

Usability Testing Pada PUSADBOT Menggunakan Black-Box dan System Usability Scale (SUS)

Salim Nurul Umam¹, Raden Bagus Bambang Sumantri² Retno Agus setiawan³
Program Studi Sistem Informasi, Universitas Harapan Bangsa, Purwokerto ¹²³
Email: salimnurulumam852@gmail.com

Received 23 Mei 2023; Revised -; Accepted for Publication 26 Mei 2023; Published 08 Juni 2023

Abstract — Chatbot is defined as a computer program designed to simulate conversations with human users based on Natural Language Processing (NLP). NLP computing has the ability to study communication between humans and computers through natural language. This study aims to evaluate the usability of PUSADBOT to see the level of effectiveness and user experience provided. Usability testing is carried out to evaluate the usability of PUSADBOT. System evaluation is carried out using the Blackbox and System Usability Scale (SUS) to determine the level of feasibility and usability of the chatbot. The results of usability testing on PUSADBOT which was carried out using the Black-Box method and the System Usability Scale (SUS) went well.

Keywords — Chatbot, Bblack-box, System Usability Scale (SUS).

Abstrak—Chatbot didefinisikan sebagai program komputer yang dirancang untuk mensimulasikan percakapan dengan pengguna manusia dengan berbasis Natural Language Processing (NLP). Komputasi NLP memiliki kemampuan mempelajari komunikasi antara manusia dengan komputer melalui bahasa alami. Penelitian ini bertujuan untuk melakukan evaluasi usability pada PUSADBOT untuk melihat tingkatan efektifitas dan pengalaman pengguna yang diberikan. Usability testing dilakukan untuk mengevaluasi usability dari PUSADBOT. Evaluasi sistem dilakukan menggunakan Blackbox dan System Usability Scale (SUS) untuk mengetahui tingkat kelayakan dan kegunaan dari chatbot. Hasil usability testing pada PUSADBOT yang telah dilakukan menggunakan metode Black-Box dan System Usability Scale (SUS) berjalan dengan baik.

Kata Kunci—Chatbot, Bblack-box, System Usability Scale (SUS).

PENDAHULUAN

Dewasa ini perkembangan teknologi Artificial Intelligence (AI) berkembang begitu cepat dan banyak diaplikasikan pada berbagai bidang. Salah satu teknologi AI yang populer digunakan adalah Chatbot [1]. Chatbot didefinisikan sebagai program komputer yang dirancang untuk mensimulasikan percakapan dengan pengguna manusia dengan berbasis Natural Language Processing (NLP). Komputasi NLP memiliki kemampuan mempelajari komunikasi antara manusia dengan komputer melalui bahasa alami. Model komputasi seperti ini berguna untuk memudahkan komunikasi antara manusia dengan komputer dalam hal pencarian informasi, sehingga dapat terjadi suatu interaksi antara keduanya dengan menggunakan bahasa alami [2].

Penerapan bahasa alami (*Natural Language*) pada chatbot memungkinkan interaksi se-natural mungkin, dan mendapatkan banyak perhatian dari berbagai industri. Chatbot tidak hanya memungkinkan cara yang lebih cepat dan lebih

alami untuk mengakses informasi, tetapi juga akan menjadi faktor kunci dalam proses memanasikan mesin dalam waktu dekat [3].

Usability didefinisikan sebagai sejauh mana suatu program dapat digunakan untuk mencapai tujuan terukur dengan efektivitas, efisiensi, dan kepuasan dalam konteks penggunaan tertentu [4]. Usability adalah aspek penting dalam sistem perangkat lunak interaktif dan karenanya penting untuk menggabungkan kegunaan dalam chatbot, untuk meningkatkan pengalaman pengguna.

Chatbot telah meresap dan digunakan di banyak area, seperti pemesanan semua jenis layanan, belanja online, dan untuk mendapatkan nasihat medis [5][6][7]. PUSADBOT merupakan salah satu bot pintar yang dikembangkan untuk meningkatkan sistem pelayanan puskesmas. Bot ini memungkinkan pengguna berinteraksi langsung dengan bot komputer untuk memudahkan pencarian informasi terkait pelayanan yang diberikan pada puskesmas.

Chatbot bukanlah konsep baru. Banyak perusahaan telah menggunakan chatbots untuk berinteraksi dengan sistem yang sudah matang. Sementara industri mendapatkan banyak pengalaman dalam desain chatbot, akan tetapi hanya ada sedikit penelitian yang membahas usability chatbot secara terintegrasi dan formal [3]. Beberapa penelitian tentang usability chatbot pernah dilakukan, di antaranya Arief [8] menguji efektifitas chatbot pembelajaran myHardware menggunakan System Usability Scale (SUS). Hasil pengujian SUS menunjukkan chatbot dapat berfungsi dengan baik dan dapat diterima oleh siswa sebagai pengguna. Julia Saenz [9] menggunakan SUS untuk menguji beberapa platform chatbot yang diperuntukkan untuk para teknisi. Hasil pengujian menunjukkan hasil skor usability yang beragam untuk masing-masing platform chatbot. Selanjutnya, Kabetta [10] menguji usability menggunakan metode Black-box dan SUS untuk chatbot pendaratan elektronik. Hasil pengujian blackbox berjalan dengan baik. Hasil perhitungan dengan SUS, menunjukkan chatbot dalam kategori baik dan layak untuk digunakan.

Dari literatur di atas menunjukkan bahwa metode SUS efektif digunakan untuk menguji usability pada chatbot yang dikembangkan pada berbagai bidang. Penelitian ini bertujuan untuk melakukan evaluasi usability pada PUSADBOT untuk melihat tingkatan efektifitas dan pengalaman pengguna yang diberikan. Pengujian usability menggunakan Black-Box dan System Usability Scale yang dipelopor oleh Brooke [10].

METODE PENELITIAN

Usability testing dilakukan untuk mengevaluasi *usability* dari PUSADBOT. Evaluasi sistem dilakukan menggunakan Blackbox dan System *Usability Scale* (SUS) untuk mengetahui tingkat kelayakan dan kegunaan dari *chatbot*.

Pengujian dengan metode black-box yaitu pengujian terhadap cara kerja *chatbot*[11]. Pengujian dilakukan dengan berdasarkan kebutuhan pengguna dengan membuat tabel skenario pengujian. Tabel skenario akan dikategorikan sesuai dengan perancangan *chatbot*.

System *Usability Scale* (SUS) yang diciptakan oleh John Brooke pada tahun 1986, merupakan salah satu metode yang dapat digunakan untuk mengukur persepsi kegunaan sebuah perangkat keras maupun perangkat lunak. Responden kuesioner SUS berjumlah sepuluh orang yang dilakukan secara daring menggunakan kuesioner dari Google Forms.

Peserta

Penelitian menunjukkan bahwa pengujian *usability* dengan hanya delapan peserta mengungkap sekitar 80% masalah kegunaan utama [12]. Total sepuluh peserta untuk mengambil bagian dalam studi penelitian. Dalam upaya mendiversifikasi rentang pengalaman dan perspektif pengguna, penelitian yang dilakukan memilih untuk memasukkan variasi jumlah orang.

Prosedur

Pengujian blackbox dilakukan berdasarkan tabel skenario pengujian dengan fokus pada pengujian fungsional *chatbot* dan pengujian untuk spesifikasi illegal message handling. Selanjutnya untuk pengujian SUS, penelitian ini menghitung skor SUS dari platform *chatbot*. Pertanyaan SUS diajukan dengan peringkat 1 sampai 5 dengan 1 mewakili "sangat tidak setuju" dan angka 5 untuk "sangat setuju". Sepuluh pertanyaan SUS ditunjukkan di bawah ini:

Tabel 1. Pertanyaan dan skala

No	Pertanyaan	Skala
1	Saya sepertinya akan sering menggunakan <i>chatbot</i> ini.	1-5
2	Ada fitur pada <i>chatbot</i> yang sebenarnya tidak perlu.	1-5
3	Saya merasa mudah menggunakan <i>chatbot</i> ini.	1-5
4	Saya sepertinya perlu bantuan teknis untuk mengoperasikan <i>chatbot</i> ini.	1-5
5	Saya menemukan berbagai fungsi dalam <i>chatbot</i> ini telah terintegrasi dengan baik.	1-5
6	Saya pikir ada terlalu banyak ketidaksesuaian dalam <i>chatbot</i> dan sistem pendukungnya.	1-5
7	Saya pikir ada terlalu banyak ketidaksesuaian dalam <i>chatbot</i> dan sistem pendukungnya.	1-5
8	Saya merasa <i>chatbot</i> ini sangat tidak praktis ketika digunakan.	1-5
9	Saya sangat yakin dapat menggunakan <i>chatbot</i> ini.	1-5
	.	

10	Lanjutan tabel 1 Sepertinya saya harus belajar banyak hal terlebih dahulu sebelum saya dapat menggunakan <i>chatbot</i> ini	1-5
----	--	-----

SUS terdiri dari 10 pertanyaan, dengan 5 pertanyaan positif dan 5 pertanyaan negatif serta setiap pertanyaan memiliki bobot 1 sampai 5. Pertanyaan nomor ganjil merupakan pertanyaan positif, skor setiap pertanyaan dihitung dengan cara bobot tiap pertanyaan dikurangi dengan nilai 1 (bobot - 1). Pertanyaan genap yang merupakan pertanyaan negatif, skor dihitung dengan cara 5 dikurangi bobot setiap pertanyaan (5 - bobot). Total skor diperoleh dari jumlah rata-rata skor setiap pertanyaan, kemudian total skor dikalikan 2,5 untuk mendapatkan skor SUS antara 0-100.

HASIL DAN PEMBAHASAN

Demografi peserta pada penelitian ini dapat dilihat pada Tabel 2 berikut.

Tabel 2. Demografi Peserta

Umur	
20-25	9
26-30	1
31-40	-
Jenis kelamin	
Laki	6
perempuan	4
Profesi	
Apoteker	-
Tenaga kefarmasian	-
Lainnya	10

Total peserta keseluruhan dalam penelitian ini berjumlah 10 peserta, dengan usia terbanyak di 20-25 tahun berjumlah 90%, sisanya usia 26-30 tahun berjumlah 10%. Untuk jenis kelamin terbanyak yaitu laki-laki dengan jumlah 60% dan yang sedikit perempuan dengan jumlah 40%. Untuk Profesi yang banyak di profesi lainya berjumlah 100%. Total 10 peserta dengan usia rata-rata 23,2 diambil untuk studi penelitian.

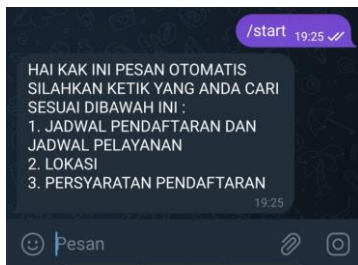
Interaksi Sistem

PUSADBOT memuat informasi Pendaftaran, Lokasi, dan pendaftaran. Tampilan awal *chatbot* ketika pertama kali diakses atau ditambahkan pada chatroom Telegram ditunjukkan pada Gambar 1.



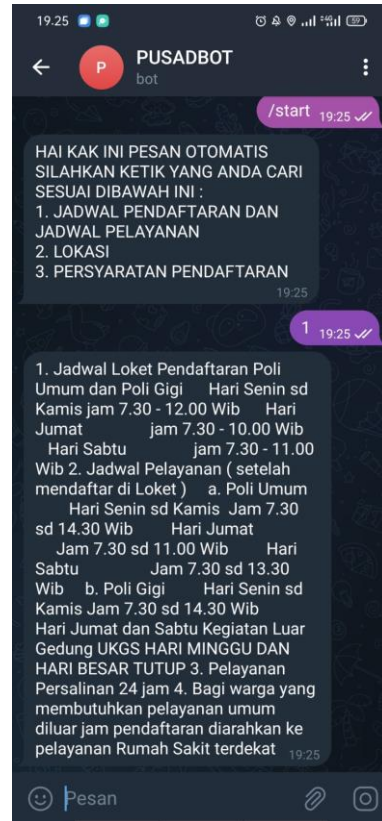
Gambar 1. Tampilan awal *chatbot* ketika pertama kali diakses atau ditambahkan pada chatroom Telegram.

Pada tampilan awal, pengguna akan diberikan informasi sekilas mengenai kegunaan *chatbot* dan disediakan tombol “MULAI” untuk memulai obrolan dengan *chatbot*. Selanjutnya Gambar 2 menunjukkan tampilan ketika pengguna menekan tombol “MULAI”.



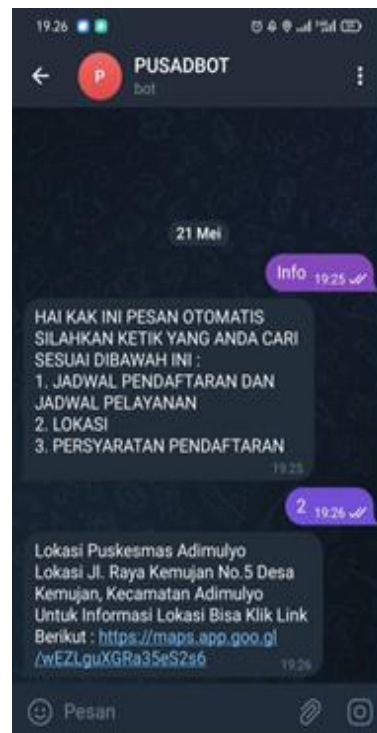
Gambar 2. Tampilan ketika pengguna telah menekan tombol “MULAI”.

Untuk menampilkan menu ke-1 pengguna mengetik “Jadwal” atau ketik “1” kemudian *chatbot* akan merespon informasi jadwal. Kemudian pilihan menu ke-2 dengan mengetik “Lokasi” atau mengetik angka “2” *chatbot* akan merespon informasi tentang lokasi puskesmas. Menu yang ke-3 dengan mengetik “Pendaftaran” atau ketik “3” *chatbot* akan merespon tentang informasi pendaftaran di puskesmas



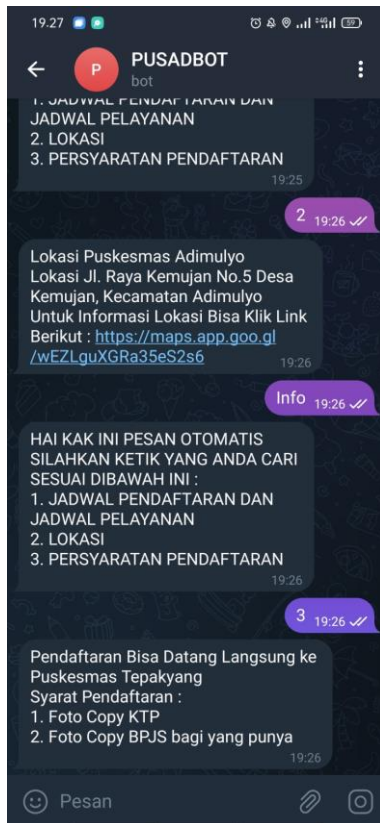
Gambar 3. Tampilan ketika mengetik angka 1.

Tampilan gambar 3 tampilan saat percakapan pemilihan menu informasi tentang puskesmas. Untuk menampilkan menu ke-1 pengguna mengetik “1” kemudian *chatbot* akan merespon informasi jadwal.



Gambar 4. Tampilan ketika mengetik angka 2.

Kemudian pilihan menu ke-2 dengan mengetik angka “2” chatbot akan merespon informasi tentang lokasi puskesmas



Gambar 5. Tampilan percakapan ketika mengetik angka 3.

Hasil Evaluasi blackbox

Pengujian sistem dilakukan untuk memastikan chatbot dapat berjalan sesuai dengan spesifikasi dan ekspektasi pengguna. Menurut Ul Haq (2019), terdapat banyak variasi pengujian perangkat lunak, salah satunya dapat dilakukan dengan metode black-box. Berikut ini skenario pengujian fungsional dengan metode black-box yang dilakukan terhadap sistem chatbot dengan daftar disajikan pada Tabel 3.

Tabel 3. Daftar skenario pada pengujian fungsional chatbot

NO	SPESIFIKASI	TES CASE ID	TES CASE
1	MENU AWAL	C01	Klik Mulai atau ketik Mulai
2	MENU 1	C02	Pengetikan angka “1” valid
3	MENU 2	C03	Pengetian angka “2” valid
4	MENU 3	C04	Pengetikan angka “3” valid
5	Illegal message handling	C05	Menuliskan pesan

	Lanjutan tabel 3		perintah ilegal
--	------------------	--	-----------------

Pada spesifikasi pertama yaitu proses mulai. Masing-masing skenario mewakili tiap tahap percakapan yang berjalan secara berurutan, percakapan akan berlanjut ke menu awal apabila percakapan sebelumnya tidak valid. Hal ini berlaku pula pada spesifikasi pengujian yang lainnya. Urutan tahap percakapan user pada pengetikan “MULAI” chatbot menampilkan menu awal. Pengetikan angka “1” chatbot menampilkan informasi jadwal. Pengetikan angka “2” chatbot menampilkan informasi lokasi. Pengetikan angka “3” chatbot menampilkan informasi pendaftaran. Masing-masing tahap percakapan kemudian diuji dengan data valid dan tidak valid sehingga terdapat dua skenario yang diuji, daftar skenario secara lengkap disajikan pada Tabel 4.

Tabel 4. Skenario pengujian untuk spesifikasi chatbot

ID	Tes Case	Step Detail / Input Value	Expected Result	Actual Result
C01	Klik Mulai atau ketik Mulai	Mengirimkan pesan teks “MULAI”	Chatbot memberikan pesan balasan “HAI KAK INI PESAN OTOMATIS SILAHKAN KETIK YANG ANDA CARI SESUAI DIBAWAH INI : 1. JADWAL PENDAFTARAN DAN JADWAL PELAYANAN 2. LOKASI 3. PERSYARATAN PENDAFTARAN”	Sesuai
C02	Pengetikan angka “1” valid	Mengirimkan pesan teks “1”	Chatbot memberikan pesan balasan “1. Jadwal Loker Pendaftaran Poli Umum dan Poli Gigi Hari Senin sd Kamis jam 7.30 - 12.00 Wib Hari Jumat jam 7.30 - 10.00 Wib Hari Sabtu	Sesuai

pengujian perangkat lunak, salah satunya dapat dilakukan dengan metode Black-Box dan System Usability Scale (SUS). Berikut ini skenario pengujian fungsional dengan metode System Usability Scale (SUS). yang dilakukan terhadap sistem *chatbot* dengan daftar disajikan pada Tabel 6. Hasil kuesioner. Hasil pengumpulan data disajikan pada tabel 6.

Tabel 6. Hasil kuesioner System Usability Scale

No	Pertanyaan	Skala
1	Saya sepertinya akan sering menggunakan <i>chatbot</i> ini.	2,9
2	Ada fitur pada <i>chatbot</i> yang sebenarnya tidak perlu.	3,2
3	Saya merasa mudah menggunakan <i>chatbot</i> ini.	3
4	Saya sepertinya perlu bantuan teknis untuk mengoperasikan <i>chatbot</i> ini.	3,6
5	Saya menemukan berbagai fungsi dalam <i>chatbot</i> ini telah terintegrasi dengan baik.	3,4
6	Saya pikir ada terlalu banyak ketidaksesuaian dalam <i>chatbot</i> dan sistem pendukungnya.	3,2
7	Saya pikir ada terlalu banyak ketidaksesuaian dalam <i>chatbot</i> dan sistem pendukungnya.	3,1
8	Saya merasa <i>chatbot</i> ini sangat tidak praktis ketika digunakan.	2,8
9	Saya sangat yakin dapat menggunakan <i>chatbot</i> ini.	2,9
10	Sepertinya saya harus belajar banyak hal terlebih dahulu sebelum saya dapat menggunakan <i>chatbot</i> ini.	3,2
Total		3,1
Skor SUS (2,5 * Total)		77,5

SUS terdiri dari 10 pertanyaan, dengan lima pertanyaan positif dan 5 pertanyaan negatif serta setiap pertanyaan memiliki bobot 1 sampai 5. Pertanyaan nomor ganjil merupakan pertanyaan positif, skor setiap pertanyaan dihitung dengan cara bobot tiap pertanyaan dikurangi dengan nilai 1 (bobot - 1). Pertanyaan genap yang merupakan pertanyaan negatif, skor dihitung dengan cara 5 dikurangi bobot setiap pertanyaan (5 - bobot). Total skor diperoleh dari jumlah rata-rata skor setiap pertanyaan, kemudian total skor dikalikan 2,5 untuk mendapatkan skor SUS antara 0-100. Menurut Baumgartner [14], standar minimum skor SUS adalah 65 untuk memastikan bahwa produk dapat diterima oleh pengguna. Dari hasil perhitungan diperoleh skor SUS sebesar 77,5 yang menunjukkan bahwa sistem yang dibangun berada dalam kategori baik dan layak untuk digunakan.

KESIMPULAN

Hasil usability testing pada PUSADBOT yang telah dilakukan menggunakan metode Black-Box dan System Usability Scale (SUS) berjalan dengan baik. Pengujian dilakukan pada komponen fungsional dan non-fungsional. Hasil pengujian terhadap tiga fungsi utama (fungsional) *chatbot* yaitu jadwal, lokasi dan pendaftaran, menunjukkan hasil yang sesuai dan *chatbot* dapat berjalan dengan baik. Hasil pengujian terhadap komponen non-fungsional menunjukkan *chatbot* dapat berjalan dengan baik. Evaluasi

sistem menggunakan System Usability Scale (SUS) menghasilkan skor sebesar 77,5 yang menunjukkan bahwa *chatbot* berada dalam kategori baik dan layak untuk digunakan.

DAFTAR PUSTAKA

- [1] M. Melinda, "... Intelligence *Chatbot* Tarra (Toyota Interactive Virtual Assistant) Dalam Meningkatkan Customer Relationship Management Di ...," no. 4647, 2022, [Online]. Available: <http://repository.uin-suska.ac.id/58876/>
- [2] S. Mulyatun, H. Utama, and A. Mustopa, "Pendekatan Natural Language Processing Pada Aplikasi *Chatbot* Sebagai Alat Bantu Customer Service," *J. Inf. Syst. Manag.*, vol. 3, no. 1, pp. 12–17, 2021, doi: 10.24076/joism.2021v3i1.404.
- [3] R. Ren, J. W. Castro, S. T. Acuña, and J. De Lara, "Evaluation Techniques for *Chatbot* Usability: A Systematic Mapping Study," *Int. J. Softw. Eng. Knowl. Eng.*, vol. 29, no. 11–12, pp. 1673–1702, 2019, doi: 10.1142/S0218194019400163.
- [4] "ISO 9241-11 Ergonomic requirements for office work with visual display terminals (VDTs)–Part II Guidance on Usability," 1998.
- [5] S. Rubaeah, T. T. W. Ningrum, Z. W. Fandol, and R. A. Setiawan, "SISPAC: *Chatbot* Untuk Diagnosis dan Penanganan Hipertensi," *Pros. Semin. Nas. Din. Inform. 2021*, pp. 100–105, 2021.
- [6] F. Fitra Ramadhan, "*Chatbot* pada E-Commerce berbasis Android dengan Pendekatan Natural Language Processing," *JCSE J. Comput. Sci. an Eng.*, vol. 2, no. 1, pp. 27–39, 2021, [Online]. Available: <http://icsejournal.com/index.php/>
- [7] Patel, "Implementasi *Chatbot* Untuk Pemesanan Produk Yang Terintegrasi Web Pada Kedai Ibaraki," pp. 9–25, 2019, [Online]. Available: <https://eprints.umm.ac.id/87456/>
- [8] Arief Hidayat; Agung Hidayat; Safa'ah Nurfa'izin, "Usability Evaluation on Educational *Chatbot* Using the System Usability Scale (SUS)," 2022, [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10006991>
- [9] J. Saenz, W. Burgess, E. Gustitis, A. Mena, and F. Sasangohar, "The usability analysis of *chatbot* technologies for internal personnel communications," *67th Annu. Conf. Expo Inst. Ind. Eng. 2017*, pp. 1357–1362, 2017.
- [10] H. Kabetta, "Desain dan Implementasi Penandatanganan Elektronik Sertifikat X509 Menggunakan Platform Bot Telegram," *Telematika*, vol. 13, no. 1, pp. 22–35, 2020, doi: 10.35671/telematika.v13i1.936.
- [11] U. Ul Haq, S., & Qamar, "Ontology Based Test Case Generation for Black Box Testing," 2019.
- [12] "How Many Test Users in a Usability Study?" [Online], 2017, [Online]. Available: <https://www.nngroup.com/articles/how-many-test->

- users/
- [13] and B. Jordan, B. Thomas, I. L. McClelland, "Weerdmeester, Usability Evaluation In Industry," 1996.
- [14] A. Baumgartner, J., Frei, N., Kleinke, M., Sauer, J., & Sonderegger, "Pictorial System Usability Scale (P-SUS) Developing an Instrument for Measuring Perceived Usability," 2019.

PENULIS



Salim Nurul Umam, prodi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Harapan Bangsa.



Raden Bagus Bambang Sumantri, prodi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Harapan Bangsa



Retno Agus Setiawan, prodi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Harapan Bangsa