

VIETNAM LEGAL QA SYSTEM

Tran Dang Khoa

University of Information Technology - UIT

What ?

We introduce a framework for building a legal question-answering system for Vietnamese law, in which we have:

- Constructed a comprehensive legal dataset that consolidates various Vietnamese legal documents, including laws, decrees, and circulars.
- Developed a hybrid model that integrates Retrieval-Augmented Generation (RAG) and Large Language Models (LLM) to enhance the accuracy and contextual relevance of responses.

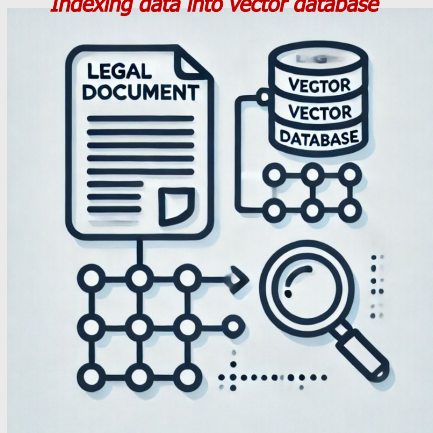
Why ?

Legal documents are one of the most crucial sources of information, as they serve as the foundation for legal interpretation and decision-making. Therefore, efficiently retrieving and understanding legal texts is essential for legal research and practical applications. However, most existing approaches focus on keyword-based search rather than leveraging advanced AI techniques for contextual understanding and precise legal information retrieval.

Overview

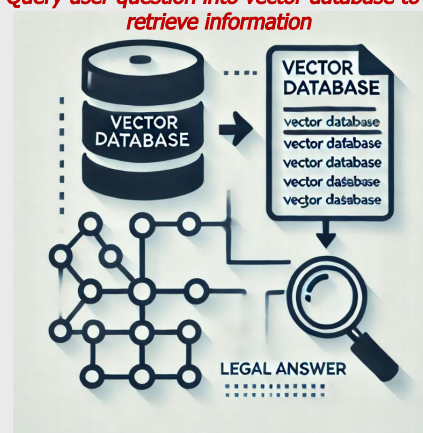
Indexing

Indexing data into vector database



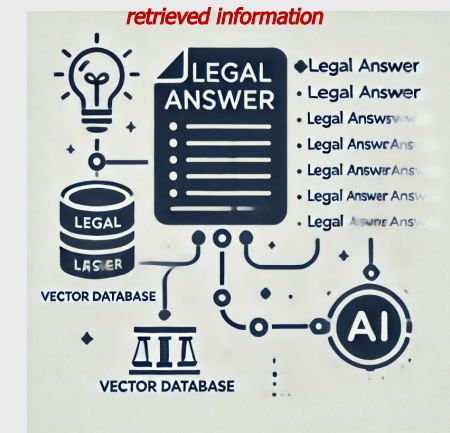
Retrieval

Query user question into vector database to retrieve information



Generation

Generate and select the best answer from retrieved information



Description

1. Indexing

- Indexing legal documents are segmented into smaller passages and indexed in a vector database to facilitate efficient retrieval. This structured segmentation ensures that relevant legal information can be accessed quickly and accurately.
- In addition to legal document indexing, question-answer pairs are stored alongside the text embeddings in the vector database. This enhances the system's ability to retrieve precise responses by leveraging both direct document retrieval and pre-defined legal

3. Generation

- Synthesizes a well-structured legal response based on retrieved legal texts and question-answer pairs.
- Utilizes Large Language Models (LLM) to enhance contextual understanding and generate comprehensive answers.
- Ensures accuracy and legal consistency by integrating retrieved information with AI-driven text generation.

2. Retrieval

- Retrieves relevant legal text passages from the vector database based on the user's query.
- Utilizes embedding-based similarity search to find the most contextually relevant legal documents.
- Ensures efficient and scalable retrieval for large legal datasets.
- Reorders retrieved results to prioritize the most legally relevant and precise answers.
- Applies advanced ranking models to improve the accuracy of legal information retrieval.
- Reduces irrelevant or ambiguous results by refining search outputs.

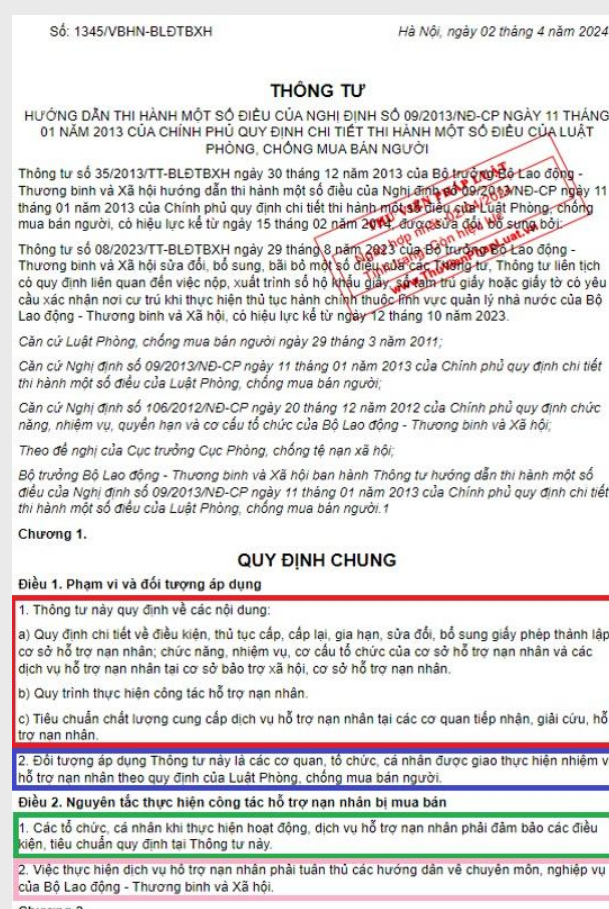


Figure 1. Legal documents are chunked into meaningful sections before being indexed.