

# 시맨틱 웹을 활용한 초등학교 학습자료 검색 시스템

이 희 경\*

전 우 천\*\*

## ◆ 목 차 ◆

1. 서 론
2. 이론적 배경
3. 학습자료 온톨로지 설계
4. 학습자료 검색시스템
5. 결 론

## 1. 서 론

웹은 팀 버너스리에 의해 1989년 처음 소개된 이래 짧은 시간 동안 괄목할만한 성장을 거듭해 왔다. 웹 사용자 증가로 웹상의 정보는 정보의 홍수로 불릴 정도로 막대한 양이 되었으며 개인이 필요한 정보를 쉽고 정확하게 찾기란 쉬운 일이 아니다. 현재의 웹에서는 기계가 정보자원을 읽을 수는 있으나 이해할 수는 없기 때문에 최종 정보 취득을 위해 사용자의 추가적인 작업이 필요하기 때문이다. 초등학교 학습에서도 웹 사용이 보편화 되어 있으며 학습자료 검색 서비스를 많이 이용하고 있는데, 이 때의 문제도 너무나 많은 검색결과가 나온다는 점과 원하지 않는 검색결과로 인해 시간을 낭비하게 된다는 점이다. 이러한 문제를 해결하고자 제시되고 있는 차세대 웹이 바로 시맨틱 웹(Semantic Web)이다. 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].

시맨틱 웹 검색이 가능해지면 기존 웹과 같이

단어를 식별해 관련 자료를 찾아주는 것은 물론, 검색어 및 사물간의 관계를 파악하여 추가적인 정보를 제공하고, 독립된 정보들을 통합하여 새로운 결론을 추론하여 제시하는 것도 가능해진다. 이러한 시맨틱 웹 검색을 위한 기초적인 연구로서 본 연구는 초등학교의 학습자료를 검색하고자 할 때 시맨틱 웹의 주요개념중 하나인 온톨로지(Ontology)를 기반으로 학습자료를 검색할 수 있는 시스템을 제안한다. 초등학교의 광범위한 자료를 위한 기본적 온톨로지를 제안하고 실험적으로 초등학교에서 다루는 역사에 대한 세부 온톨로지를 구축하여 이를 바탕으로 자료를 검색함으로써 유의미한 자료들만을 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 연구과제를 제시한다.

\* 서울여의도초등학교 교사

\*\* 서울교육대학교 컴퓨터교육과 부교수

## 2. 이론적 배경

### 2.1 검색엔진

현재 사용되고 있는 검색엔진은 검색방식에 따라 주제별 검색엔진, 키워드 검색엔진, 메타 검색엔진으로 분류할 수 있다[2].

야후와 같은 주제별 검색엔진은 대분류에서 하위분류로 분류항목을 축소하여 순차적으로 검색하는 방법이다. 이 검색엔진은 서비스 제공자가 웹페이지를 일정한 분류기준에 의거하여 항목별로 미리 정리해서 제공하게 된다.

네이버나 심마니 같은 키워드 검색엔진은 찾고자 하는 정보와 관련된 핵심어 즉 키워드를 입력하여 정보를 찾는 방법이다. 이 검색엔진은 인터넷에 있는 홈페이지의 내용과 URL (Uniform Resource Locator)을 자체 데이터베이스로 구축해 둔 것으로 간단한 키워드를 통해 원하는 정보를 찾을 수 있다는 장점을 갖는다.

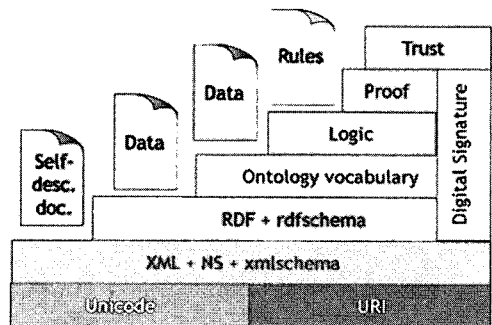
메타서치나 서치닷컴 같은 메타 검색엔진은 로봇 에이전트를 이용해 동시에 여러 검색엔진에서 정보를 찾아 그 결과를 통합하거나 각 검색엔진별로 출력해준다. 한번에 다양한 검색결과를 얻을 수 있다는 장점이 있으나 검색속도가 느리고 검색결과가 많아 원하는 정보를 가려내기 어려울 수 있다는 단점을 갖는다.

위에서 살펴본 검색방식들이 지금까지는 많이 이용되고 있었으나 웹자원의 양이 점차 반대해짐에 따라 검색결과물의 질이 만족스럽지 못해진 것이 사실이다. 이러한 문제점을 해결하고자 제안되는 것이 시맨틱 웹 기술을 활용한 검색방식이다. 온톨로지 기반 검색은 웹에서 자원을 검색할 때 온톨로지를 사용한다. 온톨로지에 정의된 개념과 규칙을 활용하여 검색하며 때로는 검색의 의미를 확장한 추론의 기반으로 온톨로지를 이용하기도 한다. 따라서 웹 온톨로지를 이용하여 의미적인 추론을 통한 검색방법이 현재 연구중이며 차세대 검색방법으로 자리잡을 것으로 기대된다.

### 2.2 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 웹의 창시자인 팀 버너스리에 의해 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 OWL 에 의해 보다 용이하게 그리고 신뢰성이 높은 정보 검색과 재이용이 가능하다고 소개하고 있다[3]. 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. 팀 버너스리가 제시한 시맨틱 웹의 계층구조는 그림 1과 같다[4].



(그림 1) 시맨틱 웹의 계층구조(1)

## 2.3 온톨로지

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

온톨로지에 대한 정의는 여러 가지가 있지만 Gruber는 온톨로지를 “공유된 개념화에 대한 정형화되고 명시적인 명세”[5]라고 정의하였다. 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 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[3]. 또한 온톨로지에서는 개념과 개념 사이의 관계를 정의하는데, 관계술어논리를 이용하며 온톨로지에 직접적으로 표현되지 않는 개념간의 관계를 추론을 통해 제시하거나 이에 대한 질의에 답을 제시할 수 있다[6].

최근 온톨로지 기술은 시맨틱 웹을 구현하기 위한 중요 요소가 되었다. 사용자가 원하지 않는 데이터를 미리 걸러주며, 관련성이 높은 정보의 발견 가능성을 높여주기 위해 사용자가 입력한 검색 단어에 대한 의미와 관계들이 온톨로지로 저장된다. 시맨틱 검색을 위해 모든 분야에 사용되는 개념들을 수집하여 이들간의 관계를 표현하는 것은 현실적으로 불가능하므로 범위를 한정하여 특정 분야나 특정 응용영역을 위한 온톨로지 개발이 추진되고 있다[7].

## 2.4 온톨로지 구축언어

온톨로지를 표현하기 위한 언어도 HTML과 마

찬가지로 SGML에 그 출발점을 두고 있다. HTML이 가진 표현의 제약을 뛰어넘고자 하였던 XML에서 자원, 자원간의 관계, 자원의 위치 정보 등을 표현한 RDF로 발전하였고, 이것은 다시 용어를 정의하는 RDFS로 발전하였다. 여기에 보다 확장된 어휘들을 다룰 수 있는 본격적인 온톨로지 표현 언어들인 개발되었으며 DAML, OIL이 이에 해당한다. 각각의 언어의 장점을 살려 만들어진 DAML+OIL이 온톨로지 구축언어로 널리 사용되다가 이를 더욱 발전시킨 OWL이 최근 온톨로지 언어의 표준으로 자리잡게 되었다.

OWL은 온톨로지의 계층적 구조나 정보자원의 속성표현 및 속성에 대한 제약조건, 정보자원 사이의 의미적관계와 추론관계 표현이 가능한 시맨틱 웹 온톨로지 언어이다[8]. 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[3].

## 2.5 교육관련 온톨로지 및 메타데이터

온톨로지는 사용 영역과 필요에 따라 한 분야에 특화된 온톨로지를 구성하는 경우도 있고 이미 구축되어 있는 온톨로지를 통합하여 온톨로지를 구성하기도 한다. 교육과 관련된 온톨로지 구축과 메타데이터에 대한 연구로는 다음과 같은 것들이 있었다.

### 2.5.1 GEM 통제어휘의 온톨로지 변환 연구(4)

GEM (Gateway to Educational Materials)은 우리나라의 교육정보 종합시스템 에듀넷과 비슷한 성격의 미국의 교육관련 사이트로 교육정보를 공유하는 역할을 담당하고 있다. 이 연구는 GEM사이트를 온톨로지기반 교수학습자료 정보 게이트웨이로서 구축하기 위한 기초작업으로 GEM에서 사용하던 계층구조를 갖는 어휘들을 온톨로지 구조로 변

환하고 있다. 이 연구는 기존에 있던 어휘에 보다 깊은 의미를 부여했다는 것에 의미가 있다.

### 2.5.2 EDUONTO : Educational Resource Ontology(9)

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.

### 2.5.3 LOV : Learning Objective Vocabulary(10)

이 연구는 앞서 살펴본 GEM 온톨로지 구축 및 EDUONTO를 기반으로 진행중인 연구이다. 이 연구에서는 학습자료의 성격을 위주로 온톨로지를 구성하고 있으며 Learning Objective, Learning Object, Learning Content, Learning Practice, Assessment, Learning Model, Learning Context로 상위 클래스가 구성된다.

### 2.5.4 KEM(Korea Educational Metadata) 2.0(11)

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 KERIS (Korea Education & Research Information Service)가 주축이 돼 하나의 규격으로 통합한 것이다.

KEM 메타데이터 정보모델은 일반사항, 생명주기, 메타메타데이터, 기술, 교육, 저작권, 관계, 주석, 분류의 9가지 범주로 분류된다.

## 2.6 관련연구

시맨틱 웹이 소개된 이후 아직까지는 실제적으로 서비스가 이루어지는 곳은 없으나 10여 년 간 세계적으로 다양한 연구가 이루어지고 있으며 국내에서도 시맨틱 웹과 시맨틱 웹에 필요한 온톨로지에 대한 연구가 점차 확산되고 있다.

[6]의 연구에서 원광대학교 박물관이 소장중인 역사유물에 관한 온톨로지를 DAML 언어로 작성하고 이를 기반으로 한 지능형 검색 시스템을 설계하였다. 유물 검색에 사용될 질문의 유형을 분석하여 온톨로지의 클래스와 관계 및 속성들을 정의하고 해당 유물의 전체 정보를 검색 결과로 제시한다. 온톨로지 작성을 위한 절차를 상세히 소개하고 있으나 사용자가 검색에 직접적으로 질의어를 입력하지 않고 온톨로지서 읽어들이는 정보에서 선택하도록 하고 있어 사용자의 선택에 한계점이 있다.

[8]의 연구에서 OWL을 이용해 디지털 카메라에 대한 온톨로지를 작성하고 RDQL로 질의하여 고객들의 상품 결정을 돕는 시스템을 구현하였다. 각 디지털카메라 제조사마다 다르게 사용되는 용어들의 동의어 및 디지털카메라의 용도와 가격대별 분류 등을 온톨로지로 정리하여 고객에게 제품정보를 비교할 수 있는 검색결과를 제공한다. 검색은 주로 용도별 구분에 의해 이루어진다고 볼 수 있으나 용도별 구分的 구체적 기준은 사용자에게 제시되지 않고 용도별 구분 외에 사용자가 카메라를 분류할 수 있는 방법이 없다는 단점이 있다.

[12]의 연구에서 제주도청에서 제공하는 여행정보에 대한 온톨로지를 DAML+OIL을 이용해 작성하고 이를 직접적으로 검색하는 시스템을 구축하였다. 이 시스템은 하나의 용어에 대해 한국어, 한자, 영어로 동의어 처리를 할 수 있도록 하였으나 의미적 연관관계를 갖는 대상을 결과로 보이는 것이 아니라 키워드 매칭 방식의 검색시스템과 다른 결과는 제시하고 있다.

최근 웹 온톨로지 표준언어로 자리잡은 OWL을 이용하여 온톨로지를 구축한 사례는 매우 적고 교육분야에서 시맨틱 웹에 대한 연구는 아직 활발하

지 않은 상태로 학습자료 검색을 위한 온톨로지나 시맨틱 검색 연구는 전무한 실정이다. 따라서 본 연구에서는 의미적 상관관계를 갖는 결과를 제시함으로써 학습효과를 높일 수 있는 학습자료 온톨로지를 OWL로 작성하고 이를 검색할 수 있는 검색 시스템을 구축한다.

### 3. 학습자료 온톨로지 설계

#### 3.1 시맨틱 웹 기술 지원 툴

시맨틱 웹을 실현시키기 위한 주요 요소인 온톨로지 구축작업을 위한 툴로 *OntoEdit*, *OilEd*, *Protégé 2000* 등의 제품들이 개발되어 사용되고 있다.

*OntoEdit*은 *Ontoprise*사에서 개발한 툴로 개념과 인스턴스를 연결하여 그래프로 만들어주는 방식이나 인터페이스가 사용하기 쉽지 않은 단점이 있다. *OilEd*는 영국의 *Manchester* 대학에서 개발된 *Oil* 기반의 단순한 온톨로지 편집기이나 *OWL* 온톨로지 편집에는 적당하지 않다. *Protégé 2000*은 스탠포드대학의 의학정보그룹인 *SMI*에서 이용자 그룹과 공동으로 개발해 나가는 온톨로지 편집 툴로 플랫폼 독립적인 환경이다. 직관적이고 사용하기 쉬운 *GUI* (*Graphic User Interface*)를 제공하며 다양한 플러그-인을 통해 손쉬운 개발환경을 제공한다.

*GUI* 편집환경을 갖는 *Protégé*는 온톨로지 저작 툴로서 국내외에서 널리 이용되고 있다[13]. 따라서 본 연구에서는 *Protégé 2000*을 이용하여 온톨로지를 구축함으로써 향후에 이를 이용한 지식처리가 가능하도록 하였다.

한편 *KERIS*에서도 *OWL* 저작툴로 *ezOWL*을 2004년 개발하여 발표하였다. *ezOWL*은 온톨로지 저작툴로 널리 이용되는 *Protégé*의 플러그인 형태로 제공되며 온톨로지 개념 및 속성간의 관계를 그래프로 표현하여 주어 *Protégé*만으로는 확인하기 쉽지 않던 온톨로지의 전체적인 구조를 입체적으로 확인할 수 있도록 도와주고 있다.

#### 3.2 온톨로지 설계

##### 3.2.1 온톨로지 설계의 기초

*Gruber*는 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 논리적으로 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 심볼 수준의 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.

또한 온톨로지를 설계할 때에는 일반적으로 다음과 같은 순서를 거친다[15].

- ① 온톨로지의 도메인과 범위 결정  
온톨로지를 설계함에 있어 가장 기본적으로 결정할 내용이 '무엇'에 대한 온톨로지를 만들 것인가 하는 것이다. 온톨로지란 무엇을 할 것인가, 어떤 질문에 대한 답을 제공할 수 있어야 하는가 등에 대한 답을 구하여 개발할 온톨로지의 도메인과 범위를 결정한다.
- ② 미리 개발된 온톨로지의 재사용 고려  
웹 상에는 다양한 온톨로지 라이브러리가 존재한다. 개발하고자 하는 온톨로지가 이미 개발되어 있는지, 혹은 비슷한 분야의 온톨로지가 개발되어 있다면 이를 재사용할 수 있는지의 여부도 확인하여 작업의 효율을 높인다.

③ 온톨로지의 주요 용어 나열

사용자가 사용할 용어들의 리스트를 작성하여 봄으로써 온톨로지 작성의 방향을 잡을 수 있으며 어떠한 관계가 필요한지 확인할 수 있다. 주요 개념, 개념간의 관계, 개념이 가질 속성 등을 나열한다.

④ 클래스와 클래스 계층 정의

클래스 계층을 개발하는 데에는 몇 가지 접근 방식이 있을 수 있다. 가장 일반적인 개념 정의에서 출발하여 점차 세부적인 개념을 정의하는 탑-다운(Top-down) 방식과 특정 영역의 개념 정의에서 출발하여 보다 일반적인 개념으로 정의해나가는 바텀-업(Bottom-up) 방식, 두 방식을 혼합한 컴비네이션(Combination) 방식으로 나눌 수 있다. 세 방식은 개발자의 편의에 따라 선택될 수 있다.

⑤ 클래스의 속성 정의 - 슬롯

클래스 자체만으로는 충분한 정보를 표현할 수 없다. 개념의 내부적 구조를 기술하기 위해 속성의 정의가 필요하다. 해당 속성이 어느 클래스에 해당되는지도 고려하여 정의한다. 클래스에 정의된 속성은 서브클래스에도 상속된다.

⑥ 슬롯의 패킷 정의

슬롯은 값의 타입, 허용된 값, 값의 수, 기타 설정 등에 따라 서로 다른 패킷을 갖게 된다.

⑦ 인스턴스 생성

가장 마지막 단계로 클래스의 개별적인 인스턴스를 생성하게 된다. 인스턴스 생성은 클래스 선택, 해당 클래스의 개별적인 인스턴스 생성, 생성된 인스턴스의 슬롯 값 작성의 순으로 이루어지게 된다.

이상의 순서에 따라 온톨로지를 개발함으로써 온톨로지 개발의 효율을 높이고 온톨로지 구조의 체계를 공고히 할 수 있다. 본 연구에서도 위에서 제시한 온톨로지 개발 절차에 의하여 온톨로지를 설계하였다.

3.2.2 학습자료 온톨로지 설계

본 연구에서는 앞서 2장에서 살펴본 EDUONTO를 비롯한 교육관련 온톨로지 연구와 교육용 메타데이터인 KEM 2.0을 참고로 하여 온톨로지를 설계하였다. 온톨로지는 광범위한 영역을 개발하기 보다는 전문적인 도메인별로 개발하여 이를 통합하는 방향으로 개발되는 것이 일반적이며, 또한 온톨로지의 재사용성을 높일 수 있다. 따라서 본 연구에서도 전체적인 학습자료 검색의 기초가 될 온톨로지를 제안하고 세부적인 학습분야별로 영역에 알맞게 개념을 설정하여 온톨로지를 설계할 것을 제안한다.

기존 검색시스템을 이용하여 학습자료를 검색할 경우에는 입력한 검색어와 일치하는 내용을 담은 페이지 뿐만 아니라 검색어를 포함하기는 하나 분야가 전혀 다른 내용까지 검색되기도 한다. 즉, 검색어로 입력되지는 않았으나 검색어와 연관된 분야의 내용은 전혀 검색할 수 없다는 단점이 있었다. 이러한 문제를 해결하기 위한 연구의 일환으로 본 연구에서는 학습자료간의 의미적 연관관계가 정의된 온톨로지를 설계하고 이를 바탕으로 검색을 하고자 하였다.

학습내용간 의미적 연관관계를 설정하여 연관된 정보를 함께 검색결과로 얻음으로써 학습효과를 높일 수 있는 분야로는 인물, 사건, 유물 및 유적 등이 서로 유기적으로 연관된 역사분야가 있다. 초등학교에서 다루는 역사학습 내용은 시대적 흐름에 따라 인물, 사건 위주로 전개가 되며 연관된 유물 및 유적이 소개되고 있다. 따라서 학습자료 온톨로지 설계를 통해 검색시스템을 구성할 경우 학습효과를 높일 수 있을 것으로 기대되는 분야이기 때문에 본 연구에서는 초등학교 학습내용 수준에 맞춘 역사학습자료 온톨로지를 설계하고 이를 바탕으로 검색을 하고자 하였다.

초등학교에서 사용되는 학습자료는 크게 학년, 학기, 과목으로 분류될 수 있는데, 중·고등 교육의 기초가 되는 초등학교 교육의 특성상 각 학년 및 과목별로 포함하고 있는 범위는 사회, 경제, 문화 등 생활의 거의 모든 영역을 다루고 있다. 이러

한 우리나라 초등학교 교육과정에 적합한 온톨로지 설계를 위해서는 각 교과에서 다루고 있는 내용에 대한 분석이 이루어져야 하며, 교과 내용에 바탕을 둔 온톨로지 설계가 이루어져야 한다.

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 기관 클래스가 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 서브클래스를 갖는다. Term 클래스는 1학기, 2학기, 공통학기 서브클래스를 갖는다.

표 1은 초등학교 학습자료의 대분류 클래스구조를 보여준다.

표 2는 초등학교 학습자료 위계에 따라 구축한 학습자료 온톨로지 일부이다.

온톨로지의 세부적인 의미관계는 슬롯 편집창과 인스턴스 편집창 등을 통해 입력하게 된다.

위의 학습자료 온톨로지 중 초등학교에서 다루고

(표 1) 초등학교 학습자료 온톨로지 클래스 구조

Top class	subclass
Person	<ul style="list-style-type: none"> <li>• Name</li> <li>• Organization</li> </ul>
Organizations	<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>• 공통학년</li> </ul>
Term	<ul style="list-style-type: none"> <li>• 1학기</li> <li>• 2학기</li> <li>• 공통학기</li> </ul>

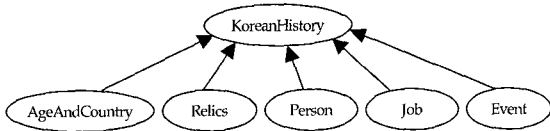
(표 2) 초등학교 학습자료 온톨로지(일부)

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-
  rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000
  /01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002
  /07/owl#"
  xmlns="http://www.owl-ontologies.
  com/unnamed.owl#"
  xml:base="http://www.owl-ontologies.
  com/unnamed.owl">
  <owl:Ontology rdf:about=""/>
  <owl:Class rdf:ID="바른생활">
    <owl:disjointWith>
      <owl:Class rdf:ID="체육"/>
    </owl:disjointWith>
  </rdf:subClassOf>
    <owl:Class rdf:ID="Subject"/>
  </rdf:subClassOf>
    :
  <owl:Class rdf:ID="Grade">
    <owl:disjointWith>
      <owl:Class rdf:ID="Pedagogy"/>
    </owl:disjointWith>
    :
  </owl:Class>
  <owl:Class rdf:ID="School">
```

있는 역사 내용에 중점을 두어 구체화한 온톨로지를 설계한 내용은 다음과 같다.

초등학교에서 다루는 역사 내용은 선사시대에서부터 현대에 이르기까지의 국사분야로 시대적 순서에 따라 주요 사건이나 인물을 중심으로 전개된다. 역사분야 학습자료를 위한 세부적인 온톨로지를 설계하는데 있어서는 이 점을 고려하여 시대 및 국가, 유물 및 유적, 인물, 사건을 중심으로 하였다. 또한 인물에 대한 자세한 정보를 추가하기 위한 직업 클래스가 설정되었다. 또한 온톨로지에 포함되는 정보들은 초등학교에서 다루고 있는 역사내용을 기준으로 설정한다. 그림 2는 역사 온톨로지의 최상위 클래스구조를 도식화 한 것이다.

시대 및 국가는 구석기시대부터 대한민국까지 18개로 구분하고 시대 및 국가 이름, 해당 연도, 관련 유물 및 유적, 관련 인물, 관련사건 정보를



(그림 2) 역사 온톨로지 최상위 클래스구조

답게 된다. 유물 및 유적은 이름, 유물 및 유적의 해당 시대 및 국가, 관련된 인물 정보를 포함한다.

인물은 이름, 해당 시대 및 국가, 인물의 직업, 관련사건 등의 정보를 포함한다. 사건은 이름, 해당 시대 및 국가, 관련 인물 정보를 포함하게 된다. 표 3은 클래스별 속성 정의의 내용을 요약한 것이다.

표 4의 시대 및 국가의 연대 정보는 초등학교 교과서 및 국사편찬위원회의 역사자료를 참고하여 작성하였다.

유물 및 유적 정보 역시 초등학교 교과서를 참고하여 작성하였으며 표 5는 유물 및 유적 정보 중 일부인 조선시대 유물 정보이다.

(표 3) 클래스별 속성 정의

클래스	속 성
AgeAndCountry	hasRelics, hasPerson, hasEvent
Relics	isRelicsOf, isMadeBy, isMadeWhen
Person	isPersonOf, hasJob, hasMade, hasRelationWith
Job	isJobOf
Event	isEventOf, isHappenedBy, hasRelationWith

(표 4) 시대 및 국가 연대정보

시대 및 국가	연대	시대 및 국가	연대
구석기시대	B.C. 70000년 경부터	통일신라	676-935
신석기시대	B.C. 8000년 경부터	후백제	892-935
청동기시대	B.C. 1000년 경부터	후고구려	901-918
철기시대	B.C. 300년 경부터	발해	698-926
고조선	B.C. 2333-B.C. 238	고려	918-1392
고구려	B.C. 37-668	조선	1392-1897
백제	B.C. 18-663	대한제국	1897-1910
신라	B.C. 57-676	일제시대	1910-1945
가야	42-532	대한민국	1945-현재

인물 정보는 사회 교과서 및 사회과 탐구 자료를 기초로 작성하였으며 인물의 특성에 관계된 직업정보도 함께 가질 수 있도록 하였다. 인물은 크게 정치인, 군인, 학자로 구분되며 정치인은 다시 왕, 왕비, 신하로, 군인은 장군, 승병장, 의병장, 민병으로 구분하였다. 표 6은 조선시대 인물정보를 정리한 자료 중 일부이다.

역사적 사건들 역시 교과서의 자료에 기초하여 작성하였으며 표 7은 사건 정보 중 일부인 조선시대의 사건과 그에 관련된 정보를 나타낸 것이다.

본 연구에서는 위와 같은 형식으로 구석기시대부터 현재의 대한민국에 이르기까지의 역사관련 정

(표 5) 유물 및 유적 정보(조선시대)

이름	관련인물	시대 및 국가
당백전	홍선대원군	조선(1868)
광혜원		조선(1885)
거북선	이순신, 선조	조선(1592)
발해고	유득공	조선(1784)
동학	최계우, 최시형	조선(1860)
대동여지도	김정호	조선(1861)
독립신문	서재필	조선(1896)
행주산성	권율	조선

(표 6) 인물 정보(조선시대)

이름	연대	직업	관련 유물 및 유적, 사건
장영실	1423	학자	
김정호	1861	학자	
이성계	1335-1408	왕	위화도회군, 척화비
태종	1367-1422	왕	호패법
세종대왕	1397-1450	왕	집현전, 훈민정음, 영도확장
서산대사	1520-1604	의병장	임진왜란
고경명	1533-1592	의병장	임진왜란
권율	1537-1599	군인	
곽재우, 홍의장군	1552-1617	군인	임진왜란
김시민	1554-1592	군인	임진왜란
인조	1595-1649	왕	병자호란
유형원	1622-1673	학자	반계수록



(표 7) 사건 정보(조선시대)

이 름	시대 및 국가	관련인물
위화도회군	조선(1388)	이성계
조선건국	조선(1392)	이성계
임진왜란	조선(1592-1598)	이순신
병자호란	조선(1636-1637)	
한산도대첩	조선(1592)	이순신
행주대첩	조선(1593)	권율
강화도조약	조선(1876)	
단발령	조선(1895)	홍선대원군
윤요호사건	조선(1875)	

(표 8) 역사자료 온톨로지(일부)

```

<owl:disjointWith rdf:resource="#백제"/>
<owl:disjointWith rdf:resource="#고려"/>
</owl:Class>
<owl:Class rdf:about="#양부일구">
<owl:disjointWith rdf:resource="#담헌서"/>
<owl:disjointWith>
<owl:Class rdf:about="#발해고"/>
</owl:disjointWith>
<owl:disjointWith rdf:resource="#반계수록"/>
<owl:disjointWith rdf:resource="#집현전"/>
<owl:disjointWith>
<owl:Class rdf:about="#집강소"/>
</owl:disjointWith>
<owl:disjointWith rdf:resource="#열하일기"/>
<owl:disjointWith rdf:resource="#경세유표"/>
<rdfs:subClassOf rdf:resource="#Relics"/>
    
```

보들을 조직화하여 초등학교 역사자료 온톨로지를 구축하였다. 표 8은 구축된 온톨로지의 일부분을 나타낸 것이다.

## 4. 학습자료 검색 시스템

### 4.1 시스템 구현 환경

본 초등학교를 위한 학습자료 검색시스템을 구현하기 위하여 사용한 환경은 다음과 같다.

- Platform

- OS : Windows XP
- Language : Java (J2SDK 1.4.2)
- WebServer : Tomcat 4.0.3
- Tool
  - Ontology Builder : Protégé 2000
  - OWL Parser : Jena2 (HP)
  - Client Application : MS Explorer 6

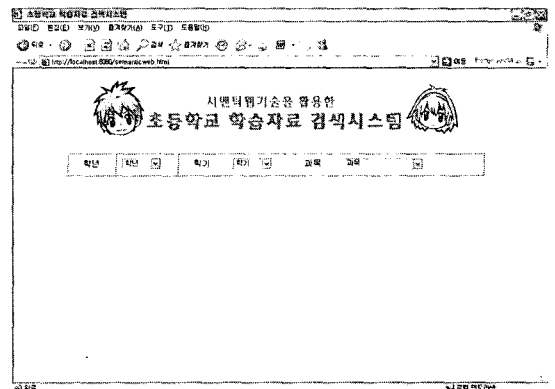
Jena는 휴렛패커드사의 시맨틱 웹 팀이 개발한 공개 자바 툴킷으로 RDF 워킹그룹의 공동의장이 제작자중 하나로 참여해 현존하는 RDF 및 OWL 파서 및 API (Application Program Interface)중 인지도가 높으며[16] Jena API는 질의언어를 통해 온톨로지에 대해 질의작업을 수행할 수 있다.

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

#### 4.1.1 초기화면 구성

본 시스템의 초기 화면은 학습자료 검색을 위해 사용자의 질의를 입력받기 위한 초기단계로 그림 3과 같다.

초기화면에서는 대분류로서 학년, 학기, 과목을 차례로 선택할 수 있도록 되어있으며 보다 정확한 검색을 위해 해당 분야의 온톨로지 선택이 이루어



(그림 3) 초등학교 학습자료 검색시스템 초기화면

질 수 있도록 직접 질의어를 입력하는 폼은 초기화면에 설정되어 있지 않다. 학년을 선택하면 각 학년에 맞게 과목이 재설정되며 학기 선택 후 과목을 선택하면 선택내용을 종합적으로 검토하여 해당 온톨로지로 연결한다. 해당 온톨로지가 없는 경우 온톨로지 설정이 필요하다는 메시지가 나타나며, 온톨로지가 설정된 경우에는 해당 온톨로지에 적합한 질의화면이 나타난다.

이 연구에서는 실험적 설계로 역사부분의 온톨로지를 설계하였으므로 해당 내용을 다루는 6학년 1학기 사회과를 선택하면 다음의 그림 4와 같은 역사관련 학습내용 질의 입력 화면이 나타난다.

질의 입력 화면에서는 초등학교에서 다루고 있는 역사 내용에 알맞게 시대 및 국가를 선택할 수 있도록 하였고 검색항목으로 유물, 인물, 사건, 기타의 네 가지 유형 중에서 원하는 항목을 복수 선택할 수 있도록 설계하였다. 마지막으로 학습자가 검색하고자 하는 검색어를 직접 입력할 수 있도록 하였다.

질의 입력 화면에서 입력된 질의내용은 서버릿으로 넘겨져 Jena API를 통해 온톨로지에 질의하게 된다. 선택 내용에 따른 질의 결과화면이 사용자에게 주어진다.

#### 4.1.2 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 5와 같이 구성된다.

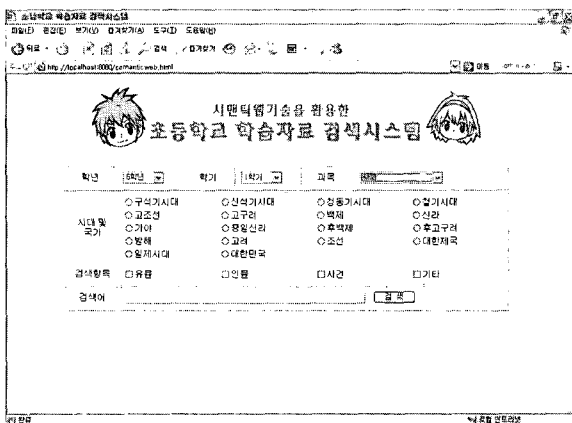
사용자가 입력한 질의정보가 선택사항 표로 표시되고 하단에는 사용자 질의어와 관련된 항목별로 테이블 형태로 검색결과가 제시된다. 시대 및 국가별로 유물, 인물, 사건, 기타 항목 순으로 제시되며 연관된 정보들이 검색결과로 나타난다.

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.

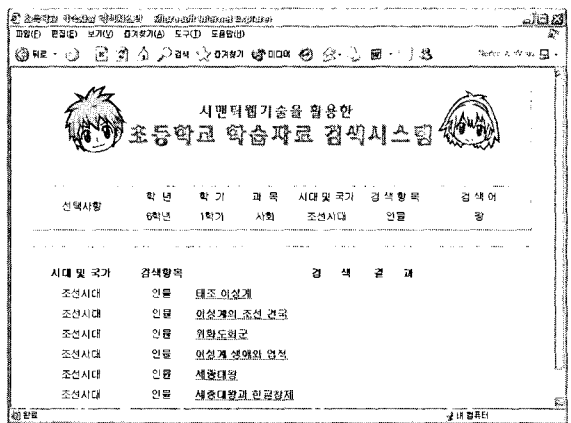
또한 역사자료들의 연관관계에 의해 질의어와 관련된 자료들을 함께 검색결과로 제시하기 때문에 학습자가 미처 파악하지 못했던 연관관계들을 학습하기에도 도움이 될 것으로 기대된다.

### 5. 결론 및 향후 연구과제

본 연구에서는 웹의 미래로 주목받고 있는 시맨틱 웹의 개념과 시맨틱 웹 구현을 위한 기술에 대해 살펴보고 시맨틱 웹의 주요 기술인 온톨로지와 온톨로지 언어에 대해 살펴보았다. 기존 검색엔진의 한계를 뛰어넘을 수 있는 대안으로서 시맨틱 웹



(그림 4) 역사관련 학습내용 질의 입력 화면



(그림 5) 학습자료 검색시스템 검색결과화면

의 활용의 한 측면으로 학습자료 온톨로지를 구축하여 온톨로지에 질의하여 결과를 얻는 학습자료 검색시스템을 구현하였다.

먼저 시맨틱 웹 구현을 위한 학습자료 온톨로지를 설계하였으며 그 중 역사학습자료에 중점을 두어 세부적인 학습자료 온톨로지를 구축하였다. 초등학교에서 다루는 역사분야의 학습자료 특성에 맞게 시대 및 국가, 유물 및 유적, 인물, 역사적 사건을 중심으로 온톨로지 구조를 설계하였으며 각각의 특성에 알맞은 속성을 정의하고 값을 정의하였다.

이 검색시스템은 의미적으로 보다 정확한 검색 결과를 얻을 수 있으며 검색결과를 해석하고 걸러내는 수동적인 작업이 불필요하므로 사용자가 결과를 얻는데 걸리는 시간을 단축시킬 수 있다는 장점을 갖는다. 또한 키워드 매칭 방식과 달리 검색어와 의미적으로 연관된 정보를 함께 제시하기 때문에 초등학교 역사학습을 하는 데 있어서도 도움이 될 수 있다.

이 검색시스템의 특징을 정리하면 다음과 같다. 첫째, 학습자료와 연관된 사용자 질의어를 보다 상세하게 입력받는다. 둘째, 사용자 질의어를 바탕으로 학습자료 온톨로지에 질의하여 검색어가 포함된 결과 및 검색어와 연관된 정보를 검색결과로 얻는다. 셋째, 검색하고자하는 내용의 의미를 분석하여 요구된 의미에 적합한 자료만을 검색결과로 제시한다. 넷째, 질의어와 연관된 내용을 함께 제시함으로써 학습의 효율을 높일 수 있다.

향후에는 초등학교에서 다루는 각 분야의 학습자료를 온톨로지로 구성하기 위해 학습자료의 특성에 알맞은 의미적 연관관계를 분석하는 연구가 이루어져야 할 것이다.

## 참 고 문 헌

- [1] Berners-Lee, T., Hendler, J. and Lassila, O. (2001), "The Semantic Web", Scientific American.
- [2] 정성희, 김갑수(2002), "초등학생 언어수준에 적합한 검색엔진", 한국정보교육학회지, 7(2).
- [3] World Wide Web Consortium Issues RDF and OWL Recommendations, <http://www.w3.org/2004/01/sws-pressrelease.html>.
- [4] 최중민(2003), "시맨틱 웹의 개요와 연구동향", 정보과학회지, 21(3), 4.
- [5] Jian Qin, Stephen Paling(2001), "Converting a controlled vocabulary into an ontology : the case of GEM", Information Research, Vol. 6 No.2.
- [6] 이진중(2003), "온톨로지 기반 지능형 검색 시스템", 석사학위논문, 원광대학교
- [7] 이은주(2003), "시맨틱 웹을 위한 문서검색시스템", 석사학위논문, 계명대학교.
- [8] 임인숙(2004), "시맨틱 웹 기술을 활용한 인터넷 쇼핑몰에서의 상품 검색 시스템 구현", 석사학위논문, 숙명여자대학교.
- [9] Jian Qin, "EDUONTO"  
<http://web.syr.edu/~jqin/eduonto/eduonto.html>
- [10] Jian Qin, "LOV",  
<http://web.sry.edu/~jqin/LO/LOV2/>
- [11] 한국교육학술정보원, <http://keris.or.kr>
- [12] 정은경(2004), "시맨틱 웹 환경에서의 온톨로지 기반 정보검색 시스템", 석사학위논문, 제주대학교
- [13] 양정진(2003), "시맨틱 웹에서의 온톨로지공학", 정보과학회지, 21(3).
- [14] Gruber, T.R.(1993), "Toward principles for the design of ontologies used for knowledge sharing".
- [15] Natalya F. Noy, Deborah L. McGuinness, "A Guide to Creating Your First Ontology".

[1] Berners-Lee, T., Hendler, J. and Lassila,

## ◎ 저자 소개 ◎



### 이 희 경

1999년 이화여자대학교 초등교육과 졸업(학사)  
1999년~2003년 서울상도초등학교 교사  
2003년~현재 서울교육대학교 대학원 컴퓨터교육과 석사과정  
2004년~현재 서울여의도초등학교 교사  
관심분야 : 데이터 마이닝, 시맨틱 웹



### 전 우 천

1985년 서강대학교 전자계산학과 졸업(공학사)  
1987년 서강대학교 대학원 전자계산학과 졸업(공학석사)  
1997년 University of Oklahoma, Dept. of Computer Science(공학박사)  
1998년~현재 서울교육대학교 컴퓨터교육과 부교수  
관심분야 : 초등컴퓨터교육, 데이터마이닝, 정보영재