Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
16/2 (2023), 125-140. DOI: http://dx.doi.org/10.21609/jiki.v16i2.1152

A Hybrid Virtual Assistant for Legal Domain Based on Information Retrieval and Knowledge Graphs

Douglas Raevan Faisal[1], Fariz Darari[1], Muhammad Ilham Al Ghifari[1],
Muhammad Zuhdi Zamrud[1], Marcellino Chris O'Vara[1],
Berty Chrismartin Lumban Tobing[2], On Lee[3]

[1] Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
[2] CATAPA, Jakarta, Indonesia
[3] GDP Labs, Jakarta, Indonesia

Email:douglas.raevan11@ui.ac.id

Abstract

Virtual assistants have gained popularity across various domains, including the legal field, where they serve to offer guidance and aid in the form of law retrieval. In this research, our aim is to develop a legal virtual assistant that combines knowledge graphs (KGs) and information retrieval (IR) techniques. This hybrid approach allows us to provide accurate answers extracted from structured interconnected data while simultaneously cater to a diverse range of legal inquiries. We categorize these inquiries into a few distinct use cases: definition lookup, law component lookup, sanctions, and domain knowledge. Our system encompasses a chatbot platform, knowledge graph querying, and information retrieval. Specifically, we construct a VA system over a legal knowledge graph pertaining to the Indonesian Act concerning Manpower or Labor (*UU Ketenagakerjaan*) and the Indonesian Act concerning the Creation of Jobs (*UU Cipta Kerja*). This marks the creation of the first legal virtual assistant in the Indonesian context that combines KG and IR methodologies. To evaluate the effectiveness of our prototype system, we conduct tests using a variety of labor law-related questions, ranging in difficulty. The integration of knowledge graphs and information retrieval proves to significantly improve the support provided for a wide range of potential applications in the legal field.

Keywords: *Law; Virtual Assistant; Information Retrieval; Knowledge Graph*

1. Introduction

According to Gartner,[1] virtual assistants (VAs) have emerged as essential tools in assisting users with tasks that were traditionally performed manually. By harnessing artificial intelligence (AI) technologies, VAs are able to comprehend user inquiries and make optimal decisions to accomplish the assigned tasks. VAs are available in various forms, including chatbots in companies and smart assistants. Notably, such applications that have recently spurred active discourses around chatbots are OpenAI's ChatGPT and Google's Bard.

[1]https://www.gartner.com/en/information-technology/glossary/virtual-assistant-va

Within the legal field, virtual assistants (VAs) have the ability to offer valuable legal information and can even extend their services to provide professional legal consultation [1]. Various companies have developed legal VA applications in the form of legal document automation and legal chatbot for professional consultation [2]. Most of those applications utilize AI and natural language processing (NLP) techniques. Moreover, they empower individuals with diverse circumstances and needs to effortlessly access legal advice or legal documents [2].

The primary function of a legal virtual assistant (VA) is to respond to user inquiries within the legal domain. Figure 1 demonstrates the capabilities of a legal VA, such as retrieving specific sections from

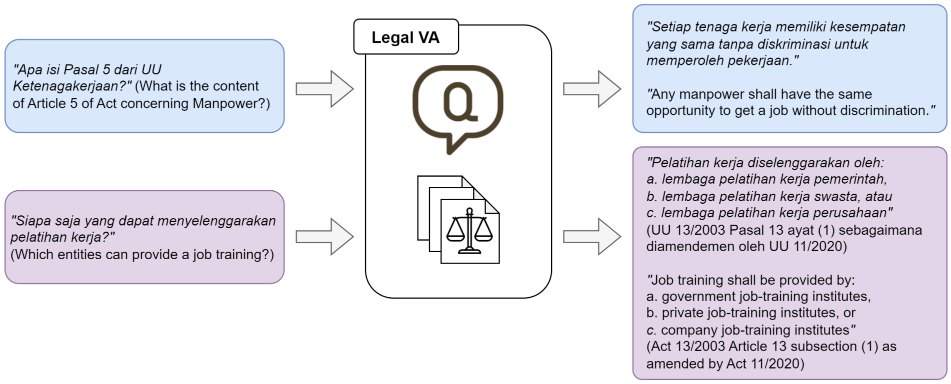

Figure 1. Legal Virtual Assistant Capabilities

legal documents and providing answers to open-ended questions. Additionally, it is imperative for the system to be accessible at all times. Given that legal documents are primarily text-based, the legal VA should also incorporate text processing and comprehension capabilities.

Nevertheless, the majority of legal expert systems and virtual assistants heavily rely on natural language processing (NLP) and information retrieval (IR) systems. While these systems offer substantial information, they are susceptible to errors when confronted with unfamiliar data and often encounter difficulties in delivering precise and structured responses. Notably, legal documents possess a crucial characteristic of being well-structured, enabling experts to comprehend them with minimal bias or ambiguity.

To address this challenge, one possible solution is to adopt knowledge graphs (KGs) as a means of representing legal documents [3]. KGs utilize graph structures to explicitly depict knowledge. Another way to define KG is as "a graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and that edges represent relations between entities" [4].

The concept of "Knowledge Graph" gained prominence when Google introduced it in 2012.[2] Google's Knowledge Graph enhances search results for its search engine by providing contextual information. Since then, other prominent companies, including e-commerce platforms (eBay [3]), social media platforms (Facebook [3]), and financial services providers like Bloomberg[3] and MasterCard,[4] have also adopted Knowledge Graphs. KGs prove to be valuable in capturing and analyzing the semantic relationships between concepts, benefiting various applications such as personalizer systems and knowledge bases.

In the legal domain, the application of knowledge graphs is evident in initiatives such as EUR-Lex[5] and project Lynx [5].[6] The latter aims to standardize legal documents in Europe by incorporating the European Legislation Identifier (ELI). On the other hand, the Lynx project focuses on addressing challenges in legal structured data, such as content enrichment and semantic annotations. Schema.org standardizes legal terms, extending its base vocabulary to include the "Legislation" type.[7] Previous efforts, such as Lex2KG, have concentrated on converting Indonesian law documents into a legal KG [6]. However, Lex2KG primarily focuses on the conversion process and lacks consideration for practical applications. In this case, Lex2KG necessitates manual and cumbersome KG querying, which is not user-friendly, to access and utilize the data effectively.

Since laws and regulations govern many aspects of a country's citizens, it would be interesting to find some areas that apply to most citizens. One such area would be labor and manpower affairs. Specifically in Indonesia, recent developments in regulating these affairs sparked controversy with the promulgation of the Act of Job Creation (UU *Cipta Kerja*) which amended many previously in-force regulations, including the Act of Manpower (UU *Ketenagakerjaan*) [7]. Hence, focusing on these laws would serve as a viable use case of legal VA for workers, employers, HR management teams, and also citizens in general.

[2]https://www.blog.google/products/search/introducing-knowledge-graph-things-not/

[3]https://www.bloomberg.com/company/stories/using-tables-to-build-better-knowledge-graphs

[4]https://thenewstack.io/tigergraph-graph-dbs-to-become-a-must-have-in-2022/

[5]https://eur-lex.europa.eu/

[6]https://lynx-project.eu/

[7]https://schema.org/Legislation

In this paper, we develop a feature-rich legal VA based on a hybrid approach of KGs and IR. Our primary contribution is on effectively harnessing the capabilities of knowledge graphs (KGs) and information retrieval (IR), leveraging the advantages of each technology to enhance the performance of the legal virtual assistant (VA) task. The IR system specializes in processing textual queries and generating a collection of relevant documents using fuzzy matching techniques. On the other hand, SPARQL, the query language designed for KGs, operates by processing structured queries and providing a result set consisting of pertinent entities or items. By strategically combining these technologies, we aim to optimize the functionality and effectiveness of our legal VA. The contributions of this paper are two-fold:

- We design and implement a legal VA system combining the use of knowledge graphs and information retrieval.
- We introduce the first legal virtual assistant incorporating both knowledge graphs and information retrieval for laws in Indonesia related to labor affairs. The system offers a non-trivial solution over Lex2KG [6], enabling better access to KG-based law representations. A trivial solution, in this case, is an IR-only or a KG-only system, each of which has its own drawbacks. For this study, we (along with stakeholders) discuss how to maximize the benefits of each technology in our system. The working prototype includes two major laws related to labor affairs. The demo video of the system is available online.[8]

The subsequent sections of this paper are organized as follows: In Section 2, we present a review of related works that substantiate our ongoing research. Section 3 outlines the requirements for our legal virtual assistant. We then elaborate on the architecture and the modules in our legal VA in Section 4. Section 5 provides an explanation of how the system handles various types of questions. The evaluation of the proposed system is presented in Section 6. In Section 7, we discuss several aspects concerning research opportunities and limitations. Finally, Section 8 concludes the research and outlines potential future endeavors.

2. Related Work

In this section, we delve into relevant studies that provide a foundation for our research in developing a hybrid legal virtual assistant (VA). The discussion is divided into three key components: information retrieval (IR) techniques for virtual assistants, the utilization of knowledge graphs (KGs) in virtual assistants, and the landscape of legal technology research in Indonesia.

2.1. Information Retrieval for Virtual Assistants

The primary objective of VAs is to retrieve information, which is why the development of VAs commonly incorporates IR techniques. Information retrieval is characterized as the process of "finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers)" [8].

A crucial aspect of VAs lies in their ability to accurately answer questions. Consequently, researchers and practitioners also build question-answering (QA) datasets to train their VA systems using machine learning (ML) techniques. In the legal domain, the JEC-QA team created a QA dataset specifically focused on the reading comprehension aspect of QA, centered around the National Judicial Examination of China [1]. Additionally, they developed a question-answering system tailored for addressing legal questions. The system contains the following modules: document retrieval utilizing Elasticsearch,[9] topic classification employing a neural network, and answer extraction utilizing several question answering models.

LAW-U is a virtual assistant dedicated to offering legal guidance for victims of sexual abuse in Thailand [9] based on Thailand court judgments. The flow of the VA was devised to gather comprehensive information from the user, aiming to capture all relevant details. Once the system determined that sufficient information was obtained, it referred to the court decisions dataset and provided the most credible decision based on the available data.

2.2. Knowledge Graph for Virtual Assistants

Several studies have explored the development of VAs using KGs as their underlying framework. An example is showcased in [10] which utilized knowledge graphs for question answering. The model employed Resource Description Framework (RDF) datasets and SPARQL queries.

In light of Indonesian legislation, there are two studies on restructuring legal documents into KGs:

[8]https://s.id/Lex2KGVADemo

[9]https://www.elastic.co/

Lex2KG [6] and LexID [11]. The flow of both systems can be summarized as follows. Beginning with legislation PDF documents as the input, the systems parsed these legislation documents to extract the underlying text data. Following this, through a set of rule-based techniques, the text data was organized into distinct legal components, including chapters (*bab*), articles (*pasal*), and subsections (*ayat*). Additionally, LexID performed semantic information extraction from legal components, enabling it to capture more fine-grained semantics, such as legal clauses and legal entities.

Both [6] and [11] explored a similar use case, that is, legal QA. By supplying manually-created SPARQL queries, a user can retrieve legal answers from the queries. However, none of them has explored the consumption of legal KGs with only natural language as the input. As such, their implementations of legal QA can only be utilized by expert users who are able to formulate their questions as SPARQL queries.

Another research endeavor [12] integrated semantic data with BERT [13]. Prior to retrieving the semantic data, a few modules were executed: entity identification, relation classification, and entity linking. The BERT model played a significant role in facilitating both entity identification and relation classification.

Another virtual assistant, Lynx, utilizes knowledge graphs technologies to manage an extensive collection of law documents from Europe [5]. Their project began with a focus on addressing the multilingual nature of the laws by employing KG technologies [14]. Over time, the project expanded its scope to encompass various use cases concerning legal analysis, geothermal energy, and labor regulations. The resulting knowledge graph was subsequently utilized in several NLP modules, including language translation, subject linking, and summarization.

2.3. Legal Research in Indonesia

Given the abundance of Indonesian legal documents, numerous studies have explored the use of NLP and knowledge extraction from these extensive document collections. As of June 2, 2023, the Indonesian Supreme Court alone has issued over 7.8 million court decisions. One study [15] employed deep learning and NLP techniques to retrieve court decisions and extract legal entities mentioned within them. The annotation process involved identifying and categorizing ten specific entities, including individuals involved in the legal proceedings (e.g., advocates, judges, prosecutors) as well as various document-related entities (e.g., laws, decision numbers, punishments).

In another research endeavor, a VA was created specifically for the Act concerning Information and Transactions in the Electronic Form (Act 11/2008). As a crucial component of the virtual assistant, a knowledge base (KB) consisting of a collection of inquiries and their corresponding answers was constructed. They employed text processing and similarity techniques to retrieve answers from their KB.

To the best of our knowledge, we have not found any legal virtual assistant on Indonesian laws that is capable of effectively addressing law-related questions and providing up-to-date information on the prevailing legal framework. Given that laws undergo amendments over time and govern numerous aspects of citizens' lives, it is essential for a legal VA to stay up-to-date with the latest laws and regulations. Furthermore, the existing Indonesian legal VAs fall short in accommodating the wide range of user questions and intentions, which will be discussed further in the subsequent section.

Some of the systematics of Indonesian legislation need to be understood prior to applying technologies to laws. Indonesian laws and regulations are written in compliance with the Act 12/2011 on Legislation Making. They are categorized and structured in a hierarchy, including (but not limited to) the 1945 Constitution (*Undang-Undang Dasar*), Law/Act (*Undang-Undang*), Government Regulation (*Peraturan Pemerintah*), Provincial Regulation (*Peraturan Daerah (Provinsi)*), and Regency/Municipal Regulation (*Peraturan Daerah (Kabupaten/Kota)*). Each law document consists of several articles (*pasal*) which are defined as a unit of rules that contain a norm in the form of a sentence. Articles can be broken down into some subsections (*ayat*). Articles and subsections can also contain details of elements that are tabulated into letters (*huruf*).

3. Legal Virtual Assistant Requirements

When constructing a virtual assistant (VA) specifically designed for the legal domain, it is crucial to consider various aspects and requirements, which depend on the specific legal tasks that need to be addressed, such as document automation, legal compliance, and question answering. After conducting a preliminary examination of the jurisdiction in Indonesia and relevant studies, we have categorized law-related VA use cases into 4 distinct groups, as outlined in Table 1.

The first use case, definition lookup, pertains to situations where users seek the legal definition of

Table 1. General Legal VA Use Cases.

#	Use Cases	Example Question	Expected Answer
			Indonesian
1	Definition Lookup	"*Apa arti dari pelatihan kerja*?"	"*Pelatihan kerja adalah keseluruhan kegiatan untuk memberi, memperoleh, meningkatkan, serta mengembangkan kompetensi kerja, produktivitas, disiplin, sikap, dan etos kerja pada tingkat keterampilan dan keahlian tertentu sesuai dengan jenjang dan kualifikasi jabatan atau pekerjaan.*"
2	Law Component Lookup	"*Apa isi Pasal* 5 UU *Nomor* 13 *Tahun* 2003?"	"*Setiap tenaga kerja memiliki kesempatan yang sama tanpa diskriminasi untuk memperoleh pekerjaan.*"
3	Sanction	"*Apakah ada hukuman bagi pengusaha yang mempekerjakan pekerja lebih dari waktu kerja tanpa ada persetujuan*?"	"(1) *Barang siapa melanggar ketentuan sebagaimana dimaksud dalam Pasal* 38 *ayat* (2), *Pasal* 63 *ayat* (1), *Pasal* 78 *ayat* (1), *Pasal* 108 *ayat* (1), *Pasal* 111 *ayat* (3), *Pasal* 114, atau *Pasal* 148 *dikenai sanksi pidana denda paling sedikit* Rp5.000.000,00 (*lima juta rupiah*) *dan paling banyak* Rp50.000.000,00 (*lima puluh juta rupiah*). (2) *Tindak pidana sebagaimana dimaksud pada ayat* (1) *merupakan tindak pidana pelanggaran.*"
4	Domain Knowledge	"*Apa yang harus dipersiapkan jika ingin mempekerjakan tenaga kerja asing*?"	"*Setiap pemberi kerja yang mempekerjakan tenaga kerja asing wajib memiliki rencana penggunaan tenaga kerja asing yang disahkan oleh Pemerintah Pusat.*"
			English (Translated)
1	Definition Lookup	"What is the definition of job training?"	"Job training (*pelatihan kerja*) shall refer to the whole activities of providing [workers or potential workers with, and paving the way for them to] acquire, enhance and develop job competence, productivity, discipline, work attitude and ethics until a [desired] level of skills and expertise that match the grade and qualifications required for a position or a job is reached."
2	Law Component Lookup	"What is the content of Article 5 of Act (*Undang-Undang*) Number 13 Year 2003?"	"Every person available for a job shall have the same opportunity to get a job without discrimination."
3	Sanction	"Are there penalties for companies who employ workers to work for more than their working hours without approval?"	"(1) Whoever violates the provisions referred to in Article 38 subsection (2), Article 63 subsection (1), Article 78 subsection (1), Article 108 subsection (1), Article 111 subsection (3), Article 114, or Article 148 shall be subject to criminal sanctions of a minimum fine of Rp5,000,000.00 (five million rupiahs) and a maximum of Rp50,000,000.00 (fifty million rupiahs). (2) The crime referred to in subsection (1) is a criminal offense."
4	Domain Knowledge	"What should be prepared if an employer wants to hire foreign workers?"	"Every employer who employs foreign workers is required to have a plan for using foreign workers which is approved by the Central Government."

specific entities mentioned within legal documents. Legal definitions hold significant importance as they offer an official and precise description of these entities. By accessing these precise definitions, users obtain reliable and solid definitions of the subjects in question.

The next one, law component lookup, facilitates retrieval of law components. When provided with a particular reference, such as "Article 8 subsection (2) of *UU Ketenagakerjaan*", the virtual assistant should be capable of retrieving the precise information associated with the reference. Moreover, it should be capable of identifying any changes or amendments that may have been made to the law component. Typically, when users search for a law component, they require the most up-to-date and currently applicable version.

The subsequent use case, sanction, focuses on situations where users inquire about the penalties for a violation. It requires a couple of resolution process: (*i*) the VA needs to accurately identify the specific article that has been violated, followed by (*ii*) locating the corresponding punishments associated with the violated component. The punishments provided should be comprehensive, encompassing the monetary fines and potential imprisonment penalties.

Lastly, the domain knowledge use case involves providing accurate law components in response to free-form inquiries. This usage is broad in nature since it has no particular limitations and the inquiries can differ significantly based on the users' requirements. In this scenario, users may ask questions related to the legal domain and expect the retrieval of the most relevant law and its corresponding components. For instance, when presented with a labor-related question, the system must locate the appropriate law component as the answer.

In addition to handling the four main use cases, there are several supplementary aspects that require attention. Firstly, in order to comply with the most up-to-date regulations, the system needs to identify the amended sections of the laws and provide the presently effective versions. Furthermore, it should also provide relevant metadata, such as identifying "which law components are referenced or being referenced by the provided answer","the chapter in which the law component is located", and "the specific domain to which the law pertains".

Considering the aforementioned requirements, we suggest a blend of methodologies that incorporate the semantic data representation of knowledge graphs along with the adaptability of information retrieval. Knowledge graphs offer a structured and interconnected representation of laws, which can result in accurate responses. On the other hand, IR enables the retrieval of answers that closely align with user queries, allowing for greater flexibility in handling open-ended questions.

Our proposed technique for developing a legal VA system is motivated by the limited research on hybrid approaches to retrieval techniques in Indonesian legal VAs. One example of a method that combines graph data and IR to create a QA dataset is the HotpotQA project [16]. In this project, a linked data structure was built using Wikipedia hyperlink data. In the legal domain, the Lynx project utilized a combination of information retrieval and knowledge graphs to develop their virtual assistant for legal use [5]. However, our system is dissimilar from Lynx in that our data comprises Indonesian language while theirs utilizes multilingual data. and other specific technical details, which we will explain in more detail in the subsequent parts of the paper.

4. Lex2KG-VA

We now discuss the overall development of our hybrid legal VA system, which we will refer to as Lex2KG-VA. The section is divided into two parts: the architecture of Lex2KG-VA and the composition of the question-answering (QA) component.

4.1. Lex2KG-VA Architecture

Figure 2 illustrates the architecture and process flow of Lex2KG-VA. The initial step involves feeding the legal documents into Lex2KG, which converts them into a legal knowledge graph (KG). Our current prototype primarily focuses on the Act concerning Manpower (UU 13/2003) and its significant amendments, particularly Article 81 of the Act concerning the Creation of Jobs (UU 11/2020). These acts establish the fundamental regulations concerning labor and manpower affairs that are presently enforced in Indonesia.

The two laws present some interesting use cases and challenges. Labor and manpower regulations are the core reference for employers, workers, and HR teams, which accounts for the majority of citizens. In legal terms, the two acts also represent amendments between two different regulations, where the Act of Job Creation amends several articles in the Act of Manpower. This challenge needs to be accounted for when democratizing a legal VA to make sure that it always provides the latest regulation in force.

The Act concerning Manpower encompasses a total of 193 articles, while Article 81 of the Act concerning the Creation of Jobs introduces 68 changes to the prior law. Each amendment point represents the alteration made to a single article in the amended law. The combination of these two acts results in a legal KG comprising over 13,000 RDF triples. To guarantee the reliability of the legal knowledge graph, we conduct a thorough quality check and address various problems, including segment duplicates, inaccurate references, and parsing errors. These measures are taken to enhance the overall quality and integrity of the legal KG.

Subsequently, the legal knowledge graph generated by Lex2KG is fed into our system, which comprises 3 primary modules. The first one–general flow module–receives user inquiries, extracts the intents and parameters, and directs them to the corresponding retriever modules. Next, the knowledge graph retrieval module generates queries in SPARQL to retrieve law components. Last, the text-based retrieval performs keyword-based queries to provide relevant answers to user inquiries. Additionally, the answers from the IR module can be enhanced using contextual information from the legal knowledge graph, if necessary. After the law component has been acquired, the system formats the output to present it as answers from the legal virtual assistant to the user's inquiries.

We compare our methods and techniques with LAW-U [9] and JEC-QA [1] in Table 2. In terms of the use cases, jurisdictions, and data sources, all three VAs differ from one another. However, we can still compare the techniques employed by each VA. JEC-QA [1] applies deep-learning-based embeddings for its retrieval, such as CNN and BERT [13]. Our system employs a different approach by leveraging KGs to better structure and query the results, enabling the provision of precise answers. We also perform intent classification and

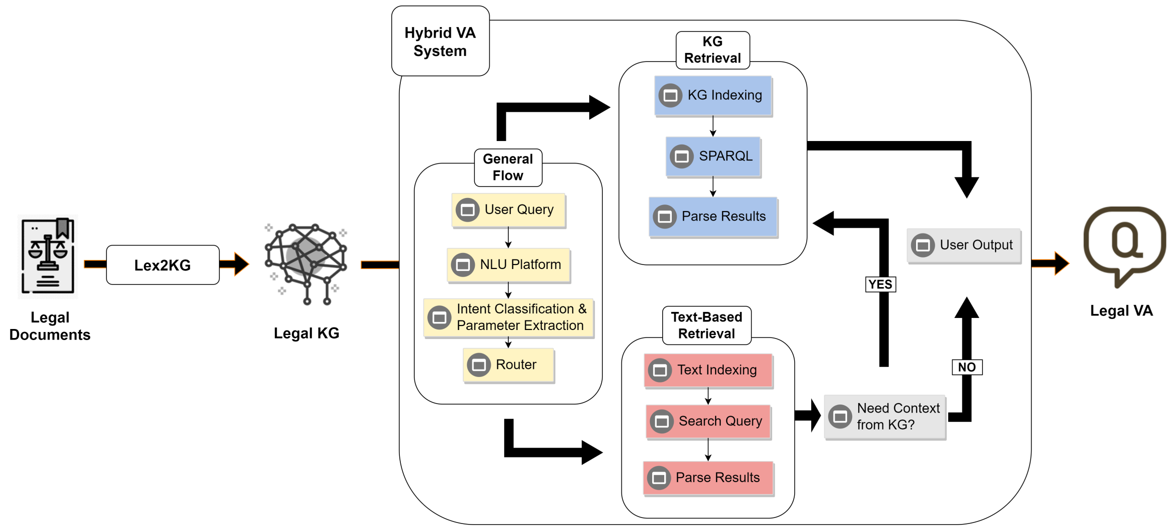

Figure 2. Lex2KG-VA Architecture

parameter extraction in order to differentiate four legal intents and several important parameters which will be discussed in Subsection 4.2.

4.2. QA Composition

Now we will provide a more detailed explanation of the fundamental elements of Lex2KG-VA that enable it to answer user questions effectively.

4.2.1. Intent Classification. This component is responsible for understanding user questions and identifying their intents. Based on the four question types outlined in Table 1, we define corresponding intents for each type. Each intent is then processed in a manner that ensures accurate and concise answers to user questions. For instance, when dealing with law component questions, the system strives to provide the exact and complete text of the specified component. On the other hand, domain knowledge questions typically involve matching domain-related terms with the most relevant text in the laws, without explicitly specifying a law reference. We employ a machine learning approach to train the intent classification model, utilizing an established conversational AI platform. We select a few question samples (approximately 5-10 questions per intent) to enable accurate intent classification. Here are some examples of the training phrases for each intent:

1) Definition lookup
   - “What does the entity ‘Minister of Manpower’ refer to?” (“*Apakah itu Menteri Tenaga Kerja*?”)
   - “What is meant by apprenticeship?” (“*Apa yang dimaksud dengan pemagangan?*”)
2) Law component lookup
   - “What is the content of Article 10 of Act 13/2003?” (*“Apa isi Pasal 10 UU 13/2003?”*)
3) Sanction
   - “What are the sanctions for employers who do not fulfill their employees’ severance pay?” (“*Apa sanksi bagi pengusaha yang tidak memberikan pesangon bagi karyawannya?*”)
4) Domain knowledge
   - “Which jobs use fixed term employment contract?” (“*Pekerjaan apa saja yang dapat menggunakan PKWT*?”)
   - “Which entity is responsible for work protection?” (“*Menjadi tanggung jawab siapa terkait perlindungan kerja?*”)

Selecting a set of training phrases involves intense discussions with domain experts from our team to populate initial sentences. These sentences are then fed into training for the intent classification module. We iteratively update the training phrases based on the evaluation results and the inputs of the aforementioned domain experts.

Table 2. Comparison between our system (Lex2KG-VA) with the other legal VAs. The technique advantages are highlighted in bold.

Aspects	**LAW-U [9]**	**JEC-QA [1]**	**Lex2KG-VA (Ours)**
Use Case(s)	Legal Action Guidance	General Legal QA Dataset	Legal VA on Labor Affairs
Jurisdiction	Thailand	China	Indonesia
Data Source	Supreme Court Decisions	Law Articles and Subsections.	Law Articles, Subsections, and Letters.
QA Techniques	Text Retrieval (TF-IDF)	Text Retrieval (TF-IDF, **Deep Learning**), **Reading Comprehension**	**Knowledge Graphs**, Text Retrieval (BM25), **Intent Classification**, and **Parameter Extraction**

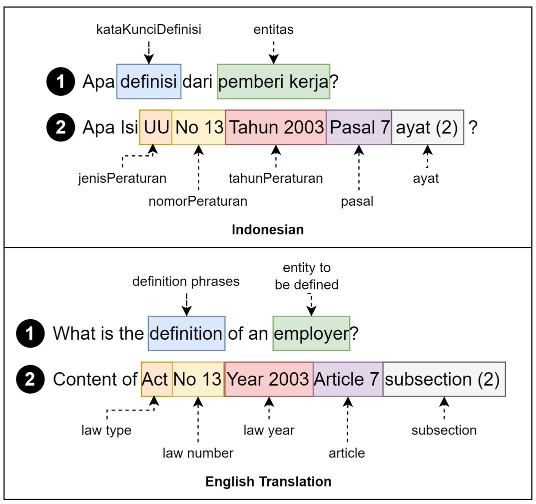

Figure 3. Parameter Extraction Examples

4.2.2. Parameter Extraction. Parameter extraction involves identifying specific parts of questions or queries that are relevant for further processing during the fulfillment stage. It also aids in intent classification by helping to differentiate between different types of intents. Figure 3 provides examples of parameter extraction. For instance, consider the first question ("*Apa definisi dari pemberi kerja?*" – "What is the definition of an employer?"), that asks for a definition of an entity. In this case, we extract two parameters: *kataKunciDefinisi* (definition phrases) and *entitas* (entity). The next question ("*Isi UU No 13 Tahun 2003 Pasal 7 ayat (2)*" – "Content of Act No 13 Year 2003 Article 7 subsection (2)") requires extensive parameter extraction, as it pertains to a law component lookup that necessitates precise answers.

In our most recent development iteration, we manually define 11 parameter categories. As presented in Table 3, we apply distinct techniques for extracting each type (or category). The simplest one involves using synonyms, in which we curate a collection of terms to match a specific category. For parameters with fixed syntactical patterns, we utilize regular expression (regex) patterns. On the other hand, for parameters without syntactical patterns or exact matches, we employ fuzzy matching. Fuzzy matching involves generating various permutations of provided terms/phrases to match during the extraction process. For instance, if we define the term "*tenaga kerja asing*" ("foreign workers"), it creates a set of permutations, such as "*tenaga*", "*tenaga kerja*", "*kerja asing*", "*tenaga asing*", and so on, to find matches.

4.2.3. KG-Based Retrieval Fulfillment. Fulfillment refers to providing accurate answers to user questions based on identified intents. In this subsection, we will discuss the fulfillment method using KG approaches. Leveraging the legal KG, we establish SPARQL query templates to get the relevant law components associated with various intents.

KG-based retrieval encompasses four primary use cases. The initial use case involves retrieving a particular law component based on given parameters. The second use case utilizes SPARQL queries to retrieve the most recent version of law articles. The third use case involves retrieving references to law components mentioned within the text. For instance, when examining sanctioning articles, it is crucial to identify the corresponding article to the sanction. Lastly, KG retrieval facilitates regex matching to some extent, enabling the search for definitions of specific entities mentioned within the law document.

For Definition Lookup and Law Component Lookup intents, the fulfillments rely solely on KG-based approaches. Law Component Lookup utilizes parameters associated with law components (such as *pasal* for article, *ayat* for subsection, *tahunPeraturan* for the law year, *nomorPeraturan* for the law number) to retrieve the corresponding KG entity. On the other hand, Definition Lookup queries search for articles within the "General Definitions" ("*Ketentuan Umum*") chapter of the law that contain the terms to be defined.

4.2.4. Text-Based Retrieval Fulfillment. This fulfillment method employs information retrieval techniques to retrieve relevant text passages that match

Table 3. Extracted Parameter Types.

#	Parameter Type	Method*	Examples	Translated Examples
1	Subsection (*ayat*)	Reg.	*ayat (2), Ayat 10*	subsection (2), Subsection 10
2	Domain Knowledge Phrases (*bagianPertanyaan-Substansi*)	FM	*apa saja persyaratan, apa kewajiban dari*	what are the requirements, what are the obligations
3	Entity to be Defined (*entitas*)	FM	*ketenaga-kerjaan, tenaga kerja, pemberi kerja*	manpower, worker, employer
4	Letter (*huruf*)	Reg.	*huruf a, Huruf c*	letter a, Letter c
5	Law Type (*jenisPeraturan*)	Syn.	UU (*Undang-Undang*), Perda (*Peraturan Daerah*)	Act, Regional Regulation
6	Law Title (*judulPeraturan*)	Syn.	*ketenaga-kerjaan*	manpower
7	Definition Phrases (*kataKunciDefinisi*)	FM	*definisi, pengertian, apa itu*	definition, meaning, what is
8	Law Number (*nomorPeraturan*)	Reg.	*Nomor 13, no 20*	Number 13, no 20
9	Law Article (*pasal*)	Reg.	*Pasal 25, pasal ke-40*	Article 25, 40th article
10	Sanction Questions (*pertanyaan-Hukuman*)	Syn.	*denda, hukuman, sanksi*	fine, punishment, sanction
11	Law Year (*tahunPeraturan*)	Reg.	*thn 2005, Tahun 2020*	yr 2005, Year 2020

***Notes:** FM: Fuzzy Matching, Reg.: Regex, Syn.: Synonyms

a given text input. It complements the knowledge graph retrieval by retrieving law components based on similarity rather than exact matches. In our prototype, we implement simple preprocessing steps such as lowercasing and tokenization using unigrams. For the retrieval, we utilize Okapi BM25 [17], which is a document ranking model that considers the frequency of words in a document to retrieve the most relevant matching documents [8].

We generate an index collection for the textual content of the law components. We extract all the text segments from the law documents, associating them with their respective Uniform Resource Identifiers (URIs) retrieved from the KG data. The indexing process follows a granularity level similar to that of the KG. Our aim is to achieve a fine-grained indexing of the documents to ensure comprehensive coverage.

Text-based retrieval is utilized for two intents: Domain Knowledge and Sanction. In both cases, the system searches the entire content of the law documents to find relevant answers. In the Domain Knowledge intent, the system retrieves the law component that best matches the user's question. As for the Sanction intent, the process closely resembles the prior intent, but with an extra step of knowledge graph querying to obtain the punishments for the identified article or subsection.

5. Step-by-Step Intent Fulfillments

We will now explore the fulfillment process of Lex2KG-VA for various use cases and intents. The current prototype addresses a total of four intents, derived from the requirements elaborated previously. The flowchart diagram presented in Figure 4 depicts the fulfillment process for each intent, highlighting the extraction of relevant parameters from user questions, in-depth fulfillment procedures, and identifying the relevant sections of the laws involved at each stage.

5.1. Definition Lookup

The fulfillment process for the Definition Lookup intent primarily relies on KG-based retrieval, with the fallback option of IR-based retrieval if the former method fails to provide results. In Indonesian laws, it is customary for definitions of legal terms within the laws themselves to be located in the chapter titled "General Definitions" ("*Ketentuan Umum*"). Therefore, the necessary data can be obtained from that specific chapter.

In the definition lookup intent, the primary information required is the entity (*entitas*) that needs defining. For instance, in the question "what is the definition of worker?" ("*apa definisi dari tenaga kerja?*"), the entity would be "worker" ("*tenaga kerja*"). Subsequently, it is utilized to generate KG queries, which restrict the scope to the "General Definitions" ("*Ketentuan Umum*") chapter of the law. The SPARQL queries employ a regex pattern to find the definition in the form of "`[ENTITY]` is ..." ("`[ENTITAS]` *adalah* ..."). It is important to note that there are certain exceptional cases where multiple entities share the same definition, including "A worker (*pekerja*)/laborer (*buruh*) is every person who works for a wage or other forms of remuneration" (Article 1 number 3 of Act concerning Manpower). We account for such cases by considering

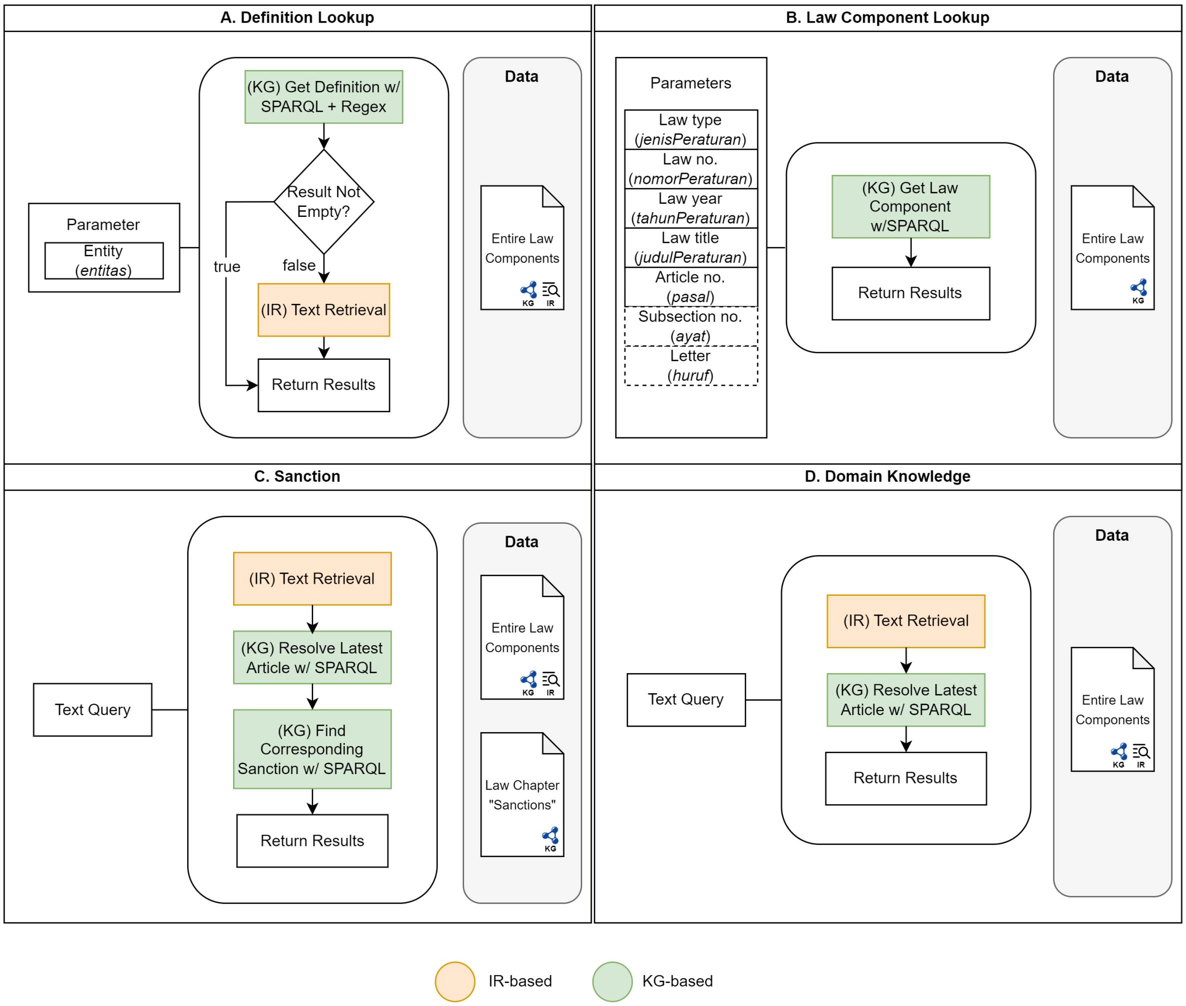

Figure 4. Intent Fulfillments for Hybrid Legal VA

the slash character (/) as a token separating entities that have similar meaning.

In case the SPARQL query does not provide any results, the system utilizes text-based retrieval as an alternative method. Likewise, the text-based retrieval specifically searches within the "General Definitions" ("*Ketentuan Umum*") chapter. The query incorporates the entity parameter, followed by the word "*adalah*" (which has a close meaning to "is/are").

5.2. Law Component Lookup

This intent exclusively utilizes the knowledge graph retrieval method due to its structuredness. With the assumption that the query parameters are adequately provided, the process is relatively straightforward. Since all the law components are interconnected in the KG data, constructing SPARQL queries with the given parameters allows us to obtain the component of the law with high precision. There is also another need to obtain the law component that is the most recent, which can be addressed by utilizing queries in SPARQL once again.

Law Component Lookup intent involves extracting multiple types of parameter, ranging from law type (*jenisPeraturan*) to article (*pasal*). Additionally, subsection (*ayat*) and letter (*huruf*) are considered as extra parameters that depend on the desired level of specificity for the answer. On the other hand, law title (*judulPeraturan*) can be used in exchange to the year (*tahunPeraturan*) and number (*nomorPeraturan*), for instance, the Act on Manpower is equivalent to Act Number 13 Year 2003.

Moreover, this intent sets a limitation on the level of granularity, ranging from a letter (*huruf*) or a subsection (*ayat*) up to an article (*pasal*). This restriction is in place because most VA interfaces are not designed to respond with long text passages (i.e., law chapters). The SPARQL queries associated with this intent also incorporate a feature to retrieve the up-to-date article. By default, if an amendment is available for the given answer, then that version is provided to the user. In addition to that, if the amendment entails the deletion or cancellation of the original law component, the user is informed that the law component cannot be found.

5.3. Sanction

The Sanction intent represents a key example of the hybrid approach combining IR and KGs. The IR aspect enables the identification of the specific law component that has been violated. Leveraging the advantages of linked data, the KG component facilitates the discovery of the corresponding sanction associated with the violated law.

The initial step in fulfilling the intent is to determine the law component that has been violated. To achieve this, we utilize the indexed text documents of the law and perform text-based retrieval. This process involves identifying the law components with the highest match and retrieving their corresponding identifiers (URIs). Subsequently, we conduct a verification process to check for any changes to these components. The answer is updated accordingly based on the latest version, ensuring that the punishments are applicable to the presently effective article. Once the violated component is resolved to the currently effective version, the next stage is to obtain the associated punishments. This is accomplished by leveraging the "refers to" link ("*merujuk*") that points to components within the "Sanctions" chapter, allowing us to obtain the correct sanctions.

5.4. Domain Knowledge

This last intent is the broadest one compared to others and heavily relies on the utilization of information retrieval techniques. However, it also incorporates the use of knowledge graph queries to reconcile the obtained components with the currently effective law. The process begins by inputting the question into the information retrieval module, which returns the most relevant documents based on the input. Subsequently, the component that ranks the highest is employed in the knowledge graph querying step to obtain the most recent version of the component.

In certain scenarios, law components may undergo reordering in amendments, leading to cases where the system fails to find the latest version. An example of this is the relocation of the contents of Article 48 of Act Number 13 Year 2003 to Article 45 subsection (1) letter c following the updates in Act Number 11 Year 2020. To address this issue, we implement a solution by retrieving multiple top results from the IR method. If the best result is determined to have been reordered by a subsequent amendment, the system proceeds to utilize the next best result until it successfully retrieves the currently effective law version.

6. Evaluation

This section focuses on evaluating the performance of the prototype and analyzing the results. We first explain on the benchmark dataset and the evaluation metric that we use to evaluate our system. Afterwards, we elaborate on the results which is divided into a few stages.

We generate a dataset comprising 201 test cases for evaluating the system's performance.[10] To ensure impartiality, the individuals that contributed in generating the dataset are not involved in programming the VA. These cases are entirely separate from the training phrases used to develop the intent classification and parameter extraction modules. When provided with a question, the system is evaluated based on its ability to accurately return the corresponding law component by verifying its URI (identifier). The comparison between the expected and the actual answer is evaluated using the accuracy (Acc.) of the law component reference (e.g., article no., subsection no.), formulated as follows:

$$\texttt{Acc.} = \frac{\texttt{\# of Correctly-Answered Questions}}{\texttt{\# of Questions}}$$

A similar evaluation metric, called exact match, is also leveraged by [9] in evaluating LAW-U's performance. A question is considered to be correctly answered only if the expected answer and the actual answer are the exact match.

To conduct a comprehensive evaluation of the system, we assess its performance in various specific areas to identify potential issues. Our evaluation consists of three key assessments: the evaluation of clean parameters, the evaluation of intent classification, and the end-to-end evaluation. Each assessment focuses on different aspects of the system's functionality and provides valuable insights into its performance.

[10] https://s.id/IDLaborLawVATestCases

Table 4. Clean Parameters Evaluation Results

#	Intent	Accuracy(%)
1	Definition Lookup	100.00
2	Law Component Lookup	100.00
3	Sanction	100.00

6.1. Clean Parameters Evaluation

During this evaluation process, we make the assumption that both the parameter extraction and classification of intents have been performed flawlessly. Our dataset already includes the correct parametersand the correct intent as the “clean” inputs for this evaluation purposes. This approach enables us to assess the system’s performance based on the anticipated correct inputs. For each test case, the parameters are fed into a SPARQL query template according to which intent the test case belongs. The resulting SPARQL query is then used to retrieve the law component. However, there is an exception for definition lookup, in which if there exists a case where the KG querying fails to yield any results, the IR comes into play by retrieving the most relevant component in the chapter of “General Definitions” (as elaborated in Subsection 5.1).

Next, we conduct a comparison between the retrieved URIs of law components and the expected URIs as they serve as references for the respective law components. This evaluation specifically targets three use cases that involve parameter extraction: Definition Lookup, Law Component Lookup, and Sanction. Specifically for the latter intent, the “clean” input consists of the identifier for the article/subsection that has been sanctioned, while the expected output is the identifier of the corresponding sanctions for that specific article/subsection. Aside from those intents, we do not include the domain knowledge in this evaluation since it does not require parameter extraction.

Through the utilization of the legal KG, we successfully attain a perfect accuracy rate of 100% in three of the intents when provided with accurate parameters and the correct intent for fulfillment, as demonstrated in Table 4. The fulfillment process in the clean parameters evaluation is nearly-deterministic. For most parts, the process can be resolved by a deterministic approach, i.e., KG querying using SPARQL. Out of 150 test cases of the three intents, only one definition lookup case that requires the fulfillment to fall back to using IR, due to an exceptional difference in the written structure. Nevertheless, these findings largely confirm that the KG part of our system is capable of obtaining the right components.

6.2. Intent Classification Evaluation

Next, we assess how well the system can distinguish the different intents. Each test case is fed into the intent classification module to obtain the predicted intent, which is then compared with the true intent. Our metrics for classification includes: accuracy, precision, recall, and F1 score. Apart from accuracy, these metrics are evaluated individually for each intent to provide a comprehensive analysis.

Table 5. Intent Classification Evaluation Results

#	Intent	Prec.(%)	Rec.(%)	F1 (%)
1	Definition Lookup	69.23	90.00	78.26
2	Law Component Lookup	98.36	100.00	99.17
3	Sanction	97.83	92.00	94.84
4	Domain Knowledge	91.89	66.67	77.27
				Accuracy: 87.56%

In our most recent prototype, the intent classification achieves an accuracy score of 87.56%, indicating the percentage of test instances where the intents are correctly guessed as shown in Table 5 . Notably, law component lookup and sanction intents exhibit scores that surpass 90% in all aspects. Those types are comparatively straightforward to differentiate syntactically among the four intents.

However, in the case of the Definition Lookup and Domain Knowledge intents, the module falls short of achieving F1 scores above 90%. These two intents pose a challenge as they are syntactically similar, with the distinction lying in their semantic meaning. To address this issue, we have made recent efforts to enhance these scores. We introduced new types of parameters, such as domain knowledge phrases and definition phrases, to establish a syntactic differentiation. Additionally, we have experimented with a variety of training phrases sets for the two intents.

6.3. End-to-End Evaluation

To comprehensively assess the performance of our virtual assistant, we conducted a holistic evaluation. It encompasses all of the steps, from classifying intents and extracting parameters to providing the relevant answer. By evaluating the system as a whole, we gain insights into its overall capability to effectively address user queries and function as a reliable virtual assistant.

When considering accuracy alone, our system achieves an accuracy rate of 60.20% for all test cases, as indicated in Table 6. With the exception of the Domain Knowledge intent, each intent can be

Table 6. End-to-End Evaluation Results

#	Intent	Lex2KG-VA Accuracy (%)	Baseline IR Accuracy (%)
1	Definition Lookup	**60.00**	7.50
2	Law Component Lookup	**69.49**	0.00*
3	Sanction	**68.00**	0.00*
4	Domain Knowledge	42.31	**46.15**†
	All Test Cases	**60.20**	13.43

*KG structures and links are required for resolving these intents, and relying solely on IR methods is insufficient to address them.;
†The baseline model only relies on IR (exactly how Domain Knowledge questions are answered), which is why it appears to have a higher score.

predicted accurately ($\geq 60\%$). As for the excluded intent, it is challenging to handle due to the breadth of the intent itself. However, if the task is to be treated as an information retrieval task where the correct document needs to rank as the first result (i.e., Hit@1), then the score of 42.31% can be considered satisfactory. Most IR tasks evaluate multiple top results, but we only focus on the top-1 result as showing more than that may not align with typical virtual assistant interfaces. Based on these evaluations, we can conclude that our system is capable of answering most of the inquiries regarding law retrieval.

We have also conducted a comparison between our system and the baseline, that is an information retrieval system, as presented in Table 6. It relies solely on text retrieval to respond to queries. When it comes to looking for concrete definitions, it often fails to obtain the component within "General Definitions" chapter. Similarly, when attempting to answer questions related to law components and sanctions, the baseline IR system also falls short. There are two main reasons for these failures. Firstly, the IR system is unable to handle the structured questions involved in Law Component Lookup. Secondly, the baseline IR approach lacks the necessary metadata to access the punishments associated with violated articles or subsections.

In contrast, the baseline system achieves a slightly higher score for Domain Knowledge (46.15% vs. 42.31%) because it relies on a pure IR approach to answer questions, which aligns with the fulfillment of Domain Knowledge queries. By avoiding the classification errors made by the intent classifier, the baseline IR system performs better in this specific intent compared to Lex2KG-VA. However, in the bigger picture, the scores for all test cases show that the baseline IR is still generally not capable enough to perform as a legal VA. Some of these test cases (i.e., sanctions, law component lookup) require precise answers and explicit links in which using only pure IR would face severe obstacles. In other words, our system performs significantly better in general as a legal VA, that is, 60.20% vs. 13.43% of overall accuracy comparison between our Lex2KG-VA vs. baseline IR.

As a preliminary comparison, we briefly assess the system against Google search since we have not found a similar publicly available system for Indonesian labor laws. Our system surpasses Google search results, particularly in intents supported by the knowledge graph, while also demonstrating competitiveness in answering queries related to domain knowledge. When presented with the question "What is the content of Article 7 subsection (2) of Act No. 13 Year 2003?" ("*Apa isi Pasal 7 ayat (2) UU 13/2003?*"), our virtual assistant manages to obtain the correct response, whereas Google search ranks the same component below top-1.

7. Discussions

Now we discuss several aspects encompassing hybrid legal VAs: combining KGs and IR, industrial use cases, legal documents conversion to KGs, legal KG metadata and granularity, and improvements of the legal VA.

Combining KGs and IR. Our hybrid KG-IR approach utilizes the benefits of both techniques. KGs support not only the law data representation, but also in precisely retrieving the required law components and are capable of inferring the semantics of amendments and law references. On the other hand, the IR technique allows the VA to perform close to a search engine and retrieves relevant results to user questions with better flexibility. Such an approach can be applied to other legal KGs to deliver a more intuitive user experience in accessing legislation contents.

Industrial Use Cases. Many corporations, particularly the human resources teams, rely on the Act concerning Manpower (UU 13/2003) and its associated regulations to ensure compliance with the current legislation. For future advancements, incorporating other relevant laws (e.g., Tax laws) can further enhance the breadth of the legal virtual assistant's coverage. Additionally, to cater to the public's requirements, the legal virtual assistant can explore other domains of law, such as marriage, inheritance, and land. It is worth noting that the Indonesian legal system follows a hierarchical structure and adheres to the principle of "*lex specialis derogat legi generali*",[11] indicating the existence of implementing regulations for higher-order laws. Including these

[11]https://www.trans-lex.org/910000/_/lex-specialis-principle/

implementing regulations can deepen the system's scope.

Another related point from the industry perspective is that laws are updated over time. Such a case very recently occurred when Government Regulation in Lieu of Law Number 2 of 2022 (*Perpu 2/2022*) concerning Job Creation was promulgated on December 30, 2022.[12] This would make an interesting case study on how legal VAs should stay up to date with the latest laws and regulations.

In this study, we also present a set of intents as use cases: definition lookup, law component lookup, sanction, and domain knowledge. This set of intents was proposed after going through a brainstorming process and incorporating business requirements as conveyed by the project stakeholders. Although this is not an exhaustive list, we believe these intents represent typical questions from the users.

Legal Documents Conversion to KGs. In preparation for the expansion of law documents, we have undertaken efforts to gather and analyze a comprehensive collection of legal materials in Indonesia. As of June 2022, this collection comprises 4,746 government regulations (*Peraturan Pemerintah*) and 1,717 laws (*Undang-Undang*). Through parsing these documents, we were able to convert 3,864 government regulations and 1,353 laws into knowledge graphs, accounting for a total of $\sim$80.72% of the collected materials. Theoretically, adding more laws and regulations to the VA system would benefit from the topic coverage standpoint. However, it is also worth noting that adding more laws would also mean reduced accuracy since the system has to pick from a larger set of documents. Fortunately, our system is designed with minimum needs for refactoring when applying it to different sets of regulations since nearly all the parameters and intents apply globally to the law schematics in Indonesia. The schematics, as ruled in Act 12/2011, contains strict guidelines in drafting law documents which result in the exact same structures for all Indonesian laws. On the other hand, we are already capable of handling amendments in the VA which is one of the most difficult challenges in developing technologies for legislation.

Legal KG Metadata and Granularity. We sought the insights of a legal practitioner regarding our system, and her feedback has been valuable. According to her, the inclusion of law metadata, such as chapter titles and references, proves highly beneficial when conducting comprehensive searches on specific subjects. Additionally, she commends the effectiveness of KGs in offering structured linked data for law documents. With this linked data representation, she can easily trace law references and access the complete context of topics discussed within the law documents. Her endorsement underscores the system's ability to support legal practitioners in their research and analysis.

There is an ongoing concern regarding the level of granularity in legal KGs. Our implementation primarily focuses on article and subsection references. However, for more effective semantic search capabilities, it would be worthwhile to explore fine-grained ontologies that encompass legal entities, as exemplified by the QAnswer KG [10] and the system developed by Fawei et al. [18]. Nevertheless, it is important to note that building a highly granular ontology increases the risk of false ontology extraction, which can pose challenges within the legal domain. Careful consideration and validation are necessary to address this potential issue.

Improvements for the Legal VA. In addition to adopting a fine-grained ontology, the integration of language models offers another approach to explore semantic search functionalities. Legal documents contain abundant textual information that can be leveraged for training language models. Existing language models tailored for the legal domain, like LEGAL-BERT [19] trained on European legislative documents, can serve as valuable references for further study. It is worth noting that, as of now, there is a lack of publications on large language models specifically designed for Indonesian laws, highlighting an area for future research.

Regarding the range of question types, our proposed four intents generally encompass the majority of cases. Nonetheless, future endeavors could involve additional customized intents. An instance of this would be an "obligations lookup" intent. This intent would entail the system searching for multiple points of obligations across various articles and subsections.

For future improvements in Lex2KG-VA, we have identified potential technical enhancements. Firstly, it would be beneficial to enable the system to retrieve older versions of law components upon request for the Law Component Lookup intent. Currently, the system only provides the most recent in-force version. The second area of improvement pertains to parameter extraction, as there are instances of incorrect extractions. For example, instead of correctly extracting the entity "*informasi ketenagakerjaan*" ("manpower information"), the module only brings out "*ketenagakerjaan*" ("manpower"). Lastly, as observed during evaluating the "clean" input performance and the end-to-end score, the model still struggles to accurately answer Domain

[12]https://setkab.go.id/pemerintah-terbitkan-perppu-cipta-kerja/

Knowledge questions. To address this, future implementations may explore semantic-based searching utilizing techniques such as text embeddings and other NLP methods.

8. Conclusions

We have developed a legal virtual assistant (VA) system using a hybrid approach that combines information retrieval (IR) and knowledge graphs (KGs). Our system is the first of its kind for Indonesian laws, capable of addressing various common legal-related inquiries by leveraging the strengths of both IR and KGs. Through its implementation, we have demonstrated the system's effectiveness in providing accurate and precise answers, as well as handling typical questions encountered in a legal VA. The prototype system successfully handles a wide range of user queries related to two of the Indonesian laws: Act concerning Manpower (UU 13/2003) and its updates in the Act concerning the Creation of Jobs (UU 11/2020).

According to our assessments, the knowledge graphs have demonstrated the ability to provide accurate answers when supplied with the appropriate parameters. Additionally, the intent classification module has achieved an accuracy rate of over 87% in correctly classifying various types of user queries. Lastly, through the end-to-end evaluation, our system has exhibited a significant improvement in accuracy compared to the baseline IR system, successfully answering a substantial number of user questions across all test cases (over 40% improvement in accuracy).

We have discussed several core aspects encompassing the legal VA, some of which are industrial use cases and improvements for the legal VA. In general, delivering law knowledge as a service poses significant challenges, but it also presents opportunities for further research and development.

Acknowledgment

The authors express their gratitude to the previous contributors, namely Muhamad Abdurahman, Gibran Brahmanta, Nanda Ryaas Absar, and Ardhani Dzaky Ralfiano for their valuable contributions in the development of Lex2KG and the hybrid legal virtual assistant.

References

[1] H. Zhong, C. Xiao, C. Tu, T. Zhang, Z. Liu, and M. Sun, "JEC-QA: A Legal-Domain Question Answering Dataset," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, 2020, pp. 9701–9708.

[2] R. Dale, "Law and Word Order: NLP in Legal Tech," *Natural Language Engineering*, vol. 25, no. 1, pp. 211–217, 2019.

[3] N. Noy, Y. Gao, A. Jain, A. Narayanan, A. Patterson, and J. Taylor, "Industry-scale knowledge graphs: lessons and challenges," *Communications of the ACM*, vol. 62, no. 8, pp. 36–43, 2019.

[4] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. de Melo, C. Gutiérrez, S. Kirrane, J. E. Labra Gayo, R. Navigli, S. Neumaier, A.-C. Ngonga Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. F. Sequeda, S. Staab, and A. Zimmermann, *Knowledge Graphs*, ser. Synthesis Lectures on Data, Semantics, and Knowledge. Morgan & Claypool, 2021, no. 22. [Online]. Available: https://kgbook.org/

[5] J. M. Schneider, G. Rehm, E. Montiel-Ponsoda, V. Rodríguez-Doncel, P. Martín-Chozas, M. Navas-Loro, M. Kaltenböck, A. Revenko, S. Karampatakis, C. Sageder *et al.*, "Lynx: A knowledge-based AI service platform for content processing, enrichment and analysis for the legal domain," *Information Systems*, vol. 106, p. 101966, 2022.

[6] M. Abdurahman, F. Darari, H. Lesmana, M. Hartopo, I. Rhesa, and B. C. L. Tobing, "Lex2KG: Automatic Conversion of Legal Documents to Knowledge Graph," in *2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. IEEE, 2021, pp. 1–5.

[7] A. Hamid and Hasbullah, "Legal Hermeneutics of the Omnibus Law on Jobs Creation: A Case Study in Indonesia," *Beijing Law Review*, vol. 13, no. 3, pp. 449–476, 2022.

[8] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008.

[9] V. Socatiyanurak, N. Klangpornkun, A. Munthuli, P. Phienphanich, L. Kovudhikulrungsri, N. Saksakulkunakorn, P. Chairaungsri, and C. Tantibundhit, "LAW-U: Legal Guidance Through Artificial Intelligence Chatbot for Sexual Violence Victims and Survivors," *IEEE Access*, vol. 9, pp. 131 440–131 461, 2021.

[10] D. Diefenbach, J. Giménez-García, A. Both, K. Singh, and P. Maret, "QAnswer KG: Designing a portable question answering system over RDF data," in *European Semantic Web Conference*. Springer, 2020, pp. 429–445.

[11] N. S. Muninggar and A. A. Krisnadhi, "LexID:

The Metadata and Semantic Knowledge Graph Construction of Indonesian Legal Document," *Jurnal Ilmu Komputer dan Informasi*, vol. 16, no. 1, pp. 15–46, 2023.

[12] M. Vegupatti, M. Nickles, B. R. Chakravarthi, L. Longo, L. Rizzo, E. Hunter, and A. Pakrashi, "Simple Question Answering Over a Domain-Specific Knowledge Graph using BERT by Transfer Learning," in *AICS*, 2020, pp. 289–300.

[13] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 2019, pp. 4171–4186.

[14] V. Rodríguez-Doncel and E. Montiel-Ponsoda, "Lynx: Towards a legal knowledge graph for multilingual Europe," *Law in Context*, vol. 37, no. 1, pp. 175–178, 2020.

[15] E. Q. Nuranti and E. Yulianti, "Legal Entity Recognition in Indonesian Court Decision Documents Using Bi-LSTM and CRF Approaches," in *2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. IEEE, 2020, pp. 429–434.

[16] Z. Yang, P. Qi, S. Zhang, Y. Bengio, W. Cohen, R. Salakhutdinov, and C. D. Manning, "HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 2369–2380.

[17] S. E. Robertson, S. Walker, S. Jones, M. M. Hancock-Beaulieu, M. Gatford *et al.*, "Okapi at TREC-3," *Nist Special Publication Sp*, vol. 109, p. 109, 1995.

[18] B. Fawei, J. Z. Pan, M. Kollingbaum, and A. Z. Wyner, "A Semi-automated Ontology Construction for Legal Question Answering," *New Generation Computing*, vol. 37, no. 4, pp. 453–478, 2019.

[19] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, and I. Androutsopoulos, "LEGAL-BERT: The Muppets straight out of Law School," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020, pp. 2898–2904.