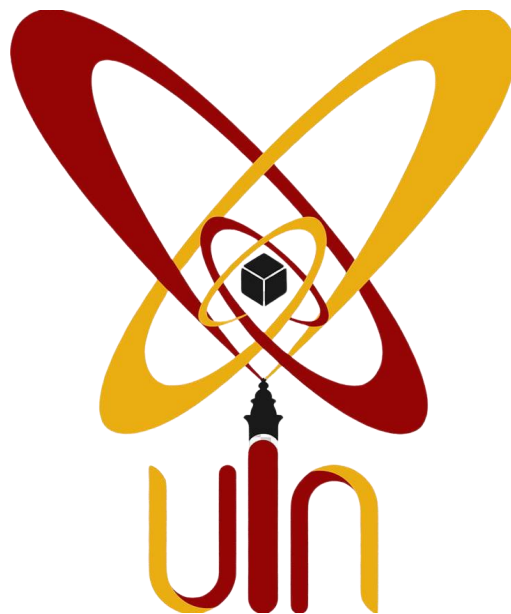


LAPORAN HASIL PELAKSANAAN PKL/MAGANG
Ekstraksi Data dan Orkestrasi Agen menggunakan
LangGraph pada Sistem Chatbot Retrieval-Augmented
Generation (RAG) untuk SRE

Disusun untuk memenuhi salah satu syarat pelaksanaan PKL/Magang
pada Program Studi Informatika



Disusun oleh:

Nama : Robiatul Adawiah

NIM : 231730013

PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI
SULTAN MAULANA HASANUDDIN BANTEN
2026

HALAMAN PENGESAHAN

Laporan hasil pelaksanaan PKL/Magang ini telah disusun oleh:

Nama : Robiatul Adawiah
NIM : 231730028
Program Studi : Informatika
Judul PKL/Magang : Data Extraction and Agent Orchestration using
LangGraph on a Chatbot Retrieval-Augmented
Generation (RAG) System for SRE
Tempat PKL/Magang : BRIN KST Cisit Bandung Samaun Samadikun
Waktu Pelaksanaan : 18 Februari – 17 Juli 2026 (6 Bulan)

Laporan ini telah diperiksa dan disahkan sebagai laporan hasil pelaksanaan
PKL/Magang.

Bandung, 25 Juni 2026

Dosen Pembimbing PKL/Magang,



M. Iman Wahyudi, S.Pd.I., M.Kom

NIP. 1998507142020121003

Pembimbing Lapangan,



Rio Nurtantvana, S.Pd., M.Pd., M.Sc., Ph.D

Peneliti Ahli Muda

Mengetahui,

Ketua Program Studi Informatika



Ahmad Tabrani M.T.I

NIP. 198509172018011002

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas rahmat dan karunia-Nya, penulis dapat menyelesaikan laporan hasil pelaksanaan PKL/Magang yang berjudul "Ekstraksi Data dan Orkestrasi Agen menggunakan LangGraph pada Sistem Chatbot Retrieval-Augmented Generation (RAG) untuk SRE" dengan baik.

Laporan ini disusun sebagai salah satu bentuk pertanggungjawaban atas kegiatan PKL/Magang yang telah dilaksanakan di Pusat Riset Sains Data dan Informasi, Badan Riset dan Inovasi Nasional (BRIN). Melalui kegiatan ini, penulis memperoleh pengalaman, pengetahuan, serta pemahaman secara langsung mengenai pengembangan sistem kecerdasan buatan berbasis Large Language Model (LLM) dan Retrieval-Augmented Generation (RAG).

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan bantuan, bimbingan, dan dukungan selama pelaksanaan PKL/Magang hingga penyusunan laporan ini, terutama kepada:

1. Bapak Ahmad Tabrani M.T.I
2. Bapak M. Iman Wahyudi, S.Pd.I., M.Kom
3. Rio Nurtantyana, S.Pd., M.Pd., M.Sc., Ph.D
4. Badan Riset dan Inovasi Nasional (BRIN)
5. Semua pihak yang telah membantu dalam pelaksanaan PKL/Magang.

Penulis menyadari bahwa laporan ini masih memiliki kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan agar laporan ini dapat menjadi lebih baik. Semoga laporan ini dapat memberikan manfaat bagi pembaca dan pihak-pihak yang membutuhkan.

Bandung, 25 Juni 2026



Robiatul Adawiah

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR	iii
DAFTAR ISI.....	iv
RINGKASAN	vi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	1
1.3 Tujuan PKL/Magang	2
1.4 Manfaat PKL/Magang	2
BAB II KAJIAN PUSTAKA	3
2.1 Kajian Pustaka	3
2.1.1 Retrieval-Augmented Generation (RAG)	3
2.1.2 LangGraph dan Agentic RAG.....	3
2.1.3 Arsitektur Dual-Database (pgvector dan Neo4j).....	3
2.1.4 DeepEval sebagai Framework Evaluasi RAG	4
2.1.5 Siklus Kognitif PRAL	4
2.2 Kerangka Pikir Teoritis	4
BAB III	5
3.1 Lokasi PKL/Magang	5
3.2 Alat dan Bahan	5
3.3 Penentuan Objek PKL/Magang	5
3.4 Metode Pengumpulan Data	6

3.5	Teknik Analisis Data	6
BAB IV HASIL PKL/MAGANG DAN PEMBAHASAN.....		7
4.1	Deskripsi Kegiatan PKL/Magang	7
4.2	Permasalahan yang Ditemukan	8
4.3	Metode Penyelesaian Masalah	9
4.4	Hasil Pelaksanaan PKL/Magang	9
4.5	Pembahasan	10
4.6	Pelajaran yang Diperoleh	11
BAB V KESIMPULAN DAN SARAN.....		12
5.1	Kesimpulan.....	12
5.2	Keterbatasan	13
5.3	Implikasi	13
5.4	Saran.....	13
DAFTAR PUSTAKA		15
LAMPIRAN.....		16

RINGKASAN

Laporan ini menyajikan hasil pelaksanaan PKL/Magang di Pusat Riset Sains Data dan Informasi, Badan Riset dan Inovasi Nasional (BRIN), yang berlangsung dari 18 Februari hingga 30 Juli 2026. Kegiatan berfokus pada pengembangan sistem MySRE, sebuah chatbot berbasis Retrieval-Augmented Generation (RAG) agentik yang menggunakan arsitektur dual-database (PostgreSQL pgvector dan Neo4j) yang diorkestrasi melalui LangGraph.

Sistem MySRE dirancang berdasarkan siklus kognitif PRAL (Perceive, Reason, Act, Learn) untuk mengontrol perilaku chatbot secara deterministik. Pipeline data ingestion mengekstraksi dokumen ilmiah menggunakan PyMuPDF (fitz) dan menerapkan algoritma Sliding Window Chunking untuk menghasilkan segmen teks dengan ukuran $K=1500$ karakter dan overlap $O=150$ karakter. Setiap segmen dikonversi menjadi vektor embedding 1536 dimensi menggunakan model text-embedding-3-small dan disimpan di PostgreSQL dengan indeks HNSW, sementara jaringan relasi antardokumen dipetakan ke dalam Neo4j.

Alur kerja agen dikelola oleh empat node LangGraph: Router Node (klasifikasi layer), Retrieval Node (pencarian semantik dan ekspansi graf), Compressor Node (pemfilteran konteks), dan Generator Node (pembangkitan respons). Evaluasi dilakukan menggunakan framework DeepEval dengan 24 sesi pengujian berdasarkan 12 query unik. Hasil menunjukkan MySRE mencapai akurasi klasifikasi 75,00% dan F1-Score 76,92%, melampaui model Baseline ChatGPT (66,67%). Selain itu, MySRE mampu menekan biaya API operasional sebesar 8 kali lebih hemat (\$0,00034867 vs. \$0,00281068 per query).

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pesatnya pertumbuhan publikasi ilmiah menghasilkan volume informasi riset yang sangat besar. Meskipun hal ini memberikan peneliti akses ke berbagai sumber pengetahuan, kondisi ini juga menimbulkan tantangan dalam menemukan dan memanfaatkan informasi yang relevan. Proses tinjauan literatur yang dilakukan secara manual menjadi tidak efisien ketika peneliti harus menyaring ratusan dokumen.

Large Language Model (LLM) menawarkan pendekatan baru untuk membantu proses pencarian informasi melalui chatbot berbasis AI. Namun, penggunaan LLM secara mandiri masih memiliki keterbatasan karena pengetahuannya bersifat statis sesuai data pelatihan dan rentan menghasilkan jawaban yang tidak sesuai fakta (halusinasi).

Pendekatan Retrieval-Augmented Generation (RAG) dikembangkan untuk mengatasi keterbatasan tersebut dengan mengintegrasikan sistem pengambilan informasi eksternal ke dalam proses generasi LLM. Dalam Smart Research Environment (SRE), sumber pengetahuan umumnya berupa dokumen ilmiah yang mengandung teks, tabel, gambar, dan rumus matematika yang harus diekstraksi secara akurat. Lebih lanjut, implementasi RAG konvensional masih mengandalkan alur kerja linear sehingga membutuhkan mekanisme orkestrasi agen yang dapat mengelola setiap tahap secara terstruktur dan adaptif.

1.2 Rumusan Masalah

1. Bagaimana membangun pipeline ekstraksi dan ingesti data dokumen ilmiah yang efisien untuk mendukung sistem RAG pada Smart Research Environment?
2. Bagaimana merancang arsitektur dual-database (PostgreSQL pgvector dan Neo4j) untuk menyimpan representasi semantik dan jaringan relasi antardokumen ilmiah?

3. Bagaimana mengimplementasikan orkestrasi agen berbasis LangGraph dengan siklus kognitif PRAL yang mampu meminimalkan halusinasi dan mengoptimalkan efisiensi biaya API?

1.3 Tujuan PKL/Magang

1. Membangun pipeline ingesti data menggunakan PyMuPDF dan Sliding Window Chunking untuk menghasilkan representasi vektor dokumen ilmiah yang akurat.
2. Mengimplementasikan arsitektur dual-database dengan PostgreSQL (pgvector, indeks HNSW) untuk pencarian semantik dan Neo4j untuk pemetaan jaringan relasi antardokumen.
3. Merancang dan mengimplementasikan State Graph LangGraph berbasis siklus PRAL dengan empat node fungsional (Router, Retrieval, Compressor, Generator) beserta UX Mode Guardrails.
4. Mengevaluasi performa sistem MySRE menggunakan framework DeepEval dengan metrik Faithfulness, Answer Relevancy, Contextual Precision, Recall, dan F1-Score.

1.4 Manfaat PKL/Magang

Manfaat yang diperoleh dari pelaksanaan PKL/Magang ini adalah sebagai berikut.

1. Bagi mahasiswa, meningkatkan kompetensi teknis di bidang Natural Language Processing (NLP), rekayasa sistem RAG agentik, arsitektur dual-database, serta evaluasi sistem AI menggunakan framework DeepEval.
2. Bagi BRIN, tersedianya prototipe sistem MySRE yang mampu mengotomatisasi proses pencarian dan sintesis informasi dari dokumen SRE dengan akurasi lebih tinggi dan biaya API 8 kali lebih hemat dibandingkan pendekatan konvensional.
3. Bagi Program Studi Informatika, memperkuat relevansi kurikulum di bidang kecerdasan buatan dan rekayasa perangkat lunak modern, serta memperdalam kemitraan dengan lembaga riset nasional

BAB II

KAJIAN PUSTAKA

2.1 Kajian Pustaka

2.1.1 Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) adalah pendekatan yang menggabungkan mekanisme pengambilan informasi dari sumber eksternal dengan proses generasi *Large Language Model*. Pendekatan ini dikembangkan untuk mengatasi keterbatasan LLM yang pengetahuannya bersifat statis. Chatbot berbasis RAG terbukti dapat meningkatkan kualitas tinjauan literatur dengan mengambil dan meringkas makalah penelitian yang relevan secara signifikan [1].

2.1.2 LangGraph dan Agentic RAG

LangGraph adalah kerangka kerja berbasis graf yang memungkinkan pembangunan sistem agen AI dengan alur kerja yang fleksibel dan dinamis. Berbeda dengan pipeline RAG linear konvensional, LangGraph mendukung pembuatan State Machine di mana agen dapat melakukan analisis query, seleksi konteks, retrieval, dan pembangkitan jawaban secara adaptif. Arsitektur Agentic RAG yang diorkestrasi LangGraph mampu mengelola interaksi antarproses secara lebih terstruktur [2], [3].

2.1.3 Arsitektur Dual-Database (pgvector dan Neo4j)

Sistem MySRE mengimplementasikan arsitektur penyimpanan hibrida yang memisahkan representasi vektor semantik dari pemetaan relasi antardokumen. PostgreSQL dengan ekstensi pgvector menyimpan embedding vektor berindeks HNSW untuk pencarian semantik berbasis cosine similarity. Neo4j sebagai graph database menyimpan metadata dan jaringan relasi antardokumen (CONNECTED_TO, HAS_PILLAR) untuk mendukung analisis komparatif lintas dokumen [3].

2.1.4 DeepEval sebagai Framework Evaluasi RAG

DeepEval adalah framework evaluasi standar industri untuk mengukur performa sistem RAG secara komprehensif. Metrik yang digunakan meliputi Faithfulness (kemampuan LLM menjawab berdasarkan konteks tanpa mengarang fakta), Answer Relevancy (relevansi jawaban terhadap query pengguna), Contextual Precision (presisi chunk yang diambil), serta Recall dan F1-Score klasifikasi berdasarkan confusion matrix. Framework ini memberikan landasan evaluasi yang objektif dan terukur.

2.1.5 Siklus Kognitif PRAL

Siklus PRAL (Perceive, Reason, Act, Learn) merupakan siklus kognitif agentik yang diterapkan dalam sistem MySRE untuk mengontrol perilaku chatbot secara deterministik. Tahap Perceive membaca query dan status sistem, Reason mengklasifikasikan query untuk menentukan jalur retrieval, Act mengeksekusi pencarian dan pembangkitan respons, serta Learn mencatat detail percakapan dan metrik eksekusi sistem.

2.2 Kerangka Pikir Teoritis

Kerangka pikir dalam PKL/Magang ini dimulai dari identifikasi permasalahan utama: variabilitas format dokumen ilmiah, keterbatasan RAG linear, dan risiko halusinasi LLM. Solusi yang dikembangkan mencakup tiga lapisan: (1) lapisan ingesti data menggunakan PyMuPDF dan Sliding Window Chunking yang menghasilkan embedding vektor dan peta relasi Neo4j; (2) lapisan orkestrasi menggunakan LangGraph State Graph berbasis siklus PRAL dengan empat node fungsional dan UX Mode Guardrails; dan (3) lapisan evaluasi menggunakan framework DeepEval. Ketiga lapisan ini bekerja sinergis untuk menghasilkan sistem chatbot RAG agentik yang akurat, efisien, dan siap produksi.

BAB III

METODE PKL/MAGANG

3.1 Lokasi PKL/Magang

Penelitian ini dilaksanakan di Pusat Riset Sains Data dan Informasi, Badan Riset dan Inovasi Nasional (BRIN) KST Samaun Samadikun, Jl. Sangkuriang, Dago, Kecamatan Coblong, Kota Bandung, Jawa Barat 40135. Kegiatan berlangsung selama 6 bulan, mulai 18 Februari 2026 hingga 30 Juli 2026.

3.2 Alat dan Bahan

Alat dan bahan yang digunakan selama PKL/Magang adalah sebagai berikut.

1. Perangkat keras: Laptop dan komputer server BRIN untuk menjalankan model embedding dan proses ingesti.
2. Bahasa pemrograman: Python.
3. Library utama: LangGraph, LangChain, PyMuPDF (fitz), MarkdownHeaderTextSplitter, RecursiveCharacterTextSplitter, pgvector, dan DeepEval.
4. Database: PostgreSQL dengan ekstensi pgvector (indeks HNSW) dan Neo4j Graph Database.
5. Model AI: OpenAI gpt-4o-mini (Router Node: $\tau=0,0$; Generator Node: $\tau=0,3$) dan text-embedding-3-small (embedding 1536 dimensi).
6. Platform kolaborasi: Git untuk version control dan Slack untuk komunikasi tim.

3.3 Penentuan Objek PKL/Magang

Objek Prosedur pelaksanaan PKL/Magang dibagi menjadi dua tahap utama sebagai berikut.

1. Tahap Pelatihan Pengantar (1 minggu): Pembekalan materi meliputi building learning commitment, pengenalan BRIN, research integrity, kerja tim, komunikasi persuasif, dan teknik presentasi.

2. Tahap Pembimbingan (24 minggu): Pelaksanaan kegiatan riset dan pengembangan sistem MySRE di bawah bimbingan pembimbing lapangan, mencakup studi literatur, perancangan arsitektur sistem PRAL, implementasi pipeline data ingestion, pengembangan agentic flow LangGraph, dan evaluasi menggunakan DeepEval.

3.4 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan selama PKL/Magang adalah sebagai berikut.

1. Observasi, yaitu pengamatan langsung terhadap kegiatan pengembangan sistem di lingkungan BRIN, termasuk partisipasi dalam sprint review mingguan.
2. Wawancara, yaitu pengumpulan informasi melalui diskusi teknis dengan pembimbing lapangan dan anggota tim penelitian Kelompok Riset HCIV.
3. Dokumentasi, yaitu pengumpulan data berupa logbook harian, kode program, log eksekusi sistem, serta hasil evaluasi DeepEval (confusion matrix, metrik RAG, dan data biaya API).

3.5 Teknik Analisis Data

Teknik analisis data yang digunakan adalah analisis deskriptif kualitatif dan kuantitatif. Secara kuantitatif, performa sistem MySRE dievaluasi menggunakan framework DeepEval melalui dua pendekatan: (1) confusion matrix untuk mengukur akurasi klasifikasi (TP, TN, FP, FN) dengan threshold skor relevansi dan faithfulness $\geq 0,5$; dan (2) metrik RAG meliputi Faithfulness, Answer Relevancy, Contextual Precision, Recall, dan F1-Score. Analisis efisiensi operasional dilakukan dengan membandingkan latensi respons dan biaya API antara MySRE dan model Baseline (ChatGPT dengan Direct PDF).

BAB IV

HASIL PKL/MAGANG DAN PEMBAHASAN

4.1 Deskripsi Kegiatan PKL/Magang

Selama pelaksanaan PKL/Magang di Pusat Riset Sains Data dan Informasi BRIN, penulis terlibat aktif dalam pengembangan sistem MySRE (My Smart Research Environment). Kegiatan dilaksanakan secara daring dan luring mengikuti metodologi Agile dengan sprint review mingguan bersama Kelompok Riset Human-Computer Interaction and Visualization (HCIV).

Tabel 4.1. Kegiatan PKL/Magang

No	Tanggal	Kegiatan	Keterangan
1	18 Feb - 24 Feb 2026	Pelatihan Pengantar MBKM BRIN	Building learning commitment, pengenalan BRIN, research integrity, kerja tim, teknik presentasi
2	25 Feb - 02 Mar 2026	Studi Literatur RAG dan LangGraph	Kajian metode RAG, LangGraph, arsitektur chatbot, pgvector, dan Neo4j
3	03 Mar - 30 Mar 2026	Perancangan Arsitektur MySRE dan Pipeline Data Ingestion	Desain siklus PRAL, skema dual-database (PostgreSQL pgvector dan Neo4j), sliding window chunking
4	31 Mar - 27 Apr 2026	Implementasi Agentic Flow LangGraph	Pengembangan Router Node, Retrieval Node, Compressor Node, Generator Node, dan UX Mode Guardrails

5	28 Apr - 25 Mei 2026	Pengujian dan Evaluasi dengan DeepEval	Eksekusi 24 sesi pengujian (12 query x 2 dokumen), analisis metrik Faithfulness, Answer Relevancy, Contextual Precision menggunakan framework DeepEval
6	09 Jun 2026	Demonstrasi Fitur kepada Expert	Presentasi dan demonstrasi MySRE kepada tim ahli BRIN
7	10 Jun - 30 Jul 2026	Finalisasi dan Penyusunan Laporan	Dokumentasi teknis, penyusunan laporan akhir, dan persiapan publikasi jurnal

Sumber: Data diolah oleh penulis, 2026

4.2 Permasalahan yang Ditemukan

Selama Permasalahan yang ditemukan selama pelaksanaan PKL/Magang adalah sebagai berikut.

1. Dokumen ilmiah dalam sistem SRE memiliki format yang sangat bervariasi, termasuk dokumen PDF digital dan dokumen yang mengandung tabel serta rumus. Hal ini menyebabkan pipeline ekstraksi teks konvensional menghasilkan output yang tidak terstruktur.
2. Implementasi RAG konvensional yang bersifat linear tidak mampu menangani query komparatif lintas dokumen maupun query di luar cakupan (out-of-scope) secara adaptif.
3. Penggunaan LLM secara langsung dengan seluruh konten PDF (Baseline) menghasilkan konsumsi token yang sangat tinggi, sehingga biaya operasional API menjadi tidak efisien untuk skala produksi.

4. Tanpa mekanisme self-correction dan UX guardrails, sistem rentan menghasilkan jawaban berdasarkan pengetahuan umum LLM untuk query yang tidak tercakup dalam dokumen (halusinasi).

4.3 Metode Penyelesaian Masalah

Metode penyelesaian masalah yang diterapkan adalah sebagai berikut.

1. Untuk mengatasi variabilitas dokumen, penulis mengimplementasikan pipeline ingesti data menggunakan PyMuPDF (fitz) yang mempertahankan struktur heading dan paragraf, dilanjutkan dengan Sliding Window Chunking menggunakan MarkdownHeaderTextSplitter dan RecursiveCharacterTextSplitter dengan ukuran $K=1500$ karakter dan overlap $O=150$ karakter.
2. Untuk menangani berbagai tipe query secara adaptif, penulis merancang Router Node berbasis LangGraph yang mengklasifikasikan query ke dalam tiga layer: Layer 1 (query dokumen tunggal), Layer 2 (query komparatif multi-dokumen dengan ekspansi Neo4j), dan Layer 3 (query di luar cakupan, short-circuit langsung ke Generator).
3. Untuk mengoptimalkan biaya API, penulis mengimplementasikan Compressor Node yang memfilter chunk mentah dari PostgreSQL menjadi maksimal 4 chunk paling relevan ($C_{max}=4$) sebelum dikirim ke Generator Node, sehingga input token LLM tetap stabil antara 1.329 hingga 2.552 token terlepas dari panjang dokumen sumber.
4. Untuk menekan halusinasi dan memberikan batas konteks yang jelas, penulis mengimplementasikan UX Mode Guardrails yang memblokir alur kerja apabila tidak ada dokumen aktif (STRICT mode) atau mengarahkan pengguna untuk beralih ke RESEARCH mode untuk query komparatif.

4.4 Hasil Pelaksanaan PKL/Magang

Hasil yang diperoleh selama kegiatan PKL/Magang adalah sebagai berikut.

1. Sistem MySRE lengkap dengan pipeline data ingestion (PyMuPDF + Sliding Window Chunking + text-embedding-3-small) yang menghasilkan

embedding vektor 1536 dimensi tersimpan di PostgreSQL dengan indeks HNSW.

2. Arsitektur dual-database yang mengintegrasikan PostgreSQL pgvector untuk pencarian semantik (cosine similarity, top-K=15 chunk) dan Neo4j untuk pemetaan relasi antardokumen (CONNECTED_TO, HAS_PILLAR) berdasarkan lima pilar SRE: Background, Methodology, Gap, Objective, dan Future Work.

3. State Graph LangGraph berbasis siklus PRAL dengan empat node fungsional (Router Node, Retrieval Node, Compressor Node, Generator Node), conditional routing tiga layer, dan UX Mode Guardrails (STRICT dan RESEARCH mode).

4. Hasil evaluasi DeepEval menunjukkan MySRE mencapai akurasi klasifikasi 75,00% dan F1-Score 76,92%, dengan Faithfulness 0,9254 dan Answer Relevancy 0,4438, melampaui model Baseline ChatGPT (akurasi 66,67%, Faithfulness 0,9054).

5. Efisiensi biaya API sebesar 8 kali lebih hemat: MySRE rata-rata \$0,00034867 per query vs. Baseline \$0,00281068 per query, dengan pengurangan input token lebih dari 90% berkat Compressor Node.

4.5 Pembahasan

Hasil pelaksanaan PKL/Magang menunjukkan relevansi yang kuat antara implementasi di lapangan dengan kajian pustaka. Penerapan Sliding Window Chunking dengan overlap $O=150$ karakter terbukti mencegah fragmentasi semantik pada batas chunk, sesuai prinsip dasar text segmentation pada sistem RAG.

Arsitektur dual-database yang memisahkan penyimpanan vektor (pgvector) dan penyimpanan relasi (Neo4j) terbukti menghasilkan kemampuan retrieval yang lebih kaya, terutama untuk query komparatif (Layer 2) yang memanfaatkan ekspansi graf Neo4j. Mekanisme conditional routing tiga layer pada Router Node secara efektif mengoptimalkan latensi dan biaya: query out-

of-scope (Layer 3) melewati database sepenuhnya sehingga menghemat latensi retrieval.

Evaluasi menggunakan DeepEval mengkonfirmasi bahwa Compressor Node merupakan komponen kunci efisiensi sistem. Dengan membatasi input ke maksimal 4 chunk terkompres, LLM terlindungi dari information overload yang dialami model Baseline saat memproses seluruh dokumen PDF, sekaligus menghasilkan Faithfulness lebih tinggi (0,9254 vs. 0,9054). Skor Contextual Precision (0,2944) dan Recall (0,3958) yang tampak rendah disebabkan oleh inklusi query adversarial dalam dataset pengujian; pada query adversarial ini, sistem secara tepat mengembalikan nol chunk karena tidak ada konteks relevan dalam database, yang justru merupakan perilaku yang benar.

4.6 Pelajaran yang Diperoleh

Selama pelaksanaan PKL/Magang di Badan Riset dan Inovasi Nasional (BRIN), penulis memperoleh pelajaran berharga sebagai berikut.

1. Hard Skills: (a) Kemampuan merancang pipeline ingesti data dokumen ilmiah menggunakan PyMuPDF dan Sliding Window Chunking; (b) Keahlian mengimplementasikan arsitektur dual-database dengan pgvector dan Neo4j; (c) Kemampuan merancang State Graph LangGraph berbasis siklus PRAL dengan routing kondisional dan UX guardrails; (d) Keahlian melakukan evaluasi sistematis menggunakan framework DeepEval (confusion matrix, metrik RAG, analisis biaya API).
2. Soft Skills: (a) Kolaborasi antarmodul dalam ekosistem sistem berskala industri bersama tim database Neo4j; (b) Pentingnya penulisan Clean Code dan logging terstruktur dalam proyek multi-komponen; (c) Pendekatan problem solving berbasis data ketika menghadapi kendala teknis; (d) Komunikasi teknis dalam sesi sprint review dan demonstrasi kepada expert BRIN.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil pelaksanaan PKL/Magang di Badan Riset dan Inovasi Nasional (BRIN), dapat disimpulkan bahwa sistem MySRE yang dikembangkan berhasil membuktikan keunggulan arsitektur RAG agentik berbasis dual-database dibandingkan pendekatan konvensional. Pipeline data ingestion menggunakan PyMuPDF dan Sliding Window Chunking (K=1500, O=150 karakter) terbukti menghasilkan representasi vektor embedding 1536 dimensi yang akurat dan tersimpan efisien di PostgreSQL dengan indeks HNSW.

Arsitektur dual-database yang memadukan pgvector dan Neo4j berhasil mengintegrasikan kemampuan pencarian semantik dengan pemetaan relasi antardokumen melalui lima pilar SRE, sehingga mendukung query tunggal maupun komparatif lintas dokumen secara bersamaan. State Graph LangGraph berbasis siklus PRAL dengan empat node fungsional dan conditional routing tiga layer terbukti mampu mengontrol perilaku chatbot secara deterministik, di mana UX Mode Guardrails menghasilkan True Negative sempurna (TN=8) dan False Positive nol pada seluruh pengujian query adversarial.

Evaluasi komprehensif menggunakan framework DeepEval membuktikan MySRE mencapai akurasi klasifikasi 75,00% dan F1-Score 76,92%, melampaui Baseline ChatGPT yang hanya mencatatkan 66,67%. Dari sisi efisiensi operasional, MySRE mampu menekan biaya API hingga 8 kali lebih hemat (\$0,00034867 vs. \$0,00281068 per query) berkat Compressor Node yang membatasi input token LLM secara konsisten antara 1.329 hingga 2.552 token, terlepas dari panjang dokumen sumber.

5.2 Keterbatasan

Terdapat beberapa keterbatasan yang dihadapi selama pelaksanaan PKL/Magang. Durasi magang selama 6 bulan membatasi ruang eksplorasi optimasi lebih lanjut, seperti perbandingan metrik jarak vektor (Euclidean, Manhattan) maupun integrasi model bahasa lokal seperti Llama 3 atau Mistral untuk kebutuhan privasi data. Selain itu, dataset evaluasi yang terbatas pada 12 query unik di dua dokumen (24 sesi total) menyebabkan generalisasi hasil ke domain dokumen SRE yang lebih luas masih memerlukan pengujian tambahan. Latensi respons rata-rata MySRE (16,44 detik) juga masih lebih tinggi dibandingkan Baseline (11,12 detik) akibat overhead orchestrator LangGraph, meskipun sebagian telah diatasi dengan implementasi token streaming via Server-Sent Events (SSE).

5.3 Implikasi

Hasil PKL/Magang ini membawa implikasi yang berarti bagi berbagai pihak. Bagi penulis, pengalaman langsung mengembangkan sistem AI produksi di lingkungan BRIN memberikan fondasi kompetensi yang kuat di bidang arsitektur RAG agentik dan evaluasi sistem AI berbasis data, yang menjadi bekal berharga untuk karier di industri teknologi. Bagi BRIN, sistem MySRE secara konkret dapat mempercepat siklus penelitian dengan efisiensi biaya API 8 kali lebih hemat dan kemampuan analisis komparatif lintas dokumen yang tidak dimiliki pendekatan Direct PDF konvensional. Bagi Program Studi Informatika, hasil ini mengonfirmasi bahwa topik RAG agentik, arsitektur dual-database, dan evaluasi AI merupakan kompetensi yang relevan dan dibutuhkan industri riset dan teknologi saat ini.

5.4 Saran

Berdasarkan pengalaman selama pelaksanaan PKL/Magang, terdapat beberapa saran yang dapat disampaikan kepada pihak-pihak terkait.

Bagi mahasiswa yang akan mengikuti program serupa, disarankan untuk mempersiapkan kompetensi di bidang framework LangChain, LangGraph,

serta dasar-dasar vector database sebelum memulai magang di lembaga riset. Kemampuan membaca dokumentasi teknis secara mandiri dan inisiatif belajar yang tinggi sangat krusial untuk mempercepat adaptasi di lingkungan kerja profesional.

Bagi BRIN dan instansi mitra, sinkronisasi jadwal akademik dengan siklus sprint pengembangan perangkat lunak perlu lebih ditingkatkan agar mahasiswa dapat berkontribusi secara optimal sejak awal penugasan. Pemberian pembekalan materi teknologi terkini seperti RAG, Agentic AI, dan graph database sebelum mahasiswa mulai bertugas juga akan sangat membantu memperpendek kurva pembelajaran.

Bagi Program Studi Informatika, evaluasi kinerja mahasiswa magang sebaiknya mempertimbangkan dampak teknis nyata yang dihasilkan, seperti perbaikan performa sistem atau penyelesaian modul fungsional, tidak hanya dari kelengkapan laporan. Membangun komunikasi tiga arah yang berkala antara mahasiswa, dosen pembimbing, dan mentor industri juga penting untuk menyelaraskan ekspektasi dan memastikan pengalaman magang memberikan manfaat maksimal bagi semua pihak.

DAFTAR PUSTAKA

- [1] H. Ghanadian, A. Kamali, and M. H. Tekieh, “Enhancing Scientific Literature Chatbots with Retrieval-Augmented Generation: A Performance Evaluation of Vector and Graph-Based Systems,” 2026, [Online]. Available: <http://arxiv.org/abs/2602.17856>
- [2] S. Bubeck *et al.*, “Sparks of Artificial General Intelligence: Early experiments with GPT-4,” 2023, [Online]. Available: <http://arxiv.org/abs/2303.12712>
- [3] T. T. Procko and O. Ochoa, “Graph Retrieval-Augmented Generation for Large Language Models: A Survey,” *Proc. - 2024 Conf. AI, Sci. Eng. Technol. AIXSET 2024*, pp. 166–169, 2024, doi: 10.1109/AIXSET62544.2024.00030.

LAMPIRAN



Gambar 1. Pengenalan lingkungan MBKM BRIN



Gambar 2. Silaturahmi perdana



Gambar 3. Laporan Progres Magang



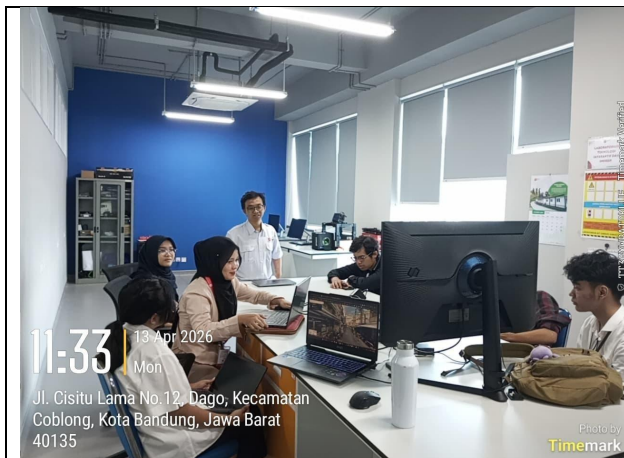
Gambar 4. Laporan progres



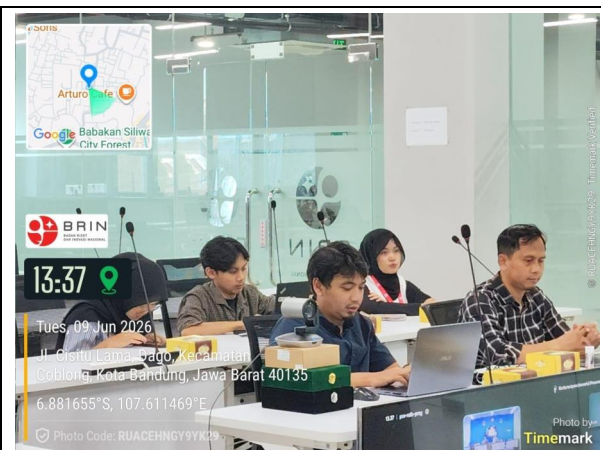
Gambar 5. Laporan Progres Magang



Gambar 6. Laporan progres



Gambar 7. Laporan progres



Gambar 8. Demonstrasi web SRE