Asisten Virtual Berbasis LLM untuk Layanan Administrasi Bidang Dikdas Dinas Pendidikan Minahasa

Rivaldo Jonathan Paath[1] , Kevin Leatemia[2], Ade Yusupa[3], Yaulie Deo Y. Rindengan[4]
[1]Universitas Sam Ratulangi, Manado, Indonesia
Email : rivaldopaath026@student.unsrat.ac.id, kevinleatemia026@student.unsrat.ac.id, ade@unsrat.ac.id, rindengan@unsrat.ac.id

Abstrak

Sektor pelayanan publik sering menghadapi kemacetan informasi akibat tingginya volume pertanyaan administratif yang berulang. Meskipun chatbot konvensional berbasis aturan telah digunakan, sistem tersebut cenderung kaku dalam merespons masukan yang dinamis. Generative AI menawarkan solusi yang fleksibel, namun penerapannya di sektor e-government memiliki tantangan besar terkait risiko halusinasi AI. Penelitian ini bertujuan mengembangkan dan mengevaluasi asisten virtual cerdas berbasis Large Language Model (LLM) untuk mengotomatisasi komunikasi layanan publik di Bidang Pendidikan Dasar (Dikdas) Dinas Pendidikan Kabupaten Minahasa, melalui pendekatan inovasi Quadruple Helix. Metodologi yang digunakan meliputi pengumpulan data untuk membangun basis pengetahuan tekstual dari Prosedur Operasional Standar (SOP), perancangan arsitektur tiga lapis yang mengintegrasikan Telegram Bot API dan Google Gemini API, serta implementasi prompt engineering dan guardrails yang ketat. Evaluasi fungsional dilakukan menggunakan metode Blackbox testing pada lima skenario interaksi nyata, pengukuran akurasi terhadap 15 pertanyaan uji, pengukuran waktu respons, dan komparasi analitis dengan arsitektur chatbot konvensional berbasis aturan. Hasil pengujian menunjukkan tingkat keberhasilan 100% pada kelima skenario uji Blackbox; sistem mampu memproses bahasa alami yang ambigu dan kesalahan ketik sesuai protokol yang diharapkan. Pengukuran akurasi menghasilkan tingkat akurasi 93,33% (14/15 pertanyaan tepat), dengan rata-rata waktu respons ~13,92 detik. Komparasi dengan chatbot rule-based menunjukkan peningkatan akurasi sebesar 20 poin persentase (73,33% menjadi 93,33%) serta kemampuan menangani variasi linguistik yang tidak terprogram. Implementasi guardrails terbukti efektif memitigasi risiko halusinasi dengan mengunci jawaban secara eksklusif pada basis pengetahuan dan menolak kueri di luar yurisdiksi instansi secara aktif. Meskipun demikian, evaluasi ini masih terbatas pada skala pengujian terkondisi dan memerlukan pengujian lebih lanjut yang melibatkan pengguna aktual dalam jumlah besar. Kesimpulannya, integrasi LLM pada platform perpesanan yang umum digunakan menghadirkan solusi E-government yang responsif, terkendali, dan hemat biaya, sekaligus mendongkrak inklusivitas birokrasi tanpa mewajibkan masyarakat mengunduh aplikasi baru.

Kata Kunci: Asisten Virtual, *E-government,* Kecerdasan Buatan Generatif, Rekayasa *Prompt, Quadruple Helix*

Abstract

Public service sectors frequently experience severe information bottlenecks due to high volumes of repetitive administrative inquiries. While conventional rule-based chatbots have been used, they often struggle with dynamic user inputs. Generative AI offers a flexible solution, but its deployment in e-government faces significant challenges, particularly the risk of AI hallucination. This study aims to develop and evaluate an intelligent virtual assistant utilizing Large Language Models (LLM) to automate public service communication at the Basic Education Division (Dikdas) of the Education Agency of Minahasa Regency, framed

within the Quadruple Helix innovation model. The methodology encompassed data collection to build a textual knowledge base of Standard Operating Procedures (SOPs), the design of a three-tier architecture integrating the Telegram Bot API and Google Gemini API, and the implementation of strict prompt engineering and guardrails. Functional evaluation was conducted using Blackbox testing across five real-world scenarios, accuracy measurement against 15 test questions, response time measurement, and analytical comparison with a conventional rule-based chatbot architecture. The results demonstrated a 100% success rate across the five Blackbox tested scenarios; the system effectively processed ambiguous natural language and typographical errors as per expected protocols. Accuracy measurement yielded 93.33% (14/15 correct answers), with an average response time of ~13.92 seconds. Comparison with the rule-based chatbot showed a 20 percentage point accuracy improvement (73.33% to 93.33%) and the ability to handle unprogrammed linguistic variations. The implemented guardrails successfully mitigated hallucination risks by confining responses exclusively to the knowledge base and politely declining out-of-jurisdiction queries. Nevertheless, this evaluation remains limited to a controlled testing environment and requires further large-scale testing with actual users. In conclusion, integrating LLMs with widespread messaging platforms provides a highly responsive, controlled, and cost-effective E-Government solution, significantly boosting bureaucratic inclusivity without requiring citizens to download new applications.

Keywords*: V*irtual Assistant, E-government, Generative AI, Prompt Engineering, Quadruple Helix

PENDAHULUAN

Menyediakan pelayanan publik yang prima merupakan kewajiban fundamental pemerintah, sebagaimana diamanatkan secara eksplisit dalam Undang-Undang Republik Indonesia Nomor 25 Tahun 2009 tentang Pelayanan Publik (Republik Indonesia, 2009). Guna memenuhi tuntutan masyarakat akan proses birokrasi dan administrasi yang cepat, transparan, dan efisien, instansi pemerintah di berbagai tingkatan tengah melakukan transformasi digital. Di Indonesia, upaya ini diakselerasi melalui penerapan Sistem Pemerintahan Berbasis Elektronik (SPBE) yang diatur dalam Peraturan Presiden Nomor 95 Tahun 2018 (Presiden Republik Indonesia, 2018). Namun, terlepas dari dorongan digitalisasi tersebut, masih banyak sektor administrasi publik, seperti Dinas Pendidikan Kabupaten Minahasa, yang menghadapi kendala *information bottleneck* atau kemacetan informasi. Ketidakefisienan ini terutama terjadi karena staf dan operator manusia sering kali kewalahan menghadapi tingginya volume pertanyaan berulang mengenai prosedur operasional standar (SOP), regulasi institusi, dan persyaratan administrasi rutin. Oleh karena itu, inovasi berkelanjutan yang didorong oleh kerangka *Quadruple Helix* yang mensinergikan pemerintah, akademisi, industri, dan masyarakat sipil sangat diperlukan untuk memodernisasi akses informasi publik.

Untuk mengatasi keterbatasan staf manusia dalam menangani pertanyaan yang masuk secara bersamaan, berbagai penelitian sebelumnya telah mengusulkan implementasi *chatbot* atau asisten virtual. Sebagai contoh, sistem *chatbot* yang menggunakan *Natural Language Processing* (NLP) tradisional telah berhasil diterapkan pada platform Telegram untuk mengotomatisasi respons dan mengelola alur informasi secara efektif (Furqan et al., 2023; Marshal & Soelistijadi, 2023). Demikian pula, asisten virtual berbasis situs web telah dirancang untuk memfasilitasi layanan pengambilan data di lembaga pemerintah daerah seperti Badan Pusat Statistik (Ningsih et al., 2026), serta digunakan secara luas untuk meningkatkan layanan pelanggan pada aplikasi

perpesanan dan sektor akademik (Ardiansyah & Sulaksono, 2023; Nurfiyah & Ramadhandi, 2023). Meskipun memberikan kemajuan, *chatbot* konvensional yang mengandalkan sistem pencocokan pola atau berbasis aturan (*rule-based*) sering kali terbukti kaku dan kesulitan merespons masukan pengguna yang sangat dinamis, santai, dan tidak terstruktur.

Di sisi lain, perkembangan pesat *Generative Artificial Intelligence* (AI) dan *Large Language Models* (LLM) telah mengubah paradigma interaksi manusia dan sistem karena kemampuannya memproses bahasa alami yang kompleks. Namun, implementasi teknologi LLM di Indonesia saat ini sebagian besar masih difokuskan pada sektor pendidikan, seperti membantu beban mengajar dan menyediakan sistem pembelajaran yang dipersonalisasi (Mambu et al., 2023; Mardeli et al., 2025). Penggunaan model AI terbuka seperti ChatGPT yang diintegrasikan dengan Telegram juga telah mulai dieksplorasi (Avisyah et al., 2023). Dalam lingkup global, adopsi AI generatif telah melangkah lebih jauh menjadi komponen fundamental dalam arsitektur *Smart Government*. Berbagai literatur internasional terbaru menekankan bahwa implementasi *AI Assistant* pada layanan publik tidak hanya memodernisasi administrasi birokrasi, tetapi juga memperkuat transparansi dan interaksi proaktif antara pemerintah dan warga negara (de Souza Bastos et al., 2025; Pathan & Shah, 2025). Integrasi LLM dengan dokumen pemerintahan yang kompleks terbukti mampu mendongkrak efisiensi instansi serta mewujudkan layanan informasi publik yang responsif dan interaktif (Khadija & Nurharjadmo, 2025; M. Erlangga Fauzi & Tata Sutabri, 2025; Macrì, 2026; Majithia et al., 2026). Meskipun demikian, penerapan LLM di ranah *E-Government* publik memunculkan tantangan besar, yakni risiko "halusinasi AI" sebuah anomali di mana sistem memberikan jawaban logis namun secara faktual salah (Hosseini & Tupasela, 2025). Oleh karenanya, metode *Prompt Engineering* sangat krusial diimplementasikan untuk menentukan persona sistem dan memasang batasan operasional (Amastini et al., 2023).

Berdasarkan kesenjangan tersebut, penelitian ini mengajukan rumusan masalah utama: "Bagaimana integrasi LLM melalui arsitektur Telegram Bot API dapat memfasilitasi asisten virtual yang tangguh dan aman dari risiko halusinasi informasi dalam mengotomatisasi pelayanan publik daerah?" Kebaruan dari studi ini terletak pada penerapan arsitektur *Prompt engineering* yang disesuaikan dengan hierarki yurisdiksi dan Prosedur Operasional Standar (SOP) lokal, mengubah model generatif serbaguna menjadi agen spesifik-domain yang secara aktif menolak kueri di luar yurisdiksi hukum instansi, sebuah pendekatan yang belum dieksplorasi secara memadai dalam literatur *E-government* di Indonesia.

Tujuan dari penelitian ini adalah untuk mengembangkan, mengimplementasikan, dan mengevaluasi asisten virtual cerdas berbasis *Generative AI* untuk layanan administrasi Dinas Pendidikan Kabupaten Minahasa. Melalui pendekatan inovasi *Quadruple Helix*, penelitian ini berupaya menjawab tantangan akses birokrasi digital dengan menawarkan cetak biru (*blueprint*) *E-government* yang skalabel, sangat adaptif, responsif 24/7, dan hemat biaya.

METODE

Penelitian ini menggunakan pendekatan pengembangan sistem yang terstruktur untuk merancang, mengimplementasikan, dan mengevaluasi asisten virtual cerdas yang secara khusus disesuaikan untuk mengoptimalkan layanan *E-government*. Untuk memastikan proses pengembangan yang komprehensif dan sistematis, metodologi dalam penelitian ini dibagi secara berurutan ke dalam empat fase utama. Fase-fase tersebut meliputi: pengumpulan dan pemrosesan

data kontekstual layanan publik, perancangan arsitektur sistem tiga lapis (three-tier architecture) yang mengintegrasikan AI generatif, implementasi *prompt engineering* bersamaan dengan batasan operasional (*guardrails*), dan terakhir adalah evaluasi fungsional sistem secara menyeluruh. Pendekatan sekuensial yang ketat ini diterapkan untuk menjamin bahwa solusi teknologi yang dihasilkan tidak hanya tangguh secara fungsional dalam merespons pengguna, tetapi juga secara ketat selaras dengan persyaratan administratif dan birokrasi yang kaku dalam operasi sektor publik.

A. Metode Pengumpulan Data

Metodologi dimulai dengan tahapan pengumpulan data sistematis yang bertujuan untuk membangun konteks operasional yang realistis bagi asisten virtual yang dikembangkan. Secara spesifik, penelitian ini menggunakan Dinas Pendidikan Kabupaten Minahasa sebagai sumber data primer. Pemilihan instansi ini didasarkan pada pertimbangan bahwa lembaga tersebut memberikan model yang kuat dan representatif dari lingkungan sektor publik tipikal yang sering mengalami "kemacetan informasi" (*information bottleneck*s) yang parah. Kemacetan ini utamanya disebabkan oleh tingginya volume pertanyaan administratif berulang yang harus ditangani secara manual oleh staf.

Proses pengumpulan data difokuskan pada pengumpulan dan digitalisasi berbagai dokumen administratif instansi yang pada awalnya tidak terstruktur. Material primer yang dikumpulkan mencakup Prosedur Operasional Standar (SOP) layanan, peraturan resmi kelembagaan, serta daftar Pertanyaan yang Sering Diajukan (Frequently Asked Questions / FAQs) yang secara rutin ditanyakan oleh masyarakat sipil maupun staf pengajar. Data yang dihimpun secara komprehensif ini mencakup domain layanan publik dengan permintaan tinggi, seperti persyaratan untuk penerbitan Nomor Unik Pendidik dan Tenaga Kependidikan (NUPTK), proses dan kelengkapan sertifikasi guru, hingga prosedur mutasi kepegawaian.

Setelah dokumen-dokumen mentah yang tersebar tersebut dikumpulkan, tahapan selanjutnya adalah melakukan fase pra-pemrosesan data yang ketat untuk memastikan kompatibilitasnya dengan mesin *Generative AI*. Informasi yang telah diekstraksi dibersihkan dari format-format yang tidak relevan, kemudian direstrukturisasi secara sistematis ke dalam sebuah knowledge base berbasis teks yang terpadu dan seragam. Langkah penataan struktur data ini sangat krusial, karena secara langsung mengoptimalkan context window pada *Large Language Model* (LLM) yang digunakan. Dengan format teks yang bersih dan terstruktur, kapasitas sistem meningkat secara signifikan dalam melakukan pencarian data secara efisien serta menghasilkan respons yang akurat dan sadar konteks.

B. Desain Arsitektur Sistem

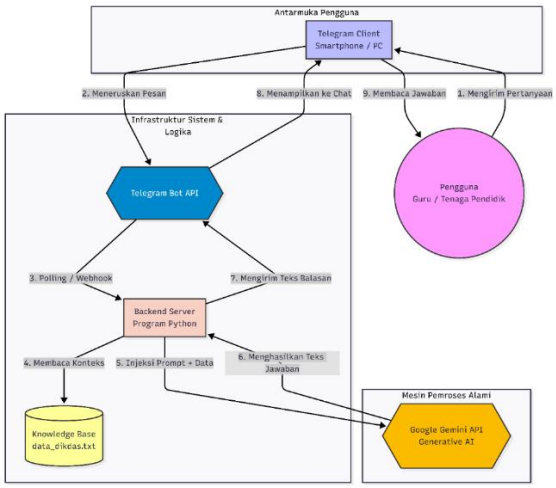
Arsitektur sistem dalam penelitian ini dirancang menggunakan model tiga lapis yang kohesif untuk memastikan jalur komunikasi yang efisien antara pengguna akhir dan mesin kognitif berbasis *Generative AI*. Penggunaan model ini memungkinkan isolasi tanggung jawab antara antarmuka pengguna, logika pemrosesan data, dan mesin kecerdasan buatan, sehingga mempermudah proses pemeliharaan dan skalabilitas sistem di masa depan.

Lapis pertama merupakan antarmuka pengguna yang dideploy pada platform pesan instan Telegram. Pemilihan Telegram dilakukan secara strategis karena platform ini memiliki Application Programming Interface (Bot API) yang sangat mumpuni, infrastruktur yang ringan, serta tingkat adopsi yang luas di kalangan masyarakat sipil. Keputusan arsitektural untuk menggunakan platform yang sudah populer ini secara signifikan

menghilangkan hambatan bagi pengguna dalam mengakses layanan *E-government*, karena mereka tidak perlu lagi mengunduh atau menginstal aplikasi khusus milik pemerintah yang sering kali membutuhkan ruang penyimpanan dan memori yang besar.

Lapis kedua berfungsi sebagai pusat logika operasional dan perutean data yang dikelola oleh server back-end yang dibangun khusus menggunakan bahasa pemrograman Python. Server ini bertindak sebagai middleware kritis yang menjembatani interaksi antara antarmuka Telegram dan mesin kognitif. Saat menerima muatan pesan teks dari Telegram Bot API, server Python menangkap dan memproses kueri tersebut secara asinkron. Di lapisan inilah basis pengetahuan tekstual (data_dikdas.txt) dan pedoman perilaku sistem disimpan. Sebelum berinteraksi dengan AI, server bertanggung jawab untuk menyusun input mentah pengguna dan menggabungkannya dengan instruksi sistem yang spesifik serta Standar Operasional Prosedur (SOP) instansi terkait.

Lapis ketiga adalah mesin pemroses bahasa alami yang didukung oleh Google Gemini API. Sebagai inti kognitif dari keseluruhan sistem, *Large Language Model* di dalam Gemini API menerima permintaan data yang telah diagregasi dari back-end Python. AI kemudian menganalisis maksud bahasa alami pengguna, melakukan referensi silang dengan basis pengetahuan institusional yang telah disuntikkan, dan merumuskan respons yang sangat kontekstual serta akurat. Seluruh alur data ini beroperasi dalam lingkaran dua arah yang berkesinambungan dan instan, memastikan birokrasi digital tetap responsif dan terjaga kualitas informasinya.



Gambar 1. Diagram Alir Arsitektur Sistem

C. Implementasi *Prompt engineering* dan *Guardrails*

Untuk memastikan AI Generatif berfungsi secara ketat sebagai instrumen layanan publik yang andal dan bukan sekadar agen percakapan umum, *prompt engineering* yang terperinci diimplementasikan di dalam logika backend. Rekayasa prompt berfungsi sebagai mekanisme kontrol fundamental yang menentukan bagaimana *Large Language Model* (LLM) menafsirkan masukan pengguna dan merumuskan jawabannya. Fase utama dari proses ini melibatkan pembentukan persona sistem yang definitif. Melalui instruksi sistem, AI secara eksplisit diarahkan untuk mengadopsi persona sebagai asisten administratif yang formal, empatik, dan sangat profesional yang mewakili Dinas Pendidikan Kabupaten Minahasa. Penanaman persona ini memastikan bahwa seluruh respons yang dihasilkan menggunakan terminologi birokrasi yang tepat dan mempertahankan nada yang berorientasi pada pelayanan, sebuah elemen mutlak untuk menjaga kredibilitas institusi dan kepercayaan publik.

Selain mendefinisikan persona, kerangka *prompt engineering* dimanfaatkan untuk menambatkan pemrosesan kognitif AI pada basis pengetahuan yang telah disusun sebelumnya. Teknik yang menyerupai

contextual injection diterapkan, di mana isi dari fail data_dikdas.txt disematkan secara sistematis ke dalam instruksi dasar model. LLM diberikan arahan yang ketat untuk menyintesis jawaban secara eksklusif hanya menggunakan data administratif yang telah disediakan tersebut. Dengan menginstruksikan model untuk memprioritaskan Standar Operasional Prosedur (SOP) dan regulasi institusi di atas pengetahuan global bawaannya, sistem secara efektif dapat memitigasi risiko "halusinasi AI" sebuah kerentanan kritis di mana model generatif merekayasa informasi yang terdengar masuk akal namun faktanya keliru.

Secara krusial, implementasi prompt juga mengintegrasikan batasan operasional yang kaku untuk menegaskan batas yurisdiksi asisten virtual. Mengingat bahwa Dinas Pendidikan di tingkat kabupaten (Dikdas) secara eksklusif hanya menaungi sektor pendidikan dasar seperti Sekolah Dasar (SD) dan Sekolah Menengah Pertama (SMP), AI diprogram dengan instruksi kendala negatif. Sistem diajarkan untuk secara aktif mengidentifikasi dan menolak dengan sopan berbagai pertanyaan pengguna yang berada di luar yurisdiksi hukumnya, seperti urusan Sekolah Menengah Atas (SMA) yang merupakan wewenang pemerintah provinsi, maupun topik di luar konteks pendidikan. Saat dihadapkan pada pertanyaan di luar cakupan, *guardrails* akan memicu mekanisme fallback atau prosedur cadangan yang telah ditentukan, sehingga AI akan meminta maaf atas keterbatasannya dan secara akurat mengarahkan pengguna ke otoritas provinsi atau instansi eksternal yang relevan. Lingkungan operasional yang sangat dibatasi ini menjamin bahwa inovasi *E-government* yang dihasilkan tetap akurat, bertanggung jawab, dan patuh pada prosedur administrasi yang berlaku.

D. Prosedur Pengujian dan Evaluasi

Fase akhir dari metodologi penelitian ini melibatkan evaluasi fungsional yang ketat terhadap asisten virtual menggunakan metode pengujian Blackbox. Pengujian Blackbox dipilih secara khusus karena metode ini sepenuhnya berfokus pada evaluasi perilaku masukan-keluaran dan pemenuhan persyaratan fungsional sistem tanpa perlu menginspeksi struktur kode sumber di dalamnya. Objektif utama dari evaluasi ini adalah untuk memvalidasi secara empiris kapasitas sistem dalam memahami masukan bahasa alami dan kemampuannya mempertahankan pengalaman pengguna yang mulus sesuai dengan protokol layanan E-government.

Evaluasi dirancang dan distrukturkan berdasarkan tiga pendekatan pengujian yang saling melengkapi: (1) evaluasi fungsional Blackbox Testing pada lima skenario interaksi, (2) pengukuran akurasi jawaban terhadap 15 pertanyaan uji, dan (3) pengukuran waktu respons sistem. Ketiga pendekatan ini bertujuan untuk mengukur sistem dari perspektif yang berbeda fungsionalitas, presisi informasi, dan performa teknis.

1. Evaluasi Fungsional Blackbox Testing
   Evaluasi fungsional dirancang berdasarkan lima skenario pengujian utama yang menyimulasikan berbagai interaksi pengguna dan perilaku sistem di dunia nyata. Kelima skenario tersebut meliputi:
   a) Validasi Inisiasi dan Salam: Skenario pertama mengevaluasi kemampuan inisiasi percakapan sistem, memastikan AI secara konsisten merespons sapaan dasar dengan sikap yang sopan dan profesional khas perwakilan layanan publik.
   b) Penanganan Kueri Kontekstual yang Valid: Skenario kedua menilai fungsi inti sistem, yakni pengambilan informasi. Pengujian ini memverifikasi bahwa AI mampu mengekstraksi instruksi persyaratan administrasi (misalnya: syarat sertifikasi) secara akurat dan eksklusif dari basis pengetahuan

tanpa merekayasa detail tambahan.

c) Penanganan Masukan Ambigu dan Kesalahan Ketik: Skenario ketiga dirancang untuk menguji ketangguhan *Natural Language Processing* (NLP) sistem. Dengan sengaja memberikan masukan menggunakan bahasa gaul/informal dan kesalahan pengetikan, pengujian ini bertujuan membuktikan kapasitas AI Generatif dalam menyimpulkan maksud inti pengguna meskipun sintaksis yang dimasukkan tidak sempurna.
d) Pengujian Batasan Kontekstual dan *Guardrails*: Skenario keempat mengevaluasi mekanisme pertahanan sistem. AI secara sengaja diberikan pertanyaan di luar yurisdiksinya untuk mengonfirmasi bahwa sistem akan menolak permintaan tersebut dengan sopan dan memberikan arahan yang akurat kepada pengguna.
e) Validasi Indikator Proses Asinkron: Skenario kelima menguji penanganan komunikasi asinkron. Karena pemrosesan LLM melalui Gemini API melibatkan sedikit jeda, pengujian ini memverifikasi bahwa bot Telegram berhasil menampilkan indikator "sedang mengetik...". Indikator ini memastikan pengguna mengetahui bahwa pertanyaan mereka sedang diproses, mencegah pengiriman pesan berulang, dan meningkatkan kenyamanan antarmuka.

2. Pengukuran Akurasi Jawaban
Selain evaluasi fungsional, dilakukan pengukuran akurasi untuk mengukur presisi faktual respons sistem terhadap basis pengetahuan institusional. Pengukuran ini menggunakan 15 pertanyaan uji yang merepresentasikan kondisi penggunaan nyata di lapangan, mencakup variasi: (a) pertanyaan administrasi standar dengan bahasa formal, (b) pertanyaan dengan bahasa informal/singkatan, (c) pertanyaan dengan kesalahan ketik, dan (d) pertanyaan di luar cakupan yurisdiksi.
Setiap respons dinilai berdasarkan kriteria berikut:
a) Tepat: Respons memberikan informasi yang sesuai secara faktual dengan SOP yang tertuang di knowledge base, lengkap, dan tidak merekayasa detail tambahan.
b) Kurang Presisi: Respons mengandung ketidaktepatan faktual, kelalaian informasi penting, atau penambahan detail yang tidak terdapat di knowledge base.

Penilaian dilakukan oleh peneliti dengan verifikasi silang langsung terhadap dokumen SOP instansi. Tingkat akurasi dihitung menggunakan rumus:

$$\text{Akurasi} = \left(\frac{Jumlah\ Jawaban\ Benar}{Total\ Pertanyaan}\right) \times 100\% \quad (1)$$

3. Pengukuran Waktu Respons
Pengukuran waktu respons dilakukan sebanyak 15 kali pengujian yang mencakup berbagai jenis kueri meliputi pertanyaan kontekstual, masukan dengan kesalahan ketik, dan pertanyaan di luar yurisdiksi. Waktu diukur sejak pesan dikirimkan pengguna hingga balasan muncul sepenuhnya di antarmuka Telegram. Pengukuran ini bertujuan untuk mengukur latensi teknis sistem dalam kondisi penggunaan nyata.

Perlu dicatat bahwa metode evaluasi dalam penelitian ini memiliki keterbatasan inheren yang perlu diakui secara transparan. Pengujian Blackbox yang dilakukan bersifat terkondisi dan sepenuhnya dikendalikan oleh peneliti, sehingga belum mencerminkan kompleksitas interaksi nyata dengan pengguna aktual dalam jumlah besar. Selain itu, evaluasi ini berfokus pada aspek fungsionalitas dan presisi teknis sistem tanpa mengukur dimensi pengalaman pengguna secara komprehensif, seperti kepuasan pengguna (user satisfaction),

kemudahan penggunaan (usability), atau persepsi kepercayaan terhadap informasi yang diberikan. Keterbatasan ini menjadi pertimbangan penting dalam menginterpretasikan hasil, mengingat keberhasilan teknis pada lingkungan pengujian terkontrol tidak serta-merta menjamin efektivitas sistem pada skenario deployment nyata dengan demografi pengguna yang lebih heterogen.

E. Komparasi Analitis dengan Arsitektur Chatbot Konvensional

Untuk mengukur nilai tambah teknis dari integrasi LLM, penelitian ini melakukan komparasi analitis antara asisten virtual berbasis LLM yang dikembangkan dengan arsitektur chatbot konvensional berbasis aturan (rule-based). Chatbot konvensional yang digunakan sebagai pembanding adalah model hipotetis berbasis pencocokan kata kunci (keyword matching) yang mengandalkan daftar pertanyaan-pertanyaan tetap dengan respons pradefinisi, sebuah pendekatan yang umum ditemui pada literatur chatbot layanan publik sebelumnya (Furqan et al., 2023; Marshal & Soelistijadi, 2023).

Komparasi dilakukan secara analitis menggunakan 15 pertanyaan uji yang identik dengan pengukuran akurasi sistem LLM. Setiap pertanyaan dievaluasi terhadap dua arsitektur: (1) chatbot rule-based, yang dianggap gagal jika variasi pertanyaan tidak cocok persis dengan kata kunci yang diprogram, dan (2) chatbot berbasis LLM, yang dinilai berdasarkan respons aktual dari sistem. Kriteria keberhasilan untuk chatbot rule-based adalah kecocokan eksak antara input pengguna dengan pola yang telah ditentukan, sementara untuk chatbot LLM adalah keakuratan dan relevansi respons terhadap basis pengetahuan. Hasil komparasi ini disajikan pada bagian Hasil dan dianalisis pada bagian Pembahasan.

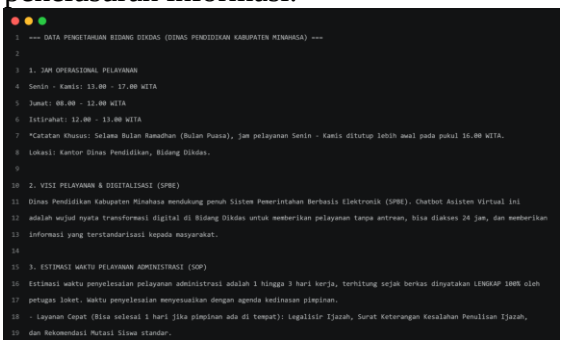
HASIL DAN PEMBAHASAN

A. Hasil Implementasi Sistem

Tahap implementasi dari penelitian ini menghasilkan sebuah asisten virtual cerdas berbasis *Generative AI* yang beroperasi penuh secara fungsional melalui platform perpesanan Telegram. Hasil pengembangan sistem ini dibagi ke dalam tiga komponen implementasi utama: pembuatan basis pengetahuan, konfigurasi *prompt engineering*, dan integrasi antarmuka pengguna.

1. Pembuatan Knowledge Base

Langkah implementasi pertama yang berhasil dilakukan adalah digitalisasi dan restrukturisasi dokumen administratif mentah dari Dinas Pendidikan Kabupaten Minahasa. Dokumen-dokumen yang mencakup Standar Operasional Prosedur (SOP) dan regulasi layanan pendidikan dasar (SD dan SMP) dikonversi dan dibersihkan dari format yang tidak relevan menjadi satu dokumen teks terpadu (data_dikdas.txt). Pendekatan ini berhasil mengubah data birokrasi yang sebelumnya tersebar menjadi format yang optimal untuk dibaca oleh *Large Language Model* (LLM), sehingga meminimalkan beban komputasi dan mengoptimalkan retensi memori (context window) AI saat melakukan penelusuran informasi.

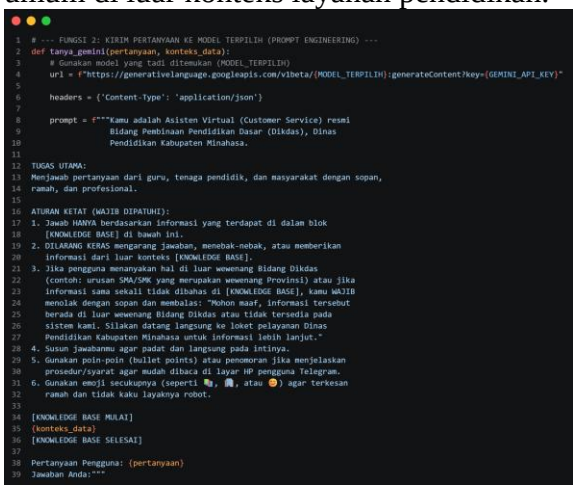


Gambar 2. Cuplikan Basis Pengetahuan Berbasis Teks

Gambar 2 menampilkan struktur knowledge base yang telah didigitalisasi. File data_dikdas.txt diorganisir ke dalam blok-blok tematik mencakup: (1) jam dan lokasi pelayanan, (2) visi digitalisasi SPBE, (3) estimasi waktu pelayanan per jenis SOP, dan (4) daftar persyaratan dokumen per layanan. Struktur teks yang bersih dan terurut ini memungkinkan LLM membaca konteks secara efisien tanpa overhead parsing.

2. Konfigurasi *Prompt engineering* dan Guardrail

Inti dari logika operasional sistem berhasil diimplementasikan pada server backend menggunakan bahasa pemrograman Python. Pada tahap ini, *prompt engineering* diterapkan secara hardcoded untuk mengunci persona AI sebagai Asisten Virtual resmi Bidang Pembinaan Pendidikan Dasar (Dikdas) Kabupaten Minahasa. Lebih lanjut, sistem berhasil mengimplementasikan *guardrails* yang ketat. Kode program secara eksplisit memerintahkan AI untuk (1) hanya menjawab berdasarkan file data_dikdas.txt, dan (2) menolak dengan sopan segala bentuk pertanyaan yang berada di luar wewenangnya, seperti urusan SMA/SMK (wewenang Provinsi) atau pertanyaan umum di luar konteks layanan pendidikan.



```
# --- FUNGSI 2: KIRIM PERTANYAAN KE MODEL TERPILIH (PROMPT ENGINEERING) ---
def tanya_gemini(pertanyaan, konteks_data):
    # Gunakan model yang tadi ditemukan (MODEL_TERPILIH)
    url = f"https://generativelanguage.googleapis.com/v1beta/{MODEL_TERPILIH}:generateContent?key={GEMINI_API_KEY}"

    headers = {'Content-Type': 'application/json'}

    prompt = f"""Kamu adalah Asisten Virtual (Customer Service) resmi
                Bidang Pembinaan Pendidikan Dasar (Dikdas), Dinas
                Pendidikan Kabupaten Minahasa.

TUGAS UTAMA:
Menjawab pertanyaan dari guru, tenaga pendidik, dan masyarakat dengan sopan,
ramah, dan profesional.

ATURAN KETAT (WAJIB DIPATUHI):
1. Jawab HANYA berdasarkan informasi yang terdapat di dalam blok
   [KNOWLEDGE BASE] di bawah ini.
2. DILARANG KERAS mengarang jawaban, menebak-nebak, atau memberikan
   informasi dari luar konteks [KNOWLEDGE BASE].
3. Jika pengguna menanyakan hal di luar wewenang Bidang Dikdas
   (contoh: urusan SMA/SMK yang merupakan wewenang Provinsi) atau jika
   informasi sama sekali tidak dibahas di [KNOWLEDGE BASE], kamu WAJIB
   menolak dengan sopan dan membalas: "Mohon maaf, informasi tersebut
   berada di luar wewenang Bidang Dikdas atau tidak tersedia pada
   sistem kami. Silakan datang langsung ke loket pelayanan Dinas
   Pendidikan Kabupaten Minahasa untuk informasi lebih lanjut."
4. Susun jawabanmu agar padat dan langsung pada intinya.
5. Gunakan poin-poin (bullet points) atau penomoran jika menjelaskan
   prosedur/syarat agar mudah dibaca di layar HP pengguna Telegram.
6. Gunakan emoji secukupnya (seperti 📄, 📝, atau 😊) agar terkesan
   ramah dan tidak kaku layaknya robot.

[KNOWLEDGE BASE MULAI]
{konteks_data}
[KNOWLEDGE BASE SELESAI]

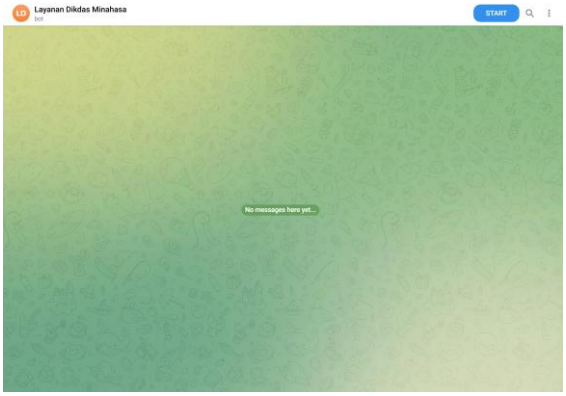
Pertanyaan Pengguna: {pertanyaan}
Jawaban Anda:"""
```

Gambar 3. Cuplikan Kode Implementasi System Instruction

Gambar 3 menampilkan implementasi system instruction pada fungsi tanya_gemini(). Terdapat tiga komponen utama: (1) baris 8–10 mendefinisikan persona AI sebagai asisten resmi Dikdas Kabupaten Minahasa, (2) baris 16–27 memuat aturan ketat *guardrails* yang melarang AI menjawab di luar knowledge base dan mewajibkan penolakan sopan untuk kueri di luar yurisdiksi, serta (3) baris 34–36 adalah mekanisme injeksi konteks di mana isi file data_dikdas.txt disematkan langsung ke dalam prompt sebelum dikirim ke Gemini API.

3. Integrasi Antarmuka Pengguna

Hasil akhir dari implementasi arsitektur adalah terhubungnya backend Python dan mesin kognitif Google Gemini API ke dalam antarmuka Telegram Bot API. Pengujian implementasi menunjukkan bahwa jalur komunikasi dua arah ini berjalan dengan sangat stabil. Masyarakat, guru, maupun staf sekolah kini dapat langsung berinteraksi dengan layanan *E-government* selama 24 jam penuh dalam seminggu (24/7) tanpa perlu mengunduh aplikasi pemerintahan baru. Antarmuka ini terbukti responsif dan mampu menampilkan balasan teks, bullet points untuk daftar syarat administrasi, serta emoji formal untuk membangun interaksi yang ramah namun tetap profesional. Keberhasilan integrasi antarmuka ini sekaligus membuktikan bahwa pendekatan tanpa aplikasi khusus (app-less approach) sangat efektif dalam menjembatani kesenjangan akses digital. Pengguna tidak lagi dibebani dengan proses instalasi, alokasi memori perangkat, atau registrasi akun yang rumit, melainkan cukup memanfaatkan ekosistem aplikasi perpesanan yang sudah familier bagi mereka. Pendekatan ini secara langsung mendongkrak tingkat aksesibilitas layanan *E-government* dan memastikan bahwa setiap pembaruan informasi dari backend dapat langsung tersalurkan kepada masyarakat secara real-time tanpa adanya hambatan teknis.



Gambar 4. Antarmuka Asisten Virtual Layanan Dikdas pada Platform Telegram.

B. Hasil Pengujian Sistem

Fase evaluasi fungsional dari asisten virtual ini dilakukan menggunakan

metode *Blackbox Testing*. Pengujian ini berfokus murni pada kesesuaian antara input pengguna dan respons output yang dihasilkan oleh sistem, tanpa memeriksa kode internal. Pengujian dilakukan secara langsung melalui antarmuka Telegram Bot dengan menggunakan lima skenario interaksi nyata yang mewakili dinamika komunikasi publik sehari-hari.

Hasil evaluasi dari kelima skenario tersebut menunjukkan tingkat keberhasilan 100% dan dirangkum secara komprehensif pada Tabel 1 di bawah ini.

Tabel 1. Hasil Evaluasi Fungsional Skenario *Blackbox Testing*

Input	Expected Output	Actual Output
/start atau /help	Sapaan formal, pengenalan persona, dan penawaran bantuan administratif.	Sistem langsung membalas dengan salam formal, memperkenalkan diri sebagai Asisten Virtual Dikdas Minahasa, dan menanyakan bantuan yang diperlukan.
"Siapa Kabid bidang dikdas" / "jadwal TKA kapan"	Jawaban berisikan daftar persyaratan akurat yang diekstrak murni dari knowledge base (SOP).	Sistem menjawab nama pimpinan dan merincikan jadwal menggunakan pemformatan teks (bullet points).
"bagaimana cra utk mngurus rehabilitas apa sja syaraynya"	AI memahami niat pengguna dan memberikan prosedur sertifikasi yang benar.	Sistem berhasil mentoleransi sintaksis informal dan typo, tetap memahami maksud (intent) inti, dan menampilkan SOP kelengkapan dokumen perbaikan ruang kelas secara terstruktur dan lengkap.
"kalau ingin mengurus ijazah SMA yang hilang bagaimana"	Penolakan secara sopan dan memberikan arahan ke instansi yang berwenang.	Sistem meminta maaf, menjelaskan bahwa SMA adalah wewenang Provinsi (bukan Dikdas), dan menyarankan pengguna ke instansi terkait.
[Pengguna mengirim pesan apa saja]	Indikator "sedang mengetik..." (typing...) muncul selama sistem memproses API.	Status "typing..." berhasil muncul di bawah nama bot selama beberapa detik saat menunggu Gemini API merumuskan jawaban, mencegah spam pesan.

Untuk memvalidasi rincian keberhasilan pada Tabel 1 di atas, berikut adalah dokumentasi cuplikan antarmuka yang mendemonstrasikan kinerja aktual sistem pada masing-masing skenario pengujian. Dokumentasi ini membuktikan bagaimana asisten virtual berhasil menangani berbagai variasi pertanyaan kontekstual sekaligus secara ketat mempertahankan batasan yurisdiksinya melalui *guardrails*.

1. Validasi Inisiasi dan Salam

Pengujian pertama ini secara khusus mengevaluasi kemampuan inisiasi percakapan sistem. Seperti yang ditunjukkan pada Gambar 5, hasil pengujian membuktikan bahwa AI secara konsisten mampu merespons sapaan dasar maupun perintah awal (seperti /start) dari pengguna. Sistem membalas dengan sikap yang sopan, empatik, dan profesional, yang

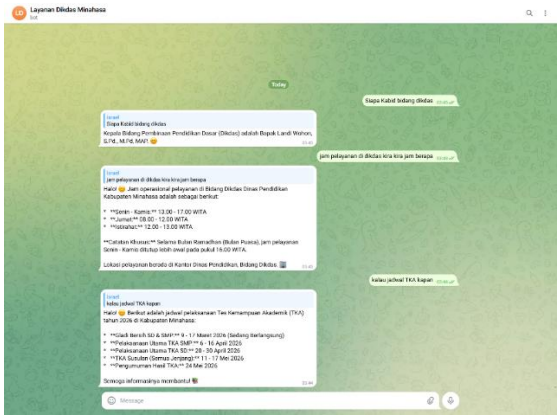
merupakan karakteristik fundamental dari representasi layanan publik yang prima.



Gambar 5. Pengujian Skenario 1 Inisiasi dan Salam

2. Penanganan Kueri Kontekstual yang Valid

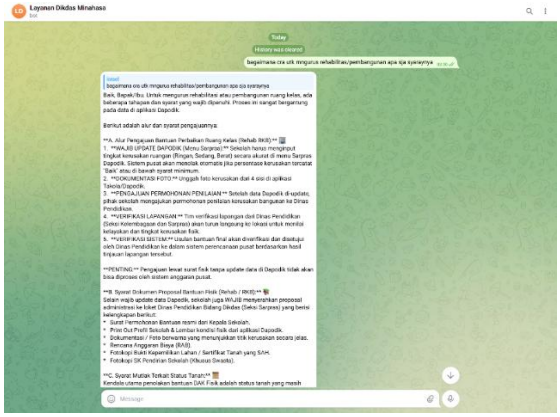
Skenario kedua menilai fungsi inti dari pengambilan informasi. Dengan memberikan pertanyaan administratif standar, pengujian ini memverifikasi bahwa AI mampu mengekstraksi informasi secara akurat dan eksklusif dari basis pengetahuan yang ditanamkan tanpa merekayasa detail tambahan.



Gambar 6. Pengujian Skenario 2 Penanganan Kueri Kontekstual

3. Penanganan Masukan Ambigu dan Kesalahan Ketik

Skenario ketiga menantang ketangguhan *Natural Language Processing* (NLP) dari sistem. Dengan menggunakan bahasa sehari-hari dan kesalahan pengetikan, pengujian ini membuktikan kapasitas AI Generatif dalam menyimpulkan niat inti pengguna secara akurat meskipun sintaksis yang digunakan tidak sempurna.



Gambar 7. Pengujian Skenario 3 Penanganan Masukan Ambigu dan Kesalahan Ketik

4. Pengujian Batasan Kontekstual dan *Guardrails*

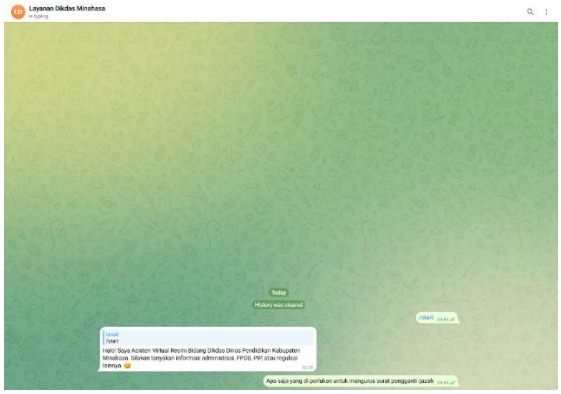
Skenario keempat mengevaluasi pengamanan prompt sistem. AI secara sengaja diberikan pertanyaan di luar yurisdiksinya untuk memastikan bahwa sistem akan menolak permintaan tersebut dengan sopan dan secara akurat mengarahkan pengguna ke otoritas yang berwenang.



Gambar 8. Pengujian Skenario 4 Pengujian Batasan Kontekstual dan *Guardrails*

5. Validasi Indikator Proses Asinkron

Skenario kelima menguji penanganan komunikasi asinkron oleh sistem. Karena pemrosesan LLM melalui Gemini API melibatkan sedikit latensi, pengujian ini memverifikasi bahwa Bot Telegram berhasil menampilkan indikator "sedang mengetik..." (typing...). Hal ini memastikan pengguna menyadari bahwa kueri mereka sedang diproses, mencegah pengiriman pesan yang berlebihan, dan meningkatkan pengalaman pengguna secara keseluruhan.



Gambar 9. Pengujian Skenario 5 Validasi Indikator Proses Asinkron

C. Hasil Pengukuran Kuantitatif Performa Sistem

Selain pengujian fungsional, dilakukan pula pengukuran kuantitatif terhadap waktu respons sistem, tingkat akurasi jawaban, dan komparasi dengan arsitektur chatbot konvensional untuk mengukur performa sistem secara objektif.

1. Pengukuran Waktu Respons

Pengukuran Waktu Respons dilakukan sebanyak 15 kali pengujian yang mencakup berbagai jenis kueri meliputi pertanyaan kontekstual, masukan dengan kesalahan ketik, dan pertanyaan di luar yurisdiksi. Waktu diukur sejak pesan dikirimkan pengguna hingga balasan muncul sepenuhnya di antarmuka Telegram.

Tabel 2. Hasil Pengukuran Waktu Respons Sistem

Pengujian Ke-	Waktu Respons
1	13,75 detik
2	17,00 detik
3	8,75 detik
4	15,25 detik
5	17,75 detik
6	9,50 detik
7	14,50 detik
8	16,25 detik
9	10,25 detik
10	16,25 detik
11	18,75 detik
12	9,00 detik
13	14,75 detik
14	17,25 detik
15	9,75 detik
Rata-Rata	~13,92 detik

Hasil pengukuran menunjukkan rata-rata waktu respons sistem sebesar ~13,92 detik. Latensi ini dipengaruhi oleh panjang konteks yang diinjeksikan ke dalam prompt serta proses inferensi pada Gemini API. Meskipun demikian, waktu respons ini masih dalam batas yang dapat diterima untuk layanan publik non-real-time dan didukung oleh indikator typing... yang menjaga kenyamanan pengguna selama proses berlangsung.

2. Pengukuran Waktu Respons

Pengukuran Akurasi dilakukan dengan menyiapkan 15 pertanyaan uji yang merepresentasikan kondisi penggunaan nyata di lapangan. Dari 15 pertanyaan tersebut, sebanyak 14 pertanyaan dijawab secara tepat dan relevan oleh sistem, sementara 1 pertanyaan menghasilkan respons yang kurang presisi. Tingkat akurasi dihitung menggunakan rumus berikut:

$$\text{Akurasi} = \left(\frac{Jumlah\ Jawaban\ Benar}{Total\ Pertanyaan}\right) \times 100\%$$
$$\text{Akurasi} = \left(\frac{14}{15}\right) \times 100\% = 93{,}33\% \quad (1)$$

Tingkat akurasi sebesar 93,33% mengindikasikan bahwa sistem asisten virtual ini mampu memberikan informasi yang tepat dan relevan pada sebagian besar skenario interaksi nyata, menegaskan kelayakannya sebagai instrumen layanan publik digital yang andal.

3. Pengukuran Waktu Respons

Untuk mengukur nilai tambah teknis dari integrasi LLM, penelitian ini melakukan komparasi sistematis antara asisten virtual berbasis LLM yang dikembangkan dengan arsitektur chatbot konvensional berbasis aturan (rule-based). Chatbot konvensional yang digunakan sebagai pembanding adalah model sederhana berbasis pencocokan kata kunci (keyword matching) yang mengandalkan daftar pertanyaan-pertanyaan tetap dengan respons pradefinisi, sebuah pendekatan yang umum ditemui pada literatur chatbot layanan publik sebelumnya (Furqan et al., 2023; Marshal & Soelistijadi, 2023). Komparasi dilakukan menggunakan 15 pertanyaan uji yang identik, mencakup variasi bahasa formal, informal, dengan kesalahan ketik, dan pertanyaan di luar cakupan.

Tabel 3. Komparasi Performa antara Chatbot Rule-Based dan Chatbot Berbasis LLM

Dimensi Evaluasi	Chatbot Rule-Based	Chatbot Berbasis LLM (Sistem Ini)
Akurasi Jawaban	73,33% (11/15 pertanyaan terjawab tepat)	93,33% (14/15 pertanyaan terjawab tepat)
Toleransi Bahasa Informal	Tidak mampu; gagal memahami singkatan dan gaya bahasa santai	Mampu menyimpulkan maksud dari bahasa informal dan singkatan
Toleransi Kesalahan Ketik	Tidak mampu; gagal total jika kata kunci tidak cocok persis	Mampu mengoreksi dan memahami intent meski terdapat typo
Penanganan Pertanyaan di Luar Cakupan	Menghasilkan respons generik atau tidak merespons	Menolak dengan sopan dan memberikan arahan ke instansi yang berwenang
Fleksibilitas Format Respons	Respons tetap (template)	Respons dinamis dan kontekstual dengan pemformatan adaptif
Waktu Respons Rata-Rata	< 1 detik (lokal)	~13,92 detik (tergantung API eksternal)
Kebutuhan Pemeliharaan Knowledge Base	Tinggi; setiap variasi pertanyaan harus diprogram manual	Rendah; penambahan SOP baru cukup dengan update file teks
Biaya Infrastruktur	Rendah (server lokal sederhana)	Moderat (bergantung pada penggunaan API eksternal)

4. Pembahasan

Berdasarkan hasil pengujian, asisten virtual berbasis *Generative AI* ini menunjukkan kinerja yang sangat baik dan menjadi lompatan teknologi yang signifikan dalam inovasi *E-government*. Berbeda dengan *chatbot* layanan publik dari penelitian sebelumnya (Furqan et al., 2023; Marshal & Soelistijadi, 2023) yang sebagian besar didominasi oleh sistem *Natural Language Processing* (NLP) berbasis aturan atau pencocokan kata kunci yang kaku, integrasi *Large Language Model* (LLM) melalui Gemini API pada sistem ini mampu mengakomodasi sifat komunikasi manusia yang dinamis dan tidak terstruktur. Pengujian secara spesifik membuktikan ketangguhan sistem dalam menoleransi bahasa informal, singkatan daerah, hingga kesalahan ketik tanpa kehilangan konteks maksud administratif inti pengguna. Kemampuan ini secara langsung menyelesaikan masalah "kemacetan informasi" dan secara efektif mengubah proses kueri manual yang lambat menjadi layanan digital yang otomatis dan instan selama 24 jam sehari.

Selain fleksibilitas linguistik, keunggulan utama dari implementasi ini adalah terjaminnya keamanan informasi melalui efektivitas rekayasa prompt dan batasan operasional. Penerapan batasan yang ketat terbukti berhasil memitigasi risiko "halusinasi AI" yang sering menjadi kendala pada model generatif, sehingga menjadikannya teknologi yang layak dan aman untuk penggunaan resmi pemerintah. Sistem secara disiplin mengunci responsnya hanya pada basis pengetahuan institusi dan terbukti mampu menolak permintaan di luar wewenangnya seperti urusan jenjang SMA secara halus, sekaligus memberikan arahan yang akurat kepada otoritas provinsi yang tepat. Pendekatan tertutup ini menjadikan sistem jauh lebih relevan untuk sektor tata kelola publik dibandingkan implementasi LLM terbuka (Avisyah, Putra, & Hidayat, 2023) yang rentan memberikan jawaban spekulatif di luar pedoman institusi. Lebih lanjut, inovasi ini mengukuhkan temuan dari penelitian sebelumnya (Macrì, 2026) mengenai pentingnya *chatbot* AI yang inklusif untuk layanan publik; penggunaan platform Telegram dalam sistem ini terbukti berhasil mendongkrak inklusivitas tanpa hambatan infrastruktur aplikasi yang berat.

Hasil komparasi pada Tabel 3 menunjukkan bahwa chatbot berbasis LLM

mengungguli pendekatan konvensional pada hampir seluruh dimensi kualitas interaksi, terutama dalam hal pemrosesan bahasa alami yang dinamis. Peningkatan akurasi sebesar 20 poin persentase (dari 73,33% menjadi 93,33%) secara kuantitatif membuktikan bahwa kemampuan inferensi LLM dalam menyimpulkan intent pengguna jauh melampaui keterbatasan pencocokan pola statis. Lebih fundamental lagi, chatbot konvensional sama sekali tidak mampu menangani variasi linguistik yang tidak terprogram, sebuah kelemahan fatal pada layanan publik di mana pengguna sering menggunakan bahasa sehari-hari. Namun demikian, komparasi ini juga mengungkapkan kompromi teknis yang perlu dipertimbangkan. Waktu respons chatbot LLM secara signifikan lebih lambat akibat latensi API eksternal, sebuah aspek yang dapat menjadi hambatan pada layanan yang memerlukan respons instan. Selain itu, ketergantungan pada infrastruktur API pihak ketiga memperkenalkan risiko kontinuitas layanan dan biaya operasional skala besar yang tidak dimiliki oleh sistem lokal. Oleh karena itu, pilihan arsitektur LLM bukanlah solusi absolut, melainkan keputusan trade-off yang menempatkan kualitas interaksi dan kemudahan pemeliharaan di atas kecepatan respons dan kemandirian infrastruktur.

Meskipun hasil pengujian menunjukkan kinerja yang memuaskan pada lingkungan terkontrol, implementasi sistem ini memiliki beberapa keterbatasan teknis yang perlu dipertimbangkan secara kritis sebelum deployment skala besar. Pertama, ketergantungan pada Gemini API memperkenalkan risiko kontinuitas layanan. Perubahan kebijakan API, penyesuaian harga, atau downtime layanan Google dapat mengganggu operasional layanan publik yang seharusnya tersedia 24/7. Berbeda dengan sistem lokal yang mandiri, ketergantungan pada infrastruktur pihak ketiga ini memerlukan strategi mitigasi berupa fallback system atau caching mechanism.

Kedua, latensi respons rata-rata ~13,92 detik meskipun masih dalam batas toleransi untuk layanan non-real-time, dapat menjadi hambatan pada skenario yang memerlukan respons instan atau pada jam sibuk dengan volume kueri tinggi. Latensi ini bersifat eksternal dan tidak sepenuhnya dapat dikendalikan oleh pengembang, sehingga perlu dievaluasi apakah masih acceptable untuk deployment produksi dengan ratusan pengguna simultan.

Ketiga, context window limitation menjadi pertimbangan penting seiring pertumbuhan knowledge base. Saat ini basis pengetahuan hanya mencakup SOP Dikdas Kabupaten Minahasa, namun jika di masa depan ingin diperluas ke seluruh dinas atau kabupaten lain, ada risiko melebihi batas token model. Hal ini dapat menurunkan akurasi atau memerlukan strategi chunking dan retrieval yang lebih kompleks.

Keempat, risiko keamanan data pada alur Telegram → server backend → Google API membuka potensi exposure data pengguna. Untuk implementasi produksi, diperlukan enkripsi end-to-end dan kepatuhan terhadap regulasi perlindungan data pribadi (misalnya Undang-Undang Perlindungan Data Pribadi di Indonesia). Selain itu, meskipun guardrails telah diimplementasikan, risiko adversarial prompt injection oleh pengguna yang sengaja mencoba "membobol" batasan sistem masih ada dan memerlukan mekanisme validasi input tambahan.

Kelima, sifat statis knowledge base saat ini menjadi keterbatasan fundamental. Sistem hanya mengakses dokumen teks yang di-update secara manual, belum terintegrasi dengan database real-time institusi. Artinya, jika terjadi perubahan regulasi atau SOP, sistem akan tetap memberikan informasi lama sampai knowledge base diperbarui secara manual. Hal ini berbeda dengan sistem enterprise yang terhubung langsung dengan sistem informasi manajemen kepegawaian atau pelayanan.

Keberhasilan dari solusi E-

government ini juga merupakan wujud nyata dari sinergi kerangka inovasi Quadruple Helix dalam memodernisasi birokrasi publik. Pemerintah, yang diwakili oleh Dinas Pendidikan Kabupaten Minahasa, bertindak sebagai pakar domain dan regulator dengan menyediakan data operasional administratif serta menetapkan batas yurisdiksi. Pihak Akademisi memainkan peran krusial dalam merancang arsitektur sistem, melakukan riset empiris, dan mengeksekusi rekayasa prompt yang kompleks agar LLM aman terintegrasi dalam layanan publik. Sektor Industri memberikan dukungan infrastruktur fundamental melalui API Google Gemini dan Telegram yang tangguh sekaligus hemat biaya, sementara Masyarakat Sipil termasuk guru dan publik secara umum menjadi penerima manfaat akhir. Pemanfaatan platform perpesanan yang sudah familier ini terbukti mampu menghapus kesenjangan digital, memungkinkan masyarakat mengakses layanan krusial tanpa beban mengunduh aplikasi pemerintahan baru yang rumit.

Meskipun sistem ini menawarkan solusi praktis dan andal, kerangka kerja ini masih memiliki keterbatasan fungsional. Saat ini, sistem asisten virtual masih bersifat statis karena hanya mengambil data dari dokumen teks dan belum terintegrasi dengan pencarian basis data waktu-nyata milik institusi. Oleh karena itu, penelitian selanjutnya direkomendasikan untuk memperluas kemampuan asisten virtual ini dengan mengintegrasikan fungsionalitas dinamis, seperti pelacakan status permohonan aplikasi secara otomatis, serta terus memperbarui basis pengetahuan agar tetap adaptif terhadap perubahan kebijakan institusi di masa mendatang.

SIMPULAN (PENUTUP)

Penelitian ini berhasil membuktikan hipotesis bahwa integrasi Large Language Model (LLM) tingkat lanjut pada platform perpesanan publik, yang diperkuat dengan guardrails operasional yang ketat, dapat secara efektif mengotomatisasi layanan publik tanpa memicu risiko halusinasi informasi. Pengembangan asisten virtual cerdas untuk Dinas Pendidikan Kabupaten Minahasa menggunakan antarmuka Telegram dan Google Gemini API terbukti berhasil mengatasi masalah kemacetan informasi administratif. Hasil pengujian menunjukkan bahwa sistem memiliki ketangguhan pemrosesan bahasa alami yang tinggi, mampu menoleransi kesalahan ketik serta masukan ambigu secara efektif. Evaluasi fungsional Blackbox Testing mencatat tingkat keberhasilan 100% pada kelima skenario uji, sementara pengukuran akurasi terhadap 15 pertanyaan lapangan menghasilkan tingkat akurasi 93,33%, menandakan bahwa sistem mampu memberikan informasi yang tepat pada sebagian besar skenario namun masih memiliki ruang untuk perbaikan presisi.

Lebih lanjut, keberhasilan kerangka prompt engineering dalam mengunci respons AI secara eksklusif pada basis pengetahuan (SOP) dan menolak kueri di luar yurisdiksi institusi menunjukkan bahwa teknologi generatif ini berpotensi menjadi instrumen yang terkendali dan dapat diandalkan untuk tata kelola birokrasi pemerintahan. Komparasi dengan arsitektur chatbot konvensional berbasis aturan menunjukkan peningkatan akurasi sebesar 20 poin persentase (73,33% menjadi 93,33%), sekaligus kemampuan menangani variasi linguistik yang tidak terprogram. Namun, potensi tersebut perlu diikuti dengan validasi lebih lanjut melalui pengujian skala besar yang melibatkan pengguna aktual, serta mitigasi terhadap risiko teknis seperti ketergantungan API eksternal, latensi respons, dan keamanan data, sebelum dapat direkomendasikan untuk implementasi penuh.

Meskipun inovasi kolaboratif berbasis kerangka Quadruple Helix ini menghadirkan cetak biru solusi E-government yang skalabel, inklusif, dan hemat biaya, sistem ini masih memiliki keterbatasan. Saat ini, asisten virtual yang dikembangkan masih bersifat statis dan belum terintegrasi dengan basis real-time database milik institusi. Oleh karena itu,

penelitian selanjutnya disarankan untuk memperluas kapabilitas sistem dengan mengintegrasikan pencarian data dinamis, seperti fitur pelacakan status berkas permohonan masyarakat secara otomatis. Selain itu, perancangan mekanisme pembaruan basis pengetahuan secara otomatis juga sangat direkomendasikan agar sistem dapat terus beradaptasi dengan perubahan regulasi dan kebijakan institusi di masa mendatang.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Dinas Pendidikan Kabupaten Minahasa yang telah memfasilitasi program magang serta menyediakan data institusional dan Standar Operasional Prosedur (SOP) yang menjadi landasan studi kasus pada penelitian ini.

DAFTAR PUSTAKA

Amastini, F., Aprijani, D. A., Prasetyo, D. A., & Summerlie, A. J. (2023). *Kajian Literatur Model Prompt Engineering ChatGPT untuk Pengajar di Institusi Pendidikan Jarak Jauh*.

Ardiansyah, R. H., & Sulaksono, A. G. (2023). Layanan pelanggan berbasis Natural Language Processing melalui chatbot pada aplikasi pesan. *Journal of Information System and Application Development*, *1*(1), 29–37. https://doi.org/10.26905/jisad.v1i1.9858

Avisyah, G. F., Putra, I. J., & Hidayat, S. S. (2023). Open Artificial Intelligence Analysis using ChatGPT Integrated with Telegram Bot. *Jurnal ELTIKOM*, *7*(1), 60–66. https://doi.org/10.31961/eltikom.v7i1.724

de Souza Bastos, A. L., Cirqueira dos Santos, L. F., Ribeiro dos Santos, S. R., Silva, M. V. S., Barbosa dos Santos, M. C., Santos, M. V., da Silva Santos, M. F., Mendonça, M. F., & Rocha, F. G. (2025). Large Language Models in Open Government Data Analysis: A Systematic Mapping Study. *International Conference on Web Information Systems and Technologies, WEBIST - Proceedings*, 404–411. https://doi.org/10.5220/0013777300003985

Furqan, M., Sriani, & Shidqi, M. N. (2023). Chatbot Telegram Menggunakan Natural Language Processing. *Walisongo Journal of Information Technology (WJIT)*, *5*(1), 15–26. https://doi.org/10.21580/wjit.2023.5.1.14793

Hosseini, Z., & Tupasela, A. (2025). Inclusive AI-Based Chatbots for Public Services in Finland: Potential of Constructivism. *The Barcelona Conference on Arts, Media & Culture 2024: Official Conference Proceedings*, 159–175. https://doi.org/10.22492/issn.2435-9475.2024.16

Khadija, M. A., & Nurharjadmo, W. (2025). PDF-Document Chatbot Responses using Large Language Models to Enable Smart City Engagement. *Jurnal Informatika: Jurnal Pengembangan IT*, *10*(3), 574–584. https://doi.org/10.30591/jpit.v10i3.8262

M. Erlangga Fauzi, & Tata Sutabri. (2025). PublicTalk: Sistem Chatbot Pintar Berbasis Natural Language Processing untuk Layanan Pemerintahan Digital. *JOURNAL SAINS STUDENT RESEARCH*, *3*(2), 426–433. https://doi.org/10.61722/jssr.v3i2.4325

Macrì, G. (2026). *LLM-Based Tools for Workers in the Public Sector: Supporting Public Knowledge Work Support through Fine-Tuning, RAG, and Prompt Chaining* [Sapienza University of Rome]. https://iris.uniroma1.it/retrieve/8d4f170f-79b7-42d8-87fe-b4e02f60665c/Tesi_dottorato_Macr%C3%AC.pdf

Majithia, N., Shinde, R., Chapman, Z., Trital, P., Decker, J., Maskey, M.,

Simperl, E., & Shadbolt, N. (2026). *The CitizenQuery Benchmark: A Novel Dataset and Evaluation Pipeline for Measuring LLM Performance in Citizen Query Tasks*. http://arxiv.org/abs/2602.04064

Mambu, J. G. Z., Pitra, D. H., Ilmi, A. R. M., Nugroho, W., Leuwol, N. V., & Saputra, A. M. A. (2023). Pemanfaatan Teknologi Artificial Intelligence (AI) Dalam Menghadapi Tantangan Mengajar Guru di Era Digital. *Journal on Education*, *06*(01), 2689–2698.

Mardeli, Sari, J. P., Setyaningsih, K., Fauzi, M., Fitri, E., Ballianie, N., & Novianti, R. (2025). Generative AI in Personalized Learning: A Systematic Review of Implementation in Indonesia. *EDUKASIA Jurnal Pendidikan Dan Pembelajaran*, *6*(2), 989–1002. https://doi.org/10.62775/edukasia.v6i2.1672

Marshal, G. W., & Soelistijadi, R. (2023). ChatBot SPBU 44.501.01 Menggunakan Telegram Bot API Dengan Metode Long-Polling. *Jurnal Ilmiah Teknologi Informasi Asia*, *17*(2).

Ningsih, G. A. D. S. A., Pramartha, C. R. A., & Astawa, I. G. S. (2026). Rancang Bangun Chatbot sebagai Virtual Assistant dalam Pelayanan Pencarian Data di Badan Pusat Statistik (BPS) Kabupaten Badung berbasis Website. *Jurnal Pengabdian Informatika*, *4*(2).

Nurfiyah, & Ramadhandi, N. (2023). Penerapan Metode Natural Language Processing (NLP) Pada Question Answering System Untuk Media Informasi Mahasiswa Universitas. *Journal of Information and Information Security*.

Pathan, M. K., & Shah, A. (2025). Ethical Considerations and Responsible Governance of Generative AI: A Systematic Review. *Premier Journal of Artificial Intelligence*. https://doi.org/10.70389/pjai.100016

Presiden Republik Indonesia. (2018). *Peraturan Presiden Nomor 95 Tahun 2018 tentang Sistem Pemerintahan Berbasis Elektronik*.

Republik Indonesia. (2009). *Undang-Undang Republik Indonesia Nomor 25 Tahun 2009 Tentang Pelayanan Publik*.