

ABSTRAK

Regulasi ketenagakerjaan di Indonesia memiliki peran penting dalam melindungi hak pekerja dan menjaga keseimbangan hubungan kerja antara pengusaha dan tenaga kerja. Namun, kompleksitas aturan, bahasa hukum yang teknis, serta penyebaran informasi dalam berbagai dokumen peraturan seringkali menjadi hambatan dalam pemahaman dan penerapan ketentuan hukum. Permasalahan ini diperparah oleh rendahnya literasi hukum dan disinformasi, sehingga pekerja kerap kesulitan mengetahui hak dan kewajibannya. Pemanfaatan *Large Language Model* (LLM) dalam sistem *Question Answering* (QA) berpotensi meningkatkan akses informasi hukum, namun model ini memiliki keterbatasan berupa fenomena *hallucination*, yakni pemberian jawaban yang meyakinkan tetapi tidak berbasis fakta. Untuk mengatasi permasalahan tersebut, diterapkan pendekatan *Retrieval-Augmented Generation* (RAG) yang mengintegrasikan proses pencarian dokumen relevan dengan kemampuan generatif LLM agar jawaban yang dihasilkan lebih akurat dan kontekstual.

Penelitian ini menggunakan dokumen Undang-Undang Nomor 13 Tahun 2003 tentang Ketenagakerjaan yang telah disesuaikan dengan Undang-Undang Nomor 6 Tahun 2023 Pasal 81 sebagai sumber data utama. Proses pengolahan data meliputi *cleansing*, *case folding*, penyesuaian pasal, serta *chunking*. Setiap potongan teks diubah menjadi vektor menggunakan *text-embedding-3-small* dari OpenAI dan disimpan pada vektor store ChromaDB. Sistem RAG dibangun dengan metode *hybrid retrieval* yang menggabungkan *sparse retrieval* dengan TF-IDF dan *dense retrieval* dengan *cosine similarity* berbasis embedding, lalu dilakukan penggabungan peringkat menggunakan *Reciprocal Rank Fusion* (RRF). Model LLM yang digunakan adalah Qwen3 8B, yang dipilih karena memberikan keseimbangan antara kinerja dan efisiensi komputasi. Evaluasi dilakukan menggunakan *Retrieval-Augmented Generation Assessment* (RAGAs) untuk mengukur *faithfulness*, *answer relevancy*, *context recall*, dan *context precision*.

Berdasarkan hasil pengujian, sistem *Question Answering* (QA) berbasis *Retrieval-Augmented Generation* (RAG) yang dikembangkan menunjukkan peningkatan relevansi dan kesesuaian jawaban yang dihasilkan oleh model *Large Language Model* (LLM) Qwen3 8B. Pengujian terhadap 18 pertanyaan uji yang telah divalidasi oleh pakar hukum menghasilkan tingkat akurasi sebesar 94,44%, di mana 17 jawaban dari sistem berbasis RAG dinyatakan sesuai dengan substansi peraturan perundang-undangan, sedangkan seluruh jawaban dari sistem tanpa pendekatan RAG dinyatakan tidak sesuai. Evaluasi menggunakan metrik *Retrieval-Augmented Generation Assessment* (RAGAs) menunjukkan nilai rata-rata *faithfulness* sebesar 0,926 yang merepresentasikan kesesuaian jawaban terhadap konteks yang diberikan, *answer relevancy* sebesar 0,860 yang mencerminkan relevansi jawaban terhadap pertanyaan, *context precision* sebesar 0,935 yang mengukur ketepatan pemilihan konteks paling relevan, serta *context recall* sebesar 0,949 yang menunjukkan cakupan konteks yang berhasil diambil.

Kata Kunci: *Retrieval Augmented Generation, Large Language Model, Question Answering, Undang-Undang Ketenagakerjaan, Hybrid Retrieval*

ABSTRACT

Labor regulations in Indonesia play a crucial role in protecting workers' rights and maintaining a balanced relationship between employers and employees. However, the complexity of legal provisions, the use of technical legal language, and the dispersion of information across various regulatory documents often hinder the understanding and implementation of labor laws. This issue is further exacerbated by low legal literacy and misinformation, making it difficult for workers to fully understand their rights and obligations. The use of Large Language Models (LLMs) in Question Answering (QA) systems offers the potential to improve access to legal information; however, such models face a limitation known as hallucination, in which convincing but factually inaccurate answers are generated. To address this limitation, the Retrieval-Augmented Generation (RAG) approach is applied, integrating the retrieval of relevant legal documents with the generative capabilities of LLMs to produce more accurate and contextually relevant answers.

This study utilizes the Indonesian Law No. 13 of 2003 on Manpower, as amended by Law No. 6 of 2023 Article 81, as the primary data source. The data processing stage includes cleansing, case folding, article adjustment, and chunking. Each text segment is converted into a vector representation using OpenAI's text-embedding-3-small model and stored in the ChromaDB vector store. The RAG system is implemented using a hybrid retrieval method that combines sparse retrieval with TF-IDF and dense retrieval with embedding-based cosine similarity, followed by rank aggregation using Reciprocal Rank Fusion (RRF). The selected LLM is Qwen3 8B, chosen for its balance between performance and computational efficiency. System evaluation is conducted using the Retrieval-Augmented Generation Assessment (RAGAs) metrics, namely faithfulness, answer relevancy, context recall, and context precision.

Based on the experimental results, the developed RAG-based QA system demonstrates improved relevance and accuracy in the answers generated by the Qwen3 8B LLM. Testing on 18 evaluation questions validated by legal experts achieved an accuracy rate of 94.44%, with 17 answers from the RAG-based system deemed compliant with the substance of labor legislation, while all answers from the non-RAG system were found to be non-compliant. The RAGAs evaluation produced an average faithfulness score of 0.926, indicating answer accuracy relative to the given context; an answer relevancy score of 0.860, reflecting answer relevance to the query; a context precision score of 0.935, measuring the accuracy in selecting the most relevant context; and a context recall score of 0.949, indicating the system's ability to retrieve the necessary context in its entirety. These findings confirm that integrating RAG with LLMs can enhance the reliability of QA systems in the legal domain while mitigating hallucination risks.

Keywords: Retrieval Augmented Generation, Large Language Model, Question Answering, Labor Law, Hybrid Retrieval