Towards Intelligent Legal Consultancy based on Integrating of Ontology and Knowledge Graph and Compare with LLMs

Vuong T. Pham[1,2,3], Thanh Cao[4⋆], Xuan-Hau Pham[5], Hung Q. Ngo[6], and Hien D. Nguyen[7,2]

[1] Faculty of Mathematics and Computer Science, University of Science, Ho Chi Minh city, Vietnam
[2] Vietnam National University, Ho Chi Minh city, Vietnam
[3] Institue of Data Science and Artificial Intelligence, Sai Gon University, Ho Chi Minh city, Vietnam
vuong.pham@sgu.edu.vn
[4] Faculty of Information Technology, Sai Gon University, Ho Chi Minh city, Vietnam
thanh.cao@sgu.edu.vn
[5] Quang Binh University, Dong Hoi city, Quang Binh, Vietnam
pxhauqbu@gmail.com
[6] Technological University Dublin, Dublin, Ireland
hung.ngo@tudublin.ie
[7] University of Information Technology, Ho Chi Minh city, Vietnam
hiennd@uit.edu.vn

Abstract. Nowadays the necessity for employing an AI system to support the search for legal documents is increasingly evident. In this paper, a method for integrating of ontology Legal-Onto and knowledge graph, called Legal Onto-Graph model, is studied. This integration, built upon laws and legal documents, facilitates various applications, such as question answering, document similarity, and search functionality. The proposed method is useful for representing knowledge spanning multiple documents within the legal domain. Organized as a knowledge base, the integrated model powers a Questions-Answering system tailored for legal documents. The experimental results conducted on Vietnamese Land Law demonstrate the system's effectiveness in handling common queries, outperforming other large language models in this domain based on the criteria of an intelligent legal consultancy system.

Keywords: Knowledge-based Systems · Legal AI · Legal Linked Data · Legal Documents · Knowledge Graph.

1 Introduction

In the forth industrial revolution, the using of intelligent systems is necessary in many fields, especially in searching the knowledge of legal documents [1].

⋆ Corresponding author: Thanh Cao (thanh.cao@sgu.edu.vn)

Artificial Intelligence (AI) in Law has over 30 years for studying and developing [12, 26, 30]. There are many methods to build AI system for supporting of legal research [6, 4].

In Vietnam, the virtual assistant for the people's courtis is a Question-Answering (QA) system for judges in courts [29]. This system is trained on 1.3 million judgments, 170,000 legal documents, helps reduce the workload for judges by 30%. However, it was only used by experts and not to be used in residence.

The research in [16] compares traditional reviewers like junior lawyers and outsourcing services to cutting-edge Large Language Models (LLMs) for legal contract review. The results are striking: LLMs can complete reviews in seconds, significantly faster than the hours it takes humans. However, this study did not propose a method for efficiently organizing legal documents as a knowledge base for LLMs. This limitation necessitates frequent retraining whenever the underlying legal landscape changes, hindering the system's long-term viability.

In the real-word, a legal domain has many related documents with their relationships [3]. However, the methods for organizing the intellectual collected from various legal documents are limitations. They cannot represent their content completely and relations between related documents.

This paper explores the integration of the Legal-Onto ontology with a knowledge graph, called *Legal Onto-Graph model*. This integration, which draws on laws and legal texts, enables diverse applications like question answering, document comparison, and search capabilities. The suggested approach proves to be valuable for capturing knowledge across multiple documents in the legal field. The combined model plays as a knowledge repository to drive a legal document-focused Question-Answering system. Experimentation with the Vietnamese Land Law 2013 [2] showcases the system's efficacy in addressing typical queries. It is also better performance than Large Language models, such as ChatGPT, Bard and Copilot, for intelligent legal advisory systems in this aspect.

The next section presents related works with this research. Section 3 describes the integration model between ontology Legal-Onto and the knowledge graph. It also presents the method for constructing the legal knowledge map. Section 4 solving the problems for retrieving the knowledge to answer an inputted query. The proposed method is applied to design a QA system for Vietnamese Land Law 2013 in Section 5. This section also presents the experimental results of the designed system on practical queries. Its effectiveness is compared with other LLMs in the legal domain of land law. Finally, the article is concluded and given some future research directions in Section 6.

2 Related work

Nowadays there are many solutions for building intelligent query systems in legal domains [27]. These methods achieve their approaches by employing advanced technologies like Natural Language Processing (NLP), data mining, and information extraction [15, 17].

The study in [31] designed a method using MapReduce to efficiently answer complex queries on large RDF databases (like linked data). This method breaks down complex queries into simpler components, leveraging the inherent structure of the data. Additionally, they incorporated techniques to improve efficiency, such as filtering irrelevant information. In Vietnamese legal repository, it has been research on building linked data for legal documents using RDF graphs [18], it has yet to explore applications utilizing this valuable resource.

Nguyen et al. (2020) proposes a BERT-based approach for selecting answers in Vietnamese legal contexts [5]. This method, fine-tuned on a specific Vietnamese legal QA dataset, achieves an F1 score of 87%. Interestingly, pre-training the base BERT model on a domain-specific legal corpus further improves performance, reaching a score of 90.6%. While demonstrating promising effectiveness, this approach requires substantial training data and faces hurdles in keeping knowledge updated and handling intricate legal relationships.

Ontology is a useful method to organize relations between entities in legal documents [11, 25]. The study in [24] proposed an ontology Legal-Onto to represent the content of legal documents in Vietnam. However, this study did not mention the performance between documents in a field. Ngo et al. (2024) built an ontology knowledge map to represent the system of legal repository with relationships between documents [19]. Nevertheless, this model is very complex to apply for building a legal queries system which was used by residence with commonly queries. The ability to process natural language for analyzing the meaning of an inputted query should be further enhanced.

In general, existing legal query systems often struggle with limitations or require significant resources to function effectively. This paper proposes a groundbreaking solution that leverages the combined power of ontology and knowledge graph to create a robust legal information system. By integrating these structures, the proposed system enhances both the efficiency and accuracy of legal knowledge retrieval. Users can then seamlessly query and access relevant legal information through the knowledge base and knowledge graph, significantly reducing the time and effort required for knowledge management.

3 Integrating of Ontology and Knowledge Graph for legal documents

This section present the knowledge model integrating ontology Legal-Onto and Knowledge graph. Legal-Onto is an ontology for representing the knowledge of legal document, which is built based on the structure of ontology Rela-model. The integrating with knowledge graph can be used to organize the knowledge of multiple legal documents as a knowledge base for a legal consultancy system as QA system.

3.1 Ontology for a legal document

Legal-Onto is a useful ontology to represent the knowledge of a legal document [23, 24]. This ontology is built based on the knowledge model of relations, Rela-model [10].

Definition 3.1: Ontology Legal-Onto is the tube:

$$(\mathbb{C}, \mathbb{R}, \mathbb{RULES}) \tag{1}$$

In which, these components is improved from the structure of Rela-model [23]:

- $\mathbb{C}$ is a set of concepts, each concept c has the structure *(Name, Content, Attrs, InnerRel, Phrases)*, where *Name* is the name of concept, *Content* is the meaning of concept c, *Attrs* is a set of attributes of concept c, *InnerRel* is the collection of relations between attributes, and *Phrases* is a set of keyphrases related to concept c.
- $\mathbb{R}$ is a set of relations between concepts in $\mathbb{C}$
- $\mathbb{RULES}$ is set of deuctive rules in the legal document.

3.2 Integrating of Knowledge graph and Legal-Onto

This section proposes a knowledge model combining ontology Legal-Onto and knowledge graph to represent knowledge in various legal documents. The designed model is built on the components of previous models, including Rela-model [23] and the OAK model [20]. Essentially, the OAK model is proposed to manage mined knowledge in the form of knowledge maps derived from data mining tasks within a particular domain, such as digital agriculture [20].

Definition 3.2: The integrating of ontology Legal-Onto and Knowledge graph, called *Legal Onto-Graph model*, has the structure as follows:

$$\mathbb{K} = (\mathbb{C}, \mathbb{R}, \mathbb{RULES}) + (Key, Rel) \tag{2}$$

In which:

- $(\mathbb{C}, \mathbb{R}, \mathbb{RULES})$ is a structure of ontology Legal-Onto as Definition 3.1.
- *(Key, Rel)* is a knowledge graph representing the relationships between entities or keyphrases of legal documents. Here, Key signifies a collection of entities or key phrases extracted from the legal document, while *Rel* represents a set of arcs indicating the relationships between these entities or key phrases. This tube also represents relationships between documents based on their concepts and relations.

The core of the proposed model for representing legal documents is ontology Legal-Onto. This essentially acts as a knowledge graph built from legal documents themselves. The structure of Legal-Onto is based on two sets: concepts (like legal terms) in $\mathbb{C}$ and relations between those concepts stored in $\mathbb{R}$. While the model extracts additional elements like key phrases (Key) and relationships (Rel) from the documents.

3.3 The linked model of multiple legal documents through Legal Onto-Graph model

The legal document system in Vietnam, established by the 2015 law on legal promulgation [3] (updated in 2020), is structured into three tiers [22]:

- Laws of great consequence passed by the National Assembly.
- Regulations and directives established by the President and Ministers to enforce key legislation.
- Local regulations are established by local People's Councils to enforce national laws.

In the ontology of Legal Onto-Graph model, several classes have been referred to concepts in the Document Components Ontology (DoCO) [7]. It was presented in [18]. The knowledge graph is structured to include two components for each legal document: (1) a section dedicated to the structured items and relationships found within the documents, and (2) another section representing conceptual graphs of legal knowledge contained in the documents [19].

According to the proposed model, every legal document is processed using NLP toolkits to construct a legal knowledge graph. All elements within the knowledge map, such as keys and relations, are assigned URIs (Uniform Resource Identifiers). Subsequently, all maps extracted from raw data are converted into RDF triples[8], a widely recognized format for linked data and the basis of graph databases.

4 Querying on Legal Onto-Graph model

When a user asks a question, the system's question-answering module first identifies key elements and their relationships within the query. This is done by leveraging the knowledge base built using ontology Legal-Onto and with the help of NLP toolkits like PhoBERT[9] and VNCoreNLP[10]. Based on this analysis, the system creates a visual representation of the query's meaning using concepts and connections. Similarly, the legal knowledge stored in the system is also transformed into a visual format. These representations are then streamlined to simplify the search process. Finally, the system matches the query's representation with the knowledge base to find the most relevant answer. The two main problems that need to be addressed when searching for information in legal documents to answer an inputted query are:

- **Problem 1:** *Building legal knowledge graph.* Using the information in the ontology, this problem is set up to build legal knowledge graph, especially relations between documents. It will enhance the performance for extracting information based on relationships between entities on the graph.

[8] https://www.w3.org/TR/n-triples/
[9] https://github.com/VinAIResearch/PhoBERT
[10] https://github.com/vncorenlp/VnCoreNLP

- **Problem 2:** *Searching on the legal knowledge domain.* This problem will search and match the meaning of an inputted query with content in the knowledge base. Comparing the matching is done with the structure of the knowledge model's constituent parts, particularly the relations and ideas and the structure of graphs. Further relationships connected to the query's content can be inferred using the knowledge model's inference rules.

4.1 Building legal knowledge graph

Each document will have conceptual legal knowledge graphs extracted using NLP toolkits. These graphs will be converted into a knowledge map format to represent the legal information within the document. The process begins with gathering legal texts and other data from the source, followed by pre-processing. This pre-processing step cleanses the data by removing errors, ensuring consistency in formatting, and verifying its accuracy. Next, the study focuses on building the ontology by collecting and defining key concepts in $\mathbb{C}$, identifying the relations ($\mathbb{R}$-set and *Rel*-set) between them, and establishing inference rules ($\mathbb{RULES}$) based on the established ontology Legal-Onto.

After developing the ontology and utilizing it as a foundation, the research extracts keyword phrases from legal texts [18]. Phrases associated with concepts serve to pinpoint nodes within the knowledge graph, whereas those linked to relationships between nodes form connections, depicting semantic associations. These steps lay the groundwork for a sturdy knowledge repository rooted in the ontological structure, resulting in a versatile knowledge graph that facilitates intelligent extraction of semantic insights from legal materials.

By removing nonsensical triples, which are lacking real connections between entities in legal documents, this optimization method reduces the complexity and size of the text [9]. It also merges redundant relationships to streamline the content further. Knowledge graph optimization simplifies and shrinks the graph by eliminating unnecessary triples unique to a document item. Similar relations are merged to avoid repetition [8].

4.2 Searching problem on legal knowledge domain

The process of legal retrieval involves converting the query into a representation within a knowledge graph, then searching through knowledge graphs that match the representation of the input query. The knowledge graph captures the semantic intricacies of the inquiry and consists of various triples with different subject types. Often, multiple questions may require additional details to create a comprehensive triple.

Given the knowledge base $\mathbf{K}$ for law documents as Legal Onto-Graph model and an input query q, the following algorithm extracts the corrseponding information for query q.

Algorithm 4.1: *Knowledge retrieval for a query*

Input: A query q.

Output: Search the content in Knowledge base **K** being suitable the meaning of query q.

Step 1: Convert the meaning of query q to a graph:

- This decomposition approach allows us to subdivide the complex question into simpler, answerable sub-questions.
- The graph structure is resolved into individual star subgraphs, in which, each subgraph represents a set of triples that all share the same subject.

Step 2: Extract the knowledge that matches the meaning of query q.

- The system examines each group of three elements (triples) in the star graph, one by one. It searches the knowledge base for similar groups, considering the following:
  - If any part of a triple in the query is missing (represented by "*"), the search focuses on the remaining parts for a match.
  - If there's a connection ("is-a" relationship) between something in the query and something in the knowledge base, the system considers them equivalent.
- Once the system finds matching relations in the knowledge base for each subgroup in the query, it combines the results. This final combination identifies subgraphs within the larger knowledge base that perfectly match the entire original star graph.

Step 3: Take the intersection set of the answer sets of the star graphs

- Determine the elements that are present in all the sets by finding their intersection.
- The process concludes by returning a list of elements that serves as the solution to the given query q.

4.3 Advantages of Legal Onto-Graph model in legal QA system

The structure of the Legal-Onto ontology provides a powerful way to represent knowledge within legal documents, making it easier to access, understand, and work with legal information. One key challenge in the legal domain is the fragmentation of knowledge between different sources. Legal-Onto helps address this by integrating knowledge from these platforms and documents. This integration, supported by knowledge graphs (KGs), creates a unified knowledge base of legal information [21]. Instead of searching disparate sources, users can access a centralized resource, simplifying and streamlining the process of finding the relevant legal knowledge they need.

Knowledge graphs take legal information retrieval a step further by enabling users to automatically ask specific legal questions. These powerful tools leverage their vast network of structured legal knowledge and existing relationships to find relevant and accurate answers within their databases. This functionality significantly benefits both legal professionals and researchers, streamlining their search for precise answers. In short, knowledge graphs make legal knowledge more accessible and user-friendly by reducing the time and effort required for legal research.

5 Experimental Results

In this section, the proposed method is applied to represent the knowledge of Vietnamese Land Law 2013 [2]. The knowledge also covers the information from Decrees instructing to do this law in [13, 14]. This knowledge base is utilized to build a demonstrate program for answering the query in land law. Moreover, this system is compared with other LLMs as chatbots in this domain. These are ChatGPT, Google Bard and Microsoft Copilot.

5.1 Experiment Setup

The knowledge of land law in Vietnam is collected from [2]. This law includes 14 chapters with 212 articles. It is a general rule for all working in the land domain. Besides, the information from two decrees, Decree on Guiding the implementation of the 2013 land law [13] and Decree on land prices [14], have been crawled. Ontology Legal-Onto integrating knowledge graph is used to represent their content and meaning.

Example 5.1: Table 1 represents the concept *"Certificate of land use rights"* by ontology Legal-Onto:

Table 1: The structure of the concept *"Certificate of land use rights"*.

Component	**Content**
Name	Certificate of land use rights.
Content	Certificate of land use rights is a legal certificate in which the State certifies the lawful land use rights and ownership of houses and land-attached assets of the person who has land use rights and ownership of houses and land-attached assets.
InnerRel	Article 3, Point 16. Article 11, Point 1. Article 75, Point 1. Article 97, Point 1,2.
Phrases	land-attached houses, land use rights, land-attached asset, sownership of houses.
Attrs	land_use_rights, Inheritance_land_attached_assets, Donation_land_attached_assets, Transfer_land_use_rights.

To fulfill the needs of ontology Legal-Onto, NLP toolkits must perform several tasks. These include *NP Chunking* breaking down sentences into smaller chunks to identify key phrases, *Named Entity Recognition* for recognizing and classifying entities like locations, organizations, people, dates, and specific terms within documents, and *Dependency Parsing* for analyzing the relationships between words to understand the meaning of sentences. Additionally, legal knowledge,

which is stored in this ontology, can be integrated with graph databases to enable efficient retrieval and analysis of this information [19].

Using the knowledge base, the question-answering (QA) system for Vietnamese Land Law 2013 is designed. The architecture of this system is shown in Figure 1. This system has two main functions: Search engine for retrieval the suitable results, and QA for a content which is inputted as a query. The second function is integrated the NLP toolkit to process natural queries.

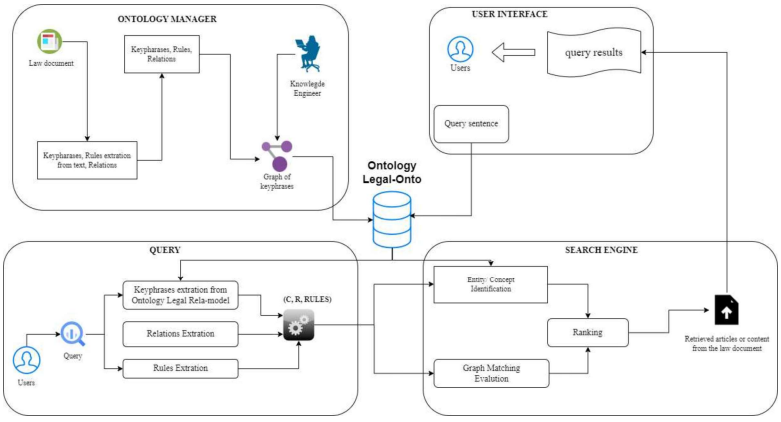

Fig. 1: The architecture of a QA system for Vietnamese Land Law.

Example 5.2: Input query q_1 = *"How are the rights and obligations of households and individuals who are leased land by the State regulated?"*

The system will extract keywords from the query q_1: *"How"*, *"rights and obiligations"*, *"households"*, *"inviduals"*, *"leased land by the State"*. From this, it returns the following results:

"According Article 166, Article 170, Article 179 of Vietnamese Land Law 2013:

- *Article 166: General rights of land users...*
- *Article 170: General obligations of land users...*
- *Article 179: Rights and obligations of households and individuals using land..."*

In this query, the system can retrieve information from Article 166 and Article 170 which are General rights and General obligations of land users. This thing can be worked based on relations of the knowledge graph in Legal-Onto.

5.2 Comparing with LLMs

Nowadays, there are many lagre language models, which are designed as a chatbot to support in many domains. This section compares the answer of the de-

signed system and some LLMs in the answering of legal questions. Those LLMs are: ChatGPT of OpenAI[11], Google Bard[12], and Microsoft Copilot[13].

The answer to a legal question must individualize the existing legal norms to specific cases, for specific individuals and organizations. This purpose is similar to that of law application. The application of law is an activity that expresses the organization and power of the state. It is carried out by state agencies, organizations, or individuals with the authority to do so in accordance with the law. The aim of this activity is to individualize the existing legal norms to specific cases, for specific individuals and organizations.

Based on Article 156, Law on Promulgation of Legislative Documents [28], the criteria for evaluation of law application includes:

- **Accuracy:** The answer must be accurate and not misleading. It must comply with the provisions of the current law. Besides, the answer must provide specific and detailed information. It should not be vague or general.
- **Completeness:** The answer must include all the information necessary to answer the question. It should not omit any important information.
- **Relevance:** The answer must be relevant to the context of the question. It should not contain irrelevant or unnecessary information.
- **Timeliness:** The answer must be based on the latest legal information. Laws change frequently, so the answer should reflect those changes.

Following these criteria, citation of a legal document is understood to be the act of presenting or citing something as evidence, illustration, or support for an argument. At this time, the person presenting the viewpoint needs to cite evidence and data to prove that their viewpoint is correct, and feasible. The act of citation must ensure the accuracy of the cited document. Therefore, when answering a legal question, the citing content is the most important thing to increase the authenticity of the content and ensure accuracy. Thus, the conditions for evaluation the answer of LLMs in legal applications are:

- Condition 1: Assess the accuracy of the citation of regulations, decrees, circulars, etc. issued by the National Assembly, including verifying the specific clause, the effective time of the cited document, etc.
- Condition 2: The completeness of the content compared to the cited legal documents.
- Condition 3: The accuracy of the content of the answers of the systems.

The set of question-answering for testing has 118 questions about Vietnamese Land Law [2]. The testing set includes questions about agricultural land, non-agricultural land and documents of Land-related procedures. Table 2 summarises the results for answering these questions of LLMs and the designed system.

Table 3 makes the comparison between our system, ChatGPT, and LLMs based on the criteria: *Accuracy, Completeness*, and *Usability*, where the criterion

[11] ChatGPT: https://chat.openai.com/
[12] Bard: https://bard.google.com/
[13] Copilot: https://copilot.microsoft.com/

Table 2: Compare the proposed method with other LLMs.

Kinds	**Quant.**	**ChatGPT**	**Bard**	**Copilot**	**Our system**
Questions about aricultural land	22	3	8	10	15
Questions about non-aricultural land	10	2	4	5	8
Questions about documents of land-related procedures	86	2	42	47	55
Total	**118**	**7**	**54**	**62**	**78**

of *Usability* performs the relevant to the context of the question and the reflection of the changing on latest legal information.

6 Conclusions and Future Works

This research introduced a new model, Legal Onto-Graph model, which combines the strengths of ontologies and knowledge graphs. This innovative approach effectively organizes and presents legal information in a clear and consistent manner across diverse legal documents. As a result, legal professionals and researchers can now analyze and comprehend complex legal content much more efficiently.

Legal Onto-Graph model is utilized to represent the knowledge in Vietnam Land Law 2013 [2] and two decrees, Decree on Guiding the implementation of the 2013 land law [13] and Decree on land prices [14]. This structured knowledge graph allows the system to answer specific legal questions, with the quality of the answers hinging on the accuracy and comprehensiveness of the underlying data. The findings of the research suggest that knowledge bases hold significant promise in enhancing land law. They provide easier access and understanding of legal information.When compared to other LLMs, the developed system demonstrates superior performance as an intelligent consulting tool in the legal domain.

In the future, a method for updating the knowledge base is studied to apply with legal documents for revision some old ones. In addition, the utilization of AI in legal matters needs to be controlled by humans. Based on the proposed model and designed ontology for presenting legal knowledge, legal documents are transformed into the linked data format and imported into the knowledge management system for retrieval. Moreover, the enhancement of NLP tookits helps to create ontology and knowledge graph more faster.

Acknowledgement

This research is supported by Sai Gon University under grant number CSB2023-04.

References

1. Armour, J., Sako, M.: Ai-enabled business models in legal services: from traditional law firms to next-generation law companies? Journal of Professions and Organization **7**(1), 27–46 (2020)
2. Assembly, V.N.: Land Law, Law No. 45/2013/QH13 (2013)
3. Assembly, V.N.: Law on Promulgation of Legislative Documents, No. 80/2015/QH13 (2015)
4. Atkinson, K., Bench-Capon, T., Bollegala, D.: Explanation in ai and law: Past, present and future. Artificial Intelligence **289**, 103387 (2020)
5. Chau, C.N., Nguyen, T.S., Nguyen, L.M.: Vnlawbert: A vietnamese legal answer selection approach using bert language model. In: 2020 7th NAFOSTED Conference on Information and Computer Science (NICS). pp. 298–301 (2020). https://doi.org/10.1109/NICS51282.2020.9335906
6. Collenette, J., Atkinson, K., Bench-Capon, T.: Explainable ai tools for legal reasoning about cases: A study on the european court of human rights. Artificial Intelligence **317**, 103861 (2023)
7. Constantin, A., Peroni, S., Pettifer, S., Shotton, D., Vitali, F.: The document components ontology (doco). Semantic web **7**(2), 167–181 (2016)
8. Dang, D., Pham, V., Cao, T., et al.: A practical approach to leverage knowledge graphs for legal query. In: Proceedings of the International Conference on Intelligent Systems and Data Science (ISDS 2023), Can Tho, Vietnam. pp. 271–284 (November 2023)
9. Dang, D., Nguyen, H., Ngo, H.e.a.: Information retrieval from legal documents with ontology and graph embeddings approach. In: 36th International Conference on Industrial, Engineering Other Applications of Applied Intelligent Systems (IEA/AIE 2023), Shanghai, China. In press. (2023)
10. Do, N.V., Nguyen, H.D., Selamat, A.: Knowledge-based model of expert systems using rela-model. International Journal of Software Engineering and Knowledge Engineering **28**(08), 1047–1090 (2018)
11. Garcia-Godinez, M.: A deflationary approach to legal ontology. Synthese **203**, Article number 64 (2024)
12. Governatori, G., Bench-Capon, T., Verheij, B., et al.: Thirty years of artificial intelligence and law: the first decade. Artificial Intelligence and Law **30**(4), 481–519 (2022)
13. Government, V.: Decree on Guiding the implementation of the 2013 land law, No. 43/2014/ND-CP (2014)
14. Government, V.: Decree on land prices, No. 44/2014/ND-CP (2014)
15. Le, H.H., Nguyen, C.T., Ngo, T.P., et al.: Intelligent retrieval system on legal information. In: Proceedings of 15th Asian Conference on Intelligent Information and Database Systems (ACIIDS 2023), Phuket, Thailand, July 2023. pp. 97–108. Springer (2023)
16. Martin, L., Whitehouse, N., Yiu, S., et al.: Better call gpt, comparing large language models against lawyers (2024), arXiv preprint arXiv:2401.16212v1
17. Martinez-Gil, J.: A survey on legal question–answering systems. Computer Science Review **48**, 100552 (2023)
18. Ngo, H.Q., Nguyen, H.D., Le-Khac, N.A.: Building legal knowledge map repository with nlp toolkits. In: Conference on Information Technology and its Applications. pp. 25–36. Springer (2023)

19. Ngo, H.Q., Nguyen, H.D., Le-Khac, N.A.: Ontology knowledge map approach towards building linked data for vietnamese legal applications. Vietnam Journal of Computer Science pp. 1–20 (2024). https://doi.org/10.1142/S2196888824500015
20. Ngo, Q.H., Kechadi, T., Le-Khac, N.A.: OAK: Ontology-based knowledge map model for digital agriculture. In: Future Data and Security Engineering: 7th International Conference (FDSE). vol. LNCS 12466. Springer (2020)
21. Nguyen, H., Pham, V., Ngo, H.Q., Huynh, A., Nguyen, B., Machado, J.: Intelligent search system for resume and labor law. PeerJ Computer Science **10**, e1786 (2024)
22. Nguyen, P.K.: How to conduct research in vietnamese law: Overview of the legal system of the socialist republic of vietnam. International Journal of Legal Information **27**(3), 307–331 (1999)
23. Nguyen, T.H., Nguyen, H.D., Pham, V.T., Tran, D.A., Selamat, A.: Legal-Onto: An ontology-based model for representing the knowledge of a legal document. In: Proceedings of 17th Evaluation of Novel Approaches to Software Engineering (ENASE 2022), Online streaming. pp. 426–434 (2022)
24. Pham, V.T., Nguyen, H.D., Le, T., et al.: Ontology-based solution for building an intelligent searching system on traffic law documents. In: Proceedings of 15th International Conference on Agents and Artificial Intelligence (ICAART 2023), Lisbon, Portugal. pp. 217–224 (2023)
25. Phan, T.T., Lam, H.P., Hashmi, M., Choi, Y.: Towards construction of legal ontology for korean legislation. In: Proceedings of 12th International Conference on Knowledge Engineering and Ontology Development (KEOD 2020). pp. 86–97 (2020)
26. Sartor, G., Araszkiewicz, M., Atkinson, K., et al.: Thirty years of artificial intelligence and law: the second decade. Artificial Intelligence and Law **30**, 521–557 (2022)
27. Szostek, D., Zalucki, M.: Legal Tech: Information Technology tools in the administration of justice. Nomos (2021)
28. of Vietnam National Assembly, O.: Law on Promulgation of Legislative Documents, No. 23/VBHN-VPQH (2020)
29. Viettel: Virtual assistant for courts. https://trolyao.cyberbot.vn/ (2024), accessed: 2024-02-26
30. Villata, S., Araszkiewicz, M., Ashley, K., et al.: Thirty years of artificial intelligence and law: the third decade. Artificial Intelligence and Law **30**, 561–591 (2022)
31. Wang, X., Chai, L., Xu, Q., Yang, Y., Li, J., Wang, J., Chai, Y.: Efficient subgraph matching on large rdf graphs using mapreduce. Data Science and Engineering **4**, 1–20 (03 2019). https://doi.org/10.1007/s41019-019-0090-z

 ISBN: 978-604-80-9774-5

Table 3: Comparison of the designed system with ChatGPT and Bard in Legal Domain

The system	**Accuracy**	**Completeness**	**Usability**
Chat GPT	ChatGPT effectively grasps natural language, aiding in the understanding and addressing of various legal inquiries, even those involving intricate situations. it cannot furnish precise, detailed figures or specifics regarding the content of the legal documents it analyzes. Users should exercise caution and confirm information from authoritative legal sources.	It lacks specialized legal knowledge within specific domains. The deficiency in interpreting and applying legal concepts may lead to inaccuracies. It does not offer precise details extracted from legal texts.	The knowledge of ChatGPT may not be consistently up-to-date because it lacks a mechanism for real-time updates and continuous monitoring of changes in legal regulations and interpretations. Users should cross-check and verify information from reliable sources to ensure the accuracy and reliability of the answers provided.
Bard	Bard possesses the capability to access and provide an overview of legal answers available online. It can extract information from diverse sources and offer specific values directly from legal texts. Because the information has not been structured into a knowledge base, Bard's effectiveness in responding to queries that necessitate data from numerous legal documents is compromised. It tends to provide simplistic answers rather than comprehensive insights.	The system relies on online sources and may not guarantee the accuracy, currency, or completeness of legal answers.	The responses provided by Bard are regularly updated to reflect the latest legal documents, ensuring accuracy in terms of the substantive content of fines. However, there may be inaccuracies in the positioning of citations within the legal text, and the supplementary penalties presented might also be incorrect.
Our system	The integration of knowledge fields into the legal domain has been shown to significantly improve information retrieval accuracy. This improvement facilitates easier access to legal information for non-experts, thereby enhancing transparency and enabling informed decision-making. By combining ontology and knowledge graphs, the system identifies inconsistencies and errors, thus ensuring that the legal information remains accurate and dependable.	The knowledge base integrates knowledge from diverse legal sources, organized directly from the source of law documents. This approach ensures centralized and easily accessible legal information. The relations within the knowledge graph enable the extraction of necessary information from multiple documents to address complex queries effectively.	By leveraging the knowledge extracted from legal documents, users can reference regulations with confidence in the accuracy and reliability of the responses. The system is evolving to launch a practical application supporting individuals in searching for legal knowledge.