

A Framework for Legal Information Retrieval based on Ontology

Dae Woong Jo *, Myung Ho Kim **

Abstract

Professional knowledge such as legal information is commonly not accessible or cannot be easily understood by the public. By using the legal ontology which is previously established, the legal information retrieval based on ontology is to use for the information retrieval. In this paper, we propose the matters required for the design and develop of the framework for the legal information retrieval based on ontology. The framework is composed of the query conversion engine of SPARQL base for query to OWL ontology and user query type engine and return value refinement engine and web interface engine. The framework does the role as the infrastructure which retrieval the legal ontology effectually and which it serves and can be used in the semantic legal information retrieval service.

▶ Keyword : Legislation, Ontology, Legal Information Retrieval, Semantic Web, Ontology Visualization

I. Introduction

법령은 전문지식 영역에 해당하는 분야로 전문 지식이 없는 일반 사람들이 법령 정보를 검색, 활용하기에는 한계가 있다. 기존의 법령 정보 검색은 법률과 관계된 정확한 키워드 및 법률 번호 등을 이용해서 검색을 하는 방식이다[2, 3]. 정보 검색에 필요한 전문지식이 없는 일반 사람들은 원하는 결과를 찾는데 시간과 노력이 소모된다. 온톨로지는 하나의 개념을 설명하기 위한 명시적이고, 정형화된 명세서로 정의[4]할 수 있으며, 그래프 기반의 자료구조를 통해 연쇄적인 정보 탐색이 가능하다. 따라서 법령 정보를 OWL과 같은 온톨로지 형태로 변환 구축하고, 구축된 법령 온톨로지를 질의하고, 결과를 반환할 수 있다면 기존의 정보 검색을 통한 방식과는 다른 형태의 결과를 도출 할 수 있다.

본 논문은 법령 온톨로지에 질의를 하고, 사용자 중심의 결과를 반환할 수 있는 온톨로지 기반의 법령 정보 검색 프레임

워크를 제안한다. 제안하는 프레임워크는 기존의 정보 검색 엔진에서 수행하던 RDB에 질의해서 정보를 반환하는 형태가 아니라 OWL 온톨로지에 특정 질의어인 SPARQL(SPARQL Protocol and RDF Query Language)[5]을 이용해서 정보를 탐색하는 형태로 검색을 수행한다. 검색 엔진은 OWL 온톨로지에 질의하기 위한 SPARQL 기반의 질의 변환 엔진과 사용자 질의 유형을 검사하기 위한 유형 검사 엔진, 결과 값 정제 엔진으로 구성된다.

본 논문은 온톨로지 기반의 법령 정보 검색을 위해 대한민국 법령의 구조를 바탕으로 법령 온톨로지를 구축하고, 구축된 온톨로지를 기반으로 법령 정보 검색을 한다. 정보 검색은 OWL 온톨로지의 특성을 반영하여 그래프 기반의 자료구조를 탐색할 수 있는 방법 및 온톨로지 기반의 법령 검색을 위한 세 가지 검색 방법을 제안하며, 제안된 검색 방식을 프레임워크 내에서 엔진으로 구축한다. 또한, 법령 정보를 요약된 형태로 표현하고, 그래프 형태의 온톨로지를 표현하기 위한 시각화 기술이 적용

• First Author: Dae Woong Jo. Corresponding Author: Myung Ho Kim

*Dae Woong Jo (jodw@ssu.ac.kr), School of Software, Soongsil University

**Myung Ho Kim (kmh@ssu.ac.kr), School of Software, Soongsil University

• Received: 2015. 06. 10, Revised: 2015. 07. 21, Accepted: 2015. 08. 04.

• This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2014R1A1A2058695).

• This research summarizes the thesis[1].

된 웹 인터페이스도 프레임워크 내에서 동작 가능하도록 한다. 이와 같은 온톨로지 기반의 법령 정보 검색 프레임워크는 기존의 텍스트로 나열되던 법령 정보를 시각화하고, 기존보다 효과적으로 법령 정보를 검색할 수 있는 검색 방법을 제안하여, 시맨틱 법령 정보 검색 서비스로 활용될 수 있다.

II. Related Works

1. The Difference between OWL and RDB

OWL은 개념 간에 연결된 그래프 기반의 자료구조를 바탕으로 정보들이 저장되고, RDB는 테이블 기반으로 데이터를 처리한다. 표 1은 OWL과 RDB의 주요 특징에 대해 비교 정리한 것이다.

Table 1. Comparison of OWL and RDB

	OWL	RDB
모델	온톨로지	관계형 모델
주요 자료구조	그래프	테이블
질의 언어	SPARQL	SQL
결과 반환 형태	OWA	CWA

OWL 온톨로지 질의를 위해 W3C에서는 SPARQL이라는 표준 질의 언어를 정의하고, 질의를 위한 문법적인 제약을 표준화하고 있다. SPARQL은 기본적으로 그래프 패턴 매칭(Graph Pattern Matching)에 따라 결과를 도출한다. SQL 문법과 비슷한 형식의 SPARQL은 SELECT, WHERE 절에 해당 트리플을 기술하여 원하는 결과 값을 얻어낸다.

OWL과 RDB는 질의 언어뿐만 아니라 결과를 반환하는 형태도 다른 양상을 나타낸다. OWL은 OWA(Open World Assumption)로 추후, 지식이 있을 것으로 예상하고 결과가 없더라도 오류를 반환하지 않는다. 하지만 RDB는 CWA(Closed World Assumption)로 질의 후, 결과가 없을 시에는 오류를 반환한다[6]. 이와 같은 차이점은 온톨로지 기반의 정보 검색 프레임워크를 구축하는데 기존과는 다른 형태로 접근해야 함을 뜻한다.

2. Method of SPARQL Query

SPARQL 질의는 SELECT 절에 찾고자 하는 값을 X, Y1, Y2, Y3과 같은 변수로 명시하고 WHERE 절에 해당 변수를 찾을 트리플(주어, 술어, 목적어)을 나열하여 각각의 관계에 의해 원하는 X, Y1, Y2, Y3 값을 질의하여 결과 값을 찾도록 한다. 그림 1은 이와 같은 SPARQL 질의 예시를 나타낸 것이다.

SPARQL 문법 내에는 복잡한 그래프 패턴 질의를 위해 다양한 방법의 패턴 매칭 방법을 정의하고 있다. 대표적인 것이 BGP(Basic Graph Patterns)에 해당하는 패턴이고, BGP 패턴은 WHERE 절에 주어진 트리플 기반의 조건이 반드시 매칭이

되어야 하는 형태이다. BGP 패턴에 의한 질의 방법이 SPARQL 질의 문 작성 시 일반적인 형태가 된다.

```
SELECT ?X ?Y1 ?Y2 ?Y3
WHERE {
    ?X rdf:type decree:Act .
    ?X decree:Act_Name ?Y1.
    ?X decree:Enforcement_Date ?Y2.
    ?X decree:Date_Of_Promulgation ?Y3.
}
```

Fig. 1. Example of SPARQL Query

그 외, GGP(Group Graph Pattern), OPG(Optional Graph patterns), AGP(Alternative Graph Pattern), PNG(Patterns on Names Graphs)에 해당하는 질의 패턴들이 존재한다[7]. 다양한 질의를 처리하기 위해선 패턴에서 정의된 방법과 일치되도록 SPARQL을 사용자 키워드를 기반으로 변환, 매칭하기 위한 방법이 필요하다.

3. Semantic Web Framework

다양한 오픈소스 진영과 시맨틱 웹 연구자들에 의해 이론적으로 가능했던 기술들이 점차 많은 사람들이 이용 가능한 수준으로 발전되고 있다. 표 2는 RDF 트리플 및 OWL 온톨로지를 파싱, 질의, 추론 등의 일을 처리할 수 있는 기존의 시맨틱 웹 프레임워크의 종류와 특징들을 정리한 것이다[8].

Table 2. Types of Semantic Web Framework

이름	특징	언어
Jena	RDF, RDFS, OWL, SPARQL를 위한 API, 룰 기반 추론엔진, 트리플 저장소	Java
OWL API	OWL 기반의 프로그래밍 인터페이스	Java
Sesame	RDFS 추론 및 질의, RDF 트리플 저장소	Java, Python
KAON2	OWL DL, SWRL, F-Logic 프레임워크, SPARQL 질의 처리	Java
FaCT++	OWL DL 추론엔진	C++
Pellet	OWL DL 추론엔진, Jena 및 OWL API에서 사용가능	Java

온톨로지 처리 방법은 기존의 RDF 트리플 기반의 처리 방식에서 OWL 온톨로지 처리가 가능한 수준의 프레임워크까지 개발이 이루어지고 있고, 온톨로지 저장소 및 추론엔진에 이르기까지 다양한 언어로 개발되어 있다. 본 논문에서는 이러한 시맨틱 웹의 기술적 성숙도를 이용하여 법령 시맨틱 검색 서비스가 가능한 프레임워크를 설계 및 구현 한다.

4. Legal Semantic Web Retrieval Services

한국법제연구원에서는 그림 2와 같은 시소러스(thesaurus) 검색 서비스를 하고 있다[9]. 온톨로지를 이용한 시맨틱 검색 서비스를 위한 준비 형태로 법률에서 쓰이는 검색 용어집을 시소러스 형태로 구축해서 서비스 하고 있다.

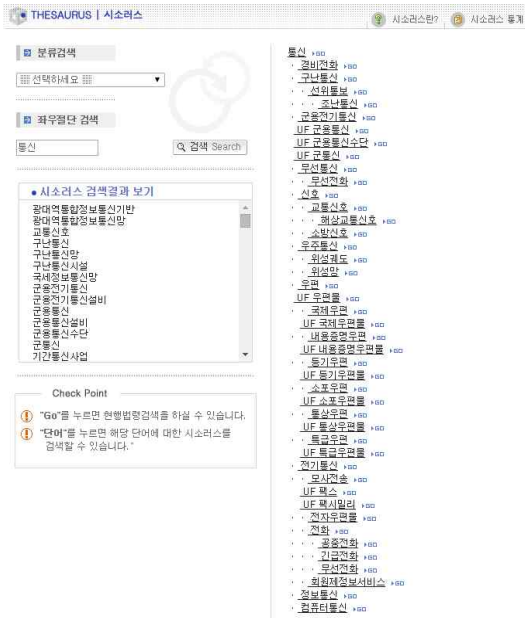


Fig. 2. Thesaurus Search Screen of Korea Legislation Research Institute

현행 법령의 주제가 되는 중요한 용어를 모아 등가관계, 계층관계, 연관관계, 참조관계 등으로 체계화 하였고, 일반 사람들이 쓰는 용어와의 연관관계 구축을 통해 원하는 법령 정보를 찾을 수 있도록 한다. 하지만 실제 사용되는 서비스는 용어 간의 관계들은 구축되어 있으나 해당 용어와 관계된 법령과는 아직 연계가 되어 있지 않은 모습이다.

그림 2의 시소러스 검색 화면은 ‘통신’이라는 검색어를 입력 시 ‘통신’과 관계된 용어의 계층 관계가 표현되고 해당 용어의 우선어, 비우선어, 최상위어, 상위어 등이 약어로 설명되고 있다. 그림에서 해당 용어 옆의 검색을 클릭 시에는 관련 법령 정보와 연결되어 법령 정보를 볼 수 있다고 하지만 아직 다른 연관 서비스는 되고 있지 않다.



Fig. 3. Screen of iLaw Service

법무부에서는 입법지원시스템을 위한 iLaw[10]를 개발하여 서비스 하고 있다. iLaw는 법령 시맨틱 검색을 위한 검색 모델을 제시하고 있으며 기존의 법령 자원을 시각화된 형태로 보여 주기 위한 웹 인터페이스가 강화된 검색 서비스이다. 그림 3은 iLaw 서비스 화면을 나타내고 있다. 검색 결과는 기존의 법제

처에서 하는 서비스와 같은 형태이나 차트나, 그래프 등으로 표현하여 해당 용어의 빈도와 법령의 수가 시각화 되어 표현되고 있다.

서비스는 국내 법령 어휘와 다국적 어휘와의 매핑 온톨로지를 구축하고, 같은 뜻을 가진 용어를 그래프 형태로 표현한다 [11]. 특징으로는 법률 용어와 관련된 논문이나 미디어 결과를 함께 보여주고 있다. 하지만 법령 자원간의 시맨틱 질의 및 온톨로지의 추론을 이용한 검색 방식은 지원하지 않고 있다.

III. The Proposed Scheme

1. Legal Ontology

본 논문에서 법령 정보 검색에 사용될 온톨로지는 법률 스키마 온톨로지와 법률 영역 온톨로지다. 법률 스키마 온톨로지는 법률과 관련된 메타 정보들이 들어간 온톨로지이다. 즉, 법령, 조문, 판례, 연혁과 같은 법률 개념들을 온톨로지로 만드는 것이며, 영역 온톨로지는 하나의 법 영역과 관련된 요소들을 개념화해서 표현한 온톨로지로 구체적인 법률의 형태를 트리플 기반으로 구축한다.

1.1 Legal Schema Ontology

법률 스키마 온톨로지의 관계는 그림 4와 같은 그래프 형태로 표현 가능하다. 각각의 원은 클래스를 나타내며, 화살표 방향이 아래를 향하는 원은 하위 클래스를 나타낸다. 점선으로 된 네모 박스는 데이터타입 프로퍼티(Datatype Properties) 관계를 나타내고 있다.

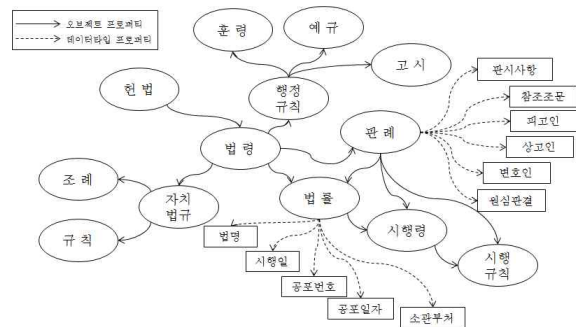


Fig. 4. Legal Schema Ontology Graph

헌법 클래스 아래 법령 클래스를 중심으로 법률, 판례, 행정규칙, 자치법규 클래스가 오브젝트 프로퍼티(Object Properties) 관계로 클래스 인스턴스 간 연결을 하며 표 3은 정의역(Domain), 공역(Range) 간의 오브젝트 프로퍼티 관계를 나타낸다. 법률 아래 시행령과 시행규칙이 IS-A 관계의 rdfs:subClassof 형태로 상위 개념을 상속 받고 있다. 이와 같은 관계기술은 [12]의 법령의 구조 및 관계 매핑에 관한 연구를 기반으로 한다.

Table 3. Object Properties Relationship of Legal Schema Ontology

#	정의역	술어	공역
1	법령	hasAct	법률
2	법령	hasPrecedent	판례
3	법령	hasAdministrativeRule	행정규칙
4	법령	hasSelfGovernment	자치법규
5	판례	relatedAct	법률
6	판례	relatedEnforcementDecree	시행령
7	판례	relatedEnforcementRule	시행규칙

1.2 Legal Domain Ontology

법률 영역 온톨로지의 관계는 그림 5와 같은 그래프로 나타낼 수 있다. 법률 스키마 온톨로지서 법령 클래스는 법률 영역 온톨로지의 법령 계층 클래스와 관계있는 클래스로 구축한다. 법령계층 클래스는 크게 2개의 상위 클래스를 가지고 있다. 규정 클래스와 편 클래스인데 규정 클래스는 법률 문서로부터 추출 가능한 규정들의 상위 클래스로 하위 클래스로는 목적, 정의, 계획, 위원회, 벌칙, 과태료, 시행일과 같은 법률로부터 공통적으로 나타나는 7가지 규정들을 클래스로 표현한다. 각 클래스의 인스턴스는 법률 조문으로부터 해당 규정 목적에 맞는 트리플을 주어, 술어, 목적어 형태로 추출, 구축 한다[12].

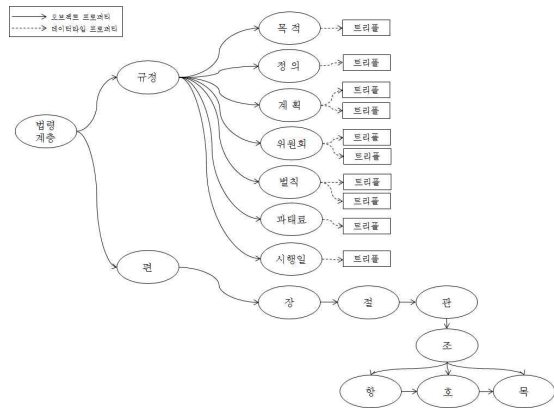


Fig. 5. Legal Domain Ontology Graph

편 클래스는 법률 문서 계층의 가장 최상위 위치에 있는 항목으로서 편 클래스 하위로 법률 문서 체계에 맞는 항목들이 장, 절, 관, 조, 항, 호, 목 단계의 클래스로 표현된다. 각각의 클래스에는 해당 항목과 관련된 항목정보 인스턴스가 매핑 된다.

2. Legal Information Retrieval Engine

본 논문에서 제안하는 법령 정보 검색 엔진은 사용자 질의 처리 엔진, 유형 검사 엔진, 결과 값 정제 엔진, SPARQL 질의 변환 엔진, SPARQL 질의 처리 엔진, 추론 엔진으로 구성된다. 그림 6은 법령 정보 검색 엔진의 구조를 나타낸 것이다.

법령 정보 검색 엔진은 사용자의 요청을 처리하고, 원하는

결과를 반환하는 역할을 수행 한다. 각 엔진은 사용자의 요청을 분석하고, 분석된 결과를 바탕으로 OWL 온톨로지에 질의하기 위한 SPARQL 질의 문 자동 변환과 질의 결과의 IRI(Internationalized resource identifier) 정보를 제거한 형태의 키워드만을 리턴 하는 결과 값 정제 작업을 수행한다.

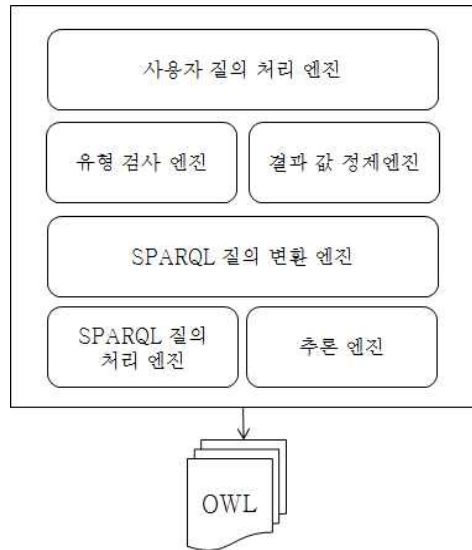


Fig. 6. The Structure of the Legal Information Retrieval Engine

법령 정보 검색 엔진의 구성 요소 중 사용자 질의 처리 엔진에서는 사용자로부터 받아온 요청 값을 분석해서 유형 검사 엔진에게 보내주는 역할을 한다. 유형 검사 엔진에서는 사용자 요청 질의에 대한 유형을 판단하는데, 본 논문에서는 법률 기반, 키워드 기반, 질의 셋 기반의 세 가지 검색 방식을 제안하며, 각각의 검색 방식을 토대로 서로 다른 SPARQL 질의 변환이 일어난다. 법률 기반은 일반적인 법률명을 기반으로 그에 준하는 메타데이터들을 리턴 하는 형태이면서 가장 기본적인 방법이다. 키워드 기반은 사용자가 입력한 키워드를 기반으로 해당 키워드의 조문과 규정종류, 관련 법률을 연관 지어서 질의 문을 만들게 된다. 마지막 질의 셋 기반은 미리 정의된 법률 질의 문을 토대로 SPARQL 질의어가 생성되어 처리되는 방식이다. 서로 다른 형태의 검색 방식은 사용자 입장에서 검색에 필요한 정보 처리의 선택을 가능하게 하여 원하는 정보를 찾는 데 도움을 줄 수 있다. 또한, SPARQL 질의어로 자동 변환하기 위한 기술적인 측면에서도 검색방식에 따라 정해진 질의 문을 생성 가능토록 하여 시스템의 복잡도를 낮출 수 있다.

SPARQL 질의 변환 엔진은 그래프 패턴 매칭을 위해 BGP 기반으로 사용자 키워드를 변환하는 역할을 수행하고, 술어 기반 및 다중 그래프 탐색을 위한 매칭도 함께 수행한다. 또한, 키워드와 일치하는 그래프를 탐색하기 위해 필드 기반의 스트링 매칭도 SPARQL 질의 변환 엔진에서는 수행된다.

그림 7은 법령 정보 검색의 형태를 나타낸 것이다. 그림은 유형 검사 엔진에 의해 구별되는 법률 기반 방식과 키워드 기

반 방식이 그래프로 연결되어 결과를 반환하는 형태를 표현하고 있다. 법률 기반의 탐색 방식에서는 그림과 같이 법률과 관련된 메타데이터를 추출하여 나열한다.

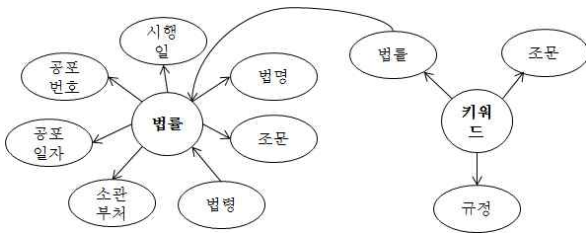


Fig. 7. A Form of Legal Information Retrieval

그림 8은 법률 기반 검색 방식의 SPARQL 변환 예시를 나타낸 것이다. 찾고자 하는 법률명을 기반으로 관계된 메타 정보들을 한 개의 질의 문으로 구성하여 표현하고 있다. 그림 7의 우측 그래프의 키워드는 사용자로부터 입력된 정보의 관련 조문과 규정, 관련 법률을 나열하고, 법률의 메타 정보와 연결될 수 있도록 질의 문을 구성한다. 이와 같은 법령 정보 검색 형태는 웹 인터페이스의 시각화 기술에 적용하여 정보를 표현할 수 있다.

```

SELECT ?law ?an ?ed ?dp ?np ?rd
WHERE {
    ?law rdf:type decree:Act .
    ?law decree:Act_Name ?an.
    ?law decree:Enforcement_Date ?ed.
    ?law decree:Date_Of_Promulgation ?dp.
    ?law decree:Act_No ?np.
    ?law decree:Relationship_Department ?rd
    FILTER regex(?law, "법률");
}
    
```

Fig. 8. Example of Transformation of SPARQL Query based on Law

그림 9는 정보 검색 엔진이 수행되는 순서도를 나타낸 것이다. 순서도는 사용자 질의어를 시작으로 질의 결과 값의 반환까지의 흐름을 나타낸다. 사용자 질의어는 입력된 요청 값을 의미한다. 먼저 입력된 요청 값을 분석하고, 유형을 검사하게 된다. 첫 번째 유형 검사에서 질의 셋 기반이라면 정해진 질의 셋으로 사용자가 검색을 수행한 것이며, 그에 따른 SPARQL 질의 변환을 수행한다. 유형검사에서 질의 셋 기반이 아니라면 법률명을 입력하여 검색을 수행 한 것으로 법률 기반으로 SPARQL 질의 변환이 수행된다. 마지막 유형의 키워드 기반의 검색이라면 키워드와 관련된 조문과 규정 관련 법률들을 찾는 SPARQL 질의 변환 문을 자동으로 생성하여 SPARQL 질의 처리기에 보내게 된다. SPARQL 질의 처리기에서는 변환된 질의 문을 기반으로 온톨로지에 질의를 하게 되고, 결과를 반환한다. 질의 엔진에 의해 OWL 온톨로지로부터 받아온 결과는 IRI 뒤에 리소스 값이 오게 된다. 따라서 IRI 값을 제거하는 정제 작업을 통해 사용자에게 해당 리소스만을 리턴하고, 정보 검색 엔진은 수행을 마친다.

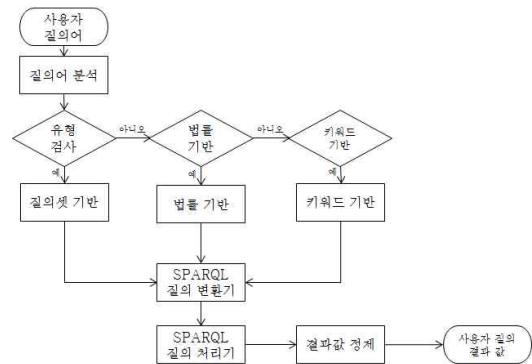


Fig. 9. A Flow Chart of Information Retrieval Engine

표 4는 본 논문에서 정의한 법령 온톨로지의 정의된 술어 리스트의 일부와 그에 준하는 의미들을 나타낸 것이다. 질의 셋 기반 정보 검색 시에 사용될 수 있는 술어들과 의미들을 조합하여 미리 정의된 질의 셋 검색을 가능하게 한다.

Table 4. Predicate List

술어	의미
klaw:isPurposeTo	법률의 제정 목적
owl:isDefinedBy	법률 용어의 정의
klaw:kindOf	해당 법률 용어와 관계된 하위 개념의 종류
klaw:minister	계획을 진행하는 장의 소속
klaw:year	계획이 갱신되는 기간
klaw:planType	계획의 종류 (기본계획 이거나 시행계획)
klaw:include	계획에 포함되는 사항의 종류
klaw:plIndicator	계획의 제정 및 개정을 위한 명령권자

3. Web Interface Engine

정보 검색은 기존의 얼마나 많은 정보를 빠르게 수집하고, 찾아내는 쪽에서 사용자가 원하는 정보를 좀 더 쉽게 찾고, 이해하기 위한 기술로 발전하고 있다[13]. 이런 측면에서 정보 검색 결과를 시각화해서 보여주는 것은 정보를 조금 더 쉽게 찾고, 이해하는데 도움을 준다.

기존의 법률 검색 서비스는 텍스트 기반의 정보 요약 및 목록, 트리형태의 탐색방식으로 키워드와 매치되는 검색결과를 보여주고 있다. 키워드와 관련된 대부분의 정보를 화면에서 텍스트로 나타내고, 사용자로 하여금 선택해서 결과를 찾아가는 식이다. 이러한 방식은 법률 지식이 없거나, 전문화된 정보를 모르는 일반사람들이 이용하기에 적합한 인터페이스로 볼 수 없다.

본 논문에서는 직관적으로 법령 정보를 찾을 수 있도록 웹 기반의 시각화 인터페이스를 적용한다. 시각화 기술과 관련된 연구는 오래전부터 많은 정보를 쉽게 표현하기 위해 연구되어 온 분야로 그림 10은 주요 시각화 기술의 종류를 나타낸 것이다. 트리(Tree) 형태의 정보표현은 개념을 중심으로 옆으로 혹은 아래로 확장되어 가는 구조이다. 선버스트(Sunburst) 방식

은 해당 개념이 최상위 개념으로 자리 잡고, 원이 커질수록 계층화된 정보 구조를 나타내고 있다. 그래프 (Graph) 표현방식은 노드와 에지형태로 정보를 표현하며, 방사형의 비방향성으로 표현하거나 해당 개념을 중심으로 원형으로 표현한다.

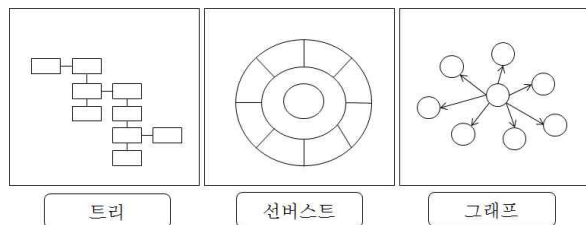


Fig. 10. Type of Visualization

본 논문에서는 방사형 그래프 방식으로 법령 정보를 표현한다. 제시하는 시각화 기술은 온톨로지가 가지고 있는 본연의 그래프 기반의 자료 구조를 바탕으로 웹 공간에 시각화해서 보여 주므로 데이터를 직관적으로 표현 가능하다는 장점을 가진다. 표현된 각 개념의 탐색 방법은 마우스를 이용해서 드래그 및 확장 이벤트를 통해 시각적으로 정보를 보면서 전문지식 영역의 데이터를 함축해가며 찾아갈 수 있도록 한다. 시각화 처리는 웹 화면의 특정 부분을 HTML5의 캔버스(canvas) 공간을 통해 차지하며, 캔버스 위에 노드(node)와 에지(edge) 형태의 그래프로 표현이 된다.

웹 인터페이스 엔진은 사용자 요청 결과를 트리플에 해당하는 주어-술어-목적어 형식으로 결과를 설명한다. 주어와 목적어는 노드가 되고 술어는 에지가 되어, 주어와 목적어의 관계를 이어주는 역할로 표현된다. 이와 같은 그래프 기반의 표현방식은 온톨로지가 가지고 있는 트리플 기반의 정보구조의 형식과 비슷한 구조를 가지고 있다는 점으로 웹으로 시각화해서 보여 주는데 다른 자료구조 변경 작업 없이 직관적으로 표현 가능하다는 장점을 가진다.

IV. Experiments

1. Example of the Legal Information Retrieval

법령은 조문별로 나타나는 규정이 명확하게 드러나는 영역이다. 본 논문에서는 법률 검색 서비스를 위한 조문별 특징을 기반으로 한 술어 기반 정보 추출 방식을 통해 의미적으로 해석 가능한 정보를 추출 한다. 표 5는 7가지 규정으로부터 추출 가능한 정보를 나타내고 있다. 추출 가능한 정보의 형태는 법률 검색을 수행 할 때 질의를 하거나 사용자가 알고자 하는 정보의 형태가 되며, 규정에서 말하고자 하는 주요 정보가 된다.

Table 5. List of Information by Regulation

규정 목적	정보
정의	법률의 목적
	법률에서 사용되는 용어
	용어의 구체적인 정의 용어와 관계된 법률외 정의
계획	계획의 유형
	계획을 진행하는 장의 소속
	계획의 재작성 기간
	계획에 포함되는 사항 리스트
	계획 명령권
	계획 통보 형태
위원회	위원회 이름
	위원회에서 심의되는 사항 리스트
	위원회 명령권
	위원회 구성 인원
	위원회 자격요건 리스트
벌칙	벌칙에 해당하는 사람(사항)
	징역 기간
	벌금 액수 벌칙 형태
과태료	과태료를 내야하는 사람(사항) 과태료 액수
시행일	법률이 공포되고 시행되는 날
법률 계층	법률 규정 종류
	규정별 항, 호, 목의 내용
	규정의 키워드

표5의 규정 별 추출 가능한 정보 리스트를 기반으로 온톨로지 질의 가능한 법률 검색 서비스의 형태는 다음 6가지 질의문의 형태로 응용할 수 있다.

- Q1. 현재 법률에서 제정된 규정의 종류는 무엇이 있는가?
- Q2. 현재 법률의 제정 목적 및 법률에서 사용된 용어 및 용어의 정의는 어떻게 되는가?
- Q3. 현재 법률의 위원회 이름과 위원장, 위원회 인원, 위원회 자격요건은 어떻게 되며, 위원회에서 처리하는 사항들은 어떤 것이 있는가?
- Q4. 현재 이러한 행위를 했는데, 벌칙에 해당하는 사례인가, 과태료에 해당하는 사례인가?
- Q5. 해당 사항에 따라 벌금 및 과태료는 얼마인가? 다른 처벌이 가능한 요소가 있는가?
- Q6. 여권의 유효기간 중 일반여권, 외교관 여권의 기간은 어떻게 되는가?

Q1부터 Q6 질의문의 해당하는 SPARQL 변환을 그림 11 ~ 16까지 나타내고 있다. 이러한 형태의 질의문은 기존의 정보 검색에서 정보를 찾던 방식과는 다른 형태의 검색을 가능하게 한다. 자연어 처리 및 지리, 위치 정보와 같은 다른 형태의 정보 자원과 연결을 하게 되면 법령정보를 이용한 다양한 서비스를 도출 할 수 있다.

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX decree:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree#>
PREFIX law1:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree/Law1#>

SELECT ?law ?head ?contents
WHERE {
    ?law rdf:type decree:Act.
    ?head decree:rule ?contents
}
    
```

Fig. 11. Example of SPARQL Transformation of Q1

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX decree:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree#>
PREFIX law1:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree/Law1#>

SELECT ?penalty ?fine ?penaltyRegulations
WHERE {
    ?p rdf:type decree:Penalty.
    ?f rdf:type decree:Fine.
    ?p law1:penaltyAbove ?penalty.
    ?p law1:penaltyBelow ?penalty.
    ?p law1:penaltyBelow ?penaltyRegulations
    ?f law1:fine ?fine
    FILTER regex(?p, "유형" || ?f, "유형")
}
    
```

Fig. 15. Example of SPARQL Transformation of Q5

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX decree:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree#>
PREFIX law1:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree/Law1#>

SELECT ?purpose ?term ?define
WHERE {
    ?law rdf:type decree:Act.
    ?law law1:isPurposeTo ?purpose.
    ?term owl:isDefineBy ?define
}
    
```

Fig. 12. Example of SPARQL Transformation of Q2

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX decree:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree#>
PREFIX law1:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree/Law1#>

SELECT ?term
WHERE {
    ?law law1:rule ?rule.
    ?rule law1:ruleContents ?x.
    ?x law1:hasSubParagraph ?y.
    ?y law1:seProvision ?keyword.
    ?y law1:seProvision ?term
    FILTER regex(?rule, "여권의 유효기간" &&
    ?keyword, "일반여권", "외교관 여권")
}
    
```

Fig. 16. Example of SPARQL Transformation of Q6

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX decree:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree#>
PREFIX law1:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree/Law1#>

SELECT ?committeeName ?chairMan ?number ?list
?deliberation
WHERE {
    ?committee law1:committeeName ?committeeName.
    ?committee law1:chairPerson ?chairMan.
    ?committee law1:committeeNumber ?number.
    ?committee law1:eligibilityCommittee ?list
    ?committee law1:deliberation ?deliberation
}
    
```

Fig. 13. Example of SPARQL Transformation of Q3

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX decree:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree#>
PREFIX law1:
<http://dss.ssu.ac.kr/complete/ontologies/2014/8/Korea_Decree/Law1#>

SELECT ?p ?f
WHERE {
    ?p rdf:type decree:Penalty.
    ?f rdf:type decree:Fine
    FILTER regex(?p, "유형" || ?f, "유형")
}
    
```

Fig. 14. Example of SPARQL Transformation of Q4

2. Implementation Framework

본 논문에서 제안하는 법령 정보 검색 프레임워크의 전체 구조 및 흐름은 그림 17과 같다. 프레임워크의 온톨로지 구축 엔진에서는 법률 문서로부터 데이터를 추출, 정제하여 OWL DL 수준의 디스크립션 로직에 따라 온톨로지로 변환하여 저장소로 저장된다. 웹 인터페이스는 HTML5 기반의 표준 인터페이스를 유지하며 CSS3의 변경으로 디자인 변경이 가능한 통합된 테마 인터페이스 저장소를 프레임워크 내에서 동작하도록 하여 웹 디자인의 변환작업을 손쉽게 한다. 웹 서버는 아파치 기반의 서블릿 형태를 기본으로 하고 있으며, 검색 엔진과 웹 인터페이스와의 검색 엔진 질의 및 결과 반환 작업을 수행한다.

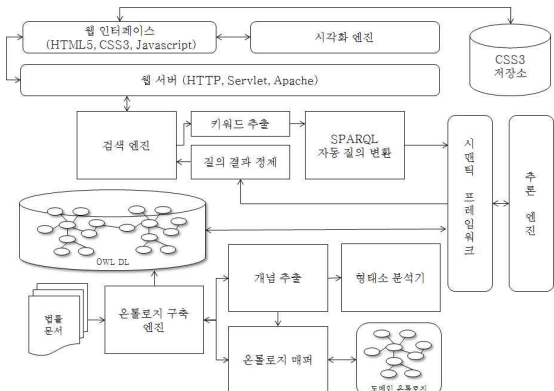


Fig. 17. The Structure of the Framework

검색 엔진은 웹 서버로부터 전달받은 사용자 질의 중 검색 유형을 분석하여 SPARQL 형태의 트리플 질의 문으로 자동 생성한다. 그 후, 시맨틱 프레임워크인 Jena를 통해 법령 온톨로지 저장소에 질의를 하고, 추론엔진을 거쳐 결과를 반환한다. 질의 결과 정제 엔진에서는 웹에 키워드만을 표현하기 위해 리소스가 가지고 있는 IRI 값을 제거한 데이터로 변환 작업하여 웹 서버에 보내게 되고, 웹 서버에서 시각화 엔진을 통해 결과를 사용자에게 보여 지는 구조를 가지게 된다.

프레임워크 내의 웹 인터페이스 엔진은 사용자 중심에서 정보를 시각화하여 보여주는 역할을 담당하며, 크게 시각화 엔진과 표현 엔진으로 구성된다. 그림 18은 웹 인터페이스 엔진의 흐름을 나타내고 있다.

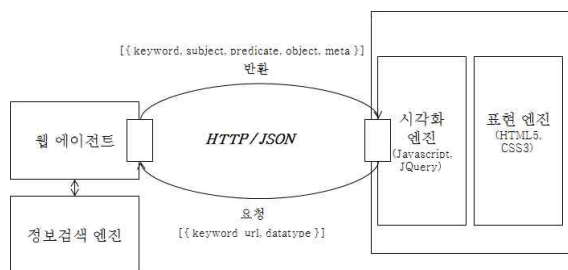


Fig. 18. A Flow of Web Interface

시각화 엔진은 Javascript, JQuery를 이용하여 사용자 정보 처리 및 통신, 시각화와 관련한 드로잉을 담당하며, 표현 엔진에서는 HTML5, CSS3의 웹 표준 기술을 이용하여 시각화 엔진에서 처리한 정보를 웹 브라우저로 보여주는 역할을 수행한다. 시각화 엔진과 서버의 웹 에이전트는 HTTP 기반의 POST 방식으로 정보를 주고받으며 데이터 포맷은 경량화된 형태의 JSON 포맷으로 한다. 요청 시에는 사용자 질의어와 서버 URI, 데이터 타입이 명시되어 에이전트에게 보내지고, 웹 에이전트는 다시 JSON 포맷으로 사용자가 요청한 질의어와 주어, 술어, 목적어의 결과 값에 해당하는 법령 트리플 정보와 메타정보가 함께 이동되어 처리되는 구조이다.

그림 19와 그림 20은 구현된 프레임워크의 실행 결과 화면을 나타내고 있다. 그림 19는 법령 온톨로지의 TBox를 로딩 후, 정보를 ABox 형태로 매칭 하여 온톨로지로 출력된 콘솔 결과 화면의 일부를 나타내고 있다. 그림 20은 본 논문에서 구축한 법령 온톨로지를 protege[14]의 그래프 툴을 이용하여 구축된 형태를 트리 기반으로 나타내고 있다.

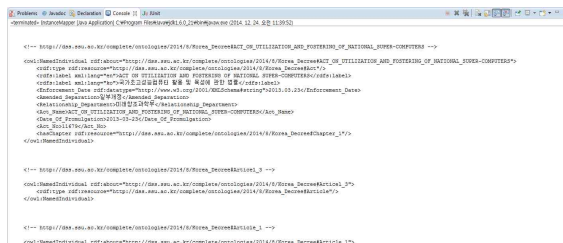


Fig. 19. Screen of OWL Ontology Construction

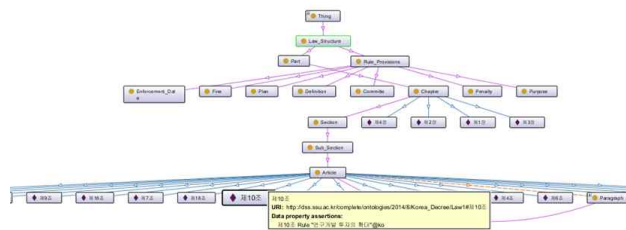


Fig. 20. Graph of Legal Ontology Construction

그림 21은 구축된 법령 온톨로지에 웹 기반의 인터페이스로부터 사용자가 ‘자동차전용도로’를 검색어로 입력 후, 나온 결과를 시각화 하여 보여 주고 있다. 결과는 ‘자동차전용도로’가 ‘정의’ 규정에 포함되는 사항이고, 관련 법률은 ‘도로교통법’으로부터 나오고 있음을 나타낸다. ‘도로교통법’은 법률과 관련된 메타 정보들이 방사형의 시각화로 ‘시행일’과 ‘법명’, ‘소관부처’, ‘공포번호’, ‘공포일자’ 등이 매칭 되어 결과로 나타나고 있다.

검색 결과는 그림 7에서 설명한 법령 정보 검색 형태 방식에 따라 웹 기반으로 시각화하여 표현하고 있으며, 프레임워크를 통해 자동 생성되어 나타난다. 이와 같은 검색 결과는 키워드 기반 방식의 결과이며, 프레임워크 내의 정보 검색 엔진에 의해 SPARQL 질의 문으로 자동 생성된다. 정보 검색 엔진에서는 해당 키워드를 중심으로 키워드와 관계된 법률을 찾고, 찾은 법률 정보들을 질의 문 생성을 통해 추출하여 프레임워크내의 웹 인터페이스 엔진을 통해 시각화하여 표현하는 방식으로 작동된다.

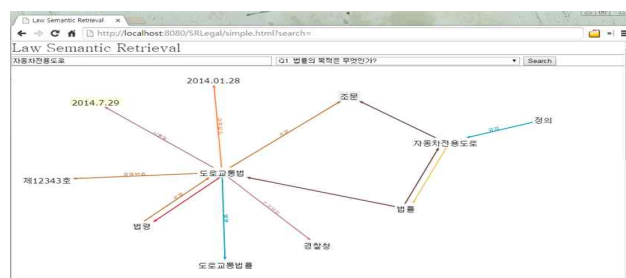


Fig. 21. Screen of Web Interface

3. Evaluations

본 논문에서 제안한 법령 정보 검색 프레임워크는 법령과 같은 전문 지식 영역을 시맨틱 웹 기술을 이용하여 기존의 정보 검색 결과를 개선하고, 일반 사용자 입장에서 검색 결과를 해석할 수 있는 기반 기술을 구축하는데 연구의 목표가 있다. 표 6은 기존의 대한민국 법령 정보 검색의 연구 및 서비스들과 제안하는 프레임워크를 비교한 표이다. 각각의 비교 항목은 온톨로지 구축 방법부터 정보검색, 법령계층, 온톨로지 구축 대상, 시맨틱 질의, 구축 언어, 법률 영역 온톨로지 구축, 시각화 인터페이스의 항목으로 비교하였다.

Table 6. Comparison of Legal Retrieval Framework

	1	2	3	4	제안
온톨로지 구축	X	X	수동	수동	반자동
정보 검색	O	O	O	O	O
법령 계층 반영	O	X	X	O	O
온톨로지 대상	X	전체	제한적	전체	전체
시맨틱 질의	X	X	X	X	O
구축 언어	한국어	한국어	한국어	한국어	한국어
영역 온톨로지 구축	X	X	O	X	O
시각화	X	O	X	O	O

대한민국 법령 분야의 기존의 구축된 서비스나 연구된 결과를 확인하면, 시맨틱 웹 기술을 이용한 정보 검색에 관한 연구가 미진한 것으로 확인 할 수 있다. 표에서 1은 법제처에서 제공하는 국가법령정보센터[2]이고, 2는 법제연구원의 시소러스 검색[9]이며, 3은 장인호의 ‘온톨로지 기반 법률 검색시스템의 구축 및 평가에 관한 연구’[15]이며, 4는 법무부의 iLaw[10]를 나타낸다. 1과 2는 시맨틱 웹 기술 방식이 고려되지 않았으며, 3은 특정 법률을 영역 온톨로지로 구축하고, 질의 문 셋을 이용해서 온톨로지를 검색하고, 평가를 하고 있으며 정보 검색 결과에 있어서 시각화 및 온톨로지 구축이 제한적인 형태로 구축하였고, 법령 계층은 고려되지 않았다. 4는 법률 영역에 관한 온톨로지 구축은 고려되지 않았으며, 검색 결과 또한 기존의 법제처의 결과와 동일한 형태를 나타내고 있다. 다만, 검색 결과에 있어서 시각화된 형태가 강화되었고, 법률 정보와 연관된 다른 미디어 정보 및 다국적 법률 용어 및 정보를 검색 결과로 보여준다. 하지만 법률을 위한 시맨틱 질의 방식에 관한 부분은 고려가 되어 있지 않다.

제안하는 프레임워크는 온톨로지 구축 방식에 있어서 법률 문장으로 추출 가능한 정보를 기반으로 반자동 형태로 온톨로지를 구축하였으며, 구축된 온톨로지를 기반으로 정보 검색 및 SPARQL 기반의 시맨틱 질의로 자동 변환 가능한 인터페이스를 갖추고 있으며, 각 법률 영역별 온톨로지 구축을 통해 방사형의 웹 기반의 시각화 인터페이스를 제공할 수 있다. 제안된 프레임워크를 기반으로 자연어 처리의 강화 및 웹 인터페이스의 개선을 통해 법률 영역과 같은 전문 지식을 이전과는 다른 형태의 서비스로 제공할 수 있다.

V. Conclusions

본 논문은 법령 온톨로지에 질의를 하고, 사용자 중심의 결과를 반환할 수 있는 온톨로지 기반의 법령 정보 검색 프레임

워크를 제안하였다. 제안하는 프레임워크의 역할은 온톨로지로 구축된 법령 도메인의 지식을 검색하고, 검색된 결과를 반환하는 역할을 수행한다. 또한, 사용자로부터 입력된 키워드를 온톨로지로 질의하기 위한 질의어 자동 변환 역할과 검색결과를 다른 해석 없이 직관적으로 이해할 수 있는 수준의 정보제공 역할을 함께 수행하여 시맨틱 법령 정보 검색 서비스를 위한 기반 구조로서의 역할을 할 수 있다.

본 논문에서 구현한 프레임워크는 법령 시맨틱 검색 서비스를 구축하기 용이한 기능적인 부분들이 존재한다. 프레임워크는 대한민국 법령을 온톨로지로 생성하고, 웹 기반에서 법령 정보 서비스가 가능한 엔진들이 함께 있다. 사용자 질의어 처리 및 온톨로지 질의가 가능한 정보 검색 엔진과 웹 기반에서 시각화된 형태의 정보 표현이 가능한 웹 인터페이스 엔진이 프레임워크 내에서 동작한다. 향후, 웹 인터페이스의 개선과 법령 정보가 기술된 온톨로지와 다른 정보 간 통합을 통해 법령 매쉬업 서비스를 위한 연구가 필요하다.

REFERENCES

- [1] D. W. Jo, "A Framework for Automatic Legal Ontology Construction and Semantic Retrieval," PhD Thesis, Soongsil University, Dec, 2014.
- [2] National Legislation Information Center, <http://www.law.go.kr/>
- [3] Legal information services, <http://glaw.scourt.go.kr/wsjo/intesrch/sjo022.do>
- [4] Thomas R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Ssharing," International journal of human-computer studies, vol.43, no.5, pp.907-928, 1995.
- [5] E. Prud'Hommeaux, and A. Seaborne, "SPARQL Query Language for RDF," W3C Recommendation, Jau, 2008.
- [6] Y. T. Park, C. M. Choi, "Overview and Research Trends in Ontology Reasoning," Communications of the Korea Information Science Society, Vol. 24, No. 4, pp.17-23, Apr, 2006.
- [7] M. Stocker, A. Seaborne, A. Bernstein, C. Kiefer, and D. Reynolds, "SPARQL basic graph pattern optimization using selectivity estimation," Proceedings of the 17th international conference on World Wide Web, Apr, 2008.
- [8] SemanticWebTools, <http://www.w3.org/wiki/SemanticWebTools>

- [9] Korea Legislation Research Institute,
http://61.81.114.30/the_web/thesa_content.asp
- [10] iLaw, <http://www.ilaw.go.kr/moj/2010/index.do>
- [11] H. M. Jung, Y. H. Choi, "MOJ's iLaw System for Intelligently Supporting Legislation," HCI Korea 2011, vol.2011, No. 1pp.2030-2035, Jau, 2011.
- [12] D. W. Jo and M. H. Kim, "A Study on Legal Ontology Construction," Journal of The Korea Society of Computer and Information, Vol. 19, No. 11, pp. 105-114, Nov, 2014.
- [13] D. W. Jo, J. W. Choi and M. H. Kim, "The Design and Implementation of OWL Ontology Construction System through Information Extraction of Unstructured Documents," Journal of The Korea Society of Computer and Information, Vol. 19, No. 10, pp. 23-33, Oct, 2014.
- [14] Protege, <http://protege.stanford.edu/>
- [15] In. Ho. Chang, "Developing and Evaluating an Ontology-based Legal Retrieval System," Journal of the Korean Society for Library and Information Science, Vol. 45, No. 2, pp. 345-355, May 2011.

Authors



Dae Woog Jo received the B.S. in Computer Engineering from Hallym University, Korea, in 2008. M.S. and Ph.D. degrees in Department of Computer Science and Engineering from Soongsil University, Korea, in 2010 and 2015, respectively. He is currently a Research Professor in the School of Software, Soongsil University. He is interested in semantic web, ontology engineering, distributed system and cloud computing.



Myung Ho Kim received the B.S. in Department of Computer Science from Soongsil University, Korea, in 1989. M.S. and Ph.D. degrees in Department of Computer Science and Engineering from POSTECH, Korea, in 1991 and 1995, respectively. He is currently a Professor in the School of Software, Soongsil University. He is interested in parallel computing, distributed computing, system software and information security.