MLCQA: A Case Retrieval-Based Question Answering System for Macao Law

Io Cheng Tong, Zhe Li, Yuhong Zhang, Hong Io Chu, Chuxin Ouyang, Ang Li

Faculty of Law, University of Macau

College of Computer Science and Technology, Zhejiang Univerity

Backgrounds

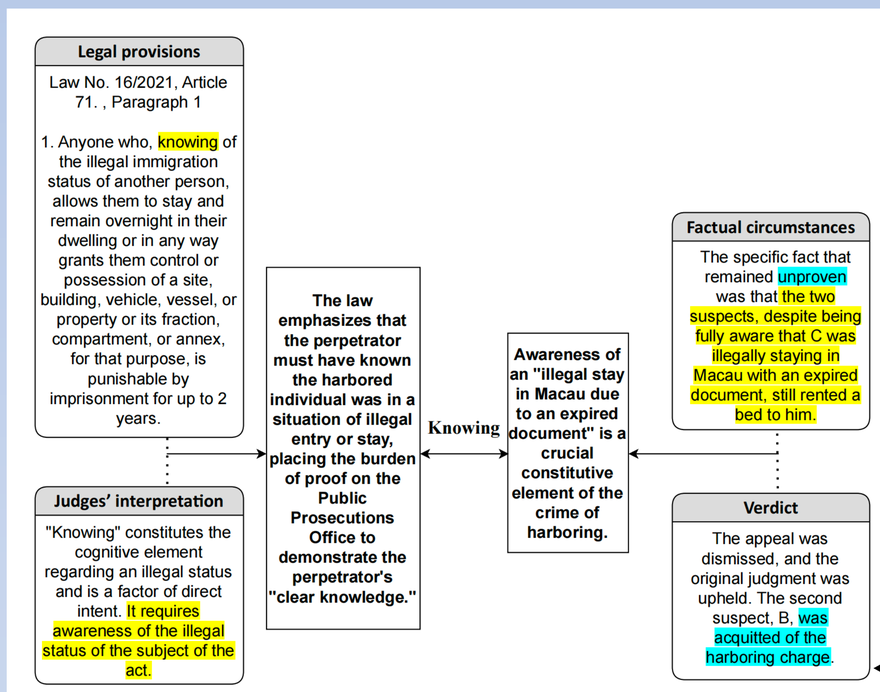

Fig1. Partial field content extracted by the model from Case No.409/2025 (Criminal Appeal) based on prompts.

1. Problem

Large language models (LLMs) continue to struggle with processing court judgments from jurisdictions marked by historical legal pluralism, such as Macao.

2. Challenge

The rigid translated legal terms, absence of unified structure and complex but varied legal reasoning styles, leading to model hallucinations and comprehension difficulties.

3. Findings

Legal Syllogism-guided LLM enabling expert reasoning to reconstruct the legal provision-interpretation-fact-verdict chain.

Methodology

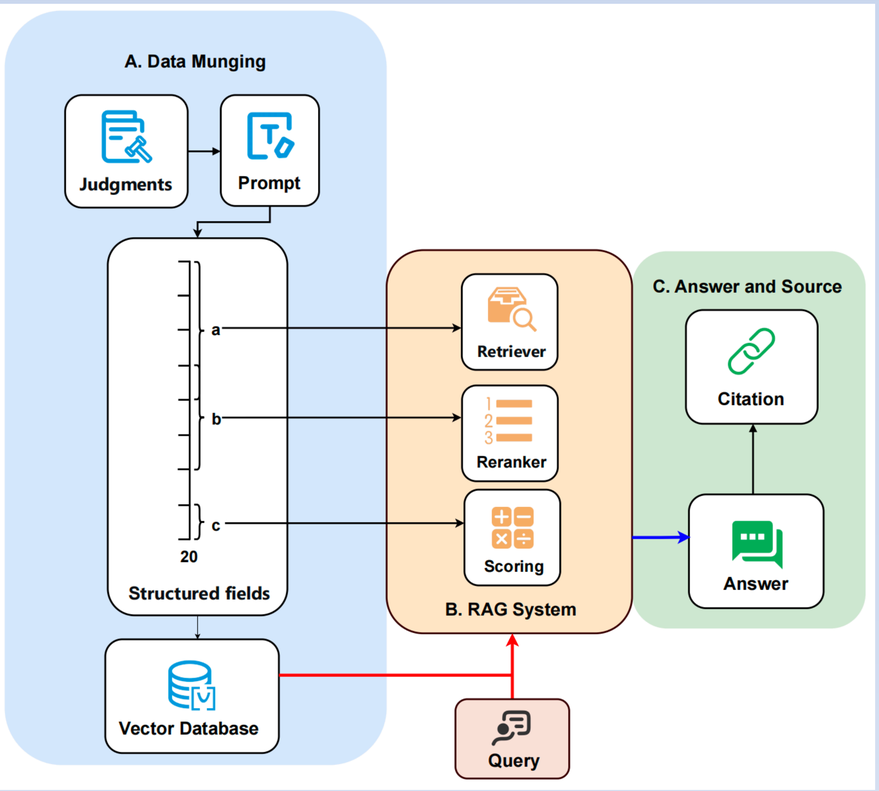

Figure 2: Development process of MLCQA

Expert-Level Legal Information Extraction

Regex-based extraction and LLM-guided extraction.

Ensuring Relevance and Authority in Case Retrieval

Vector retrieval with retrieval fields, cross-encoder reranking with rerank fields, and rule-driven importance scoring.

Interpretable Answer Generation with Traceability

Only exposing the generation fields of reranked cases to the LLM during the answer generation stage, then employ the legal expert LLM to perform case-based legal reasoning and output the answer.

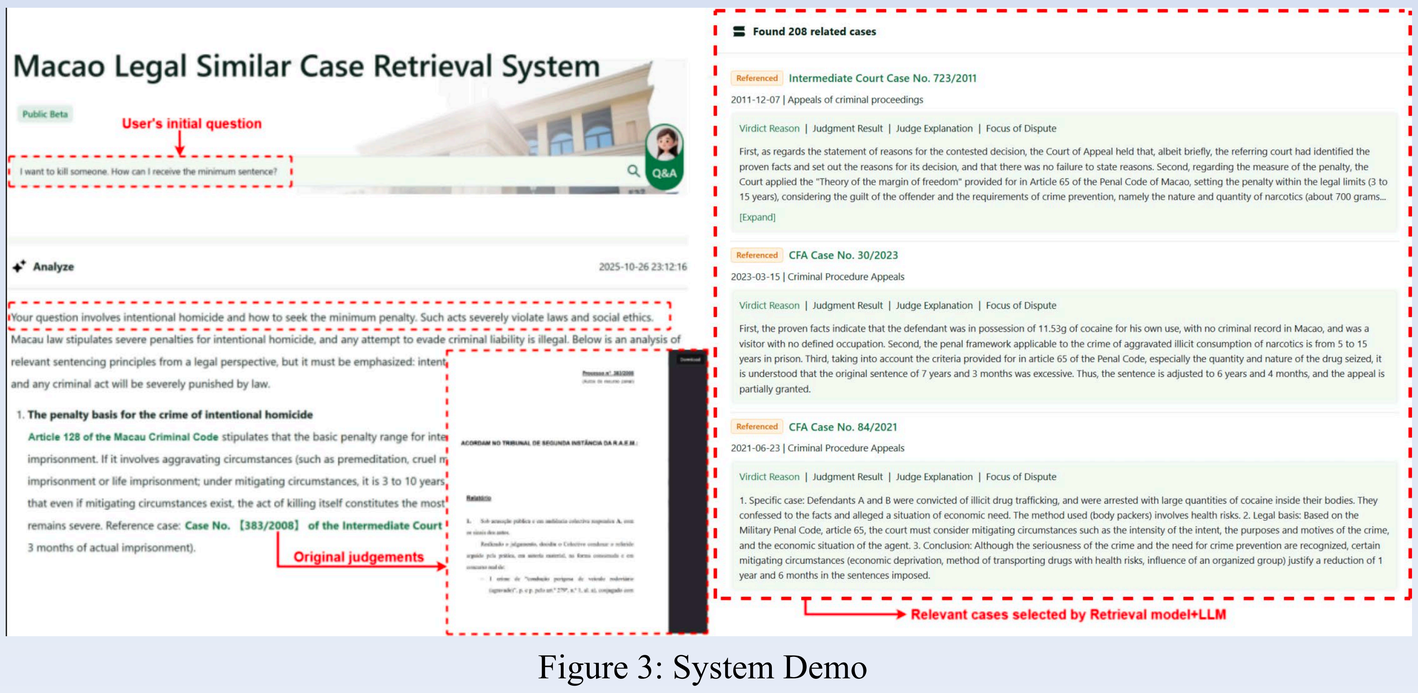

Figure 3: System Demo

1.Legal Understanding

- Interprets real intent from informal queries
- Flags legal/ethical risks & reframes as legal needs

2. Verified Legal Citations

- Answers grounded in laws & precedents
- Clickable links to full source documents

3. Case Recommendations

- Suggests relevant cases for references
- Highlights verdict, reasoning, and dispute focus

The Experiment Results

Model	Accuracy (S^L/S^H)	Term precision (S^L/S^H)	Clarity (S^L/S^H)	Alignment Ratio
MLCQA	3.92 / 3.4	3.97 / 3.9	3.76 / 3.7	94.6%
GPT4	3.80 / 3.1	3.88 / 3.2	3.75 / 3.8	88.1%
Qwen-turbo	3.73 / 3.0	3.73 / 3.1	3.58 / 3.3	86.3%