

# 시멘틱 웹 기술을 활용한 초등학교 학습자료 검색 시스템

이희경<sup>0</sup>, 전우천  
서울여의도초등학교<sup>0</sup>, 서울교육대학교 컴퓨터교육과  
lakesis@edunet4u.net<sup>0</sup>, wocjun@snue.ac.kr

## Elementary Educational Contents Retrieval System Using Semantic Web Technology

Hee-Kyoung Lee<sup>0</sup>, Woo-Chun Jun

Seoul Youido Elementary School<sup>0</sup>, Dept. of Computer Education, Seoul National University of Education

### 요 약

웹의 활용이 보편화 되면서 웹을 통한 자료의 검색이 증가하고 있으나, 웹상의 방대한 자료 중에서 학습자가 꼭 필요한 학습자료를 찾는 것은 쉬운 일이 아니다. 검색엔진을 이용하면 원하는 정보를 어느 정도 찾을 수 있으나 사용자 의존적인 검색엔진의 특성상 결과가 만족스럽지 못한 경우도 있으며 연관이 없는 정보를 필터링하기 위해 최종적인 내용을 찾기까지 많은 시간을 낭비하는 경우가 많다. 이에 본 연구에서는 웹 자원의 의미정보를 구조화하여 정보의 효율적인 검색, 통합, 재사용을 가능하도록 하는 시멘틱 웹(Semantic Web) 기술을 활용하여 초등학교 학습자료에 적합한 온톨로지(Ontology)를 구축하여 이를 기반으로 초등학교 학습자료를 검색할 수 있는 시스템을 설계하고 구현하였다. 본 검색시스템의 특징은 다음과 같다. 첫째, 학습자료와 연관된 사용자 질의어를 보다 상세하게 입력받는다. 둘째, 사용자 질의어를 바탕으로 온톨로지에 질의하여 검색결과를 얻는다. 셋째, 검색하고자 하는 내용의 의미를 분석하여 요구된 의미에 적합한 자료만을 검색결과로 제시한다.

### 1. 서 론

웹은 팀 버너스리에 의해 1989년 처음 소개된 이래 짧은 시간 동안 괄목할만한 성장을 거듭해 왔다. 웹상에서는 쉽게 익힐 수 있는 간단한 HTML(HyperText Markup Language)을 이용하면 흩어져 있는 자원들을 한데 묶어 하나의 페이지로 보여줄 수도 있으며 관련된 다른 사이트로의 이동을 클릭 한번으로 쉽게 할 수 있어 산재된 정보자원들을 연결하여 주는 역할을 할 수 있었다. 또한 HTML문법을 모르더라도 손쉽게 웹자원을 생성할 수 있는 웹문서 편집기 등이 등장하면서 웹은 보다 폭넓은 사용자를 확보하고 있으며 또한 보다 많은 정보를 쏟아내게 되어 다양한 분야에서 활용되게 되었다.

사용자 증가로 웹상의 정보 또한 정보의 홍

수로 불릴 정도로 막대한 양이 됨으로써 개인이 필요한 정보를 쉽고 정확하게 찾고자 하는 요구가 늘어났고 자연스럽게 검색엔진이 생겨나게 되었다. 어떤 정보가 웹상에 있고, 그 정보의 정확한 위치가 무엇인지 모르더라도 서비스 제공자가 분류하여 놓은 목록 중에서 원하는 목록의 링크를 따라 찾아가거나 사용자가 필요로 하는 정보의 일부-문서의 제목 또는 문서내용의 일부-를 포함하는 자료들을 찾아주는 서비스를 통해 원하는 자료에 접근하기가 한층 쉬워졌다.

이러한 검색서비스가 각광을 받으면서 다양한 검색기술을 활용한 검색서비스가 생겨나 활용되었으나 날로 늘어가는 정보 속에서 사용자가 원하는 정보를 바로 찾기란 더 이상 쉬운 일이 아니다. 검색엔진을 이용한 검색결과 또한 방대한 양인 경우가 많아 이를 일일이 확인하며 원하는 정보를 필터링하는 수동

적인 작업이 필요하기 때문이다. 이러한 오늘날 검색엔진의 문제점은 모순되게도 보다 많은 범위와 다양한 자료검색이 가능한 검색엔진의 우수성으로 검색결과가 많아지기 때문에 발생하게 된다. 현재의 웹에서는 기계가 정보 자원을 읽을 수는 있으나 이해할 수는 없기 때문에 위와 같은 문제들이 발생한다. 이러한 문제를 해결하고자 제시되고 있는 차세대 웹이 바로 시맨틱 웹이다.

현재의 웹이 사람이 보고 이해하도록 표현하는 데에 초점을 맞추고 있다면 시맨틱 웹은 웹 문서에 잘 정의된 의미정보를 부여해 컴퓨터와 사람이 협동적으로 작업을 수행할 수 있도록 하는 역할을 하는 것으로 볼 수 있다[1]. 초등학교에서도 정보화의 물결을 타고 웹의 이용이 보편화되어 있는데, 초등학교 학습에 관한 정보검색을 할 때의 문제도 너무나 많은 검색결과가 나온다는 점과 원하지 않는 검색결과로 인해 시간을 낭비하게 된다는 점이다.

본 연구는 초등학교의 학습자료를 검색하고자 할 때 시맨틱 웹의 주요개념중 하나인 온톨로지를 기반으로 학습자료를 검색할 수 있는 시스템을 제안한다. 본 시스템의 특징은 다음과 같다. 첫째, 학습자료와 연관된 사용자 질의어를 보다 상세하게 입력받는다. 둘째, 사용자 질의어를 바탕으로 온톨로지에 질의하여 검색결과를 얻는다. 셋째, 검색하고자 하는 내용의 의미를 분석하여 요구된 의미에 부합되는 자료만을 검색결과로 제시한다.

본 논문의 구성은 다음과 같다. 먼저 1장에서 연구의 배경 및 목적을 제시하고, 2장에서는 시맨틱 웹과 시맨틱 웹을 구현하기 위한 온톨로지, 온톨로지언어 OWL (Web Ontology Language) 등의 기술적 배경과 국내 검색엔진에 대해 살펴보고 교육관련 온톨로지에 관련된 연구를 소개한다. 3장에서는 초등학교 학습자료 검색시스템의 기초가 되는 온톨로지 구축 과정을 소개하고, 4장에서는 온톨로지를 기반으로 하는 검색시스템 구현과정에 대해 소개한다. 마지막으로 5장에서는 연구 결과를 요약하고 향후 연구방향을 제시한다.

## 2. 이론적 배경

### 2.1 시맨틱 웹

시맨틱 웹은 웹상의 정보에 잘 정의된 의미를 부여함으로써 사람뿐만 아니라 컴퓨터도 쉽게 문서의 의미를 해석할 수 있도록 하여 컴퓨터를 이용한 정보의 검색 및 해석, 통합 등의 업무를 자동화하기 위한 목적으로 제안되었다[1]. W3C (World Wide Web Consortium)에서는 시맨틱 웹은 이미 웹상에 있는 데이터나 문서에 기계판독 가능한 기술을 추가한 것으로서 XML (eXtensible Markup Language), RDF (Resource Description Framework), 그리고 OWL에 의해 보다 용이하게 그리고 신뢰성이 높은 정보검색과 재이용이 가능하다고 소개하고 있다[2]. 즉, 시맨틱 웹은 기존의 웹을 대체하는 새로운 형태가 아니라 현재의 웹에서 좀더 진화된 형태의 웹이라고 할 수 있다. 현재의 웹은 초기의 웹에서 실현하기 어려웠던 기술들이 보편화되어 있는데, 과거시점에서 보면 현재의 웹이 시맨틱 웹이라고 할 수 있는 것이다.

최근 시맨틱 웹을 실현하기 위한 다양한 접근방법이 연구되고 있는데, HTML을 기반으로 한 현재의 웹을 개선하는 취지에서 보면 시맨틱 웹을 달성하기 위해 웹 프로토콜과 같은 하위 레벨의 개념을 정의하고 이 하위레벨을 이용하여 다음 레벨의 개념을 정의하는 계층구조를 설정하는 것이 일반적인 연구방법이다. 시맨틱 웹의 계층구조는 그림 1과 같다[3].

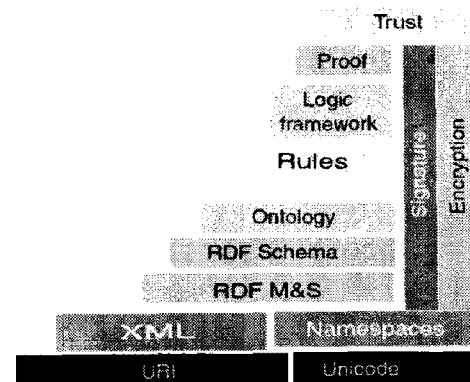


그림 1 시맨틱 웹의 계층구조

가장 하위에는 자원을 명시적으로 지칭하는 URI (Uniform Resource Identifier)와 UNICODE가 밀받침이 된다. 다음 층은 임의의 개념을 모듈방식으로 정의할 수 있는 XML과 Namespace이며, 그 다음으로 자원을 기술하기 위한 RDF와 RDF 스키마가 위치한다. 그 상위에는 온톨로지가 위치하고 규칙, 논리, 증명을 위한 기술 요소의 층들이 위치한다. 시맨틱 웹 분야에서는 기존의 웹 기술을 위해 이루어진 연구에 이어 RDF, 온톨로지에 대한 연구가 진행되고 있다. 그 위의 계층인 Rule과 Logic에서는 온톨로지에 기반하여 질의와 추론을 가능하게 한다. 이에 대해서는 인공지능의 추론연구를 밀받침으로 일부 연구가 이루어지고 있다. 또한 보다 상위 계층인 Proof와 Trust는 시맨틱 웹 정보의 신뢰성과 보안에 관한 내용으로서 아직 개념정도만 얘기되고 있으며 차후 연구과제로 제시되고 있다[3,4].

웹과 관련한 제반 기술과 관련한 연구로 인정받고 있는 W3C의 워킹그룹에 의해 시맨틱 웹 기술에 대한 연구가 꾸준히 이루어져 왔는데 2001년 활동을 시작한 Web-Ontology (WebOnt) Working Group은 2004년 5월 31일자로 종결되고, 2004년 Web Ontology Language (OWL)사이트가 개설되어 연구를 계속하고 있다. 국내에서도 정부차원에서 시맨틱 웹에 대한 연구가 시작되었으며 학계나 산업계에서 시맨틱 웹에 대한 관심이 높아지고 있다. 그러나 아직까지는 도입을 위한 기술적 검토단계로서 시맨틱 웹 기술을 활용하고 있지는 못하다. 기업들도 시맨틱 웹 기술이 대중화되기까지 시일이 필요할 것으로 보고 연구 개발에 대한 투자를 꺼리고 있는 실정이다.

## 2.2 온톨로지

시맨틱 웹의 구현을 위해 필요한 것은 주제 또는 도메인 영역에 따른 어휘를 개발하는 것이다. 이것이 바로 온톨로지의 역할이다.

온톨로지는 그리스어인 ontos[존재하는 것]와 logos[단어]에서 유래하였다. 온톨로지에

대한 정의는 여러 가지가 있지만 Gruber는 온톨로지를 “공유된 개념화에 대한 정형화되고 명시적인 명세”[5]라고 정의하였다. ‘공유’는 용어 사용이 이와 관련된 사람들의 합의에 의한다는 것으로서 객관적이고 명확한 성질을 가짐을 의미한다. ‘개념화’는 모델링의 모형이 세계가 된다는 것을 의미하며 ‘정형화’는 술어 논리와 같은 지식표현을 의미한다. ‘명시적’이라는 것은 유형과 개념 사용에 대한 규칙을 명확하게 정의함을 의미한다[6].

온톨로지는 특정 도메인의 지식을 명시적으로 표현하기 위한 개념들, 개념 사이의 관계, 개념의 속성 및 속성에 부여된 제약조건 및 객체들로 표현된다. 이 온톨로지를 통하여 특정 도메인의 단어를 공통으로 정의하고, 지식을 공유할 수 있게 된다.

온톨로지는 의료, 기계, 부동산, 재무 관리 등의 주제 또는 도메인 정보의 공유가 필요한 사람, 데이터베이스, 응용프로그램들에 의해 사용된다. 온톨로지는 기본개념을 기계가 처리할 수 있도록 정의함과 동시에 개념간의 관계도 정의함으로써 개별적 분야의 지식뿐만 아니라 분야를 넘어서는 지식도 기술할 수 있다. 현재 비교적 많이 이용되고 있는 데이터베이스는 하나의 응용프로그램에 적합하게 설계되어 적용되므로 재사용이 어려운데 비해 온톨로지는 다양한 응용프로그램에 적용할 수 있어 지식의 재사용이 가능하다는 특징을 갖는다[2].

온톨로지의 또 다른 큰 특징 중 하나는 추론기능의 제공이다[6]. 온톨로지에서는 개념과 개념 사이의 관계를 정의하는데, 관계술어 논리를 이용하며 온톨로지에 직접적으로 표현되지 않는 개념간의 관계를 추론을 통해 제시하거나 이에 대한 질의에 답을 제시할 수 있다. 이를 통해 사용자가 수동적으로 해야 했던 작업들이 기계를 통해서 상당 수준 이루어질 수 있게 된다. 따라서 온톨로지를 기반으로 하는 시스템에서는 현재까지는 불가능했던 질의 검색 및 추론작업이 가능하다.

## 2.3 OWL

OWL은 온톨로지의 계층적 구조나 정보자원의 속성표현 및 속성에 대한 제약조건, 정보자원 사이의 의미적 관계와 추론관계 표현이 가능한 시맨틱 웹 온톨로지 언어이다[7]. OWL은 10년 이상 웹 온톨로지 언어의 개발에 종사한 수많은 그룹의 성과가 결집된 것으로 미국의 DARPA와 유럽위원회의 IST의 계획에 의한 자금지원을 받은 국제 연구 그룹에 의하여 개발된 DAML+OIL (DARPA Agent Markup Language + Ontology Interface Language)을 기초로 하고 있다[2]. 한편 W3C은 2004년 시맨틱 웹의 중요한 요소기술인 RDF와 웹 온톨로지 언어 OWL을 권고안으로 인정한다고 발표했다[2]. 이로써 온톨로지 개발분야에서 기존에 주로 사용하던 DAML+OIL과 완전히 호환되면서 의미적 표현이 강화된 OWL 위주의 온톨로지 개발연구가 활성화되었다.

OWL은 서브언어로서 OWL Lite, OWL DL (Description Logic), OWL Full로 구성되어 있다. OWL Lite는 기본적인 OWL 기능을 포함하는 어휘 집합으로 구성된 언어로서 온톨로지 관련 도구 개발자를 위해 제정된 언어이다. OWL DL은 OWL Lite에 기술논리와 호환되는 어휘를 추가한 언어이며, OWL Full은 OWL DL과 동일한 어휘집합으로 구성되어 있으나 좀 더 보강된 어휘의 의미를 갖는 언어이다[8].

## 2.4 검색엔진

검색엔진은 검색방식에 따라 주제별 검색엔진, 키워드 검색엔진, 메타 검색엔진으로 분류할 수 있다[9].

주제별 검색엔진은 대분류에서 하위분류로 분류항목을 축소하여 순차적으로 검색하는 방법이다. 이를 디렉토리 검색방법이라고 하며 야후와 같은 검색엔진이 대표적이다.

키워드 검색엔진은 찾고자 하는 정보와 관

련된 핵심어 즉 키워드 (Keyword)를 입력하여 정보를 찾는 방법이다. 이 검색엔진은 인터넷에 있는 홈페이지의 내용과 URL (Uniform Resource Locator)을 자체 데이터베이스로 구축해 둔 것이다. 이는 간단한 키워드를 통해 원하는 정보를 찾을 수 있다는 장점을 갖는다. 그러나 일반적인 내용의 키워드를 사용할 경우에는 검색결과가 너무 방대하여 효율이 떨어질 수 있다는 단점을 갖는다. 네이버, 심마니, 라이코스 등의 검색엔진이 대표적이다.

메타 검색엔진은 로봇 에이전트를 이용해서 동시에 여러 검색엔진에서 정보를 찾아 그 결과를 통합하거나 각 검색엔진별로 출력해주는 검색엔진으로 지능형 검색엔진이라고도 불린다. 한번에 다양한 검색결과를 얻을 수 있다는 장점이 있으나 검색속도가 느리고 검색결과가 많아 원하는 정보를 가려내기 어려울 수 있다는 단점을 갖는다.

위에서 살펴본 검색방식들이 현재까지 많이 이용되고 있으나 웹 온톨로지를 이용하여 의미적인 추론을 통한 검색방법이 현재 연구중이며 차세대 검색방법으로 자리잡을 것으로 기대된다.

## 2.5 교육관련 온톨로지 및 메타데이터

온톨로지는 사용 영역과 필요에 따라 구성되는데 교육과 관련된 온톨로지 구축과 메타데이터에 대한 연구로는 다음과 같은 것들이 있었다.

### 1) GEM 통제어휘의 온톨로지 변환 연구 [6]

GEM (Gateway to Educational Materials)은 우리나라의 교육정보 종합시스템 에듀넷과 비슷한 성격의 사이트로 교육정보를 공유하는 역할을 담당하고 있다. 이 연구는 GEM사이트를 온톨로지기반 교수학습자료 정보 게이트웨이로서 구축하기 위한 기초작업으로 GEM에서 사용하던 계층구조를 갖는 어휘들을 온톨로지 구조로 변환하고 있다. 이 연구는 기존에 있던 어휘에 보다 깊은 의미를 부여했다는 것에 의미가 있다. 표 1은 이 연구를 통해 구축

된 온톨로지 구조이다.

Top-Level Class	Slot	Relation
Resource	Audience	resource For-Audience slot slot Audience-Of resource
	Educational Level	resource For-EducationalLevelslot slot EducationalLevel-Ofresource
	Format	resource In-Format slot slot Format-Of resource
	Grade	resource For-Gradeslot slot Grade-Ofresource
	Language	resource In-Language slot slot Language-Of resource
	Pedagogy	resource Used-Forslot slot Use-Of resource
	Relation Type	resource Has-RelationType slot1 [slot2] slot2 Related-To resource by slot1
	Subject	resource Is-About slot slot Subject-Of resource

표 1 GEM ontology

## 2) EDUONTO : Educational Resource Ontology [10]

EDUONTO는 GEM 메타데이터 스키마와 GEM의 controlled vocabulary에 기반한 온톨로지이다. 클래스 타입은 재사용이 가능한 클래스 (Person, Organization, Contact), 자원객체 클래스 (Instructional, Informational, Research) 그리고 어휘클래스 (Subject categories and Terms)를 포함한다.

그림 2는 온톨로지 편집 툴인 Protégé를 통해 본 EDUONTO의 클래스구조이다.

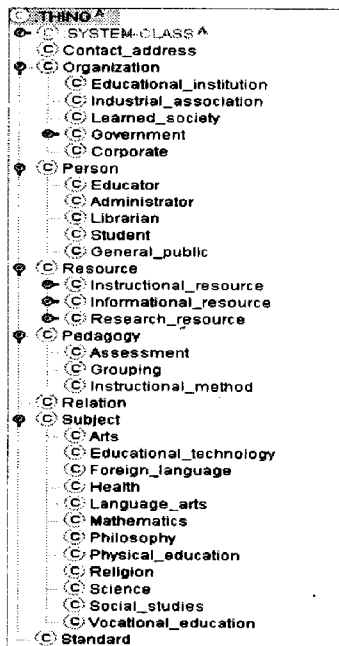


그림 2 EDUONTO 클래스구조

## 3) LOV : Learning Objective Vocabulary [11]

이 연구는 앞서 살펴본 GEM 온톨로지 구축 및 EDUONTO를 기반으로 진행중인 연구이다. 이 연구에서는 학습자료의 성격을 위주로 온톨로지를 구성하고 있으며 표 2는 LOV에서 개발중인 온톨로지 구조이다.

Top class	subclass
Learning objective	<ul style="list-style-type: none"> <li>• Knowledge</li> <li>• Comprehension</li> <li>• Application</li> <li>• Analysis</li> <li>• Synthesis</li> <li>• Evaluation</li> </ul>
Learning object	<ul style="list-style-type: none"> <li>• Curriculum</li> </ul>
Learning content	<ul style="list-style-type: none"> <li>• Fact</li> <li>• Concept</li> <li>• Principle</li> <li>• Procedure</li> <li>• Process</li> </ul>
Learning practice	<ul style="list-style-type: none"> <li>• Areas_of_practice</li> <li>• Assessment_of_practice</li> <li>• Best_practice</li> <li>• Methods_of_practice</li> </ul>
Assessment	<ul style="list-style-type: none"> <li>• Assessment_area</li> <li>• Assessment_method</li> <li>• Assessment_tool</li> </ul>
Learning model	<ul style="list-style-type: none"> <li>• Collaborative_learning</li> <li>• Experience_based_learning</li> <li>• Interactive_learning</li> <li>• Performance_support_and_resources</li> </ul>
Learning context	<ul style="list-style-type: none"> <li>• Formal_education</li> <li>• On-the-job_training</li> <li>• Continuing_education</li> <li>• Leisure_learning</li> </ul>

표 2 LOV 클래스 구조

## 4) KEM(Korea Educational Metadata)2.0 [12]

우리나라 교육분야에 적합한 온톨로지 구축을 위해서 교육분야에서 이루어진 메타데이터에 대한 연구로서 KEM 2.0을 살펴보았다.

KEM은 교육정보의 종류·제작자·용도·자료의 소재지 등을 담은 교육정보 메타데이터이다. 이는 전국 16개 시·도 교육청과 에듀넷·한국교육과정평가원·한국직업능력개발원 등 관련 기관에서 교육정보를 공유하면서 생성된 메타데이터를 KERIS가 주축이 돼 하나의 규격으로 통합한 것이다.

KERIS에서는 ICT활용 교수용 S/W, 멀티미디어 교육자료, ICT활용 교수·학습 과정안, 일반 분야로 나누어 메타데이터 기술항목을 제시하고 있다.

### 3. 학습자료 온톨로지 설계

#### 3.1 시맨틱 웹 기술 지원 툴

본 연구에서는 Protégé 2000을 이용하여 온톨로지를 구축함으로써 향후에 이를 이용한 지식처리가 가능하도록 하였다.

Protégé 2000은 스탠포드대학의 의학정보그룹인 SMI에서 이용자 그룹과 공동으로 개발해 나가는 온톨로지 편집툴로 플랫폼 독립적인 환경이다. 직관적이고 사용하기 쉬운 GUI(Graphic User Interface)를 제공하며 다양한 플러그인을 통해 손쉬운 개발환경을 제공한다.

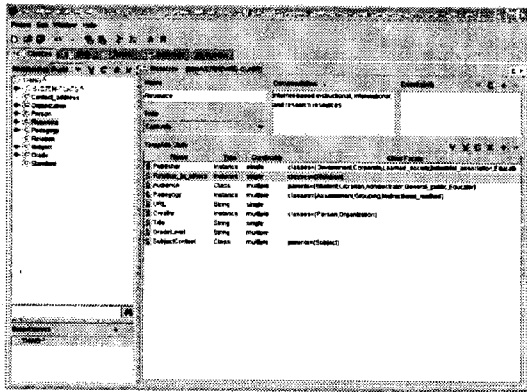


그림 3 Protégé의 편집환경

그림 3과 같은 GUI 편집환경을 갖는 Protégé는 온톨로지 저작툴로서 국내외에서 널리 이용되고 있다[13]. 한편 KERIS (Korea Education & Research Information Service)에서도 OWL 저작툴로 ezOWL을 개발하였는데 Protégé의 플러그인 형태로도 제공하고 있다.

#### 3.2 온톨로지 설계시 고려할 점

Gruber는 온톨로지 설계시 고려할 점으로 명확성, 통일성, 확장성, 인코딩 편중 최소화, 온톨로지 수행 최소화를 들고 있다[14].

- ① 명확성 (Clarity) : 정의는 객관적이고 사회적 전산학적 문맥에서 자유로워야 하며 자연어로 기술되어야 한다.
- ② 통일성 (Coherence) : 정의원칙이 논리

적으로 통일성이 있어야 한다.

- ③ 확장성 (Extendibility) : 온톨로지는 앞으로 공유될 것임을 고려해야 한다.
- ④ 인코딩 편중 최소화 (Minimal encoding bias) : 표기의 편이성을 위해 표현을 선택할 때 인코딩 편중이 발생할 수 있다. 특정 심볼수준의 인코딩에 의존하지 않고 지식레벨에서 개념화가 이루어져야 한다.
- ⑤ 온톨로지 수행 최소화 (Minimal ontological commitment) : 에이전트의 작업이 온톨로지의 정의에 잘 부합될 수 있도록 실재를 모델링해야 한다.

#### 3.3 학습자료 온톨로지

본 연구에서는 앞서 살펴본 EDUONTO를 비롯한 교육관련 온톨로지 연구와 교육용 메타데이터인 KEM 2.0을 참고로 하여 온톨로지를 설계하였다. 초등학교에서 사용되는 학습자료는 크게 학년, 학기, 과목으로 분류될 수 있는데, 고등교육의 기초가 되는 초등학교 교육의 특성상 각 학년 및 과목별로 포함하고 있는 범위는 사회, 경제, 문화 등 실생활의 모든 영역을 다루고 있다. 이러한 우리나라 초등학교 교육과정에 적합한 온톨로지 설계를 위해 먼저 광범위한 초등학교의 학습자료를 위한 온톨로지를 제안하고 이중 시험적으로 초등학교 6학년 1학기의 사회과 학습자료에 초점을 맞추어 세부적인 온톨로지를 설계하였다.

초등학교 학습자료를 위한 온톨로지의 최상위 클래스 구조는 다음과 같이 이루어져 있다. 우선 학습자료를 제작한 사람에 대한 정보로 이름 및 연락처, 기관 등을 하위클래스로 갖는 Person 클래스가 있다. Person과 관련된 클래스로 연락처와 기관클래스가 있다. 다음으로 초등학교에서 다루는 과목에 대한 Subject 클래스로 우리들은 1학년을 비롯한 14개의 교과목과 재량활동, 특별활동의 서브클래스를 갖는다. Grade 클래스는 학년을 나타내는 클래스로 1학년에서부터 6학년과 공통학년의 서브클

래스를 갖는다. Term 클래스는 1학기, 2학기, 공통학기 서브클래스를 갖는다.

표 4는 대분류된 클래스구조를 보여준다.

Top class	subclass
Person	<ul style="list-style-type: none"> <li>Name</li> <li>Organization</li> </ul>
Organization	<ul style="list-style-type: none"> <li>Educational_institute</li> <li>Government</li> <li>School</li> </ul>
Contact_address	-slot
Subject	<ul style="list-style-type: none"> <li>우리들은 1학년</li> <li>바른생활</li> <li>슬기로운생활</li> <li>즐거운생활</li> <li>국어</li> <li>도덕</li> <li>사회</li> <li>수학</li> <li>과학</li> <li>실과</li> <li>체육</li> <li>음악</li> <li>미술</li> <li>영어</li> <li>재량활동</li> <li>특별활동</li> </ul>
Grade	<ul style="list-style-type: none"> <li>1학년</li> <li>2학년</li> <li>3학년</li> <li>4학년</li> <li>5학년</li> <li>6학년</li> <li>공통</li> </ul>
Term	<ul style="list-style-type: none"> <li>1학기</li> <li>2학기</li> <li>공통</li> </ul>

표 4 학습자료 온톨로지 클래스 구조

위의 학습자료 온톨로지 중 사회과목의 6학년 1학기 내용에 중점을 두어 구체화한 사회과 온톨로지를 설계한 내용은 다음과 같다.

초등학교 6학년 1학기 사회과에서 다루고 있는 내용은 선사시대에서부터 현대에 이르기까지의 국사분야로 시대적 순서에 따라 주요 사건이나 인물을 중심으로 전개된다. 사회과 학습자료를 위한 세부적인 온톨로지를 설계하는데 있어서는 이 점을 고려하여 시대 및 국가, 유물, 인물, 사건을 중심으로 하였다. 시대 및 국가는 구석기시대부터 대한민국까지 15개로 구분하고 나라 또는 시대 이름, 해당 연도

정보를 담게 된다. 유물은 이름, 연대, 용도 정보를 포함하며 인물은 이름, 연대, 직업, 한 일 등의 정보를 포함한다. 사건은 이름, 시기, 주요 인물 정보를 포함하게 된다.

온톨로지의 세부적인 의미관계는 그림 4의 슬롯 편집과 그림 5의 인스턴스 편집창 등을 통해 입력하게 된다.

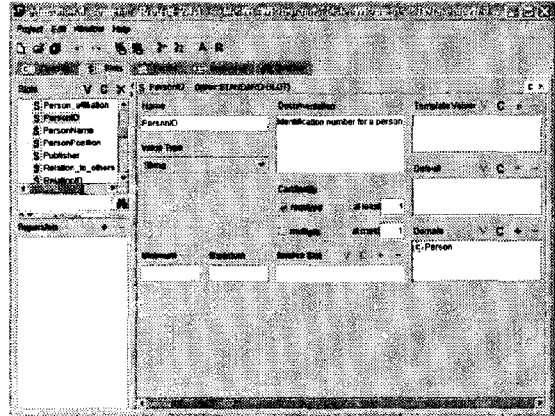


그림 4 슬롯 편집창

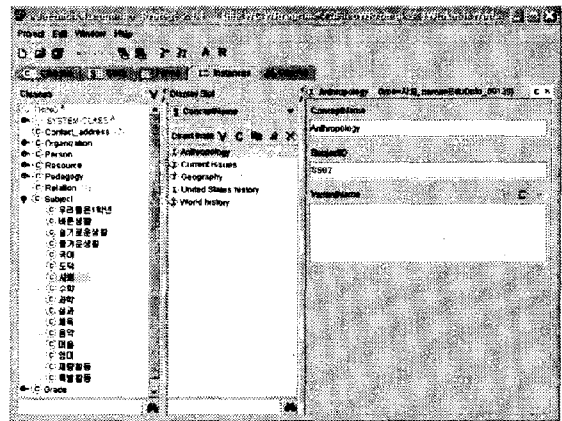


그림 5 인스턴스 편집창

## 4. 학습자료 검색 시스템

### 4.1 시스템 구현 환경

본 시스템을 구현하기 위하여 사용한 환경은 다음과 같다.

- Platform
  - OS : Windows XP
  - Language : Java (J2SDK 1.4.2)

■ Tool

□ RDF Parser : Jena2 (HP)

Jena는 휴렛패커드사의 시맨틱 웹 팀이 개발한 공개 자바 툴킷으로 RDF 워킹그룹의 공동의장이 제작자중 하나로 참여해 현존하는 RDF 및 OWL파서 및 API (Application Program Interface)중 인지도가 높다[15]. Jena API는 SQL과 유사한 문법을 가진 질의언어로 RDQL을 포함하고 있는데, 이를 통해 온톨로지에 대해 질의작업을 수행할 수 있다.

본 초등학교 학습자료 검색시스템은 학습자료에 관련된 정보를 사용자로부터 선택 및 직접입력 방법을 통해 입력받아 RDQL질의로 온톨로지에 질의를 수행하고 이를 바탕으로 검색한 결과를 HTML형태로 사용자에게 제시하는 절차로 이루어진다.

4.2 초기화면 구성

본 시스템의 초기 화면은 사용자의 질의를 입력받기 위한 단계로 그림 6과 같다.

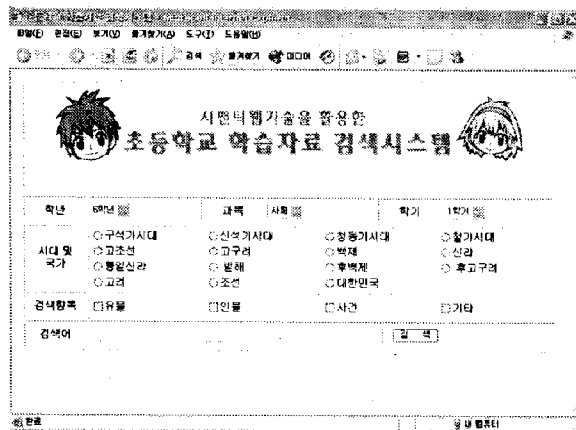


그림 6 학습자료 검색시스템 초기화면

초기화면에서는 대분류로서 학년, 학기, 과목을 차례로 선택할 수 있도록 되어있으며 이 연구에서는 시험적 설계로 6학년 1학기 사회과가 기본적으로 선택되어 있다. 다음으로 6학년 1학기 사회과에서 다루고 있는 국사 내용에 알맞게 시대 및 국가를 선택할 수 있도록 하였고 검색항목으로 유물, 인물, 사건, 기타의 네 가지 유형 중에서 원하는 항목을 복수선택

할 수 있도록 설계하였다. 마지막으로 학습자가 검색하고자 하는 검색어를 직접 입력할 수 있도록 하였다.

4.3 검색결과 화면구성

사용자가 검색정보를 입력하면 온톨로지에 질의를 걸쳐 검색 결과 목록이 테이블 형태로 제시되어 사용자가 목록 중 원하는 자료를 선택하여 상세정보로 이동할 수 있게 한다.

검색결과 화면은 그림 7과 같이 구성된다. 사용자가 입력한 질의정보가 선택사항 표로 표시되고 하단에는 시대 및 국가별, 검색항목별 순으로 정렬되어 검색결과가 제시된다.

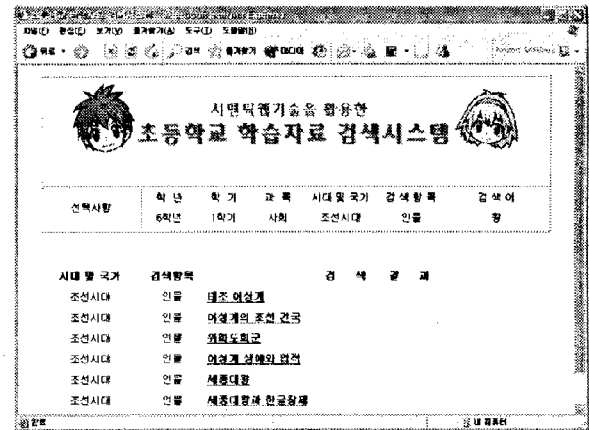


그림 7 학습자료 검색시스템 검색결과화면

검색결과는 사용자 질의어를 바탕으로 의미적 연관관계를 가지고 있는 내용만을 보여주기 때문에 단순히 검색어만 일치하는 무의미한 정보가 필터링되어 검색결과를 사용자가 수동으로 필터링하는 과정이 불필요하게 된다. 따라서 검색에 소요되는 시간낭비도 막을 수 있게 된다.

5. 결론 및 향후 연구과제

본 연구에서는 웹의 미래로 주목받고 있는 시맨틱 웹의 개념과 시맨틱 웹 구현을 위한 기술에 대해 살펴보고, 시맨틱 웹의 활용의 한 측면으로 학습자료 검색시스템을 간략하게 구현하였다.



먼저 시맨틱 웹 구현을 위한 학습자료 온톨로지를 설계하고 그 중 초등학교 6학년 1학기 사회과에 구체화된 세부적인 온톨로지를 시험적으로 구축하였다. 온톨로지 구축에는 OWL 언어를 지원하는 온톨로지 구축 툴인 Protégé 2000을 이용하였다. 구축된 학습자료 온톨로지에 RDQL질의어를 통해 학습자료를 검색하여 결과를 제시하는 검색시스템을 시험적으로 구현하여 의미적으로 보다 정확한 검색결과를 얻을 수 있으며 검색결과를 해석하고 걸러내는 수동적인 작업이 불필요하므로 사용자가 결과를 얻는데 걸리는 시간을 단축시킬 수 있다.

이 검색시스템의 특징은 다음과 같다. 첫째, 학습자료와 연관된 사용자 질의어를 보다 상세하게 입력받는다. 둘째, 사용자 질의어를 바탕으로 온톨로지에 질의하여 검색결과를 얻는다. 셋째, 검색하고자하는 내용의 의미를 분석하여 요구된 의미에 적합한 자료만을 검색결과로 제시한다.

그러나 본 연구에서 구축한 온톨로지는 광범위한 초등학교 학습자료를 모두 포용하기에는 미흡한 점이 있다. 향후 연구과제는 다음과 같다. 첫째, 교과별 학습내용을 분석하여 교과에 적합한 온톨로지를 설계한다. 둘째, 교과별 온톨로지를 통합하는 학습자료 온톨로지를 구축한다.

## 6. 참고문헌

- [1] Berners-Lee, T., Hendler, J. and Lassila, O., "The Semantic Web", Scientific American, 2001
- [2] World Wide Web Consortium Issues RDF and OWL Recommendations <http://www.w3.org/2004/01/sws-pressrelease.html.en>
- [3] 최중민, "시맨틱 웹의 개요와 연구동향", 정보과학회지, 제21권, 제3호, p.4, 2003.
- [4] 조성정, 김진형, "시맨틱 웹의 응용 사례 연구", 정보과학회지 제21권 제3호, p.11, 2003
- [5] Jian Qin, Stephen Paling, "Converting a controlled vocabulary into an ontology : the case of GEM", Information Research, Vol. 6 No.2, January 2001 <http://informationr.net/ir/6-2/paper94.html>
- [6] 이진중, "온톨로지 기반 지능형 검색 시스템", 원광대학교 교육대학원 석사학위논문, p.7, 2003.
- [7] 임인숙, "시맨틱 웹 기술을 활용한 인터넷 쇼핑몰에서의 상품 검색 시스템 구현", 숙명여자대학교 정보통신대학원 석사학위논문, p.17, 2004
- [8] OWL Web Ontology Language Guide, W3C Recommendation 10 Feb 2004. Smith, Welty, McGuinness, eds. <http://www.w3.org/TR/owl-guide/>
- [9] 전성희, 김갑수, "초등학생 언어수준에 적합한 검색엔진", 한국정보교육학회지, 7권, 2호, 2002.
- [10] Jian Qin, "EDUONTO" <http://web.syr.edu/~jqin/eduonto/eduonto.html>
- [11] Jian Qin, "LOV", <http://web.sry.edu/~jqin/LO/LOV2/>
- [12] 한국교육학술정보원, <http://keris.or.kr>
- [13] 양정진, "시맨틱 웹에서의 온톨로지공학", 정보과학회지, 제21권 제3호, p.33, 2003.
- [14] Gruber, T.R., "Toward principles for the design of ontologies used for knowledge sharing", 1993.
- [15] O'Reilly (July 2003) Practical RDF. Sample Chapter 8, Jena: RDF in Java <http://www.oreilly.com/catalog/pracrdf/chapter/index.html>