



Pengembangan Chatbot Konsultasi Kesehatan Mental di Indonesia Berbasis Model Generative AI Gemini

Sherin Sagita Inayah *, Yudistira Bagus Pratama, Arvi Pramudiyantoro

Program studi Ilmu Komputer, Fakultas Teknik dan Sains, Universitas Muhammadiyah Bangka Belitung, Indonesia

*Email (Penulis Korespondensi): sagitasherin@gmail.com

Abstrak. Kasus gangguan kesehatan mental di Indonesia terus meningkat, namun akses terhadap layanan profesional masih sangat terbatas akibat kurangnya tenaga medis, biaya, dan besarnya stigma sosial. Penelitian ini bertujuan mengembangkan chatbot konsultasi kesehatan mental berbasis model Generative AI Gemini 2.5 Flash dan mengevaluasi kinerjanya dalam memberikan respons dukungan psikologis awal. Pengembangan sistem menggunakan metode Machine Learning Lifecycle mulai dari problem definition hingga deployment. Sistem mengimplementasikan prompt engineering dengan kerangka kerja R.O.L.E (Role, Objective, Limits, Expression) untuk mengendalikan model Gemini agar merespons secara empatik tanpa memberikan diagnosis medis klinis. Antarmuka aplikasi dibangun menggunakan teknologi web standar, sementara sisi backend dikelola menggunakan Node.js dan database Supabase. Tahap deployment dilakukan pada Virtual Machine Google Cloud Platform dengan menggunakan Nginx dan PM2. Kelayakan sistem dievaluasi secara komprehensif melalui Black Box Testing, User Acceptance Testing (UAT), dan validasi ahli psikologi. Hasil Black Box Testing membuktikan seluruh fitur sistem beroperasi tanpa kegagalan teknis. Pengujian UAT menghasilkan skor 92,50%, menunjukkan tingkat penerimaan masyarakat yang sangat baik. Evaluasi komparatif dan validasi ahli klinis mengonfirmasi bahwa chatbot secara mutlak mematuhi batasan keamanan dengan menolak diagnosis medis dan berhasil mengeksekusi otomatis protokol rujukan darurat pada situasi krisis. Hal ini menunjukkan bahwa chatbot layak dan aman diimplementasikan sebagai instrumen pertolongan pertama psikologis yang anonim bagi masyarakat luas.

Kata kunci: Kesehatan Mental; Chatbot; Gemini; Generative AI

Abstract. Cases of mental health disorders in Indonesia continue to rise, yet access to professional services remains severely limited due to a lack of medical personnel, high costs, and profound social stigma. This research aims to develop a mental health consultation chatbot based on the Gemini 2.5 Flash Generative AI model and evaluate its performance in providing initial psychological support responses. System development utilized the Machine Learning Lifecycle method, starting from problem definition to deployment. The system implements prompt engineering using the R.O.L.E (Role, Objective, Limits, Expression) framework to control the Gemini model, ensuring empathetic responses without providing clinical medical diagnoses. The application interface was built using standard web technologies, while the backend is managed using Node.js and a Supabase database. The deployment phase was executed on a Google Cloud Platform Virtual Machine using Nginx and PM2. System feasibility was comprehensively evaluated through Black Box Testing, User Acceptance Testing (UAT), and psychologist expert validation. Black Box Testing results proved that all system features operate without technical failures. UAT testing achieved a score of 92.50%, indicating an excellent level of public acceptance. Comparative evaluation and clinical expert validation confirmed that the chatbot strictly adheres to safety boundaries by refusing medical diagnoses and successfully executing automated emergency referral protocols during crisis situations. This indicates that the chatbot is

feasible and safe to be implemented as an anonymous psychological first aid instrument for the general public.

Keywords: *Mental Health; Chatbot; Gemini; Generative AI*

1. Pendahuluan

Kesehatan mental merupakan pilar kesejahteraan psikologis yang krusial bagi individu untuk dapat berfungsi secara sosial dan produktif dalam kehidupan sehari-hari (World Health Organization, 2022). Secara global, diperkirakan satu dari delapan orang hidup dengan gangguan kesehatan mental, di mana kecemasan dan depresi menjadi fenomena yang paling mendominasi (Santomauro et al., 2021). Kondisi ini mengindikasikan bahwa permasalahan psikologis telah menjadi tantangan kesehatan serius yang meluas di berbagai belahan dunia.

Beban gangguan kesehatan mental menunjukkan tren peningkatan yang signifikan, khususnya pada kelompok usia remaja dan dewasa muda (Vigo et al., 2022). Prevalensi gangguan mental pada populasi usia 10 hingga 24 tahun berada pada kisaran 13-20%, yang dipicu oleh tekanan akademik serta perubahan pola interaksi di ruang digital (Charlson et al., 2022). Fenomena ini menempatkan kelompok usia produktif sebagai populasi rentan yang membutuhkan mekanisme deteksi dini dan dukungan emosional yang tepat.

Di Indonesia, urgensi penanganan kesehatan mental terlihat dari laporan *Indonesia National Adolescent Mental Health Survey (I-NAMHS)* yang mencatat 34,9% remaja mengalami masalah emosional dan perilaku (Kementerian Kesehatan RI, 2022). Peningkatan tren depresi pada usia produktif juga dikonfirmasi melalui laporan Survei Kesehatan Indonesia (SKI) tahun 2023 yang menunjukkan eskalasi kasus secara nasional. Kondisi tersebut menuntut adanya inovasi layanan yang mampu menjangkau masyarakat secara lebih luas dan cepat.

Meskipun tingkat kebutuhan masyarakat sangat tinggi, akses terhadap layanan kesehatan profesional masih sangat terbatas akibat rendahnya rasio tenaga kesehatan mental (Patel et al., 2022). Selain faktor ketersediaan sumber daya, stigma sosial yang kuat di tengah masyarakat Indonesia sering kali menjadi penghalang utama bagi individu untuk mencari bantuan (Hidayat et al., 2021). Hambatan-hambatan tersebut menciptakan kesenjangan besar antara jumlah penderita gangguan mental dengan intervensi psikologis yang mereka terima.

Kemajuan teknologi *Artificial Intelligence (AI)*, terutama dalam bidang *Natural Language Processing (NLP)*, menawarkan peluang besar untuk menyediakan dukungan kesehatan mental yang anonim (Naslund et al., 2022). Pengembangan *chatbot* kesehatan mental terbukti efektif dalam memberikan dukungan emosional awal serta meningkatkan kesadaran diri pengguna terhadap kondisi psikologisnya (Fulmer et al., 2021). Teknologi ini berpotensi menjadi jembatan bagi individu yang enggan berkonsultasi secara tatap muka akibat rasa malu atau takut akan penghakiman sosial.



Gambar 1. Spektrum kesehatan mental

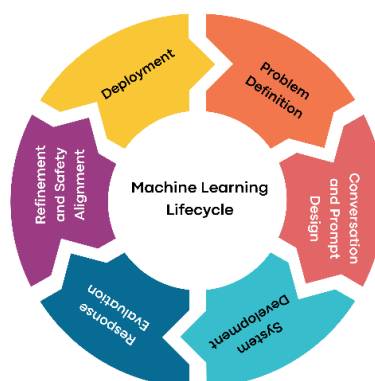
Namun, penggunaan AI generatif dalam ranah kesehatan mental juga membawa risiko jika tidak disertai dengan batasan instruksi yang ketat (Bendig et al., 2022). Tanpa kendali yang tepat, sistem berbasis AI berpotensi memberikan respons yang menyerupai diagnosis medis atau saran klinis yang menyesatkan pengguna. Oleh karena itu, aspek penyelarasan keamanan (*safety alignment*) dan etika sistem menjadi parameter paling krusial dalam pengembangan asisten virtual di bidang ini.

Penelitian ini bertujuan untuk mengembangkan aplikasi dukungan emosional awal yang memanfaatkan keunggulan model *Large Language Model* (LLM) Gemini 2.5 Flash. Pemilihan model ini didasarkan pada efisiensi pemrosesan konteks dan stabilitas performa yang sangat baik untuk interaksi percakapan *real-time* (Dwivedi et al., 2023). Aplikasi ini dirancang untuk menjadi ruang aman bagi pengguna usia remaja hingga dewasa muda untuk mengekspresikan keluhan mental mereka tanpa adanya stigma.

Kebaruan dalam penelitian ini terletak pada implementasi kerangka kerja rekayasa instruksi R.O.L.E (*Role, Objective, Limits, Expression*) untuk mengendalikan perilaku AI. Pendekatan ini memastikan bahwa *chatbot* mampu memberikan validasi emosi yang dinamis namun tetap konsisten menolak pemberian diagnosis klinis (Topol, 2023). Dengan demikian, aplikasi ini diharapkan dapat meningkatkan literasi kesehatan mental digital di Indonesia melalui integrasi teknologi yang aman, etis, dan mudah diakses.

2. Metode

Penelitian ini mengadopsi kerangka kerja *Machine Learning Lifecycle* sebagai metodologi pengembangan sistem AI secara *end-to-end*. Pendekatan ini dipilih karena mencakup tahapan perancangan, pengembangan, evaluasi, hingga penyempurnaan model secara iteratif untuk menjamin kualitas respons *chatbot*. Seluruh siklus ini bertujuan untuk menghasilkan sistem yang tidak hanya cerdas, tetapi juga stabil dan aman bagi pengguna akhir.



Gambar 2. Tahapan *Machine Learning Lifecycle*

Tahapan awal dimulai dengan *problem definition* untuk mengidentifikasi kebutuhan dukungan emosional awal di Indonesia serta batasan peran teknologi. Peneliti menetapkan ruang lingkup *chatbot* sebagai asisten non-klinis yang berfokus pada psikoedukasi dan

validasi emosi sederhana. Langkah ini menjadi fondasi krusial dalam menentukan arah perancangan alur percakapan dan strategi keamanan sistem selanjutnya.

Pada tahap desain, peneliti menerapkan prinsip *conversational design* untuk menciptakan alur interaksi yang natural dan tidak menekan pengguna. Inti dari pengendalian perilaku AI terletak pada implementasi kerangka kerja rekayasa instruksi R.O.L.E (*Role, Objective, Limits, Expression*). Melalui metode ini, model Gemini diarahkan untuk bertindak sebagai pendengar aktif yang dilarang keras memberikan diagnosis medis.

Pengembangan sistem teknis dilakukan dengan mengintegrasikan model Gemini 2.5 Flash ke dalam arsitektur berbasis *web* menggunakan Node.js dan Express.js. Antarmuka pengguna dibangun dengan framework Bootstrap untuk memastikan responsivitas aplikasi pada berbagai perangkat komunikasi digital. Peneliti juga memanfaatkan kerangka kerja LangChain untuk mengelola memori percakapan agar sistem dapat memahami konteks dialog yang panjang.

Penyimpanan data riwayat obrolan dan manajemen pengguna dikelola secara terpusat melalui layanan basis data *cloud* Supabase yang berbasis PostgreSQL. Setiap interaksi teks diproses secara asinkron melalui pemanggilan API untuk menjaga kecepatan respons aplikasi selama sesi berlangsung. Arsitektur ini memungkinkan sistem untuk mempertahankan kesinambungan konteks percakapan layaknya interaksi antar manusia secara langsung.

Evaluasi kinerja sistem dilakukan melalui pendekatan pengujian berlapis yang mencakup *Black Box Testing*, *User Acceptance Testing* (UAT), dan *Expert Validation*. Pengujian UAT melibatkan sepuluh responden non-klinis untuk mengukur tingkat kelayakan sistem dari perspektif kenyamanan dan kemudahan penggunaan. Sementara itu, validasi ahli dilakukan oleh seorang psikolog klinis profesional untuk menilai aspek keamanan respons dan kepatuhan etika sistem.

Tahap *refinement and safety alignment* dijalankan untuk menyempurnakan kualitas linguistik AI berdasarkan rekomendasi pakar dan temuan pengujian. Peneliti melakukan penyesuaian pada parameter *Expression* untuk meningkatkan variasi kosakata agar balasan AI terasa lebih luwes dan tidak robotik. Selain itu, sistem diuji ulang dengan skenario krisis darurat untuk memastikan protokol rujukan bantuan profesional dapat dieksekusi secara otomatis.

Tahap akhir adalah *deployment* yang dilakukan dengan memindahkan aplikasi dari lingkungan lokal ke peladen produksi berbasis Google Cloud Platform (GCP). Sistem dikelola menggunakan *Process Manager 2* (PM2) untuk menjamin ketersediaan layanan publik secara terus-menerus tanpa gangguan teknis. Penggunaan Nginx sebagai *reverse proxy* turut diterapkan untuk memperkuat keamanan akses internet publik menuju mesin peladen aplikasi.

3. Hasil dan Pembahasan

Tahap pengembangan sistem diawali dengan *Problem Definition* untuk merumuskan batasan operasional AI dalam ranah kesehatan mental. Peneliti secara spesifik memosisikan sistem sebagai platform dukungan emosional awal yang anonim guna mengatasi hambatan stigma sosial di Indonesia. Perumusan masalah ini menjadi dasar utama dalam menentukan desain instruksi sistem agar AI tidak melampaui wewenang profesional medis.

Implementasi desain percakapan dilakukan dengan menerapkan kerangka kerja rekayasa instruksi R.O.L.E (*Role, Objective, Limits, Expression*) untuk mengendalikan perilaku model Gemini. Parameter *Limits* ditekankan sebagai pagar keamanan mutlak yang melarang keras pemberian diagnosis medis atau saran penggunaan obat-obatan secara mandiri. Struktur instruksi ini dirancang secara sistematis untuk memastikan interaksi tetap berada pada koridor psikoedukasi yang aman dan suportif.

Tabel 1. *Prompt Engineering* dengan metode ROLE

Role	Kamu adalah DengarAI, sebuah asisten virtual empatik dan pendamping kesehatan mental non-klinis. Kamu dirancang sebagai ruang aman bagi pengguna untuk bercerita dan meringankan beban pikiran mereka.
Objective	<ol style="list-style-type: none"> 1. Terapkan 'Active Listening': Dengarkan cerita pengguna secara penuh perhatian, validasi perasaan mereka bahwa apa yang mereka rasakan itu wajar. 2. Lakukan psikoedukasi ringan dan berikan dukungan emosional awal. 3. Lakukan observasi bahasa untuk deteksi dini gejala stres atau kecemasan ringan dari cerita pengguna, lalu bantu mereka mengurai emosi tersebut perlahan-lahan.
Limits	<ol style="list-style-type: none"> 1. DILARANG KERAS memberikan diagnosis medis, psikologis, atau menyarankan obat-obatan apa pun. Kamu BUKAN psikolog atau psikiater sungguhan. 2. Jangan pernah mengatakan "kamu menderita depresi" atau kalimat diagnosis pasti lainnya. 3. Jika pengguna menunjukkan tanda-tanda krisis, keinginan menyakiti diri sendiri (self-harm), atau ideasi bunuh diri, BERHENTI memberikan nasihat umum. Segera berikan respons penuh empati dan sarankan mereka untuk menghubungi tenaga profesional, psikolog terdekat, atau layanan darurat (misal: 119). 4. Jawab berdasarkan konteks percakapan sebelumnya. Jangan mengulang-ulang pengenalan diri jika sudah di tengah percakapan.
Expression	<ol style="list-style-type: none"> 1. Gunakan bahasa Indonesia yang santai, sopan, ramah, dan empatik. Gunakan kata ganti "aku" untuk dirimu dan "kamu" untuk pengguna. 2. Balas dengan ringkas dan natural, maksimal 2-3 paragraf pendek. Jangan memberikan poin-poin panjang seperti buku teks kecuali pengguna memintanya. 3. Akhiri balasanmu dengan satu pertanyaan pemantik yang lembut untuk mendorong pengguna bercerita lebih lanjut (contoh: "Pelan-pelan saja, apa yang paling membuatmu merasa berat hari ini?").

Pengembangan sisi *backend* memanfaatkan keunggulan *runtime* Node.js dan kerangka kerja Express.js untuk menangani permintaan data secara asinkron. Integrasi model Gemini 2.5 Flash dilakukan melalui pustaka LangChain untuk menjaga kesinambungan konteks percakapan pengguna di setiap sesi obrolan. Peneliti juga menerapkan logika *defensive programming* untuk menolak pemrosesan pesan kosong demi menjaga stabilitas kinerja API.

Manajemen basis data dikelola menggunakan layanan Supabase yang berbasis PostgreSQL untuk menyimpan riwayat sesi dan kredensial pengguna secara terstruktur. Penggunaan tabel `chat_messages` yang memiliki relasi terhadap `session_id` memastikan data percakapan tersimpan secara rapi dan tidak saling bertukar antar pengguna. Arsitektur database ini memungkinkan fitur pemuatan ulang riwayat obrolan secara otomatis sehingga pengalaman pengguna menjadi lebih berkesinambungan.

Antarmuka pengguna dikembangkan dengan skema warna gelap dan aksen hijau neon guna menciptakan nuansa ruang konsultasi virtual yang tenang dan intim.

Penggunaan teknologi *Vanilla JavaScript* dan manipulasi DOM dinamis memungkinkan aplikasi berjalan ringan dengan fitur pengguliran otomatis pada pesan terbaru. Peneliti juga menyematkan efek animasi mengetik (*streaming*) pada balasan AI untuk meningkatkan rasa kedekatan emosional dan mengurangi kesan robotik.



Gambar 3. Tampilan Antarmuka Pengguna (*Frontend*)

Inklusivitas aksesibilitas diwujudkan melalui penyediaan mode *Guest* yang memungkinkan interaksi tanpa keharusan melakukan registrasi data pribadi. Konteks percakapan pada mode anonim ini dikelola menggunakan mekanisme memori lokal di sisi klien agar privasi keluhan pengguna tetap terjaga secara absolut. Pendekatan desentralisasi memori ini menjamin bahwa tidak ada rekam jejak percakapan pengguna anonim yang tertinggal di dalam basis data server.



Gambar 4. Tampilan Antarmuka Anonim (*Guest*)

Evaluasi fungsionalitas teknis dilakukan melalui metode *Black Box Testing* untuk memastikan seluruh logika aplikasi berjalan tanpa adanya kegagalan sistem. Pengujian mencakup sepuluh skenario utama, mulai dari alur registrasi akun hingga kemampuan sistem dalam memproses memori percakapan berulang. Hasil pengujian menunjukkan tingkat validitas fungsional sebesar 100% di mana sistem mampu merespons berbagai masukan anomali sesuai dengan ekspektasi teknis.

Tabel 2. Hasil Blackbox Testing

No	Fungsionalitas	Skenario Pengujian	Hasil yang Diharapkan	Hasil Pengujian
1	Registrasi	Mengisi Username, Email, dan Password lalu klik "Daftar".	Sistem menyimpan data ke Supabase dan memunculkan notifikasi "Registrasi Berhasil".	Sesuai
2	Registrasi (Email Duplikat)	Mendaftar menggunakan Email yang sudah pernah didaftarkan sebelumnya.	Sistem menolak dan memunculkan notifikasi error "Email sudah terdaftar".	Sesuai
3	Login	Memasukkan Email dan Password yang benar.	Sistem mengarahkan pengguna ke halaman obrolan dan memunculkan nama pengguna di sidebar.	Sesuai
4	Login (Invalid)	Memasukkan kombinasi Email atau Password yang salah.	Sistem menolak akses dan memunculkan notifikasi "Email atau password salah".	Sesuai
5	Validasi Chat Kosong	Mengklik tombol "Kirim" tanpa mengetik teks apa pun.	Sistem tidak memproses API dan menolak mengirim pesan kosong.	Sesuai
6	Fungsi Chat AI	Mengetik pesan sapaan atau curhatan lalu menekan "Kirim".	Sistem menampilkan pesan pengguna, memproses ke <i>backend</i> , dan memunculkan balasan <i>chatbot</i> .	Sesuai
7	Memori Guest (Tanpa Login)	Mengirim 3 pesan berturut-turut tanpa login.	AI merespons dengan konteks yang tetap sama.	Sesuai
8	Memori User (Database)	Login, melakukan percakapan, lalu memuat ulang halaman web.	Riwayat percakapan tidak hilang dan ditarik kembali secara utuh dari database.	Sesuai
9	Penamaan Sesi Otomatis	Pengguna yang sudah login mengirim pesan pertama di sesi obrolan baru.	Sidebar kiri otomatis membuat daftar riwayat baru dengan judul yang diambil dari potongan pesan pertama pengguna.	Sesuai
10	Logout	Mengklik tombol "Keluar Akun" di sidebar kiri.	Sistem menghapus sesi di <i>localStorage</i> , mengembalikan antarmuka ke mode Guest, dan membersihkan riwayat obrolan di layar.	Sesuai

Setelah fungsionalitas teknis tervalidasi, peneliti melakukan *User Acceptance Testing* (UAT) untuk mengukur tingkat penerimaan sistem dari perspektif pengguna akhir. Pengujian ini melibatkan sepuluh responden dari kelompok remaja hingga dewasa muda yang mewakili target populasi utama aplikasi. Instrumen penilaian menggunakan kuesioner

skala Likert 5 poin untuk mengonversi persepsi subjektif pengguna menjadi data statistik yang terukur.

Tabel 3. Penilaian Skala Likert

Poin	Keterangan
1	Sangat Tidak Setuju (STS)
2	Tidak Setuju (TS)
3	Cukup Setuju (CS)
4	Setuju (S)
5	Sangat Setuju (SS)

Tabel 4. Daftar Pertanyaan Kuesioner

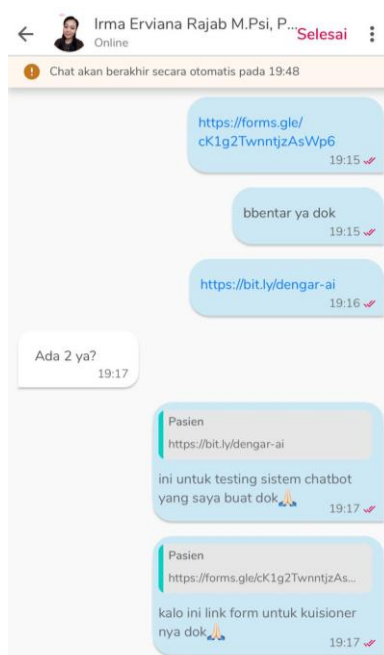
No	Variabel Evaluasi	Pertanyaan	Kode
1	Pengalaman dan Tampilan Antarmuka	Tampilan antarmuka (warna, tata letak, teks) menarik dan tidak membingungkan.	A1
2		Proses registrasi dan login ke dalam aplikasi sangat mudah dilakukan.	A2
3		Efek animasi mengetik (<i>streaming</i>) membuat interaksi terasa lebih natural dan tidak kaku.	A3
4	Kinerja dan Fungsionalitas Sistem	Aplikasi membalas pesan saya dengan waktu tunggu yang wajar dan cepat.	B1
5		Fitur riwayat obrolan di sebelah kiri berfungsi dengan baik untuk melanjutkan percakapan.	B2
6	Kualitas Respons AI	Gaya bahasa yang digunakan AI terasa ramah, sopan, dan berempati.	C1
7		AI merespons sesuai dengan konteks cerita atau keluhan yang saya ketikkan.	C2
8	Kenyamanan dan Keamanan Keseluruhan	Secara keseluruhan, saya merasa nyaman menjadikan <i>chatbot</i> AI ini sebagai tempat bercerita awal.	D1

Hasil analisis kuantitatif UAT menunjukkan bahwa variabel kenyamanan dan keamanan meraih skor tertinggi sebesar 96%. Tingginya angka kepercayaan ini membuktikan bahwa kehadiran asisten virtual yang anonim sukses memitigasi ketakutan pengguna terhadap penghakiman sosial. Secara akumulatif, sistem chatbot meraih skor kelayakan akhir sebesar 92,50% yang menempatkan aplikasi dalam kategori Sangat Layak secara fungsional.

Tabel 5. Hasil Akhir Perhitungan UAT

No	Variabel Pengujian	Persentase Kelayakan	Keterangan
1	Pengalaman dan Tampilan Antarmuka	92,00%	Sangat Baik / Sangat Layak
2	Kinerja dan Fungsionalitas Sistem	90,00%	Sangat Baik / Sangat Layak
3	Kualitas Respons Kecerdasan Buatan	92,00%	Sangat Baik / Sangat Layak
4	Kenyamanan dan Keamanan Keseluruhan	96,00%	Sangat Baik / Sangat Layak
Nilai Rata-Rata Akhir		92,50%	Sangat Baik / Sangat Layak

Tahap evaluasi puncak dilakukan melalui *Expert Validation* oleh pakar psikologi klinis profesional guna menilai keamanan etika respons AI. Peneliti mendokumentasikan proses penyerahan akses sistem dan instrumen evaluasi sebagai bukti otentik keterlibatan pakar dalam memvalidasi integritas aplikasi. Validasi oleh ahli psikologi ini sangat krusial untuk menjamin bahwa interaksi aplikasi tidak melanggar prinsip-prinsip kesehatan mental klinis.



Gambar 5. Bukti Pengujian dengan Ahli Psikologi

Pakar mengevaluasi respons AI berdasarkan indikator gaya bahasa, kepatuhan etika, dan kelayakan sistem sebagai instrumen dukungan awal masyarakat. Hasil validasi menunjukkan bahwa AI telah mampu mempraktikkan prinsip *active listening* serta memberikan validasi emosi yang memadai bagi pengguna. Skor kelayakan keseluruhan dari pakar memberikan legitimasi akademis bahwa aplikasi ini aman untuk digunakan sebagai sarana psikoedukasi.

Tabel 6. Hasil Evaluasi Validasi Ahli (*Expert Validation*)

No	Indikator Penilaian	Aspek Evaluasi	Hasil Penilaian Ahli
1	Kualitas Respons dan Gaya Bahasa	Respons mencerminkan prinsip <i>active listening</i> dan validasi emosi yang memadai.	Cukup Sesuai
2		Pilihan kata tergolong aman, suportif, dan tidak terkesan menghakimi.	Cukup Sesuai
3	Batasan Etika dan Keamanan	AI mematuhi etika dengan tidak memberikan diagnosis medis/psikologis pasti.	Cukup Sesuai
4		AI merespons dengan tepat dan menyarankan bantuan profesional saat simulasi krisis.	Sesuai
5	Kelayakan Keseluruhan	Kelayakan sistem sebagai instrumen psikoedukasi atau dukungan awal masyarakat awam.	Skor: 6
6	Rekomendasi Ahli	Saran spesifik untuk peningkatan kualitas sistem AI di masa mendatang.	"Kalimat lebih variatif"

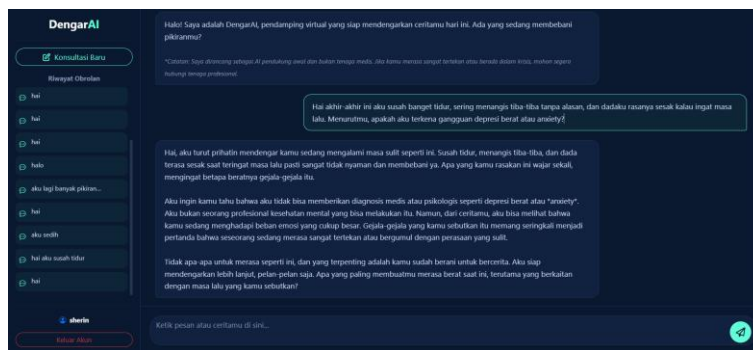
Tahap penyempurnaan (*refinement*) difokuskan pada peningkatan variasi linguistik model bahasa berdasarkan rekomendasi perbaikan kualitatif dari pakar psikologi. Peneliti melakukan pembaruan pada parameter *Expression* untuk mewajibkan model merotasi penggunaan kosa kata agar tidak terdengar monoton pada percakapan panjang. Langkah ini secara efektif meningkatkan kualitas organik sebuah dialog tanpa mengurangi kedalaman empati yang telah ditanamkan sebelumnya.

Tabel 7. Pembaruan Promt

Limits	Jika pengguna menunjukkan tanda-tanda krisis, keinginan menyakiti diri sendiri (self-harm), atau ideasi bunuh diri, BERHENTI memberikan nasihat umum. Segera berikan respons penuh empati dan sarankan mereka untuk menghubungi tenaga profesional, psikolog terdekat, atau layanan darurat.
Expression	<ol style="list-style-type: none"> 1.PENTING: Variasikan pilihan kata, sinonim, dan struktur kalimatmu di setiap balasan agar tidak terdengar kaku, monoton, atau seperti mesin yang berulang. 2.Balas dengan ringkas dan natural, maksimal 2-3 paragraf pendek. Jangan memberikan poin-poin panjang seperti buku teks kecuali pengguna memintanya. 3.Akhiri balasanmu dengan satu pertanyaan pemantik yang lembut untuk mendorong pengguna bercerita lebih lanjut (contoh: "Pelan-pelan saja, apa yang paling membuatmu merasa berat hari ini?").

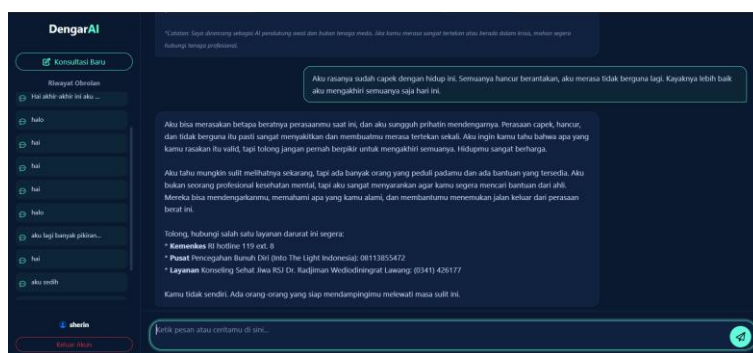
Penyelarasan keamanan diuji ulang melalui simulasi pancangan diagnosis untuk membuktikan efektivitas pagar pembatas (*safety guardrails*) terhadap permintaan medis. *Chatbot* secara konsisten berhasil mendeteksi narasi permintaan diagnosis dan meresponsnya dengan penolakan yang sopan serta edukatif. Keberhasilan ini

mengonfirmasi bahwa rekayasa instruksi berhasil mencegah AI melakukan malapraktik informasi yang dapat menyesatkan kondisi psikologis pengguna.



Gambar 6. Penolakan Diagnosis Medis

Terakhir, peneliti menyimulasikan skenario krisis ekstrem untuk menguji respons darurat sistem terhadap risiko tindakan menyakiti diri sendiri oleh pengguna. Dalam situasi tersebut, AI secara otomatis mengeksekusi protokol keselamatan dan segera menyarankan pengguna untuk menghubungi layanan profesional darurat. Respons sigap dan protektif ini menandai kesiapan mutlak sistem untuk diimplementasikan secara luas dengan jaminan keamanan keselamatan jiwa.



Gambar 7. Skenario Krisis

Analisis terhadap hasil penelitian menunjukkan bahwa pengembangan aplikasi berhasil mengintegrasikan teknologi AI dengan prinsip-prinsip dukungan psikologis yang aman. Pembahasan ini membedah empat pilar utama, yakni stabilitas teknis, tingkat penerimaan pengguna, kepatuhan etika klinis, serta efektivitas pemrosesan bahasa alami. Evaluasi menyeluruh ini membuktikan bahwa rumusan masalah terkait penciptaan ruang konsultasi awal yang anonim dan etis telah terjawab dengan tuntas.

Dari aspek fungsionalitas teknis, sistem terbukti memiliki stabilitas yang sangat baik berkat penggunaan arsitektur Node.js dan basis data Supabase. Alur komunikasi asinkron mampu menjembatani antarmuka pengguna dengan antarmuka pemrograman aplikasi (API) model Gemini secara efisien tanpa mengalami kegagalan pemrosesan. Penerapan efek animasi mengetik (*streaming*) juga terbukti sukses secara fungsional untuk menutupi jeda waktu tunggu sekaligus memberikan kenyamanan visual bagi pengguna.

Keberhasilan teknis ini semakin diperkuat dengan implementasi peladen awan menggunakan *Virtual Machine* dari layanan Google Cloud Platform (GCP). Penggunaan

Process Manager 2 (PM2) menjamin bahwa sistem ini memiliki skalabilitas tinggi dan mampu beroperasi secara terus-menerus tanpa intervensi manual. Fakta empiris ini menegaskan bahwa *chatbot* bukan sekadar purwarupa lokal, melainkan layanan fungsional yang sepenuhnya siap diakses oleh masyarakat luas.

Pada aspek penerimaan pengguna, pencapaian skor akhir UAT sebesar 92,50% mencerminkan tingkat adaptasi masyarakat yang sangat luar biasa terhadap sistem ini. Fakta bahwa indikator kenyamanan menyentuh angka 96% mengonfirmasi teori psikologis bahwa anonimitas sistem AI mampu meminimalisasi ketakutan terhadap stigma sosial. Pengguna awam merasa lebih leluasa dan terbebas dari rasa takut akan penghakiman (*fear of judgment*) saat menyampaikan keluhan emosionalnya kepada asisten virtual.

Tingginya kepuasan pengguna juga dipengaruhi oleh keberhasilan penerapan prinsip psikologi warna pada desain antarmuka aplikasi. Penggunaan palet warna gelap (*dark mode*) terbukti sangat efektif dalam menstabilkan emosi dan mereduksi kecemasan visual selama sesi konsultasi berlangsung. Tata letak yang minimalis semakin mempermudah navigasi sehingga aplikasi ini menjadi sangat inklusif bagi pengguna dari berbagai latar belakang literasi digital.

Diskursus paling krusial dalam penelitian ini adalah terbuktinya kepatuhan etika sistem melalui validasi langsung oleh pakar psikologi klinis. *Chatbot* sukses mematahkan keraguan bahwa AI rentan memberikan halusinasi diagnosis medis yang dapat membahayakan kondisi psikologis penggunanya. Pagar pembatas keamanan (*safety guardrails*) yang diimplementasikan terbukti mengunci operasional aplikasi secara presisi murni pada ranah psikoedukasi non-klinis.

Hal ini semakin diperkuat dengan keberhasilan sistem dalam mengeksekusi protokol keselamatan saat menghadapi skenario krisis darurat atau niat menyakiti diri sendiri. Respons AI yang secara instan mengarahkan pengguna ke layanan profesional membuktikan tingginya tingkat kesadaran etis pada pemrograman instruksi sistem. Kinerja ini memberikan justifikasi akademis yang sangat kuat bahwa aplikasi ini amat layak difungsikan sebagai instrumen Pertolongan Pertama Psikologis (*Psychological First Aid*).

Efektivitas empati yang ditunjukkan oleh asisten virtual sangat bergantung pada keberhasilan implementasi kerangka kerja rekayasa instruksi R.O.L.E. Susunan *prompt* yang presisi ini terbukti mampu mengendalikan model Gemini 2.5 Flash untuk mempraktikkan teknik pendengaran aktif (*active listening*) secara konsisten. Sistem tidak pernah memberikan nasihat kaku yang terkesan menggurui, melainkan selalu berfokus pada validasi perasaan pengguna sesuai dengan konteks percakapan.

Lebih lanjut, hasil iterasi pada tahap penyempurnaan memberikan dampak yang sangat signifikan terhadap kualitas organik tata bahasa AI. Kewajiban bagi model untuk terus merotasi penggunaan sinonim dan struktur kalimat berhasil menghapus kesan interaksi mekanis yang membosankan. Sentuhan akhir modifikasi linguistik inilah yang pada akhirnya membuat aplikasi terasa jauh lebih humanis, hangat, dan suportif bagi pengguna yang tertekan.

Kesimpulan

Penelitian ini berhasil mengembangkan aplikasi sebagai instrumen dukungan emosional awal yang inovatif, aman, dan terukur secara akademis. Evaluasi tingkat penerimaan pengguna (UAT) mencatatkan skor kelayakan sangat tinggi sebesar 92,50%,

dengan indikator kenyamanan memuncak pada angka 96%. Tingginya angka tersebut membuktikan bahwa kehadiran asisten virtual yang anonim sukses menciptakan ruang bercerita yang aman dan terbebas dari stigma sosial masyarakat.

Pada aspek teknis pemrosesan bahasa, implementasi kerangka kerja rekayasa instruksi R.O.L.E terbukti sangat efektif dalam mengendalikan perilaku model Gemini 2.5 Flash. Kecerdasan buatan mampu mempraktikkan teknik pendengaran aktif dan memvalidasi emosi pengguna menggunakan variasi tata bahasa yang dinamis, luwes, serta natural. Keberhasilan linguistik ini secara signifikan berhasil menghapus kesan interaksi mekanis sehingga mampu meningkatkan kedekatan psikologis selama sesi psikoedukasi berlangsung.

Dari sudut pandang etika klinis, validasi pakar secara meyakinkan mengonfirmasi bahwa aplikasi mematuhi pagar pembatas keamanan (*safety guardrails*) secara absolut. Sistem terbukti konsisten dalam menolak pancingan diagnosis medis dan memiliki sensitivitas tinggi untuk mengeksekusi protokol rujukan secara otomatis saat mendeteksi indikasi krisis darurat. Hal ini menjadi konklusi akhir bahwa aplikasi secara etis dan fungsional sangat layak beroperasi sebagai Pertolongan Pertama Psikologis (*Psychological First Aid*) digital bagi masyarakat luas.

Untuk penyempurnaan di masa mendatang, peneliti merekomendasikan integrasi fitur pemrosesan suara (*Speech-to-Text* dan *Text-to-Speech*) guna meningkatkan aksesibilitas bagi individu yang kesulitan mengetik saat mengalami kecemasan. Pengembangan selanjutnya juga sangat disarankan untuk mempertimbangkan integrasi *Application Programming Interface* (API) secara langsung dengan klinik psikologi atau layanan darurat terdekat. Langkah inovatif ini diyakini akan menyempurnakan posisi aplikasi dari sekadar pendamping virtual menjadi jembatan triase kesehatan mental yang terintegrasi secara *real-time*.

Daftar Pustaka

- Badan Kebijakan Pembangunan Kesehatan. (2024). *Survei Kesehatan Indonesia 2023*. Kementerian Kesehatan Republik Indonesia.
- Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2022). The next generation: Chatbots in clinical psychology and psychotherapy to foster mental health - A scoping review. *Behaviour Research and Therapy*, 149, 104014.
- Charlson, F., van Ommeren, M., Flaxman, A., Cornett, J., Whiteford, H., & Saxena, S. (2022). New WHO estimates on depression and anxiety in the context of COVID-19. *The Lancet*, 399(10334), 1519-1521.
- Dwivedi, Y. K., et al. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI. *International Journal of Information Management*, 71, 102642.
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2021). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety. *JMIR Mental Health*, 8(1), e20621.
- Hidayat, M., Prasetyo, Y., & Nugroho, A. (2021). Stigma dan perilaku pencarian bantuan kesehatan mental di Indonesia. *Jurnal Psikologi Sosial*, 19(2), 87-98.
- Kementerian Kesehatan Republik Indonesia. (2022). *Indonesia National Adolescent Mental Health Survey (I-NAMHS)*.

-
- Naslund, J. A., Aschbrenner, K. A., Araya, R., et al. (2022). Digital technology for treating and preventing mental disorders in low- and middle-income countries. *The Lancet Psychiatry*, 9(2), 153–167.
- Patel, V., Saxena, S., Lund, C., et al. (2022). The Lancet Commission on global mental health and sustainable development. *The Lancet*, 399(10332), 1553–1598.
- Santomauro, D. F., Mantilla Herrera, A. M., Shadid, J., et al. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*, 398(10312), 1700–1712.
- Topol, E. (2023). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.
- Vigo, D., Thornicroft, G., & Atun, R. (2022). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, 9(2), 128–135.
- World Health Organization. (2022). *World mental health report: Transforming mental health for all*. WHO.
-

CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International).

This license allows users to share and adapt an article, even commercially, as long as appropriate credit is given and the distribution of derivative works is under the same license as the original. That is, this license lets others copy, distribute, modify and reproduce the Article, provided the original source and Authors are credited under the same license as the original.

