

BAB I

PENDAHULUAN

1.1 Latar Belakang

PT Perkebunan Nusantara 1 (PTPN 1) adalah Badan Usaha Milik Negara (BUMN) yang bergerak di bidang usaha agroindustri perkebunan dengan wilayah operasional di beberapa Provinsi Indonesia. PTPN 1 memiliki sejarah panjang yang berawal pada tahun 1961 dengan nama PPN Kesatuan Aceh, yang asetnya berasal dari kebun-kebun swasta yang dinasionalisasi oleh pemerintah. Pada tahun 1968, namanya diubah menjadi PN Perkebunan. PTPN 1 secara resmi dibentuk pada 14 Februari 1996 berdasarkan Peraturan Pemerintah Nomor 6 Tahun 1996 melalui penggabungan kebun-kebun di Daerah Istimewa Aceh dari eks PTP I, PTP V, dan PT Cot Girek. Komoditas yang diusahakan PTPN 1 meliputi kelapa sawit, karet, dan kakao. Pada 1 Desember 2023, PTPN 1 menjadi bagian dari restrukturisasi besar dalam Holding Perkebunan Nusantara, PT Perkebunan Nusantara III (Persero) yang melibatkan penggabungan 13 perusahaan menjadi dua subholding utama: PalmCo dan SupportingCo. Dalam struktur baru tersebut, PTPN 1 berperan sebagai surviving entity untuk subholding SupportingCo yang menggabungkan PTPN 1, PTPN II, PTPN VII, PTPN VIII, PTPN IX, PTPN X, PTPN XI, PTPN XII, dan PTPN XIV. Penggabungan ini merupakan bagian dari skema Kementerian BUMN untuk efisiensi dan peningkatan kinerja keuangan serta operasional perseroan. Sejalan dengan restrukturisasi tersebut, PTPN 1 saat ini tengah menjalankan program transformasi digital sebagai bagian dari upaya modernisasi operasional dan peningkatan daya saing perusahaan. Transformasi ini dipandu oleh Rencana Strategis Teknologi Informasi (RSTI) 2025–2029. RSTI tersebut merupakan dokumen perencanaan strategis yang berfungsi sebagai peta jalan (*roadmap*) transformasi digital PTPN Group, mencakup berbagai *use case* dan *technology enabler* untuk mencapai sasaran strategis perusahaan menjadi Global Lighthouse.

Berdasarkan hasil wawancara dan observasi langsung terhadap karyawan Divisi Teknologi Informasi serta divisi pengguna di PTPN 1, yang juga ditegaskan dalam identifikasi kebutuhan (*requirements analysis*) yang tertuang dalam Project Charter Divisi Teknologi Informasi PTPN 1 (terlampir pada Lampiran), diketahui

bahwa salah satu tantangan operasional yang dihadapi adalah pengelolaan dan akses dokumen internal seperti Standar Operasional Prosedur (SOP), kebijakan perusahaan, dan peraturan internal. Karyawan seringkali mengalami kesulitan dalam menemukan informasi yang relevan dari berbagai dokumen yang tersebar, di mana proses pencarian manual membutuhkan waktu yang signifikan, yakni sekitar 3-10 menit untuk dokumen digital yang tersimpan di folder dan 15-30 menit untuk dokumen yang sudah diarsipkan. Waktu tersebut dihabiskan untuk menelusuri berkas satu per satu serta membaca isi dokumen guna memverifikasi kesesuaiannya dengan informasi yang dibutuhkan. Kondisi ini diperparah oleh fakta bahwa sebagian dokumen telah diarsipkan secara fisik dan belum tersedia dalam format digital, sehingga karyawan harus meluangkan waktu tambahan untuk mencari dokumen secara langsung di ruang arsip perusahaan (berdasarkan hasil wawancara dengan karyawan PTPN 1, 2025; dokumentasi lengkap pada Lampiran 7). Besarnya volume dokumen internal turut memperberat kondisi ini; berdasarkan data perusahaan, terdapat sekitar 939 dokumen tata kelola berupa SOP dan instruksi kerja, serta arsip persuratan dalam jumlah yang sangat besar yang dikelola melalui aplikasi naskah dinas elektronik (Nadine), yaitu 105.711 surat masuk internal, 21.347 surat masuk eksternal, 127.058 surat keluar internal, dan 6.221 surat keluar eksternal, atau total 260.337 surat. Temuan ini sejalan dengan hasil penelitian (Nakash & Dan, 2024) yang menyatakan bahwa sekitar sepertiga pekerja menghabiskan setengah hingga satu hari kerja penuh per minggu hanya untuk mencari informasi yang dibutuhkan.

Masalah ini diperberat dengan belum adanya sistem *chatbot* atau asisten virtual yang dapat membantu karyawan dalam menelusuri repositori dokumen secara otomatis. Metode konvensional yang mengandalkan pemeriksaan manual berisiko menghasilkan inkonsistensi (Muludi dkk., 2024), terutama ketika jumlah dokumen meningkat signifikan pasca-merger dengan bertambahnya cakupan operasional perusahaan. Temuan ini dikonfirmasi melalui hasil wawancara dengan narasumber dari Divisi Teknologi Informasi dan divisi pengguna serta keterlibatan langsung penulis dalam tahap analisis kebutuhan bersama Divisi TI, yang menciptakan urgensi akan sistem yang dapat menyediakan akses cepat dan akurat

terhadap dokumen internal melalui platform yang lebih familiar dan sejalan dengan inisiatif transformasi digital PTPN 1.

Secara lebih rinci, hasil wawancara dan analisis kebutuhan tersebut, sebagaimana juga terdokumentasi dalam Project Charter, mencatat lima permasalahan utama yang dihadapi perusahaan. Pertama, pencarian informasi dari dokumen internal seperti SOP dan kebijakan perusahaan masih dilakukan secara manual sehingga membutuhkan waktu yang lama. Kedua, karyawan sering tidak menemukan informasi yang tepat karena jumlah dokumen yang terus bertambah pasca-restrukturisasi perusahaan. Ketiga, dokumen internal tersebar di berbagai divisi dan tidak tersedia dalam satu sistem terpusat yang mudah diakses oleh seluruh karyawan. Keempat, belum ada sistem asisten virtual yang dapat membantu karyawan menelusuri dokumen internal secara otomatis. Kelima, pengelolaan dan pembaruan dokumen dalam sistem masih rumit dan sulit dilakukan oleh karyawan. Kelima permasalahan tersebut saling berkaitan dan bermuara pada satu kebutuhan yang sama, yaitu tersedianya sistem yang mampu menyediakan akses cepat, akurat, dan terpusat terhadap dokumen internal, sekaligus mudah digunakan oleh karyawan lintas divisi yang memiliki latar belakang dan kemampuan teknologi yang beragam. Rumusan kebutuhan pengguna dan kebutuhan sistem tersebut selanjutnya dituangkan secara terperinci dalam dokumen Spesifikasi Kebutuhan Perangkat Lunak (SKPL) pada Lampiran 6.

Berdasarkan laporan We Are Social dan Meltwater, Indonesia memiliki 185,3 juta pengguna internet pada awal tahun 2024, dengan WhatsApp menjadi platform yang paling banyak digunakan untuk komunikasi digital (Sue, 2025). Popularitas WhatsApp di Indonesia menjadikannya platform yang ideal untuk menjangkau pengguna dalam berbagai konteks, termasuk layanan informasi perusahaan. WhatsApp menawarkan *interface* yang familiar, dapat diakses melalui berbagai perangkat, dan mendukung berbagai jenis konten termasuk teks, gambar, dan dokumen. Integrasi *chatbot* RAG dengan WhatsApp memungkinkan karyawan mengakses informasi kapan saja dan di mana saja melalui perangkat yang sudah mereka gunakan sehari-hari, tanpa perlu mempelajari sistem baru atau menginstal aplikasi tambahan.

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) telah membawa transformasi signifikan dalam berbagai aspek kehidupan, termasuk dalam pengelolaan dokumen internal perusahaan. Dalam beberapa tahun terakhir, *Large Language Models* (LLMs) telah berkembang dari kerangka eksperimental menjadi alat praktis yang meningkatkan efisiensi dan kualitas layanan di berbagai sektor industri (Baur dkk., 2025). Kemampuan LLM dalam memproses dan menginterpretasi informasi kontekstual yang luas menjadikannya sangat relevan untuk aplikasi yang membutuhkan pemahaman mendalam terhadap dokumen dan komunikasi. Model bahasa modern saat ini mampu menghasilkan teks yang koheren dan panjang, menjawab pertanyaan, meringkas dokumen, menerjemahkan antar bahasa, hingga menulis konten kreatif serta memahami konteks dan nuansa kalimat (Fanani, 2025).

Arsitektur Transformer, yang memperkenalkan mekanisme attention, telah merevolusi pemodelan bahasa dengan memungkinkan model untuk fokus pada bagian-bagian berbeda dari teks *input* saat membuat prediksi (Vaswani dkk., 2023). Berbeda dengan model sekuensial seperti RNN dan LSTM, Transformer dapat memproses seluruh urutan sekaligus dan menentukan bagian mana yang paling relevan untuk memahami setiap kata (Sanjani dkk., 2025). Kemampuan pemrosesan paralel ini mempercepat proses pelatihan secara drastis, serta memungkinkan pengembangan model bahasa berukuran besar seperti yang saat ini dikenal sebagai *Large Language Model* (LLM) (Fanani, 2025). Perkembangan LLM seperti GPT dari OpenAI, LLaMA dari Meta, dan Gemini dari Google telah membuka peluang besar dalam pengembangan aplikasi berbasis kecerdasan buatan di berbagai domain.

Salah satu tantangan utama dalam implementasi LLM adalah kecenderungan untuk menghasilkan informasi yang tidak akurat atau *hallucinations*, terutama ketika model dihadapkan pada skenario di luar data pelatihan mereka (Sanjani dkk., 2025). Fenomena halusinasi ini dapat menimbulkan masalah serius dalam lingkungan profesional, di mana respons yang tidak akurat dapat merusak kepercayaan pengguna dan mengakibatkan keputusan yang keliru (Sanjani dkk., 2025). Model yang hanya mengandalkan pengetahuan yang tertanam dalam data latih-nya tidak dapat mengakses informasi terbaru atau

informasi spesifik domain tertentu, sehingga berisiko memberikan jawaban yang tidak tepat (Lewis dkk., 2021). Kondisi ini menjadi lebih krusial ketika LLM digunakan untuk menjawab pertanyaan terkait informasi sensitif, seperti kesehatan, hukum, dan regulasi perusahaan, yang menuntut akurasi dan keandalan tinggi (Anam dkk., 2024; Samudra dkk., 2025).

Untuk mengatasi tantangan tersebut, pendekatan *Retrieval-Augmented Generation* (RAG) telah dikembangkan sebagai solusi yang menggabungkan kemampuan generatif LLM dengan mekanisme pengambilan dokumen dari sumber terpercaya (Lewis dkk., 2021). RAG merupakan pendekatan inovatif dalam teknologi kecerdasan buatan yang menggabungkan kekuatan LLM dengan kemampuan pencarian informasi eksternal. Sistem RAG bekerja dengan mengambil informasi yang relevan dari basis pengetahuan atau basis data eksternal sebelum menghasilkan respons, sehingga dapat memberikan jawaban yang lebih akurat dan sesuai dengan konteks (Fanani, 2025). Berbeda dengan LLM tradisional, RAG memungkinkan LLM untuk mengakses dan memanfaatkan sumber informasi tambahan tanpa perlu pelatihan ulang, sekaligus mengurangi masalah halusinasi dan meningkatkan transparansi melalui kemampuannya untuk mengutip sumber informasi Lewis dkk. (2021).

RAG terdiri dari beberapa komponen utama. Pertama, basis pengetahuan atau korpus dokumen yang menjadi sumber informasi eksternal. Kedua, komponen *embedding model* yang mengubah dokumen dan *query* dari pengguna menjadi representasi vektor untuk memfasilitasi pencarian semantik. Ketiga, *retriever* yang berperan mengambil dokumen atau potongan informasi yang paling relevan dengan *query* menggunakan metode *similarity search*. Keempat, komponen *context builder* yang menyusun informasi yang diperoleh menjadi konteks terstruktur. Kelima, *generative model* berupa LLM yang menerima kombinasi *query* pengguna dan konteks hasil *retrieval* untuk menghasilkan respons yang informatif (Fanani, 2025). Pendekatan RAG ini menggabungkan memori parametrik (yang tertanam dalam parameter model) dan non-parametrik (seperti basis data vektor eksternal) untuk mengurangi kecenderungan model dalam menghasilkan informasi yang tidak akurat.

Beberapa penelitian terdahulu telah menunjukkan efektivitas implementasi RAG dalam berbagai domain. Dalam domain edukasi pasien ortopedi, pengembangan *chatbot* RAG menggunakan GPT-4o dan Qdrant berhasil mencapai *answer relevancy* 0.864, *context precision* 0.891, dan *faithfulness* 0.853 pada evaluasi RAGAS (Baur dkk., 2025). Perbandingan performa RAG dengan model Llama3, Gemma, dan Mistral untuk optimasi *chatbot call center* menunjukkan hasil terbaik dengan BERTScore F1 0.90, BLEU 0.18, dan ROUGE-L 0.36 pada kombinasi RAG+Llama3 (Sanjani dkk., 2025). Penerapan RAG dengan Mistral 7B untuk domain hukum lalu lintas Indonesia mencapai BERTScore F1 0.9151 setelah proses *fine-tuning* (Anam dkk., 2024). Implementasi *chatbot* Telegram menggunakan LLM dan LangChain untuk UU Kesehatan memperoleh BERTScore *precision* 76%, *recall* 80%, dan F1 78%, serta ROUGE-1 sebesar 60%, 45%, 50% untuk *precision*, *recall*, dan F1 (Lubis dkk., 2024).

Selain itu, pengembangan sistem tanya jawab kitab fikih berbasis web dengan Llama 3.2:3B dan metode RAG berhasil mencapai BERTScore *precision* 86.50%, *recall* 84.76%, dan F1 85.52% (Ahadi dkk., 2025). Pembangunan *chatbot* asisten kesehatan dengan LLaMA dan pendekatan RAG menggunakan data dari platform Alodokter memperoleh MRR 0.92 serta RAGAS *faithfulness* 0.87, *answer relevancy* 0.83, *context precision* 0.83, dan *context recall* 0.82 (Pratama, 2025). Pendekatan RAG untuk *document question answering* dengan GPT-3.5-turbo mencapai akurasi 96%, BERTScore F1 0.88, ROUGE *precision* 0.74, dan BLEU 0.28 pada *dataset* DocuQA (Muludi dkk., 2024).

Dalam domain kesehatan, implementasi RAG untuk perancangan *chatbot* kesehatan pencernaan menggunakan model Llama 3.1:8b dengan data dari Alodokter mencapai *Mean Reciprocal Rank* (MRR) 93%, *Faithfulness* 62%, *Answer Relevancy* 57%, dan *Semantic Similarity* 81%. Hasil penelitian tersebut menunjukkan bahwa penerapan RAG dapat mengurangi bias model dan meningkatkan kualitas respons *chatbot* dalam memberikan jawaban yang akurat dan relevan sesuai konteks (Samudra dkk., 2025). Sementara itu, implementasi RAG untuk evaluasi proposal tugas akhir mahasiswa dengan memanfaatkan 106 dokumen tugas akhir sebagai korpus mencapai *Answer Correctness* 80% menggunakan model *embedding* jinaai/jina-embeddings-v3, yang berhasil

mengotomatisasi dan meningkatkan objektivitas proses evaluasi di lingkungan akademik (Fanani, 2025).

Meskipun penelitian-penelitian tersebut menunjukkan hasil yang menjanjikan, terdapat beberapa kesenjangan (gap) yang belum terisi dalam literatur. Pertama, sebagian besar implementasi RAG berfokus pada domain publik seperti kesehatan, hukum, dan keagamaan, namun sangat terbatas untuk konteks pengelolaan dokumen internal perusahaan, khususnya di lingkungan BUMN Indonesia. Kedua, belum ada penelitian yang mengintegrasikan WhatsApp sebagai *interface* utama untuk akses dokumen internal perusahaan menggunakan pendekatan RAG. Ketiga, evaluasi yang dilakukan pada penelitian sebelumnya umumnya menggunakan metrik terbatas (hanya 1-3 metrik), belum menggabungkan evaluasi komprehensif yang mencakup BERTScore, RAGAS (*faithfulness*, *answer relevancy*, *context precision*), BLEU, ROUGE-L, validasi pakar, dan *System Usability Scale* (SUS) secara bersamaan dalam satu penelitian.

Penggabungan beberapa teknik evaluasi tersebut diperlukan karena setiap metrik mengukur dimensi kualitas yang berbeda dan saling melengkapi. BERTScore mengukur kemiripan semantik antara jawaban sistem dengan referensi menggunakan *contextual embeddings* (Zhang dkk., 2020), sementara BLEU dan ROUGE-L mengevaluasi ketepatan leksikal pada tingkat n-gram dan urutan kata (Lin, 2004; Papineni dkk., 2001). Metrik RAGAS secara khusus dirancang untuk menilai aspek-aspek kritis sistem RAG, yaitu *faithfulness* (konsistensi jawaban terhadap konteks yang di-retrieve), *answer relevancy* (relevansi jawaban terhadap pertanyaan), dan *context precision* (ketepatan dokumen yang diambil oleh *retriever*) (Es dkk., 2024). Penilaian pakar domain melengkapi metrik otomatis dengan memberikan perspektif praktis yang tidak dapat ditangkap oleh evaluasi algoritmik, seperti kesesuaian jawaban dengan konteks operasional perusahaan. Terakhir, *System Usability Scale* (SUS) memvalidasi apakah sistem benar-benar berguna dan mudah digunakan dari perspektif pengguna akhir (Brooke, 1996). Dengan menggabungkan seluruh metode evaluasi tersebut, penelitian ini dapat menilai kualitas sistem secara utuh; mulai dari seberapa akurat mesin mencari dan menyusun jawaban, ketepatan isinya bagi perusahaan, hingga seberapa mudah aplikasi ini digunakan oleh karyawan.

Kombinasi teknologi RAG dengan WhatsApp API, *vector database* Qdrant, dan *framework* LangChain dapat menciptakan solusi yang efektif untuk manajemen pengetahuan internal perusahaan. LangChain menyediakan *pipeline* pemrosesan data yang kuat yang memanfaatkan *vector database* (dalam penelitian ini Qdrant) untuk melakukan operasi *retrieval* yang efisien dalam VectorDB. Proses ini mencakup pengubahan dokumen menjadi potongan-potongan teks (*chunks*), pembentukan representasi vektor melalui teknik *embedding*, serta pencarian semantik berdasarkan kemiripan vektor (Fanani, 2025). Dengan memanfaatkan arsitektur ini, sistem dapat menghasilkan respons yang ringkas, akurat, dan informatif berdasarkan dokumen internal perusahaan.

Pemilihan LangChain sebagai *framework* orkestrasi didasarkan pada karakteristiknya sebagai *framework* modular yang menyediakan abstraksi berupa *components* dan *chains* serta LangChain Expression Language (LCEL) untuk merangkai alur *retrieval* dan *generation*, disertai dukungan integrasi yang luas terhadap berbagai penyedia model dan layanan pendukung (Topsakal & Akinci, 2023). Di antara *framework* orkestrasi lain yang tersedia, LlamaIndex memiliki orientasi desain pada pengindeksan dan pengambilan dokumen, sedangkan Haystack berorientasi pada *pipeline* pencarian dan tanya jawab kelas enterprise; LangChain dipilih karena keluasannya integrasinya paling sesuai dengan kebutuhan penelitian ini yang menggabungkan banyak komponen heterogen, yaitu WhatsApp (melalui WAHA), backend FastAPI, *vector database* Qdrant, serta LLM Google Gemini melalui Vertex AI. Kesesuaian LangChain untuk sistem RAG berbahasa Indonesia juga telah dibuktikan pada penelitian terdahulu (Lubis dkk., 2024). Perlu dicatat bahwa perbandingan antar *framework* orkestrasi ini bersifat kualitatif pada aspek kapabilitas dan cakupan integrasi, mengingat *overhead* komputasi *framework* relatif kecil dibandingkan latensi pemanggilan LLM sehingga bukan merupakan faktor pembeda utama.

Adapun pemilihan Qdrant sebagai *vector database* mempertimbangkan sejumlah alternatif yang lazim digunakan, yaitu FAISS, Pinecone, Milvus, dan Weaviate (Pan dkk., 2024). FAISS pada dasarnya merupakan *library* pencarian kemiripan vektor yang hanya menyediakan kapabilitas pencarian, bukan sistem basis data yang utuh, sehingga tidak menyediakan manajemen penyimpanan

maupun penyaringan berbasis *metadata* secara bawaan (Pan dkk., 2024). Pinecone merupakan layanan komersial berbasis *cloud* yang diakses melalui REST API (Pan et al., 2024) dan tidak dapat dikelola pada infrastruktur milik perusahaan sendiri, sehingga kurang sesuai untuk dokumen internal perusahaan yang bersifat rahasia, sementara Milvus (Wang dkk., 2021) dan Weaviate merupakan basis data vektor sumber terbuka yang andal namun memiliki kompleksitas operasional yang lebih tinggi (Pan dkk., 2024).

Qdrant dipilih karena bersifat *open source*, dapat dijalankan secara mandiri (*self-hosted*) melalui Docker, dan menggunakan algoritma pengindeksan *Hierarchical Navigable Small World* (HNSW) yang efisien untuk pencarian tetangga terdekat pada data berdimensi tinggi dengan latensi rendah dan *recall* tinggi (Aumüller dkk., 2020; Malkov & Yashunin, 2020). Keunggulan yang paling relevan bagi penelitian ini adalah kemampuan *filterable* HNSW, yaitu penyaringan berbasis *payload (metadata)* yang diterapkan langsung di dalam proses penelusuran graf sehingga hasil *retrieval* dapat dibatasi sesuai divisi pengguna; kemampuan menjawab kueri hibrida (*hybrid query*) yang memadukan pencarian vektor dan atribut ini merupakan salah satu tantangan utama basis data vektor (Pan dkk., 2024) dan secara langsung mendukung kebutuhan kontrol akses dokumen antar divisi pada sistem. Dari sisi kinerja, benchmark yang dipublikasikan Qdrant menunjukkan Qdrant mencapai *requests per second* tertinggi dan latensi terendah dibandingkan Milvus, Weaviate, Elasticsearch, dan Redis pada sebagian besar skenario (Qdrant, 2024a), dan performanya untuk beban kerja *embedding* berskala besar juga telah dievaluasi secara empiris pada lingkungan komputasi berkinerja tinggi (Ockerman dkk., 2025).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan *chatbot* RAG berbasis WhatsApp untuk akses dokumen internal PT Perkebunan Nusantara 1, serta mengevaluasi akurasi sistem menggunakan metrik komprehensif yang mencakup BERTScore, RAGAS, BLEU, ROUGE-L, penilaian pakar, serta mengukur tingkat kebergunaan sistem. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem akses dokumen internal berbasis AI untuk lingkungan BUMN sektor perkebunan di Indonesia, sekaligus menjadi langkah awal untuk mengintegrasikan teknologi AI

dalam ekosistem operasional BUMN sektor perkebunan tanpa menggantikan peran karyawan, melainkan memperkuat kapasitas akses informasi mereka.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan, dapat diidentifikasi beberapa permasalahan sebagai berikut:

1. Karyawan PT Perkebunan Nusantara 1 mengalami kesulitan dalam mengakses informasi dari dokumen internal seperti SOP dan kebijakan perusahaan secara cepat dan akurat.
2. Proses pencarian manual pada dokumen internal membutuhkan waktu yang signifikan dan rentan terhadap kesalahan interpretasi.
3. Belum tersedianya sistem *chatbot* berbasis RAG yang terintegrasi dengan WhatsApp untuk akses dokumen internal perusahaan di PTPN 1.
4. *Large Language Model* (LLM) memiliki keterbatasan berupa kecenderungan menghasilkan informasi yang tidak akurat (*hallucinations*).
5. Penelitian terdahulu belum menerapkan evaluasi komprehensif yang menggabungkan metrik BERTScore, RAGAS, BLEU, ROUGE-L, penilaian pakar dan validasi kebergunaan dari perspektif pengguna dalam satu sistem RAG.

1.3 Pembatasan Penelitian

Untuk memfokuskan penelitian ini, ditetapkan batasan masalah sebagai berikut:

1. Penelitian dilakukan pada dokumen internal PTPN 1 yang mencakup SOP dan kebijakan perusahaan dalam format PDF.
2. Sistem *chatbot* dikembangkan menggunakan arsitektur RAG berbasis *framework* LangChain yang di integrasikan dengan platform WhatsApp.
3. *Interface* pengguna menggunakan WhatsApp sebagai platform utama dan panel admin berbasis web untuk pengelolaan dokumen.
4. Evaluasi akurasi sistem dilakukan menggunakan empat metrik otomatis (BERTScore, RAGAS, BLEU, ROUGE-L) serta penilaian pakar di PTPN 1.
5. Evaluasi kebergunaan sistem dilakukan menggunakan *System Usability Scale* (SUS) terhadap karyawan PTPN 1.

6. Penelitian ini tidak mengukur perbandingan kecepatan pencarian dokumen secara manual dengan pencarian melalui *chatbot*.
7. Sistem melakukan *augmented generation* yang berlandaskan (*grounded*) pada isi dokumen internal yang telah diindeks, bukan sekadar menampilkan hasil pencarian dokumen dan tidak menghasilkan informasi di luar dokumen tersebut.
8. Bahasa yang digunakan dalam sistem adalah Bahasa Indonesia.

1.4. Rumusan Masalah

Berdasarkan identifikasi dan pembatasan masalah yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana menerapkan arsitektur *Retrieval-Augmented Generation* (RAG) berbasis LangChain yang terintegrasi dengan WhatsApp untuk menyediakan akses tanya jawab dokumen internal secara akurat di PTPN 1?
2. Apakah *chatbot* RAG memiliki tingkat akurasi yang memadai dalam menjawab pertanyaan berdasarkan dokumen internal berdasarkan evaluasi metrik BERTScore, RAGAS, BLEU, ROUGE-L, dan penilaian pakar?
3. Apakah *chatbot* RAG berbasis WhatsApp memiliki tingkat kebergunaan (*usability*) yang memadai berdasarkan penilaian karyawan PTPN 1 menggunakan *System Usability Scale* (SUS)?

1.5. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mengimplementasikan *chatbot* berbasis *Retrieval-Augmented Generation* (RAG) yang terintegrasi dengan WhatsApp untuk akses dokumen internal di PTPN 1.
2. Mengevaluasi tingkat akurasi *chatbot* RAG dalam menjawab pertanyaan berdasarkan dokumen internal menggunakan metrik BERTScore, RAGAS (*faithfulness*, *answer relevancy*, *context precision*), BLEU, ROUGE-L, serta penilaian pakar (*expert judgement*) dengan kriteria *accuracy*, *helpfulness*, *ease of use*, *clarity* dan *trust*.
3. Mengukur tingkat kebergunaan (*usability*) *chatbot* RAG berbasis WhatsApp berdasarkan penilaian karyawan PTPN 1 menggunakan *System Usability Scale* (SUS).

1.6. Manfaat Penelitian

Penelitian pengembangan sistem *chatbot* WhatsApp berbasis *Retrieval-Augmented Generation* ini diharapkan dapat memberikan manfaat sebagai berikut:

1.6.1 Manfaat Teoritis

1. Memberikan kontribusi pengetahuan dalam bidang *Artificial Intelligence* dan *Retrieval-Augmented Generation*, khususnya terkait implementasi sistem RAG untuk akses dokumen internal perusahaan berbahasa Indonesia.
2. Memperkaya literatur akademik mengenai integrasi *Large Language Model* dengan *chatbot* pada *platform WhatsApp*.
3. Menyediakan metodologi evaluasi komprehensif yang menggabungkan metrik otomatis (BERTScore, RAGAS, BLEU, ROUGE-L) dengan penilaian pakar domain sebagai acuan bagi penelitian sejenis.

1.6.2 Manfaat Praktis

1. Menyediakan *chatbot* WhatsApp fungsional yang memungkinkan karyawan PT Perkebunan Nusantara 1 mengakses dokumen internal secara akurat.
2. Memberikan *blueprint* implementasi sistem RAG yang dapat diadaptasi untuk unit bisnis lainnya.
3. Membuka peluang pengembangan penelitian lanjutan dalam bidang integrasi LLM dengan platform komunikasi untuk berbagai domain aplikasi.

Intelligentia - Dignitas