INFORMATION EXTRACTION FROM THE LEGAL DOCUMENT SYSTEM OF VIETNAM AND APPLICATION OF CIVIL LAW QUERY

Dung V.Dang

University of Information Technology, Vietnam National University - HCMC

What ?

We introduce a method for extracting information from the legal text corpus in Vietnam, in which we:

- Propose an effective approach for extracting and representing knowledge.
- Provide a knowledge-based query solution.
- Build a legal query system and conduct experiments with civil law.

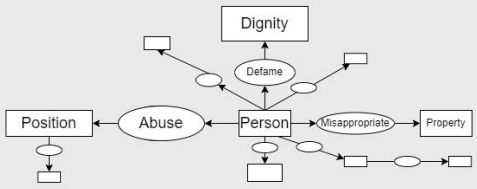
Why ?

- Legal query is a crucial element in accessing and understanding legal knowledge. It assists individuals in searching for and grasping legal regulations, while ensuring compliance with and accurate application of the law.
- Understanding legal regulations is not an easy task. Current legal information retrieval systems fall short in meeting users' search needs.

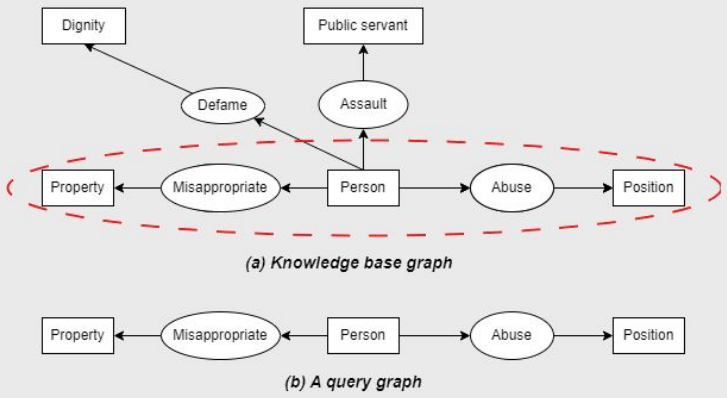
Overview

Building a knowledge graph → Legal Searching

(a) Legal document



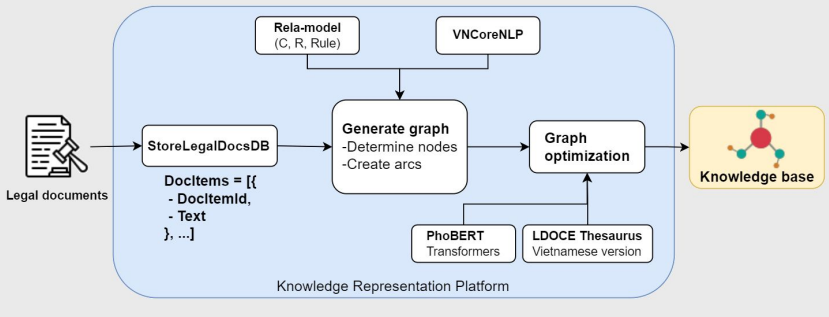
(b) Knowledge Graph



Description

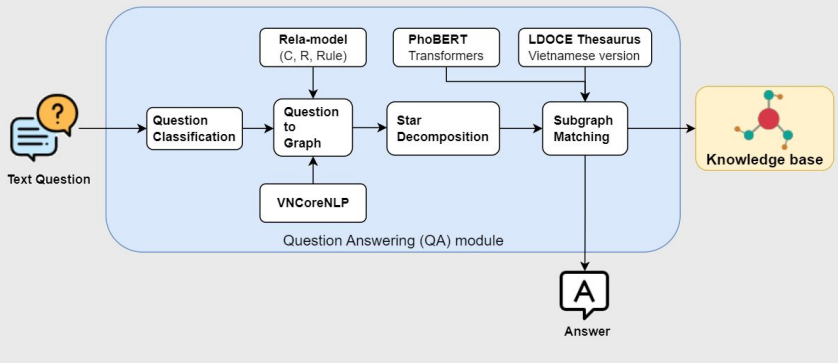
1. Building a knowledge graph

- The process of representing textual knowledge into a knowledge graph is achieved through the utilization of the Lego-Onto model in combination with natural language processing (NLP) techniques to extract and represent information as follows:
- Store legal documents in the database.
- Determine nodes of the knowledge graph by extracted key phrases.
- Create arcs of the graph through relations between keyphrases.
- Graph optimization.



2. Legal Searching

- We utilize a synonym dictionary combined with PhoBERT to perform matching and provide answers to query questions as follows:
- Categorize questions
- Represent the knowledge of the question in a knowledge graph
- Decompose the question graph
- Find knowledge that matches the star graph
- Take the intersection set of the answer sets of the star graphs



NII

Dung V.Dang – University of Information Technology

TEL : 0949930259 Email : dungdv@uit.edu.vn