

## EVALUASI LAYANAN KESEHATAN APLIKASI DEPOK SINGLE WINDOW DENGAN METODE SYSTEM USABILITY SCALE DAN HEURISTIC EVALUATION

Amelinda Kusumaningtyas<sup>\*1</sup>, Prihandoko<sup>2</sup>

<sup>12</sup>Universitas Gunadarma, Depok  
Email: <sup>1</sup>amelindafir@gmail.com, <sup>2</sup>pri@staff.gunadarma.ac.id  
<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 30 Agustus 2023, diterima untuk diterbitkan: 6 November 2023)

### Abstrak

Penerapan *Smart City* di Indonesia merupakan bentuk upaya pemerintah Indonesia dalam pengembangan dan pengelolaan daerah di era desentralisasi. Salah satu contoh implementasi ini adalah *e-government* di Kota Depok dengan meluncurkan aplikasi *mobile* bernama Depok Single Window (DSW). Aplikasi DSW merupakan aplikasi layanan publik yang dimiliki oleh Pemerintah Kota Depok untuk meningkatkan kualitas layanan publik memanfaatkan potensi teknologi telematika secara baik. Untuk meningkatkan kesejahteraan rakyat dan kepercayaan masyarakat terhadap keandalan aplikasi layanan publik, perlu diadakan evaluasi aplikasi. Untuk evaluasi ini maka digunakan 2 metode yaitu *System Usability Scale* (SUS) dan *Heuristic Evaluation* (HE). Data yang digunakan dalam Metode SUS merupakan data primer yang diperoleh langsung dari masyarakat Kota Depok melalui penyebaran kuesioner dengan Google Form. Data yang digunakan dalam Metode HE juga merupakan data primer yang diperoleh langsung dari para ahli yang memahami *usability* dari aplikasi *mobile* dan UI/UX dengan memberikan penilaian berdasarkan 10 prinsip heuristik Nielsen. Dalam Metode SUS didapatkan skor penilaian sebesar 68,75 yang menyatakan bahwa layanan kesehatan yang dievaluasi masih dibawah nilai minimal yang harus didapatkan. Sementara itu, evaluasi yang dilakukan dengan Metode HE menemukan total 13 masalah *usability* yang telah dievaluasi oleh para ahli. Masalah-masalah tersebut telah melanggar 8 dari 10 prinsip heuristik, yaitu *Consistency and Standards, Visibility of System Status, User Control and Freedom, Aesthetic and Minimalist Design, Error Prevention, Recognize, Diagnose, and Recover from Errors, Recognition Rather Than Recall, dan Flexibility and Efficiency of Use*.

**Kata kunci:** evaluasi, aplikasi mobile, DSW, system usability scale, heuristic evaluation.

## EVALUATION OF DEPOK SINGLE WINDOW MOBILE HEALTH APPLICATION USING SYSTEM USABILITY SCALE AND HEURISTIC EVALUATION METHODS

### Abstract

The application of *Smart City* in Indonesia is a form of the Indonesian government's efforts in the development and management of regions in era of decentralization. One of the implementations of it is a mobile application named Depok Single Window (DSW). DSW is a public service application owned by the Depok City Government to improve the quality of public services by utilizing the potential of telematics technology. To improve people's welfare and increase public confidence in the reliability of public service applications, it is necessary to evaluate the application. This evaluation was carried out using 2 methods namely *System Usability Scale* (SUS) and *Heuristic Evaluation* (HE). The data used in the SUS Method is primary data which obtained directly from residents of Depok City through questionnaires using Google Form. Meanwhile the data used in the HE Method is obtained directly from experts who understand usability of mobile application and UI/UX by providing an assessment with guidance from 10 principles of Nielsen's heuristic. Based on the evaluation score in SUS method, the score 68.75 stated that the health service which being evaluated is still under the minimum score. Meanwhile in HE method, a total of 13 usability problems were found and had violated 8 of 10 heuristic principles namely *Consistency and Standards, Visibility of System Status, User Control and Freedom, Aesthetic and Minimalist Design, Error Prevention, Recognize, Diagnose, and Recover from Errors, Recognition Rather Than Recall, and Flexibility and Efficiency of Use*.

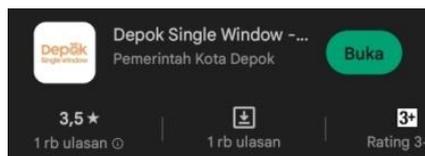
**Keywords:** evaluation, mobile application, DSW, system usability scale, heuristic evaluation.

## 1. PENDAHULUAN

Program Gerakan Indonesia Melayani merupakan salah satu Program Gerakan Nasional Revolusi Mental yang diharapkan dapat membangun perubahan menuju era masyarakat informasi (*information society*) dengan melakukan penyempurnaan standar pelayanan dan sistem pelayanan yang inovatif (*e-government*), peningkatan penyediaan sarana dan prasarana yang menunjang pelayanan publik, dan peningkatan perilaku pelayanan publik yang merupakan salah satu hal yang difokuskan dalam pelaksanaan program tersebut (Instruksi Presiden Republik Indonesia Nomor 12, 2016). Salah satu contoh implementasi ini adalah aplikasi *mobile* bernama Depok *Single Window* (DSW) pada 17 Agustus 2018.

Aplikasi DSW merupakan aplikasi layanan publik yang diluncurkan oleh Dinas Komunikasi dan Informatika Kota Depok yang telah bekerja sama dengan Perusahaan Listrik Negara (PLN), Rumah Sakit Umum Daerah (RSUD), dan Kepolisian. Pengunduhan aplikasi ini dapat dilakukan melalui *Google Playstore* untuk pengguna *android* dan melalui *App Store* untuk pengguna *iOS*. Aplikasi ini diharapkan dapat menjadi media atau jembatan bagi masyarakat Kota Depok untuk mengakses layanan informasi melalui perangkat *mobile* masing-masing.

Berdasarkan hasil data yang diperoleh pada tanggal 1 Desember 2022, dari pengguna di *Google Playstore* dapat diketahui bahwa penilaian aplikasi DSW adalah skala 3,5 dari 5 dan penilaian aplikasi di *App Store* adalah skala 3,1 dari 5. Hal ini dapat dilihat pada gambar 1 dan gambar 2 berikut ini.



Gambar 1. Penilaian Aplikasi DSW di Google Playstore



Gambar 2. Penilaian Aplikasi DSW di Appstore

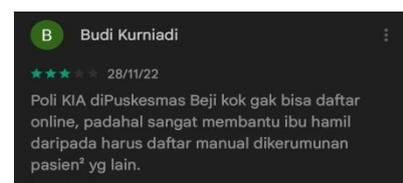
Layanan kesehatan merupakan salah satu layanan yang disediakan di aplikasi DSW untuk menunjang kesejahteraan hidup masyarakat. Pada fitur layanan ini, masyarakat diberikan 16 fitur sub layanan dengan harapan dapat membantu masyarakat dalam aspek kesehatan, seperti cari dokter, layanan RSUD, layanan puskesmas, antrian

puskesmas, informasi kamar, konsultasi Halodoc, *feedback*, berita kesehatan, Tecodep, SakuBumil, kalkulator *stunting*, layanan pembiayaan, layanan informasi BPJS kesehatan, daftar RSUD, daftar puskesmas, dan faskes kesehatan. Hal ini dapat dilihat pada gambar 3.

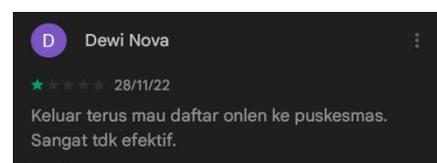


Gambar 3. Tampilan Menu Layanan Kesehatan

Kehadiran fitur layanan ini seharusnya dapat mempermudah masyarakat dalam mengakses informasi mengenai dokter, rumah sakit, puskesmas, atau informasi kesehatan lainnya. Akan tetapi, berdasarkan beberapa komentar yang diberikan oleh pengguna aplikasi di *Google Playstore*, banyak yang menyayangkan fitur-fitur yang tersedia tidak membuat pengguna dapat mencapai tujuan mereka dengan menggunakan layanan kesehatan di aplikasi DSW seperti pada gambar 4, 5, dan 6 dibawah ini.



Gambar 4. Protes Pengguna 2



Gambar 5. Protes Pengguna 3



Gambar 6. Protes Pengguna 4

Peraturan Presiden Republik Indonesia Nomor 72 Pasal 1 Ayat 2 (2012) dijelaskan bahwa Sistem Kesehatan Nasional (SKN) merupakan kegiatan mengelola kesehatan secara terpadu dengan saling mendukung untuk tercapainya tingkat derajat kesehatan masyarakat yang tinggi dan dilakukan oleh seluruh bangsa Indonesia. Kegiatan mengelola kesehatan merupakan salah satu upaya pemerintah Indonesia untuk mendukung derajat kesehatan di Indonesia supaya masyarakat dapat mencapai derajat kesehatan yang tinggi. Banyaknya protes terkait pelayanan kesehatan dari aplikasi DSW ini membuat evaluasi pada layanan kesehatan di aplikasi DSW harus dilakukan.

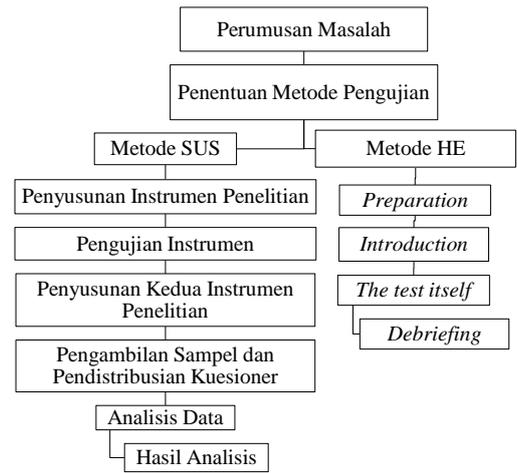
Untuk mengetahui tingkat kegunaan aplikasi maka perlu dilakukan penelitian mengenai evaluasi *usability* layanan kesehatan di aplikasi DSW. Hal ini akan menentukan tingkat penerimaan aplikasi di masyarakat Kota Depok dan faktor apa saja yang menghambat pengguna untuk mencapai tujuannya, khususnya pada fitur layanan kesehatan. Oleh karena itu, dipilih Metode *System Usability Scale* (SUS) dan *Heuristic Evaluation* (HE) sebagai metode yang digunakan untuk evaluasi tersebut.

Metode *System Usability Scale* digunakan untuk mengetahui penilaian *usability* dari pengguna akhir. Pemilihan pernyataan dilakukan dengan penuh hati-hati dan sudah terbukti keandalannya dalam menemukan pendapat pengguna mengenai *usability* aplikasi yang sedang dievaluasi (Brooke, 2013).

Metode *Heuristic Evaluation* digunakan untuk mengetahui penilaian *usability* dari Aplikasi DSW menurut para ahli atau pakar. Prasetyaningtyas, Az-Zahra, dan Brata (2018) menyatakan bahwa metode ini cocok digunakan untuk mengevaluasi aplikasi karena menghasikan daftar kesalahan *usability* pada *user interface* yang spesifik dan dapat dipercaya dari para ahli. Hughes, Ahluwalia, & Midha (2013) juga menyatakan bahwa layanan *e-government* harus terus dilakukan evaluasi untuk berkembang menjadi layanan yang lebih baik.

## 2. METODE PENELITIAN

Dalam penelitian ini, terdapat 2 kegiatan evaluasi yang dilakukan, yaitu evaluasi dengan Metode SUS dan Metode HE. Tahapan penelitian ini dijelaskan pada gambar 7.



Gambar 7. Proses Penelitian

### 2.1. Jenis Penelitian

Penelitian ini merupakan penelitian deskriptif kuantitatif. Penelitian deskriptif adalah penelitian yang memberikan gambaran atau uraian, sedangkan penelitian kuantitatif adalah penelitian yang menganalisis data-data numerikal dengan menggunakan metode statistik untuk pengolahannya (Hardani dkk, 2020). Oleh karena itu, penelitian ini merupakan penelitian yang memberikan gambaran atau uraian terhadap data-data numerikal yang didapatkan.

### 2.2. Metode SUS

Metode *System Usability Scale* (SUS) diciptakan oleh John Brooke untuk mengukur *usability* dari suatu sistem dengan mudah, cepat, dan dapat diandalkan. Pengujian *usability* merupakan pengujian yang membutuhkan pedoman atau acuan dalam menspesifikasikan *usability* dari sistem tersebut. Brooke (1995) melakukan evaluasi terhadap 50 pernyataan yang berpotensi menjadi pedoman untuk metode SUS dengan skala likert. Akan tetapi, kuesioner yang panjang akan menyebabkan kejenuhan pada responden dalam menyelesaikan kuesioner. Maka, dari 50 pernyataan yang sudah disiapkan, dipilih 10 pernyataan (pernyataan negatif dan positif).

Untuk memudahkan peneliti Indonesia dalam menggunakan pernyataan yang terdapat di dalam Metode SUS maka diterjemahkan kuesioner asli ke dalam Bahasa Indonesia (Sharfina & Santoso, 2016). Hasil terjemahan dari pernyataan SUS yang asli dapat dilihat di tabel 1 dibawah ini.

Tabel 1. Pernyataan Terjemahan

| No. | Pernyataan Terjemahan                           |
|-----|---|
| 1   | Saya berpikir akan menggunakan sistem ini lagi. |
| 2   | Saya merasa sistem ini rumit untuk digunakan.   |
| 3   | Saya merasa sistem ini mudah untuk digunakan.   |
| 4   | Saya membutuhkan bantuan dari orang lain        |

|    |   |
|----|---|
|    | atau teknisi dalam menggunakan sistem ini.                                      |
| 5  | Saya merasa fitur-fitur sistem ini berjalan dengan semestinya.                  |
| 6  | Saya merasa ada banyak hal yang tidak konsisten (tidak serasi) pada sistem ini. |
| 7  | Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat.  |
| 8  | Saya merasa sistem ini membingungkan.   |
| 9  | Saya merasa tidak ada hambatan dalam menggunakan sistem ini.                    |
| 10 | Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini.     |

Nilai akhir SUS akan dikategorikan berdasarkan *acceptability ranges*, *grade scale*, dan *adjective ratings*. *Acceptability Ranges* merupakan kategori yang menentukan apakah sistem atau produk yang dievaluasi dapat diterima oleh pengguna atau tidak. Bangor, Kortum, & Miller (2008) menyatakan bahwa suatu aplikasi harus melewati nilai 70 untuk dapat dikatakan *acceptable*. *Grade Scale* merupakan kategori yang menentukan nilai dari sistem atau produk yang. *Grade scale* dikenal juga dengan penilaian universitas analog dari *adjective ratings*. *Adjective Ratings* merupakan kategori untuk mengetahui bagaimana skala *user-friendliness* dari sistem atau produk yang dievaluasi.

Tabel 2. Kategori SUS

| Kategori SUS                | Range                   | Skor      |
|-----------------------------|-------------------------|-----------|
|                             | <i>Not Acceptable</i>   | 0 – 50    |
| <i>Acceptability Ranges</i> | <i>Marginal – Low</i>   | 51 – 62   |
|                             | <i>Marginal – High</i>  | 63 – 70   |
|                             | <i>Acceptable</i>       | 71 – 100  |
| <i>Grade</i>                | A                       | 90 – 100  |
|                             | B                       | 80 – 89   |
|                             | C                       | 70 – 79   |
|                             | D                       | 60 – 69   |
|                             | F                       | < 60      |
|                             | <i>Best Imaginable</i>  | 86 – 100  |
| <i>Adjective Ratings</i>    | <i>Excellent</i>        | 81 – 85   |
|                             | <i>Good</i>             | 71 – 80   |
|                             | <i>OK</i>               | 50,9 – 70 |
|                             | <i>Poor</i>             | 26 – 50,8 |
|                             | <i>Worst Imaginable</i> | 0 – 25    |

### 2.3. Pengujian Instrumen

Instrumen penelitian adalah alat ukur untuk mengukur variabel dalam pengumpulan data (Hardani dkk, 2020). Banyak instrumen penelitian yang disediakan oleh para peneliti terdahulu untuk dapat digunakan oleh peneliti lain, namun perlu dilakukan pengujian instrumen untuk melihat apakah instrumen penelitian tersebut sesuai dengan penelitian yang dilakukan. Pengujian instrumen penelitian dilakukan untuk mengetahui apakah instrumen penelitian tersebut valid dan reliabel. Pengujian dilakukan dengan uji validitas dan uji reliabilitas.

Tujuan dilakukan uji validitas adalah untuk mengetahui apakah pernyataan yang dibuat bernilai valid atau tidak dengan minimal 30 responden. Apabila pernyataan bernilai tidak valid maka

pernyataan tersebut dapat dihapus atau dipertahankan jika pernyataan tersebut adalah pernyataan penting (Nugroho, 2020). Analisis uji validitas diproses menggunakan SPSS versi 25 dengan 2 prinsip sebagai berikut:

1. Bila  $r$  hitung lebih besar dari  $r$  tabel maka bernilai valid.
2. Bila  $r$  hitung lebih kecil dari  $r$  tabel maka bernilai tidak valid.

Sedangkan, uji reliabilitas dilakukan untuk mengetahui keandalan instrumen penelitian yang digunakan. Pengujian ini menggunakan teknik *Cronbach's Alpha* dengan prinsip bahwa:

1. Bila *Cronbach's Alpha* lebih besar dari  $r$  tabel maka bernilai reliabel.
2. Bila *Cronbach's Alpha* lebih kecil dari  $r$  tabel maka bernilai tidak reliabel.

### 2.4. Teknik Pengambilan Sampel

Teknik penarikan sampel *purposive* merupakan teknik pengambilan sampel dimana setiap anggota populasi tidak terpilih menjadi sampel penelitian. Teknik ini merupakan teknik dengan kriteria tertentu dimana sampel sengaja dipilih untuk mewakili populasinya (Wiranata & Winoto, 2023). Pada penelitian ini, sampel yang digunakan hanyalah responden yang tinggal di Kota Depok dan merupakan warga Kota Depok berdasarkan kartu tanda penduduk yang dimiliki, serta pernah menggunakan layanan kesehatan aplikasi DSW.

Data yang digunakan merupakan data primer yang didapatkan langsung dari pengajuan kuesioner kepada 10 responden yang memenuhi kriteria penelitian yang telah ditetapkan. Kriteria ini sesuai dengan penelitian Faulkner (2003) yang menyatakan bahwa rata-rata persentase penemuan masalah *usability* yang ditemukan tidak akan naik secara signifikan untuk jumlah sampel lebih dari 10 seperti yang tertera di tabel 3.

Tabel 3. Tabel Jumlah Sampel

| Jumlah Sampel | Minimal Jumlah Penemuan (%) | Rata-Rata Jumlah Penemuan (%) |
|---------------|-----------------------------|-------------------------------|
| 5             | 55                          | 85,55                         |
| 10            | 82                          | 94,686                        |
| 15            | 90                          | 97,050                        |
| 20            | 95                          | 98,4                          |
| 30            | 97                          | 99,0                          |
| 40            | 98                          | 99,6                          |
| 50            | 98                          | 100                           |

### 2.5. Analisis Data

Metode SUS memiliki perhitungan skor penilaian kuesioner sebagai berikut:

1. Lakukan perhitungan masing-masing jawaban responden dengan persamaan (1) untuk skor pada pernyataan bernomor ganjil.

$$\text{Skor Ganjil} = \text{Bobot skor} - 1 \quad (1)$$

- Lakukan perhitungan masing-masing jawaban responden dengan persamaan (2) untuk skor pada pernyataan bernomor genap.

$$Skor\ Genap = 5 - bobot\ skor \quad (2)$$

- Totalkan semua hasil skor dari kuesioner seperti pada persamaan (3) untuk mendapatkan total skor dari masing-masing jawaban responden.

$$Total = \sum skor\ ganjil + \sum skor\ genap \quad (3)$$

- Kalikan hasil skor dengan 2,5 dengan rumus yang tertera pada persamaan (4) untuk mendapatkan skor SUS dari masing-masing jawaban responden.

$$Skor\ SUS = Total \times 2.5 \quad (4)$$

- Lakukan pencarian nilai rata-rata dari nilai skor SUS masing-masing responden untuk mengetahui hasil akhir skor SUS, dimulai dari rentang skor 0 - 100 pada penelitian ini dengan menggunakan persamaan (5). Pada persamaan ini,  $\bar{X}$  merupakan nilai rata-rata atau nilai akhir SUS dan  $n$  menunjukkan jumlah responden pada penelitian yang dilakukan.

$$\bar{X} = \frac{\sum Skor\ SUS}{n} \quad (5)$$

## 2.6. Metode HE

Metode HE merupakan metode evaluasi *usability* yang dikenalkan oleh Jakob Nielsen dan Rolf Molich yang dilakukan dengan melihat sebuah antarmuka dan memberikan opini terkait antarmuka tersebut (Nielsen & Molich, 1990). Masing-masing penilai akan melakukan eksplorasi terhadap aplikasi yang dievaluasi dan memberikan penilaian sesuai dengan prinsip Nielsen yang tertera di tabel 4 (Nielsen, 1994a).

Tabel 4. Prinsip Nielsen

| Prinsip   | Definisi  |
|---|---|
| <i>Visibility of System Status</i> (X1)             | Sistem harus selalu menginformasikan kepada pengguna mengenai informasi yang sedang terjadi sesuai dengan waktu dan kegiatan yang sedang berlangsung. |
| <i>Match Between System and the Real World</i> (X2) | Sistem harus menggunakan bahasa manusia atau bahasa yang dimengerti oleh pengguna dan konsep desain yang lazim bagi pengguna.                         |

|  |   |
|--|---|
| <i>User Control and Freedom</i> (X3)                     | Pengguna berhak diberi kebebasan untuk memilih dalam suatu keadaan yang tidak diinginkan.   |
| <i>Consistency and Standards</i> (X4)                    | Pengguna tidak perlu bertanya-tanya mengenai arti dari kata-kata, situasi, atau aksi yang tersedia.   |
| <i>Error Prevention</i> (X5)                             | Pencegahan kesalahan atau masalah yang mungkin akan dilakukan oleh pengguna.  |
| <i>Recognition Rather Than Recall</i> (X6)               | Pemberitahuan mengenai tombol, simbol, dan aksi harus dibuat dengan jelas dan terlihat untuk meminimalisir kesusahan pengguna dalam mengingat.                                  |
| <i>Flexibility and Efficiency of Use</i> (X7)            | Alternatif atau cara singkat yang dapat mempercepat kegiatan yang dilakukan.  |
| <i>Aesthetic and Minimalist Design</i> (X8)              | Antarmuka tidak boleh mengandung informasi yang tidak relevan   |
| <i>Recognize, Diagnose, and Recover from Errors</i> (X9) | Pesan mengenai kesalahan wajib menggunakan bahasa yang dapat dimengerti oleh para pengguna, menginformasikan masalah yang dialami, dan saran penyelesaian yang dapat dilakukan. |
| <i>Help and Documentation</i> (X10)                      | Bantuan dan dokumentasi dibutuhkan untuk membantu pengguna dalam memahami bagaimana menyelesaikan tugas yang ingin dikerjakan dengan sistem yang ada.                           |

*Nielsen's Severity Rating* merupakan skala penilaian yang digunakan dalam menetapkan temuan masalah yang melanggar prinsip Nielsen seperti yang terlihat di tabel 5 (Herr, Baumgartner, & Gross, 2016).

Tabel 5. Tabel *Nielsen's Severity Rating*

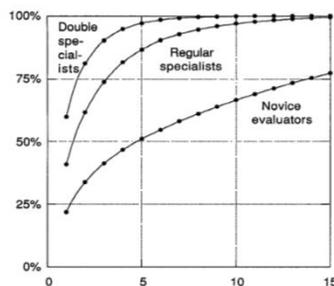
| <i>Nielsen's Severity Ratings</i> | Kategori                     | Definisi  |
|-----------------------------------|------------------------------|---|
| 0                                 | <i>Not a problem</i>         | Tidak termasuk masalah <i>usability</i> .   |
| 1                                 | <i>Cosmetic problem</i>      | Masalah dapat diperbaiki jika ada waktu luang.  |
| 2                                 | <i>Minor problem</i>         | Masalah dengan prioritas rendah.  |
| 3                                 | <i>Major problem</i>         | Masalah dengan prioritas tinggi sehingga harus diperbaiki.  |
| 4                                 | <i>Usability catastrophe</i> | Masalah wajib untuk diperbaiki, terutama untuk produk yang belum dirilis agar diperbaiki sebelum dirilis. |

Nielsen (1993) memberitahukan detail tahapan evaluasi yang dilakukan pada metode ini yaitu:

### 1. Preparation

Pada tahap ini, tempat evaluasi, instruksi, dan penentuan jumlah ahli. Terdapat 3 jenis ahli yaitu *double specialists*, *regular specialists*, dan *novice evaluators*. *Double specialists* merupakan ahli yang memiliki keahlian di 2 bidang yang berhubungan dengan *usability*. *Regular specialists* merupakan ahli yang hanya memiliki keahlian di bidang *usability*, seperti memiliki pemahaman tentang apa saja yang dibutuhkan pengguna, namun tidak memikirkan aspek lain dalam mewujudkan kebutuhan tersebut seperti tampilan *user interface*. Hal ini yang membedakan *regular specialists* dan *double specialists*. *Novice evaluators* merupakan ahli yang masih baru dalam evaluasi *usability*.

Seperti yang terlihat di gambar 8, Nielsen (1992) memberikan rekomendasi jumlah ahli untuk melakukan evaluasi. Jumlah minimal agar rata-rata persentase penemuan masalah melebihi 75% maka dibutuhkan *double specialists* berjumlah 3 – 5 orang, *regular specialists* berjumlah 2 – 3 orang, dan *novice evaluators* berjumlah 14 orang.



Gambar 8. Jumlah Ahli yang Dibutuhkan.

Adapun kriteria ahli yang dibutuhkan pada penelitian ini adalah memiliki minimal gelar S2 atau magister dan memahami, serta memiliki pengalaman di bidang teknologi informasi yang berhubungan dengan *usability* aplikasi *mobile* dan *human-computer interaction* atau UI/UX.

### 2. Introduction

Pada tahap ini, penilai akan diberikan pendahuluan mengenai aplikasi yang akan dievaluasi, tujuan dari pelaksanaan evaluasi, metode yang digunakan, dan pedoman penilaian.

### 3. The Test Itself

Pada tahap ini, masing-masing penilai akan diberikan waktu untuk menilai layanan tersebut dengan menggunakan *Nielsen's Severity Ratings*, memberikan alasan mengenai hasil nilai yang diberikan, dan menjelaskan prinsip yang dilanggar untuk setiap masalah *usability* yang ditemukan.

### 4. Debriefing

Pada tahap ini, semua penilai akan dikumpulkan dan dilakukan evaluasi bersama untuk menentukan nilai akhir dari masalah *usability* yang didapat.

## 3. HASIL DAN PEMBAHASAN

Hasil yang didapat dari Metode SUS menunjukkan bagaimana pendapat pengguna mengenai *usability* dari layanan kesehatan aplikasi *mobile* DSW, sedangkan hasil dari Metode HE menunjukkan faktor-faktor yang membuat *usability* dari layanan tersebut menjadi turun, prinsip yang dilanggar, dan rekomendasi perbaikan berdasarkan saran dari para ahli.

### 3.1. Hasil Uji Instrumen

Hasil pengujian instrumen kepada 30 responden mendapatkan hasil yang menjelaskan bahwa semua instrumen bernilai valid dan reliabel seperti yang terlihat di tabel 6 dan 7.

Tabel 6. Tabel Hasil Uji Validasi

| Kode | R Hitung | R Tabel | Keterangan |
|------|----------|---------|------------|
| X1   | 0,623    | 0,361   | Valid      |
| X2   | 0,665    | 0,361   | Valid      |
| X3   | 0,687    | 0,361   | Valid      |
| X4   | 0,686    | 0,361   | Valid      |
| X5   | 0,791    | 0,361   | Valid      |
| X6   | 0,739    | 0,361   | Valid      |
| X7   | 0,670    | 0,361   | Valid      |
| X8   | 0,786    | 0,361   | Valid      |
| X9   | 0,677    | 0,361   | Valid      |
| X10  | 0,489    | 0,361   | Valid      |

Tabel 7. Tabel Hasil Uji Reliabilitas

| Jumlah Item yang Diuji | Cronbach's Alpha | R Tabel | Keterangan |
|------------------------|------------------|---------|------------|
| 10                     | 0,905            | 0,361   | Reliabel   |

### 3.2. Hasil Metode SUS

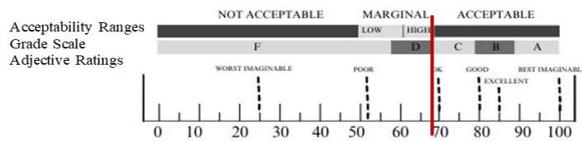
Dari evaluasi yang telah dilakukan terhadap 10 responde, didapatkan hasil bahwa nilai SUS layanan kesehatan aplikasi DSW mencapai nilai 68,75 dengan *adjective ratings* bernilai *OK* yang setara dengan nilai D pada *grade scale* dan *acceptability ranges* berada di posisi *marginal-high*. Hal ini dijelaskan pada tabel 8 dibawah.

Tabel 8. Tabel Hasil SUS

| Skor SUS | Adjective Ratings | Grade Scale | Acceptability Ranges |
|----------|-------------------|-------------|----------------------|
| 68,75    | OK                | D           | Marginal - High      |

*Adjective ratings* yang bernilai *OK* menandakan bahwa pengguna merasa tingkat *user-friendliness* dari layanan kesehatan aplikasi DSW merupakan hal yang biasa saja atau sesuai dengan kebutuhan pengguna, namun tidak ada hal yang mengesankan. Skala D pada *grade scale* menjelaskan bagaimana tingkat *adjective ratings* yang didapat dalam bentuk penilaian huruf. *Acceptability ranges* yang berada pada *marginal-high* menunjukkan bahwa banyak pengguna yang sudah dapat menerima layanan ini dengan baik, akan tetapi beberapa hal harus diperbaiki karena masih

dibawah nilai minimal, yaitu 70. Penjelasan nya dapat dilihat pada gambar 9.



Gambar 9. Peringkat Hasil Akhir SUS

### 3.3. Hasil Metode HE

Dari hasil evaluasi yang dilakukan dengan metode HE, diketahui bahwa terdapat 13 temuan masalah yang melanggar 8 dari 10 prinsip *Heuristic* yang digunakan. Dapat dilihat pada tabel 9 hasil temuan masalah, skala pelanggaran, dan prinsip yang dilanggar.

Tabel 9. Tabel Hasil Evaluasi

| No | Masalah Usability   | Severity Ratings | Prinsip yang Dilanggar  |
|----|---|------------------|---|
| 1  | <i>Hide</i> menu <i>icon</i> dan peletakan yang tidak konsisten antara satu layanan dengan layanan lain.  | 1                | 1. <i>Consistency and Standards</i>   |
| 2  | Pengguna tidak dapat langsung kembali ke menu utama setelah mengunjungi beberapa menu di layanan puskesmas.   | 2                | 1. <i>User Control and Freedom</i><br>2. <i>Flexibility and Efficiency of Use</i> |
| 3  | Infografis di bagian atas halaman utama berukuran kecil dan tidak terbaca, sehingga peletakan infografis menjadi sia-sia.                                       | 2                | 1. <i>Consistency and Standards</i>   |
| 4  | Tombol <i>keyboard</i> pada saat memasukkan nomor kartu tanda penduduk (KTP) dan <i>Captcha</i> penjumlahan masih memungkinkan pengguna untuk memasukkan huruf. | 1                | 1. <i>Error Prevention</i>  |
| 5  | Tampilan Menu Faskes Kesehatan kurang <i>clean</i> dan rapi.  | 2                | 1. <i>Aesthetic and Minimalist Design</i>   |
| 6  | Menu lambat ketika dibuka dan memerlukan waktu <i>loading</i> lebih dari 10 detik.  | 4                | 1. <i>Consistency and Standards</i>   |
| 7  | Link yang sama dengan menu yang memiliki penamaan yang serupa atau berbeda.   | 2                | 1. <i>Consistency and Standards</i><br>2. <i>Recognition Rather Than Recall</i>   |
| 8  | Menu dan sub menu berada pada tingkatan yang sama.  | 1                | 1. <i>Consistency and Standards</i><br>2. <i>Recognition Rather Than Recall</i>   |

|    |   |   |  |
|----|---|---|--|
| 9  | Pemilihan warna <i>orange</i> dan merah untuk kombinasi tulisan.  | 1 | 1. <i>Aesthetic and Minimalist Design</i>  |
| 10 | Pada sub menu Cari Dokter, ketika mengklik salah satu nama dokter, aplikasi menjadi eror.   | 4 | 1. <i>Consistency and Standards</i><br>2. <i>Recognize, Diagnose, and Recover from Errors</i>  |
| 11 | Daftar <i>online</i> RSUD – <i>Input</i> NIK tanpa batas.   | 3 | 1. <i>Error Prevention</i>   |
| 12 | Semua Menu layanan RSUD hanya untuk yang sudah pernah terdaftar atau punya rekam medis.   | 2 | 1. <i>Error Prevention</i>   |
| 13 | Layanan Puskesmas memiliki waktu <i>loading</i> yang paling lambat dan tidak ada informasi yang ditampilkan setelah menunggu diatas 30 detik. | 4 | 1. <i>Consistency and Standards</i><br>2. <i>Visibility of System Status</i><br>3. <i>Recognize, Diagnose, and Recover from Errors</i> |

Temuan masalah yang paling sering ditemukan selama evaluasi dan dinyatakan skor pelanggaran sebesar 4 adalah *loading time*. Miller (1968) menyatakan bahwa jarak waktu aplikasi antara 1 -1 0 detik memiliki dampak yang signifikan pada pengguna dalam melakukan aktivitas yang sedang dilakukan. Setiap detik akan mempengaruhi pengguna apakah aktivitas yang dilakukan akan tetap dilaksanakan atau tidak. Waktu *loading* suatu aplikasi *mobile* tidak boleh melebihi 1 detik atau pengguna akan beralih dari aplikasi tersebut. Hal ini melanggar prinsip *consistency and standard* karena *loading time* dari layanan kesehatan melebihi standar *loading time* untuk aplikasi *mobile*.

### 4. KESIMPULAN

Hasil evaluasi yang dilakukan dengan menggunakan Metode SUS menunjukkan bahwa layanan kesehatan aplikasi DSW masih dibawah nilai standar minimum suatu aplikasi. Hal ini juga didukung dengan hasil dari evaluasi yang menggunakan Metode HE bahwa terdaat 13 masalah *usability* yang melanggar 8 prinsip *Nielsen's Heuristic*. Penelitian mengenai evaluasi *usability* pada layanan kesehatan aplikasi DSW masih merupakan hal yang baru sehingga hasil dari penelitian ini akan sangat membantu pemerintah Kota Depok dalam mempertahankan kegunaan aplikasi, terutama di layanan kesehatan.

Saran yang dapat diberikan untuk pengembangan penelitian berikutnya adalah mencoba untuk menggunakan metode evaluasi lain, seperti *Think-Aloud* yang menyuarakan pendapat dari pengguna secara langsung selama evaluasi berlangsung, dan menambah jumlah responden yang

digunakan agar semakin meningkat rata-rata persentase temuan masalah.

#### DAFTAR PUSTAKA

- BANGOR, A., KORTUM, P.T., dan MILLER, J.T. 2009. Determining what individual SUS scores mean: adding an adjective rating scale. *Journal of Usability Studies archive*, 4, pp. 114-123.
- BANGOR, A., KORTUM, P.T. dan MILLER, J.T. 2008. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24(6), pp. 574–594. DOI:10.1080/10447310802205776.
- BROOKE, J. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind*, pp. 189-194.
- BROOKE, J. 2013. SUS: A Retrospective. *Journal of Usability Studies*, 08(02), pp. 29-40.
- FAULKNER, L. 2003. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behaviour Research Methods, Instruments, & Computers*, 35(3), pp. 379-383.
- HARDANI., AULIYA, N. H., ANDRIANI, H., FARDANI, R. A., AUSTIAWATY, J., UTAMI, J., SUKMANA, D. J., dan ISTIQOMAH, R. R. 2020. *Metode Penelitian Kualitatif & Kuantitatif*. Yogyakarta: Pustaka Ilmu. ISBN: 978-623-7066-33-0.
- HUGHES, J., AHLUWALIA, P., dan MIDHA, V. 2013. A heuristic evaluation instrument for e-government online software. *Electronic Government, An International Journal*, 10(1), pp. 1-18
- Instruksi Presiden Republik Indonesia Nomor 12 tahun 2016 tentang Gerakan Nasional Revolusi Mental. Jakarta: Sekretariat Kabinet Republik Indonesia.
- MILLER, R. B. 1968. Response time in man-computer conversational transactions. *Proceedings of the December 9-11, fall joint computer conference, part I (AFIPS '68 (Fall, part I)*. Association for Computing Machinery, New York, USA, pp. 267–277. DOI: 10.1145/1476589.1476628.
- NIELSEN, J. 1992. Finding usability problems through heuristic evaluation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. Association for Computing Machinery, New York, USA, pp. 373–380. DOI: 10.1145/142750.142834.
- NIELSEN, J. (1993). *Usability Engineering*. Amerika: Academic Press. ISBN: 9780125184052.
- NIELSEN, J dan MOLICH, R. 1990. Heuristic evaluation of user interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*. Association for Computing Machinery, New York, USA, pp. 249–256. DOI: 10.1145/97243.97281 69.
- NIELSEN, J. 1994. How to Conduct a Heuristic Evaluation, [online] Tersedia di:<<https://www.semanticscholar.org/paper/How-to-Conduct-a-HeuristicEvaluation-Nielsen/67259adaf6e77da5082395d6d5d043fb674c08ed#citingpapers>> [Diakses 28 Desember 2022].
- NIELSEN, J. 1994a. Enhancing the Explanatory Power of Usability Heuristics. *Conferences of Human Factors in Computing Systems (CHI '94)*. Association for Computing Machinery, New York, USA, pp. 152-158. DOI: 10.1145/259963.260333.
- NUGROHO, P. S. 2020. *Analisis Data Penelitian Bidang Kesehatan*. Yogyakarta: Gosyen Publishing. ISBN: 978-602-5411-44-1.
- PRASETYANINGTIAS, T., AZ-ZAHRA, H. M., DAN BRATA, A. H. 2018. Analisis Usability Pada Aplikasi Mobile E-Government Layanan Aspirasi dan Pengaduan Online Rakyat (LAPOR!) Dengan Heuristic Evaluation. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 02(11), 4647-4653. e-ISSN: 2548-964X.
- Peraturan Presiden Republik Indonesia Nomor 72 tahun 2012 tentang Sistem Kesehatan Nasional. Jakarta: Sekretariat Kabinet Republik Indonesia.
- HERR, S., BAUMGARTNER, N., dan GROSS, T. 2016. Evaluating Severity Rating Scales for Heuristic Evaluation. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, USA, pp. 3069–3075. DOI: 10.1145/2851581.2892454.
- SHARFINA, Z dan SANTOSO, H. B. 2016. An Indonesian Adaptation of the System Usability Scale (SUS). *ICAC SIS 2016*, pp. 145-148.
- WIRANATA, M.A., dan WINOTO, C. (2023). Evaluasi Usability Aplikasi Mobile Ojek Indralaya (Ojin) Menggunakan System Usability Scale. *Jurnal JUPITER*, 15(01), pp. 267-276.