Evaluation Of Digital Library's Usability Using the System Usability Scale Method of (A Case Study)

1st Suci Ratnawati
Information System Major
Syarif Hidayatullah State Islamic University Jakarta
Jakarta, Indonesia
suci.ratnawati@uinjkt.ac.id

2nd Luthfiyyah Widianingsih
Information System Major
Syarif Hidayatullah State Islamic University Jakarta
Jakarta, Indonesia
luthfiyyah.widianingsih16@mhs.uinjkt.ac.id

3rd Nenny Anggraini
Informatic Engineering
State Islamic Syarif Hidayatullah
Jakarta, Indonesia
Nenny.anggraini@uinjkt.ac.id

4th Imam Marzuki Shofi
Department of Informatics
UIN Syarif Hidayatullah
Jakarta, Indonesia
imam@uinjkt.ac.id

5th Nashrul Hakiem
Informatic Engineering
State Islamic Syarif Hidayatullah
Jakarta, Indonesia
hakiem@uinjkt.ac.id

6th Fenty Eka M Agustin
Department of Informatics
UIN Syarif Hidayatullah
Jakarta, Indonesia
fenty.eka@uinjkt.ac.id

***Abstract*— The benefit of Information Technology (IT) has spread to almost all fields, including in libraries. IT applications in libraries are the automation of libraries and digital libraries. Digital library or digital library system is the concept of using the internet and information technology in library management. Many libraries have implemented digital libraries, one of them is E-Perpus belongs to Depok City. E-Perpus contains digital books that can be accessed online 24 hours, anytime and everywhere. However, when an application has been launched, it is necessary to evaluate to find out how effective, efficient, and satisfying the application according to the user, what obstacles they might face and the feedback given later can be a benchmark for improving the performance of the application so that it is better in the future. Therefore, in this study, we will conduct an evaluation of usability using the System Usability Scale(SUS) to find out the effectiveness of SUS to evaluate the usability of the system. The results of the evaluation show that the SUS score of E-Perpus is 69.6. It is indicated that the application was included in the Marginal High, OK category, and a D predicate. In addition, for its relation to the Net Promoter Score, it shows that users have the potential to become Detractors where they feel less satisfied with the product or service and can have a negative effect on the brand by a negative review.**

Keywords*— *Usability, System Usability Scale, Questionnaire

I. INTRODUCTION

The benefit of Information Technology (IT) has spread to almost all fields, including libraries such as the automation of libraries and digital libraries. The library automation system is an integration between the fields of administration, procurement, inventory, cataloging, processing, circulation, statistics, management of library members, while the digital library or digital library system is the concept of using the internet and information technology in library management [1].

E-Perpus of Depok City is a library system that contains digital books. E-Perpus is expected to help users in searching for information or knowledge in the form of digital books that can be accessed 24 hours, anytime, and anywhere. E-Perpus application can be downloaded on the playstore for the android version and there is a web version for Windows. Until now, Depok City's E-Perpus

application has more than 6,000 various book titles. With a lot of book collections, E-Perpus is useful if it has a lot of users accessed it and gets the benefit. It means that E-Perpus have usability. So It's important to conduct a usability evaluation of E-Perpus. So, researchers choose E-Perpus as an object of research to explain the effectiveness of SUS to evaluate usability.

According to Nielsen, Usability is a qualitative analysis that determines how easy a user is to use the interface of an application. An application is called usable if its functions can be carried out effectively, efficiently, and satisfactorily. Effectiveness relates to the success of users in achieving goals in using the software. Efficiency refers to the smooth running of the user to achieve these goals. Satisfaction is related to user acceptance of the software [2].

When an application has been launched, a usability evaluation needs to be done to find out how effective, efficient, and satisfying an application is according to the user, what obstacles they may face and the feedback provided can be later a benchmark for improving and improving the performance of the application so that it is better in the future. There are several questionnaires that can be used in usability testing as suggested by Garcia [3], including:

1. SUS (System Usability Scale). SUS was developed by Brooke [4] as a "quick and dirty" usability measure.
2. QUIS (Questionnaire for User Interface Satisfaction). QUIS is a tool developed by a multidisciplinary research team at the University of Maryland Harper and Norman [5]
3. SUMI (Software Usability Measurement Inventory). SUMI is a licensed questionnaire consisting of 50 questions.
4. PSSUQ (Post-Study Usability Questions-naires). PSSUQ is a questionnaire with 16 question items.

Sauro [6] and Tullis and Stetson [7] research show that the System Usability Scale (SUS) is a valid and reliable usability evaluation tool. Based on the description above, the authors are interested in conducting research evaluation of E-Perpus's usability to improving Its performance and choose SUS as a tool because of its validity dan reliability.

II. THEORETICAL REVIEW

A. Usability

Nielsen defines usability as a quality that examines and measures the ease of display used by users [2]. The International Standard defines usability as the improvement of a product that is used by certain users to achieve certain goals

such as effectiveness, efficiency, and user satisfaction [8]. Nielsen defines usability into five main indicators, namely:

1. Learnability. An indicator that measures how easily users complete tasks that must be done while using the website.
2. Efficiency. Indicators that measure the speed and accuracy of users accessing a system.
3. Memorability. This indicator measures how far the user's memory is after accessing a system. Memorability usually requires research over a long period of time.
4. Errors. An indicator that measures how much the user error in using the system.
5. Satisfaction. An indicator that measures how satisfied users are with the website they use.

B. *Usability Testing*

Usability testing is done to find out how effective, efficient, and satisfying an application is according to the user. There are several questionnaires that can be used in usability testing as suggested by Garcia [3], among others:

1. SUS (System Usability Scale). SUS was developed by Brooke [4] as a "quick and dirty" usability measure. The survey consisted of 10 questions; each has 5 Likert points in response. The SUS output is a score that looks easy to understand, with a range from 0 to 100, with the higher the score, the better the usability.
2. QUIS (Questionnaire for User Interface Satisfaction). QUIS is a tool developed by a multidisciplinary research team at the University of Maryland Harper and Norman [5] QUIS is designed to assess users' subjective satisfaction with specific aspects of human interaction.
3. SUMI (Software Usability Measurement Inventory). SUMI is a licensed questionnaire consisting of 50 questions. SUMI can be used to measure perceptions of efficiency, effectiveness, usability, systems, and user learnability of the system. SUMI is available in 12 languages. SUMI is very reliable (0.92). SUMI license costs around USD $ 700 a month.
4. PSSUQ (Post-Study Usability Questions-naires). PSSUQ is a questionnaire with 16 question items. PSSUQ measures the satisfaction that users feel about the product or system. PSSUQ provides an overall satisfaction score with an average of the subscales, namely: system quality, information quality, and interface quality. PSSUQ is available free of charge. System Usability Scale

System Usability Scale (SUS) is a questionnaire that can be used to measure the usability of a computer system according to the user's subjective perspective [9]. SUS was developed by John Brookes in 1986. SUS is in the form of a questionnaire consisting of 10 question items as shown in Table 1. The SUS questionnaire uses a 5-point Likert scale. Respondents were asked to give ratings "Strongly disagree", "Disagree", "Neutral", "Agree", and "Strongly agree" on 10 items of the SUS statement according to their subjective assessment. If the respondent feels that they do not find the right response scale, the respondent must fill in the midpoint of the test scale [4].

C. *Net Promoter Score*

Net Promoter Score (NPS) is a management tool that can be used to measure the loyalty of a company's customer relationships. This serves as an alternative to traditional customer satisfaction research and is claimed to correlate with revenue growth [10]. NPS is also one of the benchmarks that are often used in measuring the usability level of a product other than SUS. NPS can be as low as -100 where each respondent is a "detractor" or as high as +100 which means each respondent is a "promoter".

III. Research Methods

The research will be conducted based on the research stages shown in Figure 1 below

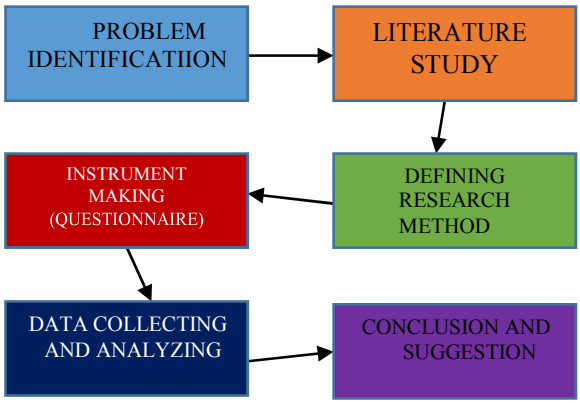

Fig. 1. Research Stages

The first stage in this research is to identify the problem by conducting a review of the research background. Then a literature study is carried out to obtain theoretical foundations related to the research being carried out, as well as similar literature on previous research.

The next stage is determining the research method that will be used in this study. The method used to determine user satisfaction is to measure the quality of application services using the System Usability Scale (SUS), and the research method used is quantitative research methods.

Next is to make research instruments as data that will be collected from users in the form of questionnaires distributed at the time of the survey. The research instrument was made consisting of 10 questions where each question uses a Likert scale of five to seven scales. After the research instrument is made, the next step is to collect and analyze data obtained from the survey to produce conclusions and suggestions for improvements to the application that is the object of research.

A. *Population and Sample*

The population in this study are all users of E- Perpus. In this study, the sample from the population was taken using a simple random sampling technique, where the right to be respondents were all people, especially those who live in Depok, with the sample calculation using the Slovin formula, namely:

$$n=\frac{N}{1+N\,(e)^2} \quad (1)$$

Where n is the number of samples, N is the total population, and e is the error tolerance limit. In this study, the total population of 500 was taken from the total number of applications downloaded at the Google Play Store and the error tolerance limit was set at 10%. Then the number of samples obtained is 83.

B. Data Collecting

- **Observation.** Observation activities were carried out directly to a research case study, namely the Department of Archives and Library of Depok City (DISKARPUS) which is located at Jl. Margonda Raya No.54 Depok to get information about the object of research, namely E- Perpus.
- **Literature study.** Literature study is carried out by looking for references from various sources such as relevant books, journals, and theses in order to clarify the problems and solutions for this research.
- **Questionnaire.** The questionnaire that is used refers to the System Usability Scale (SUS) and the measurement uses a Likert scale. The 10 SUS statements which have been validated can be seen in Table 1.

TABLE I. SYSTEM USABILITY SCALE (SUS) STANDARD QUESTIONNAIRE

No.	Original Item
1	I think that I would like to use this system.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I think that I would need the support of a technical person to be able to use this system.
5	I found the various functions in the system were well integrated.
6	I thought there was too much inconsistency in this system.
7	I would imagine that most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

C. Data Analysis

The data analysis technique used in this research is a descriptive analysis of quantitative analysis results. The quantitative data in this study was obtained by filling in the SUS questionnaire. Furthermore, the results of filling out the SUS questionnaire are used to analyze the level of user satisfaction

IV. RESULT AND DISCUSSION

Based on the results of questionnaire data collection distributed online on social media and Depok community groups from March 5 to April 5 2020, there were 85 respondents. The number of the respondent is enough according to Roscoe that the appropriate sample size in the study was between 30 and 500 [11]. Furthermore, the calculation is carried out using the formula that has been determined to get the SUS score. The results of the average SUS score is 69,6.

A. Validity test

The validity test used Pearson (2 tails) with a significance level 5%. The results are considered valid if Rcount > Rtable with Rtable of 0.361. Table 3 shows that the Rcount on the 10 questionnaire items is greater than the Rtable so that the 10 questionnaire items are considered valid.

TABLE II. VALIDITY TEST RESULT

	Rcount	Rtable	Information
R1	0.413475	0.361	Valid
R2	0.421072	0.361	Valid
R3	0.384233	0.361	Valid
R4	0.370706	0.361	Valid
R5	0.374272	0.361	Valid
R6	0.451025	0.361	Valid
R7	0.48587	0.361	Valid
R8	0.440121	0.361	Valid
R9	0.419222	0.361	Valid
R10	0.482056	0.361	Valid

B. Reliability test

Reliability test was performed using the Cronbach's Alpha method. The questionnaire will be considered reliable if the value is greater than 0.6. The reliability test results from Excel are shown in Table 4. The results show that the Cronbach's Alpha value for the 10 questionnaire items was 0.635, greater than 0.6 so that the questionnaire was considered reliable.

TABLE III. RELIABILITY TEST RESULT

Alfa Cronbach	Number of Items	Information
.635	10	Reliable

C. SUS Score Analysis

SUS is a global assessment of the usability aspects (effectiveness, efficiency, and satisfaction) that are subjectively perceived by users. The SUS score can indicate the level of user acceptance. The SUS score must be more than 70 to be included in the Acceptable category [11]. The SUS score of the E-Perpus is 69.6. This score shows that E-Perpus is in the Marginal high category as shown in Figure 2. The marginal or border value itself is divided into marginal low and high (score 50-70) for describes the user's tendency towards the application whether it is considered acceptable or not. The SUS score of e- Perpus which is included in the marginal high means that the user feels that the application is good enough and is likely to enter the Acceptable category if in the future the application experiences improvements or improvements

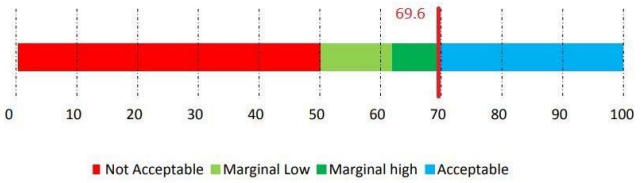

Fig. 2. E-Perpus Acceptance Rate according to Brook (2013)

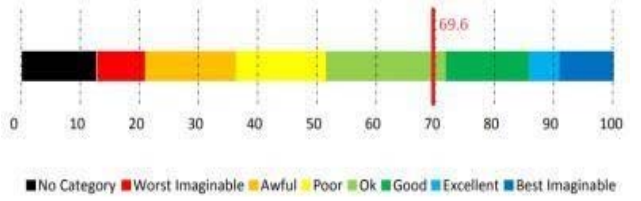

Fig. 3. Depok City E-Perpus Application Adjective value of SUS Score according to Bangor et al. (2009)

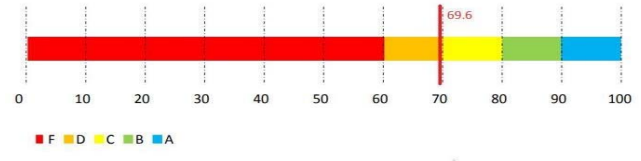

Fig. 4. E-Perpus Application Grade value of SUS Score according to Sauro (2011)

A SUS score is also considered Good if it is over 70.4 [12]. The SUS score for the E-Perpus is 69.6 that means is included in the OK category as shown in Figure 4.3. The OK category according to Babbit & Nystrom's assessment [13] is more accurately called fair or so-so, which means that the user feels that this application is mediocre and is quite satisfied with the application, but at the same time implies that the overall level of usability of the application is not yet acceptable.

Sauro's research also sheds light on the SUS score category. To get an A predicate, the SUS score must be at least 90 [6]. The SUS score for the E-Perpus application is 69.6, which is categorized as D as shown in Figure 4.

The SUS score can also indicate a tendency to become a Net Promoter [14]. An SUS score of 90 or more indicates a user has the potential to become a Promoters, while an SUS score below 70 indicates a potential user to become a Detractor. The E-Perpus City application score is 69.6, indicating that the user is potentially become a Detractor where they feel less satisfied with the product or service and can have a negative effect on the brand with negative reviews.

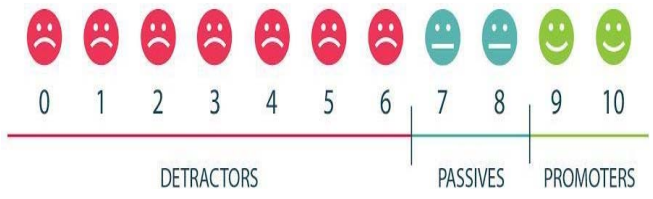

Fig. 5. Net Promoter Score Indicator Sources: (http://business2community.com)

The results of the calculation of the SUS score for the E-Perpus according to the user's subjective assessment can be seen that the usability level of the application is considered not sufficiently effective, efficient, and satisfying. This can happen due to the lack of regular updates or application improvement carried out by DISKARPUS. Therefore the recommendation that can be given is to make an evaluation team and application developers who routinely carry out surveys and improvements, as well as performance improvement by DISKARPUS in terms of service and promotion of e-Perpus application. So that more people use this application because based on The reviews given by users on Playstore, some of them experienced difficulties in the account registration process, which took a long time or were not responded to by responsible staff.

In terms of application, recommendations that can be given are changes to the appearance of the main page to make it more dynamic and attractive, the development of the 'Message Box' page to a 'Help & Feedback' page where the page contains articles of problems that users may encounter when using E-Perpus application and input fields so that users can more easily provide useful feedback in future application development, such as an example of the Help & Input page from the Google Playstore application in Figure 6

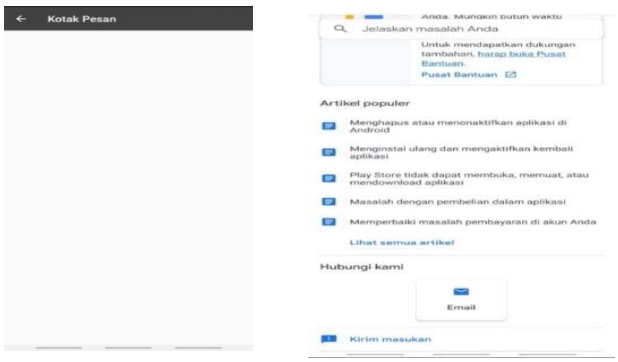

Fig. 6. 'Message Box' and "Help&Feedback' page Sources: E-Perpus Application and Google Playstore

V. Conclusion and Suggestion

The measurement results of the E-Perpus application are 69.6 which indicates that the overall usability level of the application is not effective, efficient, and satisfying enough for users. Therefore, it is necessary to re-evaluate and develop the application periodically both in terms of interface and others, as well as improving service and performance by DISKARPUS Depok City. It shows that SUS is effective to evaluate the application. This research is still possible developed both by adding variables or using other methods, and testing the website can also be done to determine the usability comparison between mobile E-Perpus and the web based E-Perpus.

References

[1] Z. Hasibuan, "Pengembangan Perpustakaan Digital: Studi Kasus Perpustakaan Universitas Indonesia," Makalah Pelatihan Pengelola Perpustakaan Perguruan Tinggi, 2005.

[2] J. Nielsen, "Usability 101: Introduction to Usability," Nielsen Norman Group, 3 January 2012. [Online]. Available: http://www.nngroup.com/articles/usability-101-introduction-to- usability/. [Diakses 10 March 2019].

[3] A. Garcia, "UX Research | Standardize Usability Questionnaire," ChaiOne, 27 November 2013. [Online]. Available: http://chaione.com/ux-research-standardizing-usability- questionnaires/. [Diakses 10 March 2019].

[4] J. Brooke, "SUS - A quick and dirty usability scale," dalam Usability Evaluation in Industry, London, Taylor and Francis, 1996, pp. 189-194.

[5] B. D. Harper dan K. L. Norman, "Improving User Satisfaction: The Questionnaire for User Interaction Satisfaction Version 5.5," dalam 1st Annual Mid-Atlantic Human Factors Conference, Maryland, 1993.

[6] J. Sauro, "Measuring Usability with the System USability Scale (SUS)," measuring, 2 February 2011. [Online]. Available: https://measuringu.com/sus/. [Diakses 10 March 2019].

[7] T. Tullis dan J. Stetson, "A Comparison of Questionnaires for Asessing Website Usability," dalam Usability Professional Association Conference, Boston, 2006.

[8] N. Bevan, "Human Computer Interaction Standards," Advance in HUman Factors/Ergonomics, vol. 20, pp. 885-890, 1995.

[9] A. Bangor, P. Kortum dan J. Miller, "Determining what individual SUS score mean: adding an adjective rating scale," Journal of Usability Studies, vol. 4, no. 3, pp. 114-123, 2009.

[10] R. Bergevin, A. Kinder, S. Winston, dan B. Simpson, Call Centers for Dummies 2nd Edition, America, For Dummies, 2010, p. 345.

[11] B. Pudjoatmodjo dan R. Wijaya, "Tes Kegunaan (Usability Testing) Pada Aplikasi Kepegawaian Dengan Menggunakan System Usability Scale (Studi Kasus: Dinas Pertanian Kabupaten Bandung)," Seminar Nasional Teknologi Informasi dan Multimedia, vol. 4, no. 1, pp. 37-42, 2016.

[12] L. S. Wiratama dan D. Sasongko, "Evaluasi Antarmuka Website SMK Muhammadiyah 2 Sragen Menggunakan Metode Usability Testing," Jurnal SIMETRIS , vol. 8, no. 1, pp. 135-140, 2017.

[13] B. A. Babbit dan C. O. Nystrom, "Questionnaire Construction Manual," Army Research Institute for the Behavioral and Social Science, Texas, 1989.

[14] J. Sauro, "Does better usability increase customer loyalty?," MeasuringU, 7 January 2010. [Online]. Available: https://measuringu.com/usability-loyalty/. [Diakses 15 February 2020