

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi informasi dan *Large Language Models* (LLM) telah mempermudah penyelesaian tugas administratif rutin, seperti peringkasan dokumen dan pengelompokan data, sehingga mengurangi kebutuhan entri data secara manual. Namun, pada sektor pengadaan digital, proses pengolahan dokumen yang bervolume tinggi dan berulang masih sangat bergantung pada tenaga manusia. Ketergantungan ini mengakibatkan rentannya terjadi inkonsistensi klasifikasi data antara satu periode dengan periode berikutnya [1].

Namun, pemanfaatan layanan kecerdasan buatan pihak ketiga yang luput dari pengawasan divisi IT berpotensi memicu risiko kebocoran data internal perusahaan [2]. Kerangka kerja mitigasi risiko untuk AI generatif sangat menekankan pentingnya penggunaan model yang dapat diaudit dan sejalan dengan kebijakan tata kelola data [3, 4]. Oleh karena itu, diperlukan pengembangan *pipeline* dan infrastruktur LLM pada platform internal guna mencegah pegawai menggunakan aplikasi eksternal yang tidak berizin.

Kegiatan magang ini dilaksanakan di PT Rekayasa Analisa Digital pada divisi *Product Development*. Laporan ini merangkum perencanaan dan implementasi *pipeline* ETL berbasis LLM (khususnya pada tahap prapemrosesan) untuk meningkatkan efisiensi pengolahan data pengadaan di dalam koridor kerja organisasi [5]. Melalui pengembangan ini, diharapkan ketergantungan pada proses manual dapat dipangkas secara signifikan. Selain itu, pengembangan ini memastikan penggunaan teknologi LLM tetap terkendali dan sejalan dengan kebijakan keamanan internal. Seluruh sistem ini berjalan di atas ekosistem Erica, yaitu sebuah platform kecerdasan buatan (AI) terpadu milik perusahaan yang berfungsi sebagai lapisan analitik untuk memfasilitasi pengolahan data bisnis secara aman.

1.2 Maksud dan Tujuan Kerja Magang

Pelaksanaan kerja magang ini memiliki maksud sebagai berikut:

1. Meningkatkan pemahaman serta keterampilan teknis sebagai *AI Engineer*

melalui pengalaman langsung dalam perencanaan dan implementasi arsitektur integrasi data.

2. Mendukung efisiensi pemrosesan data di PT Rekayasa Analisa Digital dengan merancang dan mengimplementasikan *pipeline* ETL terautomasi berteknologi kecerdasan buatan (AI).

Adapun tujuan utama dari pelaksanaan kerja magang ini adalah melakukan perencanaan dan implementasi *Pipeline* ETL Terautomasi sebagai *AI Engineer* di PT Rekayasa Analisa Digital. Secara spesifik, proyek ini berfokus pada integrasi teknologi *Large Language Models* (LLM) dan Apache Airflow untuk mengoptimalkan otomatisasi kategorisasi data pengadaan ke dalam platform Erica.

1.3 Waktu dan Prosedur Pelaksanaan Kerja Magang

Kerja magang dilaksanakan selama kurang lebih empat bulan, terhitung sejak tanggal 1 Februari 2026 - 30 Mei 2026. Seluruh pelaksanaan kerja magang dilakukan secara penuh dari rumah atau *Full Work From Home* (WFH), dengan waktu operasional kerja dimulai dari pukul 08.00 WIB hingga 18.00 WIB.

Prosedur pelaksanaan koordinasi dan evaluasi selama kerja magang didefinisikan sebagai berikut:

1. **Koordinasi Virtual Harian dan Mingguan:** Pertemuan secara virtual (*gathering*) dilakukan menggunakan platform kolaborasi virtual *gather.town* bersama *supervisor*, Akmal Dira Thursina, maupun bersama seluruh anggota tim divisi *Product Development* guna melaporkan progres harian dan mendiskusikan kendala teknis.
2. **Komunikasi dan Penugasan:** Pemberian instruksi tugas baru, diskusi teknis secara cepat, serta permintaan evaluasi kerja dilakukan secara daring melalui *gather.town* atau melalui kontak personal WhatsApp.
3. **Evaluasi Hasil Kerja:** Setiap fitur atau komponen *pipeline* yang telah diselesaikan dievaluasi langsung oleh *supervisor* secara virtual untuk memastikan kesesuaian dengan standar kualitas produk Erica di PT Rekayasa Analisa Digital.