

시맨틱 웹 기반 박물관 유물 검색을 위한 온톨로지 설계 및 구현

Design & Implementation of an Ontology for Relic Search in Museums

박진석, 양기철, 오정진*
목포대학교, 전남과학대학*

Park Jin-Seuk, Yang Gi-Chul, Oh Jeong-Jin*
Mokpo Univ., Mokpo Univ., Chunnam
Techno Collage*

요약

인터넷의 급속한 발달로 인하여 웹은 다양하고 방대한 데이터를 가지게 되었다. 이러한 방대한 정보를 가진 웹에서는 사용자가 정보의 검색, 구성 및 통합하는 작업이 쉽지 않다. 이러한 문제점을 극복하기 위해서 W3C(World Wide Web Consortium)에서는 차세대 웹으로 시맨틱 웹을 제시하였다.

시맨틱 웹은 웹 데이터의 의미(meaning)를 기계가 이해할 수 있도록 만들어 정보의 재사용성을 향상시키고, 이기종간의 상호 운용성을 보장한다. 본 논문에서는 박물관용 시맨틱 웹 구현을 위한 온톨로지를 설계하고 구현하였다. 온톨로지는 박물관 유물 검색용 표준 온톨로지를 활용될 수 있을 것이다.

Abstract

Because of the rapid development of internet and Various Types. huge amount and Various types of data are available on the web there days.

The tasks such as information retrieval, composition and integration are not easy on the web which contains such huge amount of data. In order to solve these problems W3C introduced Semantic Web as a next generation Web.

Semantic Web guarantee the cooperation between heterogeneous Systems and improve reusability of information by providing the machine understandable Web data. In this paper. We design develop an ontology for the museum. The implemented ontology can be used as a standard ontology for the relic search in a museum.

I. 서론

웹은 다양한 정보원으로써 무한한 저장 능력을 갖춘 정보의 바다로 부각되어진지 오래다. 웹을 통하여 사회의 지식이 공유됨으로써 사용자들이 쉽게 접근할 수 있고 물리적 시간과 공간을 초월한 범세계적인 단일 정보 네트워크를 구성할 수 있게 되었다. 그러나

나 방대한 데이터, 이기종간의 정보 시스템과의 통합, 정보의 의미 전달 등이 취약하다. 이러한 문제점들을 극복하기 위하여 시맨틱 웹이 등장하였다.

시맨틱 웹은 자동화된 웹 서비스의 제공을 가능하게 하며 컴퓨터의 지능적인 정보 처리가 가능하도록 온톨로지를 이용한다. 이를 통하여 사용자가 원하는

웹 서비스의 발견 및 이기종간의 상호 운용성을 보장하고 원하는 정보를 의미적으로 찾고 추론하는 기능이 가능토록 한다.

온톨로지는 개념과 개념간의 관계를 이용하여 개념을 정의한다[5][7].

II. 시맨틱 웹과 온톨로지

기존의 웹에서는 원하는 정보를 찾기 위해 키워드를 기반으로 하는 디렉토리 검색 또는 HTML 페이지의 구문에 대한 matchmaking을 통한 색인 검색을 수행하게 되는데 이러한 경우 사용자가 원하는 정확한 의미를 갖는 정보만을 찾아내기가 어려우며 동일한 의미를 갖는 다른 키워드를 포함한 정보를 찾아낼 수 없는 등의 비효율성이 존재한다.

시맨틱 웹은 웹 온톨로지를 기반으로 웹상에 존재하는 자원들에 대한 의미 마킹을 하고 검색 에이전트를 이용하여 의미 명세에 대한 추론을 통해 필요로 하는 정보를 보다 정확하게 찾아내는 것을 가능하게 해준다[5][10].



▶▶ 그림 1. 팀 버너스-리가 제안한 시맨틱 웹의 계층 구조

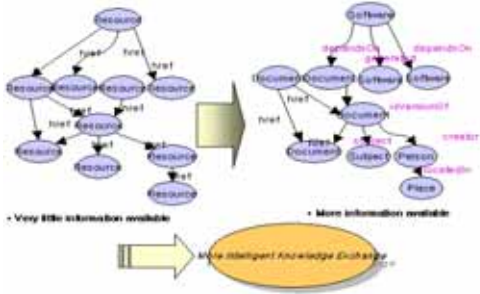
그림 1은 시맨틱 웹의 계층 구조를 보여준다. 그림에서 보는 바와 같이 시맨틱 웹은 기존의 웹 기술 및 XML 기술을 기반으로 하여 웹 자원들간의 관계를 표현하기 위한 RDF(Resource Description Framework)/RDFS(RDF Schema) 표준 위에 웹 자원을 개념으로 표현하고 각각의 개념과 개념간의 논리적인 관계를 표현하기 위한 웹 온톨로지 계층이

존재함을 알 수 있다[1][2][3].

온톨로지는 특정 개념에 대하여 의미적으로 표현하기 위해 개념과 개념들 간의 관계를 이용한 지식 표현 방법으로서 최근 차세대 웹 기술로서 시맨틱 웹이 등장하면서 그 핵심 기술로 관심이 높아지고 있다. 웹 온톨로지란 웹상에 존재하는 자원들에 대한 온톨로지로서 W3C 표준인 RDF(Resource Description Framework)를 기반으로 정의된 온톨로지 언어를 통해 기술된다. 잘 알려진 Web 온톨로지 언어로는 DAML(DARPA Agent Markup Language), OIL(Ontology Interface Layer), SHOE(Simple HTML Ontology Extension) 등이 있으며 DAML과 OIL의 기능을 합쳐서 만들어진 DAML+OIL이 있다. 현재는 W3C에 의해 DAML+OIL을 개정중인 OWL(Web Ontology Language)에 대한 표준화가 진행되고 있다[6][7].

시맨틱 웹을 실현하기 위한 단계중 하나는 바로 온톨로지다. 온톨로지란 용어 사이의 관계를 정의하고 있는 일종의 사전과 같은 것이라 할 수 있다. 온톨로지는 어휘의 정의를 다른 어휘와의 논리적 관계뿐만 아니라 가장 기본적(primitive) 어휘부터 파악해나가는(bottom-out) 구조를 통해 나타낸다.

지식 표현이 해당 분야의 전문가들이 동의하거나 공통적으로 사용하는 가장 기본적인 어휘에서 출발한다는 점에서 해당 분야에 특정적(domain-specific)이라 할 수 있다. 이는 개념규정을 말하며 용어와 용어들 간의 관계를 표현하는 컴퓨터 판독이 가능한 형식적 규정을 의미한다. 온톨로지는 해당 분야의 전문가들의 관심, 목적, 관점을 반영한다. 따라서 온톨로지의 역할은 해당 분야에서 사용되는 어휘의 해석과 의미의 한계를 정하는데 있다. 이러한 의미에서 온톨로지는 사용되는 어휘 사이의 논리적 형식보다는 어휘의 내용을 중심으로 창조된다고 할 수 있다.



▶▶ 그림 2. 시맨틱 웹의 진화단계

그림 2는 현재의 웹에서의 정보의 처리와 시맨틱 웹에서의 정보 처리에 차이를 보이는 그림이다. 현재의 웹에서 단순히 하이퍼 링크로만 정보를 기술해 보여지지만 시맨틱 웹에서의 정보는 각각의 정보 페이지들간의 개념적, 의미적으로 연결되어 있는 것을 보여준다.

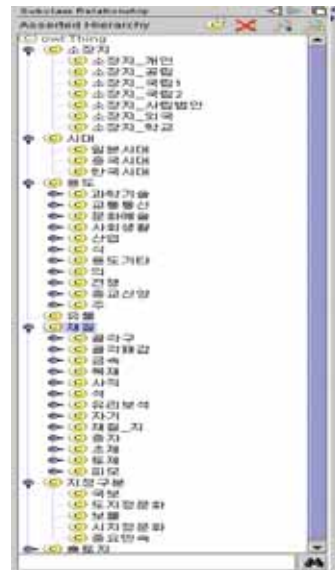
시맨틱 웹 서비스는 기존의 WSDL(Web Service Description Language)기반의 웹 서비스를 온톨로지 웹 서비스를 이용하여 보다 지능화되게 확장한다. 온톨로지 웹 서비스는 웹상에 존재하는 서비스를 좀 더 풍부한 표현력으로 기술하기 위한 semantic markup을 정의하는 온톨로지로서 서비스 도메인에 독립적인 부분(upper ontology)과 특정 서비스 도메인에 종속적인 부분으로 구성될 수 있다. 도메인에 독립적인 온톨로지는 DAML-S(DAML Service)와 최근 발표된 OWL-S(OWL Service) 등의 표준 온톨로지 웹 서비스를 통해서 작성될 수 있으며 이것을 기반으로 도메인에 종속적인 부분을 작성하게 된다 [12][13][14]. 기존의 웹 서비스는 서비스에 대한 검색과 실행 능력을 향상시키는 것에 초점을 맞추고 있으나 시맨틱 웹 서비스는 보다 자동화되고 정교한 탐색, 실행, 합성, 상호적 서비스 까지도 가능하게 한다.

III. 유물 검색을 위한 온톨로지

본 논문에서는 Protege2000을 사용하여 유물 검색을 위한 온톨로지를 설계하였다.

유물의 계층적 구조는 박물관 전산화를 위한 유물 분류 표준화방안에 따라 OWL 설계하여 작성하였다 [8][11]. 유물 온톨로지의 도메인과 범위 결정을 할 때 고려한 사항은 다음과 같이 생각하여 결정하게 되었다.

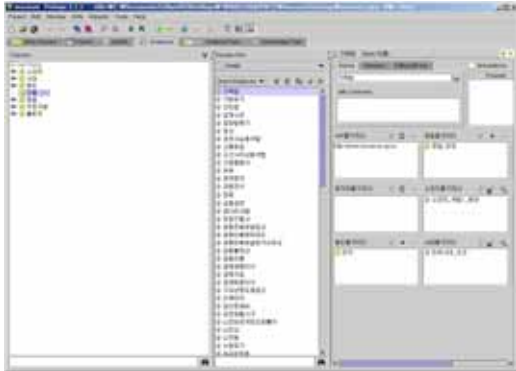
- 온톨로지가 포함하게 될 도메인은 무엇인가?
- 어디에 사용할 것인가?
- 어떠한 질문들이 있을 것이며, 그러한 질문들을 어떻게 질의 할 것인가?
- 질의에 대한 응답이 가능한가?
- 사용 대상이 누구인가(사용 대상에 따라 유물 온톨로지에 대한 적용 가능한 질의를 작성)?



▶▶ 그림 3. 박물관 유물 구조

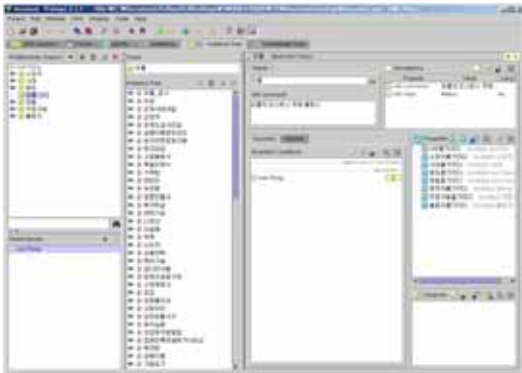
그림 3은 박물관 유물 클래스를 표현한 것이다. 최상위 계층에 있는 Thing 이하의 클래스들은 소장지, 시대, 용도, 유물, 재질, 지정구분, 출토지로 구성되어

있다. 의미적 표현을 하기 위해서 각각의 클래스들은 관계적으로 표현되어있다. 아울러 각각의 클래스들에 들어 있는 인스턴스는 또 다른 유물 클래스 안에 인스턴스로 표현되어지기도 한다.



▶▶ 그림 4. 박물관 유물 인스턴스

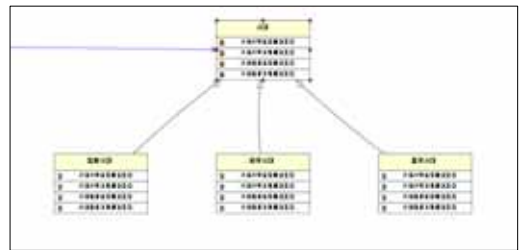
그림 4는 박물관 유물 인스턴스들을 보인 예이다. 각각의 인스턴스들은 값들을 가진다. 아울러 인스턴스는 소장지, 시대, 용도, 재질, 지정구분, 출토지 클래스에 속해있기 때문에 관계도 표현된다.



▶▶ 그림 5. 박물관 유물 인스턴스 프로퍼티

그림 5는 박물관 유물 인스턴스의 프로퍼티를 예로 보인 것이며 구현된 온톨로지에서 각 인스턴스들은 아래와 같은 프로퍼티를 가진다.

- URL
- 소장지
- 시대
- 용도
- 재질
- 제작자
- 지정구분
- 출토지



▶▶ 그림 6. 시대에 대한 구조

그림 6은 시대에 대한 클래스 부분이며 이는 ETRI에서 개발한 ezOWL을 이용하여 보인 예이다[15]. 그림 6에서 시대 클래스는 일본시대, 한국시대, 중국시대 서브클래스를 가지고 있다. 따라서 하나의 한국 유물을 검색할 때 그 유물이 가지는 시대와 다른 시대와 비교하여 검색할 수도 있다.

개발된 온톨로지는 OWL을 이용하여 작성되었다. OWL (Web Ontology Language)구문구조는 키워드 검색의 범위를 넘어서는 기능으로 지원은 웹 자원의 의미 기술을 전제 조건으로 하는 것이다. OWL은 웹 온톨로지와 연관된 지식창고(knowledge bases), 즉 추론시스템에 축적된 논리적 명제(assertions)들을 정의하는 언어로서 이러한 명제는 하나의 온톨로지에 근거하거나 또는 OWL이 명시하는 방식에 따라 수집한 다수의 분산형 온톨로지에 근거할 수 있다. OWL 명제는 클래스의 구성원 (members)들에 관한 사실과 구성원들 간의 관계를 기술하며, 이러한 명제의 집합으로 이루어진 온톨로지는 구문적으로 정의되지 않은 사실의 논리적 유추를 가능하게 한다. 클

래스 (classes)와 속성 (prosperities) 및 이에 적용할 수 있는 제약사항 (constraints)들의 집합인 OWL 온톨로지는 다음과 같은 요소들을 포함하고 있다.

- 데이터의 속성, 즉 클래스의 요소인 속성값에 관한 기술
- 객체의 속성, 즉 클래스의 요소간의 관계 기술
- 클래스들의 인스턴스
- 속성들의 인스턴스

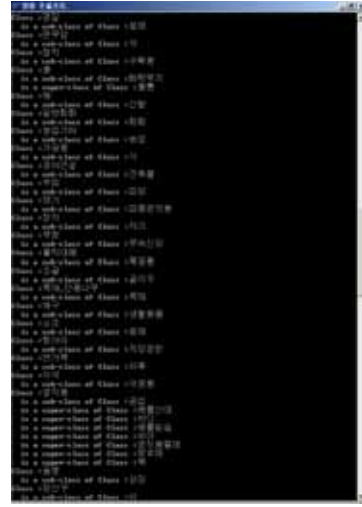
```

<owl:Class rdf:ID="산업">
<rdfs:subClassOf rdf:resource="#용도"/>
</owl:Class>
<owl:Class rdf:ID="종다래끼">
<rdfs:subClassOf>
<owl:Class rdf:about="#과종경작용"/>
</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="재질사적_마">
<rdfs:subClassOf rdf:resource="#재질_사적"/>
</owl:Class>
<owl:Class rdf:ID="안정">
<rdfs:subClassOf>
<owl:Class rdf:about="#마구"/>
</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="종이공용">
<rdfs:subClassOf rdf:resource="#공업"/>
</owl:Class>
<owl:Class rdf:ID="재질금속_은제">
<rdfs:subClassOf rdf:resource="#재질_금속"/>
</owl:Class>
<owl:Class rdf:ID="재질모재_참죽나무">
<rdfs:subClassOf rdf:resource="#재질_목재"/>
</owl:Class>
<owl:Class rdf:ID="금속공용모부">
<rdfs:subClassOf rdf:resource="#금속공용"/>
</owl:Class>

```

▶▶ 그림 7. 용도 OWL 부분

그림 7은 유물의 용도 분류 OWL 부분이다. 산업 분야에 쓰이는 유물을 사용하는 분야에 따라 분류한 것이다. 그림 8은 사용자가 찾고자 하는 유물을 검색할 때 다음과 같은 결과를 얻게 된다[9].



▶▶ 그림 8. Jena을 이용한 유물 관계 표현

IV. 결론 및 기대효과

본 논문에서는 일반적인 박물관 유물 검색에서와는 달리 단순한 유물 검색만 되는 것이 아니라 검색되어진 유물과 의미적으로 연결되어지는 유물 관계 또한 검색되어 질 수 있는 시스템 개발시 필요한 온톨로지를 설계하고 구현하였다.

이는 정형화된 의미 정보를 축적하여 확장된 검색 환경을 제공하여 웹상에서 정보의 유통을 확대하고 정보의 활용성 및 재사용성을 높일 수 있는 기반이 될 것으로 사료된다. 향후 개발되는 온톨로지는 온톨로지 언어가 가지는 개방성과 접근성을 활용하여 기존의 온톨로지와 효율적으로 융합될 수 있도록 의미적인 개념망을 형성하는 기반의 구축이 필요하고, 완결성, 정확성, 효율성을 극대화시키기 위해 추론 기능이 가능하도록 설계된 형식적 의미를 포함하도록 발전되어야 하며, 아울러 데이터들의 표준화도 해결하여야 하는 시급한 과제이다.

■ 참고문헌 ■

- [1] XML : <http://www.w3.org/XML/>
- [2] RDF : <http://www.w3.org/TR/REC-rdf-syntax/>
- [3] RDF Schema :
<http://www.w3.org/TR/2003/WD-rdf-schema-20030123/>
- [4] 온톨로지 :
<http://www.w3.org/TR/2003/WD-owl-guide-20030331/>
- [5] Semantic Web : <http://www.w3.org/2001/sw/>
- [6] <http://www.w3.org/2002/03owlt/editors-draft/draft/>
- [7] Ontology Modeler :
<http://lists.w3.org/Archives/Public/www-webont-wg/2002Aug/att-0056/01-fois-intro.pdf>
- [8] 국립중앙박물관 박물관 전산화를 위한 유물분류 표준화방안
- [9] <http://jena.sourceforge.net/ontology/>
- [10] The Semantic Web : A Guide to the Future of XML, Web Services, and Knowledge Management
Michael C.Daconta, Leo J.Obrst, Kevin T.Smith
- [11] Proceedings of Museums and the Web 2004 (MW2004), Arlington, Virginia, USA, March 31-April 3, 2004(Forthcoming), MuseumFinland-Finnish MUseums on the Semantic Web User's Perspective
- [12] <http://www.w3c.org/TR/daml+oil-reference>
- [13] <http://www.w3c.org/TR/2001/Note-daml+oil-walkthru-20011218/>
- [14] <http://www.w3c.org/TR/owl-ref>
- [15] <http://iweb.etri.re.kr/ezowl/tutorial.html>