Digital Transformation Technology (Digitech) | e-ISSN: 2807-9000
Volume 6, **Number** 1, Maret 2026
https://doi.org/10.47709/digitech.v6i1.7113

Implementasi *Indobert* dan *Large Language Model* pada Aplikasi *Chatbot* Adaptif berdasarkan Emosi Pengguna

Destira Lestari Saraswati1[1], Asep Wahyudin[2*], Yudi Ahmad Hambali[3]
[1,2,3]Universitas Pendidikan Indonesia, Indonesia
[1]dest.saras29@upi.edu, [2]away@upi.edu, [3]yudi.a.hambali@upi.edu

open access
Histori Artikel:
Diajukan: 21 Oktober 2025
Disetujui: 15 Februari 2026
Dipublikasi: 19 Februari 2026
Kata Kunci:
Chatbot, IndoBERT; Kesehatan mental; Large Language Model; System Usability Scale; User Experience Questionnaire.
Digital Transformation Technology (Digitech) *is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).*

Abstrak
Kesehatan mental menjadi isu krusial di Indonesia seiring meningkatnya prevalensi gangguan psikologis dan terbatasnya akses layanan profesional. Kolaborasi antara Kementerian Pemberdayaan Perempuan dan Perlindungan Anak (KemenPPPA) dengan UNICEF menegaskan pentingnya sinergi dukungan kesehatan mental di berbagai sektor, termasuk melalui media digital. Penelitian ini mengembangkan aplikasi kesehatan emosional berbasis *chatbot* adaptif yang memanfaatkan *IndoBERT* untuk klasifikasi emosi cinta, marah, sedih, bahagia, dan takut serta *Large Language Model* (LLM) untuk menghasilkan respons empatik dan kontekstual dalam Bahasa Indonesia. Pengembangan dilakukan menggunakan model ADDIE (*Analysis, Design, Development, Implementation, Evaluation*) dengan teknologi *Flutter*, *FastAPI*, dan *Firebase*. Evaluasi dilakukan melalui Blackbox Testing, *System Usability Scale* (SUS), dan *User Experience Questionnaire* (UEQ). Hasil menunjukkan seluruh fitur utama berfungsi sesuai spesifikasi tanpa kesalahan logika. Nilai SUS rata-rata mencapai 83,66, termasuk kategori *Excellent*, sedangkan hasil UEQ menunjukkan skor positif pada seluruh dimensi dengan rata-rata +2,045, terutama pada *Perspicuity* dan *Stimulation*. Temuan ini menunjukkan bahwa sistem mampu memberikan pengalaman interaksi yang nyaman, aman, dan adaptif terhadap kondisi emosional pengguna. Penelitian ini masih terbatas pada penerapan ilmu komputer, sehingga penelitian selanjutnya perlu ditinjau lebih lanjut dari perspektif keilmuan psikologi. Secara keseluruhan, penelitian ini membuktikan bahwa integrasi *IndoBERT* dan LLM efektif dalam membangun *chatbot* empatik yang relevan secara linguistik dan kultural, serta berpotensi menjadi dukungan awal dalam peningkatan kesadaran kesehatan emosional masyarakat Indonesia.

PENDAHULUAN

Kesehatan mental merupakan fondasi utama kesejahteraan individu yang memengaruhi cara seseorang berpikir, merasa, dan berperilaku dalam kehidupan sosialnya. Di Indonesia, isu kesehatan mental menjadi tantangan serius dan terus mengalami peningkatan. Berdasarkan laporan Riset Kesehatan Dasar (Riskesdas) yang dikutip oleh Kementerian Kesehatan Republik Indonesia (2023), sekitar 9,8% penduduk berusia di atas 15 tahun mengalami gangguan mental emosional, sedangkan 1 dari 5 orang dewasa berisiko mengalami gangguan jiwa berat seperti depresi atau skizofrenia. Namun, hanya 9% dari penderita gangguan jiwa berat yang mendapatkan penanganan profesional, menunjukkan kesenjangan besar antara kebutuhan dan akses layanan kesehatan mental di Indonesia.

Masalah ini semakin kompleks di kalangan remaja dan dewasa muda. Data UNICEF (2021) mengungkapkan bahwa lebih dari 15,5 juta anak dan remaja Indonesia mengalami gangguan psikologis, namun sebagian besar enggan mencari bantuan karena stigma sosial dan kurangnya layanan konseling yang terjangkau. Kolaborasi lintas lembaga, seperti antara Kementerian Pemberdayaan Perempuan dan Perlindungan Anak (KemenPPPA) dan UNICEF, menegaskan pentingnya sinergi dukungan kesehatan mental dan psikososial di berbagai lembaga, termasuk dunia pendidikan dan sektor *digital* (KemenPPPA, 2024).

Sementara itu, perkembangan teknologi informasi dan kecerdasan buatan membuka peluang baru dalam bidang *digital mental health*. Studi bibliometrik oleh Tresnayatna, Nurdin, & Wahyudin (2024) menunjukkan peningkatan signifikan dalam penggunaan *chatbot* untuk dukungan emosional dan administratif di sektor pendidikan selama pandemi COVID-19. Pendekatan *chatbot* memungkinkan pengguna mengekspresikan perasaan secara anonim, kapan pun, dan di mana pun tanpa rasa takut terhadap penilaian sosial. Penelitian Khairan & Habib (2025) juga menyoroti peran *chatbot* berbasis *Artificial Intelligent* (AI) dalam deteksi dini gangguan mental di Indonesia yang dapat meningkatkan kesediaan pengguna mencari bantuan hingga 40% dibandingkan metode survei tradisional.

Kemajuan teknologi *Natural Language Processing* (NLP) telah melahirkan model bahasa seperti

IndoBERT (Wilie et al., 2020) dan *IndoBERTweet* (Koto, Lau, & Baldwin, 2021), yang terbukti meningkatkan akurasi klasifikasi emosi hingga 94,9% dalam teks berbahasa Indonesia (Situmorang & Purba, 2024). Namun, sebagian besar penelitian di Indonesia masih terbatas pada klasifikasi emosi statis dan belum mengembangkan sistem percakapan adaptif yang mampu memberikan tanggapan empatik secara kontekstual.

Penelitian ini berangkat dari kesenjangan tersebut dengan mengusulkan pengembangan aplikasi kesehatan emosional adaptif berbasis *IndoBERT* dan *Large Language Model* (LLM) yang mampu mengenali, memahami, serta merespons emosi pengguna secara empatik dalam Bahasa Indonesia. Sistem ini dibangun menggunakan pendekatan *Analysis, Design, Development, Implementation, and Evaluation* (ADDIE) dengan teknologi *Flutter*, *FastAPI*, dan *Firebase*. Tujuan penelitian ini adalah untuk menghadirkan media *digital* interaktif yang aman dan adaptif dalam memberikan dukungan emosional, serta meningkatkan kesadaran masyarakat terhadap pentingnya kesehatan mental melalui teknologi berbasis AI.

STUDI LITERATUR

Perkembangan teknologi kecerdasan buatan telah memberikan dampak signifikan terhadap berbagai bidang, termasuk pendidikan, sosial, dan kesehatan mental. Studi bibliometrik yang dilakukan oleh Tresnayatna et al. (2024) menunjukkan bahwa penggunaan *chatbot* dalam pendidikan tidak hanya berfungsi sebagai alat bantu pembelajaran, tetapi juga sebagai sarana dukungan emosional bagi peserta didik. Selama pandemi COVID-19, *chatbot* terbukti mampu membantu siswa beradaptasi terhadap tekanan psikologis melalui komunikasi interaktif berbasis teks. Penelitian ini menegaskan pentingnya pendekatan berbasis empati dan kesetaraan gender dalam pengembangan sistem berbasis AI agar inklusif bagi semua pengguna.

Dalam ranah kesehatan, penerapan machine learning juga terus berkembang Hambali, Megasari, & Santoso (2023) mengimplementasikan algoritma *Evolving Artificial Neural Network* (EANN) untuk diagnosis penyakit tidak menular seperti diabetes. Hasilnya menunjukkan peningkatan akurasi prediksi, yang mengindikasikan bahwa metode machine learning memiliki fleksibilitas tinggi untuk digunakan di berbagai domain, termasuk psikologi *digital*. Sejalan dengan itu, Kaniawulan et al. (2022) melalui tinjauan sistematis menjelaskan bahwa transformasi *digital* pada sektor publik menghadapi tantangan besar, terutama dalam aspek keamanan data dan integrasi sistem. Temuan tersebut memperkuat perlunya pengembangan arsitektur aplikasi yang aman dan terukur dalam konteks *digital health*.

Kemajuan signifikan juga terjadi dalam bidang *Natural Language Processing* (NLP), terutama dengan hadirnya model bahasa berbasis *transformer* seperti *IndoBERT* (Wilie et al., 2020) dan *IndoBERTweet* (Koto et al., 2021). Kedua model ini dirancang khusus untuk memahami konteks linguistik Bahasa Indonesia, termasuk penggunaan bahasa informal dan kode campur yang sering ditemukan di media sosial. Shaw, LaCasse, & Champagne (2025) melanjutkan studi ini dengan menunjukkan bahwa pendekatan *transfer learning* berbasis *IndoBERT* meningkatkan performa analisis emosi secara signifikan dibandingkan metode tradisional.

Penerapan model NLP dalam bidang kesehatan mental telah dikaji dalam berbagai penelitian internasional. Shah et al. (2022) mengembangkan *chatbot* empatik untuk meningkatkan keterlibatan pasien dengan gangguan makan, yang menunjukkan peningkatan tingkat usability hingga 90%. Sementara itu, Khairan & Habib (2025) mengonfirmasi bahwa *chatbot* berbasis AI dapat berfungsi sebagai alat identifikasi awal gangguan mental di Indonesia, membantu pengguna mengekspresikan perasaan mereka tanpa takut stigma sosial. Esmi et al. (2025) menambahkan bahwa teknik *prompt engineering* berbasis masukan psikolog profesional dapat meningkatkan akurasi deteksi stres oleh GPT-4 sebesar 17% sekaligus menurunkan kesalahan deteksi positif palsu hingga 80%.

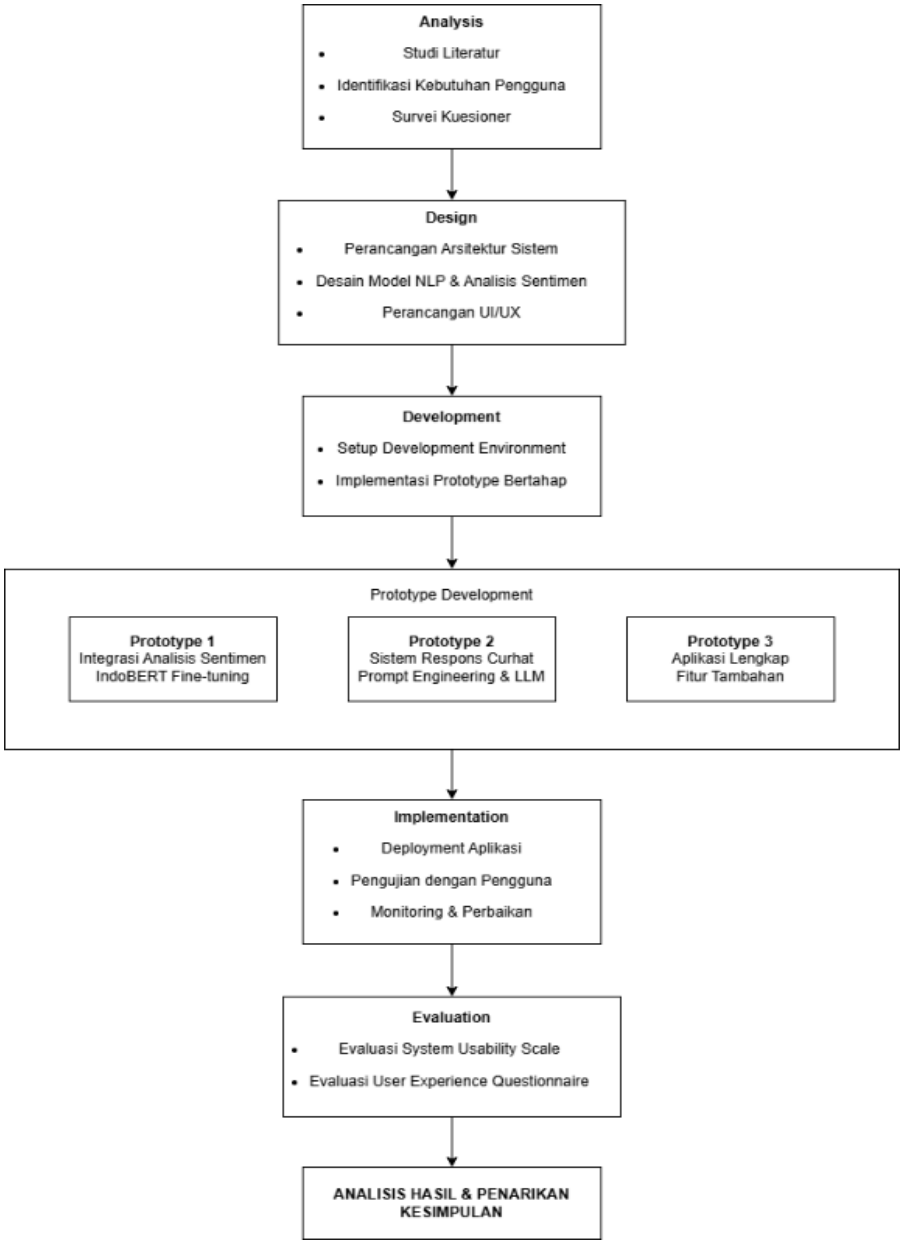
Penelitian implementatif di Indonesia juga mulai menunjukkan hasil positif Adzan, Roxana, & Cahyana (2025) mengembangkan aplikasi Curhat dengan integrasi GPT-4 untuk *mood tracking*, yang terbukti meningkatkan keterlibatan dan retensi pengguna. Di sisi lain, Situmorang & Purba (2024) melaporkan bahwa *IndoBERT* mencapai akurasi hingga 94,91% dalam mendeteksi potensi depresi berdasarkan unggahan media sosial. Hasil-hasil tersebut menegaskan efektivitas pendekatan NLP berbasis model bahasa Indonesia dalam memahami konteks emosional pengguna lokal.

Meskipun banyak penelitian telah mengangkat topik *chatbot* dan analisis emosi, sebagian besar masih berfokus pada klasifikasi emosi tanpa mempertimbangkan dinamika perubahan suasana hati dalam percakapan yang berkelanjutan. Selain itu, belum banyak penelitian yang menggabungkan kekuatan *IndoBERT* dengan *Large Language Model* (LLM) untuk menghasilkan respons adaptif yang empatik terhadap pengguna berbahasa Indonesia. Celah penelitian (*research gap*) inilah yang menjadi dasar bagi pengembangan aplikasi kesehatan emosional adaptif dalam studi ini, dengan tujuan menciptakan sistem percakapan cerdas yang tidak hanya memahami emosi, tetapi juga mampu berinteraksi secara humanis dan kontekstual sesuai budaya Indonesia.

METODE

Model pengembangan ADDIE (*Analysis, Design, Development, Implementation, Evaluation*) digunakan sebagai metode *prototyping* dalam penelitian ini (Abuhassna, Alnawajha, Awae, Mohamed Adnan, & Edwards, 2024). Setiap tahap dilakukan secara iteratif agar hasil akhir sesuai dengan kebutuhan pengguna dan dapat diuji

secara menyeluruh sebelum implementasi penuh, yang membuat model ADDIE pilihan yang tepat untuk proses pengembangan sistem berbasis pengguna. Selain itu, model ini banyak digunakan dalam pembuatan perangkat lunak interaktif dalam bidang psikologi *digital* dan pendidikan (Koto et al., 2021). Secara umum, langkah-langkah penelitian dapat diamati pada Gambar 1 yang menunjukkan proses pengembangan, mulai dari analisis kebutuhan hingga evaluasi sistem.



Gambar 1. Alur Penelitian

Analysis

Analisis tahap pertama berkonsentrasi pada penemuan masalah, kebutuhan pengguna, dan spesifikasi sistem. Mengevaluasi temuan penelitian sebelumnya tentang penggunaan model *IndoBERT* untuk klasifikasi emosi (Saputri, Mahendra, & Adriani, 2018; Shaw et al., 2025). Hasil analisis menunjukkan bahwa media berbicara *digital* yang aman, mudah digunakan, dan memberikan tanggapan empatik terhadap kondisi emosional sangat dibutuhkan oleh masyarakat.

Pada tahap ini, juga dilakukan observasi terhadap perilaku pengguna media sosial dan wawancara informal dengan individu berusia 18 hingga 25 tahun untuk mengetahui cara mereka mengungkapkan emosi secara *online*. Hasil analisis menunjukkan bahwa sebagian besar pengguna cenderung lebih suka mengungkapkan perasaan mereka melalui teks selama identitas mereka tetap anonim. Oleh karena itu, sistem yang dirancang harus dapat mengidentifikasi perasaan yang terkandung dalam teks dan memberikan respons yang sesuai tanpa

mengungkapkan data pribadi pengguna.

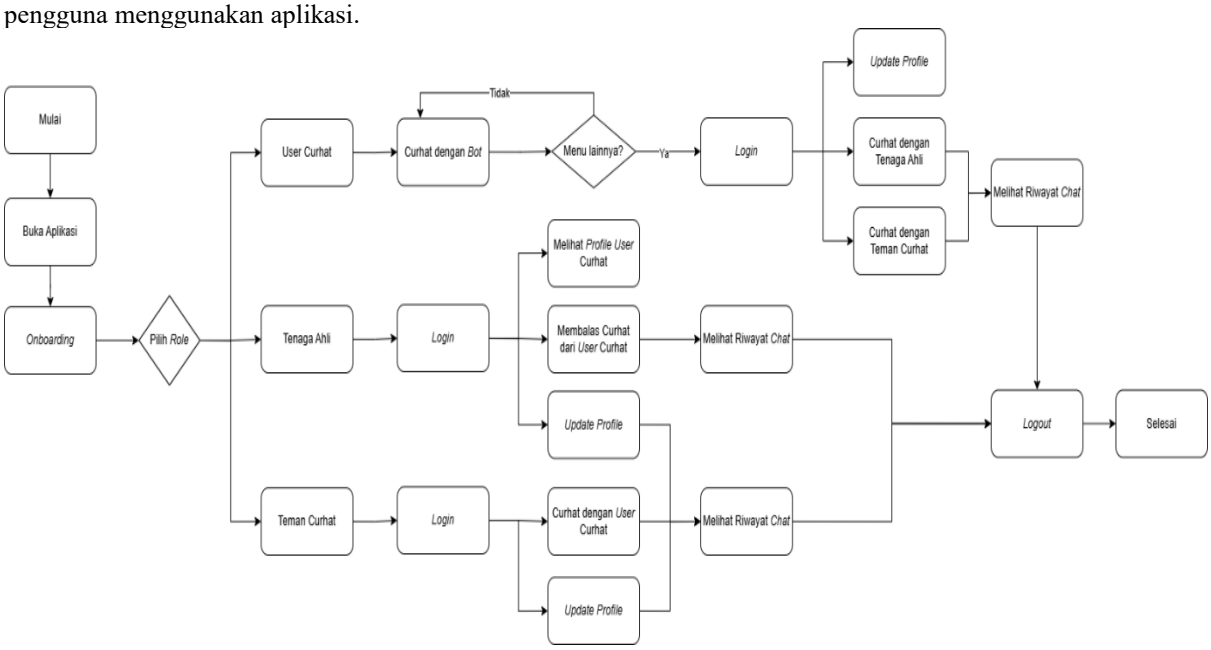
Selain persyaratan fungsional, adapula persyaratan non fungsional seperti, keamanan dan privasi data pengguna yang dijamin oleh autentikasi *Firebase* kinerja *real-time*, yang berarti sistem dapat menanggapi pesan pengguna dalam waktu kurang dari tiga detik; dan kemudahan penggunaan, sehingga aplikasi dapat diakses oleh orang-orang dengan berbagai latar belakang teknologi.

Pada tahap selanjutnya, hasil analisis kebutuhan ini akan berfungsi sebagai dasar untuk perancangan arsitektur sistem dan rancangan antarmuka pengguna.

Design

Frontend, *backend*, dan model analisis emosi diintegrasikan ke dalam desain sistem, dan tahap desain bertujuan untuk merancang struktur sistem, arsitektur data, dan alur interaksi pengguna dengan *chatbot*. Tiga elemen utama membentuk arsitektur sistem diantaranya, *Frontend* menggunakan *Flutter* yang berfungsi sebagai antarmuka pengguna; *backend* menggunakan *FastAPI* untuk menghubungkan aplikasi ke *server* dan *IndoBERT* dipakai untuk membangun model analisis emosi yang mengidentifikasi lima emosi: cinta, marah, sedih, bahagia, dan takut. Selain itu, modul *Large Language Model* (LLM) ditambahkan untuk memberikan respons yang kontekstual setelah mengenali perasaan pengguna lewat hasil analisis sentimen yang dilakukan dengan model *IndoBERT* sebelumnya.

Gambar 2 menunjukkan diagram alur sistem, yang menjelaskan bagaimana alur aplikasi berfungsi saat pengguna menggunakan aplikasi.



Gambar 2. Alur Aplikasi *Chatbot* Adaptif

Prinsip desain berpusat pada pengguna menyesuaikan tampilan dan interaksi sesuai dengan preferensi pengguna digunakan dalam desain antarmuka pengguna. Untuk memberikan kesan nyaman, gunakan warna netral yang lembut. Untuk membuat pengguna mudah berpindah antar halaman, elemen navigasi dirancang dengan sederhana. Di *Firebase*, setiap percakapan pengguna disimpan secara terenkripsi untuk menjaga privasi.

Development

Tahap pengembangan adalah proses mengubah desain menjadi aplikasi yang bermanfaat. Pengembangan menggunakan bahasa pemrograman Dart untuk *Flutter*, *Python* untuk *FastAPI*, dan API antara *server* dan model *IndoBERT*. Dataset emosi yang digunakan untuk pelatihan model *IndoBERT* merujuk pada dataset emosi berbahasa Indonesia yang diperkenalkan oleh Saputri et al. (2018). Dataset ini menyediakan data teks dengan label emosi yang jelas dan terstruktur yang terdiri atas lima kategori emosi utama. Meskipun diperkenalkan lebih awal, dataset tersebut masih digunakan dalam penelitian-penelitian terkini karena keterbatasan dataset emosi berbahasa Indonesia yang terstandarisasi dan tersedia secara publik (Heldiansyah & Winarko, 2022; Rusli, Rismala, & Nurrahmi, 2021).

Selanjutnya, model *IndoBERT* dioptimalkan dengan perbaikan tambahan melalui teknik *fine-tuning* untuk menjadi lebih sesuai dengan konteks percakapan Bahasa Indonesia sehari-hari. Modul LLM adalah model bahasa yang luar biasa yang dapat digunakan untuk membuat teks yang benar-benar mengenali dan relevan dengan emosi pengguna (Shah et al., 2022). *Chatbot* dapat menggunakan deteksi emosi dan respons berbasis konteks dengan

mengintegrasikan kedua model ini melalui API internal *server FastAPI*.

Selama tahap pengembangan, setiap modul diuji oleh unit untuk memastikan tidak ada kesalahan logika atau sintaks. Aplikasi diuji secara internal menggunakan *emulator Android Studio*, dan kemudian diperbarui hingga siap digunakan.

Implementation

Tahap implementasi mencakup pembuatan rencana uji fungsional dan penerapan aplikasi pada lingkungan nyata. Aplikasi ini dapat digunakan pada perangkat mobile berbasis *Android* karena berintegrasi penuh dengan *Flutter*, *FastAPI*, dan *Firebase*.

Metode pengujian yang digunakan adalah *Blackbox testing*, di mana setiap fungsi sistem diuji hanya dengan melihat masukan dan keluaran tanpa melihat kode sumber. Pendekatan pengujian fungsional tersebut masih digunakan secara luas pada pengembangan dan evaluasi sistem perangkat lunak(Ayuningtyas, WP, & Rachmadi, 2023). Skenario pengujian termasuk proses *login*, registrasi, mengirimkan teks dengan berbagai emosi, mendeteksi emosi oleh model, dan menyimpan hasil percakapan.

Setiap aktivitas pengujian dilakukan untuk memastikan bahwa fungsi memenuhi kebutuhan pengguna yang dibuat pada tahap analisis. Hasil pengujian ini tidak dibahas di sini, tetapi mereka dibahas secara rinci di bagian Hasil dan Pembahasan.

Evaluation

Tujuan dari tahap evaluasi ini adalah untuk mengevaluasi kualitas sistem dari perspektif kegunaan dan pengalaman pengguna setelah implementasi selesai. Dua alat utama yang digunakan untuk melakukan evaluasi dalam penelitian ini adalah instrumen usability terstandarisasi, yaitu *System Usability Scale* (SUS) dan *User Experience Questionnaire* (UEQ). Penggunaan instrumen tersebut didasarkan pada penelitian terkini yang mengevaluasi aplikasi kesehatan mental berbasis mobile dan memanfaatkan instrumen usability untuk menilai kemudahan penggunaan serta pengalaman pengguna (Høgsdal, Kaiser, & Kyrrestad, 2023; Schrepp, 2023).

Tahap ini menjelaskan rancangan proses evaluasi, bukan hasilnya. Penilaian dilakukan melalui penyusunan kuesioner dan perencanaan partisipasi responden. Kuesioner SUS digunakan untuk mengukur tingkat kemudahan penggunaan dan efisiensi antarmuka, sedangkan UEQ digunakan untuk menilai enam dimensi pengalaman pengguna (*Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation,* dan *Novelty*).

Responden terdiri dari pekerja muda dan mahasiswa, yang merupakan kelompok usia produktif yang pengguna aplikasi kesehatan mental *digital*. Selanjutnya, data dari kuesioner akan dianalisis secara deskriptif untuk mengevaluasi aspek kegunaan dan kenyamanan, yang akan dibahas pada bagian yang akan datang.

Oleh karena itu, tahap evaluasi bertanggung jawab untuk memastikan bahwa sistem tidak hanya beroperasi dengan benar, tetapi juga sesuai dengan ekspektasi pengguna dan prinsip desain empatik yang digunakan untuk membuat aplikasi ini dibuat.

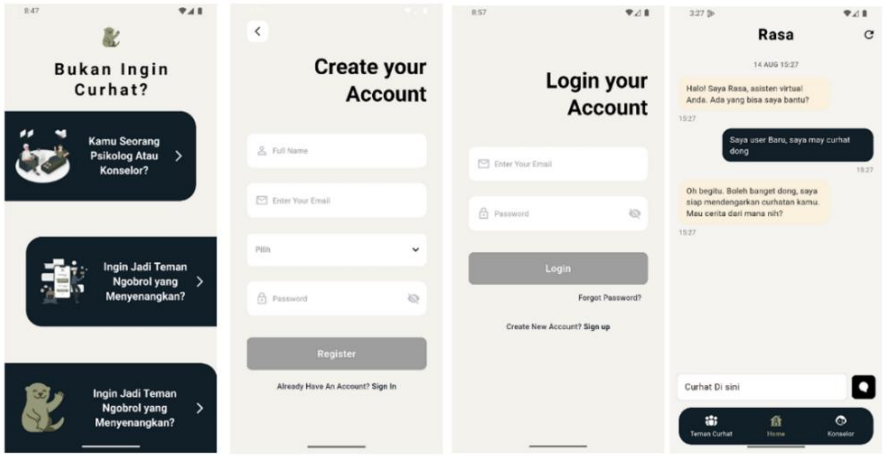
HASIL

Hasil Implementasi Aplikasi

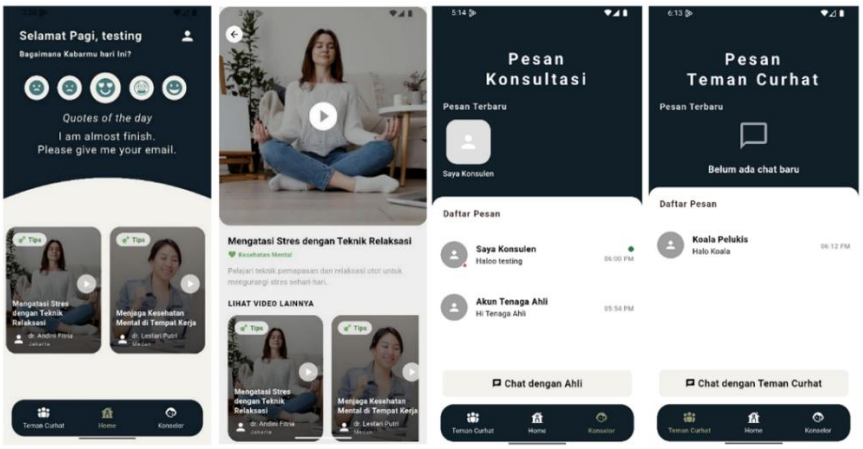
Hasil implementasi menunjukkan bahwa *chatbot* kesehatan emosional telah dirancang dengan baik pada tahap desain. Aplikasi ini dibangun dengan komponen berikut: *Flutter* UI *framework*, *FastAPI* sebagai *server backend*, *Firebase* sebagai sistem autentikasi dan penyimpanan data, dan integrasi model *IndoBERT* dan *Large Language Model* (LLM) untuk mengidentifikasi dan menanggapi emosi pengguna secara adaptif.

Aplikasi terdiri dari tiga lapisan utama: *Frontend*, yang bertanggung jawab atas interaksi pengguna, *backend*, yang bertanggung jawab atas komunikasi data, dan model analisis emosi, yang mengontrol proses klasifikasi teks. Arsitektur sistem yang diterapkan mengikuti desain pada tahap metodologi, di mana setiap komponen bekerja secara modular sehingga memudahkan proses pengujian dan pemeliharaan.

Antarmuka aplikasi terdiri atas beberapa halaman utama, yaitu halaman pilih *role*, halaman registrasi, halaman masuk, halaman percakapan *chatbot*. Tampilan antarmuka yang telah dikembangkan ditunjukkan pada Gambar 3 dan halaman beranda, halaman artikel kesehatan, halaman pesan konsultasi dengan tenaga ahli dan halaman pesan dengan teman curhat yang ditampilkan secara berurutan pada Gambar 4.



Gambar 3. Antarmuka Utama Aplikasi Bagian Pertama



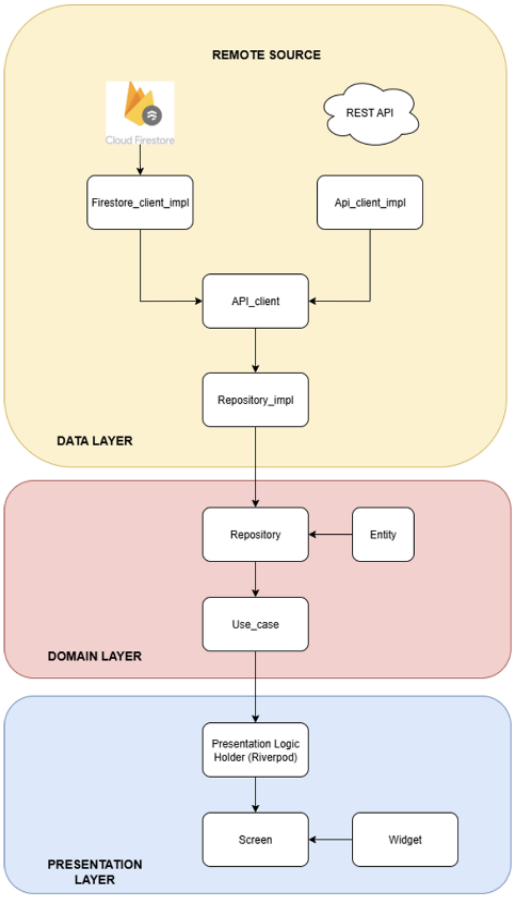
Gambar 4. Antarmuka Utama Aplikasi Bagian Kedua

Desain antarmuka menggunakan warna biru dan hijau yang menenangkan untuk menekankan kesederhanaan dan kenyamanan visual untuk menciptakan suasana yang tenang. Fitur percakapan didesain menyerupai aplikasi pesan instan agar pengguna merasa familiar. Dengan menggunakan model *IndoBERT*, *chatbot* akan menganalisis emosi dan memberikan tanggapan berdasarkan apa yang ditulis pengguna dalam Bahasa Indonesia alami.

Model *IndoBERT* mengategorikan emosi ke dalam lima kategori utama: cinta, marah, sedih, bahagia, dan takut. Response kontekstual dan empatik dihasilkan oleh modul *Large Language Model* (LLM). Sebagai contoh, ketika pengguna menulis "Aku merasa sendirian dan tidak tahu harus cerita ke siapa", sistem mengidentifikasi perasaan "sedih" dan memberikan tanggapan seperti "Terima kasih sudah berbagi, kamu bukan sendiri, dan itu wajar kok kamu merasa begitu".

Setiap percakapan disimpan secara terenkripsi di *Firebase*. Hal ini sangat penting untuk menjaga privasi pengguna dan juga menjadi fokus penelitian *chatbot* kesehatan mental lainnya seperti (Shah et al., 2022). Sistem penyimpanan berbasis cloud memungkinkan pengguna melanjutkan percakapan kapan pun mereka mau tanpa kehilangan konteks percakapan.

Implementasi aplikasi juga mempertimbangkan kinerja sistem dan responsivitasnya. Hasil deployment *chatbot* menunjukkan respons rata-rata 1.5 detik per pesan. Ini menunjukkan bahwa model analisis emosi dan *server FastAPI* bekerja sama dengan baik. Arsitektur menunjukkan bagaimana komponen sistem diintegrasikan pada Gambar 5.



Gambar 5. Arsitektur Implementasi Sistem

Secara keseluruhan, tahap implementasi membuat aplikasi stabil dan sesuai kebutuhan pengguna. Ini memberikan dasar untuk pengujian fungsional pada tahap berikutnya.

Implementation

Pengujian fungsional aplikasi dilakukan dengan metode *Blackbox* setelah diimplementasikan dengan sukses. Tujuan dari pengujian ini adalah untuk memastikan bahwa setiap fungsi sistem beroperasi dengan benar dan sesuai dengan kebutuhan yang telah ditetapkan. Autentikasi pengguna, deteksi emosi, *chatbot*, dan penyimpanan data *Firebase* adalah fitur utama yang diuji.

Dalam skenario pengujian, berbagai kondisi input, baik valid maupun tidak valid, digunakan. Misalnya, pesan kesalahan "Email dan kata sandi wajib diisi" akan ditampilkan ketika pengguna memasukkan data kosong pada kolom *login*. Sebaliknya, jika data yang dimasukkan benar, sistem akan mengarahkan pengguna ke halaman beranda. Hasil tes fungsional ditunjukkan pada Tabel 1.

Tabel 1
Hasil Pengujian Fungsional Menggunakan Metode *Blackbox*

Fitur yang Diuji	Input/Aksi	Expected Output	Result
Pilih *Role*	Memilih *role* yang tersedia	Diarahkan ke halaman *login*	Sesuai
Login (semua *role*)	Input Email dan password valid	Berhasil *login* dan diarahkan ke result halaman sesuai (user/ konsulen/teman)	Sesuai
Register User	Input data registrasi lengkap	Akun berhasil dibuat, diarahkan ke halaman *login*	Sesuai
Chat dengan *Bot* (*Guest*)	Input pesan	Bot memberikan respons	Sesuai
Melihat Halaman Beranda	User berhasil *login*	Menampilkan halaman beranda	Sesuai
Melihat Artikel Edukasi	Menekan salah satu	Menampilkan halaman artikel	Sesuai

	video edukasi pada halaman beranda	video edukasi	
Chat dengan Tenaga Ahli	Klik menu konselor	Menampilkan riwayat Chat dengan tenaga ahli	Sesuai
Chat dengan Teman Curhat	Klik menu teman curhat	Menampilkan riwayat Chat dengan teman curhat	Sesuai

Hasilnya menunjukkan bahwa semua fitur utama telah berfungsi sesuai ekspektasi. Selama pengujian, tidak ada kesalahan yang signifikan yang ditemukan. Hal ini menunjukkan bahwa desain sistem pada tahap desain telah dilakukan dengan benar dan integrasi antar komponen berjalan dengan baik.

Selain itu, uji keamanan menunjukkan bahwa orang lain tidak dapat mengakses akun *Firebase* tanpa izin pengguna. Pengujian ini sangat penting untuk mempertahankan kepercayaan pengguna dan sesuai dengan prinsip perlindungan data untuk aplikasi kesehatan *digital*.

Hasil Evaluasi *System Usability Scale* (SUS)

Setelah fungsionalitas sistem dinyatakan stabil, dilakukan evaluasi terhadap tingkat kegunaan aplikasi menggunakan instrumen *System Usability Scale* (SUS). Evaluasi dilakukan untuk mengetahui sejauh mana pengguna merasa aplikasi mudah digunakan, efisien, dan konsisten.

Kuesioner SUS terdiri atas 10 butir pernyataan dengan skala *Likert* 1–5. Evaluasi dilakukan kepada 41 responden yang mewakili kelompok usia generasi *millennial* dan Z. Responden diminta menggunakan aplikasi atau menonton demo setiap fitur aplikasi selama ±15 menit sebelum mengisi kuesioner. Hasil pengolahan data SUS ditunjukkan pada Tabel 2.

Tabel 2
Hasil Evaluasi Kegunaan Aplikasi Menggunakan SUS

Pernyataan	Skor Rata-Rata
Jumlah responden	41
Skor minimum	50.0
Skor maksimum	97.5
Rata-rata	83.66
Standar deviasi	10.92
Kategori interpretasi	***Excellent***

Skor total dihitung menggunakan rumus standar SUS, sebanyak 41 responden berpartisipasi dalam evaluasi SUS. Skor individu berkisar antara 50 hingga 97,5 dengan rata-rata 83,66 (kategori *Excellent*). Sebagian besar responden (85%) memperoleh skor di atas 80, menandakan tingkat kegunaan yang tinggi. Hasil ini menunjukkan bahwa aplikasi memiliki tingkat kegunaan tinggi dan dapat diterima pengguna tanpa kesulitan berarti.

Nilai SUS ini sejalan dengan penelitian Shah et al. (2022), yang menemukan bahwa *chatbot* kesehatan mental dengan desain sederhana cenderung memper*oleh* skor *usability* tinggi karena pengguna lebih fokus pada konten percakapan daripada kompleksitas navigasi.

Hasil Evaluasi *User Experience Questionnaire* (UEQ)

Selain aspek kegunaan, penelitian ini juga mengevaluasi pengalaman pengguna menggunakan *User Experience Questionnaire* (UEQ). Instrumen UEQ mengukur enam dimensi pengalaman pengguna: *Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation,* dan *Novelty*.

Kuesioner diisi oleh responden yang sama seperti pada pengujian SUS. Hasil rata-rata skor untuk setiap dimensi ditunjukkan pada Tabel 3.

Tabel 3
Hasil Kuesioner *User Experience Questionnaire*

Dimensi	Skor Rata-rata (−3…+3)	Interpretasi
Attractiveness	+2.07	*excellent*
Perspicuity	+2.1	*excellent*
Efficiency	+2.04	*excellent*
Dependability	+1.96	*excellent*
Stimulation	+2.09	*excellent*
Novelty	+2.01	*excellent*
Rata-rata keseluruhan	**+2.045**	***excellent***

Dari hasil tersebut, seluruh dimensi menunjukkan nilai positif, yang berarti aplikasi memberikan pengalaman yang baik bagi pengguna. Skor tertinggi terdapat pada Perspicuity dan Stimulation, menandakan bahwa aplikasi dianggap mudah dipahami dan menyenangkan digunakan.

Hasil ini mendukung temuan Abuhassna et al. (2024) bahwa model ADDIE efektif menghasilkan produk *digital* dengan tingkat kenyamanan pengguna yang tinggi karena desainnya berorientasi pada iterasi kebutuhan pengguna.

Selain itu, tingginya skor Stimulation dan Novelty menunjukkan bahwa *chatbot* yang dikembangkan memiliki unsur inovasi dan daya tarik yang memotivasi pengguna untuk berinteraksi kembali. Nilai-nilai ini penting dalam konteks aplikasi kesehatan emosional, di mana keterlibatan berulang dapat membantu pengguna dalam proses pengelolaan emosi mereka.

PEMBAHASAN

Hasil implementasi, pengujian, dan evaluasi menunjukkan bahwa aplikasi *chatbot* kesehatan emosional yang dikembangkan telah memenuhi seluruh tujuan penelitian. Integrasi model *IndoBERT* untuk analisis emosi dan LLM untuk respons adaptif menghasilkan sistem yang tidak hanya fungsional, tetapi juga empatik dalam merespons pengguna. Penerapan metode ADDIE terbukti efektif dalam pengembangan aplikasi ini karena setiap tahap dapat dievaluasi dan disesuaikan berdasarkan masukan pengguna. Pendekatan ini juga selaras dengan prinsip *user-centered design*, yang menempatkan kenyamanan dan kepuasan pengguna sebagai prioritas utama (Abuhassna et al., 2024).

Dari sisi fungsionalitas, hasil pengujian *Blackbox* menunjukkan seluruh fitur berjalan sesuai dengan spesifikasi yang ditetapkan. Hal ini menunjukkan bahwa pemisahan antara lapisan *Frontend* dan *backend* serta penggunaan API modular meningkatkan reliabilitas sistem. Dari sisi pengalaman pengguna, nilai SUS sebesar 83,66 membuktikan bahwa aplikasi mudah digunakan bahkan bagi pengguna dengan latar belakang teknologi yang terbatas. Hasil UEQ yang positif pada semua dimensi juga memperkuat bahwa aplikasi ini tidak hanya efektif secara teknis, tetapi juga menyenangkan digunakan.

Hasil penelitian ini menjadi sejalan dengan penelitian internasional oleh Shah et al. (2022), yang menggambarkan bahwa *chatbot* empatik dapat pula meningkatkan persepsi dukungan emosional pengguna. Secara keseluruhan, penelitian ini berhasil menunjukkan bahwa penerapan teknologi NLP berbasis *IndoBERT* dan LLM dapat digunakan dalam konteks kesehatan emosional di Indonesia. Aplikasi ini diharapkan dapat menjadi pelengkap layanan psikologis profesional dengan memberikan dukungan awal bagi pengguna yang membutuhkan tempat aman untuk mengekspresikan perasaannya.

KESIMPULAN

Penelitian ini berhasil mengembangkan aplikasi chatbot kesehatan emosional berbasis *IndoBERT* dan *Large Language Model* (LLM) menggunakan pendekatan ADDIE untuk membantu pengguna mengekspresikan emosi, adaptif dalam perubahan emosi, serta memperoleh respons yang kontekstual dan empatik dalam Bahasa Indonesia. Aplikasi ini dilengkapi fitur autentikasi pengguna, deteksi emosi berbasis teks, chatbot interaktif, dan penyimpanan data percakapan, bantuan langsung dengan tenaga ahli serta telah diuji secara fungsional menggunakan metode *black-box testing* dengan hasil seluruh fitur berjalan sesuai spesifikasi tanpa kesalahan logika. Evaluasi *usability* menggunakan *System Usability Scale* (SUS) menghasilkan skor rata-rata 83,66 yang termasuk kategori *Excellent*, sedangkan evaluasi pengalaman pengguna menggunakan UEQ menunjukkan bahwa aplikasi dinilai jelas, menarik, dan mampu mendorong interaksi. Secara keseluruhan, hasil penelitian menunjukkan bahwa integrasi *IndoBERT* dan LLM efektif dalam membangun *chatbot* kesehatan emosional yang empatik, mudah digunakan, dan sesuai dengan konteks bahasa serta budaya pengguna Indonesia, namun implementasi yang dilakukan masih terbatas pada perspektif ilmu komputer, sehingga penelitian selanjutnya perlu melibatkan kajian psikologi untuk menilai validitas intervensi emosional dan dampak psikologis aplikasi terhadap pengguna secara lebih komprehensif.

REFERENSI

Abuhassna, H., Alnawajha, S., Awae, F., Mohamed Adnan, M. A. Bin, & Edwards, B. I. (2024). Synthesizing technology integration within the Addie model for instructional design: A comprehensive systematic literature review. *Journal of Autonomous Intelligence*, *7*(5), 1546. https://doi.org/10.32629/jai.v7i5.1546

Adzan, M. N., Roxana, T., & Cahyana, C. (2025). Analisis Sentimen pada Chatbot untuk Mood Tracking. *Journal of Applied Engineering and Social Science (JAESS)*, *3*(1), 45–48. https://doi.org/10.25124/JAESS.V3I1.9527

Ayuningtyas, P. K., WP, D. A., & Rachmadi, P. (2023). Performance And Functional Testing With The Black Box Testing Method. *International Journal of Progressive Sciences and Technologies*. https://doi.org/10.52155/ijpsat.v39.2.5471

Esmi, N., Shahbahrami, A., Nabati, Y., Rezaei, B., Gaydadjiev, G., & de Jonge, P. (2025). Stress detection through

prompt engineering with a general-purpose LLM. *Acta Psychologica*, *260*, 105462. https://doi.org/10.1016/J.ACTPSY.2025.105462

Hambali, Y. A., Megasari, R., & Santoso, R. R. (2023). Implementasi Metode Machine Learning menggunakan Algoritma Evolving Artificial Neural Network pada Kasus Prediksi Diagnosis Diabetes. *Jurnal Aplikasi Dan Teori Ilmu Komputer*, *6*(1), 9–20. https://doi.org/10.17509/jatikom.v6i1.56536

Heldiansyah, M. F., & Winarko, E. (2022). Emotion Detection on Indonesian Tweets Using CNN and Contextualized Word Embedding. *2022 International Conference on Data and Software Engineering (ICoDSE)*, 53–58. https://doi.org/10.1109/icodse56892.2022.9972229

Høgsdal, H., Kaiser, S., & Kyrrestad, H. (2023). Adolescents' Assessment of Two Mental Health–Promoting Mobile Apps: Results of Two User Surveys. *JMIR Formative Research*, *7*. https://doi.org/10.2196/40773

Kaniawulan, I., Wibisono, Y., Wahyudin, A., Wibowo, L. A., Program, M. S., Kota, I., … Introduction, I. (2022). Systematic Literature Review : Digital Transformation Challenges and Strategies in the. *European Journal of Humanities and Educational Advancements*, *3*(03), 40–49.

KemenPPPA. (2024). siaran-pers kolaborasi-kemenpppa-dan-unicef-sinergikan-dukungan-kesehatan-mental-dan-psikososial-di-kementerian-dan-lembaga - kemenpppa-go. Retrieved February 5, 2025, from KemenPPPA website: https://kemenpppa.go.id/siaran-pers/kolaborasi-kemenpppa-dan-unicef-sinergikan-dukungan-kesehatan-mental-dan-psikososial-di-kementerian-dan-lembaga

Khairan, C. A., & Habib, M. S. (2025). Chatbot Ai Dalam Identifikasi Awal Gangguan Kesehatan Mental Di Indonesia: Tantangan Dan Prospek. *Jurnal EMPATI*, *13*(6), 498–508. https://doi.org/10.14710/EMPATI.2024.47903

Koto, F., Lau, J. H., & Baldwin, T. (2021). IndoBERTweet: A pretrained language model for Indonesian Twitter with effective domain-specific vocabulary initialization. *ArXiv Preprint ArXiv:2109.04607*.

Ministry of Health of the Republic of Indonesia. (2023). *Visualization of SKI Data 2023: Depression Among Young People in Indonesia*. Ministry of Health of the Republic of Indonesia.

Rusli, F. M., Rismala, R., & Nurrahmi, H. (2021). Emotion Classification on Indonesian Twitter Using Convolutional Neural Network (CNN). *2021 9th International Conference on Information and Communication Technology (ICoICT)*, 213–218. https://doi.org/10.1109/icoict52021.2021.9527447

Saputri, M. S., Mahendra, R., & Adriani, M. (2018). Emotion Classification on Indonesian Twitter Dataset. *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, 90–95. https://doi.org/10.1109/IALP.2018.8629262

Schrepp, M. (2023). User Experience Questionnaire Handbook v10(03.05.2023). *URL: Https://Www. Researchgate. Net/Publication/303880829{_}User{_}Experience{_}Questionnaire{_}Handbook{_}Version{_}2.(Accessed: 02.02. 2017)*, (September), 1–16. https://doi.org/10.13140/RG.2.1.2815.0245

Shah, J., DePietro, B., D'Adamo, L., Firebaugh, M., Laing, O., Fowler, L. A., … Fitzsimmons-Craft, E. E. (2022). Development and usability testing of a chatbot to promote mental health services use among individuals with eating disorders following screening. *International Journal of Eating Disorders*, *55*(9), 1229–1244. https://doi.org/10.1002/eat.23798

Shaw, C., LaCasse, P., & Champagne, L. (2025). Exploring emotion classification of indonesian tweets using large scale transfer learning via IndoBERT. *Soc. Netw. Anal. Min.*, *15*, 22. https://doi.org/10.1007/s13278-025-01439-6

Situmorang, G. F., & Purba, R. (2024). Deteksi Potensi Depresi dari Unggahan Media Sosial X Menggunakan IndoBERT. *Building of Informatics, Technology and Science (BITS)*, *6*(2), 649–661. https://doi.org/10.47065/BITS.V6I2.5496

Tresnayatna, B., Nurdin, E. A., & Wahyudin, A. (2024). Chatbot Terhadap Peran Gender Dalam Administrasi Pendidikan : Studi Bibliometrik VOSviewer. *Equalita: Jurnal Studi Gender Dan Anak*, *6*(1), 1. https://doi.org/10.24235/equalita.v6i1.19461

UNICEF. (2021). The State of the World's Children 2021 _ UNICEF. In *The state of world health*. United Nations. https://doi.org/10.18356/9789210010580

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., … others. (2020). IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding. *ArXiv Preprint ArXiv:2009.05387*.