

BAB I

PENDAHULUAN

1.1 Latar Belakang

Beberapa tahun terakhir, *Artificial Intelligence* (AI) atau kecerdasan buatan mengalami perkembangan yang sangat pesat dan telah mengubah *lifestyle* atau gaya hidup masyarakat yang membawa mentransformasi signifikan dalam zaman yang serba modern ini, dan (Wu et al., 2020) mengklaim bahwa AI menjadi salah satu peran utama yang mendorong transformasi industri di dunia. Menurut (Zhang & Lu, 2021), *Artificial Intelligence* merupakan bidang ilmu yang mempelajari bagaimana membuat sebuah komputer menjalankan tugas-tugas yang berhubungan dengan kecerdasan yang di masa lalu, hanya dapat dilakukan oleh manusia, sedangkan menurut (Pongtambing et al., 2023), kecerdasan buatan merupakan bagian dari disiplin ilmu komputer yang memanfaatkan teknik komputasi pintar yang berguna untuk menciptakan sistem yang dapat memahami, mengambil keputusan secara independen, dan belajar, sehingga teknologi *Artificial Intelligence* diklaim dapat memudahkan kehidupan manusia yang cukup kompleks. Baru-baru ini di dalam bidang *machine learning* telah menghasilkan teknologi yang lebih maju untuk membuat konten digital, seperti *Artificial Intelligence* yang bisa menciptakan konten sendiri (BAIDOO-ANU & OWUSU ANSAH, 2023). *Artificial Intelligence* menggunakan sistem pengambilan keputusan yang berbasis data, sehingga diperlukannya jumlah data yang sangat besar untuk menghasilkan model AI yang akurat (Ali et al., 2023). *Artificial Intelligence* sendiri merupakan disiplin ilmu dari bidang teknologi dan komputer yang memiliki tujuan untuk membuat perangkat lunak atau *software* dan juga perangkat keras atau *hardware* yang fungsinya digadang-gadang dapat meniru otak manusia (Ratama et al., 2019). *Artificial Intelligence* atau AI dipercaya menjadi *tools* baru yang dapat membantu manusia melakukan pekerjaannya dengan lebih efektif dan efisien dengan fitur-fitur yang sangat lengkap dan komprehensif. Menurut (Nawaz & Saldeen, 2020), manusia lebih suka menggunakan aplikasi yang berbentuk pesan daripada jenis aplikasi

serupa lainnya. Karena fitur pesan merupakan salah satu fungsi utama dari sebuah *gadget*, sehingga manusia langsung dapat beradaptasi dengan AI yang berbasis *chatbot*. Salah satu keuntungan dari penggunaan *chatbot* ialah *user* dapat dengan mudah saat mencari informasi. Oleh karena itu sejak diluncurkannya ChatGPT pada November 2022, ChatGPT menjadi teknologi yang memiliki konsumen paling cepat berkembang di dalam sejarah (Rudolph et al., 2023). Dengan perkiraan sekitar 123 juta *user* aktif bulanan kurang dari tiga bulan setelah diluncurkan. Pertumbuhan ChatGPT jauh lebih cepat dari aplikasi TikTok yang juga sedang tren pada masa peluncuran ChatGPT tetapi membutuhkan sembilan bulan untuk mencapai 100 juta *user* aktif bulanan dan Instagram membutuhkan dua setengah tahun untuk pencapaian yang sama (Wodecki, 2023). Sehingga, ChatGPT telah menjadi aplikasi dengan pertumbuhan tercepat sepanjang sejarah. Dengan perkembangannya yang sangat pesat ini, banyak *chatbot* lain meluncurkan *Artificial Intelligence* berbasis *chatbot* seperti Gemini AI yang sebelumnya Google Bard, Characteristic AI, Writesonic AI, Perplexity AI, Gemini AI, Claude AI, dan masih banyak jenis *Artificial Intelligence* yang berbasis *chatbot* lainnya. Banyaknya jenis AI membuat banyak perdebatan di kalangan penggunanya dalam menentukan AI yang akan digunakan. Di dalam penelitian ini, peneliti berfokus kepada komparasi ChatGPT dan Gemini AI atau yang sebelumnya dikenal dengan Google Bard. Pemilihan ini dikarenakan berdasarkan data *Visual Capitalist* (Conte, 2024) yang dirilis tahun 2023, ChatGPT menempati peringkat pertama sebesar 14.6 miliar kali kunjungan dan Gemini AI menempati posisi ketiga dengan kunjungan sebesar 241.6 juta kali, jika diurutkan berdasarkan jenis AI yang berbasis *chatbot* yang dikunjungi paling banyak di tahun 2023 dari periode September 2022 sampai dengan Agustus 2023.

Sehingga pada penelitian kali ini, peneliti ingin mengkomparasi kualitas dari kedua layanan *chatbot* antara *website* ChatGPT dan Gemini AI berdasarkan parameter dari Nielson Model dan ISO/IEC 9126 yang berfokus pada dimensi *usability*, karena poin penting dari pengujian *usability* kali ini yaitu pengalaman yang dirasakan pengguna saat menggunakan *website* dari kedua AI yang akan dikomparasi dan berdasarkan teori McCall didalam (Rifqatusa'Adah, 2017), tentang kualitas sebuah

software atau perangkat lunak, menurutnya faktor yang menunjang kualitas sebuah *software* (*software quality factors*) dapat diukur dari beberapa aspek, yaitu *efficiency*, *correctness*, *flexibility*, *interoperability*, *integrity*, *maintainability*, *portability*, *reliability*, *reusability*, *testability*, dan *usability* yang memiliki tujuan untuk memuaskan kebutuhan penggunanya (Rifqatusa'Adah, 2017) . Karena *software quality factors* yang dilakukan akan berfokus kepada perspektif dimensi *usability*, sehingga penelitian ini akan mengukur apakah para *pengguna* dari sebuah sistem dapat dengan mudah mengerti cara pengoperasian dan menggunakan *software* tanpa memerlukan banyak latihan, dan karena *usability* termasuk dalam kategori kebutuhan non-fungsional, sehingga ini menjadi tolak ukur dan faktor yang penting dan juga fundamental dalam sebuah *software*, karena dimensi *usability* merupakan salah satu faktor yang dapat menentukan kegagalan atau keberhasilan dari *software* pada saat digunakan oleh penggunanya. Menurut Zaphiris didalam (Utami et al., 2020) beberapa metode yang digunakan untuk mengevaluasi sebuah desain sistem ialah *Inspection*, *Model/Metric-Based*, *Inquiry* dan *Testing*, dari keempat kategori ini, metode *usability testing* merupakan metode yang saat ini paling banyak digunakan untuk menguji keakuratan sebuah sistem, oleh karena itu metode *usability testing* digunakan didalam penelitian ini, dimana metode ini dilakukan dengan cara mengamati *user* saat mereka berinteraksi dengan kedua AI yang akan dikomparasi.

Alasan peneliti menggunakan dua metode yaitu dikarenakan sub-dimensi yang diambil dari ISO/IEC 9126 hanya empat, yaitu faktor *learnability* , *understandability*, *attractiveness*, dan *operability* (Wulandari et al., 2023) sehingga faktor atau sub-dimensi tidak mencakup faktor penilaian lain yang juga penting seperti *satisfaction* pada *software* mengingat ini menjadi salah satu faktor penting guna mencapai kepuasan terhadap penggunaan *software*, sehingga peneliti menambahkan beberapa faktor atau sub-dimensi dari Nielson Model karena memiliki beberapa faktor atau dimensi yang sesuai, yaitu *efficiency*, *errors*, *memorability*, dan *user satisfaction* (Puspitasari et al., 2023) yang menjadi pelengkap dan menambah akurasi pada penilaian berdasarkan dimensi *usability* ini mengacu pada teori Rikard Edgren di dalam (Rifqatusa'Adah, 2017), yang merupakan seorang profesional di bidang *software testing*, ia menyebutkan

karakteristik apa saja yang harus diamati dalam mengukur sebuah *usability* sebuah *software*, ia menyebutkan ada 15 karakteristik, diantaranya *learnability*, *memorability*, *operability*, *errors*, *consistency*, *accessibility*, *documentation*, *tailorability*, *control*, *clarity*, *interactivity*, *discoverability*, *minimalism*, *intuitiveness*, dan *affordance*. Sehingga korelasi ISO/IEC 9126 dan Nielsen Model merupakan parameter yang dipilih karena dimensinya mencakup 15 karakteristik yang menjadi pengukur *usability* sebuah *software*. Berdasarkan permasalahan yang telah dibahas, sehingga adapun tujuan dari penelitian ini ialah untuk mengetahui hasil dari komparasi antara ChatGPT dan Gemini AI berdasarkan dimensi *usability* dengan melalui tahap pengujian, dengan memastikan faktor *usability* terhadap kualitas *software* yang dikomparasi telah sesuai dengan faktor-faktor *usability* yang perlu dipenuhi dan mengetahui tingkat persentase dari kedua *software* tersebut. Pelaksanaan evaluasi terhadap dimensi *usability* ini didasari pada sub-dimensi atau faktor dari *usability* yang ada pada Nielsen Model dan ISO/IEC 9126.

1.2 Rumusan Masalah

Mengacu pada latar belakang yang telah diuraikan, sehingga rumusan masalah penelitian ini difokuskan pada permasalahan berikut :

1. Berapakah nilai atau persentase yang dihasilkan berdasarkan penilaian *usability* dari kedua *chatbot AI* ?
2. *Chatbot AI* (ChatGPT vs Gemini AI) manakah yang lebih unggul berdasarkan penilaian dari dimensi *usability*?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah disebutkan, sehingga penelitian ini memiliki tujuan sebagai berikut :

1. Mengetahui nilai dan persentase berdasarkan *usability* dari masing-masing *chatbot*..
2. Mengetahui *chatbot AI* (ChatGPT vs Gemini AI) yang lebih unggul dari segi dimensi *usability*nya.

1.4 Manfaat Penelitian

Adapun manfaat yang ingin penulis capai pada penelitian ini, ialah :

1. Bagi pengembang *software*, agar dapat menjadi masukan dalam segi tingkat kualitas dari segi aspek *usability* dan juga mendapatkan penilaian dalam bentuk perhitungan kuantitatif dan saran yang berguna untuk mengukur tingkat kualitas dari *software* ChatGPT dan Gemini AI khususnya dari dimensi *usability*.
2. Bagi Instansi Universitas Merdeka Malang, menambah publikasi jurnal di bidang IT khususnya di lingkup *software quality control* dan penelitian ini dapat dijadikan arsip perpustakaan bagi Universitas Merdeka Malang yang akan berguna sebagai bacaan atau referensi mahasiswa/i yang kedepannya akan melakukan penelitian di bidang yang sama.
3. Bagi penulis, penelitian ini memberikan kesempatan untuk mengeksplorasi pendekatan dan juga teori-teori yang berhubungan dengan aspek *usability* dan juga dari sisi keilmuan *software quality control*.
4. Bagi pembaca, menjadi pengetahuan baru dan mendapat *insight* dari hasil penilaian dan penjabaran dari kedua *software*, sehingga lebih mudah untuk pengambilan keputusan ketika ingin memilih antara ChatGPT atau Gemini AI.

1.5 Batasan Penelitian

Terdapat batasan di dalam pengerjaan penelitian ini agar menjadi pengendali untuk berfokus pada permasalahan yang dibahas. Berikut paparan batasan masalah dari penelitian ini :

1. Berfokus kepada dimensi *usability* oleh ISO/IEC 9126 yang berdasarkan empat sub-dimensi atau faktor yaitu *attractiveness*, *learnability*, *operability*, dan *understandability* dan empat sub-dimensi dari model Nielsen Model yaitu *efficiency*, *errors*, *memorability*, dan *satisfaction*.
2. Untuk mengevaluasi kegunaan *software*, instrumen yang digunakan ialah dengan melakukan pengisian kuesioner oleh *user* secara *online* dengan teknik

purposive sampling dengan tipe pertanyaan *close-ended question* dari populasi *user* yang pernah menggunakan *chatbot AI*.

3. Saran yang akan diberikan, murni berdasarkan hasil persentase terendah atau yang tidak memenuhi syarat dari pengujian *usability software*.
4. Penelitian ini hanya menghasilkan rekomendasi perbaikan berupa rekomendasi deskriptif pada dimensi yang penilaiannya tidak mencapai kategori baik (dibawah 68.01%), karena tujuan dari peneliti murni hanya untuk mengkomparasi antara ChatGPT dan Gemini AI.