Journal of Information Technology, Computer Engineering and Artificial Intelligence (2024)
Received December 2024 / Revised December 2024 / Accepted December 2024
Journal of Information Technology, Computer Engineering and Artificial Intelligence (ITCEA)
https://journal.redtechidn.org/index.php/itcea

Exploring User Experiences of Chat GPT and Google Gemini AI : A SUS and EUCS Approach

Fausta Rizky Abriansah[1], Nanda Dwi Kurniawan[2], Praditya Rendi Ferdian[3], Daffa Nurin Nabil Ma'arif[4], Azmi Maulana Firdaus[5], Pratama Angga Buana*[6]

[1-*6]*Information Systems Study Program, Faculty of Information and Communication Technology, Universitas Semarang, Soekarno Hatta Street, Semarang, Central Java, 50196, Indonesia*

Abstract. This study evaluates user experience and satisfaction of two AI platforms, Chat GPT and Google Gemini, using the System Usability Scale (SUS) and End-User Computing Satisfaction (EUCS) approaches. The study involved 78 respondents from diverse backgrounds who provided evaluations based on five main dimensions of usability (Learnability, Efficiency, Memorability, Error, Pleasantness) and five dimensions of user satisfaction (Content, Accuracy, Format, Ease of Use, Timeliness). The results show that both platforms achieved a "Satisfied" category with an average usability and satisfaction score above 3.75. The lowest scores were recorded in Error and Timeliness variables, while Learnability and Format showed the highest scores. This study also revealed a significant positive correlation between usability and user satisfaction, providing recommendations for improvements in Error and Timeliness areas. These findings offer valuable insights for developing more user-friendly and responsive AI platforms in the future.

Keywords : Chat GPT; EUCS; Google Gemini; SUS; Usability; User Satisfaction

1. Introduction

The development of artificial intelligence (AI) technology has had a significant impact on various aspects of human life, particularly in interactions between humans and digital systems ((Rosidin, et al., 2024); (Farwati, et al., 2023)). One of the latest innovations in this field is the emergence of AI-based chatbots, which not only provide automated responses but also interact naturally with users. Two leading platforms in this area, ChatGPT and Google Gemini, have garnered attention from diverse groups, ranging from IT professionals to the general public (Heryanto, et al., 2024). Although both platforms offer advanced natural language processing capabilities, user experience (UX) in interacting with these platforms remains a critical aspect of evaluating their effectiveness and user satisfaction.

With the increasing complexity and sophistication of AI platforms, understanding the quality of user experience has become crucial. Studies on UX evaluation in the context of AI have been conducted extensively; however, they are often limited to comparisons between ChatGPT and Google Gemini in terms of specific parameters such as usability, efficiency, and overall user satisfaction. Previous research, such as the study by Tiara & Pamuji (2024), contributed to measuring the performance comparison of the two platforms using usability testing methods. Their findings revealed significant differences in efficiency and user satisfaction, with Gemini AI excelling in learnability and satisfaction. However, this study primarily employed

[1]*Corresponding author's email : pratama_angga@usm.ac.id

quantitative methods and lacked an in-depth exploration of factors influencing user experience in broader contexts.

Innovations in chatbots, such as ChatGPT, have demonstrated the potential for interactive engagement with users, making it easier for individuals to acquire knowledge and complete various tasks. In response, Google developed the Bard chatbot (Gemini) to compete with OpenAI's ChatGPT. Gemini, an evolution of Bard, offers advanced features that enhance Google's AI capabilities in text, voice, and image-based interactions (Nasrul, et al., 2024).

Today, human reliance on AI chatbots has grown significantly. Thus, understanding user satisfaction with these platforms' performance is critical. A usability approach evaluates product usability from users' perceptions, encompassing dimensions such as learnability, efficiency, memorability, error rate, and satisfaction. Additionally, the End-User Computing Satisfaction (EUCS) framework assesses content, accuracy, format, ease of use, and timeliness, aiming to measure user satisfaction (Buana, et al., 2024).

The research problems addressed in this study are : (1) How do users perceive the usability and satisfaction of ChatGPT and Google Gemini platforms ? (2) What factors influence user experience that can serve as a reference for improving AI platform quality ? The objectives of this study are to analyze the user experience of ChatGPT and Google Gemini by measuring satisfaction and usability through an empirical approach, comparing the performance of the two platforms based on proven usability and user satisfaction parameters. Primary data will be collected through a questionnaire designed to capture user perceptions, which will then be analyzed using statistical methods to gain deeper insights into the factors affecting user experience.

The methodology involves a field survey with 78 respondents from a student background. The questionnaire used is based on validated instruments such as the System Usability Scale (SUS) and End-User Computing Satisfaction (EUCS). The results of this study are expected to provide a clearer picture of the quality of user interaction with AI platforms and offer practical recommendations for developing AI technology that is more responsive to user needs.

The novelty of this research lies in its comprehensive approach, combining quantitative and qualitative methods to evaluate user experience. Unlike previous studies that focused on a single performance aspect (such as efficiency or learnability), this research aims to explore various dimensions of user experience holistically and provide data-driven recommendations for AI platform development. Thus, this study contributes not only to software quality testing but also to the development of more user-friendly AI platforms.

2. Methods

This research employs a quantitative approach to evaluate the user experience (UX) and user satisfaction of ChatGPT and Google Gemini AI applications. The research methodology is divided into four main stages: preparation, planning, data collection, and data analysis.

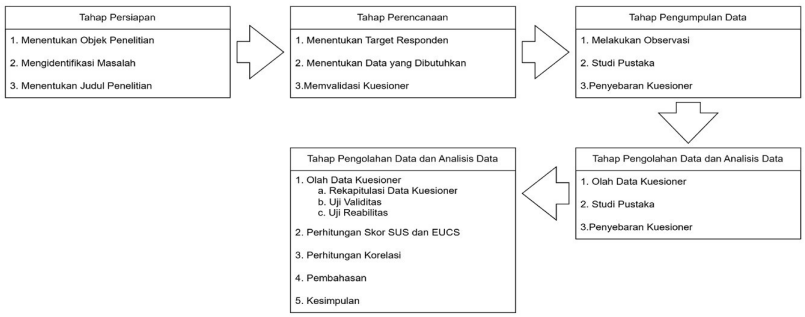

Figure 1. Metode Penelitian.

2.1. Preparation Stage

The researchers identified the objects of study, ChatGPT and Google Gemini AI applications. Respondents were selected based on the criterion that they had prior experience using both applications. The study focuses on understanding user perceptions of usability and satisfaction for these platforms.

2.2. Planning Stage

Respondents were chosen using a random sampling technique with a 5% margin of error (Fahlefi, et al., 2023). The questionnaire was designed based on two primary evaluation methods : (1) System Usability Scale (SUS) : Measures learnability, efficiency, memorability, error, and pleasantness. (2) End-User Computing Satisfaction (EUCS): Assesses content, accuracy, format, ease of use, and timeliness. The questionnaire used in this study consists of 30 items, with 15 questions for the SUS method and 15 questions for the EUCS method. Tables 1 present the list of questions for each questionnaire method used in this study.

Table 1. The List Of Questions For Each Questionnaire Method

Metode	Aspek	ID	Pertanyaan
SUS	Learnability	L1	Apakah Anda merasa mudah dalam menggunakan AI ChatGPT / Gemini ?
		L2	Apakah informasi yang diberikan oleh AI ChatGPT / Gemini mudah dipahami ?
		L3	Seberapa mudah Anda memahami alur percakapan dalam AI ChatGPT / Gemini ?
	Efficiency	E1	Seberapa cepat Anda dapat mendapatkan jawaban atau informasi yang dibutuhkan dari AI ChatGPT / Gemini ?
		E2	Apakah Anda merasa bahwa AI ChatGPT / Gemini memberikan informasi dengan cepat dan efisien ?
		E3	Seberapa mudah Anda dapat menemukan jawaban yang Anda cari dengan AI ChatGPT / Gemini ?
	Memory	M1	Seberapa mudah Anda mengingat fungsi tombol pada AI ChatGPT / Gemini setelah pertama kali mencobanya ?
		M2	Apakah Anda merasa nyaman saat menggunakan ChatGPT / Gemini kapanpun diperlukan ?
		M3	Apakah Anda merasa dapat dengan mudah mengingat percakapan yang telah dibahas sebelumnya dengan AI ChatGPT / Gemini ?
	Error	ER1	Apakah Anda menemukan kesalahan atau masalah teknis selama menggunakan AI ChatGPT / Gemini ?
		ER2	Apakah Anda merasa bahwa AI ChatGPT / Gemini jarang memberikan informasi yang salah atau tidak relevan ?
		ER3	Apakah Anda merasa AI ChatGPT / Gemini selalu memberikan informasi yang sesuai dengan konteks yang Anda ajukan ?
	Pleasant	P1	Apakah tampilan atau desain antarmuka ChatGPT / Gemini menarik dan nyaman dilihat ?
		P2	Apakah Anda merasa nyaman menggunakan ChatGPT / Gemini karena desain dan tata letak yang baik ?
		P3	Seberapa baik icon yang digunakan pada ChatGPT / Gemini dalam pengalaman Anda ?
EUCS	Content	C1	Apakah informasi yang diberikan oleh AI ChatGPT / Gemini cukup mudah dipahami ?
		C2	Seberapa baik informasi yang disediakan oleh ChatGPT / Gemini dalam membantu Anda memecahkan masalah atau menjawab pertanyaan ?
		C3	Apakah jawaban yang diberikan oleh ChatGPT / Gemini membuat Anda tertarik untuk menggali lebih lanjut tentang topik yang dibahas ?

Accuracy	A1	Seberapa akurat informasi yang diberikan oleh AI ChatGPT / Gemini terkait dengan topik yang Anda tanyakan ?
	A2	Apakah jawaban yang diberikan dalam bentuk visual (grafik, gambar, dll.) oleh AI ChatGPT / Gemini akurat dan sesuai dengan yang Anda harapkan ?
	A3	Apakah informasi yang diberikan terkait dengan harga, waktu, atau fakta lainnya sesuai dengan kenyataan ?
Format	F1	Seberapa nyaman Anda membaca teks yang ditampilkan oleh AI ChatGPT / Gemini ?
	F2	Apakah format teks (ukuran, font) yang digunakan dalam percakapan mudah dibaca dan nyaman di mata ?
	F3	Apakah gambar atau elemen visual yang ditampilkan oleh ChatGPT / Gemini jelas dan mudah dipahami ?
Ease Of Use	EU1	Seberapa mudah Anda mengakses URL ChatGPT / Gemini ?
	EU2	Apakah navigasi atau antarmuka di dalam ChatGPT / Gemini mudah dipahami dan mudah digunakan ?
	EU3	Apakah Anda merasa nyaman dan terbantu ketika menggunakan fitur-fitur tambahan (seperti pencarian, saran, atau menu navigasi) di ChatGPT / Gemini ?
Timeliness	T1	Seberapa cepat ChatGPT / Gemini merespons saat pertama kali Anda mengajukan pertanyaan atau permintaan ?
	T2	Apakah Anda pernah mengalami jeda waktu atau penundaan dalam mendapatkan jawaban atau informasi dari ChatGPT / Gemini ?
	T3	Apakah transisi antar topik atau percakapan berlangsung lancar tanpa hambatan atau gangguan ?

2.3. Data Collection Stage

The questionnaire was distributed via social media and discussion groups to reach 78 respondents who met the criteria. Respondents included university students and members of the general public.

2.4. Data Analysis Stage

Collected data were analyzed using : (1) Validity and Reliability Testing : Ensuring the accuracy and consistency of the questionnaire using Pearson Product Moment and Cronbach's Alpha methods. (2) Descriptive Statistics : Calculating the mean scores for each aspect of usability and satisfaction. (3) Correlation Analysis : Identifying the relationship between usability and user satisfaction. The analysis results were compared against a Likert scale (1-5), categorizing satisfaction levels from "Very Dissatisfied" to "Very Satisfied.

3. Results and Discussion

This study involved 78 respondents selected using a random sampling technique with a 5% margin of error. Respondents included university students and members of the general public who had experience using ChatGPT and Google Gemini AI. Data analysis assessed usability using the System Usability Scale (SUS) and satisfaction using the End-User Computing Satisfaction (EUCS) method. Describing the characteristics of respondents based on three main categories: age, gender, and occupation.

3.1. Description Based on Age

Figure 2 shows that the majority of respondents were aged 19–23 years, accounting for 89% of the total respondents. Respondents aged 15–18 years made up 7%, while those aged over 24 years constituted only 4%.

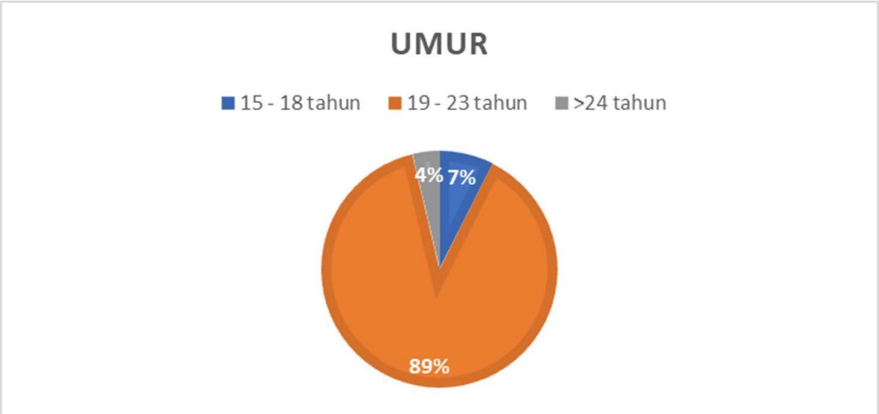

Figure 2. Respondent Diagram Based on Age

3.2. Description Based on Gender

Figure 3 indicates that male respondents were slightly more dominant, with a percentage of 52%. Meanwhile, female respondents comprised 48% of the total respondents.

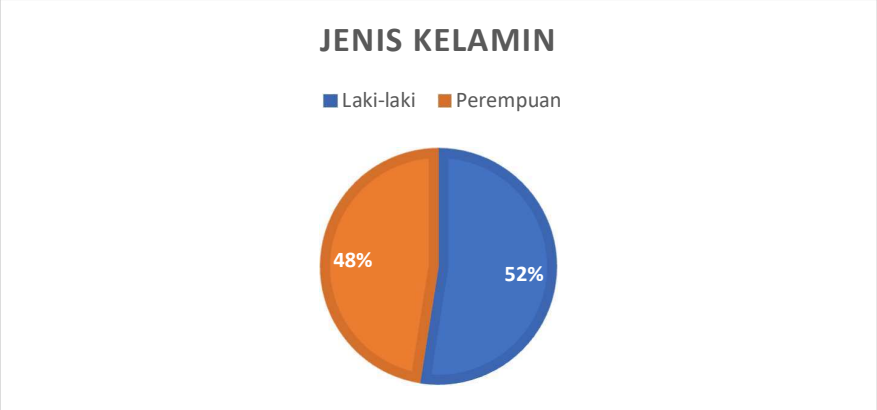

Figure 3. Respondent Diagram Based on Gender

3.3. Description Based on Occupation

Figure 4 demonstrates that the majority of respondents were students, making up 91% of the total percentage.

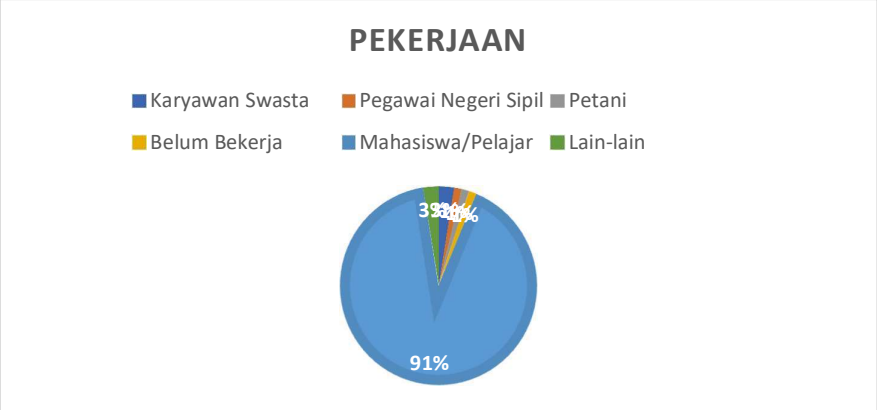

Figure 4. Respondent Diagram Based on Occupation

3.4. Validity Test Results

The validity test is conducted to evaluate the extent to which the research instrument can measure the intended variables through the distribution of questionnaires (Dewiyani & Fadila, 2023). In this study, the validity test aims to ensure that each question item in the questionnaire is relevant and accurate in measuring the variables that are the focus of the research.

Table 2. Validity Test Results ChatGPT and Gemini

Aspek	ID	R-Hitung ChatGPT	R-Hitung ChatGPT	R-Tabel	Status
Learnability	L1	0.792	0.754		
	L2	0.676	0.763		
	L3	0.797	0.697		
Efficiency	E1	0.672	0.732		
	E2	0.582	0.694		
	E3	0.678	0.774		
Memory	M1	0.568	0.754		
	M2	0.565	0.665		
	M3	0.646	0.597		
Error	ER1	0.317	0.465		
	ER2	0.274	0.410		
	ER3	0.498	0.551		
Pleasant	P1	0.626	0.679		
	P2	0.743	0.699		
	P3	0.560	0.630	0.22	Valid
Content	C1	0.738	0.754		
	C2	0.580	0.746		
	C3	0.741	0.632		
Accuracy	A1	0.653	0.736		
	A2	0.606	0.506		
	A3	0.540	0.687		
Format	F1	0.722	0.781		
	F2	0.693	0.753		
	F3	0.732	0.647		
Ease Of Use	EU1	0.658	0.694		
	EU2	0.671	0.596		
	EU3	0.607	0.672		
Timeliness	T1	0.700	0.658		
	T2	0.685	0.642		
	T3	0.652	0.681		

Based on the results in Table 2, all variables show a calculated r value greater than the r table value. Therefore, it can be concluded that all variables in this study are considered valid

3.5. Reliability Test Results

The reliability test is conducted to assess the consistency and stability of the research instrument when used under the same or similar conditions. A reliable instrument will produce trustworthy data and provide consistent results if the test is repeated. This test uses Cronbach's Alpha method to measure the internal consistency of the questionnaire, which consists of many

questions. The Cronbach's Alpha value ranges from 0 to 1, where a value greater than 0.6 indicates that the instrument has good reliability (Anggraini, et al., 2022), as shown in Tables 3.

Table 3. Reliability Test Results Chat GPT and Gemini

Aspek	R Hitung ChatGPT	R Hitung Gemini	R Tabel	Status
Learnability	0.943	0.922	0.6	Reliabel
Efficiency	0.942	0.942		
Memory	0.930	0.928		
Error	0.918	0.901		
Pleasant	0.923	0.926		
Content	0.952	0.939		
Accuracy	0.914	0.938		
Format	0.916	0.948		
Ease of Use	0.942	0.933		
Timeliness	0.932	0.934		

Table 3 shows that each variable in ChatGPT aand Gemini is considered reliable because it has a Cronbach's Alpha value greater than 0.6. The scores obtained from respondents' answers are then calculated and averaged to align with the positive measurement scale, allowing them to be compared with satisfaction levels and categories, as displayed in Table 4.

Table 4. Satisfaction Result ChatGPT and Gemini

Metode	Variabel	Mean Chat GPT	Mean Gemini	Level	Tingkat Kepuasan
SUS	Learnability	4.10	3.94	4	Puas
	Efficiency	4.03	3.88	4	Puas
	Memory	3.94	3.79	4	Puas
	Error	3.41	3.52	4	Puas
	Pleasant	3.80	3.82	4	Puas
EUCS	Content	3.91	3.80	4	Puas
	Accuracy	3.57	3.73	4	Puas
	Format	3.95	3.92	4	Puas
	Ease of Use	4.09	3.86	4	Puas
	Timeliness	3.85	3.82	4	Puas

Based on the results of the analysis and data processing presented in Table 8 and Table 9, the explanation is as follows : The usability test results using the SUS method, based on user experience, show that the variable with the lowest score is Error, with an average score of 3.41, although it still falls within the "Satisfied" category in terms of user satisfaction. Meanwhile, the highest-scoring variable is Learnability, which refers to the user's ability to learn how to use the application for the first time, with a score of 4.10. Other variables, in order, are Efficiency (the speed and ease of user interaction with the application) with a score of 4.03, Memorability (the user's ability to remember and use the application) with a score of 3.94, and Pleasantness (the comfort of the user while using the application) with a score of 3.80.

The EUCS method shows that the variables have relatively good average values, with all scores above 3.70. The Accuracy variable, which measures the accuracy of the information, has the lowest score of 3.73, followed by Content with a score of 3.80, reflecting the user's experience in obtaining relevant and quality information. The Timeliness variable, measuring the promptness and speed of access, scored 3.82, while Ease of Use, which reflects the ease of

using the system, received a score of 3.86. The highest-scoring variable is Format, which includes layout and visual design, with a score of 3.92. Based on these results, user satisfaction is in the "Satisfied" category, corresponding to level 4.

4. Conclusions

This study successfully analyzed user experience and satisfaction levels from two artificial intelligence applications, ChatGPT and Google Gemini AI. The methods used include the System Usability Scale (SUS) and End-User Computing Satisfaction (EUCS). After analyzing data from 78 participants, both applications received ratings that place them in the "Satisfied" category for all evaluated variables, both through the SUS and EUCS methods. ChatGPT recorded the highest score for the Ease of Use variable (4.09) and the lowest for the Error variable (3.41), while Google Gemini AI scored the highest for the Learnability variable (3.94) and the lowest for the Error variable (3.52). However, the Error variable score in both applications still falls within the "Satisfied" category.

This study also indicates a significant positive relationship between usability and user satisfaction, suggesting that improvements in usability have a substantial impact on user satisfaction. Overall, ChatGPT slightly outperforms Google Gemini AI in most variables, particularly in Efficiency and Ease of Use, while Google Gemini AI demonstrates competitive performance in the Learnability variable.

The Error and Timeliness aspects need further attention in order to elevate satisfaction levels from the "Satisfied" category to "Very Satisfied." Both applications should work to reduce errors and improve response speed to provide a better user experience. This study offers practical insights for AI app developers to understand key elements influencing user experience, as well as specific areas that still require improvement. With an evidence-based approach, this research highlights the importance of balancing usability and user satisfaction in designing effective and user-friendly AI applications.

Acknowledgments

With deep gratitude, the author expresses sincere thanks to Mr. Pratama Angga Buana, S.Kom., M.Kom., the supervising lecturer, for all the guidance, direction, and motivation provided during the research and article writing process. His assistance and support were invaluable in completing this work successfully. The author would also like to thank the beloved alma mater, Universitas Semarang, for providing the facilities and support in the development of knowledge, as well as the opportunity to develop personal potential. Additionally, the author extends heartfelt appreciation to all co-authors and individuals who contributed, either directly or indirectly, to the writing of this article. May all the help provided be a source of good deeds that benefit everyone. It is hoped that this article will provide value to readers and contribute positively to the world of education and research.

References

Anggraini, F. D. P., Aprianti, A., Setyawati, V. A. V. & Hartanto, A. A., 2022. Pembelajaran Statistika Menggunakan Software SPSS untuk Uji Validitas dan Reliabilitas. JURNAL BASICEDU , 6(4), pp. 6491 - 6504.

Buana, P. A., Wicaksana, D. A. & Adinugroho, S., 2024. Pengukuran Kegunaan dan Kepuasan Pengguna pada Aplikasi Virtual Tour Reality dalam Meningkatkan Pengalaman Wisata. *Jurnal Algoritma,* 21(2), pp. 101-111.

Dewiyani, A. I. C. & Fadila, R., 2023. Uji Validitas dan Reliabilitas Kuesioner Pengetahuan Masyarakat Tentang Program JKN. *Jurnal Kesehatan Qamarul Huda,* 11(1), pp. 307-315.

Fahlefi, I., Amin, S. & Fitriaty, F., 2023. Model Kinerja Melalui Gaya Kepemimpinan Transformasional Dimediasi oleh Kepuasan Kerja dan Motivasi Pegawai pada Badan Keuangan Daerah Kabupaten Batang Hari. *JIMT : Jurnal Ilmu Manajemen Terapan,* 4(6), pp. 778-789.

Farwati, M., Salsabila, I. T., Navira, K. R. & Sutabri, T., 2023. ANALISA PENGARUH TEKNOLOGI ARTIFICIAL INTELLIGENCE (AI) DALAM KEHIDUPAN SEHARI-HARI. *Jursima : Jurnal Sistem Informasi & Manajemen,* 11(1), pp. 39-45.

Heryanto, Y., Fauziah, F., Farahdinna, F. & Wijanarko, S., 2024. Evaluasi Responsivitas dan Akurasi: Perbandingan Kinerja ChatGPT dan Google BARD dalam Menjawab Pertanyaan seputar Python. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK) ,* 9(1), pp. 248-256.

Nasrul, I., Anggraini, A., Hamzah, M. L. & Saputra, E., 2024. Analysis Of User Experience Of ChatGPT And Gemini Users Using The User Experience Quistionnaire (UEQ) For Education. *Jurnal SISFOKOM (Sistem Informasi dan Komputer),* 13(3), pp. 352-359.

Rosidin, R. et al., 2024. PERAN KECERDASAN BUATAN DALAM PENGEMBANGAN SISTEM OTOMATISASI PROSES BISNIS. *Jurnal Review Pendidikan dan Pengajaran ,* 7(3), pp. 9320-9329.

Tiara, T. & Pamuji, F. Y., 2024. KOMPARASI USABILITY CHATGPT VS GEMINI AI BERDASARKAN ISO/IEC 9126 DAN NIELSEN MODEL MENGGUNAKAN METODE USABILITY TESTING. *Jusim : Jurnal Sistem Informasi Musi Rawas ,* 9(1), pp. 89-100.