subject whereas Non-repudiation is used to determine the extent to which events or actions can be proven to have occurred and Accountability is used to know that the actions of an entity can be traced to that entity in a unique way.

Non-repudiation sub-characteristics with Accountability have the same level of importance so they have priority 1. This is because these two sub-characteristics have the same two objectives, namely to prove and record activities that have been carried out so that they can be proven.

The result of determining the priority of sub-characteristics can be seen in Fig. 5. The Confidentiality gets a weight 27.3%, Integrity also gets 27.3%, Non-repudiation gets 9.1%, Accountability get 9.1%, and Authenticity get 27.3%.

4) The weighting of usability: Six sub-characteristics of Usability are Learnability, Appropriateness Recognizability, User Error Protection Operability, Accessibility, and User Interface Aesthetics. These six sub-characteristics will be included then be prioritized.
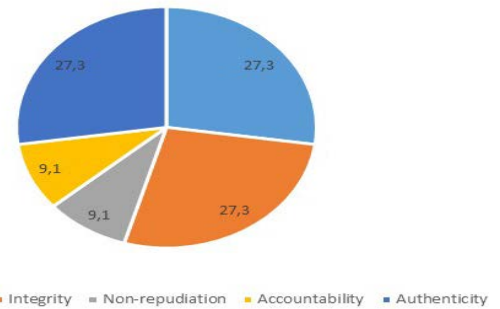

Fig. 5.   Security Weighting Result.

The Learnability and Operability sub-characteristic are more important than the Appropriateness Recognizability sub-characteristic so Learnability and Operability obtain priority 3, slightly more important. Learnability test software can be learned easily or not and Operability tests whether the software is easy to operate or use by users while Appropriateness Recognizability is used to test the extent to which users can recognize whether a software fits their needs.

User Error Protection sub-characteristics and User Interface Aesthetics with Appropriateness Recognizability sub-characteristics obtain priority 1 which means the same importance. User Error Protection tests whether the software can protect users from making mistakes, User Interface Aesthetics tests whether the software interface allows pleasant and satisfying interactions for users, and Appropriateness Recognizability is used to test the extent to which users can identify whether a software fits their needs.

The Appropriateness Recognizability, Learnability, and Operability sub-characteristic are more important than the Accessibility sub-characteristic with priority 3. Appropriateness Recognizability is used to test the extent to which users can recognize whether a software fits their needs, Learnability tests software that can be learned easily or not, and Operability sub-characteristics test whether the software is easy to operate or use by the user while Accessibility tests whether the software can be accessed or not by the user.

Learnability sub-characteristics with Operability sub-characteristics obtain priority 1 which means the same importance. Learnability sub-characteristics test the software can be learned easily or not and Operability sub-characteristics test whether the software is easy to operate or use by users.

Learnability sub-characteristics with User Error Protection and User Interface Aesthetics sub-characteristics get priority 1 which means the same importance. The Learnability sub-characteristics tests whether the software can be learned easily or not, the User Error Protection Sub-characteristics tests whether the software can protect users from making mistakes, and User Interface Aesthetics Sub-characteristics test whether the software interface allows pleasant and satisfying interactions for users.

Operability sub-characteristics with User Error Protection and User Interface Aesthetics sub-characteristics get priority 1 which means the same importance. The Operability sub-

characteristics test whether the software is easy to operate or use by the user, the User Error Protection sub-characteristics test whether the software can protect users from making errors and User Interface Aesthetics sub-characteristics test whether the software interface allows pleasant and satisfying interactions for users.

User Error Protection sub-characteristics with User Interface Aesthetics sub-characteristics obtain priority 1 which means the same importance. The User Error Protection sub-characteristic tests whether the software can protect users from making mistakes and the User Interface Aesthetics sub-characteristic tests whether the software interface allows for a pleasant and satisfying interaction for the user.

The User Error Protection sub-characteristic is more important than the Accessibility sub-characteristic with priority 3, a little more important. The User Error Protection sub-characteristic tests whether the software can protect users from making mistakes while Accessibility tests whether the software can be accessed by the user.

Sub-characteristic User Interface Aesthetics with sub-characteristic Accessibility get priority 1 which means the same importance. Sub-characteristics of User Interface Aesthetics test whether the software interface allows pleasant and satisfying interactions for users and Accessibility tests whether the software can be accessed or not by the user.

The result of determining the priority of sub-characteristics can be seen in Fig. 6. The Appropriateness Recognizability gets a weight 13.5%, Learnability and Operability each get 22.4%, User Error Protection gets 18,2 %, User Interface Aesthetics gets 15.8%, and Accessibility gets 7.7%.

C. Testing Halodoc Application

This section will describe the testing for Halodoc Application. Type of testing will be used depend of the sub-characteristic. Some of the test will be using software for testing.

1) Performance efficiency characteristic test: Time Behavior testing for Halodoc application will use Black-Box testing. This test is carried out with the aim of knowing how long the response and process time can reach the requirements when carrying out the function. This test is carried out using web-based testing software, namely testproject.io.
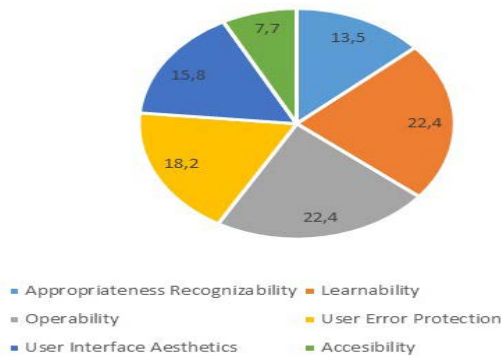

Fig. 6. Usability Weighting Result.

Resource Utilization testing will use Black-Box testing. This test is carried out with the aim of knowing the extent to which the amount and type of resources are used by the software when carrying out its functions in order to meet the requirements. The test for checking RAM allocation for the application will use Apptim Software.

Capacity testing will use Black-Box testing. This test is carried out with the aim of knowing the extent to which the maximum limit of a product or system parameter meets the requirements. The test for checking maximum RAM usage for the application will use Apptim Software.

Fig. 7 will show the score for each sub-characteristics of Performance Efficiency. The total score of Performance Efficiency characteristic will be calculated based on the accumulation sub-characteristics score. This result can be seen in Table I.

2) Reliability characteristic test: Maturity test for the Halodoc application will use stress testing. This test is carried out with the aim of knowing how much the software can meet the requirements for reliability in normal operation. This test is carried out using the Monkey feature on the Android SDK. Monkey can be used to stress testing applications by entering random but repetitive commands.

Availability test for Halodoc application will use stress testing. This test is carried out with the aim of knowing the extent to which the product can operate and can be accessed when needed for use. This test is done by accessing the website version of the Halodoc application using the Jmeter tool to create 100 virtual users. The website version was chosen to facilitate testing with Jmeter and because of the Availability characteristic; it aims to find out if the application can be accessed when needed.

Recoverability test will use Black-Box testing. This test is conducted with the aim of knowing in the event of a failure or breakdown, the system or application can recover the affected data directly and reestablish the desired system state.
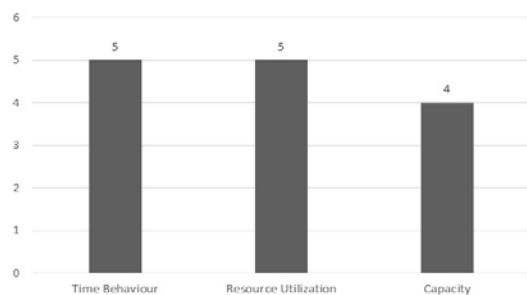
Fig. 7. Performance Efficiency Sub-characteristics Score.

TABLE I. PERFORMANCE EFFICIENCY TEST RESULT

No.	Sub-Characteristics	Weight	Score	Total Score
1	Time Behaviour	48.1%	5	48.1% * 5 = 2.405
2	Resource Utilization	40.5%	5	40.5% * 5 = 2.025
3	Capacity	11.4%	4	11.4% * 4 = 0.456
Total Score of Performance Efficiency				4.886

Fault Tolerance test will use stress testing for Halodoc application testing. This test is carried out with the aim of knowing the extent to which the system or application can operate properly despite software or hardware errors.

Fig. 8 will show the score for each sub-characteristics of Reliability. Total score of Reliability characteristic will be calculated based on accumulation sub-characteristics score. This result can be seen in Table II.

3) Security characteristic test: Confidentiality test will use Black-box testing. This test is carried out with the aim of knowing how the system or application can ensure that data can only be accessed by those authorized. In this test case, it is done by logging into the Halodoc application.

Integrity test for the Halodoc application will use Black-Box testing. This test is carried out with the aim of determining if a system or application can prevent modification of computer programs or data and prevents unauthorized access. In this case, testing will be carried out to change user profile information and modify the Halodoc program.

Non-Repudiation test will use the observation method. This test is carried out with the aim of knowing the extent to which an action or event can be proven to have occurred so that the event or action cannot be denied in the future. The test is carried out by observing the logs or history in the Halodoc application.

Accountability test will use the observation method. This test is carried out with the aim of knowing the ability of the software to distinguish the original user or not. Testing is done by observing user activities when logging into the application.

Fig. 9 will show the score for each sub-characteristics of Security. Total score of Security characteristic will be calculated based on accumulation sub-characteristics score. This result can be seen in Table III.
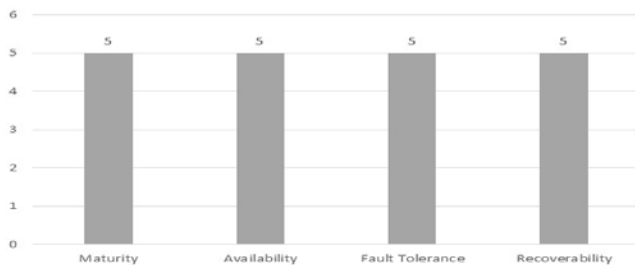

Fig. 8. Reliability Sub-characteristics Score.

TABLE II. RELIABILITY TEST RESULT

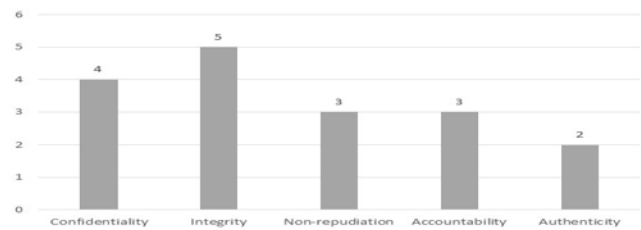| No.	Sub-Characteristics	Weight	Score	Total Score
1	Maturity	37.5%	5	37.5% * 5 = 1.875
2	Avalability	12.5%	5	12.5% * 5 = 0.625
3	Recoverability	12.5%	5	12.5% * 5= 0.625
4	Fault Tolerance	37.5	5	37.5% * 5 = 1.875
Total Score of Reliability				5

Fig. 9. Security Sub-characteristics Score.

TABLE III. SECURITY TEST RESULT

No.	Sub-Characteristics	Weight	Score	Total Score
1	Confidentiality	27.3%	4	27.3% * 4 = 1.092
2	Integrity	27.3%	5	27.3% * 5 = 1.365
3	Non-repudiation	9.1%	3	9.1% * 3 = 0.273
4	Accountability	9.1%	3	9.1% * 3 = 0.273
5	Authenticity	27.3%	2	27.3% * 2 = 0.546
Total Score of Security				3.549

4) Usability characteristic test: Usability characteristic testing for the Halodoc application will use 6 sub-characteristics. The sub-characteristics are Operability, Appropriateness Recognizability, User Error Protection, Learnability, Accessibility, and User Interface Aesthetics. The questionnaire that will be used in this test is based on J.R. Lewis Usability Questionnaire that has been modified by previous research [24]. Testing these characteristics will use a questionnaire distributed to respondents who use the Halodoc application. The results of the questionnaire test will then pass the validity and reliability testing using the IBM SPSS version 20 program.

Research sample selection will use the random sampling technique. The sampling method used refers to the solving approach. The result is 100 samples required for usability testing. The questionnaire will be distributed to 100 people who have used the Halodoc application.

The data from the questionnaire that has been collected will then be tested for validity and reliability. The results of the validity test can be said to be valid if the value obtained exceeds the value of the r table with sample (N) = 100 at a significance of 5% with an r value of 0.197. The questionnaire each question passes this validity question. The next step is to test the Reliability of the questionnaire. Reliability testing is carried out after the validity test results get valid results. This step is done after testing the validity of using SPSS Tools. The result can be seen in Fig. 10.

The results of the reliability test obtained the value of Cronbach Alpha of 0.931. If you look at Table IV, the reliability test results get the Very High category. This means the questionnaire that has been filled out by the respondent is reliable.
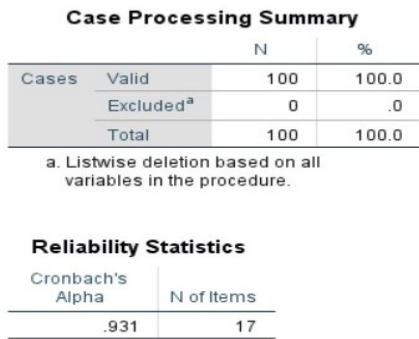
Case Processing Summary

		N	%
Cases	Valid	100	100.0
	Excluded[a]	0	.0
	Total	100	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.931	17

Fig. 10. Usability Questionnaire Reliability Test Result.
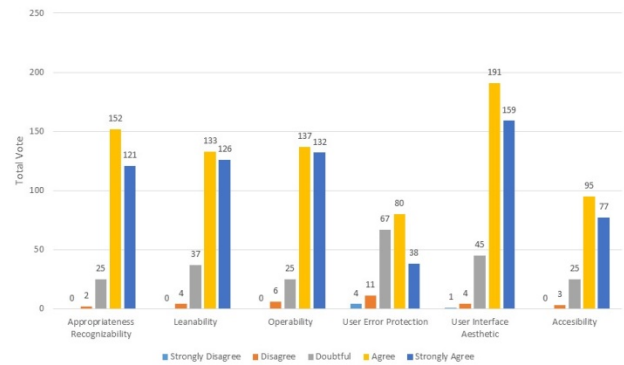
TABLE IV.   RELIABILITY INDEX CRITERIA

No.	Interval	Criteria
1	0.000 - 0.200	Very Low
2	0.201 - 0.400	Low
3	0.401 - 0.600	Enough
4	0.601 - 0.800	High
5	0.801 - 1.000	Very High

Appropriateness Recognizability is done to find out how much users can recognize whether a product fits their needs. The sub-characteristic of Appropriateness Recognizability has 3 questions out of 17 questionnaire questions asked to respondents with a total of 300 votes. From Fig. 11, the Appropriateness Recognizability sub-characteristic gets the most votes for Agree with 152 votes out of 300 votes.

Learnability is carried out to determine if the software is easy to learn by users. Learnability sub-characteristics have 3 questions out of 17 questionnaire questions asked to respondents with a total of 300 votes. From Fig. 11, the Learnability sub-characteristic gets the most votes for Agree with 133 votes out of 300 votes.

Operability is carried out to determine if the software is easy to control and use by users. Operability sub-characteristics has 3 questions out of 17 questionnaire questions asked to respondents with a total of 300 votes. From Fig. 11, the Operability sub-characteristic gets the most votes for Agree with 137 votes out of 300 votes.

User Error Protection is carried out to determine if the system protects users from making mistakes. User Error Protection sub-characteristics has 2 questions out of 17 questionnaire questions asked to respondents with a total of 200 votes. From Fig. 11 can be seen that the User Error Protection sub-characteristic gets the most votes for Agree with 80 votes out of 200 votes.

Aesthetic User Interface is carried out to determine if the user interface allows pleasant and satisfying interactions for users. User Interface Aesthetic sub-characteristics has 4 questions out of 17 questionnaire questions asked to respondents with a total of 400 votes. Fig. 11 shows that the Aesthetic User Interface sub-characteristic gets the most votes for Agree with 191 votes out of 400 votes.

Fig. 11. Usability Sub-Characteristics Votes.

Accessibility is carried out to determine if the software is easily accessed by users. Accessibility sub-characteristics have 2 questions out of 17 questionnaire questions asked to respondents with a total of 200 votes. From Fig. 11, the Aesthetic User Interface sub-characteristic gets the most votes for Agree with 191 votes out of 200 votes.

Total score of Security characteristic will be calculated based on accumulation sub-characteristics score. This result can be seen in Table V.

TABLE V.   USABILITY TEST RESULT

No.	Sub-Characteristics	Weight	Score	Total Score
1	Appropriateness Recognizability	13.5%	4	13.5% * 4 = 0.54
2	Learnability	22.4%	4	22.4% * 4 = 0.896
3	Operability	22.4%	4	22.4% * 4 = 0.896
4	User Error Protection	18.2%	4	18.2% * 4 = 0.728
5	User Interface Aesthetic	15.8%	4	15.8% * 4 = 0.632
6	Accesibility	7.7%	4	7.7% * 4 = 0.308
Total Score of Usability			4	

D. Halodoc Application Total Result

The testing and assessment of the characteristics of the ISO 25010:2011 on the Halodoc application have been carried out. The value of each characteristic can be seen in Fig. 12.

The next step is to calculate the total value for the Halodoc application. Table VI shows the calculation of the total value obtained by the Halodoc application by combining the 8 characteristics of the ISO 25010:2011 model.

The quality test assessment of the software, namely Halodoc, uses the ISO 25010:2011 method with 8 characteristics and 29 sub-characteristics, the assessment is carried out using various testing models such as Black-Box testing, stress testing, observation, and questionnaires distributed to 100 respondents. The Halodoc application gets a score of 4,515 out of a total score of 5. This score is good for Halodoc software because it is only less than 0.5 points adrift to get the maximum score.
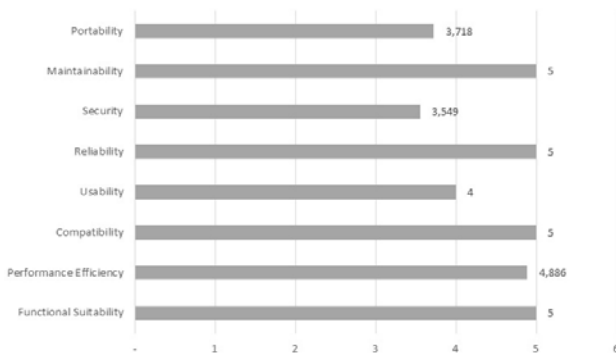
Fig. 12. ISO 25010:2011 Characteristics Score.

From Fig. 12 can be seen that the Security and Portability Characteristic result is enough but need improvement. Performance Efficiency and Usability already get a good result but still, it can be improved.

TABLE VI.   TOTAL SCORE OF HALODOC APPLICATION

No.	Sub-Characteristics	Weight	Score	Total Score
1	Functional Suitability	17.4%	5	17.4% * 5 = 0.87
2	Performance Efficiency	16.3%	4.886	16.3% * 4.886 = 0.796
3	Compatibility	6.6%	5	6.6% * 5 = 0.33
4	Usability	17.4%	4	17.4% * 4 = 0.696
5	Reliablity	17.4%	5	17.4% * 5 = 0.87
6	Security	14.7%	3.549	14.7% * 3.549 = 0.522
7	Maintainability	4.6%	5	4.6% * 5 = 0.23
8	Portability	5.4%	3.718	5.4% * 3.718 = 0.201
Halodoc Application Total Score			4.515	

V.   CONCLUSION AND SUGGESTION

A. Conclusion

Halodoc application testing using the ISO 25010:2011 model has been successfully carried out. Characteristics of Functional Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, Portability, and Maintainability with 29 sub-characteristics used to assess Halodoc applications. Tests are carried out using several testing methods ranging from Black-Box testing to testing the characteristics of Functional Suitability, Performance Efficiency, Compatibility, Security, Portability, and Maintainability. Stress testing is also carried out to test the characteristics of Reliability. The questionnaire method was used for usability testing which was distributed to 100 respondents using the Halodoc application.

Characteristics of Functional Suitability, Compatibility, Reliability, and Maintainability managed to get the best score, namely, 5. The Usability and Performance Efficiency Characteristics managed to get scores of 4 for Usability and 4.886 for Performance Efficiency. This characteristic has received good values but can still be improved. Characteristics of Portability and Security get a value of 3.718 and 3.549 respectively. The total score obtained is 4.515 from a total value of 5 for the Halodoc application. This value indicates that

the Halodoc application has good software quality, although some characteristics still need development and improvement. The results of this test can be taken into consideration for Halodoc application developers to correct deficiencies and improve what is already good.

B. Suggestion

Testing the quality of the Halodoc application using the ISO 25010:2011 method can still be developed again. The use of tools or tools for testing can be added again. Such as tools for testing the characteristics of Reliability, Security, and Portability. The testing method used can also be added, such as the penetration testing method for security characteristics. Sub-characteristics that were not used in this study can also be added for further research. This can be done so that the shortcomings that occur in the application can be known in detail.

VI.   LIST OF ABBREVIATIONS

TABLE VII.   ABBREVIATIONS LIST

No.	Abbreviations	Meaning
1	AHP	Analytical Hierarchy Process
2	ISO	International Organization for Standardization
3	SPSS	Statistical Package for the Social Sciences
4	SQA	Software Quality Assurance
5	UI	User Interface

REFERENCES

[1] R. Komalasari, "Manfaat Teknologi Informasi Dan Komunikasi Di Masa Pandemi Covid 19," Tematik, vol. 7, no. 1, pp. 38–50, 2020, doi: 10.38204/tematik.v7i1.369.

[2] M. Machmud, A. Masmuh, C. Nasirin, Salahudin, T. Haharuddin, and A. E. Z. Musma, "Artificial Intelligence In The Public Health Sector : The Use Of Telemedicine In Indonesia During Covid-19," Palarch, J. Archaeol. Egypt/Egyptology, vol. 17, no. 7, pp. 10106–10118, 2020.

[3] C. N. Mangkunegara, F. Azzahro, and P. W. Handayani, "Analysis of factors affecting user's intention in using mobile health application: A case study of halodoc," 2018 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2018, pp. 87–92, 2019, doi: 10.1109/ICACSIS.2018.8618174.

[4] M. Izzatillah, "Quality Measurement of Transportation Service Application Go-Jek using Iso 25010 Quality Model," Simetris J. Tek. Mesin, Elektro dan Ilmu Komput., vol. 10, no. 1, pp. 233–242, 2019.

[5] F. A. Alvian, A. T. Wibowo, and M. A. Izzuddin, "Software Quality Testing on the E-Sim Application Based on Iso 9126," no. January 2019, pp. 461–472, 2020.

[6] D. R. dan T. K. Syehka Larasati, "Pembangunan Sistem Ujian Harian Siswa Berbasis Web Dengan Mengacu Pada Standar Kualitas ISO 25010," J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 2, no. 11, pp. 4357–4364, 2018.

[7] H. Bayu, A. H. Brata, and A. A. Soebroto, "Analisis Pengujian Sistem RAKOON POS Berdasarkan Teori Kualitas McCall," J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 4, no. 1, pp. 207–213, 2020.

[8] B. N. Setyanto, M. Mushlihudin, and D. Y. Pradana, "Quality analysis of a learning media analog electronics on the Android platform with ISO 25010," J. Vocat. Educ. Stud., vol. 4, no. 1, p. 112, 2021, doi: 10.12928/joves.v4i1.4047.

[9] D. A. P. Wulandari and M. D. W. Aristana, "Analysis Evaluation Management Information System Audit Internal Quality," J. Electr. Electron. Informatics, vol. 5, no. 1, pp. 5–9, 2021.

[10] A. Kurniawan, D. Arifianto, and L. A. Muharom, "Analisis Kualitas Sisitem Informasi Akademik Mobile Menggunakan Iso 25010," Univ. Muhammadiyah Jember, vol. 25010, pp. 1–16, 2020.

[11] D. Galin, "SQA - Definitions and Concepts," in Software Quality: Concepts and Practice, 2018.

[12] Pressman, Software Engineering Practioner's APP 8th Ed. 2019.

[13] A. Ardian and Y. Fernando, "Sistem Informasi Manajemen Lelang Kendaraan Berbasis Mobile (Studi Kasus Mandiri Tunas Finance)," J. Teknol. dan Sist. Inf., vol. 1, no. 2, pp. 10–16, 2020.

[14] ISO, "Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE)-System and software quality models," ISO/IEC. Mar-2011.

[15] P. Ammann and J. Offutt, "Introduction to Software Testing Edition 2," 2017.

[16] T. Murnane and K. Reed, "On the effectiveness of mutation analysis as a black box testing technique," in Proceedings of the Australian Software Engineering Conference, ASWEC, 2001, doi: 10.1109/ASWEC.2001.948492.

[17] S. Masripah and L. Ramayanti, "Pengujian Black Box Pada Sistem Informasi Penerimaan Siswa Baru Berbasis Web," Inf. Syst. Educ. Prof. J. Inf. Syst., 2019.

[18] I. Sopian, "Implementasi Dan Pengujian Sistem Informasi Tracer Study Berbasis Web Menggunakan Standard Iso/Iec 9126 (Studi Kasus: Fakultas Pendidikan Matematika Dan Ilmu Pengetahuan Alam-Upi)," J. Teknol. Inf. dan Komun., vol. 8, no. 2, 2020.

[19] V. H. Pranatawijaya, W. Widiatry, R. Priskila, and P. B. A. A. Putra, "Penerapan Skala Likert dan Skala Dikotomi Pada Kuesioner Online," J. Sains dan Inform., vol. 5, no. 2, pp. 128–137, 2019.

[20] Kemdikbud, "Kelebihan dan Kelemahan Kuesioner/Angket," BPMK-Kemdikbud, 2020. [Online]. Available: https://m-edukasi.kemdikbud.go.id/medukasi/produk-files/kontenkm/km2016/KM201627/materi1.html. [Accessed: 07-Dec-2020].

[21] J. Sauro and J. R. Lewis, Quantifying the user experience: Practical statistics for user research. Morgan Kaufmann, 2016.

[22] T. E. Situmorang and D. Purba, "Perancangan Aplikasi Pengujian Validitas dan Reliabilitas Instrumen Penelitian," KAKIFIKOM Kumpul. Artik. Karya Ilm. Fak. Ilmu Komput., vol. 1, no. 2, pp. 54–58, 2020.

[23] R. S. Ilhami and D. Rimantho, "Penilaian Kinerja Karyawan dengan Metode AHP dan Rating Scale," J. optimasi Sist. Ind., vol. 16, no. 2, pp. 150–157, 2017.

[24] Y. Nuraini, "Pengujian Kualitas Perangkat Lunak Ujian Berbasis Komputer Squiz menggunakan Metode ISO 25010:2011." 2017.

APPENDIX

A. List of Question in Usability Questionnaire

TABLE VIII.   USABILITY QUESTIONS

No.	*Sub-characteristic*	Question
1	Appropriateness Recognizability	Overall I am satisfied using the Halodoc application?
2	Appropriateness Recognizability	How to use the Halodoc application is very simple?
3	Appropriateness Recognizability	The information obtained is very easy to understand?
4	Learnability	I feel very fast in getting health information on this application?
5	Learnability	I find it very easy to learn the Halodoc application?
6	Learnability	The information provided by the Halodoc application is very clear?
7	Operability	Can I get health information quickly using the Halodoc application?
8	Operability	Is it easy to get information about health on the Halodoc application?
9	Operability	The information I get on the Halodoc application helps to get the health information I want to know?
10	User Error Protection	If an error occurs, the application informs the error that occurred?
11	User Error Protection	Whenever I make a mistake, can I recover quickly?
12	User Interface Aesthetic	Am I comfortable using the Halodoc application?
13	User Interface Aesthetic	The layout in the Halodoc application display is very clear?
14	User Interface Aesthetic	Halodoc application display is very easy?
15	User Interface Aesthetic	I like the look of apps like Halodoc?
16	Accessibility	The ability of the application to complete all the functions in it make me satisfied?
17	Accessibility	Overall I am satisfied with the Halodoc application?