

BAB II

TINJAUAN PUSTAKA

Bagian ini memuat kajian dan uraian sistematis tentang informasi hasil penelitian yang pernah dilakukan peneliti lain dalam pustaka dan menghubungkannya dengan masalah penelitian. Tinjauan pustaka membahas tentang metode yang digunakan dalam penelitian *user experience*. Hal tersebut digunakan sebagai dasar membuat struktur Landasan Teori.

2.1. Penelitian Sebelumnya

Pada penelitian ini peneliti telah melakukan studi literatur terhadap beberapa jurnal penelitian terbaru sebelumnya. Jurnal tersebut dipilih berdasarkan topik yang sesuai dan berkaitan dengan tema penelitian yang diteliti yaitu mengenai evaluasi *user experience*. Berikut penjelasan lebih lanjut.

Penelitian pertama adalah penelitian mengenai evaluasi pengalaman pengguna aplikasi *mobile* PaTik Bali [13]. Penelitian ini bertujuan untuk melakukan evaluasi *user experience* maupun *user interface* aplikasi PaTik Bali dikarenakan angka *uninstall* aplikasi yang sangat tinggi yaitu sebanyak 82.000 dari sekitar 89.000 pengguna yang melakukan instalasi pada *playstore*. Penelitian ini menggunakan metode *User Experience Questionnaire* (UEQ) dengan 34 responden. Berdasarkan hasil analisis dari 6 aspek yang dievaluasi, keenam aspek tersebut memiliki penilaian yang baik. Analisis pragmatis dan hedonis juga mendapatkan nilai yang baik. Begitu juga dengan hasil analisis *benchmark* yang dilakukan. Maka dari itu, dapat disimpulkan mengenai penyebab pengguna melakukan *uninstall* aplikasi adalah bukan karena pengalaman pengguna yang buruk, melainkan hal lainnya [13].

Penelitian kedua yaitu penelitian terkait evaluasi dan perbaikan *user*

experience dengan studi kasus terhadap aplikasi Gojek [22]. Penelitian tersebut merupakan penelitian kuantitatif yang dilakukan menggunakan metode *User Experience Questionnaire* (UEQ). Responden dalam penelitian tersebut yaitu merupakan 154 mahasiswa ilmu komputer milenium. Data yang telah terkumpul diolah menggunakan *tools* yang disediakan oleh UEQ. Kemudian dilakukan UEQ *benchmarking* untuk mendapatkan gambaran analisis yang lebih baik. Hasil penelitian menunjukkan bahwa aplikasi Gojek memperoleh hasil evaluasi positif dari seluruh kategori UEQ. Kategori "*Perspicuity*" mendapatkan skor nilai tertinggi, sedangkan skor terendah diperoleh pada kategori "*Novelty*". Kemudian untuk nilai *benchmark* aplikasi Gojek dapat disimpulkan bahwa aplikasi Gojek memiliki skala penilaian "*above average*" dan "*good*" [22].

Penelitian ketiga membahas mengenai evaluasi *user experience* terhadap *game educational website Kahoot* dan *Socrative* menggunakan metode *User Experience Questionnaire* (UEQ) [23]. Terdapat 6 skala UEQ yang digunakan yaitu *Attractiveness*, *Efficiency*, *Perspicuity*, *Dependability*, *Stimulation* dan *Novelty*. Evaluasi *user experience* terhadap kedua *website* tersebut ditujukan untuk mengetahui lebih jauh bagaimana perasaan pengguna, apakah pengguna tertarik dengan *website*, sehingga dapat diketahui *website* mana yang lebih disukai oleh siswa SMKN 3 Malang kelas kelas X-TKJ 1 sebagai responden. Hasil analisis pada penelitian tersebut menunjukkan bahwa *user experience* Kahoot lebih baik berdasarkan semua skala penilaian kecuali pada skala *Dependability* pada item yang tidak dapat diprediksi. Dari segi rata-rata, Kahoot jauh di atas rata-rata nilai Socrative, sehingga dapat disimpulkan bahwa siswa lebih menyukai Kahoot daripada Socrative [23].

Penelitian keempat merupakan penelitian tentang evaluasi *user experience* pada aplikasi E-Report yang dikembangkan oleh Kementerian Pendidikan dan Kebudayaan untuk SMK/SMA dan SMP [24]. Penelitian tersebut menghasilkan data kualitatif dan kuantitatif sehingga metode yang

digunakan merupakan metode campuran, yaitu metode *Cognitive Walkthrough* (CW), *Heuristic Evaluation* (HE) dan *User Experience Questionnaire* (UEQ). Penggunaan ketiga metode secara bersamaan ini bertujuan untuk memperoleh data yang lebih komprehensif, valid, reliabel, dan objektif. Pada penerapan metode CW, peneliti melakukan evaluasi dengan memeriksa waktu pengerjaan tugas, jumlah kegagalan, dan masukan berupa kesimpulan dan saran yang diberikan responden. Sedangkan pada metode HE, peneliti mengikuti 10 prinsip Nielsen untuk melakukan evaluasi. Peneliti juga melakukan survey menggunakan UEQ yang terdiri dari 26 item pertanyaan yang terbagi menjadi 6 aspek yaitu *Attractiveness*, *Clarity*, *Efficiency*, *Satisfaction*, *Stimulation* dan *Novelty*. Hasil penelitian ini yaitu E-Report sebagai objek penelitian belum memenuhi kriteria produk dengan *usability* yang baik. Sebab, pada tingkat *effectiveness* terdapat tugas (skenario) yang memiliki persentase kesalahan yang sangat besar yaitu 100%, dan terdapat beberapa fitur dalam sistem yang gagal berfungsi [24].

Penelitian kelima yaitu penelitian mengenai evaluasi *user experience* terhadap *website e-commerce Sociolla.com* [25]. Evaluasi terhadap *website* dilakukan dengan maksud untuk mengetahui pengalaman pengguna pada saat mengoperasikan *website*. Penelitian tersebut dilakukan dengan menggunakan pengujian skenario *Usability Testing*, kuesioner SUS dan metode UEQ. Uji kuesioner menggunakan metode UEQ dilakukan oleh responden yang merupakan wanita yang belum pernah menggunakan *website Sociolla.com* dengan jumlah 25 orang. Sedangkan pengujian skenario pada *usability testing* dan SUS dilakukan oleh 10 responden yang belum pernah menggunakan *website Sociolla.com*. Terdapat 4 parameter yang dihasilkan dari pengujian senario yaitu Parameter *Task Completed*, Parameter *Error Rate during Task Performance*, Parameter *Time per Completed Task* dan Parameter *Number of Clicks*. Berdasarkan hasil dari pengujian skenario, parameter *Task Completed* memiliki nilai keberhasilan

sebesar 0.91, parameter *Error Rate* memiliki rata-rata jumlah kesalahan sebesar 0.05467, parameter *Time per Completed Task* memiliki rata-rata waktu sebesar 45.313 detik dan parameter *Number of Clicks* memiliki rata-rata sebesar 5 klik. Kemudian untuk kuesioner SUS, skor yang didapatkan adalah 75,75, dan untuk metode UEQ menghasilkan aspek *pragmatic quality* dengan nilai sebesar 3.05, dan aspek *hedonic quality* dengan nilai sebesar 2.23 [25].

Penelitian terakhir membahas mengenai perbandingan beberapa metode yang digunakan untuk melakukan evaluasi *usability*. Beberapa metode yang dibahas antara lain adalah *Heuristic Evaluation*, *Cognitive Walkthrough*, *Scenario-based*, *Remote Usability Testing* dan *User-based Testing*. Peneliti melakukan penelitian berdasarkan penelitian sebelumnya. Pada penelitian tersebut, peneliti menyajikan tabel perbandingan berdasarkan beberapa aspek diantaranya adalah kecepatan, biaya, fleksibilitas, sumber yang dibutuhkan, jumlah partisipan, jenis evaluasi, dampak yang dihasilkan, tingkat penemuan masalah, dan tujuan evaluasi. Hasil penelitian mengungkapkan bahwa setiap metode evaluasi memiliki kelebihan dan keterbatasan yang unik. Menurut penelitian tersebut, tidak ada satupun metode yang lebih unggul dibandingkan metode lainnya. Penggunaan metode bergantung pada sejauh mana metode digunakan untuk mengidentifikasi permasalahan dalam sistem pada sejumlah faktor dan tingkat kerumitannya. Pada penelitian ini, metode *Moderated Remote Usability Testing* memiliki kelebihan yaitu fleksibel untuk diterapkan [26].

Tabel 2.1 menunjukkan ringkasan dari penelitian terdahulu yang berkaitan dengan metode *Moderated Remote Usability Testing* dan metode UEQ.

Tabel 2.1 Penelitian Terdahulu

No	Peneliti	Judul	Metode	Hasil
1	I Nyoman Saputra Wahyu Wijaya, Putu Praba Santika, Ida Bagus Ary Indra Iswara & I Nyoman Alit Arsana	Analisis dan Evaluasi Pengalaman Pengguna PaTik Bali dengan Metode <i>User Experience Questionnaire</i> (UEQ)	<i>User Experience Questionnaire</i> (UEQ)	Aplikasi PaTik Bali memiliki tingkat pengalaman pengguna yang baik. Hal ini menunjukkan bahwa penyebab pengguna melakukan <i>uninstall</i> aplikasi adalah bukan karena pengalaman pengguna yang buruk, melainkan hal lainnya
2	Leon Andretti Abdillah	Analisis Aplikasi Mobile Transportasi Online Menggunakan <i>User Experience Questionnaire</i> pada Era Milenial dan Z	<i>User Experience Questionnaire</i> (UEQ)	Aplikasi Gojek memiliki skor " <i>Perspicuity</i> " yang paling tinggi dibandingkan dari kategori lainnya, sedangkan skor terendah diraih oleh kategori " <i>Novelty</i> ". Nilai <i>benchmark</i> aplikasi Gojek menunjukkan bahwa aplikasi Gojek memiliki skala penilaian yang berada pada skala "above average" dan "good"
3	Fauzan Fitrandi, Hanifah Muslimah Az-Zahra & Admaja Dwi Herlambang	Evaluasi <i>User Experience</i> pada Kahoot dan <i>Socrative</i> menggunakan Metode <i>User Experience Questionnaire</i> (UEQ) (Studi pada SMKN 3 Malang)	<i>User Experience Questionnaire</i> (UEQ)	Hasil analisis pada penelitian tersebut menunjukkan bahwa <i>user experience</i> Kahoot lebih baik berdasarkan semua skala penilaian kecuali skala <i>Dependability</i> pada item yang tidak dapat diprediksi.
4	Putu Wiwin Sintya Dewi, G R Dantes & Gedhe Indrawan	<i>User experience evaluation of e-report application using cognitive walkthrough (cw), heuristic evaluation (he) and user experience questionnaire (ueq)</i>	<i>Cognitive Walkthrough</i> (CW), <i>Heuristic Evaluation</i> (HE) dan <i>User Experience Questionnaire</i> (UEQ)	Evaluasi <i>user experience</i> yang dilakukan berdasarkan aspek <i>usability</i> dalam aplikasi yang diteliti belum memenuhi kriteria untuk produk dengan kegunaan yang baik.

No	Peneliti	Judul	Metode	Hasil
5	Shandranuur Fauziah Novitasari, Yusi Tyroni Mursityo & Alfi Nur Rusydi	Evaluasi Pengalaman Pengguna Pada E-Commerce Sociolla.com Menggunakan <i>Usability Testing</i> dan <i>User Experience Questionnaire (UEQ)</i>	<i>Usability Testing, System Usability Scale (SUS) dan User Experience Questionnaire (UEQ)</i>	Berdasarkan hasil pengujian <i>user experience</i> , dapat disimpulkan bahwa <i>website Sociolla.con</i> dapat diterima dengan baik oleh pengguna.
6	Najmeh Ghasemifard, Mahboubeh Shamsi, Abol Reza Rasouli Kenar & Vahid Ahmadi	<i>A New View at Usability Test Methods of Interfaces for Human Computer Interaction</i>	<i>Heuristic Evaluation, Cognitive Walkthrough, Scenario-based, Remote Usability Testing dan User-based Testing</i>	Hasil dari penelitian ini menyebutkan bahwa tidak ada satupun metode yang lebih unggul dibandingkan lainnya. Penggunaan metode bergantung pada sejauh mana metode digunakan untuk mengidentifikasi permasalahan dalam sistem pada sejumlah faktor dan tingkat kerumitannya.

Penelitian [13],[22],[23],[24],[25] menjadi acuan dalam penelitian ini karena terdapat kesamaan metode kuesioner yang digunakan yaitu metode *User Experience Questionnaire (UEQ)*. Sedangkan penelitian [26] membandingkan beberapa metode pengujian *usability*, salah satunya yaitu metode *remote testing*. Berdasarkan penelitian terdahulu, maka dapat disimpulkan bahwa metode UEQ dapat dilakukan untuk mengetahui tingkat pengalaman pengguna pada suatu sistem secara rinci dan juga dapat memberikan rekomendasi perbaikan sistem berdasarkan keenam aspek yang terdapat pada UEQ. Namun, terdapat beberapa kekurangan pada penelitian terdahulu, yaitu hanya berfokus terhadap penelitian kuantitatif saja. Oleh karena itu, untuk mendapatkan hasil analisis yang lebih valid dan detail, penelitian secara kualitatif ditambahkan dalam penelitian ini. Penggunaan metode *Moderated Remote Usability Testing* tepat dilakukan untuk penelitian ini karena metode ini merupakan metode yang fleksibel namun

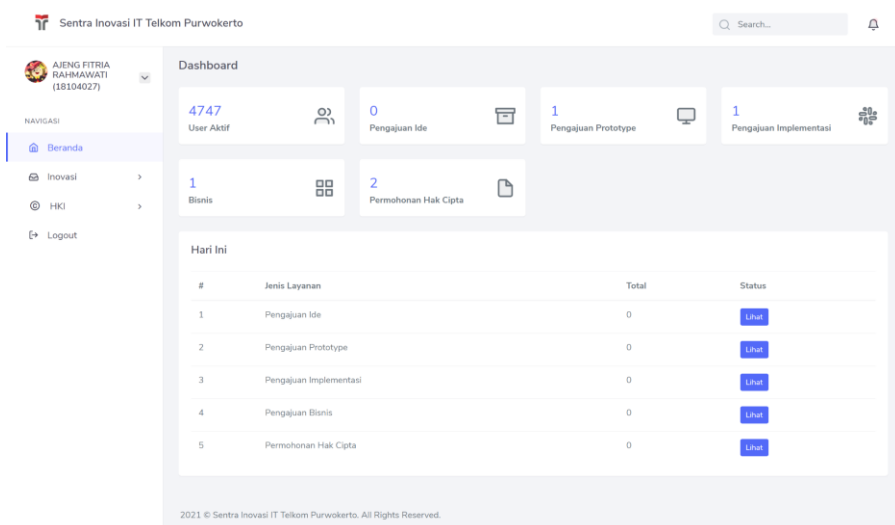
dapat menghasilkan data kualitatif yang mendetail dan representatif [19],[21],[26].

2.2. Dasar Teori

Berikut adalah kajian mengenai beberapa teori yang digunakan pada penelitian ini:

2.2.1 Website Sinovi

Website Sinovi merupakan *website* yang dikembangkan oleh Unit Sentra Inovasi dan HAKI Institut Teknologi Telkom Purwokerto. *Website* ini dikembangkan dengan tujuan untuk pengarsipan data inovasi dan HAKI (Hak Kekayaan Intelektual) secara *digital* sehingga memudahkan pengelolaan data. *Website* ini dikembangkan menggunakan bahasa pemrograman PHP dengan *framework* yang digunakan yaitu Laravel. *Website* ini berjalan secara online dan dapat diakses pada alamat sinovi.ittelkom-pwt.ac.id. Tampilan *website* dapat dilihat pada Gambar 2.1 hingga Gambar 2.4.



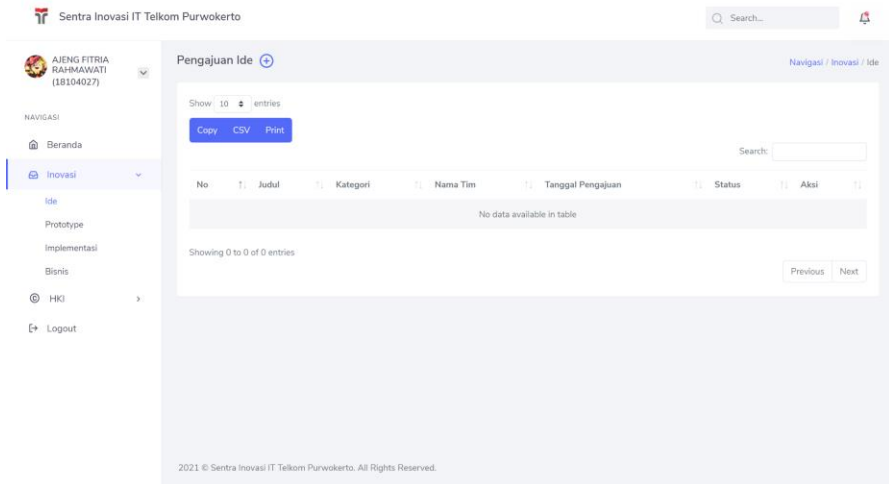
The screenshot displays the dashboard of the Sinovi website. At the top, it shows the logo of Sentra Inovasi IT Telkom Purwokerto and a search bar. The user profile for AIJENG FITRIA RAHMAWATI (18104027) is visible on the left. The dashboard features several key metrics: 4747 User Aktif, 0 Pengajuan Ide, 1 Pengajuan Prototype, and 1 Pengajuan Implementasi. Below these are two more metrics: 1 Bisnis and 2 Permohonan Hak Cipta. A table titled 'Hari Ini' lists various services with their counts and status.

#	Jenis Layanan	Total	Status
1	Pengajuan Ide	0	Lihat
2	Pengajuan Prototype	0	Lihat
3	Pengajuan Implementasi	0	Lihat
4	Pengajuan Bisnis	0	Lihat
5	Permohonan Hak Cipta	0	Lihat

2021 © Sentra Inovasi IT Telkom Purwokerto. All Rights Reserved.

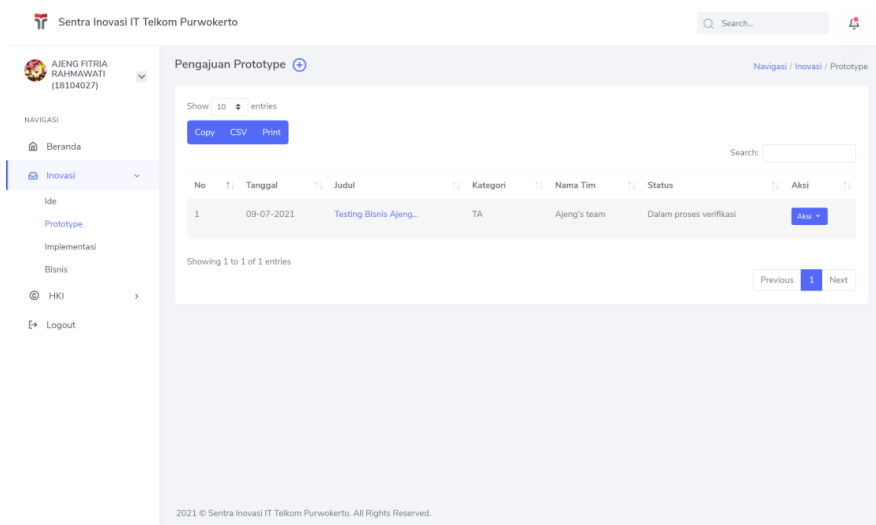
Gambar 2.1 Halaman Beranda

Gambar 2.1 menunjukkan tampilan dari halaman beranda yang menampilkan jumlah pengajuan yang dimiliki pengguna pada setiap layanan yang disediakan.



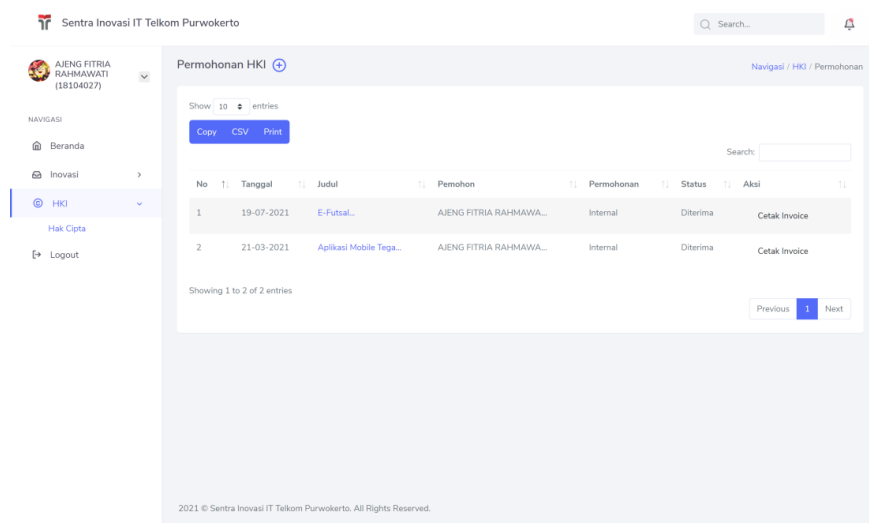
Gambar 2.2 Halaman Ide

Gambar 2.2 merupakan tampilan sub menu ide pada menu inovasi yang digunakan untuk mengelola data ide yang dimiliki pengguna mulai dari menambahkan, mengubah, lihat detail dan menghapus.



Gambar 2.3 Halaman *Prototype*

Gambar 2.3 merupakan tampilan sub menu *prototype* pada menu inovasi yang digunakan untuk mengelola data *prototype* yang dimiliki pengguna mulai dari menambahkan, mengubah, lihat detail dan menghapus.



Gambar 2.4 Halaman Hak Cipta

Gambar 2.4 merupakan tampilan sub menu hak cipta pada menu HKI yang digunakan untuk mengelola hak cipta yang dimiliki pengguna mulai dari menambahkan, mengajukan hak cipta, lihat detail dan menghapus.

2.2.2 *User Experience*

Menurut Tom Tullis dan Bill Albert dalam bukunya “*Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics: Second Edition*”, *user experience* merupakan seluruh tindakan atau interaksi antara pengguna dengan produk maupun sistem atau sesuatu apapun yang berkaitan dengan

interface [11]. *User experience* dapat digunakan sebagai tolak ukur suatu produk maupun sistem untuk dapat mengetahui apakah sistem berjalan dengan baik atau buruk sesuai dengan kebutuhan pengguna atau persepsi pengguna [15]. *User Experience* (UX) sangatlah penting dalam proses pembuatan produk maupun sistem. Dengan *user experience* maka dapat diketahui apa saja yang menjadi masalah dan apa saja yang menjadi keinginan pengguna pada suatu produk maupun sistem [27].

2.2.3 *User Experience Questionnaire* (UEQ)

User Experience Questionnaire (UEQ) merupakan kuesioner yang digunakan untuk mengukur pengalaman pengguna secara subjektif terhadap suatu produk [16]. *User experience questionnaire* (UEQ) memungkinkan penilaian cepat pengalaman pengguna untuk produk interaktif apapun [28]. Menurut Schrepp, dkk [16], terdapat 6 aspek penilaian yang digunakan pada UEQ. Berikut merupakan penjelasan mengenai aspek tersebut :

1. Daya tarik (*attractiveness*)

Attractiveness menggambarkan mengenai kesan keseluruhan dari suatu produk. Dengan aspek *attractiveness* kita dapat mengetahui apakah pengguna menyukai produk atau sebaliknya, apakah produk menarik atau menyenangkan. Pada aspek *attractiveness*, terdapat 6 indikator yang menjadi dasar penilaian, antara lain : *annoying / enjoyable, good / bad, unlikable / pleasing, unpleasant / pleasant, attractive / unattractive* dan *friendly / unfriendly*.

2. Kejelasan (*perspicuity*)

Perspicuity menunjukkan tentang kemudahan suatu produk seperti Apakah produk mudah dikenal, mudah dipelajari, dan mudah dipahami? Berikut merupakan 4 indikator yang

terdapat pada aspek *perspicuity* : *not understandable / understandable, easy to learn / difficult to learn, complicated / easy* dan *clear / confusing*.

3. Efisiensi (*efficiency*)

Efficiency menggambarkan mengenai kemampuan pengguna dalam menyelesaikan tugasnya, seperti apakah pengguna mampu menyelesaikan tugas mereka tanpa memerlukan usaha lain? Apakah interaksi suatu produk tersebut efisien dan cepat? Apakah produk bereaksi tanggap terhadap masukan pengguna? Pada aspek *efficiency*, terdapat 4 indikator yang menjadi dasar penilaian antara lain : *fast / slow, inefficient / efficient, impractical / practical* dan *organized / cluttered*.

4. Keteguhan (*dependability*)

Dependability menggambarkan mengenai kontrol pengguna terhadap suatu produk, seperti Apakah pengguna merasa dapat mengendalikan interaksi? Dapatkah pengguna memprediksi perilaku sistem? Apakah pengguna merasa aman saat bekerja dengan produk? Pada aspek *dependability*, terdapat 4 indikator yang menjadi dasar penilaian diantaranya adalah : *unpredictable / predictable, obstructive / supportive, secure / not secure* dan *meets expectations / does not meet expectations*.

5. Stimulasi (*stimulation*)

Stimulation menggambarkan seberapa jauh suatu produk dapat memberikan motivasi kepada pengguna. Pada aspek *stimulation*, terdapat 4 indikator yang menjadi dasar penilaian yaitu *valuable / inferior, boring / exciting, not interesting / interesting* dan *motivating / demotivating*.

6. Kebaruan (*novelty*)

Novelty menggambarkan mengenai seberapa inovatif suatu produk. Pada aspek *novelty* terdapat 4 indikator yang menjadi dasar penilaian yaitu : *creative / dull, inventive / conventional, usual / leading edge* dan *conservative / innovative*.

User Experience Questionnaire (UEQ) memberikan kesan yang komprehensif mengenai pengalaman pengguna mulai dari aspek klasik hingga aspek pengalaman pengguna. Metode UEQ menyajikan *data analysis tools* yang akurat untuk menafsirkan hasil dengan mudah [22]. *Tools* tersebut dapat diakses pada www.ueq-online.org. Selain itu, UEQ menawarkan tolak ukur (*benchmark*) [17] yang berisi data dari 246 evaluasi produk dengan UEQ (dengan total 9905 peserta di semua evaluasi). UEQ *benchmark* mengklasifikasikan produk ke dalam 5 (lima) kategori (per skala) yaitu:

1. Luar biasa (*Excellent*) artinya bahwa produk termasuk pada kisaran 10% hasil terbaik,
2. Bagus (*Good*) memiliki arti jika produk berada dibawah 10% dari set data benchmark dan lebih baik dari 75% data benchmark lainnya,
3. Cukup baik (*Above Average*) berarti suatu produk berada dibawah 25% dari hasil dalam benchmark dan diatas 50% dari hasil yang lebih buruk,
4. Kurang baik (*Below Average*) artinya bahwa 50% hasil dalam tolok ukur lebih baik daripada hasil untuk produk yang dievaluasi, 25% dari hasil lebih buruk, dan
5. Buruk (*Bad*) memiliki arti bahwa produk berada pada 25% kisaran hasil terburuk.

Tabel 2.2 menunjukkan bagaimana setiap kategori

berhubungan dengan nilai skala rata-rata yang diamati [16].

Tabel 2.2 Interval Benchmark Pada Skala UEQ

	<i>Att.</i>	<i>Eff.</i>	<i>Per.</i>	<i>Dep.</i>	<i>Sti.</i>	<i>Nov.</i>
<i>Excellent</i>	≥ 1.75	≥ 1.78	≥ 1.9	≥ 1.65	≥ 1.55	≥ 1.4
<i>Good</i>	≥ 1.52 < 1.75	≥ 1.47 < 1.78	≥ 1.56 < 1.9	≥ 1.48 < 1.65	≥ 1.31 < 1.55	≥ 1.05 < 1.4
<i>Above average</i>	≥ 1.17 < 1.52	≥ 0.98 < 1.47	≥ 1.08 < 1.56	≥ 1.14 < 1.48	≥ 0.99 < 1.31	≥ 0.71 < 1.05
<i>Below average</i>	≥ 0.7 < 1.17	≥ 0.54 < 0.98	≥ 0.64 < 1.08	≥ 0.78 < 1.14	≥ 0.5 < 0.99	≥ 0.3 < 0.71
<i>Bad</i>	< 0.7	< 0.54	< 0.64	< 0.78	< 0.5	< 0.3

Perbandingan menggunakan tolak ukur (*benchmark*) merupakan indikator pertama mengenai apakah produk menawarkan UX yang cukup sukses atau tidak. Biasanya 20-30 pengguna sudah memberikan pengukuran yang cukup stabil. Membandingkan hasil skala yang berbeda dengan produk di benchmark memungkinkan kesimpulan mengenai kekuatan dan kelemahan relatif produk [16].

2.2.4 *Moderated Remote Usability Testing*

Moderated Remote Usability Testing merupakan salah satu jenis *remote usability testing*. *Remote usability testing* pertama diperkenalkan oleh Hartson, dkk [30],[31] pada 1996. Metode pengujian *remote usability* memungkinkan dilakukannya pengujian jarak jauh yang tidak melibatkan pengguna secara langsung [19],[29]. *Moderated remote usability testing* sendiri adalah metode yang mengharuskan adanya moderator yang hadir untuk memandu jalannya pengujian. Moderator membacakan setiap tugas kepada peserta, kemudian peserta menggunakan metode berbicara "*think aloud*" saat berinteraksi dengan situs web atau aplikasi. *Think aloud* berarti peserta menjelaskan apa yang mereka pikirkan saat mereka

menelusuri situs web dan tugas mereka. Moderator dapat menindaklanjuti atau menyelidiki untuk mengungkap informasi tambahan dari peserta saat mereka mengerjakan tugas [29]. *Moderated Remote Usability Testing* menggunakan jumlah peserta yang relatif kecil, yaitu 5 hingga 10 peserta [10],[11].

2.2.5 *Performance Metrics*

Performance merupakan seluruh hal mengenai apa yang sebenarnya dilakukan oleh pengguna ketika berinteraksi dengan produk. Perilaku pengguna tersebut akan membentuk suatu *performance metrics*. Mengukur *performance metrics* sangat penting untuk mengetahui seberapa baik pengguna menggunakan suatu produk[11]. Selain itu, *performance metrics* juga digunakan untuk mengevaluasi efektivitas dan efisiensi dari suatu produk. Terdapat lima tipe dasar *performance metrics* yaitu [11]:

1. *Task success*

Task success atau keberhasilan tugas merupakan *performance metric* yang dinilai dengan cara mengukur seberapa efektif pengguna dapat menyelesaikan serangkaian tugas tertentu. Tentu saja kita juga dapat mengukur kegagalan tugas. Selama pengguna memiliki tugas yang cukup jelas, maka keberhasilan dapat diukur. Keberhasilan tugas adalah sesuatu yang hampir semua orang dapat hubungkan. Jika pengguna tidak dapat menyelesaikan tugas mereka, maka tentu terdapat sesuatu yang salah. Melihat pengguna gagal menyelesaikan tugas sederhana bisa menjadi bukti yang cukup meyakinkan bahwa ada sesuatu yang perlu diperbaiki. Dua jenis *task success* yang telah ditinjau yaitu keberhasilan biner dan tingkat keberhasilan.

a. *Binary Success*

Keberhasilan biner adalah cara paling sederhana dan paling umum untuk mengukur keberhasilan tugas. Pengguna

berhasil menyelesaikan tugas atau tidak. Keberhasilan biner sesuai untuk digunakan ketika keberhasilan produk tergantung pada pengguna yang menyelesaikan tugas atau serangkaian tugas. Satu-satunya hal yang penting adalah bahwa pengguna mencapai tujuan dengan tugas-tugas mereka. Setiap kali pengguna melakukan tugas, mereka harus diberi skor "berhasil" atau "gagal". Biasanya, skor ini dalam bentuk 1 (untuk sukses) dan 0 (untuk gagal). Analisis lebih mudah jika menetapkan skor numerik daripada nilai teks "berhasil" atau "gagal". Dengan menggunakan skor numerik, maka perhitungan persentase benar serta statistik lain akan lebih mudah. Cukup hitung rata-rata dari 1 dan 0 untuk menentukan persen yang benar. Cara paling umum untuk menganalisis dan menyajikan tingkat keberhasilan biner adalah dengan membandingkan hasil dari setiap tugas.

b. *Levels of Success*

Mengumpulkan dan mengukur tingkat data keberhasilan sangat mirip dengan data keberhasilan biner namun dalam tingkat keberhasilan harus ditentukan berbagai tingkat. Ada beberapa pendekatan untuk tingkat keberhasilan:

- Berdasarkan pengalaman pengguna dalam menyelesaikan tugas. Beberapa pengguna mungkin kesulitan atau membutuhkan bantuan, sementara yang lain menyelesaikan tugas mereka tanpa kesulitan.
- Berdasarkan pengguna menyelesaikan tugas dengan cara yang berbeda. Beberapa pengguna mungkin menyelesaikan tugas dengan cara yang optimal, sementara yang lain mungkin menyelesaikannya dengan cara yang kurang optimal.

Tingkat keberhasilan berdasarkan sejauh mana pengguna menyelesaikan tugas biasanya memiliki antara tiga dan enam

tingkat. Pendekatan yang umum adalah dengan menggunakan tiga tingkatan: sukses total, sukses sebagian, dan gagal total. Tingkat data keberhasilan hampir sama mudahnya untuk dikumpulkan dan diukur seperti data keberhasilan biner. Dalam menganalisis tingkat keberhasilan, hal pertama yang harus dilakukan adalah membuat diagram batang bertumpuk. Hal tersebut akan menunjukkan persentase pengguna yang termasuk dalam setiap kategori atau level, termasuk kegagalan.

2. *Time on task*

Time on task dapat dinilai dengan mengukur berapa lama waktu yang dibutuhkan untuk menyelesaikan suatu tugas. *Time on task* biasanya dinyatakan dalam menit dan detik. Secara logistik, *time on task* dapat diukur dengan berbagai cara diantaranya yaitu moderator mengukur menggunakan *stopwatch*, atau dengan mencatat waktu mulai dan waktu berakhir. Hasil *time on task* dari seluruh partisipan kemudian dijadikan satu dalam tabel *excel* untuk memudahkan perhitungan rata-rata, median, rata-rata geometrik, dan interval kepercayaan untuk setiap tugas agar dapat dilakukan analisis data.

3. *Errors*

Errors atau kesalahan mendeskripsikan kesalahan yang dibuat selama mengerjakan tugas. Kesalahan dapat berguna untuk menunjukkan bagian antarmuka yang membingungkan atau menyesatkan. Dalam beberapa kasus, mungkin ada beberapa peluang kesalahan untuk suatu tugas. Cara paling umum untuk mengatur data kesalahan adalah dengan mencatat jumlah kesalahan untuk setiap tugas dan setiap pengguna. Jika hanya ada satu peluang untuk kesalahan, angkanya adalah 1 dan 0:

0 = Tidak ada kesalahan

1 = Satu kesalahan

Jika beberapa peluang kesalahan dimungkinkan, angka akan bervariasi antara 0 dan jumlah maksimum peluang kesalahan. Semakin banyak peluang kesalahan, semakin sulit dan lebih memakan waktu untuk mentabulasi data. Perhitungan kesalahan dapat dilakukan saat mengamati pengguna selama studi lab, dengan meninjau *video* setelah sesi selesai, atau dengan mengumpulkan data menggunakan alat otomatis atau *online*.

4. *Efficiency*

Efficiency merupakan metrik yang dapat dinilai dengan memeriksa jumlah usaha yang dikeluarkan pengguna untuk menyelesaikan tugas, seperti jumlah klik di situs web atau jumlah penekanan tombol pada ponsel. Terdapat beberapa hal penting ketika mengumpulkan dan mengukur efisiensi, diantaranya adalah:

- a. Melakukan identifikasi terhadap tindakan yang akan diukur. Pada *website*, klik mouse atau tampilan halaman adalah tindakan umum. Pada *software*, dapat berupa klik mouse atau penekanan tombol.
- b. Menentukan awal dan akhir suatu tindakan. Perlu diketahui kapan suatu tindakan dimulai dan berakhir. Beberapa tindakan berlangsung sangat cepat seperti ketika menekan tombol, namun tindakan lain juga dapat memakan waktu lebih lama. Tindakan lain mungkin lebih bersifat pasif, seperti melihat halaman web.
- c. Menghitung tindakan. Tindakan harus terjadi dengan kecepatan yang dapat diidentifikasi secara visual atau, jika terlalu cepat, oleh sistem otomatis.

- d. Tindakan harus bermakna. Setiap tindakan harus mewakili peningkatan inkremental dalam upaya kognitif dan atau fisik. Semakin banyak tindakan, semakin banyak upaya.

Analisis efisiensi dapat dilakukan dengan menyajikan dengan melihat jumlah tindakan yang dilakukan setiap partisipan untuk menyelesaikan tugas. Rata-rata setiap tugas dapat dihitung untuk melihat berapa banyak tindakan yang diambil. Analisis ini sangat membantu dalam mengidentifikasi tugas mana yang membutuhkan usaha paling banyak. Namun, jika beberapa tugas lebih rumit daripada yang lain maka penting juga untuk mewakili interval kepercayaan (berdasarkan distribusi berkelanjutan).

5. *Learnability*

Learnability merupakan metrik yang digunakan untuk mengukur bagaimana kinerja meningkat atau gagal meningkat dari waktu ke waktu. *Learnability* diukur dengan cara melihat jumlah waktu dan usaha yang diperlukan untuk menjadi terbiasa menggunakan suatu produk, sehingga dapat dikatakan ahli dalam menggunakan sesuatu. Proses pengumpulan dan pengukuran *learnability* pada dasarnya sama dengan metrik kinerja lainnya, tetapi harus mengumpulkan data beberapa kali. Setiap contoh pengumpulan data dianggap sebagai percobaan. Percobaan mungkin setiap 5 menit, setiap hari, atau sebulan sekali. Waktu antara uji coba, atau saat mengumpulkan data, didasarkan pada frekuensi penggunaan yang diharapkan. Keputusan pertama adalah jenis metrik yang ingin digunakan. *Learnability* dapat diukur menggunakan hampir semua metrik kinerja dari waktu ke waktu, tetapi yang paling umum adalah yang berfokus pada efisiensi, seperti *time on task*, *errors*, jumlah

langkah, atau *task success* per menit. Saat pembelajaran terjadi, maka diharapkan efisiensi akan meningkat.