

Metode Retrieval-Augmented Generation dalam Pengembangan Chatbot Large Language Model pada Service Informasi Publik Disdukcapil Anambas

¹Muhammad Habsyi Mubarak, ²Joko Sutopo

^{1,2}Universitas Teknologi Yogyakarta, Indonesia

¹habsyimubarak@gmail.com; ²jksutopo@uty.ac.id

Article Info

Article history:

Received, 2025-12-11

Revised, 2025-12-24

Accepted, 2025-12-30

Kata Kunci:

Chatbot
Large Language Model
Retrieval-Augmented Generation
Administrasi Kependudukan
Kepulauan Anambas

Keywords:

Chatbot
Large Language Model
Retrieval-Augmented Generation
Population Administration
Anambas Islands

ABSTRAK

Layanan administrasi kependudukan di Kabupaten Kepulauan Anambas menghadapi tantangan besar terkait keterbatasan akses informasi yang disebabkan oleh kondisi *geografis*. Informasi layanan pada *website* resmi Disdukcapil yang bersifat pasif seringkali menyulitkan masyarakat untuk memperoleh jawaban yang cepat dan relevan. Kondisi ini berakibat pada penumpukan antrean layanan dan berpotensi menurunkan kepuasan publik. Sebagai solusi, sebuah aplikasi *chatbot* cerdas berbasis *Large Language Model* (LLM) dengan pendekatan *Retrieval-Augmented Generation* (RAG) telah dikembangkan. Metode ini secara efektif menggabungkan kemampuan pencarian informasi presisi dari basis data dokumen dengan kemampuan LLM dalam memahami bahasa alami untuk menghasilkan jawaban kontekstual. Proses pengembangan sistem ini dilakukan menggunakan *framework* *LangChain*, *vector store* *ChromaDB*, dan diintegrasikan ke antarmuka web sebagai *frontend*. Informasi layanan resmi Disdukcapil diproses melalui tahapan *chunking*, *embedding*, hingga pembuatan *pipeline* RAG. Hasil penelitian menunjukkan bahwa *chatbot* mampu merespons pertanyaan layanan kependudukan secara akurat dan efisien. Berdasarkan evaluasi menggunakan metrik *BERTScore*, sistem memperoleh rata-rata skor *Precision* sebesar 98,5%, *Recall* sebesar 99,1%, dan *F1-Score* sebesar 98,8%. Sistem ini dapat diakses dari mana saja, sangat membantu masyarakat di wilayah terpencil, dan menjadi prototipe awal yang potensial untuk mendukung digitalisasi pelayanan publik di daerah kepulauan.

ABSTRACT

Population administration services in the Anambas Islands Regency face significant challenges related to limited information access caused by geographical conditions. Service information on the official Disdukcapil website, which is passive, often makes it difficult for the public to obtain fast and relevant answers. This condition leads to service queue buildups and potentially decreases public satisfaction. As a solution, an intelligent chatbot application based on a Large Language Model (LLM) with a Retrieval-Augmented Generation (RAG) approach was developed. This method effectively combines precise information retrieval capabilities from a document database with the LLM's natural language understanding ability to produce contextual answers. The system's development process was carried out using the LangChain framework, Chroma vector store, and was integrated into a web interface as the frontend. Official Disdukcapil service information was processed through chunking, embedding, and RAG pipeline creation stages. The results showed that the chatbot was able to respond to inquiries about population services accurately and efficiently. Based on evaluations using the BERTScore metric, the system obtained average scores of 98,5% for Precision, 99,1% for Recall, and 98,8% for F1-Score. This system can be accessed from anywhere, greatly assisting people in remote areas, and serves as a potential initial prototype to support the digitization of public services in archipelago regions.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Muhammad Habsyi Mubarak,
Program Studi Informatika,
Universitas Teknologi Yogyakarta,
Email: habsyimumbarak@gmail.com

1. PENDAHULUAN

Administrasi kependudukan merupakan elemen fundamental dalam tata kelola pemerintahan yang berhubungan langsung dengan hak sipil masyarakat. Layanan ini berperan krusial dalam menjamin validitas data penduduk serta penyediaan dokumen legal seperti e-KTP, Akta Kelahiran, dan Kartu Keluarga yang menjadi basis layanan sektor pendidikan, kesehatan, hingga perbankan [1], [2]. Pelayanan publik yang optimal tidak hanya menuntut kecepatan, tetapi juga transparansi informasi. Oleh karena itu, ketersediaan informasi yang jelas dan terstruktur merupakan kunci untuk memenuhi partisipasi masyarakat dan menghindari kesan layanan yang lamban [3]. Namun, transformasi digital ini menghadapi tantangan besar di wilayah dengan karakteristik geografis khusus. Kabupaten Kepulauan Anambas, sebagai wilayah kepulauan terluar, memiliki hambatan aksesibilitas fisik yang signifikan. Masyarakat yang berdomisili di pulau-pulau terpisah seringkali harus menempuh perjalanan laut dengan biaya tinggi dan risiko cuaca ekstrem hanya untuk mendapatkan informasi dasar atau memverifikasi persyaratan dokumen di kantor Dinas Kependudukan dan Pencatatan Sipil (Disdukcapil).

Saat ini, kanal informasi digital yang tersedia melalui situs web resmi Disdukcapil masih bersifat satu arah. Masyarakat diharuskan menelusuri menu yang kompleks secara mandiri untuk menemukan informasi spesifik. Kendala ini sering menyebabkan ketidakpahaman masyarakat terhadap persyaratan administrasi, yang berujung pada kasus berkas tidak lengkap saat tiba di kantor dinas. Akibatnya, terjadi inefisiensi berupa antrian yang menumpuk dan keharusan masyarakat untuk melakukan perjalanan ulang antar-pulau. Kondisi ini menuntut adanya solusi teknologi yang mampu memberikan layanan informasi yang interaktif, presisi, dan dapat diakses 24 jam tanpa batasan geografis [4]. Pemanfaatan teknologi kecerdasan buatan atau Artificial Intelligence (AI), khususnya chatbot, menjadi solusi potensial untuk menjembatani kesenjangan akses tersebut.

Secara teknis, penggunaan chatbot konvensional seringkali memiliki keterbatasan dalam memahami konteks spesifik atau memberikan informasi terkini. Oleh karena itu, pendekatan *Retrieval-Augmented Generation* (RAG) pada *Large Language Model* (LLM) menjadi terobosan penting. Berbeda dengan metode *fine-tuning* yang membutuhkan pelatihan ulang model saat data berubah, RAG memungkinkan sistem untuk mengambil data dari basis pengetahuan eksternal secara real-time sebelum menghasilkan jawaban. Hal ini sangat relevan untuk layanan Disdukcapil yang memiliki dua jenis data: data statis (persyaratan, UU) dan data dinamis seperti status penyelesaian dokumen.

Sejumlah penelitian telah mengeksplorasi penerapan LLM dan RAG dalam layanan publik. Anam dkk. [5] mengembangkan asisten hukum virtual berbasis model Mistral 7B dengan metode RAG untuk Undang-Undang Lalu Lintas, yang menghasilkan F1-Score sebesar 0,9151. Penelitian ini membuktikan bahwa RAG efektif untuk mengambil informasi dari dokumen teks panjang. Selanjutnya, Muhajir dkk. [6] mengimplementasikan LLM GPT-3.5 Turbo menggunakan *framework LangChain* pada sistem layanan akademik kampus dengan akurasi mencapai 96,66%. Hidayat dkk. [7] juga merancang chatbot serupa berbasis Telegram di Universitas Mataram yang mencapai akurasi sempurna pada pengujian terbatas. Studi lain oleh Mar'arif dkk. [8] dan Permadi dkk. [9] masing-masing berfokus pada domain keagamaan yaitu fiqih dan tafsir dengan tingkat akurasi di atas 80%.

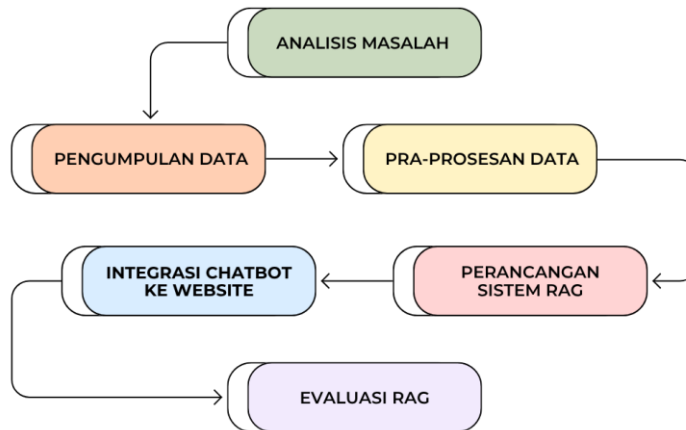
Perkembangan teknologi *Natural Language Processing* (NLP) dan *Large Language Model* (LLM) telah mendorong berbagai inovasi dalam sistem tanya-jawab cerdas dan *chatbot* [10]. Dalam skala global, peran *chatbot* dalam sektor publik juga mendapat perhatian luas. *Conversational agent* berbasis NLP memiliki potensi besar dalam meningkatkan interaksi manusia dan mesin [11]. Bakkouri dkk. [12] menekankan bahwa responsivitas chatbot berkorelasi langsung dengan peningkatan kepuasan pengguna. Lebih lanjut, El Gharbaoui dkk. [13] serta Anitha dkk. [14] menyoroti peran krusial AI dalam memperkuat kepercayaan masyarakat terhadap layanan publik digital melalui transparansi informasi. Dalam aspek edukasi pengguna, Chen dkk. [15] mengemukakan bahwa desain chatbot yang adaptif mampu mempersonalisasi pengalaman pengguna. Di sisi lain, Guo dkk. [16] membuka wawasan mengenai kemajuan *Large Language Model* berbasis multi-agent untuk menangani tugas pelayanan yang kompleks, sementara Sheikh dkk. [17] mengingatkan pentingnya penerapan etika AI dalam transformasi digital pemerintahan.

Meskipun berbagai penelitian di atas telah berhasil menerapkan LLM dan RAG, penelitian ini menawarkan pendekatan berbeda yang difokuskan pada optimalisasi tujuan pelayanan publik. Mayoritas penelitian

sebelumnya, seperti pada [3], [6], dan [7], berkonsentrasi pada akurasi pengambilan informasi dari dokumen statis melalui mekanisme tanya-jawab tunggal. Namun, tantangan utama dalam layanan birokrasi digital tidak hanya sekadar mengambil data, melainkan bagaimana sistem dapat mengenali kebutuhan spesifik pengguna secara cerdas. Oleh karena itu, research gap dalam penelitian ini berfokus pada integrasi *Intent Classification* untuk mengenali tujuan pengguna serta pemahaman konteks percakapan sebelumnya agar mesin dapat merespon secara lebih akurat.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang dirancang secara sistematis agar dapat menghasilkan sistem chatbot layanan publik berbasis Retrieval-Augmented Generation yang sesuai dengan kebutuhan Disdukcapil Kabupaten Kepulauan Anambas. Secara umum, alur penelitian ditunjukkan pada gambar 1.



Gambar 1 Alur Penelitian

Alur penelitian dimulai dari analisis masalah, yaitu mengidentifikasi kendala yang dihadapi masyarakat dalam memperoleh informasi kependudukan melalui layanan Disdukcapil. Dari hasil analisis, ditemukan bahwa sistem yang ada masih bersifat pasif dan belum menyediakan interaksi dua arah secara otomatis. Oleh karena itu, diperlukan sistem chatbot yang mampu menjawab pertanyaan publik untuk mengurangi beban kerja petugas Disdukcapil dengan tujuan meningkatkan pelayanan informasi.

Tahap pengumpulan data dilakukan dengan menghimpun sumber informasi statis dan dinamis. Data statis mencakup informasi prosedur layanan yang diekstraksi secara manual dari situs resmi Disdukcapil Kabupaten Kepulauan Anambas (<https://disdukcapil.anambaskab.go.id/>) serta dokumen regulasi Peraturan Menteri Dalam Negeri dari situs <https://www.kemendagri.go.id>. Selain itu, disusun pula kumpulan *Frequently Asked Questions* (FAQ) yang dirangkum dari situs Disdukcapil Anambas dan referensi silang dari beberapa situs Disdukcapil di wilayah Kepulauan Riau. Sementara itu, untuk kebutuhan data dinamis pelacakan status dokumen, dikembangkan basis data simulasi yang diakses melalui API (*Application Programming Interface*) khusus untuk keperluan pengujian sistem. Secara lengkap jenis-jenis data beserta deskripsi dapat dilihat pada tabel 1.

Tabel 1. Jenis-jenis Data

Kategori Data	Sumber Data	Metode	Deskripsi
Data Statis	https://disdukcapil.anambaskab.go.id/	Ekstraksi Manual	Informasi persyaratan, alur, dan prosedur layanan administrasi kependudukan (KTP-el, KK, Akta).
Data Statis	https://kemendagri.go.id/	Unduhan Dokumen	Dokumen Peraturan Menteri Dalam Negeri (Permendagri) yang menjadi landasan hukum layanan.
Data Statis	Website Disdukcapil Anambas & Portal Disdukcapil Regional Kepri	Ekstraksi Manual	Kumpulan <i>Frequently Asked Questions</i> (FAQ) yang disusun dari berbagai pertanyaan umum masyarakat di wilayah kepulauan.
Data Dinamis (Simulasi)	Basis Data Lokal	Generate via Custom API	Data simulasi (<i>mock data</i>) status pengajuan dokumen yang dibuat khusus untuk pengujian respon <i>real-time</i> sistem.

Tahap pra-prosesan data dilakukan untuk memastikan data dalam kondisi bersih dan konsisten sebelum dimasukkan ke sistem. Langkah-langkah yang dilakukan meliputi pembersihan teks, normalisasi, penghapusan karakter khusus, serta penyusunan ulang konten agar mudah diolah dan dikenali oleh model *embedding*.

Khusus untuk data statis, diterapkan strategi splitting yang disesuaikan dengan jenis dokumen untuk memaksimalkan relevansi retrieval. Dokumen regulasi diproses menggunakan pemisahan berbasis Pasal didalam dokumen peraturan presiden dan undang-undang, agar setiap unit hukum tetap utuh sebagai satu *chunk*. Sementara itu, dokumen prosedur layanan dipisahkan secara manual menggunakan penanda karakter spesial “####”. Setelah pemisahan struktural, data dipecah lebih lanjut menjadi *chunks* menggunakan *RecursiveCharacterTextSplitter* dengan *chunk overlap* sebesar 400 untuk mempertahankan konteks semantik antar *chunks* sebelum diubah menjadi *embedding*.

Pada tahap perancangan sistem RAG, integrasi antara komponen pencarian konteks dan model generatif diwujudkan menggunakan framework *LangChain* sebagai fondasi utama. Pada sisi pemrosesan dokumen, sistem memanfaatkan model *embedding* lokal *bge-m3:latest* untuk mengubah korpus informasi menjadi representasi vektor yang lebih mudah ditelusuri. Hasil *embedding* tersebut kemudian disimpan dan diorganisasi dalam *ChromaDB*, yang berfungsi sebagai basis data vektor guna mendukung proses pengambilan informasi secara cepat dan presisi. Arsitektur ini dirancang untuk memastikan bahwa setiap permintaan pengguna diproses melalui mekanisme retrieval yang mampu mengidentifikasi dokumen paling relevan dari koleksi data yang tersedia, sehingga konteks yang diberikan kepada model generatif benar-benar akurat dan sesuai kebutuhan [18].

Setelah konteks berhasil diperoleh, proses perumusan jawaban dilakukan oleh model *Gemini Flash 2.5* yang diakses melalui API, sehingga sistem dapat menghasilkan respons yang lebih informatif dan terarah berdasarkan data yang ditemukan. Infrastruktur backend dibangun dengan *FastAPI*, yang berperan sebagai server untuk mengelola permintaan, menjalankan pipeline RAG, serta menjembatani komunikasi antara berbagai komponen sistem. Sementara itu, antarmuka pengguna dikembangkan menggunakan *React.js* untuk memberikan pengalaman interaksi yang modern dan responsif. Seluruh komponen ini dijalankan secara lokal agar pengelolaan data, performa, dan keamanan sistem dapat dikontrol secara optimal selama proses pengembangan maupun pengujian. Untuk daftar teknologi yang digunakan bisa dilihat pada tabel 2.

Tabel 2. Daftar Teknologi yang Digunakan

Komponen Sistem	Teknologi	Fungsi Utama
Framework RAG	<i>LangChain</i>	Mengatur alur kerja retrieval dan generation
Database Vector	<i>ChromaDB</i>	Penyimpanan dan pencarian vektor secara efisien
Model <i>Embedding</i>	<i>Bge-m3:latest</i> (lokal)	Mengubah dokumen menjadi vektor semantik
Model Generatif	<i>Gemini flash 2.5 API</i>	Menghasilkan jawaban berdasarkan konteks
Backend/Server	<i>FastAPI</i>	Menangani <i>request</i> dan menjalankan <i>pipeline</i> RAG
Frontend	<i>React.js</i>	Antarmuka web interaktif untuk pengguna

Tahap integrasi chatbot ke website Disdukcapil bertujuan untuk menempatkan sistem RAG ke dalam antarmuka web Disdukcapil. Chatbot dikembangkan sebagai fitur tambahan di situs resmi dengan koneksi ke *backend API local* menggunakan *Fastapi*. Sistem ini dioperasikan dalam *localhost* untuk proses pengujian dan validasi.

Tahap terakhir adalah evaluasi RAG, yang dilakukan dengan mengukur kualitas dan kesesuaian jawaban chatbot terhadap pertanyaan publik menggunakan metrik *BERTScore*. Proses evaluasi RAG dalam penelitian ini mengacu pada pendekatan metrik [19], yang menilai kualitas retrieval berdasarkan kesesuaian semantik dan relevansi konteks. Evaluasi ini menilai kemampuan sistem dalam memberikan jawaban yang relevan, faktual, dan sesuai konteks, serta menjadi dasar penilaian keberhasilan penerapan metode RAG dalam meningkatkan layanan informasi publik Disdukcapil.

Evaluasi kinerja sistem *Retrieval-Augmented Generation* (RAG) pada penelitian ini dilakukan menggunakan seperangkat pertanyaan uji yang merepresentasikan kebutuhan informasi utama masyarakat terhadap layanan administrasi kependudukan. Sebanyak 10 pertanyaan uji disusun secara purposive berdasarkan kategori layanan yang paling sering diakses, meliputi informasi visi dan tugas Disdukcapil, persyaratan dokumen kependudukan, prosedur layanan, serta satu pertanyaan di luar cakupan basis pengetahuan untuk menguji kemampuan sistem dalam menjaga batas konteks (*out-of-scope handling*).

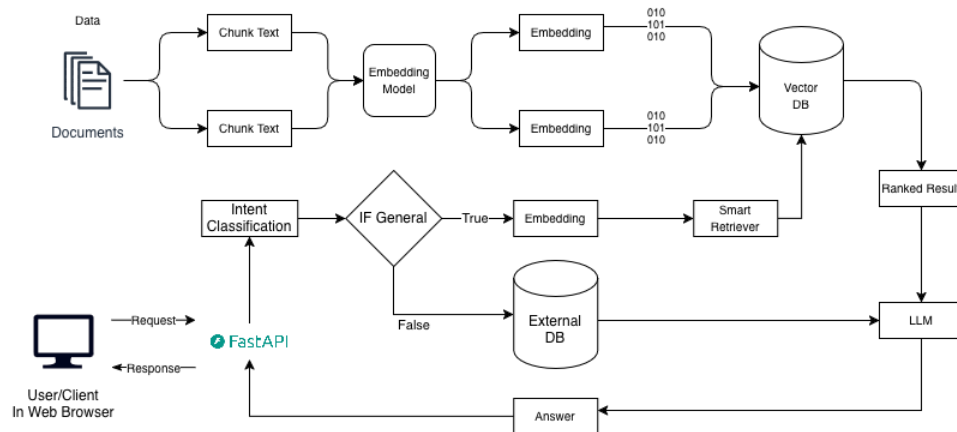
Jumlah pertanyaan uji yang relatif terbatas dipilih secara sadar karena tujuan utama penelitian ini adalah menguji kelayakan penerapan metode RAG sebagai *proof-of-concept* pada domain layanan publik kependudukan, bukan untuk melakukan generalisasi statistik berskala besar. Pendekatan serupa juga digunakan pada sejumlah penelitian awal RAG berbasis domain-spesifik, yang menitikberatkan pada kualitas semantik jawaban dan kesesuaian konteks dibandingkan kuantitas data uji. Selain itu, penelitian ini belum menyertakan perbandingan langsung dengan baseline model LLM tanpa RAG, karena fokus utama riset diarahkan pada evaluasi kemampuan RAG dalam menyajikan informasi faktual berbasis dokumen resmi Disdukcapil. Namun demikian, keberadaan pertanyaan yang tidak memiliki referensi jawaban dalam basis data digunakan sebagai kontrol kualitatif untuk mengamati kecenderungan *hallucination model*. Hasil pengujian menunjukkan bahwa

sistem tidak menghasilkan jawaban spekulatif ketika informasi tidak tersedia, yang mengindikasikan efektivitas mekanisme retrieval dalam membatasi keluaran model generatif.

3. HASIL DAN ANALISIS

Arsitektur Sistem

Arsitektur sistem yang dikembangkan terdiri atas dua komponen utama, yaitu sisi server yang menjalankan mekanisme *Retrieval-Augmented Generation* (RAG) dan sisi *client* berupa antarmuka *chatbot* berbasis web seperti terlihat pada gambar 2.

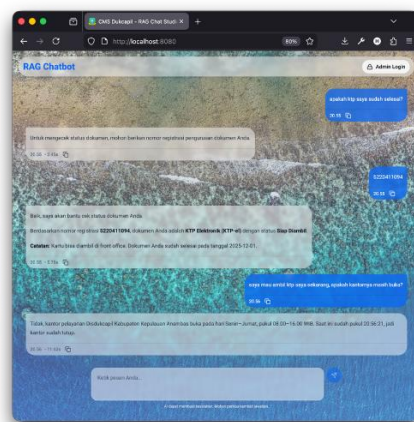


Gambar 2 Arsitektur Sistem

Sistem ini mengintegrasikan berbagai jenis data dari API yang dibuat oleh dokumen PDF, data FAQ, serta layanan pelacakan status kependudukan. Data yang bersifat statis seperti dokumen PDF dan FAQ diproses melalui tahap *chunking* dan *embedding* menggunakan model lokal *bge-m3:latest*, lalu disimpan dalam *vector* database *ChromaDB*. Sementara itu, data pelacakan yang bersifat dinamis diakses secara langsung melalui API eksternal setiap kali pengguna melakukan permintaan pelacakan. *FastAPI* berfungsi sebagai jembatan komunikasi antara *client* dan *server* sistem RAG untuk mengelola alur permintaan dan tanggapan secara efisien. Ketika pengguna mengirimkan pertanyaan, sistem akan melakukan *intent classification* [20]. Jika berupa pelacakan, sistem memanggil API eksternal, dan jika bersifat umum, sistem mengeksekusi proses *retrieval* dan *generation* melalui pipeline RAG untuk menghasilkan jawaban kontekstual.

Tampilan Antarmuka Chatbot

Antarmuka *chatbot* dikembangkan dalam bentuk halaman web yang menampilkan jendela percakapan interaktif antara pengguna dan sistem. Tampilan dirancang sederhana agar mudah digunakan tanpa memerlukan panduan teknis seperti tampilan *chatbot* pada umumnya seperti terlihat pada gambar 3.



Gambar 3 Tampilan Antarmuka Chatbot

Pada bagian bawah terdapat kolom input untuk mengetik pertanyaan, sementara area utama menampilkan riwayat percakapan dalam format *bubble chat*. Ketika pengguna mengirimkan pertanyaan, data tersebut dikirim ke *endpoint FastAPI* dan diproses oleh sistem RAG. Chatbot kemudian menampilkan hasil jawaban dari LLM

yang telah mempertimbangkan konteks dokumen atau hasil pelacakan. Selama pengujian, sistem mampu memberikan respons dalam waktu rata-rata 4–15 detik tergantung kompleksitas pertanyaan dan kecepatan internet, dengan hasil jawaban yang relevan terhadap konteks layanan kependudukan.

Evaluasi Akurasi RAG

Evaluasi RAG dilakukan menggunakan 10 pertanyaan uji yang diambil dari data statis. Dari jumlah tersebut, 9 pertanyaan memiliki jawaban referensi sehingga memungkinkan pengukuran kesesuaian semantik antara respons sistem dan jawaban rujukan. Sementara itu, 1 pertanyaan tidak memiliki jawaban dalam basis data dan digunakan untuk menilai kemampuan model dalam mempertahankan batas konteks, yakni apakah sistem tetap memberikan respons sesuai pengetahuan internal tanpa menarik informasi eksternal yang tidak tersedia. Pendekatan ini memastikan evaluasi mencakup aspek akurasi maupun kontrol terhadap keluaran model seperti terlihat pada tabel 3.

Tabel 3. Daftar Pertanyaan dan Referensi Jawaban

Pertanyaan	Referensi Jawaban
Apa Visi Disdukcapil Kepulauan Anambas?	Terciptanya masyarakat sadar akan tertib administrasi kependudukan.
Dimana alamat Disdukcapil Kepulauan Anambas?	jl. imam bonjol no. 50, tarempa, kecamatan siantan.
Apakah ada biaya yang harus dibayarkan untuk mengurus dokumen kependudukan, khususnya KTP?	layanan pembuatan ktp elektronik tidak dipungut biaya
Apa Tugas Pokok Disdukcapil Kepulauan Anambas?	Melaksanakan urusan pemerintahan yang menjadi kewenangan daerah di bidang administrasi kependudukan dan pencatatan sipil dan tugas pembantuan yang diberikan kepada daerah
Apa persyaratan pembuatan kartu keluarga yang rusak?	- Mengisi Formulir - Surat Keterangan Hilang Dari Kepolisian - Fotocopy KTP-el - KK Yang Rusak
Apa Syarat yang Diperlukan untuk Pembuatan KTP?	Persyaratan dalam pembuatan KTP Elektronik, yaitu: 1. Fotokopi Kartu Keluarga (KK) 2. KTP-el asli (Bagi KTP-el yang rusak) 3. Surat keterangan kehilangan dari kepolisian (Bagi KTP-el yang hilang)
Jika Saya ingin membuat SKPWNI, bagaimana mekanisme atau prosedurnya?	Berikut Ini Merupakan Sistem, Mekanisme, Dan Prosedur Dalam Proses Pembuatan SKPWNI: 1. menuju ke front office, untuk mengisi buku tamu, mengambil nomor antrian, menyerahkan berkas persyaratan ke front office 2. Menunggu panggilan sesuai urutan nomor antrian 3. Petugas memberikan resi pengambilan Menuju ke front office untuk pengambilan hasil (SKPWNI)
Apa Motto Pelayanan Disdukcapil Kepulauan Anambas?	Disdukcapil Kepulauan Anambas memiliki motto “BERMADAH” B: BERSIH (BERSIH DARI PUNGUTAN LIAR) E: EFISIEN (PELAYANAN CEPAT DAN TEPAT) R: RESPONSIF (CEPAT TANGGAP DALAM PELAYANAN) M: MUDAH (JENIS PELAYANAN JELAS DAN GRATIS) A: AKUNTABEL (PELAYANAN DAPAT DI PERTANGGUNG JAWABKAN SESUAI PERATURAN) D: DEDIKASI (PENGORBANAN TANPA BATAS DALAM PELAYANAN) A: AMANAH (MENJAGA KERAHASIAAN DATA PENDUDUK) H: HORMAT (HORMATI PEMOHON DENGAN SEPENUH HATI)
Apa Persyaratan Pembuatan Akta Kelahiran?	Berikut adalah persyaratan penerbitan Akta Kelahiran untuk anak yang lahir di dalam Indonesia: 1. Fotokopi surat keterangan kelahiran dari rumah sakit, puskesmas, atau bidan. 2. Fotokopi Kartu Keluarga (KK) orang tua. 3. Fotokopi Kartu Tanda Penduduk elektronik (KTP-el) orang tua. 4. Fotokopi buku nikah atau akta perkawinan orang tua. Berikut adalah persyaratan penerbitan Akta Kelahiran untuk anak yang lahir di luar Indonesia: 1. Fotokopi surat keterangan kelahiran dari rumah sakit atau instansi berwenang di negara setempat. Fotokopi paspor orang tua. 2. Fotokopi Kartu Keluarga (KK) orang tua. 3. Fotokopi buku nikah atau akta perkawinan orang tua.
Siapa Bupati Kabupaten Kepulauan Anambas tahun 2024?	Maaf, saya tidak menemukan informasi tersebut dalam data yang saya miliki.

Evaluasi sistem dilakukan menggunakan metrik *BERTScore* [21], [22], yaitu metode evaluasi berbasis *semantic similarity* yang menghitung kesamaan makna antara jawaban sistem dan jawaban referensi. *BERTScore* memanfaatkan *embedding* dari model *Bidirectional Encoder Representations from Transformers* (BERT) untuk mengukur kesamaan antar token berdasarkan *cosine similarity*. Dari hasil perhitungan tersebut diperoleh tiga komponen utama, yaitu *Recall*, *Precision*, dan *F1-score* yang merupakan rata-rata harmonik antara keduanya. Rumus perhitungan *BERTScore* ditunjukkan pada Persamaan (1)–(3) berikut.

$$R_{BERT} = \frac{1}{|x|} \sum_{x_i \in x} \max_{\hat{x}_j \in \hat{x}} x_i^T \hat{x}_j, \quad (1)$$

Persamaan (1) menunjukkan perhitungan *Recall*, yaitu rata-rata nilai kesamaan tertinggi antara setiap token pada teks referensi dengan token paling mirip pada teks prediksi. Nilai ini menggambarkan sejauh mana makna dari teks referensi tercakup dalam jawaban yang dihasilkan oleh sistem.

$$P_{BERT} = \frac{1}{|\hat{x}|} \sum_{\hat{x}_j \in \hat{x}} \max_{x_i \in x} x_i^T \hat{x}_j, \quad (2)$$

Persamaan (2) mendefinisikan *Precision*, yaitu rata-rata kemiripan setiap token pada teks prediksi terhadap token paling mirip dalam teks referensi. Nilai *Precision* yang tinggi menunjukkan bahwa jawaban model memiliki kesesuaian makna yang tinggi terhadap referensi.

$$BERTScore = 2 \frac{P_{BERT} \cdot R_{BERT}}{P_{BERT} + R_{BERT}} \quad (3)$$

Persamaan (3) menunjukkan *F1-score*, yaitu rata-rata harmonik antara *Precision* dan *Recall* yang digunakan untuk menilai keseimbangan antara kelengkapan makna dan ketepatan semantik. Nilai *F1-score* menjadi indikator utama dalam menentukan tingkat relevansi dan kesamaan semantik antara jawaban sistem dan jawaban referensi seperti terlihat pada tabel 4.

Tabel 4. Hasil Evaluasi Bertscore

Pertanyaan	F1-Score	Precision	Recall
Apa Visi Disdukcapil Kepulauan Anambas?	0.989	0.998	0.993
Dimana Alamat Disdukcapil Kepulauan Anambas?	0.981	0.988	0.984
Apakah ada biaya yang harus dikeluarkan untuk membuat KTP?	0.974	0.980	0.977
Apa Tugas Pokok Disdukcapil Kepulauan Anambas?	0.991	0.996	0.993
Apa Persyaratan Pembuatan Kartu Keluarga yang rusak?	0.986	0.988	0.987
Apa syarat yang diperlukan untuk pembuatan KTP?	0.994	0.994	0.994
Jika saya ingin membuat SKPWNI, bagaimana mekanismenya?	0.988	0.993	0.990
Apa motto pelayanan Disdukcapil Kepulauan Anambas?	0.980	0.989	0.985
Apa persyaratan pembuatan Akta Kelahiran?	0.984	0.991	0.988
Siapa Bupati Kabupaten Kepulauan Anambas tahun 2024?	0.986	0.994	0.990
Rata-rata	0.988	0.985	0.991

Hasil pengujian menunjukkan nilai rata-rata *Precision* sebesar 0.985, rata-rata *Recall* sebesar 0.991, dan rata-rata *F1-Score* sebesar 0.988, yang menandakan tingkat relevansi jawaban yang sangat tinggi. Nilai tersebut membuktikan bahwa penerapan metode RAG secara efektif meningkatkan akurasi jawaban dan mengurangi halusinasi pada model LLM. Secara keseluruhan, sistem yang dibangun mampu memberikan jawaban faktual dan kontekstual terhadap pertanyaan masyarakat serta berpotensi meningkatkan efisiensi layanan informasi publik Disdukcapil Kepulauan Anambas.

Tingginya skor *BERTScore* dalam penelitian ini dipengaruhi oleh beberapa faktor, antara lain sifat domain yang sempit dan terkontrol, penggunaan sumber data resmi yang telah terkurasi, serta kesamaan struktur bahasa antara jawaban referensi dan keluaran sistem. Oleh karena itu, skor yang tinggi lebih merefleksikan kemampuan sistem dalam menjaga konsistensi semantik dan relevansi konteks, bukan semata-mata kecerdasan generatif model secara umum.

Selain itu, metrik *BERTScore* memiliki keterbatasan dalam membedakan variasi redaksi yang masih bermakna sama dengan kesalahan faktual yang bersifat implisit. Hal ini terlihat pada beberapa kasus di mana sistem memberikan jawaban dengan formulasi berbeda namun tetap memperoleh skor tinggi karena kesamaan makna semantik. Dengan demikian, evaluasi kuantitatif dalam penelitian ini diposisikan sebagai indikator pendukung, yang dilengkapi dengan observasi kualitatif terhadap relevansi dan kepatuhan sistem terhadap batas pengetahuan yang tersedia. Pendekatan ini menegaskan bahwa keberhasilan utama sistem tidak hanya terletak pada nilai metrik yang tinggi, tetapi pada kemampuannya mengurangi *hallucination*, menjaga akurasi berbasis dokumen, dan memberikan respons yang sesuai dengan konteks layanan publik.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan *chatbot* berbasis *Large Language Model* dengan pendekatan *Retrieval-Augmented Generation* untuk mendukung peningkatan layanan informasi publik di Disdukcapil

Kabupaten Kepulauan Anambas. Sistem dibangun dengan arsitektur yang mengintegrasikan berbagai sumber data, meliputi dokumen PDF layanan Disdukcapil Anambas, regulasi seperti Peraturan Presiden Nomor 96 dan Permendagri Nomor 108 Tahun 2019 tentang persyaratan dan tata cara pendaftaran penduduk dan pencatatan sipil, himpunan FAQ, serta API pelacakan status dokumen kependudukan. Seluruh data statis diproses melalui tahapan chunking dan embedding ke dalam *vector store*, sementara *FastAPI* berperan sebagai antarmuka komunikasi antara pengguna dan server RAG. Hasil pengujian menunjukkan bahwa sistem mampu memberikan jawaban yang relevan dan kontekstual dengan waktu respons rata-rata 4–15 detik dan nilai evaluasi BERTScore dengan rata-rata *Precision* 98,5%, rata-rata *Recall* 99,1%, dan rata-rata *F1-Score* 98,8%, yang menandakan tingkat akurasi yang sangat tinggi. Penerapan metode RAG terbukti efektif dalam meningkatkan relevansi jawaban serta mengurangi potensi hallucination pada model LLM, sehingga sistem ini dapat menjadi solusi inovatif untuk mempercepat dan mempermudah akses informasi kependudukan. Pengembangan sistem selanjutnya dapat difokuskan pada peningkatan efisiensi pemrosesan konteks, khususnya melalui optimalisasi jumlah token input pada tahap retrieval maupun generation. Optimalisasi ini penting untuk mengurangi penggunaan konteks yang berlebihan, menekan beban komputasi, serta menjaga ketepatan jawaban tanpa mengorbankan konsistensi semantik.

REFERENSI

- [1] Moh. As'adi, "Transformasi Digital dalam Pelayanan Administrasi Kependudukan di Kabupaten Banyuwangi," *Konstitusi : Jurnal Hukum, Administrasi Publik, dan Ilmu Komunikasi*, vol. 1, no. 4, pp. 49–59, Sep. 2024, doi: 10.62383/konstitusi.v1i4.158.
- [2] N. Ritonga *et al.*, "Sosialisasi Administrasi Kependudukan Guna Mencapai Pelayanan Administrasi Kependudukan yang Efektif di Desa Sipaku Area, Kecamatan Simpang Empat," *Communnity Development Journal*, vol. 5, no. 2, pp. 3920–3923, 2024, doi: <https://doi.org/10.31004/cdj.v5i2.27662>.
- [3] J. Sutopo, "Pentingnya Tata Informasi Tentang Pelayanan Publik Bagi Masyarakat," *Journal of Rural and Development*, vol. 1, no. 1, pp. 15–26, 2014.
- [4] B. M. Setiawan, Z. Zulhawati, A. L. Maitri, and J. Sutopo, "Chatbot Services at Educational Institutions With Customer Relationship Management," *Journal of International Conference Proceedings*, Apr. 2019, doi: 10.32535/jicp.v2i1.404.
- [5] M. R. Anam, A. S. Akbar, and H. Saputro, "QnA Chatbot with Mistral 7B and RAG method: Traffic Law Case Study," *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, vol. 15, no. 03, p. 207, Jan. 2025, doi: 10.24843/lkjiti.2024.v15.i03.p06.
- [6] M. Dimas Arya Muhajir *et al.*, "Implementasi Chatbot Menggunakan Framework Langchain Berbasis LLM GPT (Studi Kasus : Panduan Akademik Universitas Trunojoyo)," 2025.
- [7] L. R. Hidayat, I. G. P. S. Wijaya, and R. Dwiyanaputra, "Optimalisasi Layanan Sistem Informasi Mahasis dengan Integrasi Telegram : Chatbot Retrieval-Augmented-Generation berbasis Large Language Model," *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTika)*, vol. 7, no. 1, pp. 121–131, Mar. 2025, doi: 10.29303/jtika.v7i1.459.
- [8] M. M. Mar'arif, N. Safaat Harahap, and M. Affandes, "Sistem Tanya-Jawab Berbasis Chatbot Telegram Tentang Fiqih Kontemporer Menggunakan Langchain Dan LLM," *IJCCS*, vol. 19, No.2, pp. 1–5, Apr. 2025, doi: <https://doi.org/10.5281/zenodo.15437200>.
- [9] Aji Bayu Permadi, Nazruddin Safaat H, Lestari Handayani, and Yusra, "Implementasi Question Answering System Tafsir Al-Azhar Menggunakan Langchain dan Large Language Model Berbasis Chatbot Telegram," *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, vol. 12, no. 1, pp. 62–69, Apr. 2024, doi: 10.21063/jtif.2024.v12.1.62-69.
- [10] W. Khan, A. Daud, K. Khan, S. Muhammad, and R. Haq, "Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends," *Natural Language Processing Journal*, vol. 4, p. 100026, Sep. 2023, doi: 10.1016/j.nlp.2023.100026.
- [11] E. Adamopoulou and L. Moussiades, "Chatbots: History, technology, and applications," *Machine Learning with Applications*, vol. 2, p. 100006, Dec. 2020, doi: 10.1016/j.mlwa.2020.100006.
- [12] B. El Bakkouri, S. Raki, and T. Belgnaoui, "The Role of Chatbots in Enhancing Customer Experience: Literature Review," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 432–437. doi: 10.1016/j.procs.2022.07.057.

- [13] O. El Gharbaoui, H. El Boukhari, and A. Salmi, "Chatbots and Citizen Satisfaction: Examining the Role of Trust in AI-Chatbots as a Moderating Variable," *TEM Journal*, vol. 13, no. 3, pp. 1825–1836, Aug. 2024, doi: 10.18421/TEM133-11.
- [14] C. Anitha, K. R. Prasad, G. Shrivastava, A. Kurkarni, G. Kumar, and S. Oza, "Applications of Natural Language Processing in Chatbot," *Journal of Information Systems Engineering and Management*, vol. 10, no. 33s, pp. 1099–1106, Apr. 2025, doi: 10.52783/jisem.v10i33s.5936.
- [15] D. L. Chen, K. Aaltonen, H. Lampela, and J. Kujala, "The Design and Implementation of an Educational Chatbot with Personalized Adaptive Learning Features for Project Management Training," *Technology, Knowledge and Learning*, vol. 30, no. 2, pp. 1047–1072, Jun. 2025, doi: 10.1007/s10758-024-09807-5.
- [16] T. Guo *et al.*, "Large Language Model based Multi-Agents: A Survey of Progress and Challenges," Apr. 2024, [Online]. Available: <http://arxiv.org/abs/2402.01680>
- [17] H. Sheikh, C. Prins, and E. Schrijvers, *Mission AI*. in Research for Policy. Cham: Springer International Publishing, 2023. doi: 10.1007/978-3-031-21448-6.
- [18] R. Patil and V. Gudivada, "A Review of Current Trends, Techniques, and Challenges in Large Language Models (LLMs)," *Applied Sciences*, vol. 14, no. 5, p. 2074, Mar. 2024, doi: 10.3390/app14052074.
- [19] A. Salemi and H. Zamani, "Evaluating Retrieval Quality in Retrieval-Augmented Generation," in *SIGIR 2024 - Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, Inc, Jul. 2024, pp. 2395–2400. doi: 10.1145/3626772.3657957.
- [20] B. Sharma, M. Madhavi, and H. Li, "Leveraging Acoustic and Linguistic Embeddings from Pretrained speech and language Models for Intent Classification," Feb. 2021, [Online]. Available: <http://arxiv.org/abs/2102.07370>
- [21] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, "BERTScore: Evaluating Text Generation with BERT," Feb. 2020, [Online]. Available: <http://arxiv.org/abs/1904.09675>
- [22] J. Shor, R. A. Bi, S. Venugopalan, S. Ibara, R. Goldenberg, and E. Rivlin, "Clinical BERTScore: An Improved Measure of Automatic Speech Recognition Performance in Clinical Settings," in *Proceedings of the 5th Clinical Natural Language Processing Workshop*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2023, pp. 1–7. doi: 10.18653/v1/2023.clinicalnlp-1.1.