Retrieving the Right Law: Enhancing Legal Search with Style Translation

Szu-Ju Chen*, Jing Jin*, Sheng-Lun Wei, Chien-Hung Chen, Hsin-Hsi Chen

Challenges in Legal Query-to-Provision Retrieval

Linguistic gap between user query and legal provisions

- Query: widely vary in form, direct and concise phrasing

- Provision: formally structured, technical terminology

Query2Provision Dataset

Include 1,375 legal query-to-provisions pairs

Paired with article level provisions and cover 299 legal codes

Source: Legispedia - Taiwanese online legal QA platform

Legal Query-to-Provision Retrieval System with Style Translation

Enhances legal retrieval model performance by style-specific translation, fine-tuning BiEncoder ranker, and majority voting retrieval with RAG
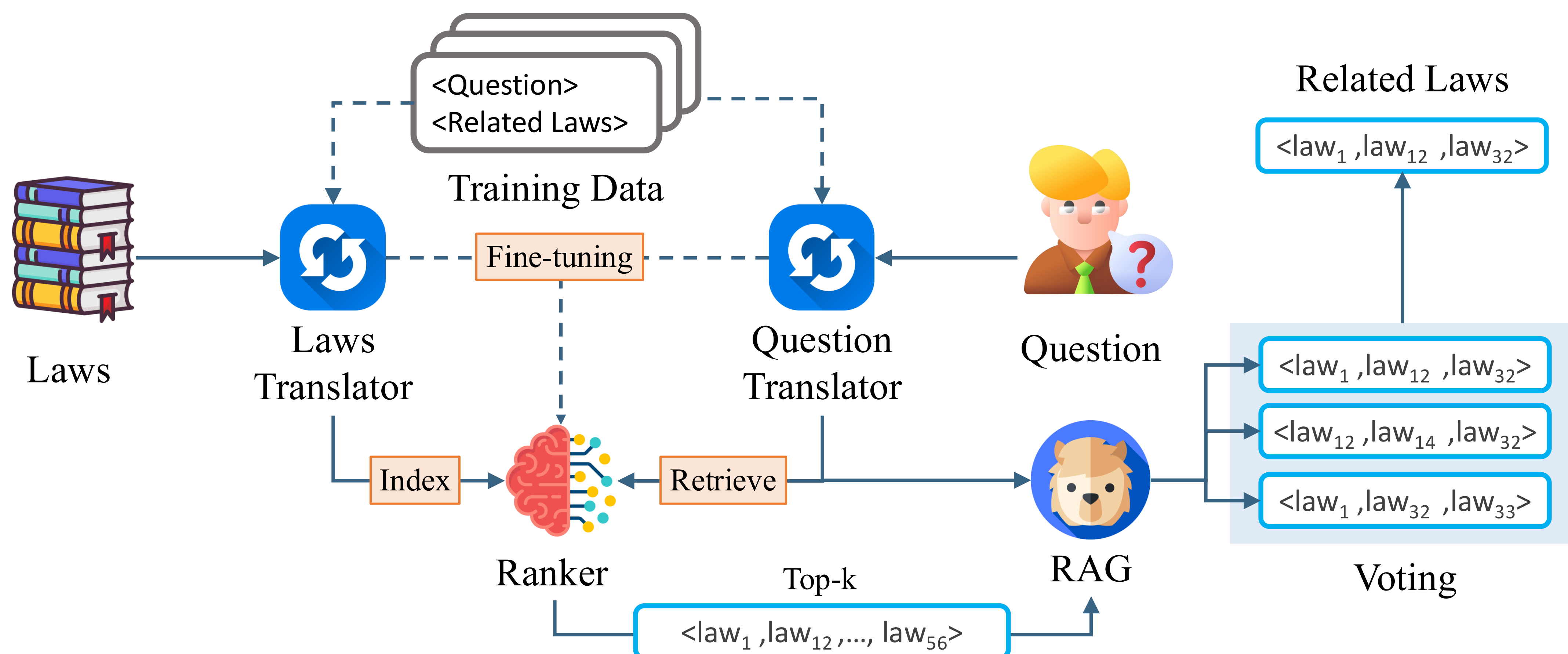
Prompt and Example of Query Translation

You are a writing assistant helping a user improve their legal question before posting online. Rewrite the following in a more professional and formal tone:

<title> Can I sell COVID test kits or oximeters online?

<question> I saw news saying we can't sell these online. Will I get fined if I do?

<translated query> I would like to inquire about the legal implications of selling COVID-19 test kits or oximeters online. Recently, I noticed media reports suggesting that such online sales might be restricted. If I proceed with selling these products online, will I face fines or other legal liabilities? I would appreciate more explicit legal advice to ensure compliance with relevant regulations.

Prompt and Example of of Provision Translation

You are a legal expert and can interpret various legal provisions. Briefly explain this law: **<provision>** Criminal Code, Article 1: An act is punishable only if expressly so provided by the law in force at the time of its commission.

Article 1 of the Criminal Code stipulates that an act may be punished only if there is a clear legal provision applicable at the time the act was committed. In other words, no individual may be penalized for conduct that was not explicitly prohibited by law at that time. This rule also applies to security measures involving restrictions on personal liberty, ensuring legal certainty and stability while protecting individual rights.

Experiment Results

- GPT-4o as translator

- fine-tuned Chinese RoBERTa as ranker

- evaluates on top 15 retrieved candidates

Models	Original			Style Translated		
	Precision	Recall	F1-score	Precision	Recall	F1-score
BM25	0.0605	0.0602	0.0604	0.1169	0.1171	**0.1170**
Ranker	0.1743	0.1758	0.1750	0.1827	0.1843	**0.1835**
LLaMA	0.1378	0.1438	0.1408	0.1552	0.2007	**0.1751**
LLaMA_voting	0.1692	0.1579	0.1633	0.1807	0.1967	**0.1884**
TaiwanLLaMA	0.2158	0.1555	0.1808	0.2470	0.1722	**0.2030**
TaiwanLLaMA_voting	0.2185	0.1943	0.2057	0.2284	0.2099	**0.2188**

- All models with style translation outperforms the original data retrieval

- **94.05%** improvements for the BM25 baseline

- Additional improvements when incorporating majority voting strategy

Discussions

Models	Original	Style Translated
BM25	0.0604	**0.1053**
Ranker	0.1750	**0.1988**
LLaMA	0.1408	**0.1590**
LLaMA_voting	0.1633	**0.1790**
TaiwanLLaMA	0.1808	**0.1837**
TaiwanLLaMA_voting	**0.2057**	0.1963

▲ Performance using Gemini-1.5-Flash as style translator

Query / Provision	Original	Translated
Original	0.20569	0.21248
Translated	0.20862	**0.21876**

▲ Performance of difference translation setting of query and provision

Models	Original	Style Translated
LLaMA	0.1601	**0.1618**
TaiwanLLaMA	0.1891	**0.2211**

▲ Performance of direct fine-tuning LLM

- Style translator with other LLMs is also effective

- Most gains when both query & provision are translated

- Style-translated data also show effectiveness in direct fine-tuning LLMs under zero-shot prompt

Conclusion and Future Work

1. Bridging linguistic gap between user queries and legal provisions is crucial in Legal Query-to-Provision Retrieval

2. Style-translated data consistently improves performance across various models and retrieval methods

3. Translating both queries and legal texts yields the best alignment and retrieval results

4. Multiple LLMs can serve as effective style translators

5. Future work includes exploring diverse prompts and expanding available datasets for style translation