



**LAPORAN PELAKSANAAN**  
**MERDEKA BELAJAR KAMPUS MERDEKA (MBKM)**  
**BADAN RISET DAN INOVASI NASIONAL**  
**SEMESTER GENAP 2025/2026**

**IMPLEMENTASI *MULTI-STAGE QUALITY GATE***  
**DAN *MULTI-LEVEL CHUNKING* DALAM**  
**PEMBANGUNAN *KNOWLEDGE GRAPH* LITERATUR**  
**ILMIAH UNTUK**  
**SISTEM *RESEARCH AGENTIC* AI MY-SRE**

**PENELITIAN/RISET**

**Disusun oleh:**

**Nama : Muhammad Zaidan Rafat**  
**NIM : 231730016**  
**Prodi/Jurusan : Informatika**  
**Fakultas : Sains dan Teknologi**  
**Universitas : UIN Sultan Maulana Hasanuddin Banten**  
**Pembimbing : 1. Rio Nurtantyana, S.Pd., M.Pd., M.Sc., Ph.d.**  
**2. Ahmad Tabrani, M.T.I., M.E.**

**PUSAT RISET SAINS DATA DAN INFORMASI**  
**BADAN RISET DAN INOVASI NASIONAL**  
**BANDUNG, JUNI 2026**

**LEMBAR PENGESAHAN LAPORAN PELAKSANAAN**  
**MERDEKA BELAJAR KAMPUS MERDEKA (MBKM)**  
**BADAN RISET DAN INOVASI NASIONAL**  
**SEMESTER GENAP 2025/2026**

**IMPLEMENTASI *MULTI-STAGE QUALITY GATE***  
**DAN *MULTI-LEVEL CHUNKING* DALAM**  
**PEMBANGUNAN *KNOWLEDGE GRAPH* LITERATUR**  
**ILMIAH UNTUK**  
**SISTEM *RESEARCH AGENTIC AI* MY-SRE**

**PENELITIAN/RISET**

**Disusun oleh:**

**Nama : Muhammad Zaidan Rafat**  
**NIM : 231730016**  
**Prodi/Jurusan : Informatika**  
**Fakultas : Sains dan Teknologi**  
**Universitas : UIN Sultan Maulana Hasanuddin Banten**  
**Unit Kerja BRIN : Pusat Riset Sains Data dan Informasi**

**Bandung, 16 Juni 2026**

**Menyetujui,**  
**Pembimbing BRIN**

**Mengetahui,**  
**Pembimbing Perguruan Tinggi**



**TT ELEKTRONIK**

**Rio Nurtantvana, M.Sc., Ph.d.**  
**NIP. 199312152024031001**

**Ahmad Tabrani, M.T.I., M.E**  
**NIP. 198509172018011002**

## DAFTAR ISI

|                                                                                       |     |
|---------------------------------------------------------------------------------------|-----|
| DAFTAR ISI.....                                                                       | 1i  |
| DAFTAR TABEL .....                                                                    | iii |
| DAFTAR GAMBAR .....                                                                   | iv  |
| BAB I .....                                                                           | 1   |
| PENDAHULUAN .....                                                                     | 1   |
| 1.1 Latar Belakang .....                                                              | 1   |
| 1.2 Tujuan Kegiatan MBKM .....                                                        | 3   |
| 1.3 Sasaran Kegiatan MBKM .....                                                       | 4   |
| BAB II.....                                                                           | 6   |
| PROFIL MITRA KEGIATAN MBKM.....                                                       | 6   |
| 2.2 Struktur Organisasi Badan Riset dan Inovasi Nasional .....                        | 7   |
| 2.3 Visi dan Misi Tujuan dan Sasaran Strategis Badan Riset dan Inovasi Nasional ..... | 9   |
| 2.4 Profil Satuan Kerja Tempat Pelaksanaan MBKM .....                                 | 11  |
| 2.5 Profil Kelompok Penelitian/Fungsi Kerja Tempat Pelaksanaan MBKM .....             | 13  |
| BAB III .....                                                                         | 15  |
| PELAKSANAAN KEGIATAN MBKM.....                                                        | 15  |
| 3.1 Uraian Umum Kegiatan MBKM yang Dilakukan .....                                    | 15  |
| 3.2 Keterlibatan Penelitian/Riset atau Magang/Praktik Kerja yang Dilakukan .....      | 16  |
| 3.2.1. Analisis Kebutuhan Sistem Aplikasi Brain.....                                  | 18  |
| 3.2.2. Perancangan Arsitektur dan Fondasi Teknologi Backend .....                     | 20  |
| 3.2.3. Spesifikasi Kontrak Komunikasi dan Antarmuka API .....                         | 22  |
| 3.2.4. Perancangan Skema Database.....                                                | 24  |
| 3.2.5. Implementasi Multi Level Chunking & Stuffed Context.....                       | 28  |
| 3.2.6. Strategi Prompt Engineering dan Ekstraksi Berbasis Skema.....                  | 32  |
| 3.2.7. Rekayasa Multi-Stage Semantic Gating.....                                      | 34  |

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| 3.3 Penugasan-Penugasan yang Telah Diselesaikan .....                   | 39 |
| 3.3.1. Penugasan Administratif dan Pelatihan Penguatan Kompetensi ..... | 39 |
| 3.3.4. Partisipasi Forum Ilmiah dan Demonstrasi Produk .....            | 43 |
| 3.4 Pembelajaran Hal Baru (Lessons Learned) yang Diperoleh .....        | 43 |
| BAB IV .....                                                            | 46 |
| PENUTUP.....                                                            | 46 |
| 4.1 Kesimpulan .....                                                    | 46 |
| 4.2 Saran.....                                                          | 46 |
| DAFTAR PUSTAKA .....                                                    | 48 |
| LAMPIRAN.....                                                           | 50 |

## DAFTAR TABEL

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| Tabel 3.2.1 : Kebutuhan Fungsional Sistem Backend .....                       | 20 |
| Tabel 3.2.2 : Kebutuhan Non-Fungsional Sistem Backend .....                   | 20 |
| Tabel 3.2.3 : Spesifikasi Teknologi dan Library .....                         | 22 |
| Tabel 3.2.4 : Skema API Infrastruktur .....                                   | 23 |
| Tabel 3.2.5 : Skema API Graph dan Agentic AI Services .....                   | 23 |
| Tabel 3.2.6 : Skema Model User (Manajemen Identitas) .....                    | 25 |
| Tabel 3.2.7 : Skema Model Project (Kontainer Kerja) .....                     | 25 |
| Tabel 3.2.8 : Skema Model DocumentChunk (Gudang Vektor) .....                 | 26 |
| Tabel 3.2.9 : Skema Model ChatSession dan ChatMessage .....                   | 26 |
| Tabel 3.2.10 : Skema Properti Identitas dan Metadata (Node :Literature) ..... | 27 |
| Tabel 3.2.11 : Skema Properti 5 Pilar Ilmiah (Node :Literature).....          | 27 |
| Tabel 3.2.12 : Skema Properti Relasi Semantik (Edge).....                     | 28 |
| Tabel 3.2.13 : Parameter Kalibrasi Threshold Dinamis per Pilar .....          | 36 |
| Tabel 3.2.14 : Aturan Dependensi Semantic Coherence Gate .....                | 37 |

## DAFTAR GAMBAR

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| Gambar 2.1.1 : Logo BRIN .....                                                | 7  |
| Gambar 2.2.1 : Struktur Organisasi BRIN. ....                                 | 8  |
| Gambar 2.4.1 : Struktur Organisasi Pusat Riset Sains Data dan Informasi ..... | 12 |
| Gambar 2.4.2 : Alur Kerja Pusat Riset Sains Data dan Informasi .....          | 12 |
| Gambar 2.5.1 : Profil Kelompok Riset HCI-V .....                              | 13 |
| Gambar 3.2.1 : Siklus Penelitian dan Pengembangan.....                        | 17 |
| Gambar 3.2.2 : Arsitektur Client-Server MySRE .....                           | 21 |
| Gambar 3.2.3 : ERD Postgresql (Supabase) .....                                | 24 |
| Gambar 3.2.4 : Pipeline Multi-Level Chunking .....                            | 29 |
| Gambar 3.2.5 : Snippet Code Purifikasi Data .....                             | 30 |
| Gambar 3.2.6 : Anchor Density Query .....                                     | 31 |
| Gambar 3.2.7 : Arsitektur Ekstraksi LLM .....                                 | 32 |
| Gambar 3.2.8 : Snippet Code Prompt Ekstraksi Pilar .....                      | 33 |
| Gambar 3.2.9 : Snippet Code Purification Output LLM.....                      | 34 |
| Gambar 3.2.10 : Multi-Stage Semantic Gating Architecture .....                | 35 |
| Gambar 3.2.11 : Snippet Code Domain Filtering .....                           | 38 |
| Gambar 3.2.12 : Snippet Code Edge Reasoning Quality Check .....               | 39 |
| Gambar 3.3.1 : Sertifikat Brilliant.....                                      | 40 |
| Gambar 3.3.2: Sertifikasi Fundamental Neo4j.....                              | 41 |

## **BAB I**

### **PENDAHULUAN**

#### 1.1 Latar Belakang

Volume publikasi ilmiah global telah mencapai titik jenuh yang melampaui kapasitas kognitif manusia dalam mengelola informasi. Data *National Science Board* (2024) mencatat produksi artikel sains dan teknik dunia menyentuh angka 3,3 juta dokumen pada tahun 2022, dengan akselerasi pertumbuhan sebesar 59% dalam satu dekade terakhir (Statistics (NCSES), 2023). Ledakan data ini bukan sekadar statistik, melainkan sebuah hambatan struktural yang melumpuhkan efisiensi riset. Peneliti kini menghadapi paradoks informasi: melimpahnya literatur justru mempersulit identifikasi kebaruan (*novelty*) dan pemetaan kaitan logis antar-studi. Proses tinjauan literatur konvensional yang mengandalkan pembacaan manual terhadap teks tidak terstruktur telah menjadi metode yang usang dan tidak skalabel bagi percepatan inovasi nasional.

Menghadapi krisis skalabilitas ini, arsitektur kecerdasan artifisial harus bergeser dari model bahasa generatif pasif menuju sistem otonom yang deterministik. Singh et al. (2026) menegaskan bahwa keterbatasan chatbot konvensional berakar pada alur kerja yang statis dan linear, sehingga gagal menangani penalaran multi-langkah yang kompleks dan rentan menghasilkan halusinasi pada dokumen teknis (Singh et al., 2026). Tantangan ini mendesak lahirnya Agentic AI, sebuah paradigma sistem yang memperkenalkan lapisan kontrol eksplisit untuk melakukan penalaran aktif, perencanaan tugas, dan penggunaan alat (*tool use*) secara mandiri. Dalam konteks rekayasa data, pendekatan agen menjadi sarana untuk mengonversi narasi ilmiah yang tersebar menjadi *Knowledge Graph* yang terstandarisasi. Melalui representasi entitas dan relasi semantik yang presisi, sistem ini memastikan bahwa setiap temuan riset tidak lagi berdiri sebagai fragmen data yang terisolasi, melainkan sebagai bagian dari jaringan pengetahuan yang utuh dan dapat diverifikasi..

Implementasi teknologi Agentic AI dan *Knowledge Graph* tersebut menemukan urgensi dan relevansinya di Pusat Riset Sains Data dan Informasi

(PR SDI), Badan Riset dan Inovasi Nasional (BRIN). Sebagai motor penggerak kedaulatan data nasional, PR-SDI melalui kelompok riset Interaksi Manusia-Komputer dan Visualisasi memiliki visi strategis untuk mentransformasi cara peneliti berinteraksi dengan informasi ilmiah yang kompleks. Melalui integrasi program Merdeka Belajar Kampus Merdeka (MBKM), visi tersebut diwujudkan menjadi aksi rekayasa yang konkret. Program ini tidak hanya berfungsi sebagai medium pembelajaran, tetapi juga sebagai jembatan untuk mensinkronkan kemajuan teori akademik dengan kebutuhan infrastruktur riset nasional. Kolaborasi ini menjadi pondasi bagi pembangunan *Smart Research Environment* (SRE) yang mampu mengakomodasi kecepatan akselerasi pengetahuan di Indonesia.

Namun, hambatan teknis yang paling persisten dalam mewujudkan visi tersebut terletak pada fragmentasi informasi di dalam artikel ilmiah itu sendiri. Pengetahuan riset saat ini terkunci di dalam narasi teks tidak terstruktur, sebuah kondisi yang didefinisikan oleh Lu et al. (2026) sebagai pelebaran jarak (*widening gap*) yang menjadi hambatan utama bagi pembentukan sistem pengetahuan (Lu et al., 2026). Hal ini menyulitkan identifikasi kaitan logis antara Latar Belakang, Metodologi, Tujuan, Gap (Kesenjangan), hingga Arahan Masa Depan antar-dokumen secara otomatis. Ketiadaan mekanisme sintesis semantik menyebabkan proses pencarian “benang merah” penelitian tetap menjadi beban manual yang repetitif dan rentan terhadap ketidakteraturan data. Tantangan ini bukan sekadar masalah ekstraksi kata kunci, melainkan kegagalan sistemik dalam memetakan posisi sebuah penelitian di dalam peta pengetahuan (*knowledge map*) yang lebih luas secara akurat dan konsisten.

Guna memutus rantai inefisiensi tersebut, dikembangkanlah proyek SAAI (*Service Agentic AI*) melalui produk MySRE (*My Smart Research Environment*) dengan fokus utama pada fitur “*Brain*”. Sistem ini dirancang untuk melakukan dekomposisi terhadap narasi artikel ilmiah dan merekonstruksinya menjadi *Knowledge Graph* yang merepresentasikan keterkaitan antar-pilar riset secara deterministik. Dalam kerangka kerja ini, peran *Backend & Knowledge Data Engineer* menjadi krusial untuk membangun

fondasi arsitektur yang mampu menangani alur pemrosesan data mulai dari *ingestion*, *smart chunking*, hingga pemodelan relasi semantik yang stabil. Dengan menghadirkan infrastruktur backend yang andal, diharapkan aplikasi “*Brain*” tidak hanya menjadi alat bantu administratif, melainkan sebuah ekosistem otak buatan yang mampu mengasimilasi pengetahuan riset secara otonom bagi kemajuan sains di Indonesia.

## 1.2 Tujuan Kegiatan MBKM

Tujuan utama dari pelaksanaan program MBKM di Badan Riset dan Inovasi Nasional (BRIN) adalah menyelaraskan visi strategis Pusat Riset Sains Data dan Informasi (PR-SDI), khususnya pada kelompok riset *Human Computer Interaction - Visualization*, ke dalam aksi rekayasa data yang konkret. Target ini mencakup pembangunan fondasi *Knowledge Graph* menggunakan Neo4j untuk memetakan keterkaitan semantik antar-literatur ilmiah secara sistematis. Melalui pemodelan graf yang terstruktur, sistem diharapkan mampu memberikan transparansi data yang lebih tinggi bagi para peneliti dalam meninjau perkembangan ilmu pengetahuan secara komprehensif.

Secara teknis, kegiatan ini bertujuan merancang arsitektur backend yang resilien berbasis FastAPI dan Python untuk menangani alur *ingestion* serta *chunking* data secara stabil. Fokus rekayasa diarahkan pada otomatisasi ekstraksi 5 Pilar Penelitian, yang meliputi Latar Belakang, Metodologi, Tujuan, Gap (Kesenjangan), dan Arahana Masa Depan, dengan tingkat akurasi yang terkalibrasi secara objektif. Langkah ini krusial untuk menyintesis relasi antar-studi dari teks tidak terstruktur guna meniadakan hambatan repetitif dalam tinjauan literatur manual di lingkungan riset nasional.

Pada aspek pengembangan kapasitas profesional, program ini menjadi sarana utama untuk memperkuat kompetensi penulis sebagai *Backend* dan *Knowledge Data Engineer*. Melalui keterlibatan langsung dalam ekosistem penelitian nasional di BRIN, penulis bertujuan menginternalisasi standar integritas riset serta kemampuan kolaborasi teknis tingkat lanjut. Pencapaian tujuan ini diharapkan dapat membentuk basis SDM yang tangguh dalam membangun infrastruktur data riset yang berdaulat di Indonesia.

### 1.3 Sasaran Kegiatan MBKM

Sasaran utama dari kegiatan MBKM di Badan Riset dan Inovasi Nasional ini adalah transformasi kapasitas penulis sebagai mahasiswa Informatika dari tataran teoretis menuju praktik rekayasa data (Knowledge Engineering) yang resilien di lingkungan profesional. Melalui internalisasi dalam Kelompok Riset Human Computer Interaction - Visualization pada Pusat Riset Sains Data dan Informasi (PR-SDI), program ini menyasar penguatan standar kompetensi mahasiswa dalam mengelola arsitektur data yang kompleks dan terukur. Fokus pengembangan diarahkan pada kesiapan penulis untuk beroperasi secara mandiri sebagai entitas teknis yang mampu mensinkronkan visi kelompok riset dengan solusi infrastruktur digital yang berdaulat.

Secara fungsional, sasaran teknis difokuskan pada kontribusi mahasiswa dalam penyelesaian lingkup pengembangan aplikasi MySRE (“Brain”). Sistem ini ditargetkan mampu menjembatani problem fragmentasi informasi melalui mekanisme ekstraksi otonom terhadap lima pilar utama penelitian, yakni Latar Belakang, Metodologi, Tujuan, Gap (Kesenjangan), dan Arahana Masa Depan. Dengan terwujudnya alur pemrosesan data yang stabil, sasaran akhir dari keterlibatan penulis adalah terciptanya repositori pengetahuan yang tidak lagi bersifat atomik, melainkan terintegrasi secara semantik dalam jaringan Knowledge Graph yang dapat dieksplorasi secara cerdas oleh para peneliti di BRIN. Sebagai parameter keberhasilan, program ini menetapkan beberapa luaran konkret yang harus diselesaikan oleh penulis, antara lain:

1. Arsitektur Backend dan Skema Knowledge Graph:

Dokumentasi teknis dan kode sumber berbasis FastAPI dan Neo4j yang telah tervalidasi secara fungsional.

2. Pipeline Agentic AI:

Implementasi mesin *ingestion* otonom yang mampu mengekstraksi komponen riset dari dokumen tidak terstruktur dengan akurasi yang terkalibrasi.

3. Draf Artikel Ilmiah:

Naskah publikasi akademis mengenai rekayasa sistem Brain sebagai

kontribusi mahasiswa terhadap literatur sains data nasional.

4. Laporan Akhir dan Dokumentasi Administratif:

Dokumen komprehensif serta catatan aktivitas harian yang mencerminkan seluruh proses pelaksanaan magang di lingkungan PR-SDI BRIN.

## **BAB II**

### **PROFIL MITRA KEGIATAN MBKM**

#### **2.1 Sejarah Badan Riset dan Inovasi Nasional**

Ekosistem riset di Indonesia sebelumnya tersebar pada berbagai instansi, meliputi Lembaga Ilmu Pengetahuan Indonesia (LIPI), Badan Pengkajian dan Penerapan Teknologi (BPPT), Badan Tenaga Nuklir Nasional (BATAN), Lembaga Penerbangan dan Antariksa Nasional (LAPAN), dan satuan kerja penelitian sektoral di berbagai kementerian. Integrasi sumber daya riset nasional bermula dari pengesahan Undang-Undang Nomor 11 Tahun 2019 tentang Sistem Nasional Ilmu Pengetahuan dan Teknologi. Pasal 48 regulasi ini mewajibkan pembentukan sebuah badan riset nasional untuk menyatukan kegiatan penelitian, pengembangan, pengkajian, dan penerapan, serta invensi dan inovasi. Pemerintah merealisasikan amanat tersebut melalui Peraturan Presiden Nomor 74 Tahun 2019. Pada fase pembentukan awal ini, Badan Riset dan Inovasi Nasional (BRIN) belum memiliki otonomi penuh dan beroperasi di bawah naungan Kementerian Riset dan Teknologi.

Struktur kelembagaan BRIN berubah drastis pada pertengahan 2021. Presiden menetapkan pemisahan BRIN dari Kementerian Riset dan Teknologi untuk membentuk lembaga pemerintah nonkementerian yang berdiri sendiri. Landasan hukum pemisahan ini tertuang dalam Peraturan Presiden Nomor 33 Tahun 2021. Kebijakan tersebut menginstruksikan peleburan empat lembaga riset utama, yakni LIPI, BPPT, BATAN, dan LAPAN, beserta seluruh unit penelitian dari kementerian dan lembaga negara lainnya ke dalam BRIN. Konsolidasi ini mengakhiri fragmentasi riset yang memecah pendanaan dan sumber daya manusia selama beberapa dekade.

Pemerintah menyempurnakan landasan operasional institusi melalui Peraturan Presiden Nomor 78 Tahun 2021. Regulasi ini menetapkan BRIN sebagai entitas tunggal yang berada dibawah dan bertanggung jawab langsung kepada Presiden. BRIN kini memegang kewenangan mutlak untuk menyelenggarakan penelitian, pengembangan, pengkajian, dan penerapan

secara terintegrasi. Institusi ini juga mewarisi mandat spesifik penyelenggaraan ketenaganukliran dari eks-BATAN dan penyelenggaraan keantariksaan dari eks-LAPAN. Peraturan Presiden ini secara definitif memposisikan BRIN sebagai pusat kendali utama bagi arah kebijakan, pendanaan, dan tata kelola infrastruktur riset di Indonesia.



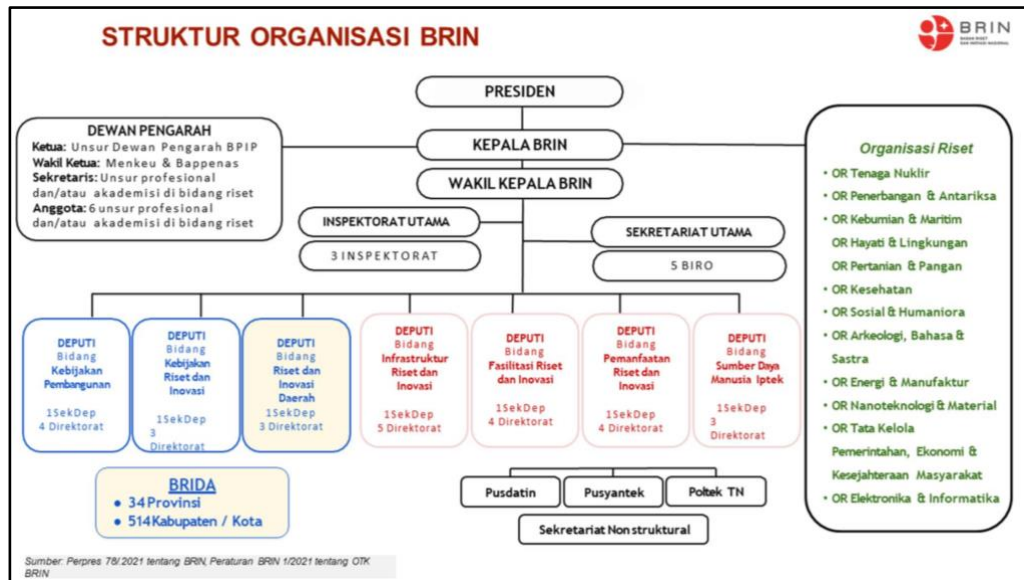
Gambar 2.1.1 : Logo BRIN

Identitas visual BRIN resmi diluncurkan pada 10 Agustus 2021, bertepatan dengan peringatan Hari Kebangkitan Teknologi Nasional ke-26. Logo institusi dirancang bukan sekadar sebagai simbol administrasi, melainkan sebagai manifestasi visual dari integrasi multi-disiplin ilmu (Alinea, 2021). Logogram BRIN tersusun atas lima elemen vektor yang membentuk sebuah bulatan utuh, dibalut warna merah yang merepresentasikan keberanian bangsa untuk memajukan inovasi (Permana, 2021). Kelima elemen tersebut memetakan cakupan ekosistem riset nasional: figur manusia menegaskan bahwa riset bertumpu pada kesejahteraan masyarakat; bentuk bintang merepresentasikan eksplorasi keantariksaan dan pencapaian ilmu pengetahuan; elemen daun mewakili kekayaan botani; siluet sirip ikan melambangkan kekayaan maritim dan zoologi; sementara garis melingkar mengikat seluruh elemen biodiversitas. Secara geometris, bentuk sirkular logo ini mengunci makna peleburan mutlak seluruh entitas riset ke dalam satu institusi terpusat (suara.com, 2021).

## 2.2 Struktur Organisasi Badan Riset dan Inovasi Nasional

Tata kelola struktural Badan Riset dan Inovasi Nasional (BRIN) berpijak pada Peraturan Presiden Nomor 78 Tahun 2021. Regulasi ini merancang hierarki institusi untuk mengakomodasi integrasi entitas riset nasional. Pada tingkat pimpinan, Dewan Pengarah memegang wewenang untuk merumuskan arahan strategis dan kebijakan makro. Di ranah eksekutif, operasional institusi dikendalikan oleh Kepala BRIN yang berkedudukan sebagai pengambil

kebijakan tertinggi sekaligus pengguna anggaran. Dalam mengeksekusi tugasnya, Kepala BRIN didampingi oleh seorang Wakil Kepala BRIN yang bertugas mengoordinasikan pencapaian strategis lintas unit kerja.



Gambar 2.2.1 : Struktur Organisasi BRIN.

Untuk menjamin akuntabilitas institusi, struktur pimpinan ditopang oleh unsur pengawas dan unsur tata usaha. Inspektorat Utama menjalankan fungsi audit dan pengawasan internal, bertugas memastikan seluruh kegiatan riset, pembiayaan, dan tata laksana kepatuhan beroperasi sesuai dengan regulasi perundang-undangan. Sementara itu, urat nadi administrasi institusi dikelola oleh Sekretariat Utama. Unit ini mengintegrasikan sistem penyusunan perencanaan program, mengelola sumber daya manusia aparatur, membina tata keuangan, serta menyelenggarakan fungsi kerumahtanggaan, kearsipan, hingga kehumasan bagi seluruh instansi.

Fungsi pelaksana utama BRIN didistribusikan kepada tujuh kedeputian yang terbagi dalam tiga domain besar pengelolaan ekosistem riset. Domain perumusan kebijakan ditangani oleh Deputi Bidang Kebijakan Pembangunan serta Deputi Bidang Kebijakan Riset dan Inovasi. Domain penyediaan sumber daya dan infrastruktur diampu oleh tiga unit, yakni Deputi Bidang Sumber Daya Manusia Ilmu Pengetahuan dan Teknologi, Deputi Bidang Infrastruktur Riset dan Inovasi, serta Deputi Bidang Fasilitas Riset dan Inovasi. Pada tahap akhir,

domain hilirisasi dan desentralisasi menjadi tanggung jawab Deputy Bidang Pemanfaatan Riset dan Inovasi serta Deputy Bidang Riset dan Inovasi Daerah, yang bertugas memastikan hasil penelitian terapan di sektor industri dan pemerintah daerah.

Ujung tombak kegiatan penelitian dikelola melalui struktur organisasi hierarkis yang mengklasifikasikan disiplin keilmuan. Makro-disiplin ilmu pengetahuan dipayungi oleh institusi yang disebut Organisasi Riset (OR), sebagai contoh Organisasi Riset Elektronika dan Informatika. Untuk keperluan operasional yang lebih spesifik, setiap Organisasi Riset menaungi berbagai Pusat Riset (PR). Pusat Riset bertindak sebagai dapur teknis tempat para peneliti, perekayasa, dan teknisi mengeksekusi langsung kegiatan penelitian dasar dan terapan, seperti halnya Pusat Riset Sains Data dan Informasi.

Arsitektur kelembagaan BRIN ditutup oleh keberadaan unsur pendukung teknis yang beroperasi secara terpusat. Unit layanan sentral ini mencakup Pusat Data dan Informasi, yang memiliki mandat untuk membangun dan memelihara pangkalan data riset nasional. Selain itu, terdapat Pusat Pelayanan Teknologi yang secara spesifik menangani urusan komersial seperti manajemen kontrak, proyek kemitraan, dan tata kelola lisensi atas produk-produk inovasi yang dihasilkan oleh para peneliti.

### **2.3 Visi dan Misi Tujuan dan Sasaran Strategis Badan Riset dan Inovasi Nasional**

Arah strategis Badan Riset dan Inovasi Nasional (BRIN) berpedoman pada kerangka pembangunan jangka panjang nasional. Sebagai otoritas tunggal penyelenggara penelitian dan inovasi, institusi ini memformulasikan target kinerjanya agar selaras dengan agenda prioritas pemerintah pusat. Arsitektur perencanaan ini tertuang secara hierarkis melalui penetapan visi, misi, tujuan, dan sasaran strategis, yang secara komprehensif dirancang untuk mendukung pencapaian cita-cita Indonesia Emas 2045.

#### **1. Visi**

Visi Presiden dan Wakil Presiden Republik Indonesia

"Bersama Indonesia Maju Menuju Indonesia Emas 2045"

## 2. Misi

- a) Memperkokoh ideologi Pancasila, demokrasi, dan hak asasi manusia (HAM);
- b) Memantapkan sistem pertahanan keamanan negara dan mendorong kemandirian bangsa melalui swasembada pangan, energi, air, ekonomi syariah, ekonomi digital, ekonomi hijau, dan ekonomi biru;
- c) Melanjutkan pengembangan infrastruktur dan meningkatkan lapangan kerja yang berkualitas, mendorong kewirausahaan, mengembangkan industri kreatif serta mengembangkan agromaritim industri di sentra produksi melalui peran aktif koperasi;
- d) Memperkuat pembangunan Sumber Daya Manusia (SDM), sains, teknologi, pendidikan, kesehatan, prestasi olahraga, kesetaraan gender, serta penguatan peran perempuan, pemuda (generasi milenial dan generasi Z), dan penyandang disabilitas;
- e) Melanjutkan hilirisasi dan mengembangkan industri berbasis sumber daya alam untuk meningkatkan nilai tambah di dalam negeri;
- f) Membangun dari desa dan dari bawah untuk pertumbuhan ekonomi, pemerataan ekonomi, dan pemberantasan kemiskinan;
- g) Memperkuat reformasi politik, hukum, dan birokrasi, serta memperkuat pencegahan dan pemberantasan korupsi, narkoba, judi, dan penyelundupan; dan
- h) Memperkuat penyelarasan kehidupan yang harmonis dengan lingkungan, alam dan budaya, serta peningkatan toleransi antarumat beragama untuk mencapai masyarakat yang adil dan makmur.

## 3. Tujuan

- a) Terwujudnya temuan, terobosan dan pembaharuan ilmu pengetahuan dari hasil penelitian, pengembangan, pengkajian dan penerapan, serta invensi dan inovasi, penyelenggaraan ketenaganukliran, dan penyelenggaraan keantariksaan dalam rangka peningkatan produktivitas dan daya saing, peningkatan kualitas lingkungan hidup dan ketahanan bencana, serta iklim (T1)

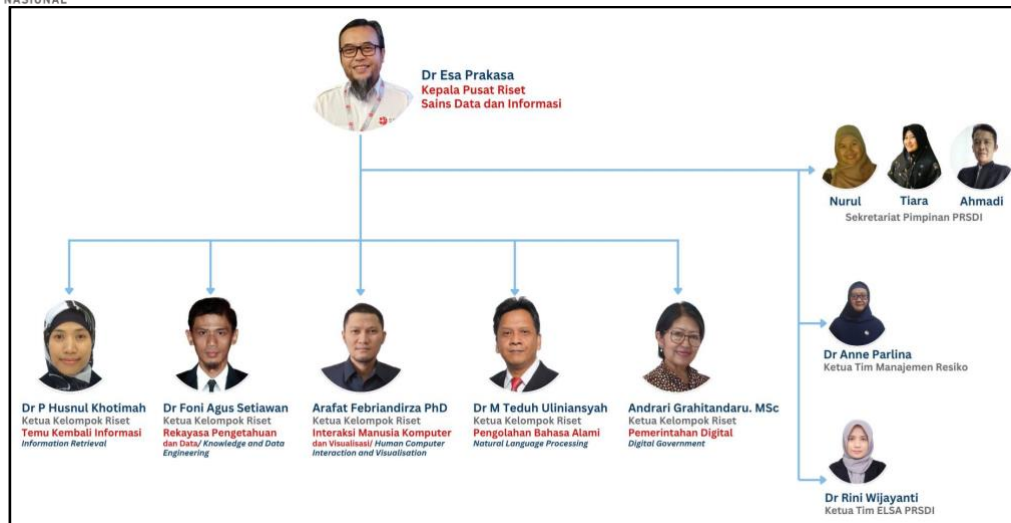
- b) Terwujudnya sumber daya manusia, infrastruktur, fasilitasi dan pemanfaatan riset dan inovasi yang unggul dan kompetitif (T2)
- c) Terwujudnya Tata Kelola Pemerintahan di Badan Riset dan Inovasi Nasional yang baik dan bersih (T3)

#### 4. Sasaran Strategis

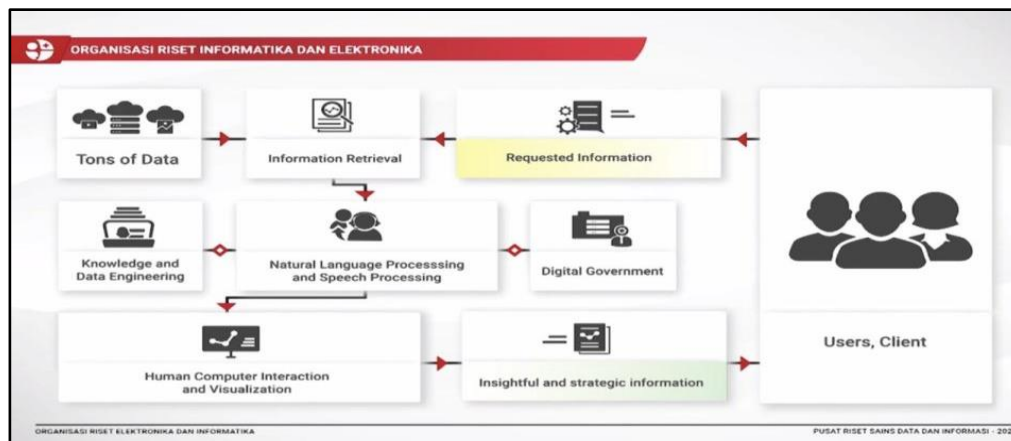
- a) Meningkatnya keunggulan riset dan inovasi ilmu pengetahuan dan teknologi, serta dapat dijadikan kebijakan berbasis bukti yang selaras dengan arah pembangunan berkelanjutan (SS1)
- b) Meningkatnya kolaborasi dalam pengembangan dan pemanfaatan produk ilmu pengetahuan berdasarkan prioritas pembangunan berkelanjutan (SS2)
- c) Meningkatnya produktivitas dan daya saing sumber daya riset dan inovasi BRIN (SS3)
- d) Meningkatnya penerapan ilmu pengetahuan dan teknologi untuk mendukung kualitas lingkungan hidup, ketahanan bencana, dan kerentanan iklim (SS4)
- e) Tata kelola BRIN yang efektif, efisien dan akuntabel (SS5)

#### 2.4 Profil Satuan Kerja Tempat Pelaksanaan MBKM

Pusat Riset Sains Data dan Informasi (PRSDI) beroperasi sebagai subsistem pelaksana teknis di bawah naungan Organisasi Riset Elektronika dan Informatika (OREI). Unit kerja ini dipimpin oleh Dr. Esa Prakasa, M.T., selaku Kepala PRSDI, dan menjalankan aktivitas operasionalnya dari Gedung BASIC Lantai 5, Kawasan Sains dan Teknologi (KST) Samaun Samadikun, Bandung. Dasar hukum pembentukan entitas ini disahkan melalui Peraturan Badan Riset dan Inovasi Nasional Nomor 10 Tahun 2022 tentang Tugas, Fungsi, dan Struktur Organisasi Riset Elektronika dan Informatika. Struktur kelembagaan PRSDI direpresentasikan secara visual pada Gambar 2.3, sedangkan mekanisme kerja teknisnya tertuang pada Gambar 2.4



Gambar 2.4.1 : Struktur Organisasi Pusat Riset Sains Data dan Informasi



Gambar 2.4.2 : Alur Kerja Pusat Riset Sains Data dan Informasi

Kewenangan operasional PRSDI diatur secara eksplisit dalam Pasal 14 regulasi tersebut. Pasal ini memandatkan unit terkait untuk memimpin penyelenggaraan penelitian teknis, mengelola pengkajian, serta merumuskan strategi penerapan metodologi. Tanggung jawab teknis ini mewajibkan PRSDI untuk secara konsisten menghasilkan invensi dan inovasi yang berfokus pada ranah sains data dan informasi.

Eksekusi tugas pokok tersebut didukung oleh lima fungsi spesifik yang dijabarkan dalam Pasal 15. PRSDI mengemban fungsi pelaksanaan teknis penelitian, yang berjalan beriringan dengan penyiapan bahan rekomendasi maupun tanggapan ilmiah bagi keperluan pimpinan. Untuk menjamin kelancaran prosedur di lapangan, unit ini diamanatkan menyediakan bimbingan

teknis serta melakukan pengawasan melalui proses supervisi. Serangkaian tahapan ini membutuhkan kerangka kerja sama lintas sektor yang bermuara pada kegiatan pemantauan, evaluasi, serta penyusunan laporan, guna memastikan seluruh luaran riset sesuai dengan target capaian institusi pusat.

## 2.5 Profil Kelompok Penelitian/Fungsi Kerja Tempat Pelaksanaan MBKM

Kelompok Riset Interaksi Manusia-Komputer dan Visualisasi (*Human-Computer Interaction and Visualisation*) merupakan unit riset spesifik yang beroperasi di bawah naungan Pusat Riset Sains Data dan Informasi (PRSDI). Kelompok ini diketuai oleh Arafat Febriandirza, Ph.D., yang didukung oleh sembilan anggota peneliti lainnya, yaitu Rio Nurtantyana, S.Pd., M.Pd., M.Sc., Ph.D., Abdurrahman Prasetyadi, M.P., Rumadi, S.T., M.T., Halim Hamadi, M.Sc., Aang Gunawan Sutyawan, S.Kom., M.T., Christine Cecylia Munthe, S.Kom., M.T.I., Muhammad Yudhi Rezaldi, Ph.D., Ridwan Suhud, S.Kom., M.T., dan Ana Hadiana. Profil representasi dari kelompok riset ini dapat dilihat pada Gambar 2.5.



Gambar 2.5.1 : Profil Kelompok Riset HCI-V

Ruang lingkup keilmuan pada kelompok riset ini dibangun dari spesialisasi multidisiplin para anggotanya. Fokus utamanya mengkaji Human-

Computer Interaction, Simulation, serta perancangan *User Interface* (UI) dan *User Experience* (UX) guna merumuskan antarmuka sistem yang andal. Area studi ini meluas pada terapan spesifik seperti rekayasa interaksi untuk edukasi (*HCI-E for Education*) dan arsitektur sistem informasi. Seluruh rangkaian kajian tersebut ditopang secara fundamental oleh keahlian esensial di bidang Data Science, Information Science, dan Artificial Intelligence.

Melalui integrasi keilmuan tersebut, kelompok riset ini memfokuskan tujuannya untuk mentransformasi kompleksitas teknis menjadi bentuk komputasi spasial dan visual yang rasional. Visi ini diimplementasikan dengan menghadirkan solusi interaktif pada ekosistem berskala masif, seperti pemodelan visual pada sistem kendali (*control system*) dan simulasi fasilitas pembangkit nuklir. Eksplorasi ini senantiasa didorong oleh pemanfaatan inovatif teknologi imersif, seperti *Augmented Reality* (AR) dan *Virtual Reality* (VR), dengan objektif akhir menghasilkan interaksi manusia dengan mesin yang presisi, aman, dan efisien.

## BAB III

### PELAKSANAAN KEGIATAN MBKM

#### 3.1 Uraian Umum Kegiatan MBKM yang Dilakukan

Program MBKM Riset/Penelitian ini diselenggarakan oleh Direktorat Manajemen Talenta Badan Riset dan Inovasi Nasional (BRIN) dengan durasi pelaksanaan selama enam bulan, terhitung sejak 18 Februari hingga 30 Juli 2026. Penulis melaksanakan tugas berdasarkan ketetapan Surat Keputusan (SK) penugasan pada Pusat Riset Sains Data dan Informasi (PR SDI). Seluruh aktivitas operasional dilakukan secara tatap muka (*Onsite/Work From Office*) di Gedung BASIC, Kawasan Sains dan Teknologi (KST) Samaun Samadikun, Bandung, di bawah supervisi tim Kelompok Riset *Human Computer Interaction – Visualization*.

Pelaksanaan kegiatan dimulai dengan fase penguatan kompetensi riset melalui serangkaian pelatihan pada platform BRILian. Penulis menyelesaikan modul-pelatihan mandiri yang mencakup teknik penelusuran referensi ilmiah secara sistematis, metodologi penulisan karya tulis ilmiah bagi talenta muda, hingga prosedur teknis pengambilan sampel data. Sebagai upaya pendalaman spesialisasi pada domain *Knowledge Data*, penulis juga menempuh pendidikan mandiri di Neo4j Academy untuk menguasai fundamental basis data graf. Perolehan sertifikasi dari kedua institusi tersebut menjadi landasan teoretis sebelum penulis terlibat langsung dalam rekayasa sistem.

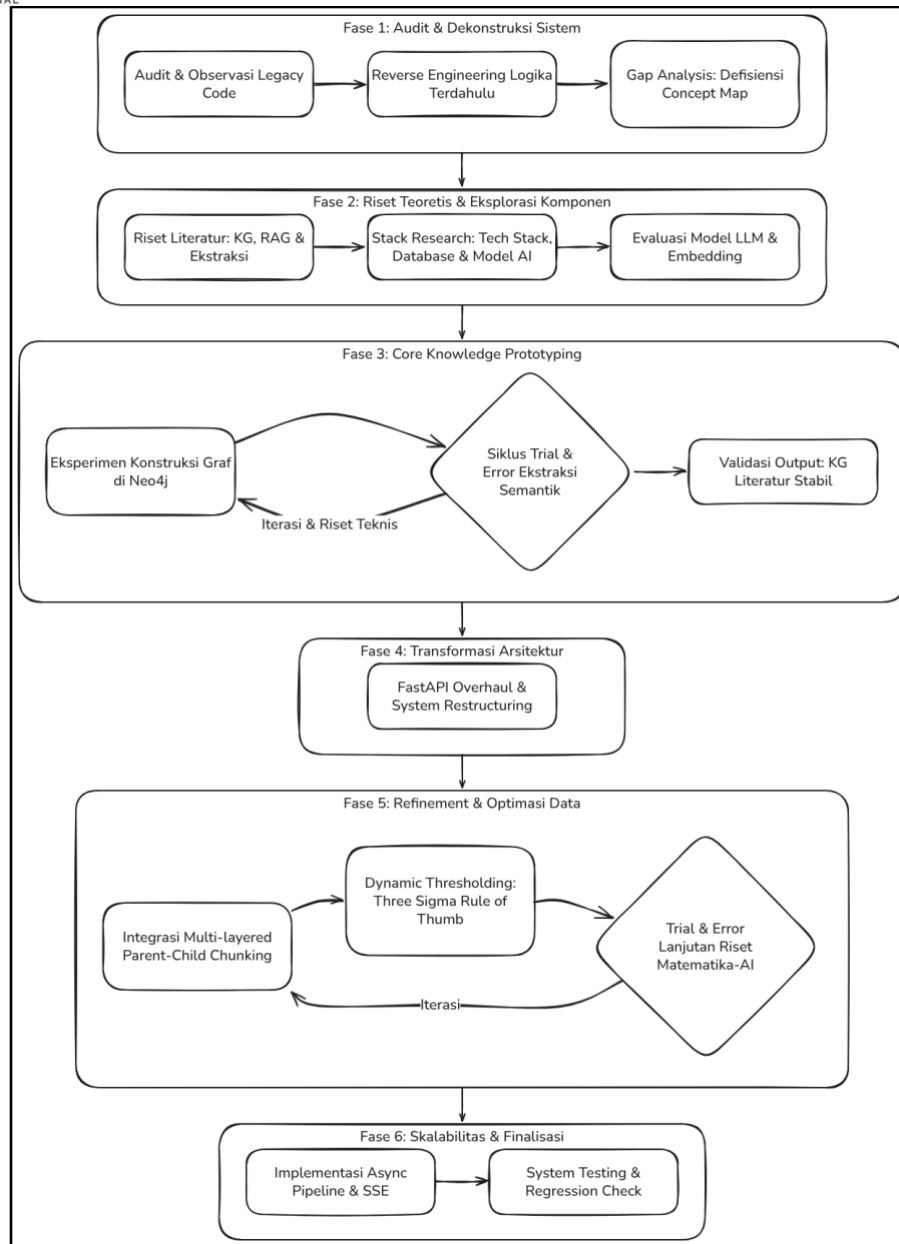
Mekanisme kerja harian dijalankan melalui siklus koordinasi yang terukur. Penulis berpartisipasi dalam pertemuan rutin (*progress meeting*) setiap hari Senin bersama pembimbing internal dari sivitas BRIN untuk menyelaraskan target mingguan. Dari sisi administrasi, penulis mendokumentasikan rincian pekerjaan dan capaian harian secara kronologis pada portal ELSA BRIN ([elsa.brin.go.id](https://elsa.brin.go.id)) Integrasi antara pertemuan tatap muka dan pelaporan digital ini menjamin adanya transparansi serta akuntabilitas performa selama masa magang.

Dalam aspek teknis, penulis menerapkan metodologi pengerjaan berbasis *Literature-Driven Development*. Aktivitas rutin mencakup penelusuran literatur dan artikel ilmiah terkini untuk mengomparasikan berbagai algoritma pembangunan arsitektur MySRE, serta melakukan riset mendalam terhadap stack teknologi yang akan digunakan, seperti evaluasi efisiensi graph database dan agentic framework. Seluruh manajemen tugas, perumusan ide, hingga penyusunan timeline teknis dikelola secara mandiri melalui workspace Notion. Pendekatan proaktif ini memungkinkan penulis untuk mengajukan inisiatif improvisasi arsitektur yang tetap bersandar pada bukti ilmiah dan validasi teknis.

### **3.2 Keterlibatan Penelitian/Riset atau Magang/Praktik Kerja yang Dilakukan**

Penulis bertugas sebagai Backend & Knowledge Data Engineer pada proyek SAAI (*Service Agentic AI*) MySRE (*My Smart Research Environment*) di Kelompok Riset *Human Computer Interaction - Visualisation*. Fokus pekerjaan terletak pada perubahan struktural sistem pengolahan literatur ilmiah. Penulis mengganti model pemetaan konsep lama (*Legacy Concept Map*) yang bersifat non-persisten menjadi arsitektur *Agentic Knowledge Graph* dengan skema hybrid Neo4j dan Supabase. Rekayasa ini membangun SAAI Engine sebagai penggerak utama aplikasi Brain untuk mengekstraksi dan menautkan entitas pengetahuan dari artikel ilmiah secara sistematis.

Pengembangan sistem mengikuti metodologi evolusioner dan iteratif dalam enam fase sebagaimana tercantum pada Gambar 3.2.1 Penulis memulai dengan audit dan dekonstruksi kode warisan (*legacy code*) untuk memetakan kelemahan alur kerja terdahulu. Temuan mengenai risiko halusinasi pada metode injeksi data langsung ke LLM mendorong penulis melakukan riset teoretis mendalam tentang *Knowledge Graph* dan Agentic RAG. Siklus uji coba (*trial and error*) pada fase *prototyping* di Neo4j dilakukan untuk mengkalibrasi logika ekstraksi semantik sebelum penulis mengonsolidasikan arsitektur ke dalam kerangka kerja FastAPI.



Gambar 3.2.1 : Siklus Penelitian dan Pengembangan

Tahap pengerjaan yang signifikan terjadi pada fase refinement. Penulis mengintegrasikan teknik Multi-layered Parent-Child Chunking untuk memperkaya konteks data. Selain itu, penulis menerapkan perhitungan ambang batas (threshold) KNN yang terkalibrasi melalui prinsip statistik Three Sigma Rule of Thumb. Secara teknis, keterlibatan penulis terbagi dalam dua bidang fokus utama:

1. Rekayasa Infrastruktur Backend dan Orkestrasi Data

Penulis membangun infrastruktur server yang responsif melalui Asynchronous Ingestion Pipeline berbasis FastAPI. Arsitektur ini memisahkan proses komputasi berat ke dalam background tasks guna menjaga stabilitas layanan. Transparansi proses bagi pengguna disediakan melalui integrasi Server-Sent Events (SSE) untuk pelaporan progres secara real-time. Keamanan data dikelola melalui isolasi multi-tenant dengan identitas komposit, didukung operasi penghapusan atomik (Double Kill deletion) untuk menjaga sinkronisasi antara aset fisik dalam penyimpanan dan representasi graf dalam basis data.

## 2. Knowledge Engineering dan Inteligensi Graf

Penulis merancang pemodelan pengetahuan melalui skema 5 Pilar Ilmiah (Latar Belakang, Metodologi, Tujuan, Gap, dan Arahana Masa Depan) pada Neo4j. Integritas relasi antar-literatur dijaga melalui mekanisme Multi-Stage Semantic Gating yang terdiri atas empat lapisan validasi: Retrieval Gate untuk validasi ambang batas KNN, *Coherence Gate* untuk konsistensi hubungan lintas pilar, *Dynamic Domain Filtering Gate* untuk memastikan relevansi bidang akademik melalui penilaian LLM, serta *Edge Reasoning Gate* untuk menjamin kualitas narasi relasi. Logika matematika-statistik diterapkan untuk memastikan setiap tautan dalam graf bersandar pada kemiripan substansial yang terukur secara ilmiah.

### 3.2.1. Analisis Kebutuhan Sistem Aplikasi Brain

Analisis kebutuhan sistem ini bertujuan memetakan batasan operasional dan kapabilitas teknis dari backend engine yang menggerakkan aplikasi Brain. Perumusan kebutuhan ini didasarkan pada hasil audit terhadap kode warisan (*legacy code*), yang secara sistematis menutup celah arsitektur lama dengan pendekatan baru berbasis Knowledge Graph dan Agentic AI.

Kategorisasi kebutuhan sistem dibagi menjadi dua domain. Kebutuhan Fungsional difokuskan pada siklus hidup data, mencakup alur dari ekstraksi dokumen mentah (PDF) hingga distribusi *Knowledge Graph* terstruktur. Sementara itu, Kebutuhan Non-Fungsional menitikberatkan

pada pemenuhan standar industri sistem Agentic AI, seperti penerapan keamanan data melalui isolasi multi-tenant, penjagaan integritas data dengan jaminan idempotensi, serta transparansi performa melalui eksekusi asinkron dan *Server-Sent Events* (SSE). Rincian spesifikasi kebutuhan sistem yang menjadi landasan rekayasa arsitektur diuraikan pada Tabel 3.2.1 dan Tabel 3.2.2 berikut :

| No.  | Fitur / Modul                                      | Deskripsi Kebutuhan                                                                                                                                                                                                 |
|------|----------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| F-01 | Ekstraksi Teks (Parsing)                           | Sistem harus menerima dokumen PDF (digital maupun pindaian) dan mengonversinya ke format teks murni (Markdown), dengan kemampuan menangani tata letak kolom tunggal maupun multi-kolom.                             |
| F-02 | Sterilisasi Data                                   | Sistem harus mengidentifikasi dan membuang teks yang tidak relevan dengan substansi penelitian (seperti daftar pustaka, ucapan terima kasih, dan lampiran) sebelum proses vektorisasi.                              |
| F-03 | Retensi Semantik Struktural                        | Sistem harus mengikat struktur judul atau sub-bab asli dari jurnal (metadata_header) pada setiap potongan teks agar hierarki informasi tidak hilang saat pemecahan.                                                 |
| F-04 | Pemecahan Hierarkis (Multi Layered Chunking)       | Sistem harus memecah teks menjadi dua lapisan: Child Chunks berukuran kecil untuk efisiensi komputasi pencarian kemiripan, dan Parent Chunks berukuran besar untuk mempertahankan konteks utuh bagi pemrosesan LLM. |
| F-05 | Ekstraksi Pilar Pengetahuan                        | Sistem harus mengekstraksi teks terkurasi ke dalam 5 format JSON baku yang mewakili Latar Belakang, Metodologi, Tujuan, <i>Gap</i> , dan Arahana Masa Depan.                                                        |
| F-06 | Penyimpanan Aset & Vektor                          | Sistem harus menyimpan dokumen mentah ke dalam <i>cloud storage</i> dan mengamankan potongan teks beserta nilai <i>embedding</i> vektornya ke dalam basis data relasional (Supabase).                               |
| F-07 | Konstruksi <i>Knowledge Graph</i>                  | Sistem harus mengubah 5 pilar pengetahuan menjadi entitas ( <i>Node</i> ) dan menghitung relasinya terhadap literatur lain ( <i>Edge</i> ) di dalam basis data graf (Neo4j).                                        |
| F-08 | Penghapusan Atomik ( <i>Double-Kill Deletion</i> ) | Sistem harus menyediakan fungsi penghapusan menyeluruh, di mana penghapusan satu literatur akan menghapus Node di Neo4j sekaligus melenyapkan data vektor dan berkas fisik di Supabase.                             |

|      |                      |                                                                                                                                                              |
|------|----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| F-09 | Distribusi Data Graf | Sistem harus menyediakan layanan (endpoint) untuk mendistribusikan data node dan edge secara terstruktur agar dapat divisualisasikan oleh aplikasi frontend. |
|------|----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|

*Tabel 3.2.1 : Kebutuhan Fungsional Sistem Backend*

| Id    | Atribut Kualitas                               | Deksripsi Kebutuhan                                                                                                                                                                                   |
|-------|------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| NF-01 | Performa (Eksekusi Asinkron)                   | Sistem tidak boleh memblokir antarmuka pengguna (UI blocking) saat memproses komputasi I/O berat (parsing, ekstraksi LLM, embedding). Pemrosesan wajib berjalan di latar belakang (background tasks). |
| NF-02 | Transparansi (Pelaporan Real-Time)             | Sistem wajib menyediakan komunikasi Server-Sent Events (SSE) untuk melaporkan status ekstraksi secara instan kepada frontend.                                                                         |
| NF-03 | Keamanan (Isolasi Multi-Tenant)                | Sistem harus menjamin graf pengetahuan satu proyek tidak terekspos ke proyek lain melalui penerapan identitas komposit ( <i>composite ID</i> ) berbasis struktur <i>hash</i> .                        |
| NF-04 | Integritas Data (Idempotensi & Anti-Duplikasi) | Sistem harus kebal terhadap insiden penggandaan data saat terjadi unggahan berulang, memanfaatkan klausa spesifik basis data graf untuk memperbarui entitas yang ada tanpa menciptakan duplikat.      |
| NF-05 | Keandalan (Integritas Relasi Graf)             | Sistem dilarang membentuk relasi antar-literatur secara acak. Pembentukan relasi wajib melewati gerbang validasi bertingkat (Multi-Stage Semantic Gating).                                            |
| NF-06 | Ketahanan (Pemulihan Mandiri)                  | Sistem harus menangani kegagalan API eksternal (contoh: limit koneksi LLM) dengan mengimplementasikan mekanisme percobaan ulang yang terkalibrasi (exponential backoff retries).                      |

*Tabel 3.2.2 : Kebutuhan Non-Fungsional Sistem Backend*

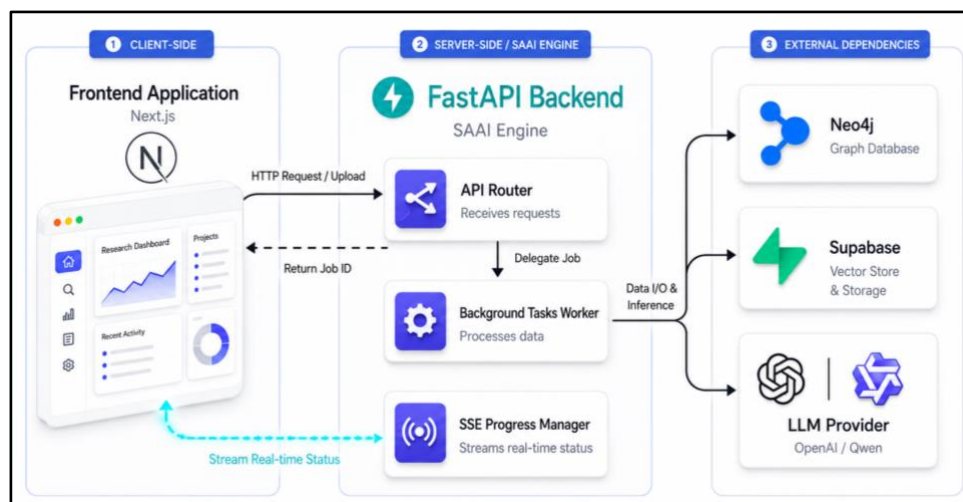
### 3.2.2. Perancangan Arsitektur dan Fondasi Teknologi Backend

Arsitektur MySRE beroperasi dengan memisahkan ranah klien dari server melalui penerapan arsitektur decoupled. Pemisahan ini merupakan

kebutuhan teknis untuk mengisolasi antarmuka pengguna (Next.js) dari beban komputasi masif pada backend. Pemrosesan *Natural Language Processing* (NLP) dan eksekusi instruksi kecerdasan buatan menyita sumber daya I/O secara intensif. Arsitektur terdistribusi memastikan interaksi pengguna tidak tertunda oleh kalkulasi data di belakang layar.

Alur kontrol sistem berpusat pada pendelegasian tugas. Lapis antarmuka program aplikasi (*API Router*) menerima permintaan dari klien. Router ini menolak menahan siklus permintaan, melainkan segera melempar komputasi mentah ke pekerja latar belakang (*Background Tasks*). Untuk memfasilitasi visibilitas proses, backend membuka saluran *Server-Sent Events* (SSE). Saluran searah ini memancarkan status progres secara instan, menyinkronkan lapisan analitik dengan antarmuka klien.

Interaksi tingkat tinggi antar-komponen ini direpresentasikan pada Gambar 3.2.2 berikut.



Gambar 3.2.2 : Arsitektur Client-Server MySRE

| Kategori              | Library                 | Justifikasi Teknis                                                                                                                                                    |
|-----------------------|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Framework             | FastAPI & Uvicorn       | Menyediakan arsitektur asynchronous berkinerja tinggi untuk menangani beban I/O berat tanpa memblokir antrean permintaan masuk.                                       |
| Komunikasi & Keamanan | SSE-Starlette & SlowAPI | sse-starlette memfasilitasi pelaporan progres real-time via Server-Sent Events. slowapi membatasi laju permintaan (rate limiting) untuk melindungi stabilitas server. |

|                                |                           |                                                                                                                                                                                                                                                                       |
|--------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ekstraksi Dokumen & OCR        | PyMuPDF & RapidOCR        | pymupdf4llm digunakan sebagai parser utama PDF. rapidocr-onnxruntime disematkan sebagai mesin <i>Optical Character Recognition</i> untuk mengonversi gambar/tabel pindaian menjadi teks.                                                                              |
| Pemrosesan Semantik Teks       | Langchain Text Splitters  | Mengatur hierarki pemecahan teks (Multi-layered Chunking) yang mempertahankan metadata judul (header) agar konteks spasial informasi tidak hilang.                                                                                                                    |
| Komputasi Embedding            | Sentence-Transformers     | Memfasilitasi kalkulasi vector embedding (seperti Multilingual-e5-small) secara lokal yang dikhususkan untuk menunjang kebutuhan pencarian semantik internal pada proses <i>Multi-layered Chunking</i> .                                                              |
| Orkestrasi AI & LLM            | OpenAI SDK & Google GenAI | Modul interkoneksi resmi untuk mengakses layanan model bahasa besar (LLM) dan model embedding eksternal. Digunakan sebagai penghasil matriks vektor utama penyusun graf, serta menangani kueri inferensi kompleks pada tahap Reasoning Gate dan klasifikasi semantik. |
| Ketahanan & Profiling AI       | Tenacity & Tiktoken       | tenacity memastikan pemulihan mandiri dari kegagalan API (exponential backoff). tiktoken melakukan kalkulasi token luring guna mencegah <i>over-context error</i> .                                                                                                   |
| Basis Data Relasional & Vektor | Supabase & Psycopg        | supabase digunakan untuk sinkronisasi aset penyimpanan. psycopg[binary.pool] merupakan adaptor PostgreSQL tingkat rendah yang mengoptimalkan manajemen koneksi (connection pooling).                                                                                  |
| Basis Data Graf                | Neo4j Driver              | Modul penghubung resmi ke graph database untuk eksekusi kueri Cypher tingkat lanjut, pemetaan relasi antar-literatur, dan perhitungan kalkulus graf.                                                                                                                  |
| Komputasi Numerik              | NumPy                     | Diandalkan untuk komputasi matriks matematika berkinerja tinggi, terutama dalam kalkulasi Cosine Similarity lokal dan perhitungan statistik Z-Score.                                                                                                                  |

Tabel 3.2.3 : Spesifikasi Teknologi dan Library

### 3.2.3. Spesifikasi Kontrak Komunikasi dan Antarmuka API

Perancangan antarmuka pemrograman aplikasi (API) pada SAAI Engine mengadopsi prinsip pemisahan tanggung jawab (separation of concerns). Arsitektur komunikasi memisahkan beban pengelolaan infrastruktur data dari beban komputasi kecerdasan buatan. Spesifikasi ini diklasifikasikan ke dalam dua ranah operasional utama, merepresentasikan

pembagian peran teknis dalam pengembangan sistem.

Tabel 3.2.4 merinci skema API infrastruktur yang berfokus pada manajemen lalu lintas data, orkestrasi tugas latar belakang, dan operasi lintas basis data.

| Method        | Endpoint                        | Deskripsi                                                                                                                            |
|---------------|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|
| <b>POST</b>   | /mcp/upload/submit              | Menerima unggahan dokumen mentah dan mendelegasikan beban komputasi berat ke background task. Merespons secara instan dengan Job ID. |
| <b>GET</b>    | /mcp/upload/stream              | Membuka jalur komunikasi <i>Server-Sent Events</i> (SSE) untuk menyiarkan progres pemrosesan secara <i>real-time</i> .               |
| <b>DELETE</b> | /api/graph/literature/{node_id} | Mengeksekusi penghapusan atomik ( <i>Double Kill</i> ); menghapus Node dan Edge di Neo4j sekaligus aset fisik di Supabase.           |
| <b>GET</b>    | /health                         | Memeriksa stabilitas operasional layanan dan konektivitas menuju basis data graf.                                                    |

Tabel 3.2.4 : Skema API Infrastruktur

Tabel 3.2.5 menjabarkan skema API inteligensi graf yang berfokus pada distribusi struktur pengetahuan dan penyediaan antarmuka bagi interaksi Agentic AI.

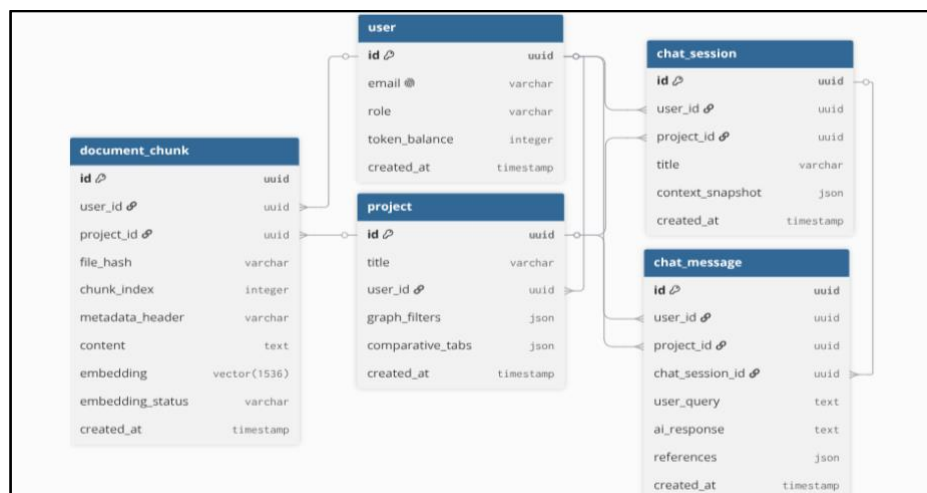
| Method | Endpoint              | Deskripsi                                                                                                                       |
|--------|-----------------------|---------------------------------------------------------------------------------------------------------------------------------|
| GET    | /api/graph/context    | Mendistribusikan data Knowledge Graph terisolasi (Node dan Edge) beserta penalaran LLM (reasoning) untuk visualisasi antarmuka. |
| POST   | /api/chat/stream      | Menangani kueri interaktif pengguna berbasis konteks graf (RAG), merespons luaran LLM secara streaming.                         |
| POST   | /api/comparative/deep | Memfasilitasi eksekusi LLM untuk melakukan analisis komparatif mendalam antar-literatur yang dipilih pengguna.                  |

Tabel 3.2.5 : Skema API Graph dan Agentic AI Services

### 3.2.4. Perancangan Skema Database

Penyimpanan data pada arsitektur MySRE mengadopsi strategi *Hybrid Data Store* untuk menyeimbangkan efisiensi komputasi dan skalabilitas penyimpanan. Sistem memisahkan beban kerja ke dalam dua platform basis data yang berbeda. Supabase yang berbasis PostgreSQL digunakan untuk mengelola data relasional serta penyimpanan massa data potongan teks (*raw chunks*). Basis data ini dioptimasi untuk menangani volume data yang besar dan operasi CRUD standar. Di sisi lain, Neo4j digunakan sebagai mesin pengelola pengetahuan yang memetakan relasi semantik kompleks antar-literatur. Penggunaan basis data graf memungkinkan sistem melakukan pencarian kedekatan vektor secara *native* dan penelusuran hubungan literatur yang sulit dicapai oleh basis data relasional konvensional.

Pondasi infrastruktur sistem dan manajemen akses pengguna dikelola melalui skema relasional pada Supabase. Integrasi antar-entitas data dalam platform ini digambarkan melalui diagram relasi antar-tabel (*Entity Relationship Diagram*) pada Gambar 3.2.3. Skema ini memastikan bahwa setiap aset data, mulai dari identitas pengguna hingga potongan teks terkecil, memiliki jalur rujukan yang jelas.



Gambar 3.2.3 : ERD Postgresql (Supabase)

Tabel 3.2.6 menjabarkan skema model User yang berfungsi sebagai pengelola identitas dan otoritas akses sistem. Model ini mencatat profil

dasar serta saldo penggunaan unit komputasi AI. Selanjutnya, Tabel 3.2.7 merinci model Project yang bertindak sebagai kontainer isolasi data. Model ini menyimpan preferensi visualisasi dan status fitur analitik yang disesuaikan untuk setiap ruang kerja penelitian.

| Kolom         | Tipe Data | Deskripsi                                 |
|---------------|-----------|-------------------------------------------|
| id            | UUID      | Pengenal unik pengguna (Primary Key).     |
| email         | String    | Alamat surel unik untuk autentikasi.      |
| role          | Enum      | Peran pengguna (USER atau ADMIN).         |
| token_balance | Int       | Saldo kuota penggunaan API LLM/Embedding. |

Tabel 3.2.6 Skema Model User (Manajemen Identitas)

| Kolom            | Tipe Data | Deskripsi                                              |
|------------------|-----------|--------------------------------------------------------|
| id               | UUID      | Pengenal unik proyek (Primary Key).                    |
| title            | String    | Nama proyek penelitian.                                |
| user_id          | UUID      | Relasi kepemilikan proyek ke tabel <i>User</i> .       |
| graph_filters    | Json      | Menyimpan preferensi filter visualisasi graf pengguna. |
| comparative_tabs | Json      | Menyimpan status tab pada fitur analisis komparatif.   |

Tabel 3.2.7 Skema Model Project (Kontainer Kerja)

Penyimpanan potongan teks ilmiah yang telah melalui proses *Multi-layered Chunking* dikelola oleh model *DocumentChunk* sebagaimana dirinci pada Tabel 3.2.8. Model ini dilengkapi dengan kolom `chunk_index` untuk menjamin urutan rekonstruksi narasi yang sistematis. Selain itu, kolom `embedding_status` digunakan untuk menjaga ketahanan sistem dalam melacak keberhasilan komputasi vektor.

| Kolom            | Tipe Data    | Deskripsi Teknis                                                                         |
|------------------|--------------|------------------------------------------------------------------------------------------|
| id               | UUID         | Pengenal unik potongan teks (Primary Key).                                               |
| project_id       | UUID         | Relasi ke tabel <i>Project</i> untuk isolasi data.                                       |
| user_id          | UUID         | Relasi ke tabel <i>User</i> untuk otorisasi pemilik data.                                |
| file_hash        | String       | Hash MD5 dokumen sumber sebagai identifikasi integritas berkas.                          |
| chunk_index      | Integer      | Nomor urut potongan teks untuk menjamin rekonstruksi narasi yang sistematis.             |
| metadata_header  | String       | Judul bab atau sub-bab guna mempertahankan konteks semantik teks.                        |
| content          | Text         | Isi potongan teks hasil proses <i>Chunking</i> .                                         |
| embedding        | Vector(1536) | Representasi numerik teks untuk kebutuhan pencarian semantik (RAG).                      |
| embedding_status | String       | Melacak status komputasi vektor ( <i>PENDING</i> , <i>SUCCESS</i> , atau <i>ERROR</i> ). |
| created_at       | Timestamp    | Catatan waktu pembuatan entitas data.                                                    |

Tabel 3.2.8 Skema Model DocumentChunk (Gudang Vektor)

Sebagai penunjang interaksi asisten riset, sistem juga menyediakan skema untuk menyimpan riwayat percakapan pengguna yang dirangkum pada Tabel 3.2.9 Model ini mencatat rujukan literatur yang digunakan oleh kecerdasan buatan dalam memberikan jawaban.

Tabel 3.2.9 Skema Model ChatSession dan ChatMessage

| Model              | Kolom Utama                         | Fungsi                                              |
|--------------------|-------------------------------------|-----------------------------------------------------|
| <b>ChatSession</b> | title, context_snapshot             | Mengelola sesi percakapan dan ringkasan graf aktif. |
| <b>ChatMessage</b> | user_query, ai_response, references | Menyimpan log tanya-jawab dan rujukan ragi (RAG).   |

Perancangan skema pada Neo4j berfokus pada efisiensi pemetaan pengetahuan saintifik. Setiap dokumen diwakili oleh Node :Literature yang memiliki identitas unik berdasarkan *composite\_id*. Pengenal ini dihasilkan melalui perhitungan MD5 di sisi *backend* untuk mengunci keamanan data tingkat tinggi. Properti metadata dan penyimpanan 5 Pilar Ilmiah pada setiap *node* diuraikan pada Tabel 3.9. Penulis menerapkan format penyimpanan berpasangan antara teks naratif dan vektor 1536-dimensi. Pendekatan ini memungkinkan sisteNode :Literature)m melakukan pencarian kandidat relasi yang spesifik pada setiap pilar riset.

**Tabel 3.2.10 : Skema Properti Identitas dan Metadata (Node :Literature)**

| Properti     | Tipe Data    | Deskripsi Teknis                                                                           |
|--------------|--------------|--------------------------------------------------------------------------------------------|
| composite_id | String       | Identifier Utama: hasil MD5 dari penggabungan identitas pengguna, proyek, dan hash berkas. |
| hash         | String       | MD5 dari berkas fisik PDF untuk verifikasi integritas data.                                |
| title        | String       | Judul artikel ilmiah hasil ekstraksi atau identitas berkas cadangan.                       |
| keywords     | List<String> | Kumpulan kata kunci yang menjadi basis validasi pada <i>Domain Compatibility Gate</i> .    |
| project_id   | String       | Pengenal unik proyek sebagai filter isolasi data.                                          |
| user_id      | String       | Pengenal unik pemilik data untuk otoritas akses.                                           |

**Tabel 3.2.11 : Skema Properti 5 Pilar Ilmiah (Node :Literature)**

| Properti          | Tipe Data    | Peruntukan Teknis                                                                                                        |
|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------|
| text_[pilar]      | String       | Menyimpan narasi pilar: <i>background</i> , <i>methodology</i> , <i>objective</i> , <i>gap</i> , dan <i>futurework</i> . |
| embedding_[pilar] | Vector(1536) | Representasi numerik pilar terkait untuk pencarian pada <i>Native Vector Index</i> .                                     |
| edges_generated   | Boolean      | Penanda status operasional untuk mengoptimalkan siklus kerja <i>background worker</i> .                                  |
| status            | String       | Melacak posisi dokumen dalam alur kerja sistem.                                                                          |

Relasi antar-literatur (*Edge*) pada sistem ini membawa atribut validasi yang membuktikan integritas hubungan tersebut. Tabel 3.11 merinci properti pada relasi semantik yang dihasilkan oleh alur kerja *Agentic AI*. Properti *gate* dan *matched\_pillars* berfungsi sebagai rekam jejak teknis yang menunjukkan bahwa relasi tersebut telah lolos tahap uji validasi yang terukur.

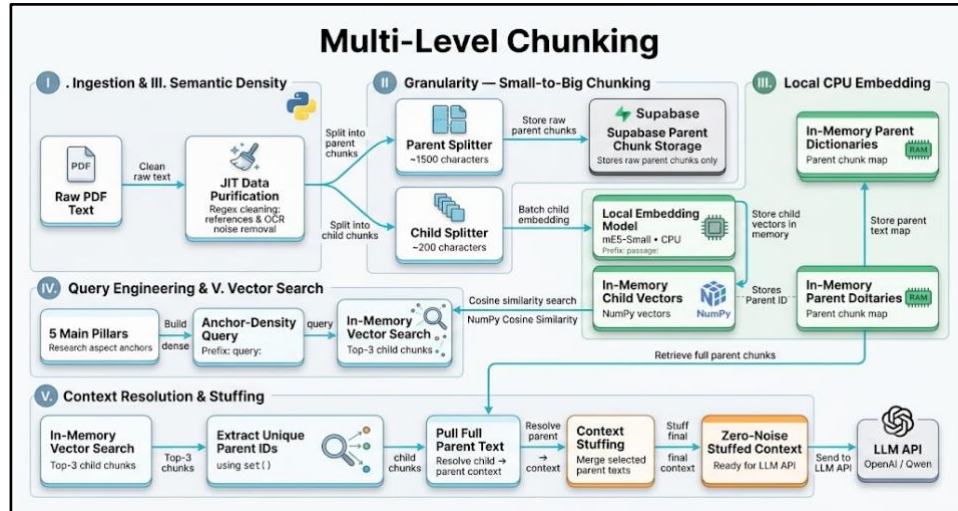
**Tabel 3.2.12 : Skema Properti Relasi Semantik (Edge)**

| Properti        | Tipe Data    | Deskripsi dan Justifikasi                                                                |
|-----------------|--------------|------------------------------------------------------------------------------------------|
| score           | Float        | Nilai <i>Cosine Similarity</i> yang memicu terbentuknya relasi (rentang 0.0 sampai 1.0). |
| reasoning       | String       | Narasi penjelasan hasil nalar LLM yang menjustifikasi keterhubungan antar-literatur.     |
| matched_pillars | List<String> | Daftar pilar yang terdeteksi selaras (bukti fisik dari <i>Coherence Gate</i> ).          |
| gate            | String       | Label pelacak logika validasi yang berhasil dilewati.                                    |

### 3.2.5. Implementasi *Multi Level Chunking & Stuffed Context*

Sistem ekstraksi pilar Literature pada modul “Brain” MySRE memerlukan pendekatan yang berbeda dari arsitektur *Retrieval-Augmented Generation* (RAG) konvensional. Pada metode *Classic RAG*, dokumen seringkali dikirimkan secara utuh atau dipotong dalam ukuran statis yang besar ke dalam model bahasa, yang berisiko memicu distorsi informasi. Arsitektur Backend mengadopsi paradigma *Small-to-Big Retrieval*, sebuah arsitektur yang memisahkan unit teks untuk kebutuhan pencarian dan kebutuhan penalaran. Pemisahan ini didasarkan pada temuan Eibich et al. (2024) melalui kerangka kerja ARAGOG, yang menunjukkan bahwa membedakan granularitas teks antara fase pencarian dan fase sintesis terbukti secara statistik memberikan hasil yang lebih unggul dibandingkan metode pencocokan vektor tunggal (Eibich et al., 2024). Validasi ini diperkuat oleh riset Lumer et al. (2025) di *PricewaterhouseCoopers* (PwC) yang menyatakan bahwa mekanisme perluasan konteks ini mutlak diperlukan untuk menyelesaikan kueri penalaran kompleks yang melintasi

batas-batas paragraf (*Multi-hop Reasoning*) tanpa menimbulkan latensi yang signifikan (Lumer et al., 2025).



Gambar 3.2.4 : Pipeline Multi-Level Chunking

## 1. Granularitas Vektor dan Pencegahan Information Loss

Dalam mengimplementasikan strategi *Small-to-Big*, teks akademik dipecah secara hierarkis. Sistem menetapkan dua ukuran berbeda: *Parent Chunk* dengan batas  $\pm 1500$  karakter untuk menjaga keutuhan narasi, dan *Child Chunk* berukuran  $\pm 200$  karakter yang dioptimalkan untuk pencarian vektor. Hanya teks pada level *Child Chunk* yang diubah menjadi representasi numerik. Desain ini merujuk pada prinsip Tan et al. (2025) bahwa memaksakan teks panjang ke dalam vektor berdimensi tetap akan menyebabkan hilangnya detail spesifik (*Information Loss*) (Tan et al., 2025). Penggunaan unit kecil setingkat proposisi menjamin sinyal semantik tetap padat. Pendekatan ini selaras dengan bukti empiris dari Chen et al. (2024) yang mendemonstrasikan bahwa pencarian pada tingkat proposisi mampu meningkatkan akurasi *Recall* hingga +10 poin, khususnya saat sistem menargetkan informasi yang bersifat anomali atau langka (*long-tailed information*) seperti pilar “Gap Penelitian” (Chen et al., 2024).

## 2. Purifikasi Data dan Vektorisasi Asinkronus

Akurasi pencarian vektor bergantung pada kemurnian teks. Sebelum tahap pemodelan vektor, sistem melakukan pembersihan teks secara *Just-In-Time (JIT) (Data Purification V2)* menggunakan aturan *Regular Expression* untuk membuang tabel hasil OCR yang rusak dan daftar pustaka. Metode ini merujuk pada konsep *Balanced Entropy Engineering* oleh Wang et al. (2025), yang memperingatkan bahwa teks asing dalam ruang vektor akan menciptakan ekspansi entropi buatan yang melarutkan atensi model terhadap konten esensial “*Attention Dilution*” (Y. Wang et al., 2025).

Vektorisasi dilaksanakan di memori lokal menggunakan model *intfloat/multilingual-e5-small*. Karena proses ini melibatkan ratusan *child chunks* per dokumen, sistem menggunakan pemrosesan asinkronus (*ThreadPool/Asyncio*) guna mencegah hambatan pada alur utama. Penggunaan model *mE5-Small* didasarkan pada kapabilitasnya dalam menangani jurnal ilmiah yang mencampur istilah bahasa Inggris dan Indonesia (L. Wang et al., 2024). Sesuai dengan spesifikasi arsitektur *asymmetric contrastive learning* pada *mE5*, instruksi prefix disematkan untuk menjaga integritas matriks:

```
def _embed_batch(self, texts: List[str], is_query: bool = False) -> List[List[float]]:
    """Embed a list of texts using the local SentenceTransformer model with E5 prefix rules."""
    prefix = "query: " if is_query else "passage: "

    # E5 models have a max sequence length of 512 tokens.
    # We truncate by character (approx 1500 chars to be safe) before encoding.
    MAX_CHARS = 1500

    prefixed_texts = []
    for t in texts:
        cleaned = (t or "").strip()
        if len(cleaned) > MAX_CHARS:
            logger.warning(f"Text exceeded {MAX_CHARS} chars, truncating before embedding.")
            cleaned = cleaned[:MAX_CHARS]
        prefixed_texts.append(f"{prefix}{cleaned}")

    # model.encode returns a numpy array or list of numpy arrays, we convert to list of floats
    embeddings = self.model.encode(prefixed_texts, batch_size=32, show_progress_bar=False)
    return [emb.tolist() for emb in embeddings]
```

Gambar 3.2.5 : Snippet Code Purifikasi Data

### 3. Rekayasa Query : Anchor Density Strategy

Untuk meningkatkan ketajaman titik temu, sistem meninggalkan format kueri tanya-jawab naratif (Q&A) dan beralih ke strategi *Anchor-Density*. Strategi ini menggabungkan proposisi dari Wang et al. (2025) dan Tan et al. (2024) bahwa kueri yang padat informasi (*information-dense formats*) lebih efektif mengunci fakta esensial. Dengan menggunakan deretan kata kunci teknis sebagai jangkar, sistem memusatkan entropi murni pada target pencarian, mencegah retriever meleset di ruang multidimensi. Berikut adalah konstanta kueri yang diterapkan:

```
query_texts = [  
    "latar belakang masalah, urgensi penelitian, motivasi mendasar, fenomena krusial"  
    "konteks situasi saat ini, dan alasan utama studi ini dilakukan.",  
    "rincian teknis metodologi, arsitektur sistem, desain eksperimen, teknik pengumpulan dataset"  
    "algoritma yang diusulkan, sampel, dan prosedur pengujian.",  
    "keterbatasan penelitian terdahulu, kelemahan metode sebelumnya, kekosongan literatur"  
    "masalah yang belum terselesaikan, namun demikian, akan tetapi, dan hasil yang belum jelas.",  
    "tujuan utama penelitian, sasaran spesifik, kontribusi inti makalah, hasil yang ingin dicapai,"  
    "mengeksplorasi, dan solusi yang diusulkan untuk menyelesaikan masalah.",  
    "saran untuk penelitian selanjutnya, rekomendasi pekerjaan di masa depan, arah pengembangan"  
    "masa depan, dan peluang studi berikutnya untuk mengatasi batasan saat ini."  
]
```

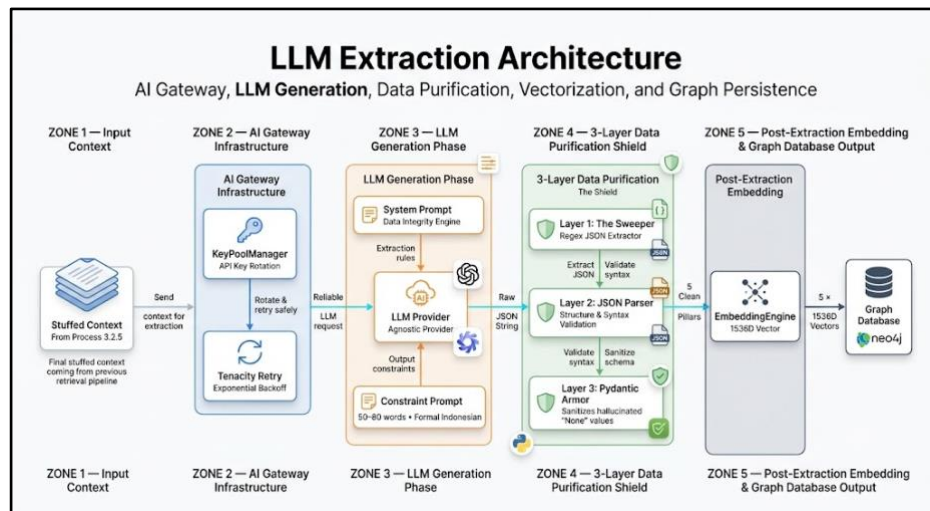
Gambar 3.2.6 : Anchor Density Query

### 4. Deduplikasi & Stuffing Context

Tahap akhir dari alur in-memory ini adalah resolusi konteks. Setelah pencarian vektor menemukan *Child Chunks* yang relevan, sistem menarik *Ids Parent* terkait untuk mengumpulkan teks utuh. Mengingat satu paragraf Parent dapat menaungi beberapa Child yang relevan, sistem menggunakan himpunan **set()** pada Id Parent untuk mengeliminasi duplikasi atau residu yang bertumpang tindih (*overlap*). Teks-teks dari ID unik tersebut digabungkan menjadi satu blok informasi yang disebut *Stuffed Context*. Penggabungan akhir ini menghasilkan *Zero-Noise Context*, yang memastikan model LLM dapat menerima landasan fakta yang padat, akurat, dan terarah untuk penalaran tahap akhir tanpa risiko halusinasi.

### 3.2.6. Strategi Prompt Engineering dan Ekstraksi Berbasis Skema

Sebelum mengeksekusi instruksi pemrosesan bahasa, sistem memerlukan lapisan infrastruktur yang menjamin stabilitas dan agnostisisme terhadap penyedia layanan. Backend MySRE mengimplementasikan *AI Gateway* yang tidak bergantung pada satu penyedia (seperti OpenAI atau DashScope) secara statis. Lapisan ini dilengkapi dengan *KeyPoolManager* untuk merotasi kunci API secara dinamis, mengeliminasi risiko pembatasan kuota (*rate-limiting*). Selain itu, untuk mengamankan eksekusi ekstraksi dari kegagalan jaringan yang tidak terduga, sistem mengadopsi mekanisme *Exponential Backoff Retry* menggunakan pustaka Tenacity. Pendekatan ini memastikan bahwa pipeline ekstraksi tetap beroperasi secara berkesinambungan dan tidak mengalami kelumpuhan akibat gangguan koneksi sesaat. Arsitektur Ekstraksi 5 pilar dapat dilihat pada gambar 3.2.7.



Gambar 3.2.7 : Arsitektur Ekstraksi LLM

#### 1. Rekayasa Prompt dan Penegakan Gaya Bahasa Akademik

Mengekstraksi pilar ilmiah dari teks yang panjang menuntut batasan yang ketat (*Constraint Prompting*). Sistem memisahkan peran instruksi dasar melalui System Prompt yang bertindak sebagai “*Senior Academic Data Integrity Engine*”, dan instruksi spesifik melalui User Prompt. Teks sumber yang telah digabungkan (*Stuffed Context*) disematkan secara aman ke dalam instruksi. Selain itu, *Large Language*

*Model* (LLM) diinstruksikan secara absolut untuk mematuhi rentang panjang tulisan (50-80 kata per pilar) dan diwajibkan menggunakan frasa pembuka akademik bahasa Indonesia yang baku. Tujuannya adalah untuk mensterilkan keluaran dari struktur kalimat generik yang sering dihasilkan oleh mesin. Berikut adalah cuplikan instruksi yang menegakkan aturan gaya bahasa tersebut :

```
prompt = f
"""You are a Senior Academic Reviewer and Data Extraction Specialist...

    CRITICAL RULE: Unlike metadata, you MUST write ALL pillar outputs STRICTLY IN FORMAL INDONESIAN with a sharp, flowing academic tone. Avoid robotic or machine-translated styles.

    - background: Berdasarkan teks, ekstrak LATAR BELAKANG PENELITIAN.
    ... Gunakan frasa pembuka seperti: 'Dilatarbelakangi oleh...', 'Menjadi krusial karena...'
    - gap: Berdasarkan teks, ekstrak GAP DAN KETERBATASAN PENELITIAN.
    ... Gunakan frasa pembuka seperti: 'Meninggalkan celah pada...', 'Belum dieksplorasi...'

    ## GLOBAL CONSTRAINTS & FORMAT
    - Each pillar text MUST strictly be between 50 and 80 words.
    - Your response MUST perfectly match the requested JSON schema.

    Compiled Text Excerpts:
    {text[:25000]}
    """
```

Gambar 3.2.8 : Snippet Code Prompt Ekstraksi Pilar

## 2. Kontrak Data Terstruktur dan Tiga Lapis Purifikasi (The Shield)

Ekstraksi yang dilakukan oleh model bahasa bersifat non-deterministik, sehingga rentan menghasilkan data yang tidak terstruktur atau mengalami halusinasi. Untuk mengamankan integritas basis data, sistem memberlakukan kontrak data berformat JSON berdasarkan skema *LiteraturePillars*. Keluaran LLM kemudian melewati tiga lapis purifikasi bertahap. Lapis pertama adalah modul Sweeper yang bertugas memotong elemen non-esensial dari struktur data. Lapis kedua memvalidasi sintaksis JSON. Lapis ketiga, yang berfungsi sebagai pelindung utama, memanfaatkan fitur *BeforeValidator* dari pustaka Pydantic untuk menetralkan halusinasi, seperti luaran bertuliskan “None” atau “Informasi tidak tersedia”, dan menggantinya menjadi string kosong yang aman. Berikut adalah cuplikan implementasi dari

ketiga lapis purifikasi tersebut :

```
1  ```python
2  def clean_llm_json_string(raw_text: str) -> str:
3      """LAPIS 1: Membersihkan teks respons LLM dari basa-basi dan markdown fences."""
4      start_idx = raw_text.find('{')
5      end_idx = raw_text.rfind('}')
6      if start_idx != -1 and end_idx != -1 and start_idx < end_idx:
7          return raw_text[start_idx:end_idx + 1]
8      return raw_text
9
10 # LAPIS 3: Armor sanitasi halusinasi (mengubah 'None'/'tidak tersedia' menjadi string kosong)
11 def sanitize_llm_field(value) -> str:
12     # ... (logika pembersihan reject_phrases) ...
13     return value.strip()
14
15 class LiteraturePillars(BaseModel):
16     background: Annotated[str, BeforeValidator(sanitize_llm_field)] = Field(...)
17     gap: Annotated[str, BeforeValidator(sanitize_llm_field)] = Field(...)
18     # ...
19
20 # Eksekusi Parsing (Lapis 2 & 3 berjalan bersamaan)
21 cleaned_json_str = clean_llm_json_string(raw_json)
22 parsed = json.loads(cleaned_json_str) # Lapis 2: JSON Load
23 validated_response = ExtractionResponse.model_validate(parsed) # Lapis 3: Pydantic Validation
24 ```
```

Gambar 3.2.9 : Snippet Code Purification Output LLM

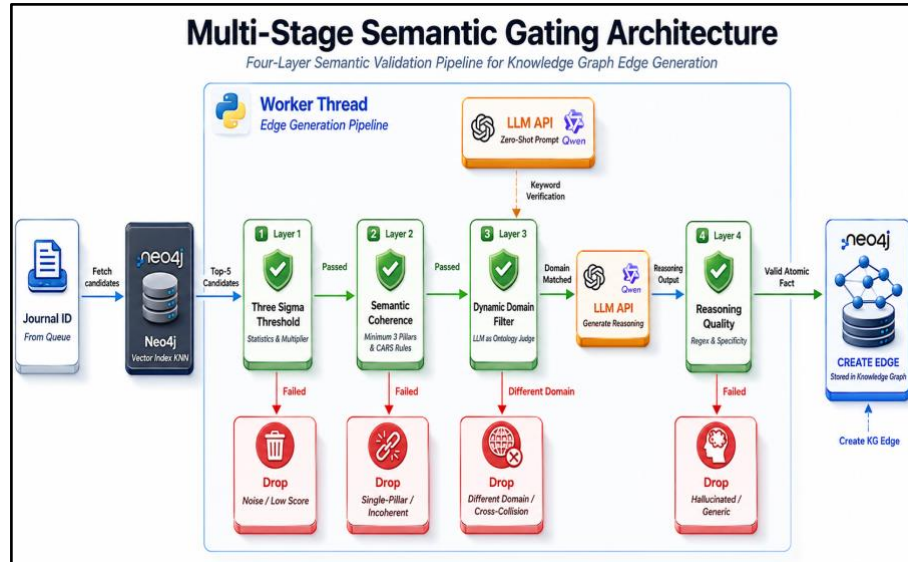
### 3. Vektorisasi *Post-Extraction*

Fase terakhir dalam alur ekstraksi ini adalah mentransformasi teks pilar yang telah disanitasi menjadi representasi matematis. Setelah kelima pilar dinyatakan lolos purifikasi Pydantic, masing-masing blok teks dikembalikan ke *EmbeddingEngine* di dalam infrastruktur *AI Gateway*. Teks pilar tersebut dikonversi menjadi vektor semantik berdimensi 1536. Vektor inilah yang nantinya akan disimpan di dalam basis data grafik (Neo4j) dan berfungsi sebagai amunisi utama untuk proses rekayasa Multi-Stage Semantic Gating pada tahap selanjutnya.

#### 3.2.7. Rekayasa Multi-Stage Semantic Gating

Implementasi Vector Search tradisional yang murni bersandar pada algoritma *K-Nearest Neighbor* (K-NN) menyimpan kelemahan matematis dalam konteks ekstraksi literatur ilmiah. Algoritma K-NN beroperasi dengan memaksakan pembentukan relasi terhadap sejumlah dokumen terdekat, terlepas dari relevansi aktualnya. Dalam skala yang masif, pemaksaan ini berisiko menciptakan *hairball graph*, dimana sebuah graf dipenuhi oleh relasi palsu (*false positives*). Untuk mereduksi risiko tersebut, arsitektur MySRE menerapkan mekanisme *Multi-Stage Semantic Gating*, sebuah sistem penyaringan empat lapis yang dirancang untuk

memverifikasi setiap relasi secara statistik, struktural, ontologis, dan leksikal sebelum disimpan ke dalam *Knowledge Graph*.



Gambar 3.2.10 : Multi-Stage Semantic Gating Architecture

#### 1. K-NN dan Ambang Batas Deteksi Anomali (Three Sigma Rule)

Lapisan pertama memodifikasi pendekatan K-NN menjadi Graf Lingkungan- $\epsilon$  ( $\epsilon$ -*Neighborhood Graph*). Sesuai dengan batasan yang diusulkan oleh Luxburg (2007), metode ini menolak pemaksaan koneksi dengan hanya mengizinkan pembentukan relasi jika skor kemiripan melampaui parameter  $\epsilon$  *Threshold* (ambang batas). Dalam sistem ini, nilai  $\epsilon$  tidak ditentukan secara statis (Von Luxburg, 2007). Mengacu pada teknik *Statistical Anomaly Detection* (Chandola et al., 2009), kemiripan semantik sejati didefinisikan sebagai “anomali” yang berada di luar batas tiga standar deviasi ( $3\sigma$ ) dari rata-rata ( $\mu$ ) skor noise vektor.

Lebih lanjut, sistem mengimplementasikan *Threshold Multiplier* sebagai Hyper Parameter untuk memitigasi kelemahan *model dense retriever* saat menghadapi entitas yang spesifik dan langka (Luan et al., 2021). Pilar dengan kepadatan informasi rendah, seperti “Gap Penelitian”, diberikan penalti pengali yang lebih tinggi untuk memperketat interval kepercayaan (*confidence interval*). Berikut adalah rumusan matematis dan nilai kalibrasi yang diterapkan pada sistem:

$$Threshold_{pillar} = (\mu + 3\sigma) \times Multiplier_{pillar}$$

Tabel 3.2.13 Parameter Kalibrasi Threshold Dinamis per Pilar

| Pilar Akademik                          | Multiplier | Nilai Threshold Statis) |
|-----------------------------------------|------------|-------------------------|
| Latar Belakang ( <i>Background</i> )    | 1.0        | 0.6807                  |
| Metodologi ( <i>Methodology</i> )       | 1.0        | 0.6914                  |
| Tujuan ( <i>Objective</i> )             | 1.05       | 0.7292                  |
| Saran Masa Depan ( <i>Future Work</i> ) | 1.05       | 0.7962                  |
| Gap Penelitian ( <i>Gap</i> )           | 1.15       | 0.8102                  |

## 2. *Semantic Coherence Gate* dan Kepatuhan Tata Bahasa Akademik

Lolos dari ambang batas statistik tidak menjamin relevansi struktural. Lapisan kedua (*Semantic Coherence Gate*) memastikan keselarasan lintas faset. Secara teoretis, sistem mengadopsi filosofi dari model ASPIRE (Mysore et al., 2022) yang membuktikan bahwa dokumen saintifik memiliki argumen yang bersifat multi-faset, sehingga evaluasi kemiripan harus dilakukan secara terperinci per-aspek.

Namun, alih-alih mengimplementasikan algoritma Optimal Transport dari ASPIRE yang menuntut komputasi kuadratik tinggi, penulis menerapkan pendekatan rekayasa heuristik. Melalui fungsi `validate_semantic_coherence`, sistem akan menggugurkan kandidat relasi secara mutlak jika jumlah pilar yang selaras tidak memenuhi syarat minimum konsensus “*len(matched\_pillars) < COHERENCE\_MIN\_MATCHED\_PILLARS*”. langkah ini bertindak sebagai mekanisme yang efisien untuk memastikan bahwa kemiripan bukanlah anomali kebetulan pada satu pilar tunggal, melainkan wujud penyelarasan struktural yang utuh sebelum sebuah relasi (Edge) terbentuk.

Selain kuantitas pilar, sistem menegakkan dependensi logis yang diturunkan dari Teori CARS (*Creating a Research Space*) oleh John Swales (1990). Dalam tata bahasa saintifik, metodologi (sebagai instrumen) tidak bermakna tanpa adanya niat, yaitu “Tujuan” atau “Gap” (Swales, 1990). Karenanya, kesamaan metodologi secara terisolasi akan ditolak oleh sistem. Aturan dependensi deterministik ini dikodifikasikan sebagai berikut:

Tabel 3.2.14 Aturan Dependensi Semantic Coherence Gate

| Pilar Utama yang Terdeteksi Mirip        | Syarat Ketergantungan Mutlak (Minimal Salah Satu)             |
|------------------------------------------|---------------------------------------------------------------|
| Metodologi                               | Harus disertai kemiripan Tujuan atau Gap                      |
| Tujuan                                   | Harus disertai kemiripan Latar Belakang, Metodologi, atau Gap |
| Gap                                      | Harus disertai kemiripan Tujuan atau Metodologi               |
| Arahan Masa Depan ( <i>Future Work</i> ) | Harus disertai kemiripan Tujuan, Metodologi, atau Gap         |

### 3. *Dynamic Domain Filtering*

Kelemahan inheren dari komputasi vektor adalah kebutaannya terhadap taksonomi keilmuan, yang dapat memicu Cross-Domain Semantic Collision. misalnya, algoritma *Random Forest* di domain meteorologi akan dinilai mirip dengan *Random Forest* di domain kardiologi. Untuk menanganinya, Zhou et al. (2023) merekomendasikan teknik Zero-Shot Prompting untuk mendelegasikan LLM sebagai “Hakim” (Zhou et al., 2023). Sistem memanggil LLM untuk mengevaluasi kata kunci dari kedua jurnal dan memastikan adanya irisan interdisipliner sebelum relasi diresmikan.

```
def evaluate_domain_compatibility_with_llm(client, keyword_source: list, keyword_target:
list) -> dict:

    """Gate 3: Evaluate domain compatibility between two nodes using LLM semantic judgmen
t."""
    user_content = (
        "Compare two sets of research keywords. Determine if they share the same "
        "overarching academic domain or have a strong interdisciplinary intersection.\n\n"
        "Rules:\n"
        "1. Focus on latent semantic meaning, not exact string matches.\n"
        "2. Treat concepts in different languages as identical if their meanings align.\n"
        "3. Resolve abbreviations to their full concepts.\n\n"
        "Return ONLY a valid JSON object with exactly two keys:\n"

        "- reason: Ultra-concise explanation of their relationship or mismatch (under 10 words).
\n"
        "- is_compatible: boolean (true if domains intersect/match, false if unrelated).
\n\n"
        f"Source Keywords: {keyword_source}\n"
        f"Target Keywords: {keyword_target}"
    )
```

Gambar 3.2.11 : Snippet Code Domain Filtering

#### 4. Validasi Kualitas Edge Reasoning Eksternal

Tahap pamungkas mengharuskan LLM menuliskan alasan (*reasoning*) keterhubungan antar-dokumen (*generation-time correction*), sesuai panduan Zhou et al. (2023). Namun, LLM rentan menghasilkan halusinasi atau teks generik. Karena LLM tidak mampu mengoreksi penalarannya sendiri (*Intrinsic Self-Correction*) (Huang et al., 2025), sistem menggunakan modul evaluasi eksternal berbasis *Regular Expression* (Regex) di Python. Modul ini bertindak sebagai *Rule-Based Expert System* yang menggugurkan relasi jika alasan yang diberikan bersifat basa-basi. Evaluasi ini diperkuat dengan pendekatan fakta atomik (Min et al., 2023), yang mewajibkan LLM untuk menyebutkan kata kunci spesifik yang dapat diverifikasi (seperti “algoritma” atau “dataset”).

```
REASONING_GENERIC_PATTERNS = [
    r"\bsama[- ]?sama membahas\b",
    r"\bmemiliki kemiripan\b",
    r"\btopik serupa\b",
]

def validate_edge_reasoning_quality(reasoning: str) -> tuple[bool, str]:
    """Runtime quality gate untuk menolak reasoning fallback/generik sebelum
    merge."""
    lowered = (reasoning or "").strip().lower()

    # 1. Pengecekan Pola Basa-basi (Halusinasi LLM)
    generic_hits = any(re.search(pattern, lowered, flags=re.IGNORECASE)
    for pattern in REASONING_GENERIC_PATTERNS)

    # 2. Pemaksaan Fakta Atomik (Alat pendeteksi kebohongan)
    specificity_hits = sum(1 for kw in REASONING_SPECIFICITY_KEYWORDS
    if kw in lowered)

    if generic_hits and specificity_hits < 2:
        return False, "generic-reasoning-detected"

    return True, "quality-ok"
```

Gambar 3.2.12 : *Snippet Code Edge Reasoning Quality Check*

### 3.3 Penugasan-Penugasan yang Telah Diselesaikan

Bagian ini menguraikan linimasa dan rincian operasional penugasan yang telah dilaksanakan selama program magang. Seluruh pelaksanaan tugas mencakup pemenuhan kewajiban administratif, prosedur bimbingan rutin, hingga pengembangan arsitektur teknis yang didasari oleh landasan literatur yang terukur.

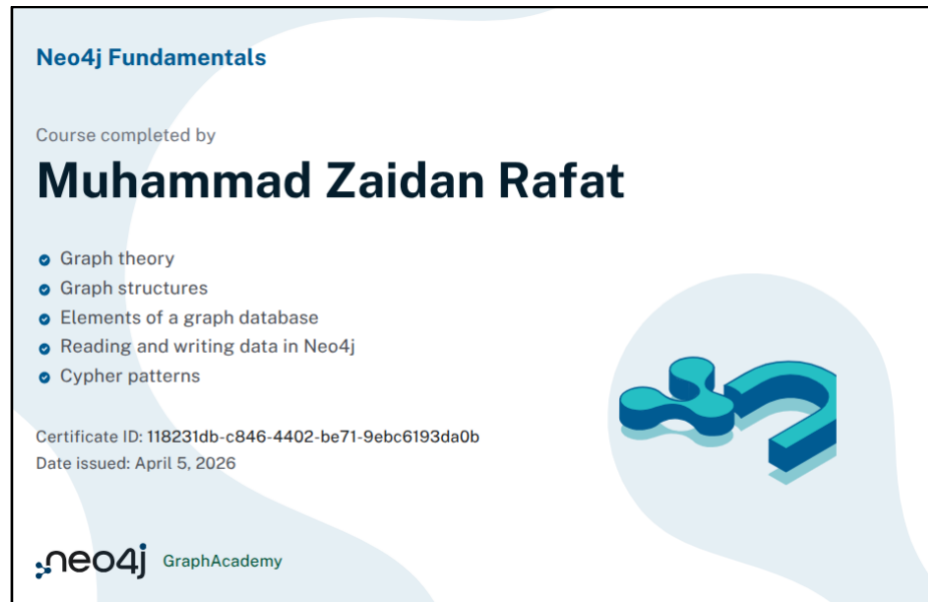
#### 3.3.1. Penugasan Administratif dan Pelatihan Penguatan Kompetensi

Pelaksanaan program diawali dengan pemenuhan kewajiban administratif berupa penyelesaian rangkaian modul pelatihan pada platform BRILian. Modul ini berfokus pada penguatan kapasitas dasar peneliti, termasuk pembentukan komitmen belajar dan integritas riset.



Gambar 3.3.1 : Sertifikat Brilliant

Selain pelatihan institusional, penugasan dilanjutkan dengan pendidikan mandiri untuk menguasai fundamental basis data graf. Kompetensi ini disahkan melalui penyelesaian sertifikasi dari Neo4j Academy, yang berfungsi sebagai landasan teoretis sebelum pengembangan pemodelan graf secara teknis.



Gambar 3.3.2: Sertifikasi Fundamental Neo4j

### 3.3.2. Sistem Bimbingan, Evaluasi Rutin, dan Pelaporan Kinerja

Pemantauan dan evaluasi kinerja dilakukan secara berkala untuk memastikan penyelesaian target proyek. Evaluasi utama diwujudkan melalui rapat progres (*progress meeting*) setiap hari Senin, yang dieksekusi secara hybrid dengan porsi kehadiran *onsite* yang lebih dominan.

Untuk memastikan akuntabilitas jam kerja dan pelaporan tugas, seluruh aktivitas operasional harian dicatat secara kronologis melalui pengisian logbook dan presensi pada portal resmi ELSA BRIN.

### 3.3.3. Penugasan Teknis Berbasis Linimasa Magang

Pengembangan dan restrukturisasi arsitektur MySRE dibagi ke dalam peta jalan teknis selama enam bulan:

1. **Bulan Pertama:** Observasi dan Analisis Kelemahan Sistem  
Penugasan dimulai dengan tahap pemahaman konteks proyek melalui audit dan observasi terhadap legacy code. Eksekusi reverse engineering dilakukan untuk membongkar dan menganalisis kelemahan struktural pada arsitektur pemetaan konsep sebelumnya, yang menjadi pijakan untuk perancangan ulang.
2. **Bulan Kedua:** Riset Literatur dan Spesifikasi Teknologi Bulan

kedua difokuskan pada pengumpulan landasan teori dan komparasi instrumen pengembangan. Proses ini mencakup riset tech stack, evaluasi kemampuan model bahasa besar (LLM) dan model embedding dari berbagai penyedia, serta pembelajaran mendalam mengenai arsitektur Neo4j.

3. **Bulan Ketiga:** Eksperimen Knowledge Graph dan Implementasi Filter Domain Pada bulan ini, dilakukan eksperimen konstruksi Knowledge Graph secara end-to-end. Untuk mencegah pembentukan relasi yang salah sasaran antar jurnal lintas disiplin, diimplementasikan mekanisme Domain Compatibility Gate. Sesuai dengan rujukan Zhou et al. (2023), model LLM diposisikan secara Zero-Shot Prompting untuk memvalidasi kesamaan ranah ontologi sebelum relasi dibentuk. Selain itu, Semantic Coherence Gate diterapkan guna memastikan kecocokan antar jurnal divalidasi pada minimal tiga aspek pilar, merujuk pada teori struktur retorika artikel ilmiah CARS Model (Swales, 1990) dan model kemiripan multi-vektor (Mysore et al., 2022).
4. **Bulan Keempat:** Restrukturisasi Asinkronus dan Perhitungan Statistik Ambang Batas Fase ini mencakup pemugaran arsitektur sistem secara menyeluruh menggunakan framework FastAPI guna membangun fondasi layanan asinkronus. Dalam pemrosesan vektor, algoritma pencarian k-NN dimodifikasi menjadi “ *$\epsilon$ -Neighborhood Graph*” dengan mengaplikasikan *Threshold Dynamic Filtering*. Ambang batas relasi dihitung melalui *Statistical Anomaly Detection* berdasarkan aturan *Three Sigma Rule of Thumb*, yang memastikan hanya vektor berstatus anomali semantik “ *$Z\text{-Score} \geq \mu + 3\sigma$* ” yang lolos menjadi relasi dalam graf.
5. **Bulan Kelima:** Pematangan Arsitektur dan Integrasi Antarmuka Fokus bulan kelima adalah stabilisasi arsitektur asinkronus dan eksekusi integrasi fungsional antara backend dan frontend. Tahap ini juga dialokasikan untuk investigasi dan mitigasi isu teknis (*bug*

*fixing*) yang ditemukan pasca-integrasi.

6. **Bulan Keenam:** Penyusunan Luaran Akhir dan Sosialisasi Bulan terakhir didedikasikan untuk kompilasi luaran program yang meliputi penyusunan draf Laporan MBKM, penulisan draf artikel ilmiah berskala IEEE, penyusunan materi presentasi untuk Laporan MBKM, serta sosialisasi serah terima penyelesaian proyek MySRE.

#### 3.3.4. Partisipasi Forum Ilmiah dan Demonstrasi Produk

Di luar penugasan pengembangan teknis, keterlibatan aktif dalam kegiatan diskusi akademik turut dilaksanakan melalui partisipasi pada forum ilmiah pekanan Pusat Riset Sains Data dan Informasi (PR-SDI) BRIN setiap hari Rabu.

Sebagai penugasan strategis, sistem MySRE didemonstrasikan untuk memaparkan kapabilitas arsitekturnya bersama pakar internasional (*International Expert*) dari Taiwan, yang menjadi sarana validasi empiris di lingkungan riset.

#### 3.4 Pembelajaran Hal Baru (*Lessons Learned*) yang Diperoleh

Berdasarkan pengalaman dan pelaksanaan tugas selama program magang di Badan Riset dan Inovasi Nasional (BRIN), terdapat serangkaian pembelajaran baru yang diperoleh oleh penulis. Pembelajaran ini mencakup aspek manajerial, pemahaman riset, hingga peningkatan kemampuan teknis, yang dijabarkan sebagai berikut:

1. **Dinamika Kolaborasi dalam Ekosistem Riset Profesional** Penulis merefleksikan bahwa kolaborasi dan bimbingan merupakan faktor penentu keberhasilan program magang. Momen berharga berupa arahan teknis dari pembimbing di BRIN sangat membantu penulis dalam merinci dan memahami tahapan pengembangan sistem. Selain panduan teknis, penulis juga mengakui bahwa dukungan mental dari pembimbing merupakan elemen penting yang menjaga konsistensi kinerja selama proses magang.
2. **Strategi Manajemen Proyek Terapan** Penulis memperoleh pembelajaran terkait tata kelola siklus pengembangan perangkat lunak secara nyata. Pengalaman dalam membedah kode bawaan (*legacy code*), mengeksekusi

*reverse engineering*, dan menentukan skala prioritas fitur memberikan pemahaman mendalam bagi penulis tentang praktik manajemen proyek yang tidak sekadar teoretis, melainkan sangat adaptif terhadap berbagai masalah di lapangan.

3. **Kedudukan Literatur Saintifik sebagai Fondasi :** Penulis menyadari secara penuh bahwa dalam pengembangan arsitektur sistem pada lingkup penelitian, baris kode tidak boleh ditulis berdasarkan asumsi mentah. Pengalaman merancang sistem penyaringan (*filtering gates*) dan perhitungan statistik membuktikan kepada penulis bahwa arsitektur yang tangguh mutlak memerlukan justifikasi literatur (seperti jurnal ilmiah) sebagai pijakan validitas operasionalnya.
4. **Integrasi Peran *Backend* dan *Knowledge Data Engineer*** Penulis mendapatkan pembelajaran praktis yang berharga dalam menghadapi dualitas beban kerja di dunia kerja nyata. Melalui proyek ini, penulis memahami betul bagaimana cara menyeimbangkan keandalan infrastruktur di belakang layar (*backend* asinkronus dengan FastAPI) dengan kompleksitas pengelolaan struktur data berbasis graf (Neo4j) tanpa mengorbankan performa sistem secara keseluruhan.
5. **Konstruksi dan Orkestrasi Arsitektur *Agentic AI*** Penulis memperoleh pemahaman fundamental bahwa membangun sistem *Agentic AI* bukanlah sekadar aktivitas memanggil API model bahasa besar (LLM). Pembelajaran penulis terfokus pada perancangan arsitektur agen otonom yang secara ketat dibatasi oleh aturan deterministik, sehingga kecerdasan buatan dapat diarahkan untuk melakukan penalaran yang aman dan terukur.
6. **Penetapan Fokus Spesialisasi Karier Profesional** Pada akhirnya, penulis menemukan jati diri karir yang jelas setelah sebelumnya sempat merasa bimbang dalam menentukan fokus akibat kecenderungan sebagai seorang generalis. Pengalaman serta tantangan yang diselesaikan selama mengerjakan proyek ini telah menumbuhkan ketertarikan yang nyata, sekaligus meyakinkan penulis untuk fokus berkarir secara mendalam di bidang *Agentic AI*, *Backend Engineering*, dan *Data Engineering*.



## **BAB IV**

### **PENUTUP**

#### **4.1 Kesimpulan**

Program MBKM Riset/Penelitian di Badan Riset dan Inovasi Nasional (BRIN) telah dilaksanakan selama enam bulan pada Pusat Riset Sains Data dan Informasi (PR-SDI), secara spesifik di bawah naungan Kelompok Riset *Human Computer Interaction - Visualization*. Kegiatan magang ini berfokus pada perancangan fondasi arsitektur untuk sistem Research Agentic AI yang bernama MySRE.

Dalam pelaksanaannya, kegiatan ini menuntut penyelesaian dua beban kerja teknis yang saling terintegrasi. Pada ranah rekayasa data pengetahuan (*Knowledge Data Engineering*), penulis berhasil membangun pipeline ekstraksi yang beralih dari metode *Classic* RAG menuju arsitektur *Small-to-Big Retrieval* melalui teknik pemecahan teks berjenjang (*Multi-Leveling Chunking*). Sistem ini menyaring keluaran teks menggunakan tiga lapis purifikasi (*The Shield*) dan mengamankan integritas *Knowledge Graph* di Neo4j melalui empat lapis *Multi-Stage Semantic Gating*. Lapis validasi ini mencakup perhitungan ambang batas *KNN Statistical Anomaly Detection* menggunakan *Three Sigma Rule of Thumb*, *Coherence Gate*, *Dynamic Domain Filtering*, hingga validasi relasi dengan *Edge Reasoning*. Pada ranah rekayasa infrastruktur backend, penulis merancang arsitektur asinkronus terpisah (*decoupled*) menggunakan kerangka kerja FastAPI. Infrastruktur ini mampu mendelegasikan komputasi I/O yang berat ke pekerja latar belakang (*background tasks*) sambil menyiarkan status operasional secara langsung via *Server-Sent Events* (SSE). Arsitektur ini terbukti mampu melayani fitur analitik dan antrean kueri Chatbot tanpa memblokir antrean pemrosesan utama (*bottleneck*).

#### **4.2 Saran**

Berdasarkan hasil arsitektur yang telah dibangun, terdapat dua aspek utama yang direkomendasikan untuk pengembangan sistem selanjutnya. Pertama, implementasi algoritma klasterisasi (*clustering*) terhadap node dan edge pada graf yang telah terbentuk. Tujuannya adalah untuk memetakan

kelompok konsep riset yang memiliki kedekatan semantik secara otomatis, sehingga pola penelitian dapat terbaca lebih jelas. Kedua, integrasi sistem dengan API eksternal penyedia literatur ilmiah dengan pendekatan pencarian sentris-entitas. Melalui fitur ini, node (entitas/dokumen) yang tersimpan di dalam *Knowledge Graph* tidak lagi bersifat pasif, melainkan dapat digunakan sebagai kata kunci referensi (*query anchor*) bagi sistem untuk melakukan penarikan otomatis artikel-artikel baru. Mekanisme ini memastikan bahwa setiap literatur yang ditarik dari pangkalan data eksternal memiliki keterkaitan konteks yang erat dengan dokumen yang sedang difokuskan oleh pengguna.

## DAFTAR PUSTAKA

- Alinea. (2021). *Harteknas, BRIN luncurkan logo baru*.  
<https://www.alinea.id>. <https://www.alinea.id/nasional/harteknas-brin-luncurkan-logo-baru-b2c7L95D0>
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), 1–58.  
<https://doi.org/10.1145/1541880.1541882>
- Chen, T., Wang, H., Chen, S., Yu, W., Ma, K., Zhao, X., Zhang, H., & Yu, D. (2024). *Dense X Retrieval: What Retrieval Granularity Should We Use?* (arXiv:2312.06648). arXiv. <https://doi.org/10.48550/arXiv.2312.06648>
- Eibich, M., Nagpal, S., & Fred-Ojala, A. (2024). *ARAGOG: Advanced RAG Output Grading* (arXiv:2404.01037). arXiv. <https://doi.org/10.48550/arXiv.2404.01037>
- Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ACM Transactions on Information Systems*, 43(2), 1–55.  
<https://doi.org/10.1145/3703155>
- Lu, Y., Wu, W., Zhao, X., Peng, R., & Wang, J. (2026). *KARMA: Leveraging Multi-Agent LLMs for Automated Knowledge Graph Enrichment* (arXiv:2502.06472). arXiv. <https://doi.org/10.48550/arXiv.2502.06472>
- Luan, Y., Eisenstein, J., Toutanova, K., & Collins, M. (2021). Sparse, Dense, and Attentional Representations for Text Retrieval. *Transactions of the Association for Computational Linguistics*, 9, 329–345.  
[https://doi.org/10.1162/tacl\\_a\\_00369](https://doi.org/10.1162/tacl_a_00369)
- Lumer, E., Melich, M., Zino, O., Kim, E., Dieter, S., Basavaraju, P. H., Subbiah, V. K., Burke, J. A., & Hernandez, R. (2025). *Rethinking Retrieval: From Traditional Retrieval Augmented Generation to Agentic and Non-Vector Reasoning Systems in the Financial Domain for Large Language Models* (arXiv:2511.18177). arXiv. <https://doi.org/10.48550/arXiv.2511.18177>
- Min, S., Krishna, K., Lyu, X., Lewis, M., Yih, W., Koh, P. W., Iyyer, M., Zettlemoyer, L., & Hajishirzi, H. (2023). *FActScore: Fine-grained Atomic Evaluation of Factual Precision in Long Form Text Generation* (arXiv:2305.14251). arXiv. <https://doi.org/10.48550/arXiv.2305.14251>
- Mysore, S., Cohan, A., & Hope, T. (2022). Multi-Vector Models with Textual Guidance for Fine-Grained Scientific Document Similarity. *Proceedings of the 2022 Conference of the North American Chapter of the*

- Association for Computational Linguistics: Human Language Technologies*, 4453–4470. <https://doi.org/10.18653/v1/2022.naacl-main.331>
- Permana, R. H. (2021). *BRIN Punya Logo Berwarna Merah, Ini Artinya*. detiknews. <https://news.detik.com/berita/d-5721352/brin-punya-logo-berwarna-merah-ini-artinya>
- Singh, A., Ehtesham, A., Kumar, S., Khoei, T. T., & Vasilakos, A. V. (2026). *Agentic Retrieval-Augmented Generation: A Survey on Agentic RAG* (arXiv:2501.09136). arXiv. <https://doi.org/10.48550/arXiv.2501.09136>
- Statistics (NCSES), N. C. for S. and E. (2023). *Publications Output: U.S. Trends and International Comparisons* (NSB-2023-33). Article NSB-2023-33. <https://nces.nsf.gov/pubs/nsb202333>
- suara.com. (2021). *BRIN Luncurkan Logo Baru di Momen Harteknas*. suara.com. <https://www.suara.com/tekno/2021/08/10/204500/brin-luncurkan-logo-baru-di-momen-harteknas>
- Swales, J. (1990). *Create a Research Space (CARS) Model of Research Introductions*.
- Tan, H., Zhan, S., Lin, H., Zheng, H.-T., & Chan, W. K. (2025). QAEA-DR: A Unified Text Augmentation Framework for Dense Retrieval. *IEEE Transactions on Knowledge and Data Engineering*, 37(6), 3669–3683. <https://doi.org/10.1109/TKDE.2025.3543203>
- Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4), 395–416. <https://doi.org/10.1007/s11222-007-9033-z>
- Wang, L., Yang, N., Huang, X., Yang, L., Majumder, R., & Wei, F. (2024). *Multilingual E5 Text Embeddings: A Technical Report* (arXiv:2402.05672). arXiv. <https://doi.org/10.48550/arXiv.2402.05672>
- Wang, Y., Ren, R., Wang, Y., Liu, J., Zhao, W. X., Wu, H., & Wang, H. (2025). *BEE-RAG: Balanced Entropy Engineering for Retrieval-Augmented Generation* (arXiv:2508.05100). arXiv. <https://doi.org/10.48550/arXiv.2508.05100>
- Zhou, H., Liu, F., Gu, B., Zou, X., Huang, J., Wu, J., Li, Y., Chen, S., Zhou, P., Liu, J., Hua, Y., Mao, C., Wu, X., Zheng, Y., Clifton, L., Li, Z., Luo, J., & Clifton, D. (2023). A Survey of Large Language Models in Medicine: Progress, Application, and Challenge. *arXiv.Org*. <https://doi.org/https://doi.org/10.48550/arXiv.2311.05112>

## LAMPIRAN

Yang perlu dilampirkan:

1. Biodata peserta MBKM (format bebas)
2. Biodata pembimbing BRIN (format bebas)
3. Biodata pembimbing universitas (format bebas)
4. Dokumentasi kegiatan MBKM
5. Luaran kegiatan MBKM (jika ada). Misalnya: artikel jurnal yang sudah dipublikasikan, hasil review jurnal, laporan hasil penelitian, laporan hasil pengujian, dsb. atau kegiatan lainnya yang dapat menjadi luaran dalam kegiatan MBKM, baik jenis penelitian/riset maupun magang/praktik kerja)

## LAMPIRAN BIODATA

### 1. Biodata Peserta MBKM

#### a. Identitas Diri



|                       |   |                                      |
|-----------------------|---|--------------------------------------|
| Nama                  | : | Muhammad Zaidan Rafat                |
| NIM                   | : | 231730016                            |
| Tempat, Tanggal Lahir | : | Serang, 19 Juli 2005                 |
| Program Studi         | : | Informatika                          |
| Fakultas              | : | Sains dan Teknologi                  |
| Universitas           | : | UIN Sultan Maulana Hasanuddin Banten |
| Email                 | : | muhammadzaidanrafat@gmail.com        |

#### b. Pendidikan

| Jenjang        | S1                                   | SMA / SMK Sederajat          |
|----------------|--------------------------------------|------------------------------|
| Nama Institusi | UIN Sultan Maulana Hasanuddin Banten | SMK Negeri 1 Kota Serang     |
| Program Studi  | Informatika                          | Teknik Komputer dan Jaringan |
| Tahun Masuk    | 2023                                 | 2020                         |
| Tahun Keluar   | -                                    | 2023                         |

#### c. Pengalaman

|                       |   |                                     |                                      |
|-----------------------|---|-------------------------------------|--------------------------------------|
| Perusahaan / Instansi | : | PT. Krakatau Information Technology | Badan Riset dan Inovasi Nasional     |
| Satuan Kerja          | : | IT Infrastructure                   | Pusat Riset Sains Data dan Informasi |
| Posisi                | : | Staff Magang                        | Magang Riset                         |
| Tahun                 | : | Juni 2022 - Agustus 2022            | Februari 2026 - Juli 2026            |
| Alamat                | : | Gedung Krakatau IT, Jl. Raya        | Jl. Sangkuriang, Dago,               |

|  |  |                                                         |                                    |
|--|--|---------------------------------------------------------|------------------------------------|
|  |  | Anyer No.Km.3, Warnasari,<br>Citangkil, Cilegon, Banten | Kecamatan Coblong, Kota<br>Bandung |
|--|--|---------------------------------------------------------|------------------------------------|

**d. Penghargaan**

| No | Jenis Penghargaan                                  | Institusi Pemberi Penghargaan         | Tahun |
|----|----------------------------------------------------|---------------------------------------|-------|
| 1  | Awardee Bank Indonesia<br>Scholarship 2024 Batch 2 | KPW Bank Indonesia<br>Provinsi Banten | 2024  |

**e. Hak Paten / HKI**

| Tahun | Judul HKI                                                                                                                             | Jenis     | Nomor P/Id     |
|-------|---------------------------------------------------------------------------------------------------------------------------------------|-----------|----------------|
| 2024  | Rancang Bangun Sistem Informasi<br>Pencatatan Jurnal Mengajar Dosen<br>Berbasis Web : Solusi Monitoring<br>Perkuliahan Di Era Digital | Hak Cipta | EC002024238345 |

## 2. Biodata Pembimbing BRIN

### a. Identitas Diri

|    |                             |                                                                             |
|----|-----------------------------|-----------------------------------------------------------------------------|
| 1  | Nama Lengkap (dengan gelar) | Rio Nurtanyana, S.Pd., M.Pd., M.Sc., Ph.D                                   |
| 2  | Jenis Kelamin               | Laki-Laki                                                                   |
| 3  | NIP/NIK/Identitas lainnya   | 199312152024031001                                                          |
| 4  | NIDN                        | -                                                                           |
| 5  | Tempat dan Tanggal Lahir    | Yogyakarta, 15 Desember 1993                                                |
| 6  | E-mail                      | rion003@brin.go.id                                                          |
| 7  | Nomor Telepon/HP            | +62 851-3215-5491                                                           |
| 8  | Nama Institusi Tempat Kerja | Pusat Riset Sains Data dan Informasi, BRIN<br>KST Samaun Samadikun          |
| 9  | Alamat Kantor               | Jl. Sangkuriang, Dago, Kecamatan Coblong,<br>Kota Bandung, Jawa Barat 40135 |
| 10 | Nomor Telepon/Faks          | (022) 2530050                                                               |

### b. Riwayat Pendidikan

|                                | S-1                                                                                                                                           | S-2 (Indonesia)                                                                                                   | S-2 (Taiwan)                                                                                                      | S-3                                                                                             |
|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| Nama Perguruan Tinggi          | Universitas Negeri Yogyakarta, Indonesia                                                                                                      | Universitas Negeri Yogyakarta, Indonesia                                                                          | National Central University, Taiwan                                                                               | National Central University, Taiwan                                                             |
| Bidang Ilmu                    | Pendidikan Teknik Elektronika dan Informatika                                                                                                 | Pendidikan Teknik Elektronika dan Informatika                                                                     | Network Learning Technology                                                                                       | Network Learning Technology                                                                     |
| Tahun Masuk-Lulus              | 2012-2016                                                                                                                                     | 2017-2019                                                                                                         | 2017-2019                                                                                                         | 2019-2023                                                                                       |
| Judul Skripsi/Tesis/Dissertasi | Pengembangan Aplikasi Mobile Direktori Smk Sebagai Informasi Lokasi Sekolah Menengah Kejuruan (Smk) Di Kota Yogyakarta Pada Platform Android. | Application of the Authentic-UG using Augmented Reality to Support Geometry Learning in Authentic Learning Design | Application of the Authentic-UG using Augmented Reality to Support Geometry Learning in Authentic Learning Design | X-Education: Education of All Things with AI and Edge Computing—One Case Study for EFL Learning |
| Nama Pembimbing/Promotor       | Handaru Jati, Ph.D.                                                                                                                           | 1. Prof. Dr. Wu-Yuin Hwang                                                                                        | 2. Suprpto, Ph.D                                                                                                  | Prof. Dr. Wu-Yuin Hwang                                                                         |

**c. Pengalaman Penelitian 5 Tahun Terakhir**

| No. | Tahun | Judul Penelitian                                                                                                                                              | Sumber Pendanaan |
|-----|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| 1   | 2023  | AI and Recognition Technologies to Facilitate English as Foreign Language Writing for Supporting Personalization and Contextualization in Authentic Contexts. | NSTC, Taiwan     |
| 2   | 2023  | Collaboration and interaction with smart mechanisms in flipped classrooms.                                                                                    | NSTC, Taiwan     |
| 3   | 2023  | Investigating anxiety and acceptance of students and parents toward authentic contextual learning with a mobile app under the COVID-19.                       | NSTC, Taiwan     |
| 4   | 2023  | X-Education: Education of All Things with AI and Edge Computing - one Case Study for EFL Learning.                                                            | NSTC, Taiwan     |
| 5   | 2023  | Augmented Reality with Authentic GeometryGo App to Help Geometry Learning and Assessments.                                                                    | NSTC, Taiwan     |

**d. Publikasi Artikel Ilmiah Dalam Jurnal 5 Tahun Terakhir**

| No. | Judul Artikel Ilmiah                                                                                                                                                                                                                                                                                                                                                                  | Nama Jurnal                                | Penerbit |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|----------|
| 1   | Hwang, W.-Y., <b>Nurtanyana, R.</b> , Purba, S.W.D., Hariyanti, U., Indrihapsari, Y., & Surjono, H. D. (2023). AI and Recognition Technologies to Facilitate English as Foreign Language Writing for Supporting Personalization and Contextualization in Authentic Contexts.<br><a href="http://dx.doi.org/10.1177/07356331221137253">http://dx.doi.org/10.1177/07356331221137253</a> | Journal of Educational Computing Research. | SAGE     |
| 2   | Hwang, W.-Y., <b>Nurtanyana, R.</b> , & Hariyanti, U. (2023). Collaboration and interaction with smart mechanisms in flipped classrooms.<br><a href="https://doi.org/10.1108/DTA-04-2022-0171">https://doi.org/10.1108/DTA-04-2022-0171</a>                                                                                                                                           | Data Technologies and Applications.        | Emerald  |

|   |                                                                                                                                                                                                                                                                                                                                                                                                    |                                            |                  |
|---|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|------------------|
| 3 | Hariyanti, U., Hwang, W.-Y., <b>Nurtanyana, R.</b> , Surjono, H. D., Rochmah, I. N., Hoang, A., Luthfi, M. I., & Maulana Putra, M. T. (2023). Investigating anxiety and acceptance of students and parents toward authentic contextual learning with a mobile app under the COVID-19. 1-13.<br><a href="https://doi.org/10.1080/1049482022.2160465">https://doi.org/10.1080/1049482022.2160465</a> | Interactive Learning Environment s.        | Taylor & Francis |
| 4 | Hwang, W.-Y., & <b>Nurtanyana, R.</b> (2022). X-Education: Education of All Things with AI and Edge Computing -one Case Study for EFL Learning., 14(19), 12533.<br><a href="https://doi.org/10.3390/su141912533">https://doi.org/10.3390/su141912533</a>                                                                                                                                           | Sustainability                             | MDPI             |
| 5 | Hwang, W.-Y., Nurtanyana, R., Purba, S.W.D., Hariyanti, & Suprpto. (2023). Augmented Reality with Authentic GeometryGo App to Help Geometry Learning and Assessments.<br><a href="https://doi.org/10.1109/TLT.2023.3251398">https://doi.org/10.1109/TLT.2023.3251398</a>                                                                                                                           | IEEE Transactions on Learning Technologies | IEEE             |

**e. Pemakalah Seminar Ilmiah (Oral Presentation) dalam 5 Tahun**

**Terakhir**

| No. | Nama Temu Ilmiah/Seminar                                                                        | Judul Artikel Ilmiah                                                                                                                                                                                                                                    | Waktu dan Tempat |
|-----|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| 1   | 6th International Conference on Information Technology (InCIT2022). IEEE.                       | Hwang, W.-Y., & Nurtanyana, R. (2022). The Integration of Multiple Recognition Technologies and Artificial Intelligence to Facilitate EFL Writing in Authentic Contexts.                                                                                | Thailand, 2022   |
| 2   | Innovative Technologies and Learning (ICITL 2021), Lecture Notes in Computer Science. Springer. | Hwang, W.-Y., Lin, Y.-J., Hoang, A., Nurtanyana, R., & Lin, O. (2022). Facilitating Geometry Learning Through Real-Time Collaborative Activities with Augmented Reality in Authentic Context. In Y.-M. Huang, S.-C. Cheng, J. Barroso, & F. E. Sandnes. | Portugal, 2021   |

|   |                                                                                                                                                                                                                            |                                                                                                                                                                                     |                  |
|---|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| 3 | Innovative Technologies and Learning (ICITL 2021), Lecture Notes in Computer Science. Springer.                                                                                                                            | Hwang, W.-Y., Nurtanyana, R., & Putra, M.T.M. (2021). Facilitating 3D Geometry Learning with Augmented Reality in Authentic Contexts. In Y.-M. Huang, C.-F. Lai, & T. Rocha (Eds.). | Portugal, 2021   |
| 4 | 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress. IEEE. | Hwang, W.-Y., Nurtanyana, R., & Hildayati, J. (2021). Investigation of students' understanding of speech with smart mechanisms and scaffolding in flipped classrooms.               | Canada, 2021     |
| 5 | International Conferences Blended Learning 2020. Atlantis.                                                                                                                                                                 | Nurtanyana, R., Suprpto, Hwang, W. Y., & Hariyanti, U. (2020). Developing a Learning Monitoring System Dashboard for Augmented Reality in Ubiquitous Geometry (Authentic-UG).       | Yogyakarta, 2020 |

**f. Karya Buku dalam 5 Tahun Terakhir**

| No. | Judul Buku                                                                                                                     | Tahun | Jumlah Halaman | Penerbit |
|-----|--------------------------------------------------------------------------------------------------------------------------------|-------|----------------|----------|
| 1   | AI and recognition technologies for authentic contextual learning (Book chapter in Education in the era of digital technology) | 2024  | 10             | FIP UPI  |

**g. Perolehan HKI**

| No. | Judul/Tema HKI                                                                                             | Tahun | Jenis     | No P/ID       |
|-----|------------------------------------------------------------------------------------------------------------|-------|-----------|---------------|
| 1   | BATIKNESIA BERBASIS AR PADA PLATFORM IOS DAN ANDROID                                                       | 2013  | Hak Cipta | C00201301766  |
| 2   | Program Komputer "APLIKASI REKOMENDASI & INFORMASI LOKASI SEKOLAH MENENGAH KEJURUAN (SMK) "DIREKTORI SMK"" | 2016  | Hak Cipta | C00201703984  |
| 3   | Hibona Smartswitch                                                                                         | 2018  | Hak Cipta | EC00201851487 |
| 4   | Arcapada                                                                                                   | 2021  | Merek     | IDM000862782  |
| 5   | HiBona                                                                                                     | 2020  | Merek     | IDM000752162  |
| 6   | Batiknesia                                                                                                 | 2017  | Merek     | IDM000577804  |

**h. Pengalaman Merumuskan Kebijakan Publik/Rekayasa Sosial Lainnya dalam 10 Tahun Terakhir**

| No. | Judul/Tema/Jenis Rekayasa Sosial Lainnya yang Telah Diterapkan | Tahun | Tempat Penerapan | Respon Masyarakat |
|-----|----------------------------------------------------------------|-------|------------------|-------------------|
| 1   | -                                                              | -     | -                | -                 |

**i. Penghargaan dalam 10 tahun Terakhir (dari pemerintah, asosiasi atau institusi lainnya)**

| No. | Jenis Penghargaan                    | Institusi Pemberi Penghargaan | Tahun |
|-----|--------------------------------------|-------------------------------|-------|
| 1   | Phi Tau Phi scholastic honor society | Phi Tau Phi                   | 2023  |

### 3. Biodata Pembimbing Universitas

#### a. Identitas Diri



|                       |   |                                                             |
|-----------------------|---|-------------------------------------------------------------|
| Nama                  | : | Ahmad Tabrani                                               |
| Tempat, Tanggal Lahir | : | Tangerang, 17 September 1985                                |
| Pekerjaan             | : | Dosen / PNS                                                 |
| Email                 | : | Ahmad.tabrani85@gmail.com;<br>ahmad.tabrani@uinbanten.ac.id |
| NIDN                  | : | 2017098503                                                  |
| SINTA ID              | : | 6694015                                                     |

#### b. Riwayat Pendidikan

| Tahun           | Jenjang / Program Studi       | Institusi                            |
|-----------------|-------------------------------|--------------------------------------|
| 2024 – Sekarang | S3 Informatika                | Universitas Amikom Yogyakarta        |
| 2022 – Sekarang | S3 Manajemen Pendidikan Islam | UIN Sultan Maulana Hasanuddin Banten |
| 2019 – 2022     | S2 Ekonomi Islam (ME)         | UIN Sultan Maulana Hasanuddin Banten |

|                |                                |                       |
|----------------|--------------------------------|-----------------------|
| 2014 –<br>2016 | S2 Teknik Informatika<br>(MTI) | STMIK Raharja         |
| 2003 –<br>2007 | S1 Teknik Informatika<br>(ST)  | Universitas Gunadarma |

**c. Pengalaman Pekerjaan**

| Tahun              | Jabatan                | Instansi / Perusahaan                |
|--------------------|------------------------|--------------------------------------|
| 2018 –<br>Sekarang | Dosen                  | UIN Sultan Maulana Hasanuddin Banten |
| 2012 –<br>2018     | Section Head IT        | PT Dharma Poliplast                  |
| 2011 –<br>2012     | Senior System Engineer | PT Intikom Berlian Mustika           |
| 2007 –<br>2011     | Network Engineer       | PT Data Sari Citra                   |

**d. Publikasi Jurnal 5 Tahun Terakhir**

| Tahun | Judul Publikasi                                                                                   | Jurnal / Publikasi                                    |
|-------|---------------------------------------------------------------------------------------------------|-------------------------------------------------------|
| 2024  | Rancang Bangun Aplikasi Manajemen Simpatisan Sebagai Optimalisasi Pemenangan Pemilu Legislatif    | SIMADA                                                |
| 2024  | Analisis Bibliometrik Penggunaan Chat GPT Berbasis Data Google Scholar                            | SIMIKA                                                |
| 2023  | Digitalisasi Pelayanan Akademik Melalui Aplikasi Simika Pada UIN Sultan Maulana Hasanuddin Banten | ICIT Journal 9                                        |
| 2022  | Analisis Kritis Kebijakan Pemerintah Tentang Undang-Undang Sisdiknas Nomor 20 Tahun 2003          | JPPM Jurnal Penelitian dan Pengabdian Masyarakat 1(1) |

|      |                                                                                                                                                |                                                                             |
|------|------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| 2022 | Analisis Faktor-Faktor Yang Mempengaruhi Keputusan Nasabah Melakukan Cicilan Emas Di Bank Syariah Indonesia KCP Pandeglang                     | MUAMALATU NA 14(1)                                                          |
| 2021 | Initiating Agropreneurship From An Islamic Perspective                                                                                         | -                                                                           |
| 2021 | Designing a Pustekipad Computer Laboratory Loan Application for Learning ICT Subjects                                                          | MOTIVECTION: Journal of Mechanical, Electrical and Industrial Engineering 3 |
| 2020 | Analisis Faktor yang Mempengaruhi Agresivitas Pajak Studi Empiris pada Perusahaan LQ 45 yang Terdaftar di Bursa Efek Indonesia Tahun 2015–2018 | Jurnal REKOMEN: Riset Ekonomi Manajemen 4(1)                                |
| 2020 | Pengaruh Tingkat Suku Bunga Deposito Terhadap Dana Pihak Ketiga Di Bank BJB Cabang Rangkasbitung                                               | Jurnal Manajemen Bisnis Krisnadwipayana 8(2)                                |

**e. Publikasi Buku**

| Tahun | Judul Buku                                                                                                | Penerbit      |
|-------|-----------------------------------------------------------------------------------------------------------|---------------|
| 2021  | Kampus Bebas Kertas: Studi atas Pelayanan Akademik Secara Digital di UIN Sultan Maulana Hasanuddin Banten | Farha         |
| 2022  | Jurus Jitu Meningkatkan Webometrics                                                                       | Harfa Kreatif |
| 2024  | Basis Data                                                                                                | Ruang Karya   |

|      |                   |             |
|------|-------------------|-------------|
| 2024 | Jaringan Komputer | Ruang Karya |
|------|-------------------|-------------|

**f. Hak Paten / HKI**

| Tahun | Judul HKI                                                                                                                    | Jenis     | Nomor P/Id     |
|-------|------------------------------------------------------------------------------------------------------------------------------|-----------|----------------|
| 2024  | Rancang Bangun Sistem Informasi Pencatatan Jurnal Mengajar Dosen Berbasis Web : Solusi Monitoring Perkuliahan Di Era Digital | Hak Cipta | EC002024238345 |
| 2021  | Kampus Bebas Kertas: Studi Atas Pelayanan Akademik Secara Digital di UIN Sultan Maulana Hasanuddin Banten                    | Hak Cipta | EC00202182858  |

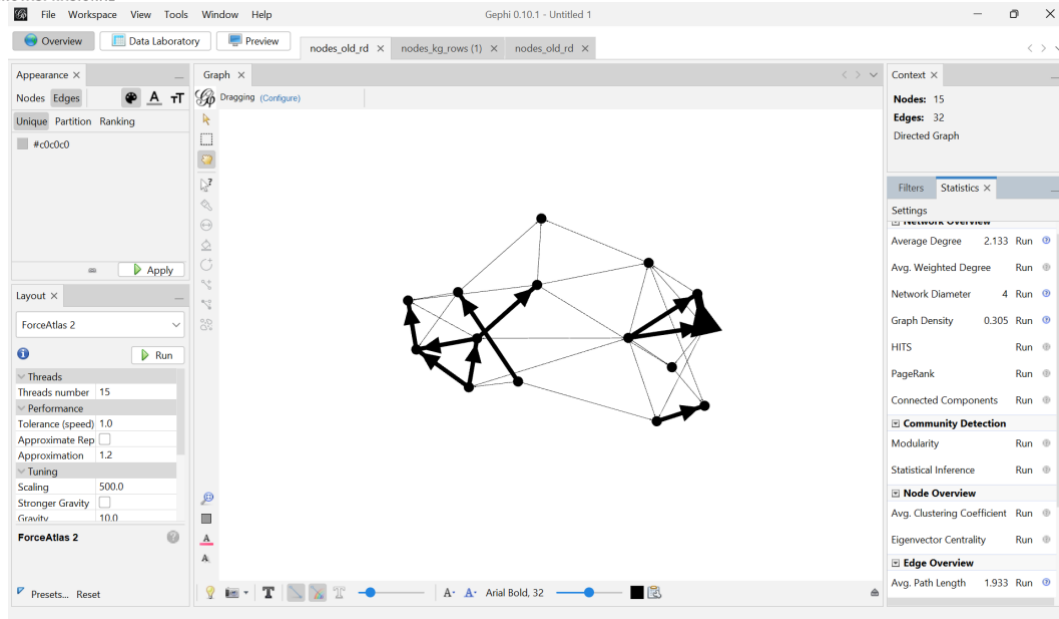
## Lampiran Uji Komparasi Knowledge Graph Menggunakan Gephi

### Hasil Analisis Gephi

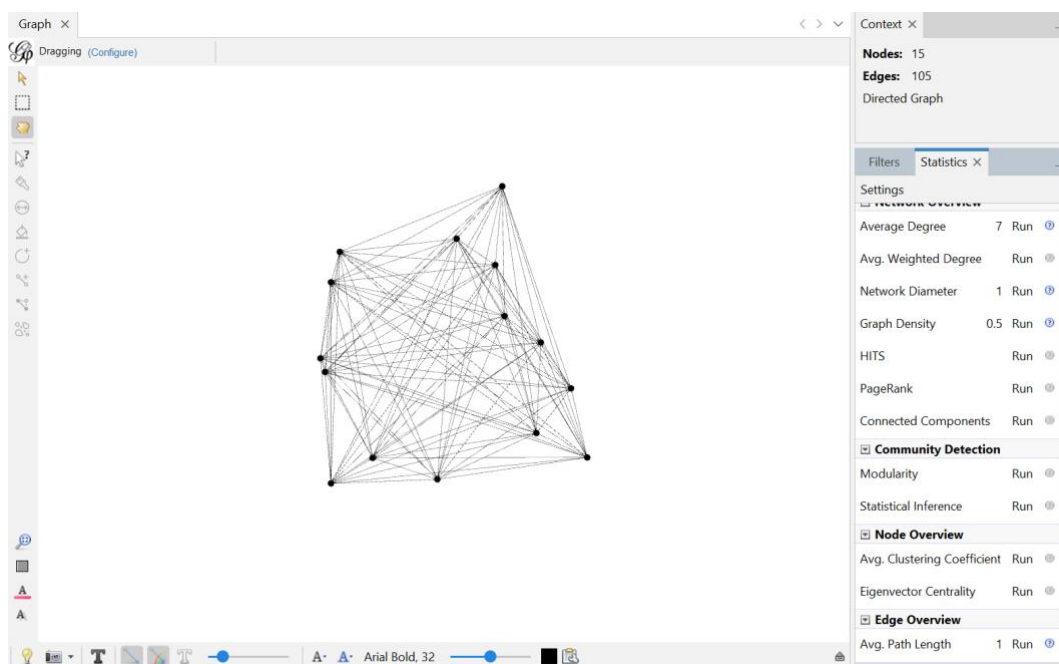
| Poin Analisis    | SRE-v.1 | SRE-v.2 |
|------------------|---------|---------|
| Nodes            | 15      | 15      |
| Edges            | 32      | 105     |
| Avg. Degree      | 2.13    | 7       |
| Network Diameter | 4       | 1       |
| Graph Density    | 0.305   | 0.5     |
| Avg. Path Length | 1.933   | 1       |

### Penjelasan Analisis

| Poin Analisis    | Penjelasan                                                                                                                                                                                   |
|------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Nodes            | Proses ekstraksi nodes menggunakan artikel yang sama dengan jumlah 15                                                                                                                        |
| Edges            | Edges yang terbentuk lebih banyak menunjukkan SRE terbaru memiliki edges lebih kompleks dan ekspansi hubungan pada antar nodes                                                               |
| Avg. Degree      | Peningkatan rata-rata jumlah koneksi antar nodes dipengaruhi oleh banyaknya edges yang terbentuk. Hal ini mengindikasikan jaringan knowledge graph yang lebih padat                          |
| Network Diameter | Penurunan jalur terpanjang antar node menunjukkan menandakan adanya peningkatan efisiensi komunikasi antar entitas artikel                                                                   |
| Graph Density    | Peningkatan kepadatan jaringan pada knowledge graph terbaru mengindikasikan pengetahuan yang lebih terkoneksi dan potensi inferensi antar node lebih luas                                    |
| Avg. Path Length | Panjang rata-rata jalur antar node menjadi satu langkah menunjukkan efisiensi yang terlihat di network diameter dan memperlihatkan bahwa Knowledge Graph baru memiliki struktur yang optimal |



SRE v1



SRE v2

## Dokumentasi



*Dokumentasi 1: Silaturahmi Perdana (Jumat, 20 Februari 2026)*



*Dokumentasi 2: Meeting Progress (Selasa, 24 Februari 2026)*



*Dokumentasi 3: Meeting Progress (Senin, 02 Maret 2026)*



*Dokumentasi 4: Meeting Progress (Senin, 9 Maret 2026)*



Dokumentasi 5: Meeting Progress (Senin 30 Maret 2026)



11:01 | 06 Apr 2026  
Mon  
Jl. Cisitua Lama No.12, Dago, Kecamatan  
Coblong, Kota Bandung, Jawa Barat  
40135

Dokumentasi 6: Meeting Progress (Senin, 06 April 2026)



Dokumentasi 7: Meeting Progress (Senin, 13 April 2026)



Dokumentasi 8: Meeting Progress (Senin, 20 April 2026)



Dokumentasi 9: Meeting Progress (Senin, 28 April 2026)



Dokumentasi 10: Meeting Progress (Selasa, 05 Mei 2026)



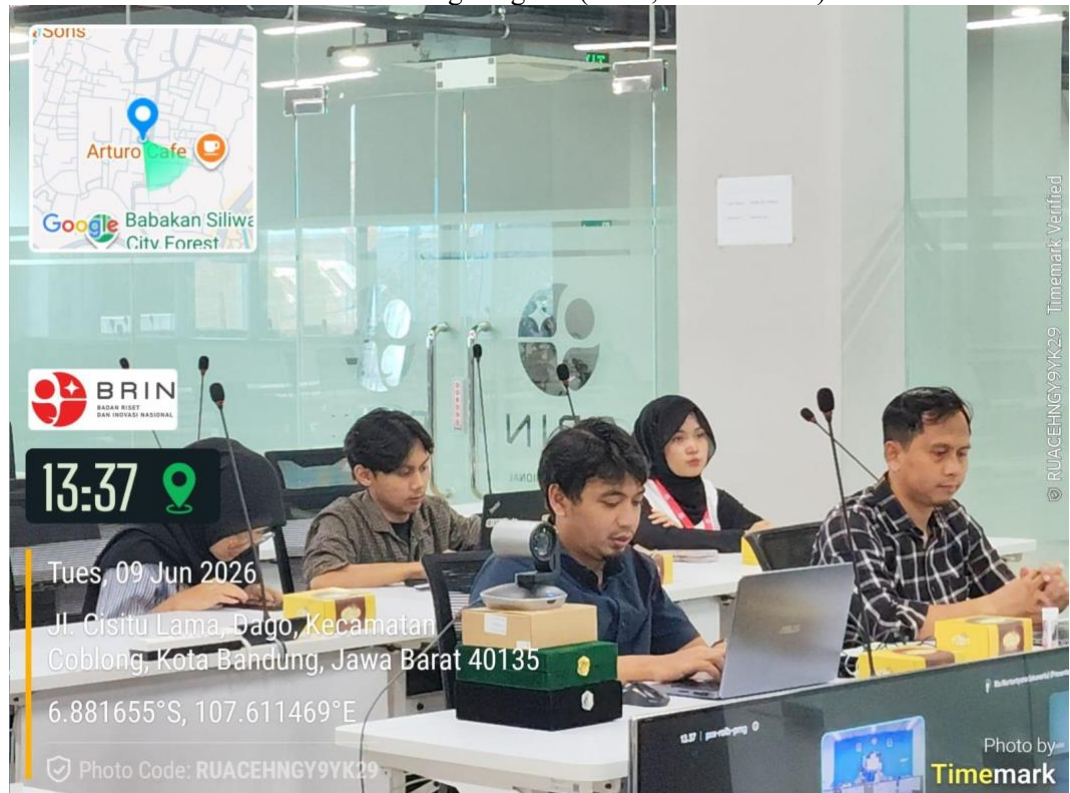
Dokumentasi 11: Meeting Progress (Senin, 11 Mei 2026)



Dokumentasi 12: Meeting Progress (Senin, 18 Mei 2026)



Dokumentasi 13: Meeting Progress (Rabu, 03 Juni 2026)



Dokumentasi : Demonstrasi feat Expert (Selasa, 09 Juni 2026)