Ontology-based Solution for Building an Intelligent Searching System on Traffic Law Documents

Vuong T. Pham, Hien D. Nguyen1,3[[1]](#footnote-2),

2Faculty of Software Engineering, University of Information Technology, Ho Chi Minh City, Vietnam

1Faculty of Computer Science, University of Information Technology, Ho Chi Minh City, Vietnam

3*Vietnam National University, Ho Chi Minh City, Vietnam*

[hiennd@uit.edu.vn](mailto:hiennd@uit.edu.vn),

Keywords: knowledge base, searching system, traffic law, law on road traffic, legal document

Abstract: In this paper, an ontology-based approach is used to organize the knowledge base of legal documents. This knowledge model is built by the improvement of ontology Rela-model. Besides, some searching problems on the designed legal knowledge base are proposed. The intelligent searching system on the Vietnam road traffic law is constructed by applying the method. The searching system can help users to find concepts and definitions in the road traffic law. Moreover, it also can determine the penalties and fines for violation in the traffic. The experiment results show that the system is effective on usual searching of users and emerging to apply in the real-world.

# INTRODUCTION

Nowadays, transportation is a need for everyone. Almost every adults have their own vehicles. The traffic more and more complicated, especially the road traffic. In Vietnam, there are more than 3 million cases of violation traffic law with more than 14,500 cases of traffic accidents in 2020 (National Traffic Safety Committee, 2020), in which, some cases have the results of injuries or deaths. The reason of those cases mostly is people have low awareness about rules of traffic law.

Ontology is an effective approach for representing the knowledge (Jakus et al., 2013). This model has been used for organizing the knowledge in education, healthcare (Do et al., 2018). There are some studies to apply ontology for representing the knowledge of legal document. Besides, there are some studies to use ontology for organizing the legal knowledge (Valente and Breuker, 1992, Fawei et al., 2019). However, they did not mention to the traffic law for searching its content and determine penalties for violations.

In this paper, a method for building the knowledge base for Vietnam road traffic law in (Vietnam National Assembly, 2008, Vietnam Government 2019) is proposed. This method is applied to construct a searching system in this law. The designed system supports users to find the content of law related to their queries, and it can determine penalties for violations in the road traffic via this law. The system helps to raise awareness of people about traffic law.

The main value of the designed system is the abilities of searching on penalties and fines for road traffic offences based on the keywords of the inputted query. The knowledge base of the system is organized as a relational ontology, which includes concepts, entities, their relations and the rules of Vietnam Law on road traffic. In order to do that, the knowledge domain about road traffic law are collected and classified into knowledge components: concepts, relations and rules.

The next section presents some related works for constructing the relational ontology, especially in the law domain, and some searching systems on legal documents. Section 3 proposed an improved model of Rela-model to represent the knowledge of road traffic code in Vietnam. Section 4 builds an architecture and searching problems of an intelligent querying system on Vietnam traffic code. The designed system can support to find the content of law related to the query and penalties for road traffic offences. The last section concludes the results of this paper and give some future works.

# RELATED WORK

There are many studies to organize the legal knowledge. Valente and Breuker (1992) stated three approaches for the legal knowledge base: the logic approach, the case-based approach and the pragmatic approach. Those approaches are used to build legal ontologies and legal documents data retrieving system (Sator et al, 2011).

Ontology LIDO for Legal Informatics Document is built based on the standard CEN Metalex (Sartor, 2019). It represents legal actions affecting the document, the legal temporal events, the structure of the legal resource, and the semantic structure of the legal document organization.

Ngo et al. (2021) proposed a method of data augmentation based on legal domain knowledge for the legal textual entailment. This method is used to design a system for Vietnamese legal text processing. However, the proposed method only shows the articles of a specified query, it does not use the legal knowledge to explain its results clearly.

Pham et al. (2019) built an ontology-L for representing the Law Public Investment and design a consultant system for estimating costs of a project based this law. Besides, an intelligent chatbot was designed to tutor some administrative procedures in printing licensing based on the ontology Rela-Ops model (Nguyen et al., 2020a). However, those methods are difficult to apply in searching the content of a law document related to the working domain.

In Vietnam, there are some legal searching system, such as National database of Legal documents (2022) of Ministry of Justice, and law library (2022). In general, those systems only allow users to search for documents or entities with keywords. However, they cannot help users to catch a deeper searching in the legal document in the real-world. For example, in the traffic law, users need to search for penalties and fines for violation based on rules in the legal document. Thus, the current systems are not suitable to support users in the practice.

This study tends to build an intelligent searching system on the Vietnam road traffic code based on ontology approach. This ontology can be used to represent the content of the law code and deduce based on the inference rules extracted from the code.

# Knowledge base of Vietnam traffic law

## The structure of the Vietnamese law on the road traffic

This section gives more details about the structure of Vietnamese law on road traffic and the knowledge model of the system. Through Vietnam National Assembly (2015), the system of legal documents in Vietnam has the followed levels:

1. The highest validity is Constitution;

2. Codes/Laws and resolutions of National Assembly;

3. Sub-law documents for instructing the detail of the law established by National Assembly.

In general, a law document has the structure: the heading, the content and the ending. The heading shows the national name, the crest, number and sign of document, enact place and date, type and name of the document, basis of document. The content is list of parts, chapters, articles, clauses and points. The ending is the signing of the person that implement the document.

Inside the content, parts is the highest level, then, in order is chapters, sections, articles, clauses and points. Through (Vietnam Ministry of Justice, 2011), based on the type of document, there will be different structures, for examples some documents have chapters, articles, clauses, points but not sections. Each part or section or chapter defines a different factor. Below chapter is articles and clauses to define concepts, principles, penalties or regulations. If a clause needs more than a sentence to define it, there will be points in addition to it.

For the concept in the legal document, it has the name of the concept and the sentence to define it. For the offences, each principle, penalty or regulation which are defined in articles and clauses of the legal document, they always have the subject (the person or organization that participate or engage in the event), a fact (or action) and penalties if there is any.

In particular, the Vietnamese traffic law has the same structure as stated. Two legal documents currently implement and have most effect in the social are: Law on road traffic (Vietnam National Assembly, 2008) which prescribes interpretation of concepts, road traffic rules, regulations for vehicles and users on the road traffic; Decree of Administrative of penalties for road traffic offences and rail transport offences (Vietnam Government, 2019) (known as Decree 100) which states penalties and fines for administrative violations of road traffic. Besides, there is National Technical Regulation on Traffic Signs and Signals (Vietnam Ministry of Transport, 2019) to define and describe the road traffic signs.

## The knowledge model for the traffic law

Ontology Rela-model is a useful ontology representing the knowledge of relations. This model includes three components about concepts, relations between concepts (Nguyen et al., 2015). It is effective to represent the knowledge domains in education, consultant the finace method based on the investment law.

For representing the knowledge of a legal document, Rela-model has been improved the structure of its concept-component being suitable the legal domain. The knowledge model for the Vietnamese road traffic law based on the concepts or entities and their relation. Each relation of them define an action or event of road traffic. Base on those relations and rules of law on road traffic, the issues about retrieving the information of offences and its penalties have been also proposed.

**Definition 2.1:** The knowledge model for representing the legal domain of road traffic is improved from ontology Rela-model. This model consists of three components as follows:

**(C, R, Rules)**

In which, **C** is the set of concepts or entities of road traffic law, **R** is the set of relations between conceps/facts, **Rules** represent the inference rules to specify the relation between concepts or determine offences and their penalties.

Set **C** is the set of concepts and entities in the road traffic law. There are three kinds of concepts in C: users and vehicles of road traffic; traffic signs and signals; road infrastructure. Based on those kinds, each concept *c* ∈ **C** has the structure:

(*Name, Meaning, Attrs, Keywords*)

where, each element has the type and meaning for specifying the corresponding concept as Table 1:

Table 1. Structure of a concept

|  |  |  |
| --- | --- | --- |
| **Element** | **Type** | **Meaning** |
| *Name* | Text | Name of the concepts |
| *Meaning* | Text | Meaning of the concepts. |
| *Attrs* | List | List of attributes of the concepts. |
| *Keywords* | Set | Set of keywords determined or related to the concepts. |

Example 1: The concepts *“Electric motorcycle”* in (Vietnam Government, 2019) is described as follows:

|  |  |
| --- | --- |
| **Element** | **Content** |
| *Name* | *Electric motorcycle* |
| *Meaning* | “a two-wheel vehicle operated by an electric engine with power not exceeding 4 kW and maximum speed not exceeding 50 km/h” |
| *Attrs* | *Attrs* = [*kind, type, legal*]  • *kind*: road traffic vehicle  • *type*: two-wheel vehicle  • *legal*: [Article 3, Clause 1, Point d, Decree No. 100/2019/ND-CP] |
| *Keywords* | Motorcycle; electric; two-wheel vehicle |

Set **R** is the set of relations between concepts in set C. These relations determine a specific fact or an action of the road traffic code. Each relation *r* ∈ **R** has the structure:

(*Name, Conc, Meaning, Prop, Keywords*)

where, each element has the type and meaning for specifying the corresponding relation as Table 2:

Table 2. Structure of a relation

|  |  |  |
| --- | --- | --- |
| **Element** | **Type** | **Meaning** |
| *Name* | Text | Name of the relation. |
| *Conc* | List | List of parameters as concepts of the relation. |
| *Meaning* | Text | Meaning of the relation. |
| *Prop* | Set | Set of propertise of the relation. This study mention two main properties on a binary relation: transitive, symmetric. |
| *Keywords* | Set | Keywords of the relation. |

Example 2: The relation “*comply*” of two concepts “*car*” (or “*car-like vehicles*”) and “*traffic light*”, denoted *simply*(*car*, *traffic light*), means “Operators of car and car-like vehicles failed to comply with the traffic lights”.

The keywords of its are “*comply*”, “*over*”

Set **Rules** is the set of inference rules. Those rules deduce relations between concepts or determine offences based on the road traffic law. Each rule *r* ∈ **Rules** has the form

*u*(*r*) → *v*(*r*)

where, *u*(*r*) is the hypothesis facts of rule *r* and *v(r*) is the result of rule *r*.

The **Rules**-set is classified two kind of rules:

**Rules = Rulinfer** ∪ **Ruloffence**

In which, **Rulinffer** is the set of rules inferring the relation between concepts, and **Ruloffence** is the set of rules determining offences and penalties.

## Some problems for searching on traffic law

Using the improved Rela-model, the knowledge base for the road traffic law hase been organized. Based on this knowledge base, the problems for searching on the law document are studied. There two issues for searching on law, which are searching the concepts or definition of the law, especially the law explanation, and determining offences and their penalties & fines through the law document. To do this, two searching problems need to be solved for designing the intelligent searching system on the law document:

**Definition 2:** The searching problems of an intelligent searching system based on the knowledge base as an improved Rela-model:

* **Problem 1:** Extracting the keywords from the inputted query to search the concepts and relations in the legal knowledge base related to the keywords.
* **Problem 2:** Retrieve the knowledge from the knowledge base matching extracted concepts and relations.

For solving problem 1, the inputted query needs to be classified. The input can be classified into two kinds: query about meaning of a concept (“*what is*?”) and query about the penalties & fines of an offence (“*how much*”, “*penalty*”, “*fines*”). After that, from the kind of the query, its main keywords are extracted. Besides, some similarity words for extracted keywords are also achieved. The similarity keywords can be collected from legal document sources, experts (as lawyers, legal lecturers), or from dictionaries. With extracted keywords and determined similarity words, concepts related to those keywords are determined by using rules in **Rulinffer**. The process also finds inference rules used to deduce concepts and their relations.

For solving problem 2, after identifying the concepts and relations, the article of legal documents that states the offence is found by using rules in **Ruloffence**. Then, the information, penalties and fines of it are retrieved through the specified content of law in the knowledge base. The process for solving this problem is as follows:

Given the knowledge base K of road traffic law in legal documents () as Rela-model. This algorithm will determine the information, penalties or fines of an inputted query *q*.

**Input:** The knowledge base K = (**C, R, Rules**) as Rela-model.

The inputted query *q*.

**Output:** information, penalties and fines of road traffic offence.

The process for solving:

**Step 1:** Extract keywords from the query *q* and find similarly words based on the knowledge base K.

**Step 2:** Determine related concepts and their relations Base on the keywords and inference rules in **Rules**.

**Step 3:**

• Using rules in **Ruloffence** to find the offence in the knowledge base K.

• Retrieve the information, penalties and fines of the determined offence from the specified law document.

• Update the results of the query *q*.

**Step 4:** Return the results of the query *q*.

# The Searching System OF vietnamese LAW ON ROAD TRAFFIC

## Requirements of a searching system on legal documents

The intelligent searching system on legal documents needs to be supported the understanding of users about the legal domain. In the road traffic law, besides the ability for solving of necessary issues of the searching system, this system has some criteria of intelligent software evaluation in searching (Nguyen et al., 2020b, Giakoumakis and Xylomenos, 1996):

* + *Portability:* This is the level of difficulty to work with the same project with different computers.
  + *Installation:* the requirements of software and hardware for the simulator, and how straightforward is the installation in a supported system.
  + *Usability:* this criterion shows the content being suitable and detail with the current law domain. It is updated and easily to use in the practice.
  + *Understandability:* this is one of the most important characteristics of the intelligent law searching software quality. This system has to help users understanding the law content in legal documents. It can influence the users’ feeling about software and reliability of software evolution in reuse or maintenance.

## The architecture of the searching system on the traffic law

The architecture of searching system on traffic law is presented in Figure 1. The system consists of the user interface, the knowledge base and the searching engine.

The knowledge of the road traffic codes is collected from (Vietnam Government, 2019, Vietnam National Assembly, 2008). These facts and entities of those documents are organized as a knowledge base by the improved ontology Rela-model. The similarity words are manual established via the collecting of intellectual experts and their experiences.

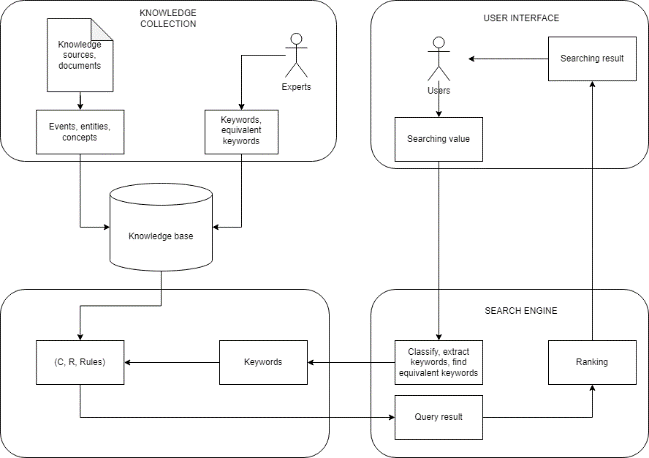


Figure 1. The architecture of an intelligent searching system on the Vietnam road traffic.

When a user inputs the query, the search engine will execute the extract keywords tasks by the problem 1, which are classifying the words, checking typo, removing stop words, checking synonyms, checking equivalent keywords, to generate the query values. From the extracted keywords, the similarity words will be determined through the knowledge base as the improved ontology Rela-model form. Those are used to search the necessary knowledge by using inference rules of the knowledge. Besides, their penalties and fines are also retrieved by the problem 2. The result will be ranked by the search engine before showing in the user interface.

## Testing Results

Based on the knowledge base has been organized in section 3 and the proposed architecture in section 4.2, an intelligent searching system on Vietnam road traffic law is designed. This section presents some testing results of the system through some kinds of the inputted queries.

Example 3: The inputted query *q*1 = “*What is motorcycle*?”

The system will extract keywords from the query *q*1: “*What is*”, “*motorcycle*”. From that, it returns the results as follows:

“*Motorcycle means a motor vehicle that has two or three wheels with a cylinder capacity of 50 cm3 or higher, maximum speed over 50 km/h and net weight not exceeding 400 kg.*”

The word “*what* *is*” used to classify the query into the kind of declaring the meaning of a concept. The keyword “*motorcycle*” helps to find the concept. The result is retrieved from Article 3, Clause 3.31 of National Technical Regulation on Traffic Signs and Signals (Vietnam Ministry of Transport, 2019).

Example 4: The inputted query *q*2 = “*The fines of operator of motorbike does not wear helmet*”

The keywords of the query *q*2 are *“fines”, “not wear”, helmet”, “operator of motorbike”*. The word “*fines*” used to classify the query into stating penalties and fines of offences. The word “*operator of motorbike*” consists of “*motorbike*” that is similarly to the word of “*motorcycle*”. The word “helmet” is in the keywords of the concept “motorcycle helmet”. Hence, the concepts of the query *q*2 are “*operator of motorcycle*” and “*motorcycle helmet*”. The relational keyword is “*not wear*”. With the concepts and relation, the rules were used to match them and find the result.

The result is returned:

“*Through article 6, Decree 100/2019/ND-CP: Penalties imposed upon operators of mopeds and motorcycles (including electric motorcycles) and the like violating road traffic rules.*

*2. A fine ranging from VND 200,000 to VND 300,000 shall be imposed upon a vehicle operator who commits any of the following violations:*

*i) The operator or the passenger on the vehicle does not wear a motorcycle helmet or does not wear it properly;*”

The designed system can do some common searching on the road traffic law. It is effective with finding usual penalties and fines in the road. This system was tested on 95 queries about the road traffic codes. The results are shown in Table 3:

Table 3. Results for testing of queries

|  |  |  |  |
| --- | --- | --- | --- |
| **Kind** | **Quantity** | **Correct** | **Rate** |
| Queries about concepts / definitions | 41 | 34 | 83% |
| Queries about penalties and fines | 55 | 42 | 76% |
| **Total** | **96** | **76** | **79%** |

# CONCLUSION AND FUTURE WORK

This paper proposed an ontology-based model for representing the legal knowledge in the Vietnam road traffic codes. This model is improved based on ontology Rela-model in the structure of concepts, relations and inference rules. Through the designed knowledge base, some searching issues on the Vietnam road traffic codes are proposed: extracting keywords and inferring the matched result for inputted query. Moreover, the architecture of an intelligent searching system on road traffic law has been constructed. This system can do some common searching, such as find concepts/definition in the law and determine penalties for violations in the road traffic. At the moment, most of knowledge collected by manual collection method. The next work is the improvement of the collection method within by using an automatically method.

In the future, the system can be involved in other legal aspects such as commercial law, civil law, etc. Further, the system can be used to provide an e-learning system for legal aspects. The abilities of using AI to identify entities and concepts from an image or using voice recognition to identify searching input are also the features considered to add more in the system.

**References**

Do, N., Nguyen, H., Selamat, A. 2018. Knowledge-Based model of Expert Systems using Rela-model. *International Journal of Software Engineering and Knowledge Engineering* 28(8), 1047 – 1090.

Fawei, B., Pan, J.Z, Kollingbaum, M., Wyner, A.Z. 2019. A Semi-automated Ontology Construction for Legal Question Answering. *New Generation Computing* 37(4), 453-478, 2019.

Giakoumakis, E.A., Xylomenos, G. 1996. Evaluation and Selection criteria for software requirements specification standards. *Software Engineer Journal*, 11(5), 307-319.

Jakus, G., et al. 2013. *Concepts, Ontologies, and Knowledge Representation*. Springer Nature.

Law library. (2022). <https://thuvienphapluat.vn/>

National database of Legal documents . 2022. <http://vbpl.vn/botuphap/Pages/Home.aspx>

National Traffic Safety Committee. 2020. *Final report the traffic accidents in 2020*. Vietnam Government.

Ngo, H., Nguyen, T., Nguyen, D., Pham, M. 2021. AimeLaw at ALQAC 2021: Enriching Neural Network Models with Legal-Domain Knowledge. In *KSE 2021, 13th International Conference on Knowledge and Systems Engineering*, Nov. 2021. IEEE.

Nguyen, H.D., et al. 2015. A Mathematical Approach for Representation Knowledge about Relations and Its Application. In *KSE 2015, 7th IEEE International Conference on Knowledge and Systems Engineering*, Oct. 2015. IEEE.

Nguyen, H., Tran, D., Pham, H., Pham, V. 2020a. Design an intelligent system to automatically tutor the method for solving problems. *International Journal of Integrated Engineering* 12(7), 211 – 223.

Nguyen, H., Do, N., Tran, N., Pham, H., Pham, V. 2020b. Some criteria of the Knowledge Representation method for an Intelligent Problem Solver in STEM education. *Applied Computational Intelligence and Soft Computing* 2020, Article ID 9834218.

Pham. H., Do, N., Nguyen, H. 2019. A Consulting System for Estimating Costs of an Information Technology Hardware Project based on Law of Public Investment. In *ICSSE 2019, 7th IEEE International Conference on System Science and Engineering*, July 2019. IEEE.

Valente, A., Breuker, J. 1992. A Model-Based Approach to Legal Knowledge Engineering. In: L*egal Knowledge Based Systems: Information Technology & Law*, Grütters, et al. (eds), JURIX'92, Koninklijke Vermande, Lelystad, NL, 1992.

Sartor, G., Casanovas, P., Biasiotti, M., Fernandéz-Barrera, M. 2011. *Approaches to Legal Ontologies*. Springer Dordrecht Heidelberg, London, New York.

Vietnam National Assembly. 2008. *Law on Road Traffic*. Law No. 23/2008/QH12.

Vietnam National Assembly. 2015. *Law on Promulgation of Legislative Documents*.No. 80/2015/QH13.

Vietnam Government. 2019. *Decree on Administrative penalties for road traffic and rail transport offences*. No. 100/2019/ND-CP.

Vietnam Ministry of Justice. 2011. *Circular of the Minister about the formats, and techniques for legal documents of the Government, the Prime Minister, Ministers, and joint legal documents.* No. 25/2011/TT-BTP.

Vietnam Ministry of Transport. 2019. *National Technical Regulation on Traffic Signs and Signals*. QCVN 41:2019/BGTVT, 2019.

1.  https://orcid.org/0000-0002-8527-0602 [↑](#footnote-ref-2)