

시맨틱 웹 응용을 위한 웹 온톨로지 구축기법

김 수 경[†] · 안 기 흥^{††}

요 약

시맨틱 웹 응용의 성공은 현재 웹 기술을 바탕으로 하여, 지식에 대한 규칙과 추론 기능이 제공되는 웹 온톨로지 구축에 달려있다. 이를 위해 본 연구는 기존에 제안된 온톨로지 구축기법들을 비교 분석하고, 시맨틱웹과 웹 온톨로지의 특징을 조사하여 시맨틱 웹 응용의 기반기술이 되기 위한 웹 온톨로지의 기술적 특징과 지식표현 단계를 정의하고, 현재 웹 기술과 관련되는 웹 온톨로지의 기술적 요소를 연구하였다. 이를 바탕으로 시맨틱 웹 응용을 위한 웹 온톨로지 구축기법을 제안하고, 제안한 구축기법을 적용하여 다양한 지식분야의 웹 온톨로지들을 구축하였으며, 구축된 웹 온톨로지들의 추론 검증 결과를 통해 제안된 구축방법이 기존 온톨로지 구축기법들에 비해 시맨틱 웹 응용의 요구 조건을 충족하는 결과가 획득됨을 확인하였으며, 따라서 제안한 웹 온톨로지 구축기법이 시맨틱 웹 응용에 적합한 구축기법임을 증명하였다.

키워드 : 웹 온톨로지, 시맨틱 웹 응용, 지식 표현, 추론

Web Ontology Building Methodology for Semantic Web Application

Sukyung Kim[†] · Keehong Ahn^{††}

ABSTRACT

Success of a semantic web application, currently base on web technology, depend on web ontology construction that provided rule and inference function about knowledge. For, this study compared the ontology construction methods that were already proposed, and analyzed, and investigated characteristics of semantic web and web ontology, investigated characteristics of semantic web and web ontology, and defined characteristics of web ontology as-based technology of a semantic web application and knowledge representation steps, and studied a technical element that related currently web technology, and proposed a web ontology construction method for a semantic web application with bases to these. Established web ontologies of various knowledge fields as applied the construction method that proposed. Also evaluate performance of web ontology through inference verification of web ontologies established, web ontologies evaluated performance of web ontology as used inference verification. According to this, we confirmed that proposed construction method that can establish the ontology suitable for semantic web application.

Key Words : Web Ontology, Semantic Web Application, Knowledge Representation, Inference

1. 서 론

시맨틱 웹 응용의 최종 단계는 현재 웹 기술을 바탕으로 하여, 웹 온톨로지의 구축을 기반으로 추론 엔진을 이용하여 지식에 대한 규칙과 추론 기능이 제공되는 시맨틱 웹의 최소 요건[1]을 만족하는 응용시스템이다. 그러나 시맨틱 웹 응용의 기반 기술인 추론 기반의 웹 온톨로지 표현과 구축은 다른 시맨틱 웹 기술 수준들에 비해 발전 수준이 부족한 상황이고 이는 시맨틱 웹 응용의 발전과 보급에 장애가 되고 있다. 또한 팀 버너스 리가 2005년 7월 AAAI에서 발표한[2] 바와 같이 시맨틱 웹은 인공지능이 아님에도 불구하고, 웹 온톨로지 구축기법은 시맨틱 웹과 웹 온톨로지의 특

징을 고려하지 않고 인공지능 분야에서 연구된 온톨로지 개념과 구축기법이나 특정 응용을 위한 온톨로지 구축기법 또는 소프트웨어공학분야에서 제안한 구축방법 등을 적용하기 때문에 추론과 같은 웹 온톨로지의 역할을 제대로 제공하지 못하는 문제점이 있으며, 온톨로지에 구축될 지식의 표현에 있어서도 형식적이고 명시적인 표현 기법을 위한 일관성을 제공하지 못함으로 인해 온톨로지 공학자나 기계가 구축된 온톨로지의 지식을 이해하거나 재활용하는데 어려움이 있다. 또한 시맨틱 웹 발전을 위해서는 표준화된 웹 온톨로지 모델링 방법의 연구 개발이 필요하나 현재 이에 대한 연구가 부족한 실정이다[3].

따라서 본 논문은 시맨틱 웹 응용의 기반 기술이 되기 위한 웹 온톨로지의 기술적 특징과 요소들을 연구하고 기존 온톨로지 구축 기법들을 비교 분석하여 시맨틱 웹 응용에 적합한 웹 온톨로지 구축 기법을 제안하고자 한다. 본 논문

[†] 정 회 원 : 한밭대학교 컴퓨터공학과 연구원

^{††} 정 회 원 : 한밭대학교 컴퓨터공학과 교수

논문접수 : 2007년 8월 7일, 심사완료 : 2007년 9월 14일

의 구성은 다음과 같다. 2장에서는 시맨틱 웹과 웹 온톨로지의 관계에서의 특징과 기반 기술 등의 관련 연구 그리고 기존 온톨로지 구축기법들의 특징들을 비교 분석한다. 3장에서는 시맨틱 웹 응용의 핵심 기반 기술로 적용되기 위한 웹 온톨로지의 기술적 요소와 지식 표현의 명시적 단계를 정리하고, 조사 분석된 결과에 따라 서술 논리(Description Logic)[4]와 규칙 언어(Rule Language) [5, 6]를 기반으로 시맨틱 웹과 웹 온톨로지의 특징을 만족하는 웹 온톨로지 구축 기법을 제안한다. 4장은 제안된 구축 기법에 따라 웹 온톨로지들을 구축한 뒤, 구축된 웹 온톨로지의 성능을 실험하여 기존 온톨로지 구축 기법보다 제안된 기법이 시맨틱 웹 응용의 기술적 요구에 만족함을 증명하고, 5장은 본 논문에 대한 결론과 앞으로의 연구 방향을 기술한다.

2. 관련 연구

2.1 시맨틱 웹에서의 웹 온톨로지

웹 온톨로지에 대해 먼저 정의하면, 웹 온톨로지는 넓은 의미에서 메타데이터 온톨로지의 한 종류로서 웹상의 지식을 표현하고, 공유하며, 재사용할 수 있도록 하는 온톨로지이다[7, 8]. 웹 온톨로지는 지식 표현의 대상이 대부분 웹상에 존재하는 자원이나 개념이다. 웹상에 존재하는 자원들이나 개념은 현재 태깅(tagging)으로 구분되어 있으며, 이와 같은 웹과 관련된 기술을 바탕으로 웹 온톨로지는 웹상의 자원 또는 개념에 대한 표현 방법과 자원 또는 개념의 의미 정의에 필요한 기준 그리고 각 자원 또는 개념의 의미론적 연결 방법을 제공한다.

시맨틱 웹이 기존 웹의 확장이라고 하는 것은 XML, URI와 같이 이제까지 개발된 웹 기술을 바탕으로 의미검색이 가능하도록 웹 온톨로지가 사전에 구축되기 때문이다. 즉, 시맨틱 웹을 실제로 구현하는 기반기술이 웹 온톨로지이다.

그리고 시맨틱 웹의 구조가 기존의 웹과 다른 점은 의미 계층에 해당하는 웹 온톨로지 지식 기반시스템과 온톨로지에 포함된 개념의 의미를 유추할 수 있는 추론엔진 시스템이 필요한 것이다. 온톨로지 추론엔진은 온톨로지에 표현된 추론 규칙을 적용하여 개념을 유추하는데 이때 질의어의 관계를 파악한 뒤 관련된 용어를 검색 엔진에 전달하며 이를

통해 추론의 기능을 실행하는 것이다. 이에 따라 시맨틱 웹 응용을 위한 기술 구성과 단계는 온톨로지 모델링을 기반으로 시맨틱 주석과 저장 그리고 질의와 추론을 거쳐 구현됨을 (그림 1)이 보여준다[9].

2.2 웹 온톨로지를 위한 기술적 요구

2.1절에서 조사 분석된 시맨틱 웹과 웹 온톨로지의 특징에 따른 시맨틱 웹 구조를 완성하기 위해 필요한 웹 온톨로지의 기술적 요구는 다음과 같다.

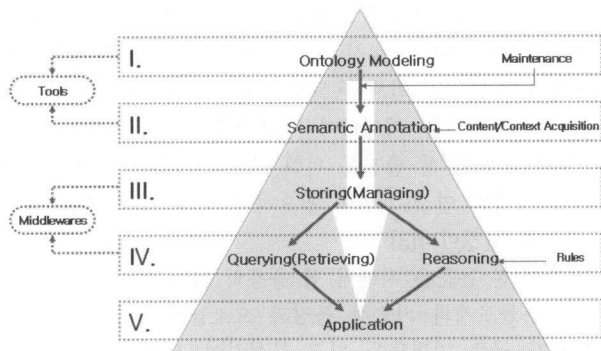
- (1) 웹 온톨로지는 온톨로지 공학자나 기계가 이해할 수 있도록 지식이 명시적이고 형식적으로 표현되어야 한다.
- (2) 구문론적 상호작용으로 파싱된 데이터의 재사용을 증가시키거나 의미론적인 상호 작용 등을 현재 웹 환경보다 향상시켜야 한다.
- (3) 웹 온톨로지의 제약 사항 중 하나인 개념적 모델을 묵시적으로 표현함에 따른 코드와 표현의 차이점을 극복하여야 한다.
- (4) 현재 웹을 바탕으로 정보의 표현과 추론이 추가되어야 한다.
- (5) 응용들 간에 웹 기반의 지식 처리와 공유, 재사용이 가능해야 한다.
- (6) 표현된 지식들을 웹 온톨로지로 구축하기 위한 언어는 OWL(Web Ontology Language)[10], RDF (Resource Description FrameWork), DAML+ OIL (DARPA Agent Markup Language+Ontology Inference Layer) 등과 같은 웹 온톨로지 언어를 이용한다.
- (7) 웹 온톨로지가 존재론에서 거론하는 모든 형상에 대한 표현을 하는 것은 매우 어렵기 때문에 웹이라는 특정 분야를 표현하고 구축한다.
- (8) 웹 온톨로지는 웹 서버를 통해 운용되고 응용시스템과 연결되며 개방형의 구조를 갖는 것이 좋다.
- (9) 웹 온톨로지는 온톨로지에 포함된 개념의 의미를 유추할 수 있도록 Jena[11]와 같이 추론 엔진의 기능을 제공하여야 한다.

2.3 기존 온톨로지 구축기법 비교 분석

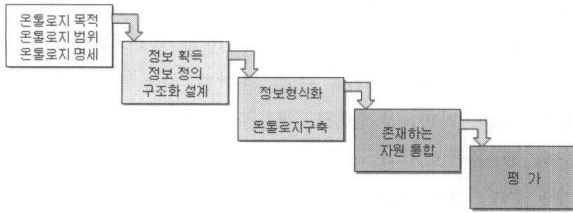
온톨로지는 추상적인 개념구조를 도출하고, 개념 요소의 추출과 그 관계를 결정하는 작업이기 때문에 일률적인 알고리즘적인 방법을 개발하는 것은 매우 어려우며, 개발하고자 하는 대상의 성격과 개발 목적에 따라 나름대로의 방법과 특징을 가진다[7].

온톨로지를 구축하는 가장 일반적인 과정은 (그림 2)의 과정에 따라 진행된다. 기존에 연구된 온톨로지 구축기법들도 대부분 이러한 과정이 일부 포함되거나 유사한 경우[8, 12]가 많이 발견된다.

온톨로지 개발과 관련된 기술은 실용적 목적에서 볼 때는 충분히 발전된 것은 아니지만 Cyc[13], KACTUS[14], TOVE [15]와 같은 다양한 온톨로지 개발기법 들이 인공지능 분야나



(그림 1) 시맨틱 웹 응용의 기술 구성과 단계



(그림 2) 일반적인 온톨로지 개발 과정[8]

지식정보 분야를 위해 개발되었으며, OTKM[16] 등은 온톨로지 기반 지식 관리 시스템에서, 기타 Ontology Development 101[17], DOLCE[18], Lifecycle of a Casual Web Ontology Development[19] 등은 시맨틱 웹의 제안된 후에 제시된 구축기법 들이다.

위에 기술된 각 온톨로지 구축기법 들을 소프트웨어공학적 기준과 온톨로지 표현을 기준으로 비교 분석한 결과는 <표 1>과 같다. <표 1>의 구분 항목은 각 온톨로지 구축기법들을 조사 분석한 결과 공통적으로 포함되는 항목이나 온톨로지 구축을 위해 중요한 요소들로 각 구축 방법들의 해당 여부를 'O' 로 표시한다.

각 구축기법들을 전체적으로 비교 분석한 결과 TOVE, KATUS, OTKM, DOLCE의 온톨로지 구축 시점(starting point)은 응용 또는 특정한 업무를 통해 시작되며, 방법론 모델은 TOVE와 같은 단계 기반 모델과 METHONTOLOGY[20]와 같은 진화형 프로토타입 모델이 있으며, 온톨로지 구축 시점은 프로젝트 개발 상황이나, 접근 방법에 따라 온톨로지가 전혀 구축되지 않은 시점에서 새롭게 온톨로지 구축을 시작하는 경우, 기존에 온톨로지를 기반으로 온톨로지를 구축해야 하는 경우, 정보 자원의 집합으로 온톨로지를 구축해야 하는 경우, 혼합된 형태로 구축해야 하는 경우로 구분된다.

이와 같은 결과에 따라 각 온톨로지 구축기법 등에 대한 특징과 장·단점은 다음과 같다.

- (1) OTKM과 METHONTOLOGY와 방법은 개발 단계의 모든 과정에 대한 절차를 제시한다. 따라서 웹상의 단위 정보를 표현하는 웹 온톨로지의 구축에 있어서는 구축 과정이 너무 방대하고 시간이 소요된다.
- (2) KATUS, DOLCE, Ontology Development Guide 101 방법은 온톨로지의 지식 표현 방법에 대한 제시가 모호하며, 온톨로지 구축에 대한 형식적인 단계가 명확하지 않으며, 특히 온톨로지 구축 이후의 온톨로지 유지 보수에 대한 단계를 제시하지 않았다.
- (3) Cyc, TOVE 등은 지식기반시스템 구축에 중점을 둔 방안으로 웹 온톨로지를 포함한 모든 온톨로지 구축에 적용하기는 적합하지 않다.
- (4) Ontology Development 101 온톨로지 구축방법은 온톨로지 구축의 필요성과 경험 사례를 바탕으로 기초적인 온톨로지 구성에 대한 과정을 제공하였으며, 체계적이고 구조적인 온톨로지 구축 방법은 제공하지 않았다.
- (5) METHONTOLOGY 기법은 반복 검토를 통한 정밀화 단계가 생략되어 있기 때문에 개발속도는 빠르지만 온톨로지의 질적 수준은 미약하다.
- (6) OTKM 방법은 소프트웨어 공학적인 접근의 프로토타입 방법을 중심으로 대규모의 온톨로지 프로젝트를 개발하는데 적합하지만 온톨로지의 개발 시간이 많이 소모될 수 있다.
- (7) Lifecycle of a Casual Web Ontology Development 방법은 웹 온톨로지를 빠르게 개발할 수 있도록 많은 과정을 생략하고 온톨로지의 표현도 간단한 기호를 사용하는 방법을 제시하였으나, 실제 온톨로지 구축에 있어서 제시된 기호를 이용할 때 온톨로지 공학자나 기계가 지식의 이해나 공유를 어려운 단점이 있다.
- (8) 온톨로지 공학자나 기계가 이해할 수 있도록 형식화된 지식 표현을 위한 문서의 제공과 추론을 위한 지식 표현의 방법 등이 부족하다.

<표 1> 온톨로지 구축 방법론별 비교

구분	국 외									국 내	
	Cyc	KACTUS	TOVE	Meth-ontology	Ontology Development 101	OTKM	DOLCE	Casual Web Ontology	BOE	IEE	
온톨로지 생성 기법	Bottom-up	O	-	-	-	-	-	O	-	-	
	Top-down	-	O	O	-	-	-	-	-	O	
	middle-out	-	-	-	O	O	O	O	O	-	
온톨로지 시작 방법	업무 중심 개발	-	O	O	-	-	O	-	O	-	
	추상적 개발	O	-	-	O	O	-	O	-	O	
모델	단계별 모델	O	O	O	-	O	-	O	-	O	
	프로토타입	-	-	-	O	-	O	-	O	-	
온톨로지 표현	형식화된 온톨로지 기술	-	-	O	O	-	O	O	O	-	
	지식 표현 언어	F-logic, CycL	-	C++	-	DL	-	FOL	-	-	
	지원 영역	일반	업무	업무/기업	일반	일반	프로젝트	프로젝트	일반	일반	
	개발년도	-	1994	1992	1997	2000	1999	2004	2004	2004	2004

3. 웹 온톨로지 구축기법 제안

앞 절에서 조사 분석된 결과를 기반으로 웹 온톨로지 구축 기법을 제시하기 위해 본 절은 먼저 시맨틱 웹 응용의 기반 기술로 웹 온톨로지가 사용되기 위해 필요한 다양한 요소들을 먼저 분석한다.

3.1 기반 기술 요소 연구

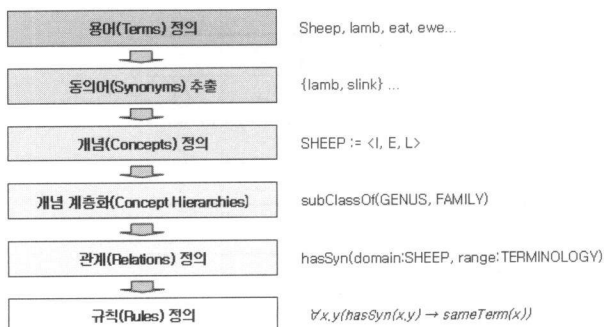
3.1.1 지식의 명시적 표현

웹 온톨로지의 제약 사항 중 하나인 개념적 모델을 묵시적으로 표현하는데 따른 코드와 표현의 차이점을 극복하기 위해 본 연구에서는 서술 논리와 규칙을 이용하여 웹상에 존재하는 단어나 용어들로부터 온톨로지에 표현되어야 할 지식, 추론 그리고 규칙의 정의에 이르는 과정을 (그림 3)과 같이 기술한다.

(그림 3)에 기술된 각각의 단계는 본 연구에서 제안할 웹 온톨로지 구축기법 단계 중 온톨로지 내부 분석과 설계 단계에서 표현될 각 온톨로지 개념, 지식, 공리 등의 명시적 표현을 위해 적용된다. 특히 본 연구에서는 (그림 3)을 각 단계별로 문서화한 양식을 제시한다.

(그림 3)에서 제시된 각 단계에 대한 구체적인 설명은 다음과 같다.

- [단계 1] 용어 계층에서는 온톨로지내 표현될 지식 영역에 포함되는 일반적인 명사형의 용어들을 나열한다.
- [단계 2] 동의어 계층은 나열된 용어들 중에서 의미가 유사하거나 같은 용어들을 동의어로 정의한다.
- [단계 3] 개념화 단계는 동의어와 용어들을 그룹화하여 서술논리의 선언적 형식(Assertional axiom) 표현에 따라 개념 또는 클래스의 형태로 표현한다.
- [단계 4] 개념계층화 단계는 나열된 개념 또는 클래스들 간의 상하 관계들을 명시적으로 표현한다.
- [단계 5] 관계 단계는 각 개념 또는 클래스들을 도메인(domain)과 영역(range)로 구분한 뒤 개념 또는 클래스들 간의 관계를 명시적으로 표현한다.
- [단계 6] 규칙 단계는 단계 3과 단계 4가 서술 논리의 선언적 형식을 사용하여 지식을 표현하였다면 이 단계는 용어적 형식(Terminology Axiom)과 규칙 언어를 사용하여 속성과 속성간의 관계 또는 속성과 개념 간의 관계 등을 표현한다.



(그림 3) 온톨로지 지식의 명시적 표현 단계

3.1.2 시맨틱 웹 응용을 위한 온톨로지 기술 요소

시맨틱 웹 응용과 웹 온톨로지는 현재 웹의 환경을 기반으로 활용된다. 이에 따라 웹과 관련된 요소들을 적용하여 시맨틱 웹의 기반 지식으로 웹 온톨로지가 동작하기 위해 단순히 웹 온톨로지의 구축만이 아닌 이들을 운용하기 위한 다양한 기술적(technology) 요소들이 필요하다. 이같은 기술적 요소들을 결정하기 위해 “Wonder Web deverable 17” [21]에서 제안한 시맨틱 웹 인프라를 위한 온톨로지 형식화 관련 내용을 참고로 시맨틱 웹 응용을 구현하기 위해 선행되어야 할 온톨로지와 관련된 기술적 요소들은 <표 2>로 제안한다. 이 기술적 요소들은 시맨틱 웹 응용을 구현하였을 때 웹 온톨로지가 시맨틱 웹 응용과의 상호작용을 위해 필요한 최소한의 기술적 요소이며, 온톨로지가 구축되기 전과 구축되는 과정 그리고 구축 완료 후까지 시맨틱 웹 응용 구현의 모든 단계에서 필수적으로 적용된다.

또한 <표 2> 요소들은 온톨로지를 기반으로 한 시맨틱 웹 응용 시스템을 구축할 때 시스템 성능의 평가 항목으로도 활용될 수 있다. 기타 온톨로지를 생성할 때 고려해야 할 요소들은 저작권 배포 문제, 온톨로지 구축 시점, 자동화 여부, 로컬 온톨로지와 글로벌 온톨로지 간의 관계, 온톨로지의 결합/통합/합병의 문제 등이다.

3.2 제안 웹 온톨로지 구축 기법

2장과 3.1절에 의해 조사 분석된 결과를 기준으로 본 연구는 (그림 4)와 같이 시맨틱 웹 응용에 적합한 웹 온톨로지 구축기법의 단계를 제안한다. (그림 4)의 각 단계는 이전 단계로 피드백 되어 구축될 온톨로지의 반복적인 수정이 가능하며, 각 단계는 단계마다 기술되어야 할 문서화 과정이 포함된다.

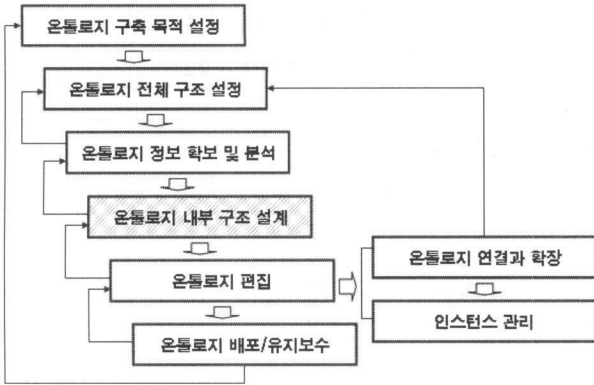
다음은 (그림 4)의 각 단계마다 온톨로지 구축을 위해 수행되어야 할 구체적인 내용과 온톨로지의 지식표현의 형식화를 위해 작성될 문서들에 대한 설명이다.

3.2.1 단계 1 : 온톨로지 구축 목적 설정

이 단계는 응용 시스템에서의 온톨로지 기능, 온톨로지 단위, 지식 표현 단위 등을 연구 조사하여 <표 3>의 온톨로지

<표 2> 시맨틱 웹 응용을 위한 기술 요소

번호	주 제	설 명	확언
1	Ontology Language	구축에 사용될 온톨로지 언어	
2	Rule Language	규칙 표현에 사용될 언어	
3	Language Extension	온톨로지 확장에 필요한 언어	
4	Ontology Server Architecture	온톨로지 서버 구조	
5	Ontology Server Demonstrator	온톨로지 서버를 통한 데모	
6	Ontology Editor	온톨로지 개발 툴	
7	Ontology Inference System	온톨로지 추론 시스템	
8	Ontology DataBase	온톨로지 데이터베이스	
9	Reasoner Demonstrator	추론 데모	
10	Ontology RoadMap	구축될 온톨로지 로드맵	
11	Ontology Methodology	온톨로지 구축 기법	
12	Ontology Library	온톨로지 라이브러리 제공	
13	Module Mechanism	모듈별 동작 구조	
14	Evaluation Report	온톨로지 평가	
15	Project Presentation	온톨로지 활용 제시	
16	Technical Implementation Plan	온톨로지 구축과 응용 시스템	



(그림 4) 제안 웹 온톨로지 구축기법 단계

<표 3> 온톨로지 개요서

온톨로지 개요서	
번호 : _____	작성일 : _____
응용 시스템명	이미지 지식 검색 시스템
응용 시스템 목적	현재의 웹은 웹에서 이미지를 검색할 때 이미지가 표현하는 내용을 중심으로 검색할 수가 없고 검색어나 이미지 파일명과 같은 단편적인 정보로 검색이 실행되거나 이미지를 검색할 때 검색어가 아닌 검색 문장을 이용해 이미지의 내용을 검색한다.
온톨로지 필요성	이미지의 내용을 검색할 때, 이미지에서 중심이 되는 주제를 온톨로지로 구축하여 일반 데이터베이스의 검색 기능보다 지식 추론에 대한 기능이 뛰어난 온톨로지를 지식 기반으로 도입한다.
응용 시스템에서의 온톨로지 기능	1. 이미지가 표현하는 주제에 대한 용어 정보 제시 2. 이미지와 관련된 용어에 대한 지식 추론 3. 동음이의어와 이음동어어 문제 해결 4.
구축 온톨로지명	이미지 지식 기반 온톨로지
지식 표현에 따른 온톨로지 구분	1. 이미지의 내용을 검색하기 위한 지식 표현 : 2. 이미지 주제에 대한 지식 표현 : 3.

개요서를 작성한다. 온톨로지개요서는 응용 시스템의 목적, 온톨로지 필요성, 응용 시스템에 있어서의 온톨로지 기능과 글로벌 온톨로지에 해당하는 대표 온톨로지에 대한 정보 그리고 지식 표현에 따라 구분될 수 있는 온톨로지들의 정보를 기술한다. 개략적인 온톨로지 정보는 <표 4>의 도메인 온톨로지 요약표를 통해 더욱 구체적으로 기술된다.

도메인 온톨로지 요약표는 온톨로지 구축 대상과 응용에 대한 좀 더 구체적인 역할과 목적, 온톨로지의 범위와 표현 수준 등에 대한 정보가 작성되며 이 온톨로지 개요서는 여러 개의 문서가 작성될 수 있다. 이 단계의 마지막 형식 문서로 구축될 응용시스템을 통해 얻고자 하는 질의와 그에 대한 결과를 기술한 <표 5>의 적격질의 목록표를 작성한다.

적격질의 목록표는 추후 온톨로지 내부적 구조 설계 단계에서 도메인 온톨로지서에서 표현될 지식과 추론규칙의 기본 정보로 활용되며, 구축된 온톨로지의 성능 검증 시 판단 정보로 이용된다. 적격질의 목록표가 명확하게 작성될수록 온톨로지가 표현하고 획득하고자 하는 지식에 대한 명확성과 형식

<표 4> 도메인 온톨로지 요약표

도메인 온톨로지 요약표	
번호 : _____	작성일 : _____
관련 응용 시스템	이미지 지식 검색
글로벌 온톨로지명	이미지 지식 기반 온톨로지
도메인 온톨로지명	양(sheep) 온톨로지
온톨로지 개발자	Hong Kil Dong
온톨로지 개발 목적	이미지 지식 검색 시스템에서 지식 정보 제공을 위해 동물의 양에 대한 온톨로지를 구축한다.
온톨로지 표현 수준	1. 서술 논리와 규칙언어를 이용한 명시적 지식 표현 2. 웹 온톨로지 언어(OWL) 사용 3.
온톨로지 표현 범위	1. 양에 대한 기본 정보 2. 양의 학명과 같은 양의 종별 분류 3. 양의 종별 특징 4.

<표 5> 적격질의 목록표

적격 질의 목록표				
번호 : _____	작성일 : _____			
관련 응용 시스템	이미지 지식 검색			
글로벌 온톨로지명	이미지 지식 기반 온톨로지			
번호	질의내용	예상답	관련도메인	비고
1	lamb과 동일한 의미를 가진 이미지는?	slink	terminology	
2	merino sheep의 종명은?	ovis aries	zoological	
3	highborn sheep이 서식하는 지역은?	Rockie Moutain	zoological map	
....

<표 6> 온톨로지명세서

온톨로지 명세서			
번호 : _____	작성일 : _____		
글로벌 온톨로지명	이미지 지식 기반 온톨로지		
외부 온톨로지 이입	1. 규칙에 대한 완전한 추론을 위해 swrl온톨로지를 이입 2. OWL과 RDF에 있는 어휘를 표현하기 위해 rdf, rdfs, owl 온톨로지를 기본으로 이입		
중심 온톨로지	온톨로지간의 지식 추론과 정보 교환에서 중심 역할을 할 온톨로지로 결정		
온톨로지 네임스페이스 설계			
도메인온톨로지	역할	URI	접두사
sheep	이미지 내용 주제	http://www.owl.org/	#Sheep
terminology	이미지 용어	http://www.owl.org/	#terminology
owl	owl 어휘 사용	http://	#owl
....

성이 강화된다.

3.2.2 단계 2 : 온톨로지 전체 구조 설정

이 단계는 먼저 기존에 구축된 온톨로지가 있는 경우 이를 재활용할 수 있는 방안을 검토한 뒤, <표 6>의 온톨로지 명세서를 작성한다.

온톨로지명세서는 [단계 1]에서 정의한 범위와 대상들에 대한 전체적인 온톨로지 명세와 그에 따른 각 온톨로지의 네임스페이스와 접두어(prefix) 등을 결정하며, 구축될 온톨

<표 7> 도메인별 용어조사표

도메인 용어 조사표				
번호 : _____		작성일 : _____		
글로벌 온톨로지명	이미지 지식 기반 온톨로지			
도메인 온톨로지명	양(sheep)			
단어명	동의어	약어	설명	유형
sheep	domestic sheep		양, 면양	class
lamb	slink, suckler		어린 양, baby sheep	class
wool	fur, hair, coat..		양모, 털...	class
mutton	hoggatt		양고기	class
wild sheep			야생종 양	class
location			양의 서식지	property
merino			양의 한 품종	instance
.....

<표 8> 도메인별 개념사전

도메인 개념 사전				
번호 : _____		작성일 : _____		
글로벌 온톨로지명	이미지 지식 기반 온톨로지			
도메인 온톨로지명	양(sheep)			
개념명	속성	인스턴스	타 개념 관계	크기
sheep	location	Map	3
	genusName	Zoological
	hasColor
	merino	domestic sheep
wild sheep	location	sheep
	genusName
	hasColor
	Bighorn	wildsheep

로지들 중에서 중심 온톨로지를 결정한다. 또한 구축될 온톨로지의 하위 온톨로지들로 각 도메인 온톨로지들과 외부에서 이입될 온톨로지에 대한 정보를 정리하여 온톨로지내 각 개념들의 유일성을 보장한다.

3.2.3 단계 3 : 온톨로지 정보 확보 및 분석 단계

이 단계는 온톨로지 도메인 범위에 포함되는 개념들의 명세를 작성한다. 개념들의 용어와 동의어 그리고 개념에 대한 분석과 설계를 한다. 분석 결과에 따라 이전의 단계로 피드백 하여 설계를 조절한다.

먼저, 도메인 온톨로지별 용어(단어)를 최종적으로 결정할 도메인별 용어조사표를 <표 7>의 형식에 따라 작성한다. 도메인별 용어조사표는 해당 온톨로지에서 지식으로 표현되어야 할 용어들과 그의 동의어, 약어, 용어 설명과 용어의 유형 등에 대한 개략적인 기술이다. 이 용어를 토대로 <표 8>의 도메인별 개념사전을 작성한다. 이때 각 개념이 가져야 할 속성과 속성의 유형 그리고 개념에 속하는 인스턴스와 크기 등에 대해 기술한다. 이 단계는 용어와 개념에 대한 정확한 지식표현의 기초가 된다.

3.2.4 단계 4 : 온톨로지 내부 구조 설계 단계

이 단계는 개념 계층, 관계, 규칙에 대한 구체적인 정의의 위해 클래스의 속성, 제한규칙, 관계의 형태를 분석하고 도메인의 논리구조를 분석하여 규칙과 공리를 추출한 뒤 수식

<표 9> 도메인 온톨로지별 개념속성표

도메인 온톨로지 개념 속성표					
글로벌 온톨로지명	이미지 지식 기반 온톨로지				
도메인 온톨로지명	양(sheep)				
개념명	domestic sheep				
속성명	속성 종류	속성 유형	측정 단위	정밀도	제한
hasLocation	Data	String	-	-	3
hasColor	Data	String	-	-	3
hasGenusName	Object	동물분류	-	-	1
.....
개념명	wild sheep				

<표 10> 도메인별 규칙공리표

도메인 규칙 공리표	
온톨로지명	이미지 지식 기반 온톨로지
도메인 온톨로지명	용어(terminology)
개념명	LifeStage
공리명	sameLifeStage
설명	lamb과 동일한 양의 성장 단계를 나타내는 용어를 추론한다
표현식	$sameLifeStage(?x,?y) \wedge sameLifeStage(?y,?z)$ $\wedge differentFrom(?x,?y)$ $\Rightarrow sameLifeStage(?x,?z)$
관련 개념	Sheep, Actor, LifeStage
관련 속성	sameLifeStage, differentFrom, LifeStage(x) ...

화 한다. 개념과 추론에 대한 기본 표현은 서술논리의 용어적 공리와 선언적 공리 구조를 이용하고 규칙 추론의 표현은 SWRL의 표기법을 이용한다.

이러한 작업을 위해 <표 9>의 도메인 온톨로지별 개념속성표를 작성하여 개념별로 구성되는 속성에 대한 정보를 명확하게 기술한다. 각 개념에 대한 더욱 구체화된 정보를 표현하기 위해서 도메인 개념 속성표, 도메인별 개념 인스턴스표를 기반으로 <표 10>의 도메인별 규칙공리표를 작성한다. 도메인별 규칙 공리표의 정확한 기술은 온톨로지 내의 지식 추론과 온톨로지 간 지식 이입이나 공유를 위해 중요하다.

3.2.5 단계 5 : 온톨로지 생성과 편집 단계

이 단계에서는 온톨로지 편집 툴을 이용하여 전 단계에서 설계된 명세에 따라 실제 온톨로지를 작성한다. 도메인 온톨로지간 연결과 확장이 필요할 때는 온톨로지 연결과 확장 작업을 하고 온톨로지의 인스턴스를 생성한 뒤 온톨로지를 검증한다. 이 단계의 작업을 위한 세부 과정을 기술하면 다음과 같다.

첫째, 온톨로지를 구축하기 위해서는 먼저 온톨로지 편집 툴을 선택한다. 온톨로지 구축 라이프사이클을 모두 지원하는 온톨로지 편집툴은 없지만 설계된 온톨로지의 지식표현과 추론 범위에 적합한 툴로 선택한다.

둘째, 온톨로지 편집툴의 제공 기능에 따라 온톨로지가 저장될 물리적 공간설계를 하여 저장될 온톨로지의 프로젝트와 작업 공간 등을 실제 구축한다.

셋째, 실제 구축된 온톨로지의 일관성(consistency) 검사를 시행하여 온톨로지의 클래스, 속성, 인스턴스, 규칙 등의 추론 결과를 통해 온톨로지의 유효성과 성능 등을 검증한다.

3.2.6 단계 6 : 온톨로지 배포/유지 보수

온톨로지 구축의 마지막 단계로서, [단계 5]까지 완성된 온톨로지는 웹상에 온톨로지의 전체 구조와 정보를 배포한다. 이 같은 온톨로지의 배포를 통해 시맨틱 웹의 특징 중 하나인 온톨로지의 공유나 재사용의 기능을 제공할 수 있다. 또한 변화하는 지식이나 환경 그리고 사용자의 추가적인 요구에 부응할 수 있도록 온톨로지의 지속적인 보완과 유지가 이뤄진다.

온톨로지의 원활한 유지 보수를 위해 이 단계에서는 앞에서 작성된 각각의 형식 문서의 정보를 기반으로 온톨로지와 관련된 <표 11>과 같이 시맨틱 웹 응용시스템 평가표를 작성한다. 이를 통해 온톨로지에 표현된 지식의 일관성과 무결성 그리고 지속성과 시맨틱 웹 응용시스템과의 활용성 등을 높일 수 있다.

<표 12>는 본 연구에서 제안한 웹 온톨로지 구축 기법의 각 단계에서 작성되고 문서화 되어야 하는 형식 문서들이며, 이 문서들은 구축된 웹 온톨로지의 유지보수 및 정보 제공을 위해 활용된다.

4. 실험

4.1 웹 온톨로지 구축 실험

본 연구의 실험을 위해 3장에서 제안된 웹 온톨로지 구축 기법에 따라 지식 기반 온톨로지인 동물분류온톨로지, 양온톨로지, 용어온톨로지를 구축하고 응용 인터페이스 온톨로지인 응용프레임온톨로지를 구축한다. 본 실험을 위해 사용한 온톨로지 개발 툴은 2006년 발표한 TopBraidComposer [22]이며, 추론 검증을 위한 프로그램은 TopBraidComposer에 플러그인(plug-in) 되어 있는 Inference-Rule 명령어를 사용한다.

설계와 구축시 앞장에서 제시한 형식화된 문서들을 각 단계에 맞춰 작성하고, 이의 내용을 적용하여 서로 다른 지식 영역을 표현한 도메인 온톨로지들의 지식을 추론한다.

4.1.1 동물분류온톨로지 구축

동물분류온톨로지는 린네가 창시한 생물분류학의 모든 시소러스를 표현하는 광범위한 온톨로지가 아니고 동물의 상하 계층을 중심으로 각 계층에 해당하는 동물학명과 이와 연관되는 양(sheep)의 동물학명에 대한 지식을 추론하기 위해 구축된다.

(그림 5)는 기술된 구축 목적에 따라 동물분류온톨로지에 구축된 개념과 속성들의 전체 구조와 관련된 인스턴스를 보여주는 그래프이다. (그림 5)의 이해를 돕기 위해 (그림 6)은 동물분류온톨로지의 개념 계층 중 “Kingdom(계)”을 중심으로 각 개념들의 연관성을 자세하게 보여준다. 이에 대한 설명은 다음과 같다.

- ①의 음영원은 클래스를 의미하고 음영원 오른쪽은 클래스 이름이다. 그리고 클래스 이름 하단의 사각형에는 클래스에 정의된 다양한 정보들을 보여준다.
- ②의 마름모는 해당 클래스의 인스턴스를 표시한다. 마름

<표 11> 온톨로지 기반 응용 시스템 평가표

온톨로지 기반 응용 시스템 평가표		
번호 :	작성일 :	
응용 시스템명	지능형 이미지 검색 시스템	
응용 시스템 목적	시스템 실행 결과	비고
검색 문장을 이용해 이미지의 내용을 중심으로 이미지를 검색한다.		
적격 질의 목록	추론 결과	비고
lamb과 동일한 의미를 가진 이미지는?	slink	
merino sheep 의 종명은?	ovis aries	
.....
응용시스템과 온톨로지 요소 평가		
주 제	확 인	비고
Ontology Language	OWL, RDF	
Rule Language	SWRL, DL	
Language Extension	Java, JSP	
Ontology Server Architecture	Apach, Tomcat	
Ontology Server Demonstrator	지능형 이미지 검색	

<표 12> 구축 단계별 작성문서

구축 단계	작성 문서	비 고
온톨로지 구축 목적	1. 온톨로지 개요서 2. 도메인 온톨로지 요약표 3. 적격 질의 목록표	
온톨로지 전체 구조 설정	1. 온톨로지 명세서	
온톨로지 정보 확보 및 분석	1. 도메인별 용어 조사표 2. 도메인별 개념 사전	
온톨로지 내부 구조 설계	1. 도메인별 개념 속성표 2. 도메인별 개념-속성 공리표 3. 도메인별 규칙 공리표	
온톨로지 배포/유지보수	1. 온톨로지 기반 응용시스템 평가표	

모 오른쪽은 인스턴스 이름이다. 이름 하단 사각형에는 인스턴스의 값들이 표시된다.

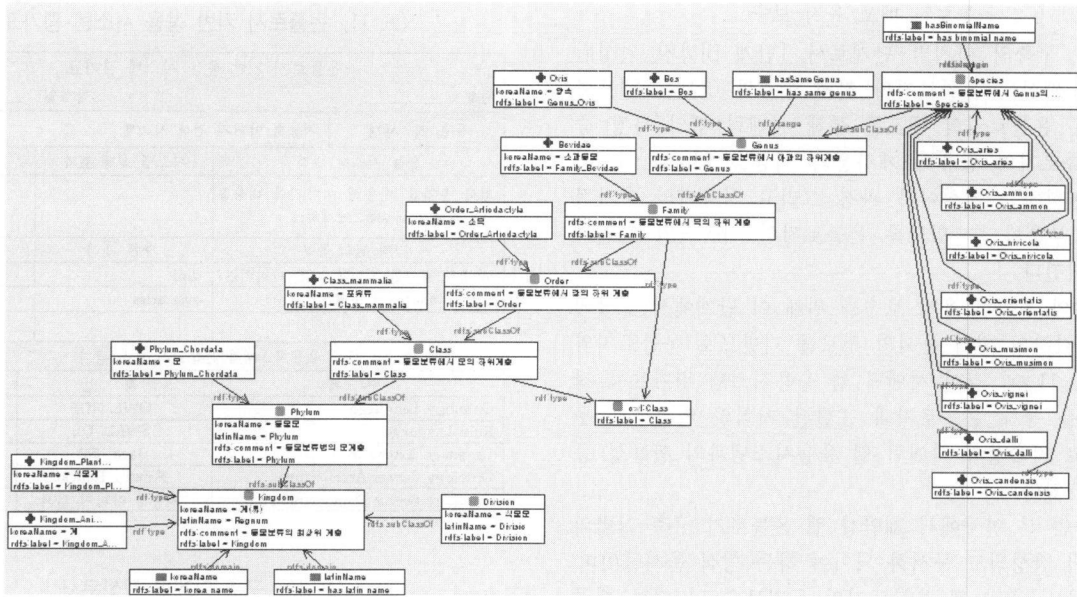
- ③의 직사각형은 데이터 프로퍼터를 표시한다.
- ④의 화살표는 술어(predicate)로서 가리키는 방향이 중요하다. 상단 “Phylum” 클래스에서 아래 “Kingdom” 클래스로 화살표 방향이 이어진다. 이때 문자열 “rdfs: subclassOf”가 화살표와 같이 표시되어 있는데 이는 “Phylum 클래스는 Kingdom 클래스의 서브클래스이다”를 의미한다.
- ⑤의 화살표에는 “rdf:type” 속성이 같이 표시되어 있는데 이는 “Kingdom_Plant 인스턴스는 Kingdom 클래스 타입을 갖는다”는 의미이다.
- ⑥의 화살표에는 “rdfs: domain”이 표시되어 있다. 이는 “속성 latinName의 도메인은 Kingdom 클래스”라는 것을 의미한다.

(그림 7)은 (그림 5)의 그래프가 OWL 형태의 코드로 표현되어 지식기반 저장소에 저장된 파일의 일부이다.

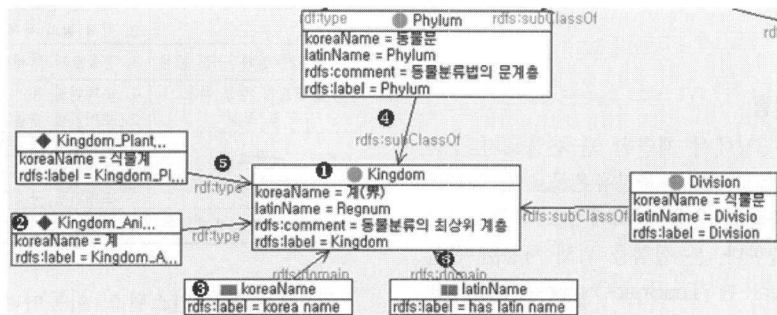
4.1.2 양온톨로지 구축

(그림 8)은 동물의 한 종인 양(sheep)에 대한 지식을 표현한 도메인 온톨로지 그래프이다. 양에 대한 기본 정보는 Wikipedia 사이트의 양정보를 이용한다.

‘Sheep’을 루트개념으로 ‘Domestic Sheep’ 개념과 ‘Wild Sheep’ 개념으로 분류되며 각 하위 개념과 개념간 속성 관계의 정보



(그림 5) 동물분류온톨로지 그래프



(그림 6) "Kingdom" 개념을 중심으로 한 개념간의 관계

와 인스턴스 정보를 확인할 수 있다. 이 온톨로지 또한 OWL 코드 형태로 저장된다. 양온톨로지의 개념 계층의 인스턴스는 동물분류온톨로지의 지식과 연관되어 추론에 활용된다.

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns="http://www.owl-ontologies.com/Zoological.owl#"
  xmlns:rdfs="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdf="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:daml="http://www.daml.org/2001/03/daml-rdf#"
  xml:base="http://www.owl-ontologies.com/Zoological.owl">
  <owl:Ontology rdf:about="">
    <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Created by Kim, Su Kyoung</owl:versionInfo>
  </owl:Ontology>

  .....

  <owl:Class rdf:ID="Phylum">
    <koreaName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">동물문</koreaName>
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Kingdom"/>
    </rdfs:subClassOf>
    <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Phylum</rdfs:label>
    <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">동물분류법의 문계층</rdfs:comment>
    <latinName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Phylum</latinName>
  </owl:Class>

  <owl:Class rdf:ID="Class">
    <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">동물분류에서 문의 하위계층</rdfs:comment>
    <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Class</rdfs:label>
    <rdfs:subClassOf rdf:resource="Phylum"/>
  </owl:Class>

  <owl:Class rdf:ID="Species">
    <owl:disjointWith rdf:resource="Genus"/>
    <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">동물분류에서 Genus의 하위 계층</rdfs:comment>
    <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Species</rdfs:label>
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Genus"/>
    </rdfs:subClassOf>
  </owl:Class>
</rdf:RDF>
```

(그림 7) OWL 형태의 동물분류온톨로지 파일

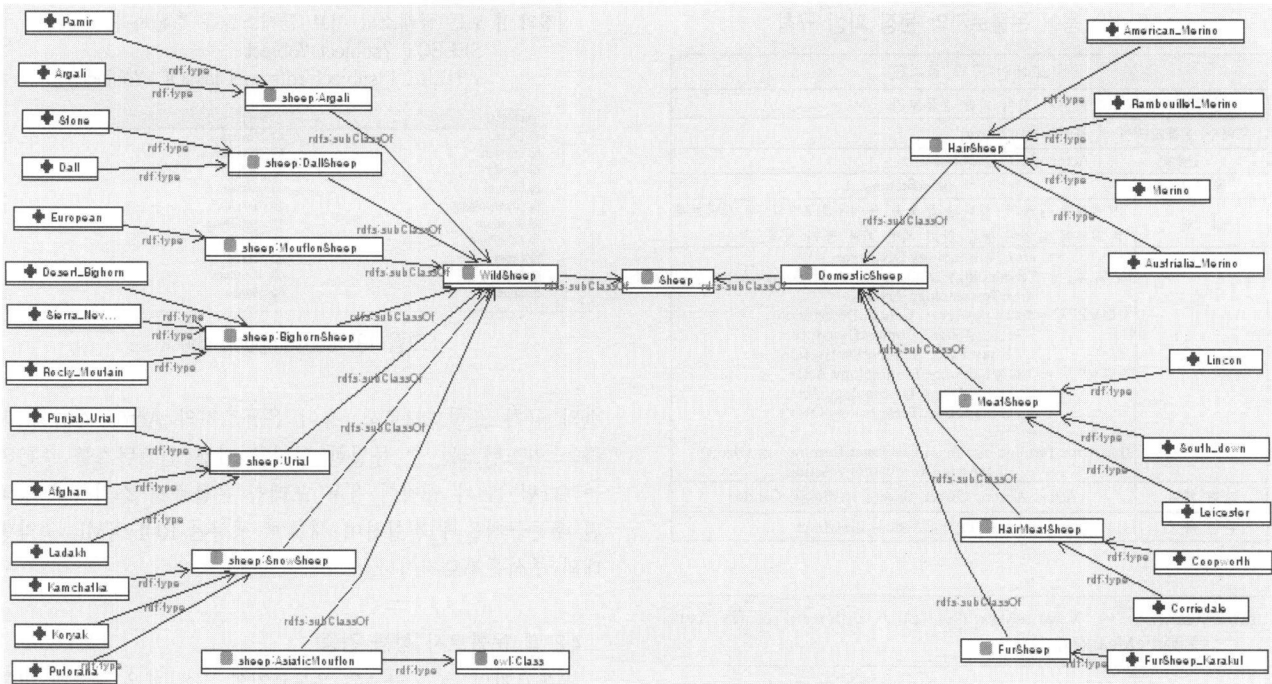
4.1.3 용어온톨로지 구축

용어온톨로지의 구조는 Audrey M.Tam[23]의 연구에서 제안된 Agent 개념 구조 분석한 결과를 실험 온톨로지의 지식표현과 목적에 맞춰 재설계한 뒤 본 논문의 구축 기법에 따라 구축한다.

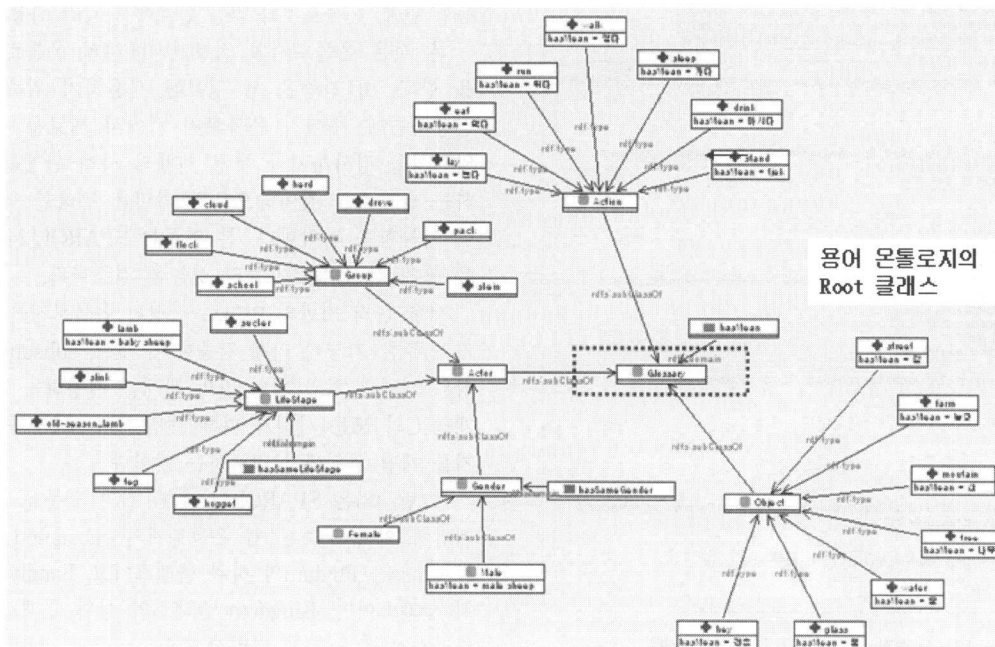
(그림 9)는 동물이나 양에 대한 다양한 용어에 대한 개념과 속성을 표현한 용어온톨로지 그래프이다. 용어온톨로지는 시맨틱 웹 응용에서 입력된 검색 문장의 각 단어를 주어-술어-목적어의 형태로 파싱한 후, <표 13>의 공리규칙표의 "parseStatement" 공리 표현식에 따라 온톨로지 내부에 정의된 추론 규칙에 의해 각 도메인 온톨로지들의 지식을 추론하는 기준 단어 역할을 한다.

4.1.4 응용프레임온톨로지 구축

(그림 10)은 웹 온톨로지의 특징에 따라 지식 기반 도메인 온톨로지의 지식을 추론 규칙에 따라 추론하여 새롭게 구축되는 응용프레임온톨로지 스키마 그래프이다. (그림 10)의 내부에 표현된 (그림 11)과 같은 SWRL 기반의 임의 추론 규칙이나 서술 논리의 선언적 지식 표현들의 틀인 스키마에 의해 (그림 12)의 결과와 같이 주석온톨로지가 생성된

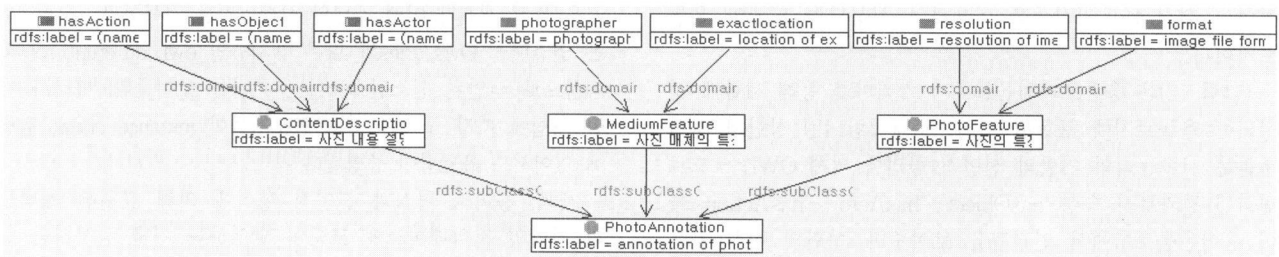


(그림 8) 양 온톨로지 그래프



용어 온톨로지의 Root 클래스

(그림 9) 용어 온톨로지의 그래프



(그림 10) 응용 프레임 온톨로지 스키마

〈표 13〉 용어 온톨로지의 문장 파싱 규칙

도메인 규칙 공리표	
온톨로지명	이미지 지식 기반 온톨로지
도메인 온톨로지명	용어(terminology)
개념명	Actor, Action, Object
공리명	parseStatement
설명	시맨틱 웹 용어에 입력된 문장을 용어 온톨로지의 각 개념들과 비교하여 동의어 등을 찾기 위한 기본 공리 정의
표현식	$iC \Rightarrow \text{lambTerminology.LifeStage}$ $C_1 \sqsubseteq C_2 \Rightarrow \text{Terminology.LifeStage} \sqsubseteq \text{Terminology.Actor}$ $\text{lambTerminology.Actor}$ $T \sqsubseteq \forall P.C \Rightarrow T \sqsubseteq \forall \text{hasActor.ContentDescription}$ $\quad \sqcup \text{hasAction.ContentDescription}$ $\quad \sqcup \text{hasObject.ContentDescription}$ $T \sqsubseteq \forall P.C \Rightarrow T \sqsubseteq \forall \text{hasActor.Terminology.Actor}$ $\quad \sqcup \text{hasAction.Terminology.Action}$ $\quad \sqcup \text{hasObject.Terminology.Object}$ $\langle i_1, i_2 \rangle : P =$ $\langle \text{hasActor.Terminology.Actor}, \text{hasObject.Terminology.Object} \rangle$ $\quad : \text{hasAction.Terminology.Action}$
관련 개념	Actor, Action, Object, Sheep, LifeStage, Gender
관련 속성	hasActor, hasAction, hasObject ...

```
sameMean(?x, ?y) ^ sameMean(?y, ?z) ^ differentFrom(?x, ?y)
=> sameMean(?x, ?z)
```

(그림 11) SWRL 기반의 임의추론규칙 일부

```
<PhotoAnnotation rdf:ID="PhotoAnnotation.1">
  <ClassObject rdf:resource="http://www.owl-ontologies.com/Terminology.owl#class"/>
  <ClassActor rdf:resource="http://www.owl-ontologies.com/Terminology.owl#lamb"/>
  <ClassAction rdf:resource="http://www.owl-ontologies.com/Terminology.owl#eat"/>
</PhotoAnnotation>

<swrl:imp rdf:ID="Cq.1">
  <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema:string">동일 hasActor 추론</rdfs:label>
  <swrl:body>
    <swrl:AtomList>
      <rdf:first>
        <swrl:ClassAtom>
          <swrl:classPredicate rdf:resource="http://www.owl-ontologies.com/Terminology.owl#Actor"/>
          <swrl:argument1 rdf:resource="#x"/>
        </swrl:ClassAtom>
      </rdf:first>
      <rdf:rest>
        <swrl:AtomList>
          <rdf:first>
            <swrl:IndividualPropertyAtom>
              <swrl:propertyPredicate
                rdf:resource="http://www.owl-ontologies.com/Terminology.owl#hasSameMean"/>
              <swrl:argument1 rdf:resource="#x"/>
              <swrl:argument2>
                <swrl:Variable rdf:ID="y"/>
              </swrl:argument2>
            </swrl:IndividualPropertyAtom>
          </rdf:first>
          <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
        </swrl:AtomList>
      </rdf:rest>
    </swrl:AtomList>
  </swrl:body>
  <swrl:head>
    <swrl:AtomList>
      <rdf:first>
        <swrl:ClassAtom>
          <swrl:classPredicate rdf:resource="http://www.owl-ontologies.com/Terminology.owl#Actor"/>
          <swrl:argument1 rdf:resource="#y"/>
        </swrl:ClassAtom>
      </rdf:first>
      <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
    </swrl:AtomList>
  </swrl:head>
</swrl:imp>
```

(그림 12) 새롭게 재생성된 주석온톨로지

다. 이 온톨로지는 미리 구축된 도메인 온톨로지들의 명시적이고 형식적인 지식들을 이용하여 생성되는 도메인 온톨로지이다.

(그림 12)은 응용프레임온톨로지 스키마를 통해 새롭게 생성되는 온톨로지에 표현된 지식으로 (그림 12) 상단의 점선 부분은 서술논리의 기반의 선언적 공리에 의해 OWL 코드로 변환된 결과로서 속성 'hasObject', 'hasActor', 'hasAction' 들이 용어온톨로지서 연결됨을 보여주며, 하단의 점선 부분은 (그림 11)에 정의된 임의 규칙에 따라 생성된 SWRL 기

[질의 1] 모든 클래스의 서브 클래스들을 추론하라.
 SELECT ?subject ?object
 WHERE (?subject rdfs:subClassOf ?object)

subject	object
● Class	● Phylum
● Division	● Kingdom
● Family	● Order
● Genus	● Family
● Green.algae	● Division
● Kingdom	● owl:Thing
● LandPlant	● Division
● Order	● Class
● Phylum	● Kingdom
● Species	● Genus

(그림 13) subsumption check

반의 규칙 표현코드로서, 용어 온톨로지의 Actor 개념에 속하는 인스턴스들 중 동일한 의미를 가진 인스턴스를 추론하는 (그림 11)의 규칙이 응용 프레임 온톨로지 스키마에 정의된 추론규칙들과 결합하여 새롭게 생성된 RDF/XML 축약형태의 주석온톨로지이다.

4.2 웹 온톨로지 성능 검증

온톨로지의 추론 기능은 온톨로지의 질에 매우 중요한 역할을 하며, 추론에 대한 결과는 온톨로지의 검증의 효과를 보여준다. 이러한 추론 기능은 온톨로지의 설계 단계, 온톨로지의 통합 단계, 온톨로지의 적용 단계에서 각각 사용될 수 있다[4].

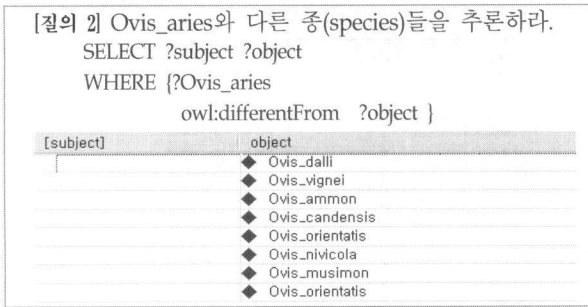
본 절은 온톨로지의 적용단계에서의 온톨로지 검증 실험을 한다. 이 검증은 웹 정보를 사용하기 위해 추론 기능을 검증하는데, 주어진 온톨로지 하에서 새로운 사실들에 대한 일관성을 검사하거나 특정 개체의 사례 관계를 파악하여 속하는 클래스를 찾아줌으로써 의미를 고려한 웹 정보의 활용이 가능함을 파악한다. 본 연구는 SPARQL[24]의 추론 구문을 통해 각 온톨로지들의 성능을 검증한다.

서술 논리 기반의 언어로 작성된 온톨로지는 이를 위해 세 가지 추론 기능에 의해 검증한다. 첫째, subsumption check는 상위 개념과 하위 개념 간의 관계를 결정하는 기능이다. 어떤 개념 C의 모든 사례가 다른 개념 D의 모든 사례에 포함되는지를 개념 차원에서 검증하는 것이다.

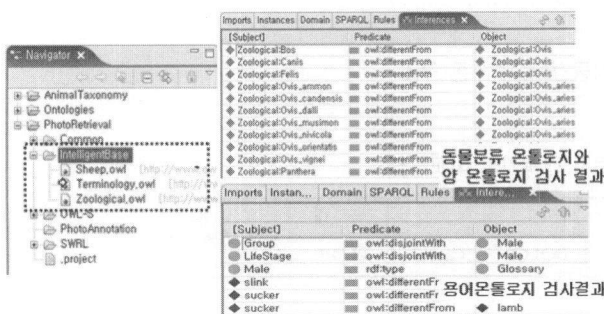
(그림 13)은 SPARQL의 추론 질의 구문을 이용하여 제안 기법에 의해 구축된 웹 온톨로지의 subsumption check 결과로 Class는 Phylum의 하위 클래스이고, Family는 Order의 하위 클래스이며, Kingdom 클래스의 상위 클래스는 owl:Thing 클래스가 추론된다. 따라서 subsumption check 검증 결과가 유효함이 증명된다.

둘째, instance check는 사례 관계를 결정하는 기능이다. 이것은 특정 개체가 어떤 개념의 사례인지를 검사한다. (그림 14)은 인스턴스 Ovis_aries와 다른 객체중인 owl:differentFrom의 속성으로 선언된 인스턴스들이 추론됨을 보여준다. 따라서 제안 기법에 따라 구축된 웹 온톨로지들이 instance check 검증에 있어서도 유효함이 증명된다.

셋째, 용어적 형식과 선언적 형식에 의해 표현된 개념들의 집합인 지식베이스에 모순이 발생하는지를 검사하는 일관성 검사로 (그림 15)는 본 연구에서 적용한 온톨로지 개



(그림14) instance check 결과



(그림 15) 온톨로지들의 일관성 check 결과

발 도구인 TopBraidComposer의 플러그인인 Racer가 제공하는 Consistency Inference 명령어를 통해 구축된 온톨로지들의 일관성 검사 결과로 지식베이스의 모순 없이 규칙이 적용된 모든 지식들이 출력됨을 보여준다.

4.3 기존 온톨로지 구축기법과의 비교 분석

본 실험을 위해 구축된 온톨로지의 목적은 단순히 온톨로지의 구축을 하는 것이 아니라 제안된 웹 온톨로지 구축기법을 사용할 때 시맨틱 웹 응용에서 요구되는 기술적 요인과 특징에 적합함을 증명하는데 있다. 특히 지식표현에 있어 형식적이고 명시적인 표현을 문서화 하고 이를 통해 구축된 온톨로지의 추론 결과가 지식의 공유와 재사용을 제공할 수 있는데 목적이 있다.

현재 온톨로지 구축기법에 대한 평가 기준과 적절한 방법에 대한 연구가 부족한 상황에서 온톨로지 구축기법들이 시맨틱 웹 응용의 기술적 요소와 특징에 어느 정도 적합한지를 분석하기 위해, 본 연구는 2.2절에서 제안한 기술적 요인과 <표 2>의 항목들을 기준으로 시맨틱 웹 응용의 기반 기술로서 웹 온톨로지가 구축되었는지를 평가하는 항목을 <표 14>의 “웹 온톨로지 특징”으로 정하고, 이를 기준으로 비교 대상이 되는 온톨로지 구축기법들과 본 논문이 제안한 웹 온톨로지 구축 기법과의 비교 결과를 기술 한다.

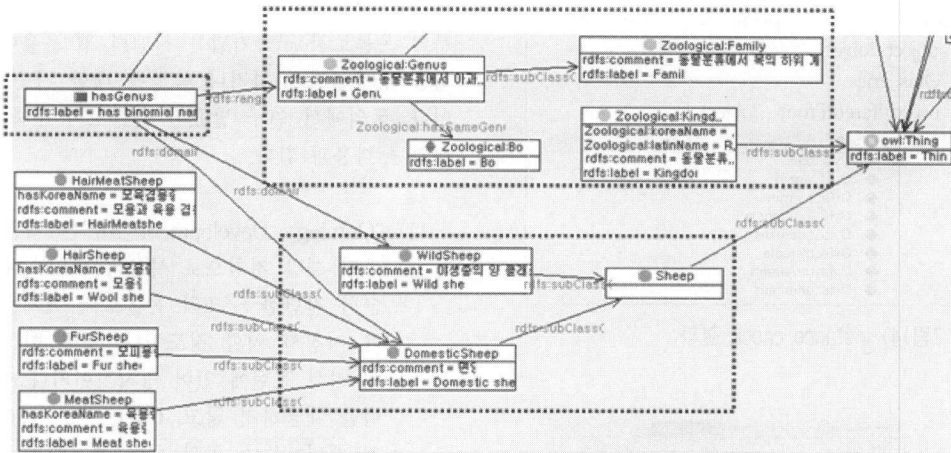
먼저, 구축기법을 비교하기 위해 비교 대상이 될 온톨로지 구축기법을 선정한다. 선정 기준으로는 첫째, 온톨로지 구축기법이 비교적 최근에 개발된 것이거나, 둘째, 시맨틱 웹이 제안된 이후 발표된 기법이거나, 셋째, Ontology Development 101과 OTKM 같이 온톨로지 공학자가 많이 적용하는 구축기

법들이 대상이다. <표 14>의 분석 결과가 보여주듯이 대부분의 온톨로지 구축기법은 시맨틱 웹 응용에서 요구하는 사항을 만족하지 못하거나 측정에 대한 기준조차 없는 상황이다. 국외에서 연구 발표된 온톨로지 구축기법들을 분석한 결과는 다음과 같다.

- (1) “Ontology Development 101”은 많은 연구에서 온톨로지 개발 지침으로 사용하고 있으나, 초보자들이 단순히 간단한 구조의 온톨로지 예제를 이용해 따라하는 지침서 역할 정도이다. 또한 이 기법은 온톨로지 내 지식 표현에 있어 명시적이거나 형식적인 표현 방법을 제공하지 않고, OWL 코드 형식 기반의 온톨로지 개념, 속성, 공리 등만 제시한다. 또한 온톨로지의 재사용과 배포와 같은 사항은 고려되지 않았다.
- (2) 소프트웨어 공학 관점의 프로토타입 모델을 적용한 OTKM의 경우 대규모의 온톨로지를 개발하는 연구기관에서 온톨로지 구축기법으로 많이 사용하고, 안정적인 온톨로지 개발 기법으로 알려져 있다. 그러나 2002년에 해당 프로젝트가 종료되면서 새롭게 변화 제시된 웹 온톨로지의 다양한 요구는 적용시키지 못하였다. 웹 온톨로지 언어는 RDF 만을 사용하고, 온톨로지 데이터들은 주석 형태 위주로 저장되며, 도메인 온톨로지의 이입이나 확장에 대한 방법 제시가 없다.
- (3) DOLCE 구축기법은 2004년도에 웹 온톨로지들의 병합을 해결하기 위한 온톨로지 구축 기법으로 제시되었으나 지능형 서비스나 지식 추론을 위한 임의 규칙의 지식 표현 방법이나 형식적인 개발 과정을 제시하지 않았다.
- (4) Casual 구축기법은 웹 온톨로지 약어를 사용해 온톨로지를 간단히 구축하는 방법을 제시하였으나 지능형 서비스, 온톨로지 확장, 이입 그리고 형식화된 지식 표현 등에 대한 내용이 제시되지 않았다.

국내 연구로 2004년 한국전산원이 제안한 EOE (Evolving Ontology Engineering) 방법론[25]은 국내에서 온톨로지 개발에 대한 지침을 최초로 제안한 연구이다. 온톨로지 설계시 형식화된 문서 작성의 필요성을 제시하였으나 그에 대한 명확한 형식화 방법을 제시하지 못하고, 단순히 온톨로지 개발에 대한 방법만을 제시하여 시맨틱 웹 응용과의 연계 방법이나 임의규칙을 위한 지식표현 방법의 제공이 부족하다.

그러나 본 논문에서 제안한 구축방법은 <표 14>에서 제시한 대부분의 요구 조건을 만족하고 있다. 특히, 시맨틱 웹 응용이 갖춰야할 가장 중요한 요소인 추론과 지식 공유 그리고 재사용을 위해 명시적이고 형식적인 지식표현을 기반으로 (그림 16)에서 보여주듯 온톨로지들 간의 지식 추론 구조가 생성된다. 아래 질의는 이같은 지식 추론 구조에 의해 양온톨로지의 인스턴스인 “Merino”의 학명을 추론하는 질의로서 동물분류온톨로지서 해당되는 학명인 “Ovis”가 출력된다.



(그림 16) 두 도메인 온톨로지들간 지식 추론 구조

<표 14> 제안기법과 기존 온톨로지 구축기법 비교 결과

웹 온톨로지 특징	국 외					제안 기법
	Development 101	OTKM	DOLCE	Casual	EOE	
지능형 서비스를 위한 표현 방법	○	○	-	△	○	○
웹 환경 구현 방법 제시	-	○	-	-	-	-
온톨로지 연결과 재사용 방법	-	-	○	-	-	○
웹 온톨로지 언어 사용	○	○	○	△	○	○
추론 엔진 제시	-	-	○	-	-	-
형식화를 위한 문서 제시	-	○	-	-	○	○
도메인 온톨로지별 구축	-	-	○	○	△	○
공유를 위한 온톨로지 배포	-	○	-	-	-	-
임의 규칙 언어 사용 제안	-	-	-	-	-	○
웹 온톨로지 구축	-	○	-	○	○	○

* ○ : 만족, △ : 부분 만족, - : 측정 불가

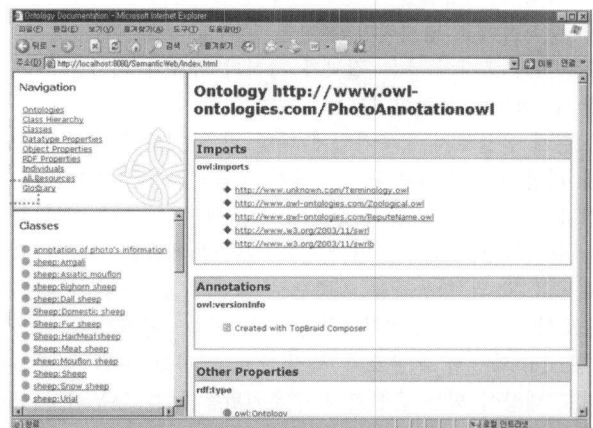
[질의] 인스턴스 Merino의 속(Genus) 명을 검색하라.
 SELECT ?subject ?object
 WHERE {?Merino :hasGenus ?object}

[결과]

[subject]	object
	◆ Zoological:Ovis

본 연구에서 제안한 기법의 또 다른 특징은 규칙언어나 선언적 지식의 표현이 온톨로지 내부 설계 단계에서 명확하게 기술되면 (그림 12)와 같이 기존 온톨로지를 기반으로 한 새로운 온톨로지가 재생성될 수 있는 점이다.

또한 제안된 웹 온톨로지 구축기법의 마지막 단계인 “온톨로지의 배포/유지보수”에 따라 (그림 17)과 같은 실험 온톨로지들이 웹에 배포된다. 이 과정에 의해 개방된 온톨로지들은 다른 온톨로지와의 관계에서도 쉽게 활용되며, 본 연구에 실험 구축된 온톨로지의 구조와 내용에 대한 지식이 부족한 사람도 이 배포 정보를 통해 온톨로지를 재활용하거나 공유할 수 있는 장점을 제공한다.



(그림 17) 구축된 실험 온톨로지의 배포

5. 결론 및 향후 연구 방향

본 연구는 시맨틱 웹 응용의 발전을 위해 시맨틱 웹 응용의 기반 기술로 웹 온톨로지가 갖춰야할 기술적 요소와 지식

표현 방법을 조사 분석하여 시맨틱 웹 응용에 적합한 웹 온톨로지 구축기법을 제안하였다. 제안된 구축기법에 따라 다양한 웹 온톨로지들을 구축하여 웹 온톨로지가 갖춰야하는 지식표현이나 추론의 검증을 통해 제안한 웹 온톨로지 구축기법이 시맨틱 웹 응용의 요구 조건을 만족하였다. 특히 웹 온톨로지들 간의 지식 공유나 재활용 그리고 기 구축된 온톨로지들을 추론 규칙을 이용하여 새로운 도메인 온톨로지 생성이 가능함도 제시하여 기존의 온톨로지 구축기법과의 비교 분석을 통해 시맨틱 웹 응용에 적합한 구축기법임을 증명하였다.

이러한 연구에 덧붙여 수동적 웹 온톨로지 구축을 보완할 수 있는 다양한 연구를 진행하고자 한다. 특히, 현재 HTML로 표현된 웹상의 정보들을 자동으로 웹 온톨로지의 지식 표현 단위로 변환하기 위한 연구와 온톨로지 학습(Ontology Learning)에 대한 깊이 있는 연구를 통해 기계가 자동으로 웹 온톨로지들간의 연관된 지식들의 추론규칙을 생성하고 이를 기반으로 온톨로지를 연결하는 방안 연구가 필요하다고 본다.

참 고 문 헌

[1] Michel Klein, Ubbo Visser. "Semantic Web Challenge 2003," IEEE Computer Society, IEEE INTELLIGENT SYSTEM, pp.31-33, 2004. 5.

[2] Tim Burners Lee. "Artificial Intelligence and the Semantic Web," website. <http://www.w3.org/2006/Talks/0718-aaai-tbl/>, AAAI, 2006.7.

[3] 한국전산원. "시맨틱 웹 발전 방향 및 표준화 개발전략 연구," pp.111-115, 2005. 12.

[4] Benjamin N. Grosz, Ian Horrocks. "Description Logic Programs : Combining Logic Programs with Description Logic," Working Paper, version of Nov., 21, 2002.

[5] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosz, Mike Dean. "SWRL: A Semantic Web Rule Language Combining OWL and RuleML," website. <http://www.daml.org/2003/11/swrl/>, 2003.11.

[6] Mike Dean. "Semantic Web Rules: Covering the Use Case," Third International Workshop, RuleML, Proceedings, pp.1-5, 2004.

[7] 옥철영. "한국어정보처리와 온톨로지," 한국어정보처리연구회 동계 튜토리얼 자료집, 2004.

[8] 이현실. "온톨로지 기반 한의학 처방 지식관리시스템 설계에 관한 연구," 중앙대 대학원, p.183, 2003.

[9] 이경일, 최광선. "산업적 관점에서의 시맨틱 기술," 한국정보과학회 학회지 제24권 제4호, pp.45-50, 2006. 4.

[10] Deborah L. McGuinness, Frank van Harmelen. "OWL Web Ontology Language Overview", W3C Recommendation, 2004.

[11] Jena2.1, jena.sourceforge.net/documentation.html

[12] Kenneth Baclawski. "UML for ontology development," The Knowledge Engineering Review archive, Vol. 17, Issue 1. pp.61-64, 2002. 3.

[13] OpenCyc.org. "OpenCyc: the project," website. <http://sourceforge.net/projects/opencyc/>, 2006.

[14] KACTUS Consortium. "KACTUS ESPRIT Project 8145," website. <http://www.swi.psy.uva.nl/projects/NewKACTUS/home.html>, 1996.

[15] Enterprise Integration Laboratory. "TOVE Ontology Project," website. http://www.eil.utoronto.ca/enterprise_modelling/tove/index.html, 2002.

[16] Dieter Fensel et al. "On-To-Knowledge: Ontology- Based Tools for Knowledge Management,," 2000.

[17] Natalya F. Noy, Deborah L. McGuinness. "Ontology Development 101:A Guide to Creating Your First Ontology,," pp.3-9, 2001.

[18] Nicola Guarino. "DOLCE: a Descriptive Ontology for Linguistic and Cognitive Engineering,," <http://www.loa-cnr.it/DOLCE.html>, 2006.

[19] Aditya Kalyanpur, Nada Hashmi, Jenifer Globeck, Bijan Parsia. "Lifecycle of a Casual Web Ontology Development Process,," Application Design, Development and Implementation Issues in the Semantic Web, 2004. 1.

[20] M. Fernández, A. Gómez-Pérez, N. Juristo. "METHONTOLOGY : From ontological art towards ontological engineering,," In Working Notes of the AAAI Spring Symposium on Ontological Engineering, Stanford, CA, Stanford University, AAAI Press. 1997.

[21] C.Masolo, S. Borp, A. Gangemi, N. Guarino, A. Oltramari. "Wonder Web deliverable D17," intermediate report 2.0, ISTC-CNR, 2002.

[22] Holger Knublauch. "TopBraidComposer Development Guide Online Help," website. <http://www.topbraid.com/documentation>, 2006.

[23] Audrey M. Tam, Clement H.C. Leung. "Structured Natural-Language Descriptions for Semantic Content Retrieval of Visual Data," JASIS, 2001. 9.

[24] Eric Prud'hommeaux, Andy Seaborne. "SPARQL Query Language for RDF," Editors working draft. website. <http://www.w3.org/2001/sw/DataAccess/rq23/>, 2006.

[25] 한국전산원. "웹 온톨로지 개발지침 연구," pp.11-45, 2004. 12.



김 수 경

e-mail : kimsk@hanbat.ac.kr
 1993년 한밭대학교 전자계산학과(학사)
 1997년 충남대학교 교육대학원
 컴퓨터교육학과(교육학석사)
 2007년 한밭대학교 정보통신전문대학원
 컴퓨터공학과(공학박사)

관심분야: 시맨틱웹, 온톨로지, 에이전트 등



안 기 흥

e-mail : khahn@hanbat.ac.kr

1981년 성균관대학교 전자공학과(학사)

1985년 성균관대학교 대학원 전산학과(석사)

1993년 성균관대학교 대학원 전산학과

(이학박사)

1988년~1996년 대전산업대학교 조교수

1997년~현재 한밭대학교 컴퓨터공학과 교수

관심분야: 지식관리, 시맨틱웹, 온톨로지 등