

# Pengembangan Chatbot Informasi Kesehatan Ibu dan Anak Jawa Barat Berbasis Hybrid RAG dan TextToSQL

Muhammad Alam Basallamah<sup>1</sup>, Lala Septem Riza<sup>2\*</sup>, Ani Anisyah<sup>3</sup>

<sup>1,2,3</sup> Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Ilmu Komputer, Universitas Pendidikan Indonesia, Bandung, Indonesia

Email: <sup>1</sup>basalamahalam@upi.edu, <sup>2,\*</sup>lala.s.riza@upi.edu, <sup>3</sup>anianisyah@upi.edu

(\* Email Corresponding Author: lala.s.riza@upi.edu)

Received: January 8, 2026 | Revision: January 10, 2026 | Accepted: January 10, 2026

## Abstrak

Tingginya Angka Kematian Ibu dan prevalensi *stunting* di Jawa Barat menghambat pencapaian SDGs. Aksesibilitas informasi menjadi kendala utama karena data tersebar dalam format narasi dokumen Profil Kesehatan dan Buku KIA. Penelitian ini mengembangkan chatbot hibrida berbasis *Retrieval-Augmented Generation* (RAG) dan *Text-to-SQL* menggunakan *Large Language Model* Google Gemini. Sistem mengintegrasikan mekanisme *chunking* hierarkis, *router* cerdas, memori percakapan, dan visualisasi data otomatis untuk menyajikan informasi faktual serta statistik secara presisi. Evaluasi sistem menunjukkan kinerja tinggi: *Text-to-SQL* mencapai akurasi eksekusi 93,33% dan akurasi *Router* 85%. Pada modul RAG, pengujian 60 pertanyaan menghasilkan skor *Faithfulness* 0,990, *Context Recall* 0,883, dan *Answer Relevancy* 0,950. Sistem mencapai skor sempurna 1,000 untuk penanganan pertanyaan di luar konteks, membuktikan keamanan dari halusinasi. Pengujian pengguna (UAT) mencatat tingkat keberhasilan 80%, di mana fitur grafik dinilai signifikan dalam membantu pemahaman data. Penelitian menyimpulkan bahwa pendekatan hibrida efektif meningkatkan aksesibilitas wawasan kesehatan bagi masyarakat dan pemangku kebijakan, meskipun ekstraksi tabel kompleks memerlukan optimalisasi lanjut.

**Kata Kunci:** Chatbot, Kesehatan Ibu dan Anak, Hybrid, RAG, Text To SQL, Large Language Model

## Abstract

The high Maternal Mortality Rate and *stunting* prevalence in West Java hinder the achievement of SDGs. Information accessibility remains a key constraint as data is scattered in narrative formats within Health Profiles and MCH Books. This study develops a hybrid chatbot based on *Retrieval-Augmented Generation* (RAG) and *Text-to-SQL* using the Google Gemini Large Language Model. The system integrates hierarchical chunking, intelligent routing, conversational memory, and automatic data visualization to present factual and statistical information precisely. System evaluation demonstrates high performance: *Text-to-SQL* reached 93.33% execution accuracy and 85% Router accuracy. In the RAG module, testing on 60 questions yielded scores of 0.990 Faithfulness, 0.883 Context Recall, and 0.950 Answer Relevancy. The system achieved a perfect score of 1.000 for out-of-context handling, proving safety from hallucinations. User Acceptance Testing (UAT) recorded a 80% success rate, where the chart feature was rated significant in aiding data understanding. The study concludes that the hybrid approach effectively enhances health insight accessibility for the public and stakeholders, although complex table extraction requires further optimization.

**Keywords:** Chatbot, Maternal and Child Health, Hybrid, RAG, Text To SQL, Large Language Model

## 1. PENDAHULUAN

Kemajuan teknologi di era globalisasi telah membawa transformasi fundamental dalam berbagai sektor, di mana kesehatan menjadi pilar utama dalam membentuk masyarakat yang produktif dan sejahtera [1]. Dalam ekosistem digital saat ini, pemanfaatan Teknologi, Informasi, dan Komunikasi (TIK) melalui Sistem Informasi Kesehatan (SIK) menjadi instrumen vital untuk meningkatkan efektivitas pelayanan kesehatan nasional. Tantangan terbesar dalam dunia medis modern adalah laju pertumbuhan informasi yang eksponensial; ribuan artikel medis dan data klinis dipublikasikan setiap tahunnya, membuat tenaga kesehatan dan masyarakat berisiko tertinggal jika tidak didukung oleh alat bantu teknologi yang memadai [2]. Oleh karena itu, teknologi informasi kini tidak hanya berfungsi sebagai wadah penyimpanan data pasif, tetapi dituntut memiliki kapabilitas cerdas untuk memfilter, mengolah, dan menyajikan data mentah menjadi wawasan yang dapat ditindaklanjuti guna mendukung pengambilan keputusan yang presisi [3].

Di tingkat regional, Pemerintah Provinsi Jawa Barat terus mengakselerasi transformasi digital kesehatan demi mencapai *Sustainable Development Goals* (SDGs) poin ke-3, yakni menjamin kehidupan yang sehat dan sejahtera bagi semua usia. Upaya ini tercermin melalui berbagai inisiatif strategis, mulai dari integrasi sistem rekam medis elektronik, layanan *telemedicine*, hingga peluncuran portal transparansi informasi seperti *Open Data Jabar* dan aplikasi Sehat Jabar [4]. Langkah ini diperkuat oleh regulasi nasional melalui Undang-Undang Kesehatan Nomor 17 Tahun 2023 yang menekankan urgensi integrasi data untuk mengatasi fragmentasi tata kelola kesehatan [5]. Harapannya, digitalisasi ini mampu memangkas inefisiensi birokrasi dan memudahkan masyarakat dalam mengakses riwayat serta layanan kesehatan secara *real-time*.

Meskipun berbagai inovasi teknologi telah diimplementasikan, Provinsi Jawa Barat masih dihadapkan pada tantangan serius terkait indikator kesehatan ibu dan anak (KIA). Data statistik menunjukkan kondisi yang memprihatinkan, di mana Angka Kematian Ibu (AKI) masih cukup tinggi dengan tercatatnya 444 kasus pada triwulan III

tahun 2023, yang didominasi oleh penyebab medis seperti eklampsia dan perdarahan [6]. Selain itu, masalah *stunting* pada balita masih menjadi isu kritis dengan prevalensi mencapai 21,7%, angka yang masih jauh dari target nasional. Kesenjangan akses layanan kesehatan di wilayah padat penduduk, seperti Kabupaten Bogor dan Bandung, memperburuk situasi ini [7]. Fenomena ini mengindikasikan bahwa ketersediaan data saja tidak cukup; diperlukan mekanisme pemantauan dan edukasi yang lebih proaktif untuk menekan angka mortalitas dan morbiditas tersebut.

Permasalahan inti yang teridentifikasi dalam penelitian ini adalah adanya kesenjangan (*gap*) signifikan antara ketersediaan data kesehatan yang masif dengan aksesibilitas informasi tersebut bagi pengguna akhir. Data kesehatan sering kali tersebar dalam dokumen naratif yang panjang (seperti Buku KIA dan PDF Profil Kesehatan) atau tersimpan dalam format tabel statistik yang kompleks, sehingga sulit dipahami secara cepat oleh masyarakat awam maupun bidan desa. Untuk menjembatani kesenjangan ini, diperlukan solusi berbasis kecerdasan buatan (*Artificial Intelligence*) berupa *chatbot* yang mampu mengumpulkan, menstandarisasi, dan mendeteksi risiko kesehatan secara dini [8]. Namun, *chatbot* konvensional sering kali memiliki keterbatasan dalam memahami konteks spesifik dan menyajikan data akurat.

Untuk mengatasi keterbatasan tersebut, pendekatan *Retrieval-Augmented Generation* (RAG) dan *Text-to-SQL* menjadi solusi teknologi yang relevan. RAG menggabungkan kemampuan model bahasa besar (*Large Language Model/LLM*) dengan mekanisme pencarian dokumen eksternal, sehingga memungkinkan sistem menjawab pertanyaan kualitatif berdasarkan fakta terkini dan mengurangi risiko halusinasi pada model [9]. Sementara itu, *Text-to-SQL* memungkinkan pengguna untuk mengekstraksi data kuantitatif dari basis data relasional menggunakan bahasa alami, tanpa perlu memahami sintaks pemrograman SQL [10]. Integrasi kedua teknologi ini memungkinkan terciptanya sistem hibrida yang mampu menangani data tidak terstruktur (teks) dan terstruktur (angka) secara simultan.

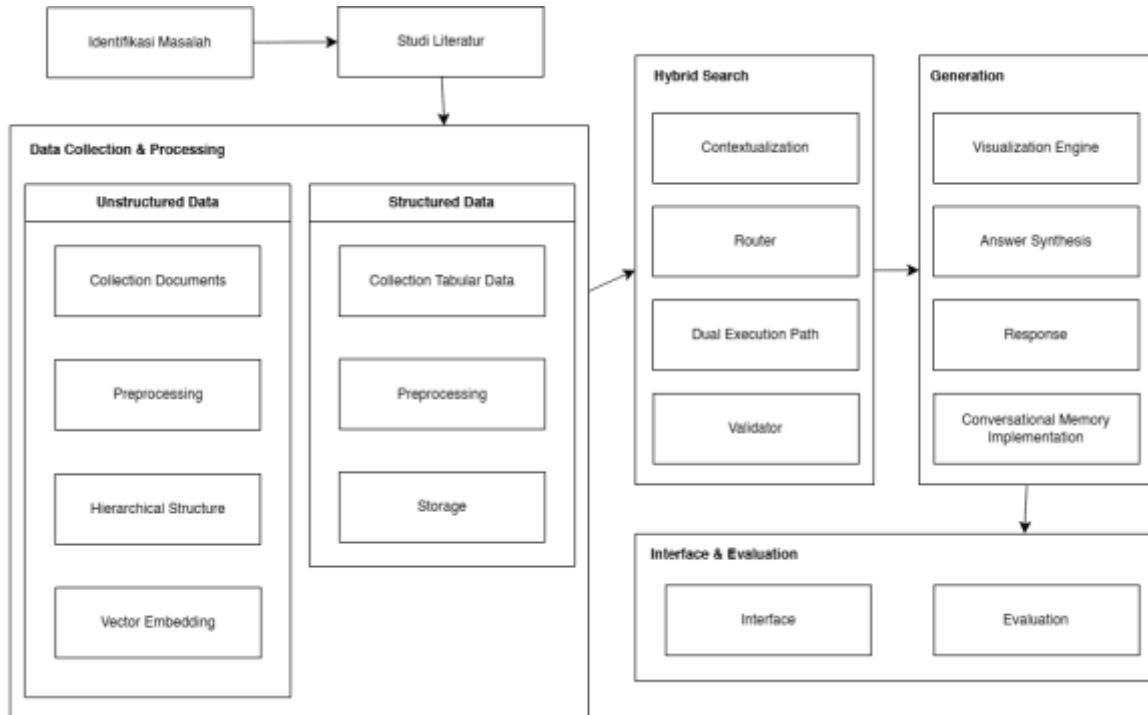
Sejumlah penelitian terdahulu telah berupaya mengembangkan *chatbot* untuk keperluan kesehatan dan layanan publik. Rohim & Zuliarso [11] menerapkan algoritma Deep Learning dalam pengembangan *chatbot* informasi Covid-19 yang mampu memberikan respons cepat dan relevan, namun sistem ini masih terbatas pada data statis dan belum mendukung integrasi sumber data eksternal secara dinamis. Purwitasari & Soleh [12] mengimplementasikan *Artificial Neural Network* (ANN) untuk pencarian peraturan perundang-undangan dengan akurasi 94,28%, namun model ini bersifat kaku terhadap variasi pertanyaan di luar data latih. Haryono & Prasetya [13] mengembangkan sistem *Text-to-SQL* berbasis aturan (*rule-based*) untuk mengidentifikasi informasi kuantitas dalam kalimat bahasa Indonesia dengan pendekatan *forward chaining*, namun sistem ini masih bersifat deterministik dan kurang fleksibel dalam menangani variasi bahasa di luar aturan yang telah didefinisikan.

Penelitian yang lebih mutakhir mulai mengadopsi teknologi *Large Language Model* (LLM) dengan metode *Retrieval-Augmented Generation* (RAG) untuk mengatasi keterbatasan konteks dan halusinasi. Samudra et al. [14] mengimplementasikan RAG pada *chatbot* kesehatan pencernaan menggunakan model Llama 3.1, yang terbukti efektif meningkatkan validitas informasi medis dengan skor *Mean Reciprocal Rank* (MRR) mencapai 93% dan *faithfulness* sebesar 62%. Meskipun metode RAG pada penelitian Samudra et al. berhasil meminimalisir kesalahan informasi pada data teks tidak terstruktur, sistem tersebut belum dirancang untuk mengolah data terstruktur (statistik) secara akurat.

Untuk menjawab tantangan pengolahan data yang heterogen, Kachhwal et al. [15] memperkenalkan sistem pencarian hibrida (*Multi-Modal Hybrid Search*) yang menggabungkan RAG dan *Text-to-SQL*. Penelitian tersebut membuktikan bahwa integrasi kedua metode ini memungkinkan pengguna mengakses informasi dari dokumen (PDF/TXT) dan basis data (PostgreSQL) dalam satu antarmuka terpadu. Namun, penelitian Kachhwal et al. masih bersifat umum (*general purpose*) dan belum diterapkan secara spesifik pada kompleksitas domain kesehatan daerah di Indonesia. Hingga saat ini, belum ada penelitian yang secara komprehensif mengadaptasi arsitektur hibrida tersebut untuk konteks Kesehatan Ibu dan Anak (KIA), di mana kebutuhan informasi mencakup penjelasan naratif pedoman medis sekaligus visualisasi data statistik vital secara *real-time*.

Berdasarkan analisis masalah dan peluang tersebut, penelitian ini bertujuan untuk merancang bangun *chatbot* hibrida yang mengintegrasikan arsitektur RAG dan *Text-to-SQL* menggunakan model Google Gemini. Sistem ini dirancang untuk memberikan informasi kesehatan ibu dan anak di Jawa Barat secara akurat, baik berupa penjelasan prosedural maupun visualisasi data statistik. Kinerja sistem akan diukur secara komprehensif menggunakan metrik *Execution Accuracy* untuk komponen SQL, kerangka kerja RAGAS (meliputi *Faithfulness*, *Context Recall*, *Answer Relevancy*) untuk komponen RAG, serta *User Acceptance Testing* (UAT) untuk mengukur keberterimaan pengguna. Penelitian ini diharapkan dapat menghadirkan alat bantu yang efektif bagi masyarakat dan pemangku kebijakan dalam mengakses informasi kesehatan berbasis data.

## 2. METODOLOGI PENELITIAN



**Gambar. 1** Tahapan Penelitian

Gambar 1 mengilustrasikan tahapan penelitian yang diterapkan dalam studi ini. Alur kerja disusun secara sistematis yang diawali dengan identifikasi masalah dan studi literatur untuk memetakan kesenjangan teknologi terkini, dilanjutkan dengan pengembangan teknis yang terbagi ke dalam empat modul utama: *Data Collection & Processing*, *Hybrid Search*, *Generation*, serta *Interface & Evaluation*.

Kerangka kerja ini dirancang khusus untuk mengimplementasikan arsitektur *Hybrid Retrieval Augmented Generation* (RAG) dan *Text-to-SQL*. Pada tahap *Data Collection & Processing*, dilakukan pemisahan jalur pemrosesan data yang ketat: data tidak terstruktur (*unstructured*) diproses melalui *hierarchical structure* dan *vector embedding* untuk dokumen, sementara data terstruktur (*structured*) dikelola melalui penyimpanan basis data relasional. Pendekatan dualistik ini dipilih untuk memastikan sistem memiliki fleksibilitas dalam menangani berbagai jenis sumber informasi kesehatan.

Tahapan inti dari sistem terletak pada *Hybrid Search*, di mana mekanisme *Router* berfungsi sebagai otak pengambil keputusan untuk menentukan *Dual Execution Path*, apakah kueri memerlukan pencarian semantik atau eksekusi perintah SQL. Alur ini bermuara pada tahap *Generation*, yang dilengkapi dengan *Visualization Engine* dan *Conversational Memory* untuk menghasilkan sintesis jawaban yang komprehensif. Rangkaian penelitian diakhiri dengan tahap *Evaluation*, yang bertujuan memvalidasi kinerja antarmuka dan akurasi sistem dalam menjawab kebutuhan pengguna.

### 2.1 Identifikasi Masalah

Tahap awal penelitian difokuskan pada pemetaan masalah fundamental terkait aksesibilitas informasi KIA. Berdasarkan observasi awal, teridentifikasi bahwa kebutuhan informasi di Provinsi Jawa Barat belum terlayani secara optimal akibat format data yang tersebar dan akses yang lambat. Sistem yang ada saat ini belum mampu menyediakan data secara *real-time* maupun jawaban yang kontekstual. Untuk mengatasi hal tersebut, studi literatur dilakukan secara mendalam guna mengeksplorasi arsitektur *Hybrid Retrieval-Augmented Generation* (RAG) dan *Text-to-SQL*. Kajian ini bertujuan menemukan celah penelitian (*gap analysis*) dan menentukan teknologi *state-of-the-art* yang paling relevan, seperti penggunaan *Large Language Model* (LLM) Google Gemini dan basis data vektor Qdrant.

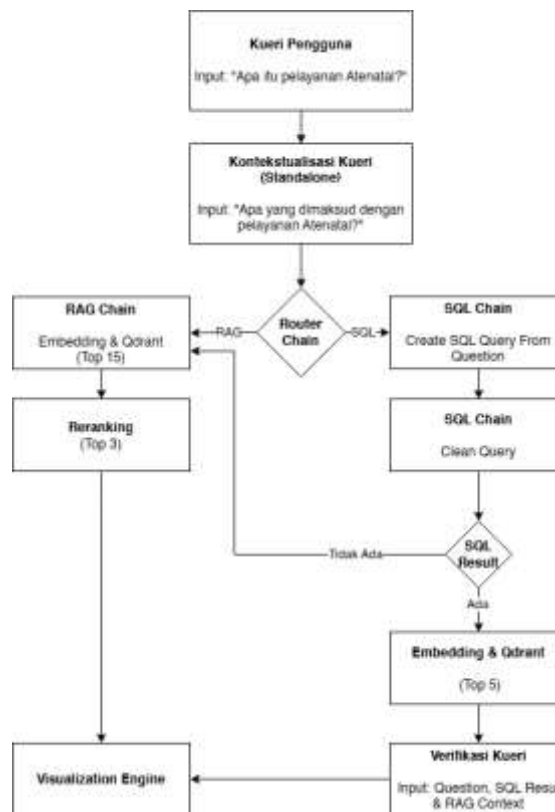
### 2.2 Studi Literatur

Studi literatur dilakukan untuk memperdalam pemahaman mengenai integrasi teknologi *Large Language Models* (LLM), khususnya Google Gemini, dengan arsitektur *Retrieval-Augmented Generation* (RAG) guna memitigasi risiko halusinasi. Selain itu, kajian difokuskan pada metode *Text-to-SQL* sebagai solusi atas keterbatasan LLM dalam analisis data statistik terstruktur. Berdasarkan tinjauan terhadap teknik pemrosesan data kompleks dan kerangka evaluasi standar seperti RAGAS, penelitian ini menetapkan penggunaan arsitektur Hibrida (penggabungan RAG dan *Text-to-SQL*) sebagai pendekatan metodologis utama yang paling efektif.

### 2.3 Data Collection & Processing

Tahap ini berfungsi sebagai fondasi utama pengembangan sistem, di mana proses akuisisi dan transformasi data dilakukan untuk mempersiapkan basis pengetahuan yang komprehensif. Proses dimulai dengan pengumpulan data yang diklasifikasikan ke dalam dua kategori utama, yaitu data tidak terstruktur (*unstructured data*) dan data terstruktur (*structured data*). Untuk data tidak terstruktur, dokumen naratif seperti Buku Kesehatan Ibu dan Anak (KIA), Laporan Profil Kesehatan Dinas Kesehatan Jawa Barat, serta regulasi terkait dikumpulkan dalam format digital. Dokumen-dokumen ini kemudian melalui proses pra-pemrosesan yang ketat untuk membersihkan elemen *noise* seperti *header* dan *footer* yang berulang. Setelah data bersih, diterapkan teknik *Hierarchical Structure Chunking* yang memecah dokumen menjadi segmen-segmen teks yang lebih kecil tanpa menghilangkan konteks bab atau sub-bab aslinya. Segmen teks tersebut selanjutnya dikonversi menjadi representasi vektor (*vector embedding*) menggunakan model *embedding* dari Google Gemini dan disimpan ke dalam basis data vektor Qdrant untuk keperluan pencarian semantik. Secara simultan, pengolahan data terstruktur dilakukan dengan mengumpulkan statistik kuantitatif, seperti Angka Kematian Ibu (AKI), prevalensi *stunting*, dan cakupan imunisasi per wilayah. Data mentah ini dinormalisasi dan distrukturisasi ke dalam skema basis data relasional (MySQL) yang dirancang khusus untuk memfasilitasi penerjemahan bahasa alami ke dalam kueri SQL, memastikan setiap entitas data memiliki relasi yang jelas dan valid.

### 2.4 Hybrid Search



Gambar. 2 Retrieval Process

Tahap ini merupakan inti pemrosesan cerdas sistem yang diawali dengan proses *Contextualization*, di mana pertanyaan pengguna diformulasi ulang menjadi kalimat mandiri (*standalone question*) untuk menghilangkan ambiguitas konteks. Pertanyaan yang telah terstruktur kemudian dianalisis oleh *Router Chain* untuk menentukan jalur eksekusi yang tepat. Jika pertanyaan bersifat kualitatif atau definisi, sistem mengaktifkan *RAG Chain* yang mengambil 15 kandidat dokumen teratas (*Top 15*) dari basis data vektor, yang kemudian disaring kembali melalui proses *Reranking* untuk mendapatkan 3 dokumen paling relevan. Sebaliknya, jika terdeteksi kebutuhan data statistik, sistem menjalankan *SQL Chain* untuk mengonversi bahasa alami menjadi kueri SQL yang bersih. Sistem ini memiliki mekanisme toleransi kesalahan yang dinamis; apabila eksekusi SQL tidak menghasilkan data (*null result*), alur secara otomatis dialihkan kembali ke *RAG Chain* sebagai langkah antisipasi (*fallback*). Namun, jika data statistik berhasil ditemukan, hasil tersebut akan diverifikasi dan diperkaya dengan konteks dokumen pendukung (*Top 3 embedding*) sebelum diteruskan ke mesin visualisasi, memastikan jawaban yang dihasilkan akurat dan komprehensif.

## 2.5 Generation

Tahap selanjutnya adalah proses sintesis jawaban akhir yang terjadi pada modul *Generation*. Pada fase ini, data yang telah berhasil diambil dari proses *retrieval* baik berupa potongan teks dokumen maupun hasil eksekusi kueri SQL diproses lebih lanjut untuk disajikan kepada pengguna. Salah satu komponen kunci dalam tahap ini adalah *Visualization Engine*, yang bertugas menerjemahkan data mentah hasil kueri SQL menjadi format visual yang intuitif, seperti grafik batang atau garis, secara otomatis. Bersamaan dengan itu, implementasi *Conversational Memory* memastikan bahwa sistem mampu menyimpan dan memanggil kembali konteks percakapan sebelumnya, sehingga interaksi dapat berjalan secara alami dan berkesinambungan. Akhirnya, seluruh komponen informasi tersebut diserahkan kepada modul *Answer Synthesis* yang memanfaatkan *Large Language Model (LLM)* untuk merangkai jawaban natural. LLM akan menggabungkan narasi penjelasan, data statistik, dan elemen visual menjadi satu respons yang koheren, akurat, dan mudah dipahami oleh pengguna.

## 2.4 Interface & Evaluation



Gambar 3. Interface Chatbot

Tahap akhir melibatkan pengembangan antarmuka pengguna berbasis web untuk memfasilitasi interaksi dengan *chatbot*. Setelah sistem terintegrasi, evaluasi menyeluruh dilakukan menggunakan metode kuantitatif dan kualitatif. Kinerja *Text-to-SQL* diukur menggunakan *Execution Accuracy*, sedangkan kualitas jawaban RAG dievaluasi menggunakan kerangka kerja RAGAS yang mencakup metrik *Faithfulness* (kesetiaan pada sumber), *Context Recall*, dan *Answer Relevancy*. Selain itu, *User Acceptance Testing (UAT)* dilakukan kepada responden relevan (masyarakat atau tenaga kesehatan) untuk memvalidasi kebermanfaatan sistem dalam skenario dunia nyata.

## 3. HASIL DAN PEMBAHASAN

Bagian ini menguraikan hasil implementasi sistem *chatbot* hibrida yang dirancang untuk menjawab tantangan aksesibilitas informasi kesehatan ibu dan anak di Provinsi Jawa Barat. Pembahasan mencakup evaluasi kinerja sistem dilakukan secara komprehensif menggunakan pendekatan kuantitatif dan kualitatif. Pengujian dibagi menjadi tiga segmen utama: evaluasi akurasi eksekusi *Text-to-SQL*, evaluasi kualitas jawaban RAG menggunakan metrik RAGAS, dan evaluasi penerimaan pengguna (*User Acceptance Testing*).

### 3.1 Evaluasi Kinerja Router

Tabel 1 Evaluasi Kinerja Router

Keterangan	Total	Persentase
Routing Benar	17	85%
Routing Salah	3	25%

Akurasi modul *Router* diuji untuk memastikan sistem tidak salah memilih jalur (SQL vs RAG). Dari 20 pertanyaan uji campuran, sistem mencatat akurasi 85%. Kesalahan klasifikasi umumnya terjadi pada pertanyaan ambigu yang mengandung angka tetapi sebenarnya menanyakan prosedur (contoh: "Apa yang harus dilakukan pada langkah ke-2 persalinan?"). Meskipun ada angka "2", ini seharusnya masuk jalur RAG, namun terkadang terdeteksi sebagai kueri statistik. Perbaikan pada *few-shot prompting* di tahap iterasi selanjutnya diharapkan dapat meningkatkan akurasi ini.

### 3.2 Evaluasi Kinerja RAG

#### 3.2.1 Hasil dari Pertanyaan Faktual

**Tabel 2.** Hasil Evaluasi Pertanyaan Faktual

Matriks Evaluasi	Skor
<i>Faithfulness</i>	0.990
<i>Context Recall</i>	0.8833
<i>Answer Relevancy</i>	0.91677

Hasil evaluasi terhadap 60 pertanyaan faktual yang dirancang untuk menguji kemampuan sistem dalam mengambil informasi spesifik dan detail dari dokumen. Hasil menunjukkan skor *Faithfulness* yang sangat tinggi mencapai 0.990, mengindikasikan bahwa jawaban yang dihasilkan sistem hampir sepenuhnya didasarkan pada konteks dokumen yang tersedia, meminimalisir risiko halusinasi. Skor *Answer Relevancy* sebesar 0.9167 menunjukkan bahwa jawaban sistem sangat relevan dan *to-the-point* terhadap pertanyaan pengguna. Namun, skor *Context Recall* tercatat sebesar 0.8833, yang meskipun tergolong kategori baik, merupakan skor terendah dibandingkan metrik lainnya. Hal ini mengindikasikan adanya tantangan minor dalam penarikan kembali (*retrieval*) informasi yang sangat spesifik, kemungkinan disebabkan oleh kompleksitas struktur tabel dalam dokumen PDF yang belum diambil sempurna oleh proses *chunking*.

#### 3.2.2 Hasil dari Pertanyaan Penalaran

**Tabel 3.** Hasil Evaluasi Pertanyaan Penalaran

Matriks Evaluasi	Skor
<i>Faithfulness</i>	0.9833
<i>Context Recall</i>	0.9167
<i>Answer Relevancy</i>	0.95

Evaluasi terhadap kemampuan penalaran sistem diuji menggunakan 20 pertanyaan yang menuntut sintesis informasi dari berbagai bagian dokumen. Berdasarkan Tabel 3, sistem menunjukkan kinerja yang sangat impresif dengan skor *Answer Relevancy* mencapai 0.95, lebih tinggi dibandingkan pertanyaan faktual. Hal ini membuktikan bahwa *Large Language Model* (LLM) yang digunakan mampu memahami nuansa pertanyaan kompleks dan menyusun argumen yang logis. Skor *Faithfulness* (0.9833) dan *Context Recall* (0.9167) yang tetap tinggi menunjukkan bahwa meskipun pertanyaan membutuhkan logika deduktif, sistem tetap berpegang teguh pada fakta yang ditemukan dalam dokumen sumber dan berhasil mengambil konteks pendukung yang lebih luas secara akurat.

#### 3.2.3 Hasil dari Pertanyaan *Out of Context*

**Tabel 4.** Hasil Evaluasi Pertanyaan di Luar Konteks

Matriks Evaluasi	Skor
<i>Faithfulness</i>	1.0
<i>Context Recall</i>	1.0
<i>Answer Relevancy</i>	1.0

Hasil pengujian ketahanan sistem terhadap 20 pertanyaan di luar domain Kesehatan Ibu dan Anak (KIA). Sistem mencapai skor sempurna 1.0 untuk ketiga metrik: *Faithfulness*, *Context Recall*, dan *Answer Relevancy*. Skor sempurna ini menegaskan bahwa mekanisme pertahanan sistem (*guardrails*) berfungsi dengan sangat baik. Sistem berhasil mengidentifikasi bahwa pertanyaan tersebut tidak memiliki relevansi dengan basis pengetahuan (*Knowledge Base*), sehingga secara konsisten menolak memberikan jawaban spekulatif atau halusinasi. Kemampuan ini sangat krusial untuk menjaga integritas *chatbot* sebagai asisten kesehatan yang terpercaya dan profesional

### 3.2 Evaluasi Kinerja TextToSQL

**Tabel 5.** Evaluasi Kinerja TextToSQL

Keterangan	Total	Persentase
Query Benar	28	93,33%
Query Salah	2	6,67%

Hasil pengujian menunjukkan bahwa sistem mencapai *Execution Accuracy* sebesar 93,33%. Dari 30 pertanyaan uji, sistem berhasil mengeksekusi 28 kueri dengan hasil data yang tepat.

### 3.2 User Acceptance Test

**Tabel 6.** Evaluasi *User Acceptance Test*

Keterangan	Total	Persentase
Jawaban Benar	32	80%
Jawaban Salah	8	20%

Selain metrik teknis, pengujian penerimaan pengguna dilakukan melibatkan 6 responden yang terdiri dari seorang ibu rumah tangga, masyarakat umum, mahasiswa, kepala keluarga, dan tenaga kesehatan (bidan). Pengujian dilakukan menggunakan kuesioner skala Likert untuk menilai aspek kegunaan (*usability*), keakuratan informasi, dan tampilan visual. Hasil UAT menunjukkan tingkat keberhasilan rata-rata 80%. Responden memberikan apresiasi tertinggi pada fitur Visualisasi Data Otomatis, yang dianggap sangat membantu dalam memahami tren kesehatan dibandingkan hanya membaca teks angka. Kemampuan *chatbot* mengingat konteks percakapan (*conversational memory*) juga dinilai membuat interaksi terasa lebih natural seperti berkonsultasi dengan manusia.

## 4. KESIMPULAN

Berdasarkan hasil perancangan, implementasi, dan pengujian *chatbot* hibrida untuk optimalisasi akses informasi Kesehatan Ibu dan Anak (KIA) di Provinsi Jawa Barat, dapat disimpulkan bahwa sistem yang dikembangkan berhasil menjawab tantangan kesenjangan antara ketersediaan data masif dan aksesibilitas informasi bagi pengguna. Integrasi arsitektur *Retrieval-Augmented Generation* (RAG) dan *Text-to-SQL* terbukti mampu mentransformasi dokumen Profil Kesehatan dan Buku KIA yang semula bersifat statis dan tersebar menjadi layanan informasi yang interaktif, terpusat, dan mudah dipahami. Fitur-fitur cerdas seperti visualisasi data otomatis yang digerakkan oleh mekanisme *Text-to-SQL* telah membawa dampak signifikan dalam penyajian data statistik, memungkinkan pengguna melihat tren kematian ibu atau prevalensi *stunting* melalui grafik intuitif alih-alih tabel mentah yang kompleks. Selain itu, modul RAG yang didukung oleh *Large Language Model* (LLM) Google Gemini terbukti andal dalam menyajikan penjelasan prosedural dan definisi medis yang akurat. Hal ini divalidasi oleh skor *Faithfulness* yang mencapai 0,990, membuktikan kemampuan sistem dalam meminimalisir risiko halusinasi dan menjaga integritas informasi medis yang disajikan. Penerapan mekanisme *Intelligent Router* berperan krusial dalam memastikan efisiensi sistem dengan mengarahkan setiap pertanyaan ke jalur eksekusi yang paling relevan, baik itu analisis statistik maupun pencarian semantik. Hasil evaluasi kinerja menunjukkan stabilitas tinggi, di mana akurasi eksekusi perintah SQL mencapai 93,33% dan tingkat keberterimaan pengguna (UAT) mencatat angka 80%. Capaian skor sempurna 1.0 pada pengujian pertanyaan di luar konteks juga menegaskan keamanan sistem dari penyalahgunaan topik. Dengan demikian, sistem *Chatbot Hybrid* ini dinyatakan telah fungsional, andal, dan siap diimplementasikan di lingkungan operasional sebagai alat bantu strategis bagi masyarakat umum maupun pemangku kebijakan dalam upaya menekan Angka Kematian Ibu dan Anak di Jawa Barat.

## REFERENCES

- [1] H. Kurniasih, K. D. Purnanti, and R. Atmajaya, "Pengembangan Sistem Informasi Penyakit Tidak Menular (Ptm) Berbasis Teknologi Informasi," *J. Teknoinfo*, vol. 16, no. 1, p. 60, 2022, doi: 10.33365/jti.v16i1.1520.
- [2] R. K. Diyan Zahra, Ambia Nurdin, Ully Fitria, Kiki Asrifa Dinen, "Pemanfaatan Teknologi dalam Bidang Kesehatan," *Promot. J. Kesehat. Masy.*, vol. 8, no. 1, p. 97, 2018.

- [3] Siti Noor Chotimah, "Implementasi Sistem Informasi Kesehatan di Fasilitas Pelayanan Kesehatan Indonesia : LiteratureReview," *J. Rekam Medis dan Manaj. Inf. Kesehat.*, vol. 2, no. 1, pp. 8–13, 2021.
- [4] N. F. Purba, F. S. Annisa, A. Syafitri, and S. H. Purba, "Jurnal Kesehatan Unggul Gemilang," *J. Kesehat. Unggul Gemilang*, vol. 8, no. 1, pp. 7–15, 2024.
- [5] M. E. dr. Siti Nadia Tarmizi, "Pemanfaatan Teknologi sebagai Penopang Transformasi Sistem Informasi Kesehatan," Kemenkes. [Online]. Available: <https://kemkes.go.id/id/pemanfaatan-teknologi-sebagai-penopang-transformasi-sistem-informasi-kesehatan>
- [6] Hilda B Alexander, "Angka Kematian Ibu Masih Tinggi, Upaya Komprehensif Diperlukan," Kompas.com. [Online]. Available: <https://lestari.kompas.com/read/2023/12/28/110000286/angka-kematian-ibu-masih-tinggi-upaya-komprehensif-diperlukan>
- [7] A. Nugraha, "Pemerintah Jawa Barat Klaim Jumlah Kematian Bayi Menurun Drastis," Liputan6. [Online]. Available: <https://www.liputan6.com/health/read/5198260/pemerintah-jawa-barat-klaim-jumlah-kematian-bayi-menurun-drastis>
- [8] B. R. Anassai and P. Josaphat, "Pembangunan Chatbot Sistem Informasi KBLI dan KBJI Berbasis LLM (Development of LLM-Based KBLI and KBJI Information System Chatbot)," *Semin. Nas. Off. Stat.* 2024, pp. 723–734, 2024.
- [9] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Adv. Neural Inf. Process. Syst.*, vol. 2020-December, 2020.
- [10] M. M. Abootorabi *et al.*, "Ask in Any Modality: A Comprehensive Survey on Multimodal Retrieval-Augmented Generation," pp. 16776–16809, 2025, doi: 10.18653/v1/2025.findings-acl.861.
- [11] Nur Rohim and E. Zuliarso, "Penerapan Algoritma Deep Learning Untuk Pengembangan Chatbot Yang Digunakan Untuk Konsultasi Dan Pengenalan Tentang Virus Covid-19," *Pixel J. Ilm. Komput. Graf.*, vol. 15, no. 2, pp. 267–278, 2022, doi: 10.51903/pixel.v15i2.777.
- [12] N. A. Purwitasari and M. Soleh, "Implementasi Algoritma Artificial Neural Network Dalam Pembuatan Chatbot Menggunakan Pendekatan Natural Language Parocessing," *J. IPTEK*, vol. 6, no. 1, pp. 14–21, 2022, doi: 10.31543/jii.v6i1.192.
- [13] B. A. Haryono and A. Prasetya, "Identifikasi Kuantitas Berbasis Rule Pada Masalah Text-to-SQL," vol. 5, no. 1, pp. 26–39, 2025.
- [14] S. Gufranaka, A. Turmudi Zy, and Ermanto, "Implementasi Retrieval Augmented Generation (RAG) Dalam Perancangan Chatbot Kesehatan Pencernaan," *JSAI (Journal Sci. Appl. Informatics)*, vol. 8, no. 1, pp. 181–188, 2025, doi: 10.36085/jsai.v8i1.7678.
- [15] A. Kachhwal, A. S. Gupta, A. Jangid, and A. Sharma, "A Multi-Modal Hybrid Search System using RAG and Text2SQL for Semantic Information Retrieval," 2025.