

Evaluation of the Usability of Meta AI in WhatsApp Using the System Usability Scale (SUS)

Alif Salsabila.A¹, Evy Nurmiati², Nuryasin³

¹²³ Universitas Islam Negeri Syarif Hidayatullah Jakarta

Corresponding Author : alif.salsabila.a24@mhs.uinjkt.ac.id

Abstract: The integration of artificial intelligence (AI) assistants into instant messaging applications has created new opportunities for users to access information and interact with conversational systems. However, the usability of these integrated features requires further evaluation. This study aimed to evaluate the perceived usability of Meta AI in WhatsApp using the System Usability Scale (SUS). A quantitative approach was employed through a questionnaire administered to 30 university students in Indonesia. Of these, 27 respondents were represented in the SUS calculation, while three respondents who had never used the feature were not included in the usability assessment. The results yielded an average SUS score of 53.0. Based on commonly used SUS interpretation frameworks, this score falls within the Marginal Low acceptability range, corresponds to Grade D on the grading scale, and receives an Okay adjective rating. The score is also below the commonly cited benchmark of 68, indicating relatively low perceived usability among the respondents included in the assessment. These findings suggest that the usability of Meta AI in WhatsApp requires further improvement. However, the results should be interpreted within the context of the study sample, which consisted of university students, and should not be generalized to all WhatsApp users. Further research involving more diverse respondents and complementary usability evaluation methods is recommended.

Keyword: Artificial Intelligence; Meta AI; System Usability Scale; Usability; WhatsApp

Copyright © 2025 Walisongo Journal of Information Technology

Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.



To cite this article (APA Style):

1. Introduction

Usability evaluation is an essential component of software development because it helps determine whether a digital product can be used effectively and understood by its intended users. As artificial intelligence (AI) continues to develop, intelligent assistants are increasingly integrated into existing digital platforms, including instant messaging applications. WhatsApp, for example, enables users to communicate and exchange information through messages, photos, and videos (Yahya & Dirgantara, 2022). The integration of Meta AI into WhatsApp further extends this functionality by allowing users to ask questions and obtain information through a conversational interface. This development highlights the importance of evaluating not only the availability of AI features but also how users perceive their usability.

WhatsApp has a substantial user base in Indonesia. According to We Are Social data reported by (Databoks, 2025), WhatsApp had a reported penetration rate of 91.7% in Indonesia at the beginning of 2025. This figure demonstrates the widespread use of the messaging platform; however, it does not directly indicate the adoption rate of the Meta AI feature within WhatsApp. Therefore, the usability of Meta AI needs to be examined independently of the popularity of its host application. As AI assistants become integrated into familiar digital environments, understanding users' perceptions of these features is important for identifying potential usability concerns and informing future improvements.

Previous studies have examined usability in various digital applications and AI-based systems. (Nurmiati & Irahman, 2023), for example, evaluated the usability of Microsoft Word, highlighting the importance of usability assessment in software applications. Research on AI-based text generation has also evaluated ChatGPT using the System Usability Scale (SUS), reporting a score of 67.44 (Mulia et al., 2023). In addition, (Caro-Álvaro et al., 2022) examined usability recommendations in the development of mobile instant messaging applications. These studies provide relevant perspectives on usability evaluation across AI-based tools and messaging applications. However, the present study focuses specifically on the perceived usability of Meta AI as an AI assistant integrated into WhatsApp, using the SUS instrument. This specific evaluation provides an initial assessment of how the surveyed users perceive the usability of this feature.

The System Usability Scale (SUS) was selected because it provides a standardized questionnaire for assessing perceived usability. By using SUS, this study aims to obtain an overall usability score for Meta AI in WhatsApp and interpret the result using established SUS interpretation dimensions. The evaluation is intended to provide an initial basis for identifying whether the perceived usability of the feature meets commonly used usability benchmarks and for formulating recommendations for future improvement. Previous research has also highlighted the relevance of user experience and perceived ease of use in the continued use of AI chatbots (Pitts & Motamedi, 2026). Nevertheless, the present study does not directly measure adoption rates, actual usage effectiveness, or the accuracy of AI-generated responses.

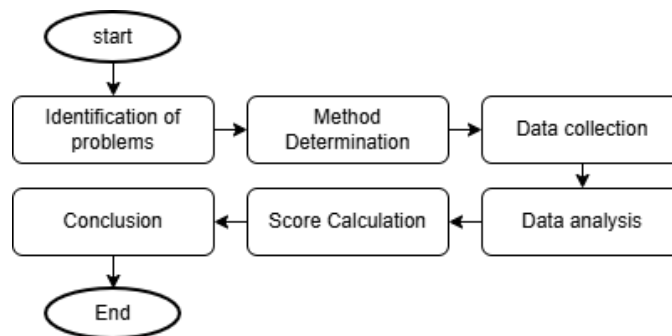
The study employs a quantitative, descriptive approach to evaluate the perceived usability of Meta AI in WhatsApp among university students in Indonesia. An initial expectation was that the feature would achieve an average SUS score above the commonly cited benchmark of 68. This expectation serves as a reference for interpreting the descriptive results rather than as a statistically tested hypothesis. Accordingly, the study addresses the following research question: What is the perceived usability of Meta AI in WhatsApp among the surveyed university students based on the System Usability Scale? The findings are expected to provide an initial understanding of the feature's perceived usability within the study sample and support recommendations for further usability improvement.

2. Method

This study employed a quantitative approach, enabling the evaluation of user satisfaction with the usability of WhatsApp Meta-AI function using the System Usability Scale (SUS). The methodology comprised five steps: problem identification, method selection, data collection via questionnaire, data analysis (validity and reliability checks, data processing, and mean calculation), interpretation of the results to assess the application's usability, and finally, the formulation of conclusions and recommendations. These steps are illustrated in Figure 1 (Hudaya & Nurmiati, 2023). The sole measured variable was the usability of WhatsApp's meta-AI function. The SUS instrument used in this study has been extensively validated for use in

various digital interfaces, including mobile applications and AI-based systems. In particular, the work of (Usman & Gustalika, 2022) confirmed the validity and reliability of the SUS questionnaire for smartphone applications, thus validating its use for evaluating mobile features such as MetaAI in WhatsApp. (Deshmukh & Chalmeta, 2024) further validated the SUS as a usability measure for AI-driven interactive interfaces, demonstrating a strong correlation between SUS scores and users' system quality ratings (Pearson correlation coefficient $r = 0.9093$). This confirms the robustness of the instrument for the present study.

Figure 1
Research Methodology Stages



In this study, the method used to collect data was to distribute questionnaires to WhatsApp users, specifically regarding the Meta AI feature (Hudaya & Nurmiati, 2023). This research questionnaire used the System Usability Scale (SUS) method, which consists of five positive and five negative questions, as shown in Table 2. The scale measurement tool used in this study was a five-point Likert scale, with respondents providing ratings as shown in Table 1. (Nurmiati & Irahman, 2023)

The respondents in this study were university students from across Indonesia, with a total of 30 respondents who met the criteria: 27 users of the Meta AI feature on WhatsApp and three respondents who did not use the Meta AI feature on WhatsApp. According to (Sugiyono, 2019) a suitable sample size for research ranges from 30 to 500 respondents. Therefore, the number of respondents in this study met the minimum criteria.

Table 1
Rating Categories on a Likert Scale

Scale	Information
1	Sangat Tidak Setuju (STS)
2	Tidak Setuju (TS)
3	Netral (N)
4	Setuju (S)
5	Sangat Setuju (SS)

Table 2
Instrumen SUS (Ardhana, 2022)

No.	Instrumen	Skor
1	Saya berpikir akan menggunakan fitur Meta AI di WhatsApp	1-5
2	Saya merasa fitur Meta AI ini rumit digunakan	1-5

3	Saya merasa fitur Meta AI ini mudah digunakan	1-5
4	Saya membutuhkan bantuan orang lain atau teknisi dalam menggunakan fitur ini.	1-5
5	Saya merasa Fitur Meta AI ini berjalan dengan semestinya	1-5
6	Saya merasa ada banyak hal yang tidak konsisten pada fitur ini	1-5
7	Saya merasa orang lain akan memahami cara menggunakan fitur ini dengan cepat	1-5
8	Saya merasa fitur Meta AI ini membingungkan, terkadang tanyanya apa jawabnya apa	1-5
9	Saya merasa percaya diri menggunakan fitur, karna membantu saya mencari materi kuliah	1-5
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan fitur ini.	1-5

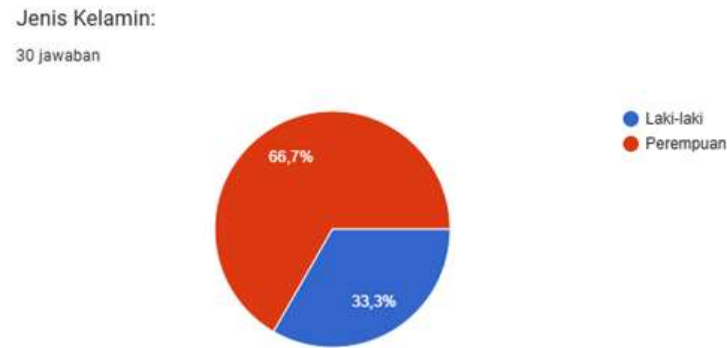
To calculate the average SUS score from the questionnaire responses, it is rated on a scale of 0 to 5. The calculation method for the SUS score is as follows (Hudaya & Nurmiati, 2023):

1. Odd-numbered questions (1, 3, 5, 7, 9) are weighted positively; 1 point is deducted from the respondent's score. Example: A score of 4 for question 9 results in a score of 3 (19 - 1).
2. Even-numbered questions (2, 4, 6, 8, 10) are weighted negatively; 5 points are deducted from the respondent's score. Example: A score of 2 for question 8 results in a score of 3 (5 - 19).
3. The sum of all individual scores is multiplied by 2.5 to obtain a total score between 0 and 100.

3. Result and Discussion

After distributing questionnaires, researchers successfully collected a sample of 30 students in Indonesia from various regions. The following is a demographic overview of the respondents based on the questionnaires collected:

Figure 2
Respondent Demographics by Gender



Based on the results of data processing, the distribution of respondents by gender is shown in the following details: 10 male respondents with a percentage of 33.3%, while the completion of this questionnaire was dominated by 20 female respondents with a percentage of

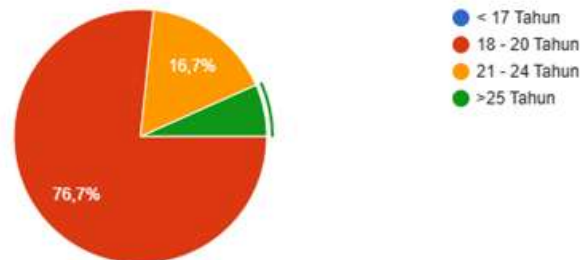
66.7% of the total number of respondents of 30 people, 3 of whom were not users of the Meta AI feature on WhatsApp.

Figure 3

Respondent Demographics by Age

Usia Anda saat ini (dalam tahun):

30 jawaban



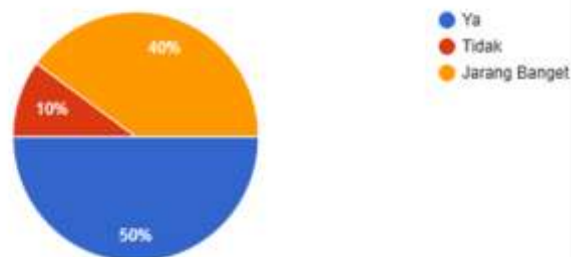
Based on the data collection results, the characteristics of respondents by age range showed quite diverse variations. Respondents in the 18-20 age category were the most dominant group, with a percentage reaching 76.7% (23 people). Furthermore, respondents in the 21-24 age category were recorded at 16.7% (5 people). Meanwhile, respondents aged 25 and over had the smallest percentage, at 6.7% (2 people) of the total 30 respondents, 3 of whom were not users of the Meta AI feature on WhatsApp.

Figure 4

Demographics of Feature User Respondents

Apakah Kamu Pernah Menggunakan Fitur Meta AI di WhatsApp?

30 jawaban



Based on the data collection results, respondents' experience with the Meta AI feature on WhatsApp varied considerably. Of the 30 respondents, 50% (15) stated they had used the feature (category "Yes"). Furthermore, 40% (12) of the respondents fell into the "Rarely Use" category. Meanwhile, 10% (3) of the respondents stated they had never used the Meta AI feature at all (category "No").

The first data was also obtained from the questionnaire results, specifically from the spreadsheet (Google Form) in Table 3.

Table 3
Questionnaire Results

Responden	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
R1	1	1	4	1	1	4	5	5	1	1
R2	3	4	3	2	3	4	3	2	2	3
R3	2	2	4	1	2	2	4	5	1	2
R4	3	4	4	1	3	2	4	2	3	1
R5	2	3	3	4	4	5	3	2	3	5
R6	3	2	4	1	3	2	4	2	2	4
R7	2	3	3	2	3	3	4	5	1	4
R8	2	2	4	2	3	2	4	5	3	4
R9	3	1	5	1	5	1	5	1	5	5
R10	2	5	2	3	2	2	2	5	1	3
R11	3	2	3	2	3	4	4	5	2	2
R12	2	2	4	2	2	5	4	5	1	2
R13	3	4	4	1	1	4	3	5	3	2
R14	4	2	4	2	2	3	3	3	3	3
R15	1	5	3	2	2	5	3	3	2	4
R16	3	2	4	1	4	3	3	4	3	3
R17	4	3	4	4	4	4	5	4	5	4
R18	3	3	4	4	4	3	4	4	4	4
R19	3	2	4	4	4	2	3	3	3	4
R20	2	3	4	2	2	4	4	4	3	4
R21	3	2	5	1	5	1	5	4	4	1
R22	3	2	5	2	3	3	4	5	2	2
R23	1	2	5	1	3	2	5	5	3	1
R24	2	3	4	1	3	2	3	4	1	2
R25	1	2	5	1	3	3	4	3	3	3
R26	5	4	4	5	2	5	3	5	4	5
R27	4	2	4	1	2	2	3	4	2	2

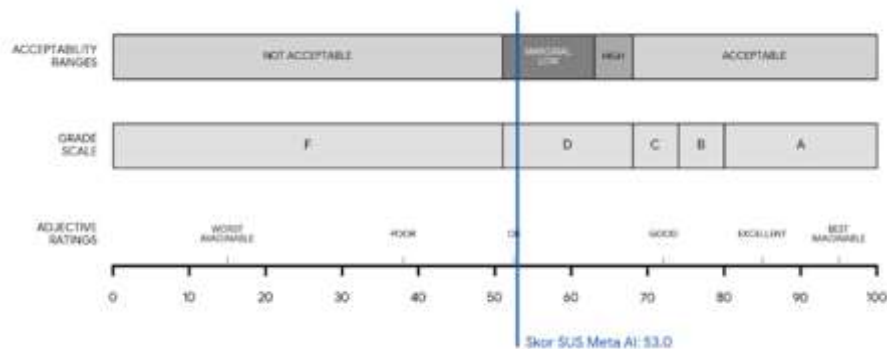
After obtaining questionnaire data from all respondents, the data was then processed using the standard System Usability Scale (SUS) calculation method. The data processing steps were as follows:

1. Odd-numbered instruments (1, 3, 5, 7, 9) are positive, and the respondent's assessment score is subtracted by 1 point. For example, if the respondent scored 4 for instrument number 9, the resulting score is 3 (4 - 1).
2. Even-numbered instruments (2, 4, 6, 8, 10) are negative, and 5 points are subtracted from the respondent's assessment score. For example, if the respondent scored 2 for instrument number 8, the resulting score is 3 (5 - 2).
3. After all scores have been determined, the sum of the scores is multiplied by 2.5 to obtain a range of 0-100.

Table
SUS Calculation Results

Responden	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	Jumlah	Total (*2,5)
R1	0	4	3	4	0	1	4	0	0	4	20	50
R2	2	1	2	3	2	1	2	3	1	2	19	47,5
R3	1	3	3	4	1	3	3	0	0	3	21	52,5
R4	2	1	3	4	2	3	3	3	2	4	27	67,5
R5	1	2	2	1	3	0	2	3	2	0	16	40
R6	2	3	3	4	2	3	3	3	1	1	25	62,5
R7	1	2	2	3	2	2	3	0	0	1	16	40
R8	1	3	3	3	2	3	3	0	2	1	21	52,5
R9	2	4	4	4	4	4	4	4	4	0	34	85
R10	1	0	1	2	1	3	1	0	0	2	11	27,5
R11	2	3	2	3	2	1	3	0	1	3	20	50
R12	1	3	3	3	1	0	3	0	0	3	17	42,5
R13	2	1	3	4	0	1	2	0	2	3	18	45
R14	3	3	3	3	1	2	2	2	2	2	23	57,5
R15	0	0	2	3	1	0	2	2	1	1	12	30
R16	2	3	3	4	3	2	2	1	2	2	24	60
R17	3	2	3	1	3	1	4	1	4	1	23	57,5
R18	2	2	3	1	3	2	3	1	3	1	21	52,5
R19	2	3	3	1	3	3	2	2	2	1	22	55
R20	1	2	3	3	1	1	3	1	2	1	18	45
R21	2	3	4	4	4	4	4	1	3	4	33	82,5
R22	2	3	4	3	2	2	3	0	1	3	23	57,5
R23	0	3	4	4	2	3	4	0	2	4	26	65
R24	1	2	3	4	2	3	2	1	0	3	21	52,5
R25	0	3	4	4	2	2	3	2	2	2	24	60
R26	4	1	3	0	1	0	2	0	3	0	14	35
R27	3	3	3	4	1	3	2	1	1	3	24	60
Rata-rata												53,0

The questionnaire responses were processed using the standard System Usability Scale (SUS) scoring procedure. Of the 30 survey respondents, 27 had used the Meta AI feature in WhatsApp and were included in the SUS calculation. Based on the individual scores presented in the SUS calculation table, the cumulative SUS score was 1,432.5, yielding a mean score of 53.06 across the 27 respondents. Rounded to one decimal place, the average SUS score was 53.1. The following are the scores from the System Usability Scale (SUS) questionnaire regarding the interpretation of the three methods (Evy Nurmiati & Saputri, 2024).

Figure 5*SUS Score position regarding interpretation*

The mean SUS score was interpreted using three dimensions: acceptability ranges, grade scale, and adjective ratings. Based on the interpretation framework used in this study, a score of 53.1 falls within the *Marginal Low* acceptability range, corresponds to Grade D, and receives an *Okay* adjective rating. These classifications represent distinct dimensions of SUS interpretation and should not be treated as interchangeable.

The results indicate relatively low perceived usability of the Meta AI feature among the surveyed respondents. However, these findings reflect the perceptions of the university students who participated in this study and should not be generalized to all WhatsApp users or interpreted as evidence of the overall acceptance of Meta AI.

These results are consistent with previous studies that have tested the usability of AI-based chatbot tools using SUS scores. For example, an evaluation of the ChatGPT platform also yielded a score below the acceptable threshold (67.44), indicating that AI-integrated features generally require significant improvements before they can be considered fully usable by the general public (Mulia et al., 2023). Similarly, a study of the Indonesian government application MyASN obtained a SUS score of 54.08, a value very close to the results of this study (53.0), confirming that scores in this range consistently indicate the need for fundamental system overhaul (Anam et al., 2025).

The low usability scores (SUS) achieved in this study can also be analyzed in terms of user motivation and technology acceptance. A study on the continued use of meta-AI on WhatsApp showed that hedonistic benefits and social appeal were the main motivating factors for use, while ease of use remained a point of criticism (Biru et al., 2025). This suggests that users are more likely to be interested in meta-AI out of curiosity or due to social influence than out of genuine satisfaction with its functionality, which explains the low usability scores observed in this study.

Furthermore, a study applying the Technology Acceptance Model (TAM) to students' use of AI chatbots showed that perceived usability plays a crucial role in the continued use of AI tools (Arisyi et al., 2025). The SUS score of 53.0 achieved in this study – in which respondents expressed confusion regarding the consistency and learnability of the meta-AI functions – is directly related to the low perceived usability, a key factor limiting wider acceptance among Indonesian students.

A more detailed analysis of individual responses on the SUS instruments revealed striking trends. The instruments for ease of use (I3) and learning from others (I7) scored relatively high, indicating that users found meta-AI functionality readily accessible within the familiar WhatsApp interface. In contrast, the instruments for consistency (I6) and the need to familiarize themselves with the tool before use (I10) performed worse. This highlights issues with predictability and response accuracy, particularly regarding the chatbot's interpretation and response to user questions. This discrepancy suggests that while the feature's accessibility is satisfactory on the surface, its deeper functional reliability remains a weakness.

Demographically, respondents aged 18 to 20, who comprised 76.7% of the sample, are digital natives who regularly use AI-based tools across various platforms. Despite their familiarity with the technology, this group achieved below-average SUS scores. This suggests that the lack of ease of use is not primarily due to a lack of digital skills, but rather to systemic limitations in functional design and response quality. This finding underscores the importance of assessing ease of use independently of users' general technological capabilities, as even technically proficient users can struggle with poorly designed AI interfaces (Deshmukh & Chalmeta, 2024)

4. Conclusion

This study evaluated the perceived usability of the Meta AI feature in WhatsApp using the System Usability Scale (SUS). Based on the responses of 27 users included in the analysis, the mean SUS score was 53.1, which is below the commonly cited benchmark of 68. Under the interpretation framework used in this study, the score falls within the Marginal Low acceptability range, corresponds to Grade D, and receives an Okay adjective rating. These results indicate relatively low perceived usability among the surveyed respondents.

The item-level results provide additional information about respondents' perceptions of specific usability aspects. Item 3 received a relatively high mean contribution score, whereas item 8 received the lowest. These findings may help identify aspects that warrant further investigation, but they do not directly establish the accuracy of AI-generated responses or the causes of usability difficulties.

The conclusions of this study are limited to the perceptions of the respondents who participated in the survey. The findings should not be generalized to all WhatsApp users or interpreted as a measurement of the overall adoption or acceptance of Meta AI. Future research could involve larger and more diverse samples and complementary evaluation methods, such as heuristic evaluation or the User Experience Questionnaire (UEQ), to obtain a more comprehensive understanding of the feature's usability.

5. Acknowledgement

The author would like to express his sincere gratitude to Ms. Evy Nurmiati and Mr. Nuryasin for their guidance, support, and valuable comments during the preparation of this article. He also extends special thanks to all parties who contributed to the distribution of the questionnaire. Finally, the author would like to thank all respondents for their time and cooperation, which made this data collection possible.

6. References

- Anam, M. K., Susanti, S., Nurjayadi, N., Zoromi, F., & Sari, A. K. (2025). Evaluation of an existing system using the System Usability Scale (SUS) as a guideline for system improvement. *Jurnal Teknik Informatika, 18*(1), 122–132. <https://doi.org/10.15408/jti.v18i1.40766>
- Anargya, L. N., Purwaamijaya, B. M., & Nugraha, M. R. (2025). Implementasi model UTAUT untuk mengidentifikasi faktor-faktor yang memengaruhi penerimaan chatbot Meta AI WhatsApp. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), 10*(4), 3851–3863. <https://doi.org/10.29100/jipi.v10i4.9337>
- Ardhana, V. Y. P. (2022). Evaluasi usability e-learning Universitas Qamarul Huda menggunakan System Usability Scale (SUS). *Journal of Informatics, Electrical and Electronics Engineering, 2*(1), 5–11. <https://doi.org/10.47065/jieeee.v2i1.430>
- Arisyi, N. D., Haykal, F. M., Setiawan, D. I., Ma'sum, R. A., & Buana, P. A. (2025). Analisis kepuasan pengguna terhadap layanan digital akademik menggunakan System Usability Scale (SUS). *Jurnal Komputer dan Teknologi Sains (KOMTEKS), 4*(1), 29–34.
- Biru, I. R., Rahmawati, R. F., & Salsabila, B. A. D. (2025). Analysis of motivation for continued use of Meta AI on WhatsApp: Uses and gratification theory approach. *Journal of Artificial Intelligence and Engineering Applications, 4*(3), 2162–2168. <https://doi.org/10.59934/jaiea.v4i3.1123>
- Caro-Álvaro, S., García-López, E., García-Cabot, A., de-Marcos, L., & Domínguez-Díaz, A. (2022). Applying usability recommendations when developing mobile instant messaging applications. *IET Software, 16*(1), 73–93. <https://doi.org/10.1049/sfw2.12039>
- Deshmukh, A. M., & Chalmeta, R. (2024). Validation of system usability scale as a usability metric to evaluate voice user interfaces. *PeerJ Computer Science, 10*, Article e1918. <https://doi.org/10.7717/peerj-cs.1918>
- Hudaya, F., & Nurmiati, E. (2023). Evaluasi usability aplikasi video conference sebagai media knowledge sharing (studi kasus: Aplikasi Zoom dan Google Meet). *JSiI (Jurnal Sistem Informasi), 10*(1), 68–74. <https://doi.org/10.30656/jsii.v10i1.5935>
- Muhamad, N. (2025, March 7). Daftar media sosial paling banyak digunakan di Indonesia awal 2025. *Databoks*. <https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/67ca734b43927/daftar-media-sosial-paling-banyak-digunakan-di-indonesia-awal-2025>
- Mulia, A. P., Piri, P. R., & Tho, C. (2023). Usability analysis of text generation by ChatGPT OpenAI using System Usability Scale method. *Procedia Computer Science, 227*, 381–388. <https://doi.org/10.1016/j.procs.2023.10.537>
- Nurmiati, E., & Irahman, M. S. (2023). Evaluasi usability aplikasi Microsoft Word menggunakan Questionnaires Nielsen's Attributes of Usability (NAU) (studi kasus: Mahasiswa Program Studi Biologi UIN Syarif Hidayatullah Jakarta tahun 2021). *Jurnal Perangkat Lunak, 5*(2), 217–223.
- Nurmiati, E., & Saputri, E. (2024). Penerapan usability pada aplikasi C-Access KAI Commuter menggunakan metode System Usability Scale. *Jurnal Perangkat Lunak, 6*(2), 275–286. <https://doi.org/10.32520/jupel.v6i2.3293>

- Pitts, G., & Motamedi, S. (2026). *What drives students' use of AI chatbots? Technology acceptance in conversational AI* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2602.20547>
- Sugiyono. (2019). *Metode penelitian pendidikan: Kuantitatif, kualitatif, kombinasi, R&D, dan penelitian pendidikan* (Edisi ke-3). Alfabeta.
- Usman, M. L. L., & Gustalika, M. A. (2022). Pengujian validitas dan reliabilitas System Usability Scale (SUS) untuk perangkat smartphone. *Jurnal Ecotipe (Electronic, Control, Telecommunication, Information, and Power Engineering), 9*(1), 19–24. <https://doi.org/10.33019/jurnalecotipe.v9i1.2805>
- Yahya, A., & Dirgantara, P. (2022). Efektivitas penggunaan WhatsApp sebagai media komunikasi organisasi di lingkungan pegawai Dinas ATR/BPN Kota Palembang. *Medium, 10*(1), 409–420. [https://doi.org/10.25299/medium.2022.vol10\(1\).9525](https://doi.org/10.25299/medium.2022.vol10(1).9525)

