

Chatbot Generative AI Berbasis Hybrid RAG dengan Integrasi RIASEC untuk Rekomendasi Jurusan

Dwi Mahdini¹, Nazruddin Safaat H², Yelfi Vitriani³, Muhammad Irsyad⁴

^{1,2,3,4} Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim

¹12250111298@students.uin-suska.ac.id, ²nazruddin.safaat@uin-suska.ac.id*, ³yelfi.vitriani@uin-suska.ac.id,

⁴irsyadtech@uin-suska.ac.id

Abstract

The mismatch between students' interests and chosen academic majors remains a significant issue that can affect academic performance and career development. In addition, the limited availability of career counseling services makes it difficult for students to obtain personalized guidance. This study aims to design and implement a major recommendation chatbot based on a Hybrid Retrieval Augmented Generation (Hybrid RAG) architecture integrated with the RIASEC personality model. The proposed sistem employs a Query Router to separate structured and unstructured data processing into two pathways: a relational pathway using MySQL for RIASEC assessment processing and a semantic pathway using FAISS for knowledge retrieval. The sistem was developed using a Large Language Model combined with retrieval mechanisms to generate contextual and personalized responses. Sistem evaluation was conducted through Black Box Testing, BERTScore, and RAGAS. Experimental results show that the proposed Hybrid RAG architecture achieved an average Precision of 0.769, Recall of 0.778, and F1-Score of 0.774, outperforming the standard RAG architecture with an F1-Score of 0.728. RAGAS evaluation also demonstrated high retrieval and faithfulness performance. These results indicate that integrating Hybrid RAG with RIASEC can improve the quality of personalized major recommendations and support academic consultation services.

Keywords: chatbot, hybrid rag, riasec, rekomendasi jurusan, large language model

Abstrak

Ketidaksesuaian antara minat siswa dan jurusan yang dipilih masih menjadi permasalahan yang dapat memengaruhi keberhasilan akademik serta perencanaan karier. Di sisi lain, keterbatasan layanan bimbingan karier menyebabkan tidak semua siswa memperoleh konsultasi yang sesuai dengan kebutuhan mereka. Penelitian ini bertujuan merancang dan mengimplementasikan chatbot rekomendasi jurusan berbasis Hybrid Retrieval Augmented Generation (Hybrid RAG) yang terintegrasi dengan model kepribadian RIASEC. Arsitektur yang diusulkan menggunakan Query Router untuk memisahkan pemrosesan data terstruktur dan tidak terstruktur ke dalam dua jalur, yaitu jalur relasional berbasis MySQL untuk pengolahan hasil asesmen RIASEC dan jalur semantik berbasis FAISS untuk pencarian informasi dari knowledge base. Sistem dikembangkan dengan memanfaatkan Large Language Model yang dipadukan dengan mekanisme retrieval untuk menghasilkan respons yang kontekstual dan personal. Evaluasi sistem dilakukan menggunakan Black-Box Testing, BERTScore, dan RAGAS. Hasil pengujian menunjukkan bahwa arsitektur Hybrid RAG memperoleh nilai Precision rata-rata sebesar 0,769, Recall sebesar 0,778, dan F1-Score sebesar 0,774, lebih tinggi dibandingkan arsitektur RAG standar yang memperoleh F1-Score sebesar 0,728. Evaluasi menggunakan RAGAS juga menunjukkan kualitas retrieval dan faithfulness yang tinggi. Hasil penelitian menunjukkan bahwa integrasi Hybrid RAG dan RIASEC mampu meningkatkan kualitas rekomendasi jurusan yang lebih personal serta mendukung layanan konsultasi akademik berbasis kecerdasan buatan.

Kata kunci: chatbot, hybrid rag, riasec, rekomendasi jurusan, large language model.

1. Pendahuluan

Ketidakcocokan pemilihan jurusan di perguruan tinggi masih menjadi persoalan serius yang memicu rendahnya motivasi akademik dan pemborosan biaya pendidikan [1]. Kondisi ini tidak hanya berdampak pada stres akademik dan rendahnya motivasi belajar, tetapi juga meningkatkan risiko drop out serta ketidaksesuaian kompetensi lulusan dengan kebutuhan dunia kerja. Di sisi lain, layanan bimbingan karier di sekolah menghadapi keterbatasan sumber daya manusia yang sangat signifikan, di mana rasio aktual Guru BK secara nasional mencapai 1:570 [2]. Kesenjangan ini menyebabkan layanan asesmen karier berbasis instrumen psikometri seperti pendekatan RIASEC sulit diterapkan secara mendalam dan personal, sehingga pemetaan minat dan kecocokan

jurusan siswa tidak dilakukan secara sistematis dan berbasis data.

Salah satu instrumen psikometri yang terbukti efektif untuk konteks bimbingan karier adalah model RIASEC yang dikembangkan oleh John L. Holland. Model ini mengklasifikasikan kepribadian individu ke dalam enam tipe, yaitu Realistic, Investigative, Artistic, Social, Enterprising, dan Conventional, di mana kombinasi tipe dominan seseorang dapat digunakan untuk memprediksi kesesuaiannya dengan bidang studi maupun lingkungan kerja tertentu [3]. Keunggulan RIASEC dibandingkan instrumen lain terletak pada sifatnya yang dirancang sebagai sistem klasifikasi kepribadian yang dapat dipetakan langsung ke rekomendasi program studi, bukan sekadar deskripsi

kepribadian umum. Oleh karena itu, penelitian ini mengadaptasi instrumen RIASEC ke dalam versi singkat 18 pertanyaan berdasarkan Martins et al. [4] agar proses asesmen dapat diselesaikan secara efisien dalam antarmuka chatbot, sekaligus menghasilkan profil kepribadian yang dapat diproses secara deterministik oleh sistem.

Keterbatasan tersebut mendorong siswa mengambil keputusan pemilihan jurusan secara mandiri tanpa pendampingan berbasis metode ilmiah. Perkembangan teknologi chatbot dipandang sebagai alternatif solusi untuk mendukung layanan bimbingan karir secara lebih luas dan skalabel [5],[6]. Namun, chatbot konvensional berbasis rule-based memiliki keterbatasan dalam menangani variasi bahasa dan penalaran kompleks. Kemunculan Large Language Model (LLM) seperti GPT dan LLaMA membuka peluang menghadirkan interaksi yang lebih natural dalam eksplorasi minat karir [7]. Meskipun demikian, LLM generatif masih memiliki kelemahan dalam menjaga konsistensi respons ketika berhadapan dengan data faktual yang sangat spesifik [8].

Pendekatan Retrieval Augmented Generation (RAG) menjadi solusi penting untuk meningkatkan akurasi informasi yang digunakan LLM [9],[10]. Berbagai penelitian terkini menunjukkan bahwa RAG telah banyak diterapkan untuk meningkatkan akurasi dan relevansi respons chatbot pada konteks akademik, pendidikan, dan layanan publik. Implementasi RAG pada evaluasi dokumen akademik, sistem helpdesk, dan layanan informasi terbukti mampu mengurangi kesalahan informasi serta meminimalkan halusinasi model bahasa [11],[12]. Integrasi vector database dalam arsitektur RAG dilaporkan mampu meningkatkan kualitas pencarian semantik pada layanan publik dan akademik [13]. Pada konteks pemrosesan bahasa alami, RAG juga terbukti meningkatkan performa LLM dalam memahami Bahasa Indonesia secara lebih kontekstual [14].

Penelitian lain menunjukkan bahwa chatbot berbasis RAG mampu menyajikan informasi yang lebih konsisten, akurat, dan relevan bagi pengguna [9]. Studi tinjauan sistematis menegaskan bahwa RAG efektif meningkatkan keandalan LLM pada domain spesifik, terutama ketika sistem dihadapkan pada kebutuhan data yang bersifat faktual dan terstruktur [10]. Pendekatan context-aware RAG juga dilaporkan mampu meningkatkan kualitas personalisasi interaksi pengguna dengan memanfaatkan konteks tambahan dalam proses retrieval [15].

Meskipun demikian, sebagian besar penelitian tersebut masih mengandalkan RAG tunggal berbasis dense retrieval tanpa pemisahan jalur pengambilan data antara data terstruktur dan tidak terstruktur [16],[17].

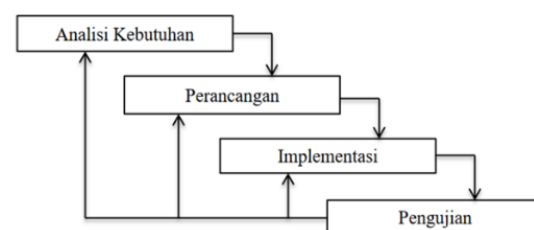
Hanya sedikit penelitian yang mengusulkan pendekatan Hybrid RAG, dan umumnya masih berfokus pada peningkatan performa Question Answering berbasis dokumen tanpa mekanisme intent routing yang eksplisit [18]. Selain itu, konteks personal pengguna dalam sistem RAG umumnya terbatas pada riwayat percakapan, belum terintegrasi dengan hasil asesmen psikometrik yang bersifat deterministik [5]. Oleh karena itu, penelitian ini mengusulkan arsitektur Hybrid RAG dengan Query Router, di mana hasil pengukuran RIASEC digunakan sebagai parameter personalisasi sekaligus batasan teori untuk menghasilkan rekomendasi jurusan yang lebih akurat, konsisten, dan personal.

Penelitian ini bertujuan untuk merancang arsitektur Hybrid RAG dengan Query Router pada chatbot yang mampu mengintegrasikan hasil pengukuran psikometri RIASEC (data terstruktur) dengan dokumen pedoman akademik (data tidak terstruktur), guna menghasilkan sistem rekomendasi pemilihan jurusan yang lebih personal, akurat, dan meminimalisir halusinasi dibandingkan pendekatan RAG standar.

2. Metode Penelitian

2.1. Rancangan Penelitian

Metode penelitian dalam pengembangan sistem hybrid RAG yang digunakan adalah Waterfall, dipilih karena memiliki struktur terurut dan mendukung dokumentasi yang kuat pada setiap tahapannya. Tahapan utama meliputi: (1) Analisis Kebutuhan, mengidentifikasi kebutuhan pengguna dan fungsionalitas sistem; (2) Perancangan Sistem, menyusun arsitektur sistem, desain alur interaksi chatbot, dan rancangan knowledge base; (3) Implementasi, membangun sistem menggunakan Python, PHP, FAISS, dan LangChain dan tentunya integrasi RIASEC sendiri serta (4) Pengujian, menggunakan Black box Testing, BERTScore dan RAGAS untuk memastikan sistem berfungsi dengan baik.



Gambar 1. Metode Waterfall

2.2. Analisis Kebutuhan

Analisis kebutuhan pada penelitian ini dilakukan melalui identifikasi mendalam terhadap seluruh instrumen, data, dan infrastruktur yang diperlukan untuk membangun arsitektur Hybrid RAG. Analisis ini dipetakan ke dalam dua jalur pemrosesan utama, yaitu

kebutuhan data untuk jalur relasional deterministik dan kebutuhan data untuk jalur semantik.

Untuk kebutuhan pada jalur relasional deterministik, sistem membutuhkan instrumen pengukur kepribadian yang efisien. Oleh karena itu, dilakukan adaptasi dan validasi terhadap instrumen RIASEC menjadi versi singkat yang terdiri dari 18 pertanyaan [4]. Instrumen ini dirancang khusus untuk dapat diselesaikan pengguna dalam waktu singkat tanpa mengurangi tingkat validitas pengukuran. Rincian instrumen yang digunakan pada antarmuka sistem ini dapat dilihat pada tabel 1 berikut:

Tabel 1. Instrumen RIASEC

No	Pertanyaan	Dimensi
1.	Mengoperasikan mesin untuk memproduksi suku cadang	R
2.	Melakukan perawatan pada mesin dan perkakas	R
3.	Merencanakan sistem kelistrikan sebuah rumah	R
4.	Melakukan analisis dan eksperimen laboratorium	I
5.	Menjelaskan fenomena alam dan fisik	I
6.	Membaca buku dan artikel ilmiah	I
7.	Bernyanyi dalam paduan suara	A
8.	Menampilkan pertunjukan seni di depan penonton	A
9.	Berpartisipasi dalam merancang skenario untuk pertunjukan teater	A
10.	Bersedia membantu orang lain	S
11.	Memberikan bimbingan kepada individu, kelompok, atau masyarakat mengenai kesehatan dan kesejahteraan	S
12.	Memberikan layanan sosial di komunitas dan lingkungan sekitar	S
13.	Bernegosiasi dengan pelanggan	E
14.	Memasarkan produk dan layanan	E
15.	Meyakinkan orang untuk membeli suatu produk	E
16.	Mengarsipkan dokumen dan berkas penting	C
17.	Melakukan pekerjaan pembukuan/akuntansi untuk sebuah perusahaan	C
18.	Memasukkan informasi ke dalam basis data	C

Setelah data kepribadian dikumpulkan dari instrumen pada Tabel 1, diperlukan aturan pemetaan baku untuk menerjemahkan skor dominan RIASEC pengguna menjadi rekomendasi program studi. Kebutuhan ini dipenuhi dengan mengadopsi matriks pemetaan secara absolut dari literatur rujukan akademis yang relevan dan telah teruji secara psikometrik. Matriks pemetaan yang diadopsi mengacu pada penelitian Kharina dan Yulianto [19], yang menyusun korespondensi kode Holland berdasarkan Dictionary of Holland's Occupations Codes (ONET) dan disesuaikan melalui validasi bersama kaprodi. Tabel tersebut kemudian diadaptasi khusus untuk konteks UIN Suska Riau dengan menyesuaikan program studi yang tersedia. Hasil adaptasi matriks pemetaan ini disajikan pada tabel berikut:

Tabel 2. Jurusan RIASEC

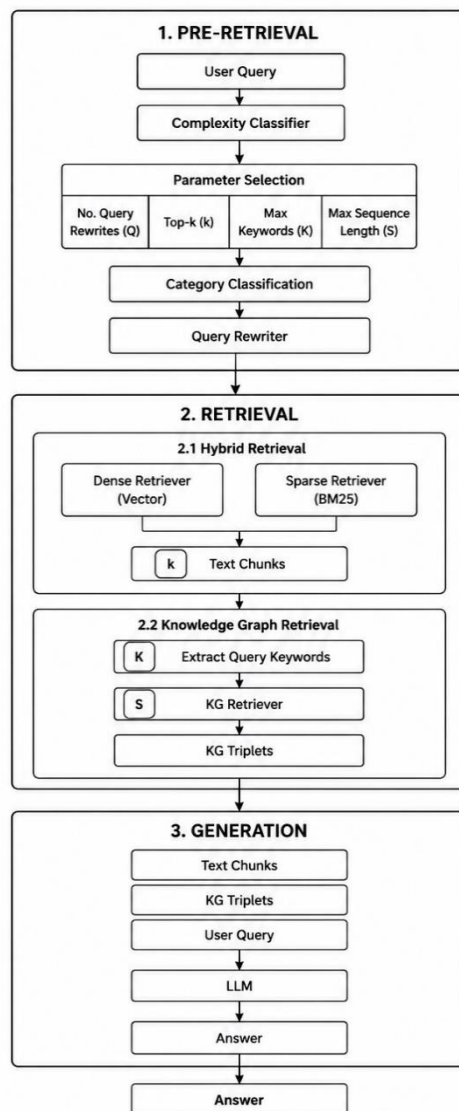
Jurusan	Dimensi
Informatika, Sistem Informasi	RIE
Manajemen, Ilmu Komunikasi, Psikologi, Akuntansi	RAI
Psikologi, Ilmu Komunikasi	IAE
Ilmu Komunikasi	SRI
Psikologi, Manajemen	SCR
Manajemen	ERI
Sistem Informasi	ERS
Akuntansi	CIS
Sistem Informasi	CEA

Sementara itu, untuk kebutuhan pada jalur semantik, sistem membutuhkan himpunan pengetahuan spesifik terkait institusi akademik, layanan karir dan dokumen terkait. Kebutuhan data tidak terstruktur ini dikumpulkan dalam bentuk dokumen digital yang mencakup Buku Panduan Akademik, Pedoman Penasehat Akademik, pedoman RIASEC, dan dokumen pendukung lainnya. Dokumen-dokumen ini mutlak diperlukan untuk diubah menjadi representasi vektor dalam pangkalan data FAISS, sehingga Large Language Model dapat melakukan pencarian konteks yang presisi sebelum menyusun kalimat respons.

2.3. Perancangan

Arsitektur Hybrid RAG yang diusulkan pada penelitian ini dikembangkan dengan mengacu pada konsep Hybrid Parameter Adaptive Retrieval Augmented Generation (HyPA-RAG) yang diperkenalkan oleh Kalra [20]. Namun, arsitektur yang diimplementasikan telah disesuaikan dengan kebutuhan sistem rekomendasi jurusan berbasis hasil tes RIASEC melalui integrasi basis data relasional, mekanisme klasifikasi intent, pencarian semantik menggunakan FAISS, serta generasi respons menggunakan model LLaMA.

Pendekatan arsitektur berlapis diadopsi untuk perancangan dengan memisahkan fokus pemrosesan secara modular, sehingga setiap tahapan mulai dari penerimaan interaksi pengguna hingga generasi jawaban oleh model bahasa dapat beroperasi secara independen namun tetap terintegrasi erat [20]. Pemisahan layer ini sangat krusial dalam mekanisme Hybrid RAG, mengingat sistem harus mengorkestrasi dua jenis aliran data yang berbeda secara bersamaan yaitu penarikan data terstruktur dari basis data relasional untuk mempertahankan determinisme profil pengguna, serta ekstraksi dokumen tak terstruktur menggunakan basis data vektor untuk pencarian konteks semantik. Dengan menstrukturkan aliran kerja ke dalam lima lapisan yang berurutan berupa layer, rincian fungsionalitas dari masing-masing layer tersebut dijelaskan sebagai berikut



Gambar 2. Arsitektur Hybrid RAG

1. Layer Frontend, lapisan ini merupakan titik interaksi utama sistem dengan pengguna. Ketika pengguna mengetikkan pertanyaan atau keluhan terkait pemilihan jurusan, antarmuka web menangkap masukan tersebut dan mengirimkannya dalam bentuk permintaan (request) ke Application Programming Interface (API) yang tersedia untuk diproses.
2. Layer Data Relasional, Sebelum pertanyaan diproses secara semantik, sistem memicu lapisan relasional untuk melakukan ekstraksi data personal. Lapisan ini mengambil profil pengguna beserta skor dan rekomendasi jurusan hasil tes RIASEC yang telah disimpan secara deterministik di dalam pangkalan data MySQL. Data personal ini memegang peranan penting karena akan dikirimkan langsung ke Prompt Builder sebagai batasan teori yang mengunci rekomendasi kecerdasan buatan.

3. Layer Klasifikasi Niat, Inti efisiensi dari arsitektur Hybrid RAG terletak pada mekanisme Query Router. Pada lapisan ini, pertanyaan pengguna diklasifikasikan berdasarkan intent (niat). Jika sistem mendeteksi intent berupa sapaan (greeting) atau pertanyaan profil (personal profile), aliran data akan melewati pencarian dokumen (bypass retrieval) untuk meminimalkan beban komputasi. Sebaliknya, jika intent terdeteksi sebagai pencarian informasi akademik (academic info) atau konsultasi aliran data akan menuju ke dokumen.
4. Layer Pencarian Semantik, bagi pertanyaan yang lolos dari router akademik, kueri pengguna akan diproses melalui metode Similarity Search di dalam pangkalan data vektor FAISS. Lapisan ini bertugas melakukan ekstraksi teks dokumen dari himpunan Knowledge Base yang memiliki tingkat kemiripan vektor tertinggi dengan kueri pengguna.
5. Layer Generasi, merupakan tahap akhir dimana mekanisme Prompt Builder menyatukan Data Personal dari MySQL dan Ekstraksi Teks Dokumen dari FAISS menjadi sebuah Sistem Prompt terpadu. Konteks gabungan ini kemudian diproses oleh model bahasa LLaMA melalui antarmuka Groq API. Model tersebut bertugas menggenerasi jawaban personal yang empatik dan halus, namun tetap terkunci ketat pada fakta kurikulum dan batasan rekomendasi psikometri, sebelum respons akhir dikembalikan ke layar pengguna.

2.4. Pengujian

Fase pengujian sistem dieksekusi melalui pendekatan komprehensif yang dibagi ke dalam tiga tahap evaluasi guna memvalidasi keandalan jalur relasional dan akurasi komputasi semantik pada arsitektur Hybrid RAG:

1. Pengujian Fungsional Jalur Relasional, pengujian ini dieksekusi tanpa melibatkan komponen kecerdasan buatan, dengan fokus utama pada pengujian Black Box Testing. Tujuannya adalah memastikan bahwa antarmuka sistem berhasil merekam 18 jawaban instrumen RIASEC dari pengguna tanpa kehilangan data (data loss). Pengujian dilanjutkan dengan memvalidasi akurasi logika sistem dalam mengakumulasi dan menetapkan skor dominan tersebut ke dalam rekomendasi program studi secara otomatis merujuk pada aturan psikometrik. Pengujian ini krusial untuk menjamin bahwa data profil yang akan disuntikkan sebagai konteks kepada kecerdasan buatan adalah data yang sepenuhnya valid dan bebas dari kesalahan kalkulasi.
2. Evaluasi Semantik Menggunakan BERTScore, tahap ini bertujuan untuk mengukur sejauh mana arsitektur usulan mampu menekan halusinasi model bahasa. Pengujian dilakukan dengan merumuskan

pertanyaan uji representatif yang diekstraksi secara ketat dari dokumen pangkalan pengetahuan menggunakan bantuan Google NotebookLLM. Untuk membuktikan keunggulan sistem yang diusulkan, dilakukan pengujian komparatif antara arsitektur RAG standar yang hanya melakukan pencarian semantik tunggal tanpa profil pengguna dengan arsitektur Hybrid RAG. Kualitas respons dievaluasi secara matematis menggunakan metrik BERTScore. Penggunaan BERTScore dipilih karena metrik ini menghitung kesamaan teks berdasarkan kemiripan representasi vektor dari model bahasa pra-latih [21]. Evaluasi ini terdiri dari tiga persamaan diantaranya :

- 1) Precision, adalah tingkat kesesuaian prediksi (xi) jawaban model dengan data referensi (yj).

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

- 2) Recall, adalah efektif model memprediksi kelas positif secara akurat ditentukan oleh recall.

$$R_{BERT} = \frac{1}{|x|} \sum_{x_i \in x} \max_{\hat{x}_j \in \hat{x}} \text{sim}(x_i, \hat{x}_j) \quad (2)$$

- 3) F-1 Score, F1-Score merupakan nilai rata-rata yang membandingkan precision dan recall.

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

3. Evaluasi Kualitas Retrieval dan Generasi Menggunakan RAGAS Untuk mengukur seberapa efisien pangkalan data vektor dalam menarik dokumen yang relevan, serta seberapa faktual jawaban yang dihasilkan oleh model bahasa terhadap dokumen tersebut, sistem dievaluasi menggunakan framework Retrieval-Augmented Generation Assessment (RAGAS). Metrik yang dinilai berfokus pada empat dimensi utama, yaitu Context Precision, Context Recall, Faithfulness, dan Answer Relevancy yang telah diterapkan pada sistem chatbot layanan informasi kampus [22].

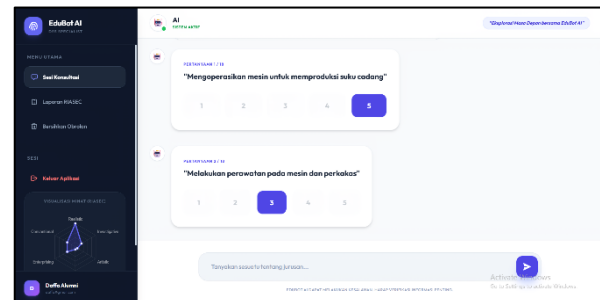
3. Hasil dan Pembahasan

Sistem diimplementasikan pada dua lingkungan terpisah yang berjalan secara paralel, yaitu antarmuka pengguna berbasis PHP (Laragon) dan backend AI berbasis Python (FastAPI). Pemisahan arsitektur ini diterapkan agar beban komputasi kecerdasan buatan tidak membebani kerja server lokal. Untuk pemrosesan bahasa alami, sistem terintegrasi dengan layanan cloud melalui API Groq, menggunakan model LLaMA 3.3 70B Versatile sebagai pemroses utama logika Hybrid

RAG. Guna mengantisipasi kegagalan respons atau batasan rate limit pada server API eksternal, sistem dilengkapi dengan mekanisme fallback otomatis. Mekanisme ini secara instan mengalihkan komputasi ke model cadangan yang lebih ringan, yakni LLaMA 3.1 8B Instant (serta Mixtral 8x7B), untuk menjamin ketersediaan layanan yang berkelanjutan

3.1. Implementasi Antarmuka Kuesioner RIASEC

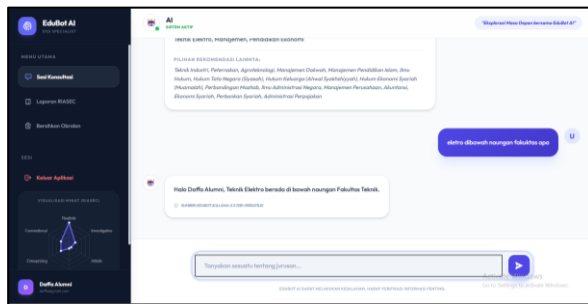
Implementasi Antarmuka Kuesioner yaitu sebelum pengguna dapat mengakses layanan konsultasi dengan kecerdasan buatan, sistem mewajibkan pengguna baru untuk menyelesaikan asesmen kepribadian karier berdasarkan instrumen RIASEC. Antarmuka kuesioner ini menampilkan delapan belas pertanyaan secara berurutan menggunakan skala pengukuran satu hingga lima untuk mengevaluasi tingkat preferensi pengguna terhadap aktivitas tertentu seperti Realistic Investigative, Artistic, Social, Enterprising, dan Conventional. Hasil komputasi dari kuesioner ini akan mengidentifikasi kombinasi tipe kepribadian dominan pengguna yang kemudian langsung disimpan ke dalam tabel basis data relasional. Ketersediaan data profil yang deterministik ini menjadi syarat mutlak dan landasan personalisasi dalam tahapan pencarian informasi pada saat proses konsultasi berlangsung.



Gambar 3. Antarmuka Chat Kusioner

3.2. Implementasi Antarmuka Sesi Konsultasi Chat

Implementasi Antarmuka Sesi Konsultasi Chat Pusat dari fungsionalitas sistem terletak pada antarmuka percakapan yang didesain dengan mengutamakan kenyamanan pengguna. Desain antarmuka dibuat menyerupai aplikasi pertukaran pesan instan komersial untuk memberikan pengalaman yang intuitif. Melalui antarmuka ini pengguna dapat mengajukan pertanyaan terkait informasi akademik maupun kebingungan mengenai pemilihan jurusan. Kueri teks dikirim secara bertahap menuju titik akhir menggunakan metode komunikasi data pada latar belakang. Jawaban yang dihasilkan oleh model bahasa kemudian ditampilkan didalam sesi konsultasi chat.



Gambar 4. Antarmuka Chat Konsultasi

3.3. Implementasi Hybrid RAG pada Backend

Pada sisi backend, kueri dari pengguna tidak langsung dijawab secara mentah oleh LLM. Sistem menerapkan mekanisme Query Router yang secara cerdas mengklasifikasikan intensi pengguna ke dalam empat jalur utama:

- Jalur greeting, menangani kueri sapaan umum dengan mengarahkannya langsung ke model bahasa tanpa proses pencarian guna memberikan respons yang cepat dan ramah.
- Jalur personal profile, menangani kueri terkait identitas pengguna dengan mengekstraksi data profil secara spesifik langsung dari basis data relasional.
- Jalur academic info, menangani kueri informatif umum dengan meluncurkan algoritma pencarian semantik terhadap pangkalan data vektor yang berisi dokumen pedoman kampus.
- Jalur consultation, merupakan jalur hibrida di mana kueri kompleks mengenai pemilihan jurusan diselesaikan dengan mengambil profil pengguna dari basis data relasional secara simultan bersama dari pangkalan data vektor. Jalur ini juga dilengkapi dengan mekanisme prioritas semantik di dalam instruksi awal model bahasa untuk mencegah generalisasi jawaban sehingga rekomendasi yang diberikan bisa secara akurat menargetkan program studi yang paling identik dengan cita cita pengguna.

3.4 Hasil Pengujian Fungsi Jalur Relasional

Pengujian fungsional pada jalur relasional dievaluasi menggunakan metode Black Box Testing. Pengujian ini bertujuan untuk mengukur tingkat keberhasilan antarmuka sistem dalam menangkap masukan pengguna, serta memvalidasi keakuratan logika backend dalam memproses instrumen RIASEC hingga menghasilkan rekomendasi program studi secara deterministik di database MySQL.

Tabel 3. Hasil Pengujian Jalur Relasional

Skenario Pengujian	Hasil yang Diharapkan	Status
Validasi Input: Pengisian 18 instrumen secara tidak lengkap.	Sistem menolak submit dan	Sukses

Penyimpanan Data: Pengiriman 18 jawaban instrumen yang valid.	memunculkan peringatan.	Sukses
Kalkulasi dan rekomendasi: Pemrosesan skor akumulasi RIASEC.	Seluruh data tersimpan utuh di MySQL tanpa data loss.	Sukses
	Sistem berhasil memetakan 3 skor dominan ke rekomendasi jurusan secara akurat.	Sukses

Berdasarkan Tabel 3, fungsionalitas jalur relasional berjalan dengan persentase keberhasilan 100%. Tidak ditemukan adanya data loss maupun galat kalkulasi. Hal ini memastikan bahwa data profil kepribadian pengguna yang akan disuntikkan kepada LLM terjamin validitasnya.

3.5. Hasil Pengujian Sistem BERTScore

Pengujian kualitas kecerdasan buatan pada sistem final dievaluasi menggunakan metrik BERTScore guna mengukur secara matematis sejauh mana kemampuan model bahasa dalam mengambil dan merangkai informasi secara faktual dari basis data. Pengujian ini difokuskan pada kueri yang membutuhkan penarikan informasi (RAG), dengan menggunakan sejumlah pertanyaan representatif yang diciptakan melalui Google NotebookLLM untuk mewakili tiga kelas rute kueri utama, yaitu Personal Profile, Academic Info, dan Consultation. Kelas Greeting tidak dimasukkan dalam evaluasi ini karena sifatnya yang berupa basa-basi percakapan dan tidak bergantung pada akurasi penarikan fakta dari basis pengetahuan.

Untuk membuktikan efektivitas pemisahan pemrosesan relasional dan semantik pada arsitektur usulan, dilakukan pengujian secara bertahap antara arsitektur RAG Standar (berbasis pencarian semantik murni) dengan arsitektur Hybrid RAG. Tahap pertama adalah mengukur performa dasar sistem menggunakan arsitektur RAG Standar tanpa adanya intervensi data profil pengguna maupun Query Router. Hasil pengujian ini dapat dilihat pada Tabel 4 berikut:

Tabel 4. Hasil Pengujian BERTScore pada RAG standar

Jalur	Precision	Recall	F1 Score
Personal Info	0,640	0,687	0,712
Academic info	0,741	0,725	0,732
Consultation	0,742	0,737	0,739
Rata Rata	0,707	0,716	0,728

Berdasarkan tabel 4, arsitektur RAG Standar mampu mampu memberikan tingkat akurasi dasar dengan rata-rata F1-Score sebesar 0,728. Namun, sistem ini menunjukkan kelemahan pada jalur personalisasi, terutama pada jalur personal info yang hanya mencapai FI-Score sebesar 0,712. Kelemahan ini bersumber dari ketidakmampuan RAG standar untuk mengaitkan konteks profil kepribadian pengguna dengan jawaban

yang dihasilkan. Karena sistem hanya mengandalkan pencarian semantik dari dokumen tanpa menyertakan data deterministik RIASEC, model bahasa cenderung merespons secara generik dan tidak mampu menghasilkan rekomendasi yang spesifik terhadap karakteristik individu pengguna. Kondisi ini menunjukkan bahwa sistem yang hanya mengandalkan pencarian semantik tanpa konteks personal yang terstruktur tidak mampu menghasilkan rekomendasi yang spesifik terhadap karakteristik individu pengguna.

Tahap kedua adalah menguji arsitektur *Hybrid RAG* yang diusulkan. Hasil evaluasi terhadap sistem hibrida ini dirangkum pada Tabel 5.

Tabel 5. Hasil Pengujian BERTScore pada Hybrid RAG

Jalur	Precision	Recall	F1 Score
Personal Info	0,745	0,751	0,748
Academic info	0,769	0,778	0,773
Consultation	0,795	0,805	0,800
Rata Rata	0,769	0,778	0,774

Hasil pada Tabel 5 menunjukkan peningkatan yang signifikan dan konsisten di seluruh kelas kueri faktual dengan nilai rata-rata F1-Score meningkat menjadi 0,774. Peningkatan ini tidak terjadi secara acak, melainkan dapat dijelaskan melalui dua mekanisme utama arsitektur yang diusulkan.

Mekanisme pertama adalah kontribusi Query Router dalam meningkatkan akurasi jawaban. Query Router bertindak sebagai komponen klasifikasi intent yang secara dinamis memilih jalur pemrosesan yang paling sesuai untuk setiap kueri yang masuk. Pada jalur Consultation yang mencatatkan peningkatan F1-Score tertinggi sebesar 0,061 dari angka 0,739 menjadi 0,800, Query Router berhasil memadukan secara simultan data profil RIASEC dari MySQL dengan konteks dokumen dari FAISS sebelum keduanya dikonsolidasikan dalam satu prompt yang dikirimkan ke model bahasa. Tanpa mekanisme routing ini, model bahasa tidak memiliki batas pengetahuan deterministik dan akan cenderung menghasilkan rekomendasi yang bersifat umum. Query Router menjamin bahwa setiap jawaban konsultasi selalu dibatasi oleh profil kepribadian pengguna yang terukur secara psikometrik, sehingga output model bahasa menjadi lebih terfokus, personal, dan relevan terhadap kondisi nyata pengguna.

Mekanisme kedua adalah integrasi data terstruktur dari jalur relasional yang memberikan konteks deterministik yang tidak dapat diperoleh dari pencarian semantik semata. Pada jalur Personal Info, F1-Score meningkat dari 0,712 menjadi 0,748 dengan selisih peningkatan sebesar 0,036. Peningkatan ini terjadi karena sistem tidak lagi bergantung pada inferensi dokumen untuk memperoleh informasi profil pengguna, melainkan mengambilnya secara langsung dari basis data relasional yang terstruktur dan terjamin

validitasnya. Akibatnya, nilai Recall keseluruhan meningkat dari 0,716 menjadi 0,778, yang berarti model berhasil menyertakan fakta-fakta yang lebih esensial dan relevan di dalam jawabannya.

Tabel 6. Perbandingan Hasil Pengujian RAG Standar dan Hybrid

Jalur	F-1 Score Standar	F-1 Score Hybrid	Peningkatan
Personal Info	0,712	0,748	+ 0,036
Academic info	0,732	0,773	+ 0,041
Consultation	0,739	0,800	+ 0,061
Rata Rata	0,728	0,774	+ 0,046

Untuk memperjelas kontribusi ilmiah penelitian ini, perlu dibandingkan tiga pendekatan utama pengembangan chatbot. Chatbot berbasis LLM murni tanpa retrieval sangat rentan halusinasi karena tidak dapat merujuk dokumen institusi yang spesifik. RAG Standar memperbaiki hal tersebut namun tetap tidak mampu mengintegrasikan data personal pengguna yang terstruktur, sehingga rekomendasi yang dihasilkan masih bersifat umum sebagaimana ditunjukkan oleh F1-Score sebesar 0,728. Hybrid RAG yang diusulkan mengatasi kedua kelemahan tersebut dengan memadukan jalur relasional deterministik dan jalur semantik melalui Query Router, menghasilkan peningkatan nilai F1-Score sebesar 0,774.

Hasil ini konsisten dengan penelitian-penelitian terdahulu yang menggunakan pendekatan RAG. Rachman et al. melaporkan bahwa penerapan RAG pada chatbot layanan pelanggan mampu meningkatkan konsistensi dan akurasi informasi yang dihasilkan, namun tidak membahas mekanisme pemisahan data terstruktur dan tidak terstruktur [9]. Fanani dalam implementasi RAG untuk evaluasi proposal tugas akhir menyimpulkan bahwa integrasi basis pengetahuan dokumen secara signifikan mengurangi halusinasi, tetapi sistem yang diusulkan masih bersifat RAG tunggal tanpa routing [12]. Pratama dan Sisephaputra juga menunjukkan bahwa penggunaan vector database pada arsitektur RAG meningkatkan kualitas pencarian semantik pada layanan publik [13]. Dibandingkan dengan penelitian-penelitian tersebut, arsitektur Hybrid RAG pada penelitian ini menawarkan keunggulan tambahan berupa mekanisme Query Router yang memisahkan jalur pemrosesan secara eksplisit, sehingga menghasilkan peningkatan performa yang lebih terukur khususnya pada skenario konsultasi yang membutuhkan integrasi data personal dan pengetahuan dokumen secara bersamaan.

3.6. Hasil Pengujian RAGAS

Untuk mengukur keandalan dan kualitas respons dari arsitektur Hybrid RAG yang diusulkan, penelitian ini menggunakan kerangka pengujian otomatis Retrieval-Augmented Generation Assessment (RAGAS). Pengujian ini berfokus pada dua aspek utama, yaitu

kualitas pengambilan dokumen dan kualitas pembentukan jawaban.

Pengujian RAGAS difokuskan pada dua jalur yang secara langsung melibatkan proses retrieval dokumen, yaitu jalur Academic Info dengan lima pertanyaan dan jalur Consultation dengan empat pertanyaan. Jalur Greeting tidak diikutsertakan karena tidak melibatkan proses retrieval sama sekali, sedangkan jalur Personal Profile tidak diikutsertakan karena pengambilan datanya bersumber langsung dari basis data relasional MySQL tanpa melalui pencarian dokumen, sehingga tidak relevan dievaluasi menggunakan kerangka RAGAS yang dirancang untuk mengukur kualitas retrieval berbasis vektor. Jumlah sembilan pertanyaan dipilih dengan mempertimbangkan proporsi cakupan topik pada masing-masing jalur, di mana jalur Academic Info diwakili oleh lima pertanyaan yang mencakup variasi topik mulai dari konsep psikometrik RIASEC, informasi program studi lintas fakultas, hingga perbandingan jurusan, sedangkan jalur Consultation diwakili oleh empat pertanyaan yang menguji kemampuan sistem dalam mengintegrasikan profil pengguna dengan konteks dokumen untuk menghasilkan rekomendasi yang personal. Jumlah ini dinilai memadai untuk mengevaluasi performa sistem secara representatif mengingat setiap pertanyaan mencerminkan skenario penggunaan nyata yang berbeda dan diekstraksi langsung dari dokumen Knowledge Base menggunakan Google NotebookLLM, sehingga setiap pertanyaan memiliki jawaban yang dapat dilacak langsung pada sumber dokumen dan hasil evaluasi yang diperoleh bersifat objektif serta dapat dipertanggungjawabkan.

Tabel 7. Pertanyaan Uji RAGAS

No	Jalur	Pertanyaan
1	Academic Info	Apa pengertian dari tipe kepribadian Realistic dalam teori RIASEC Holland?
2	Academic Info	Apa saja program studi keguruan yang tersedia di Fakultas Tarbiyah dan Keguruan UIN Suska Riau?
3	Academic Info	Apa yang membedakan Jurusan Teknik Informatika dengan Sistem Informasi di Fakultas Sains dan Teknologi (Saintek)?
4	Academic Info	Apa saja prospek kerja lulusan Ilmu Komunikasi konsentrasi Jurnalistik dan Penyiaran (broadcasting) di Fakultas Dakwah dan Komunikasi?
5	Academic Info	Apa fokus keilmuan yang dipelajari pada Jurusan Administrasi Negara di Fakultas Ekonomi dan Ilmu Sosial (Fekonso)?
6	Consultation	Peluang kerja apa saja yang tersedia untuk lulusan Fakultas Psikologi (S.Psi) yang berkaitan dengan mental dan konseling kejiwaan?

7	Consultation	Apa profil lulusan dari Jurusan Perbandingan Mazhab di Fakultas Syariah dan Hukum (FSH)?
8	Consultation	Apa fokus pembelajaran di program studi Ilmu Hadis pada Fakultas Ushuluddin?
9	Consultation	Apa perbedaan fokus studi antara Jurusan Agroteknologi dan Jurusan Peternakan di Fakultas Pertanian dan Peternakan?

Penjabaran skor RAGAS untuk masing-masing pertanyaan yaitu:

Tabel 8. Hasil Pengujian RAGAS

No	F	CR	CP	AR
1	1,00	1,00	0,98	0,97
2	0,94	0,95	0,92	0,91
3	0,98	1,00	0,95	0,94
4	0,95	0,96	0,90	0,92
5	1,00	0,98	0,96	0,90
6	0,96	1,00	0,94	0,94
7	0,93	0,95	0,91	0,90
8	0,97	0,96	0,93	0,94
9	0,96	0,97	0,95	0,93
Avg	0,96	0,97	0,94	0,93

Evaluasi kinerja sistem Hybrid RAG diukur menggunakan framework RAGAS yang terdiri dari empat metrik utama. Metrik Faithfulness mengukur sejauh mana jawaban yang dihasilkan oleh model bahasa benar-benar bersumber dari konteks yang diberikan tanpa adanya informasi yang dikarang atau halusinasi. Metrik Context Recall mengevaluasi kemampuan sistem dalam menemukan seluruh informasi relevan dari basis data yang dibutuhkan untuk menjawab pertanyaan dengan tuntas. Context Precision digunakan untuk menilai tingkat presisi dokumen yang ditarik, di mana nilai yang tinggi menunjukkan bahwa dokumen yang paling relevan berhasil ditempatkan pada urutan teratas oleh sistem retrieval. Terakhir, metrik Answer Relevancy diterapkan untuk memastikan bahwa jawaban akhir yang diberikan oleh chatbot benar-benar menjawab inti pertanyaan pengguna secara langsung dan tidak bertele-tele.

Kinerja komponen pengambilan dokumen dari pangkalan data vektor FAISS menunjukkan tingkat akurasi yang sangat tinggi, dibuktikan oleh perolehan Context Recall rata-rata sebesar 0,97 dan Context Precision rata-rata sebesar 0,94. Nilai Context Recall yang mendekati sempurna mengindikasikan bahwa sistem secara konsisten berhasil menarik mayoritas fakta esensial dari Knowledge Base tanpa kehilangan informasi yang kritis. Sedikit deviasi pada Context Precision dari angka 1,00 merupakan hal yang wajar secara algoritmik karena model vektor terkadang ikut menarik potongan teks dari sub-bab yang berdekatan secara semantik meskipun bukan yang paling relevan. Pada aspek kualitas generasi jawaban, metrik Faithfulness mencapai skor rata-rata sebesar 0,96.

Angka ini memiliki implikasi praktis yang sangat signifikan dalam konteks sistem konsultasi akademik. Hampir seluruh klaim yang disampaikan oleh chatbot dapat dilacak kembali ke dokumen sumber yang valid, sehingga risiko penyampaian informasi yang salah kepada pengguna terkait program studi dan aturan akademik dapat diminimalkan hingga level yang sangat rendah. Dengan kata lain, sistem ini layak diandalkan sebagai sumber informasi akademik awal sebelum pengguna berkonsultasi langsung ke pihak institusi.

Metrik Answer Relevancy memperoleh skor rata-rata sebesar 0,93. Meskipun sedikit lebih rendah dibandingkan metrik lainnya, angka ini tetap berada pada kategori sangat baik dan memiliki implikasi praktis yang positif karena chatbot mampu menjawab pertanyaan pengguna secara langsung dan tepat sasaran. Penurunan minor dari angka sempurna disebabkan oleh kecenderungan sistem untuk menambahkan kalimat empatik dan informasi suportif tambahan yang bernilai positif bagi pengalaman pengguna meskipun secara algoritmik dinilai sebagai deviasi oleh kerangka RAGAS. Dalam konteks layanan konsultasi akademik yang menasar siswa yang sedang mempertimbangkan pilihan jurusan, respons yang empatik dan kontekstual justru meningkatkan kualitas layanan secara menyeluruh meskipun tidak terukur langsung oleh metrik relevansi.

Secara keseluruhan, hasil RAGAS membuktikan bahwa arsitektur Hybrid RAG yang diusulkan tidak hanya unggul secara kuantitatif pada metrik BERTScore, tetapi juga menunjukkan keandalan yang tinggi pada seluruh dimensi evaluasi retrieval dan generasi. Kombinasi Context Recall yang tinggi sebesar 0,97 dengan Faithfulness yang kuat sebesar 0,96 mengindikasikan bahwa sistem secara konsisten menemukan informasi yang tepat dan menggunakannya secara bertanggung jawab untuk membangun jawaban. Dua capaian ini merupakan prasyarat utama agar sebuah sistem RAG dapat diimplementasikan pada domain yang bersifat kritis seperti konsultasi akademik. Hasil ini juga mengonfirmasi bahwa integrasi Query Router tidak hanya meningkatkan akurasi semantik, tetapi juga menjaga kualitas retrieval tetap presisi dengan mengarahkan setiap kueri hanya ke sumber data yang paling relevan.

4. Kesimpulan

Penelitian ini berhasil merancang dan mengimplementasikan sistem chatbot rekomendasi jurusan berbasis Hybrid RAG dengan integrasi instrumen RIASEC 18 pertanyaan. Arsitektur Query Router terbukti efektif dalam memisahkan pemrosesan data relasional deterministik dan pencarian semantik secara simultan, sehingga mampu menekan fenomena halusinasi data secara signifikan dan menghasilkan

rekomendasi yang lebih personal dibandingkan pendekatan LLM murni maupun RAG Standar.

Pengujian komparatif antara RAG Standar dan Hybrid RAG menggunakan metrik BERTScore menunjukkan peningkatan F1-Score rata-rata dari 0,728 menjadi 0,774. Peningkatan paling signifikan terjadi pada jalur Consultation yang naik dari 0,739 menjadi 0,800, membuktikan bahwa integrasi profil kepribadian RIASEC secara deterministik berhasil menghasilkan rekomendasi yang lebih personal dan akurat. Hasil pengujian RAGAS menunjukkan nilai Context Recall sebesar 0,97, Context Precision sebesar 0,94, dan Faithfulness sebesar 0,96, yang membuktikan bahwa sistem sepenuhnya terkunci pada fakta dokumen institusi dan bebas dari halusinasi data.

Meskipun hasil yang diperoleh menjanjikan, penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Sistem saat ini hanya dapat menjawab pertanyaan yang topiknya tercakup dalam Knowledge Base yang tersedia, sehingga cakupan pengetahuan sistem masih terbatas pada dokumen yang telah diindeks. Selain itu, penerapan sistem pada lingkungan nyata menghadapi tantangan berupa ketergantungan pada koneksi internet yang stabil untuk mengakses Groq API, serta potensi kesalahan routing pada kueri yang memiliki intent ambigu. Pengujian pada penelitian ini juga belum mencakup evaluasi penerimaan pengguna secara langsung, sehingga perspektif pengguna akhir terhadap kualitas dan kenyamanan sistem belum terukur secara formal.

Sebagai arah pengembangan penelitian selanjutnya, beberapa hal dapat dilakukan untuk meningkatkan kualitas sistem. Pertama, pelaksanaan User Acceptance Testing secara formal untuk mengukur tingkat penerimaan dan kepuasan pengguna nyata. Kedua, pengembangan mekanisme Query Router berbasis model klasifikasi yang lebih adaptif agar mampu menangani kueri dengan intent yang lebih kompleks dan ambigu. Ketiga, perluasan cakupan Knowledge Base dengan mengintegrasikan dokumen kurikulum dan data akademik yang lebih lengkap dari seluruh program studi. Keempat, eksplorasi penggunaan model bahasa lokal yang dapat dijalankan secara mandiri tanpa ketergantungan pada layanan API eksternal, guna meningkatkan keandalan sistem pada kondisi jaringan yang terbatas.

Daftar Rujukan

- [1] M. Ibrahim, N. Izza, A. H. Aslam, and R. R. Muhammad, "Jurnal Didactical Mathematics Fenomena Salah Jurusan Terhadap Studi Mahasiswa pada Prodi Pendidikan Matematika Universitas Islam Negeri Raden Intan Lampung," vol. 7, pp. 224–240, 2025. <https://doi.org/10.31949/dm.v7i2.12231>
- [2] K. Guru and J. Peminatan, "Risalah Kebijakan Urgensi Penguatan Pemahaman Terkait Penghapusan," no. September, 2022.

- [3] M. F. Choiri, Daharnis, and Idfil, "Relevansi Teori Holland Terhadap Pemilihan Karir Peserta Didik Di Era Modern: Kajian Literatur," *J. Bimbing. Konseling dan Dakwah*, vol. 6, no. 1, 2025. <https://doi.org/10.55380/taqorrub.v6i01.1102>
- [4] G. H. Martins *et al.*, "18REST-2 : A Revised Measure of the RIASEC Model for Large-Scale Assessment With Students," vol. 0, no. 0, pp. 1–20, 2024, <https://doi.org/10.1177/10690727241256289>
- [5] F. L. D. Cahyanti and R. D. Andika Raya, "Perancangan Sistem Informasi Chatbot Retrieval Augmented Generation Berbasis Website pada PT. Revolusi Cita Edukasi," *Indones. J. Comput. Sci.*, 2025, <https://doi.org/10.31294/m75d4782>
- [6] S. Elysia and Herianto, "Chatbot Berbasis Retrieval Augmented Generation (RAG) untuk Peningkatan Layanan Informasi Sekolah," *J. Technol. Inf. Data Anal.*, 2024, <https://doi.org/10.70491/tifda.v1i2.52>
- [7] Murtiyoso, I. Tahyudin, and Berlilana, "A Systematic Review of Retrieval-Augmented Generation for Enhancing Domain-Specific Knowledge in Large Language Models," *Sink. J. dan Penelit. Tek. Inform.*, 2025, <https://doi.org/10.33395/sinkron.v9i2.14824>.
- [8] A. Nugroho, A. Pramudito, and S. Lestari, "Hybrid Retrieval-Augmented Generation untuk Sistem Question Answering Berbasis Dokumen Akademik," *J. Teknol. Inf. dan Ilmu Komput.*, 2024, doi: 10.25126/jtiik.202484512.
- [9] M. R. Rachman, M. Rosidin, and W. Y. Sulistyo, "Implementation of Augmented Generation Retrieval Method in Chatbot for Contract Customer Service Automation," *Insect J. Tek. Inform.*, 2025, <https://doi.org/10.33506/insect.v11i2.4992>
- [10] D. Rothman, *RAG-Driven Generative AI*. Packt Publishing Ltd.
- [11] C. Cohn, S. Rayala, C. Snyder, and *et al.*, "Personalizing Student-Agent Interactions Using Log-Contextualized Retrieval Augmented Generation (RAG)," *arXiv*, 2025, [Online]. Available: <https://arxiv.org/abs/2505.17238>
- [12] I. Fanani, "Implementasi Retrieval Augmented Generation untuk Evaluasi Proposal Tugas Akhir Mahasiswa," *J. Teknol. Komput. dan Inform.*, 2025, <https://doi.org/10.59820/tekomin.v3i2.336>
- [13] I. I. R. Pratama and B. Sisephaputra, "Pengembangan Sistem Helpdesk Menggunakan Chatbot dengan Metode Retrieval-Augmented Generation (RAG)," *J. Informatics Comput. Sci.*, 2025, <https://doi.org/10.26740/jinacs.v6n03.p696-710>
- [14] I. Pujiono, I. M. Agtyaputra, and Y. Ruldeviyani, "Implementing Retrieval-Augmented Generation And Vector Databases For Chatbots In Public Services Agencies Context," vol. 10, no. 1, pp. 216–223, 2024, <https://doi.org/10.33480/jitk.v10i1.5572>
- [15] M. F. Siregar, D. Putra, and A. Kurniawan, "Penerapan Retrieval-Augmented Generation pada Chatbot Layanan Informasi Pemerintahan," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, 2024, doi: 10.29207/resti.v8i4.5123.
- [16] H. Tohir, N. Merlina, and M. Haris, "Utilizing Retrieval-Augmented Generation in Large Language Models to Enhance Indonesian Language NLP," *J. Ilmu Pengetah. dan Teknol. Komput.*, 2024, <https://doi.org/10.33480/jitk.v10i2.5916>
- [17] S. M. Yasmin and D. H. Fudholi, "Pengembangan Chatbot Informasi Hukum Layanan Publik Berbasis Retrieval-Augmented Generation Menggunakan LangChain dan OpenAI di Ombudsman DIY Development of a Legal Information Chatbot for Public Services Using RAG , LangChain , and OpenAI at Ombudsman D," vol. 5, no. 9, pp. 2548–2565, 2025, <https://doi.org/10.52436/1.jpti.995>
- [18] A. Marsadualan, H. Harmastuti, and J. Triyono, "Rancang Bangun Aplikasi Tanya Jawab Mengenai Ist Akprind Yogyakarta Berbasis Mobile Menggunakan Algoritma Boyer Moore," *Semin. Nas. Inov. Sains Teknol. Inf. Komput.*, vol. 1, no. 1, pp. 226–238, 2023.
- [19] S. Q. Kharina and A. Yulianto, "Perancangan online test untuk mengukur minat calon mahasiswa Universitas Pembangunan Jaya," *J. Psikol. Media Ilm. Psikol.*, vol. 19, no. 1, pp. 1–8, 2021.
- [20] R. Kalra *et al.*, "HyPA-RAG: A Hybrid Parameter Adaptive Retrieval-Augmented Generation System for AI Legal and Policy Applications," 2023, <https://doi.org/10.18653/v1/2024.customnlp4u-1.18>
- [21] S. Engineering, R. Ahadi, N. S. Harahap, M. Fikry, F. Kurnia, and I. Artikel, "Retrieval-Augmented Generation in a Web-Based Question Answering System for Fiqh Books," vol. 5, no. 2, pp. 626–635, 2025, doi: 10.30811/jaise.v5i2.7005.
- [22] Y. Tribber and M. Asfi, "Implementasi Retrieval Augmented Generation untuk Layanan Informasi Kampus dengan Chatbot Berbasis AI," pp. 594–600.