

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Transformasi menuju ekonomi digital telah memosisikan data sebagai sumber daya yang keberadaannya kini sukar untuk digantikan. Sepanjang tahun 2025, volume data global yang tercipta diperkirakan mencapai 173,4 zettabyte dan diproyeksikan terus bertumbuh hingga kisaran 230–240 zettabyte pada tahun 2026; lonjakan eksponensial tersebut menjadi cerminan dari pesatnya digitalisasi aktivitas manusia di berbagai belahan dunia [1]. Tidak mengherankan apabila dalam literatur bisnis dan teknologi data kerap dianalogikan sebagai *the new oil*, yakni komoditas paling bernilai pada abad ke-21 yang, ketika diekstraksi dan diolah secara tepat, mampu melahirkan wawasan strategis untuk menopang pengambilan keputusan perusahaan, baik dalam hal peningkatan pendapatan, penguatan loyalitas pelanggan, maupun efisiensi belanja iklan [2]. Himpunan data yang dicirikan oleh volume yang besar, variasi yang beragam, serta kecepatan produksi yang tinggi inilah yang selanjutnya tergolong sebagai *big data* [3]. Signifikansi pengelolaan data bagi dunia bisnis modern turut tergambar dari nilai pasar analitik *big data* global yang saat ini berada pada kisaran \$348,21 miliar dan diperkirakan menembus \$924,39 miliar pada tahun 2032 [1]. Kendati demikian, di balik besarnya potensi tersebut, karakteristik *big data* sekaligus memunculkan sejumlah tantangan nyata yang mencakup aspek keamanan, kualitas, integritas, kapasitas penyimpanan, hingga kompleksitas analisisnya [4].

Sebagai konglomerasi multisektor yang menaungi unit bisnis pada bidang manufaktur, industrial, ritel, properti, makanan dan minuman, hingga teknologi komersial, Kawan Lama Group pun tidak terlepas dari kompleksitas *big data* sebagaimana telah diuraikan sebelumnya. Keberagaman ekosistem bisnis tersebut menjadikan volume serta variasi data yang dihasilkan sarat dengan tantangan, mulai dari persoalan integritas, kebersihan, sampai keamanan data di seluruh lini usaha. Menyikapi kondisi ini, perusahaan membentuk tim *Master Data Management* (MDM) Analyst Officer yang diberi mandat untuk menjaga kualitas data secara menyeluruh sekaligus memastikan data berperan sebagaimana mestinya, yaitu sebagai aset strategis penopang pengambilan keputusan korporat. Pemilihan Kawan Lama Group sebagai lokasi pelaksanaan kerja magang dilandasi oleh kesadaran bahwa kualitas data yang tinggi merupakan prasyarat utama agar *big data* mampu menghasilkan wawasan yang bernilai, sebuah urgensi yang tercermin pula dari

proyeksi pasar analitik global yang terus tumbuh pesat. Melalui posisi *Master Data Management* (MDM) Analyst Officer, terbuka kesempatan untuk terlibat secara langsung dalam menghadapi tantangan pengelolaan data berskala korporat konglomerasi, sekaligus membangun fondasi kompetensi yang kuat pada bidang *data quality management* sebagai bekal pengembangan karier di ranah data pada masa mendatang.

1.2 Maksud dan Tujuan Kerja Magang

Pelaksanaan kerja magang di Kawan Lama Group pada posisi *Master Data Management* (MDM) Analyst Officer merupakan pemenuhan salah satu kewajiban akademik yang harus ditempuh mahasiswa Universitas Multimedia Nusantara untuk menyelesaikan program studinya. Program ini membuka peluang bagi mahasiswa untuk memperoleh pengalaman profesional secara langsung di lingkungan kerja, sekaligus mengaplikasikan pengetahuan dan keterampilan yang telah diperoleh selama perkuliahan, terutama pada ranah pengelolaan dan analisis data. Keterlibatan tersebut memungkinkan mahasiswa memahami praktik pengelolaan data pada perusahaan berskala besar yang menaungi beragam unit bisnis dengan kebutuhan data yang bervariasi. Lebih lanjut, peran sebagai *Master Data Management* (MDM) Analyst Officer memberi ruang untuk turut serta dalam proses pengelolaan, pengolahan, serta peningkatan kualitas data yang dimanfaatkan oleh berbagai unit bisnis perusahaan. Aspek ini menjadi krusial mengingat kualitas dan konsistensi data merupakan salah satu faktor penentu efektivitas pengambilan keputusan bisnis.

Secara lebih terperinci, kegiatan kerja magang ini diarahkan pada beberapa tujuan berikut. Tujuan pertama adalah membantu tim *Master Data Management* dalam membenahi sekaligus meningkatkan kualitas data perusahaan melalui aktivitas penggabungan, pembersihan, serta pengelolaan data yang bersumber dari beragam unit bisnis. Tujuan kedua adalah menunjang tim *Master Data Management* (MDM) dalam memenuhi kebutuhan kolaboratif para *stakeholder*, yakni dengan menyajikan informasi dalam bentuk *dashboard*, skenario analisis, maupun visualisasi data lain yang menopang proses *monitoring* dan pengambilan keputusan. Di samping itu, keterlibatan pada berbagai proyek lintas unit bisnis turut membuka kesempatan untuk memahami kompleksitas pengelolaan data dalam lingkungan korporasi. Dari rangkaian pengalaman tersebut, mahasiswa tidak sekadar mengasah kemampuan teknis dalam pengolahan data, tetapi juga

menumbuhkan kompetensi profesional seperti adaptabilitas, kolaborasi dalam tim lintas fungsi, serta daya pikir kritis dan kecakapan pemecahan masalah ketika menghadapi beragam tantangan selama proyek berlangsung.

1.3 Waktu dan Prosedur Pelaksanaan Kerja Magang

Kegiatan kerja magang di Kawan Lama Group dijalankan selama lima bulan, terhitung sejak bulan Februari hingga Juni 2026. Adapun jam operasional kerja mengikuti ketentuan perusahaan, yakni pada setiap hari kerja mulai pukul 08:00 sampai pukul 17:00 WIB atau mulai pukul 09.00 sampai pukul 18.00 WIB.

Prosedur pelaksanaan magang dititikberatkan pada tahapan riset, perancangan, serta implementasi automasi untuk pengayaan data (*data enrichment*) dan pengelolaan *Master Data*. Pendekatan yang ditempuh memanfaatkan sejumlah model *Artificial Intelligence* (AI), antara lain *Retrieval-Augmented Generation* (RAG), SBERT, TF-IDF, serta *Large Language Models* (LLM). Rincian *timeline* dari masing-masing tahapan tersebut dipaparkan pada Tabel 1.1.



Tabel 1.1. Timeline Kegiatan Implementasi Pipeline Data & AI

Deskripsi Kegiatan	Februari				Maret				April				Mei				Juni			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. Mengembangkan arsitektur automasi end-to-end menggunakan n8n dan Python untuk ekstraksi dan klasifikasi data vendor secara otomatis.																				
2. Mengimplementasikan Semantic Matching dan pendekatan Hybrid Ensemble (SBERT, TF-IDF) untuk pencocokan deskripsi bisnis secara cerdas.																				
3. Mengintegrasikan berbagai Large Language Model (OpenAI GPT, Claude Haiku) dan API Pencarian web untuk proses pengayaan data.																				
4. Membangun pipeline Retrieval-Augmented Generation (RAG) dengan indeks FAISS dan BM25 untuk pemetaan aktivitas bisnis ke kode KBLI.																				
5. Mengoptimalkan pipeline preprocessing teks dan normalisasi data master menggunakan Regular Expression (Regex) pada Python.																				
6. Menyusun visualisasi dan laporan untuk mengevaluasi tingkat akurasi model AI dalam proses standarisasi klasifikasi industri.																				
7. Mengintegrasikan <i>Knowledge Graph</i> , <i>Confidence Classification</i> , dan pelacakan <i>Token Usage</i> API untuk validasi akhir pipeline.																				