

## 시맨틱 웹의 온톨로지 언어<sup>†</sup>

서울시립대학교 이재호\*

### 1. 개요

시맨틱 웹(semantic web)은 웹(web) 상의 정보에 잘 정의된 의미(semantic)를 부여함으로써 사람뿐만 아니라 컴퓨터도 쉽게 문서의 의미를 해석할 수 있도록 하여 컴퓨터를 이용한 정보의 검색 및 해석, 통합 등의 업무를 자동화하기 위한 목적으로 제안되었다. 이러한 “잘 정의된 의미”를 다루고자 하는 것이 바로 시맨틱 웹 온톨로지 언어의 역할이다[4]. 시맨틱 웹은 이러한 지식의 정의와 관련된 온톨로지 연구와, 웹 자원을 서술하기 위한 RDF 및 RDFS와 같은 연구[15], 그리고 이를 활용하기 위한 자동화된 자율적 프로그램인 에이전트[16][29]에 관한 연구 등을 포괄한다.

시맨틱 웹의 문서는 자연어 위주의 기존 웹 문서와 달리 컴퓨터가 해석하기 쉽도록 부여한 의미를 가지고 있기 때문에 자동화된 에이전트나 정교한 검색 엔진들이 부여된 의미를 이용하여 고 수준의 자동화와 지능화를 이룰 수 있게 된다. 웹은 기본적으로 웹 자원(resource)의 위치를 URI(Universal Resource Identifier)를 써서 정적으로 지정하는 공유된 공간이다. 웹에 의미(semantic)를 부여한다는 것은 사용자 인터페이스를 위한 자연어 처리 기능을 부가하고자 하는 것이 아니라 컴퓨터가 처리하기 용이하게 하고자 선언적인 추가 정보를 부여하는 것을 뜻한다.

아직까지 웹 상의 대부분의 정보는 컴퓨터가 아닌 사람이 읽고 해석하기 편리하도록 구성되어 있기 때문에 이를 컴퓨터가 지능적으로 해석하여 처리하기에는 어려움이 많은 것이 현실이다. 일상의 컴퓨터

사용을 보더라도 사용자가 검색 엔진을 이용하여 원하는 정보들을 검색하고 이를 사용자의 눈으로 일일이 확인하여 필요한 정보를 선택한 후 이를 다시 마우스나 키보드 조작을 거쳐 선택된 내용을 해석하여 확인하는 과정을 반복적으로 수행하게 된다. 원하는 정보를 검색하고 해석하여 핵심 정보만을 추출한 후 이를 가공하고 종합하여 사용자에게 제공하고 사용자의 의도에 맞게 예약이나 구매와 같은 일을 행할 수 있는 컴퓨터 프로그램이 사용자의 신뢰 속에 일상적으로 사용될 수 있다면 사용자는 반복적이고 소모적인 작업에서 벗어날 수 있을 것이다.

특히 웹의 사용자와 정보량이 증가함에 따라 사용자가 필요한 정보를 얻는 것이 오히려 힘들어지는 현상이 발생한다. 이러한 현상은 웹 상에 존재하는 문서들이 문서의 “내용(content)”이 아니라 글자체, 글자크기, 문단의 모양과 같은 문서의 “외관(format)”을 표현하기 위해서 HTML이나 SGML과 같은 마크업 언어를 사용하기 때문이다. 문서의 의미를 담고 있는 문서의 내용은 실상 자연어로 나열되어 있어 자동화된 컴퓨터 프로그램, 즉 에이전트가 의미에 적합한 처리를 하기 위해서는 많은 노력을 들여 의미 해석을 해야 한다. 더욱이 이러한 과정은 노력에 비하여 부정확한 해석을 종종 낳는 것이 현실이다. 웹 상의 문서가 증가함에 따라, 대부분의 웹 검색에 사용자는 검색 엔진이 제시한 수많은 검색 결과에서 자신의 원하는 정보를 찾기 위해 시간을 허비하게 되는 것이다.

온톨로지를 사용하면 어떠한 효용이 있는지를 예를 들어 Google[39]과 Yahoo[38] 검색 엔진을 비교해 볼 수 있다. Google은 기본적으로 사용자가 제시한 주제어를 가지고 이 주제를 포함한 웹 문서를 기계적으로 검색하여 제시한다. 주제와의 관련 정도를 수치로 제시하기도 하지만 이러한 수치 역시 사용

<sup>†</sup> 본 연구는 첨단정보기술 연구센터를 통하여 과학재단의 지원을 받았음.

\* 종신회원

자의 “의도”를 반영하기 보다 문서의 외형적 특징에 의존한 기계적 계산 결과인 경우가 많다. 반면 Yahoo는 사람이 미리 정의한 주제들과 이들을 상속 관계나 부분-전체 관계를 써서 계층적으로 분류한 주제 계통을 가지고 웹 문서들을 이 주제에 맞게 분류하고 이들 간의 연관 관계를 미리 정의하여 놓았다. 여행을 준비하기 위해 여행 정보를 찾고자 하면, Yahoo는 “1. 여가생활과 스포츠 > 여행, 관광 2. 비즈니스와 경제 > 기업간거래(B2B) > 여행, 관광 3. 엔터테인먼트 > 음악 > 음악감상실 > 테마별 감상 > 여행 4. 비즈니스와 경제 > 취업, 채용 > 회사별 > 여행, 교통”과 같은 4개의 카테고리를 제시하고 각 카테고리에 하위 카테고리를 계층적으로 두어 해당 정보를 찾아 갈 수 있게 한다. 반면에 같은 “여행”이라는 검색어를 가지고 Google에서 찾게 되면 2,130,000개의 관련 문서를 제시한다. 검색 목적에 따라 Google과 같은 검색이 유용한 경우도 많지만 특정 “의도”를 가지고 이 의도에 적합한 문서를 찾고자 할 때는 Yahoo와 같은 주제 분류가 유용할 것이다. 바로 Yahoo의 계층적 카테고리가 온톨로지의 한 예라 할 수 있다. 주제들과 주제들 간의 관계를 계층적으로 정의한 온톨로지를 기반으로 문서를 분류하고 검색하는 것이다. 앞의 예는 단편적인 온톨로지의 예이며 시맨틱 웹에서 제공하는 온톨로지는 다음과 같은 특징을 갖는다.

- 온톨로지의 작성 주체로 국가나 대규모 단체와 같은 집단 뿐만 아니라 개인이나 그룹 같은 소규모의 다수 집단을 지향한다. 현재의 웹도 대규모 단체 뿐만 아니라 개인을 포함한 다수의 소규모 집단에 의해서 자발적으로 수많은 문서가 작성되고 있다. 시맨틱 웹의 온톨로지도 이와 같이 다수의 소규모 집단에 의한 자발적 온톨로지 작성을 전망하고 있다[3].
- 온톨로지의 확장과 수정을 전제로 작성된다. Yahoo의 카테고리는 이미 정해진 카테고리를 따라야 하며 사용자가 카테고리를 확장하거나 수정하는 일을 전제로 만들어진 것이 아니다. 시맨틱 웹에서는 이미 정의된 다른 온톨로지를 검색하여 이를 수정하고 확장하여 개별적 목적에 맞게 진화시키는 기능을 추구한다. 또한 하나 이상의 온톨로지를 결합하여 새로운 온톨로지를 만들어 낼 수도 있다. 이렇게 수정되거나 만들어진 온톨로지는 다시 웹을 통하여 다른 사용자가 검색하고 내려 받아

사용할 수 있게 된다[16].

- 사용자에게 맞게 특화되거나 전문 분야에 적합하게 세분화되고 경우에 따라서는 상충되는 온톨로지의 모임을 허용한다. 일상의 가장 보편적인 주제와 분류를 위해 만들어진 Yahoo와 같은 독점적이고 일관성 있는 카테고리와는 달리 전문 분야 또는 극히 개인적인 시각을 반영하는 다수의, 경우에 따라서는 서로 상충되는 온톨로지들이 자연스럽게 공존하며 산재하게 될 것이다. 시맨틱 웹 이전의 CYC[30][31]나 WordNet[32] 등이 모순이 없는 일체의 온톨로지를 추구하는 것과는 대조적이다.
- 시맨틱 웹의 목적과 부합되도록 기계에 의한 처리를 전제로 만들어진다. 다른 온톨로지도 물론 컴퓨터에 의한 처리를 전제로 하겠지만 시맨틱 웹의 온톨로지는 지능화된 에이전트 간의 주요한 의사소통 수단으로 이용된다. 이를 위해 웹의 표준 언어인 XML[33]을 근간으로 개발이 이루어지고 있다. XML은 기본적으로 계층구조를 순차적으로 표현하는 문법을 제공한다. 이를 바탕으로 에이전트에 의한 의미 처리가 수월하도록 데이터 모델과 문법을 추가한 RDF[15][34]가 시맨틱 웹 온톨로지의 바탕이 된다.
- 온톨로지의 작성이 자발적 무의식적으로 이루어진다. 현재의 웹 문서도 초기에는 HTML 태그를 직접 이용하는 전문가에 의해 제한적으로 만들어졌으나 점차 이러한 작업을 지원하는 소프트웨어의 개발로 인하여 단순한 문서 작업만으로도 손쉽게 작성할 수 있게 되었다. 사용자는 HTML, 구문에 대해 전혀 이해할 필요가 없고 또한 이미 만들어진 웹 문서를 수정하여 새로운 문서를 작성하게 되었다. 시맨틱 웹의 온톨로지도 같은 과정을 거칠 것으로 전망된다. 사용자는 온톨로지에 대한 이해 없이 일반 문서 작업과 유사한 작업을 하면서 자연스럽게 이미 만들어진 온톨로지를 이용하고 수정하고 확장하는 작업을 하게 될 것이다. 사용자는 문서 작업 소프트웨어의 일부로 편집된 기능을 써서 온톨로지를 사용하게 될 것이다.
- 온톨로지의 추론 기능이 제공된다. 온톨로지의 중요한 역할의 하나는 개념과 개념 간의 관계를 정의하는 것이다. 대표적으로는 상속과 부분-전체 관계를 들 수 있다. 이러한 관계와 Description Logic[28]을 이용하면 온톨로지에 직접적으로 표현되지 않은 개념 간의 관계를 추론을 통하여 제시

하거나 이에 대한 질의에 답할 수 있게 된다.

시맨틱 웹에서는 자연어 위주의 기존 웹 문서와 달리 컴퓨터가 해석하기 쉽도록 의미를 부여한 계층을 가지고 있기 때문에 자동화된 에이전트나 정교한 검색 엔진들이 부여된 의미를 이용하여 고 수준의 자동화와 지능화를 이룰 수 있게 된다. 온톨로지를 핵심으로 "의미"를 부여하고자 하는 시맨틱 웹은 현재의 "링크의 그물"인 웹을 "의미의 그물"로 변환시킬 것이다[1].

## 2. 온톨로지의 개요

웹의 발전과 더불어 온톨로지라는 용어를 자주 하게 된다. 검색 엔진을 써서 온톨로지라는 단어를 키워드로 입력하면 454,000건의 문서가 검색되는 것만 보아도 사용 빈도를 짐작할 수 있다. 시맨틱 웹에서의 온톨로지는 "개념화의 규정"(a specification of a conceptualization)을 말한다. 즉 해당 영역의 개념들과 이들 개념 간의 상호 관계를 정의하는 것을 온톨로지라고 한다. 개념화란 특정 목적을 위해서 표현하고자 하는 대상 세계의 단순화, 추상화 된 관점을 말한다. 지식을 다루는 지식기반 시스템이나 에이전트 시스템은 명시적 또는 암묵적으로 개념화를 하게 마련이다. 개념화는 관심 대상 내에 개체와 개념, 그리고 이들 간의 관계가 존재한다는 가정 하에 이루어진다.

본래 철학의 주된 연구 대상인 만큼 철학만큼 오랜 역사를 가지고 있는 온톨로지는 존재의 본질을 연구하는 추상적 철학 분야를 일컫는다고 하나 시맨틱 웹에서의 정의를 좀더 부연하면 "공유된 개념화의 형식적 명시적 서술 (a formal explicit specification of a shared conceptualization)"이라고 할 수 있다[19]. 개념화란 앞서 언급한 대로 대상으로 삼고 있는 세계에서 일어나는 현상에 연관된 개념들을 파악하기 위한 추상적 모델을 말한다. 명시적이란 것은 개념의 사용 유형과 사용된 유형의 제약조건이 명시적이란 것을 의미한다. 형식적이란 것은 기계가 읽을 수 있어야 한다는 것을 말한다. 공유는 온톨로지가 표현하는 개념이 개별적이 아닌 해당 그룹 구성원 간에 합의된 지식에 바탕을 두고 있다는 것을 의미한다.

이러한 맥락에서 어휘사전(thesaurus)이나 데이터베이스 스키마, 또는 Yahoo의 주제 분류와 같은 계층 분류 카테고리(taxonomy)도 온톨로지의 일종

으로 볼 수 있다. 온톨로지는 이러한 단순한 개념의 분류와 단순한 클래스-서브클래스 관계 외에도 개념과 개념간의 다양한 관계와 개념의 속성과 제약조건, 그리고 추론규칙을 가지게 된다.

시맨틱 웹의 입장에서 앞의 정의를 다시 풀어 쓰면 광범위한 구성원 간에 합의되어 애매모호하지 않게 통용되며 컴퓨터 프로그램에 의해 처리하기 용이한 지식 체계를 온톨로지라고 말할 수 있을 것이다. 웹 프로그램이나 데이터베이스에 익숙한 사용자를 위해 비유적으로 정의하자면, "표준화된 지식체계의 데이터베이스"라고도 할 수 있다. 표준화는 공유를 전제로 합의된 표현 및 사용 규칙이 정의되어 있다는 것을 말하며 이를 데이터베이스라고 하는 것은 관련 지식과 관련된 개념들이 명시적 스키마에 따라 저장되어 특정 개념을 기존의 개념과 연관 지어 추가, 삭제, 갱신할 수 있고 질의 조건과 관계를 지정하여 질의, 검색, 추출, 더 나아가서는 새로운 개념과 관계의 추론[24]까지 할 수 있음을 나타낸다. 추후에 4장에서 논의할 온톨로지 관련 도구들[14]은 바로 이러한 작업들을 효과적으로 수행하는 기능을 가진다.

온톨로지는 응용프로그램의 개발에 매우 유용하게 이용할 수 있으며 이러한 추세는 웹을 기반으로 한 자동화된 시스템의 수요 증가와 더불어 다수의 프로그램 간의 의사소통이 필요해짐에 따라 더욱 증가할 것으로 예상된다. 에이전트 간의 의사소통을 하기 위한 가장 기본적인 전제가 언어와 같은 의사소통 수단이며 이러한 언어의 의미를 정의하는 것이 바로 온톨로지의 역할이기 때문이다. 에이전트 시스템이나 지식관리 시스템, 전자상거래등에는 이미 일정한 형태의 온톨로지를 기반으로 플랫폼이 개발되어 이용되고 있다. 예를 들어 전자상거래 플랫폼의 경우 주문, 구매, 입찰등에 필요한 제품 정보와 규격, 제한 조건 등을 온톨로지로 구축하여 거래 상대들과 공유된 표준화된 방법으로 거래가 이루어질 수 있도록 플랫폼이 구현되고 있다. ISO/IEC[40]를 중심으로 개발 중인 XML Topic Maps (XTM)[13]도 이러한 온톨로지의 필요성을 반영하고 있다.

온톨로지와 관련하여 온톨로지 공학(ontological engineering)은 온톨로지의 설계, 평가, 검사, 수정, 유지보수, 통합, 적용 등의 온톨로지 생명주기 전 과정에서 온톨로지를 효율적으로 개발하고 사용하는 방안을 제시한다. 이에 대해서는 4장에서 간단히 논하기로 한다.

### 3. 시맨틱 웹의 온톨로지

이장에서는 시맨틱 웹의 온톨로지를 실현하기 위한 다양한 방안 중에서 DAML+OIL[37]과 이어 개발된 OWL[22][23]을 중심으로 현황을 소개한다. 이를 위해서 DAML+OIL이 바탕으로 하고 있는 XML[33]과 RDF[34]를 먼저 소개한다.

#### 3.1 XML

의미 부여를 위해서는 데이터 뿐만 아니라 데이터에 대한 추론을 규정하는 규칙(rule)을 표현할 수 있는 언어가 필요하다. 시맨틱 웹을 위한 의미 부여 위한 언어로서 이미 사용중인 XML(eXtensible Markup Language)과 RDF(Resource Description Framework)를 이용한다.

XML은 컴퓨터 프로그램에 의한 관독이 용이한 문서 작성 규정이다. 마크업(markup)이란 문서 내용의 역할을 지정하는 추가 정보를 지정한다는 것을 뜻하며 문서에 논리적 구조를 부여하고 정보 자체를 기술하는 역할을 한다. 마크업은 HTML에서 사용된 것과 유사한 태그(tag)를 써서 이루어진다. 반면에 HTML과는 달리 XML에서는 사용자가 필요한 태그(tag)를 직접 만들어 웹 문서에 사용할 수 있다. 사용자는 임의의 태그를 지정하여 문서를 구조화 할 수 있으며 컴퓨터 프로그램이나 스크립트(script)는 이렇게 사용자가 정의한 태그를 이용하여 다양한 작업을 수행할 수 있게 된다. XML의 이러한 확장성으로 인하여 XML은 언어를 표현하는 언어 역할인 메타언어(metalanguage)의 기능을 갖는다.

XML은 임의의 문서 구조를 지정할 수 있는 반면에 문서 구조가 갖는 의미에 대해서는 정의하지 않으며 따라서 표현한 문서의 의미를 해석하는 역할은 하지 않는다. RDF는 이러한 문서의 의미를 지정하는 기본적인 역할을 한다.

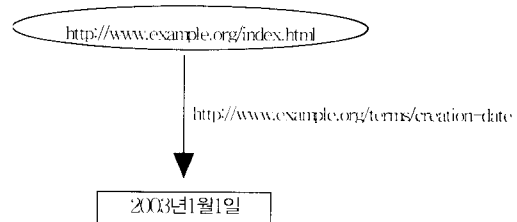
#### 3.2 RDF와 RDFS

RDF(Resource Description Framework)는 W3C의 가장 기본적인 시맨틱 웹 언어로서 웹에 있는 자원(resource)에 관한 메타정보를 표현하기 위한 언어이다[36]. 특히 웹 자원을 표현하는데 기본이 되는 제목, 저자, 최종 수정일, 저작권과 같은 웹 문서에 관한 메타데이터를 표현할 목적으로 개발되었

으나, 웹 자원의 개념을 웹 상에서 다른 것과 구별하여 식별할 수 있는 대상으로 일반화하면 다른 목적으로도 RDF를 활용할 수 있다. 이러한 식별 대상이 반드시 웹에 존재할 필요는 없다. XML을 기반으로 하고 있는 RDF는 매우 간단한 구조를 가지고 있다. URI를 써서 식별할 수 있는 대상으로 문장을 구성하며 구성된 문장은 노드와 화살표를 써서 도식적으로 표현할 수도 있다.

RDF는 기본적으로 세 개의 정보를 지닌 쌍들을 정의한다. 이 세 개의 정보는 일반 문장의 주어, 동사(또는 서술어), 목적어에 해당하는 것으로서 사람이나 웹 문서 등 특정 대상(object)이 특정 속성(attribute)에 대하여 특정 값(value)을 가지고 있는 상태를 표현한다.

RDF는 XML의 메타언어 기능을 사용하여 표현된다. 주어에 해당되는 대상과 목적어에 해당되는 값, 그리고 동사나 서술어에 해당되는 속성들을 모두 URI로 지정할 수 있다. 따라서 사용자는 새로운 개념이나 동사를 URI를 써서 손쉽게 정의할 수 있는 것이다. RDF Primer[36]에 나오는 예로서 "http://www.example.org/index.htm가 2003년1월1일에 만들어 졌다"라는 문장을 RDF로 표현하는 경우를 살펴보자. 이 문장은 "http://www.example.org/index.htm은 creation-date라는 속성을 가지고 있으며 그 속성의 값은 2003년1월1일이다"라고 표현될 수 있다. 이는 RDF의 도식화 규칙에 의해 다음과 같이 표현된다.



앞의 문장의 구성 요소를 분석하면 다음과 같다.

|         |                                            |
|---------|--------------------------------------------|
| 주어(자원)  | http://www.example.org                     |
| 서술어(속성) | http://www.example.org/terms/creation-date |
| 목적어(값)  | "2003년1월1일"                                |

이를 RDF를 써서 다음과 같이 표현할 수 있다.

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/
1999/02/22-rdf-syntax-ns#"
xmlns:extterms="http://www.
example.org/terms/">
<rdf:Description rdf:about="http://
www.example.org/index.html">
<extterms:creation-date>2003년1
월1일</extterms:creation-date>
</rdf:Description>
</rdf:RDF>
    
```

위의 RDF 문서는 앞서 표현한 도식을 XML로 표현한 것이다. RDF는 이러한 표현에 필요한 잘 정의된 구문을 제공한다. 위 문서의 첫번째 줄은 이 문서 내용이 XML 되어 있다는 것과 XML 버전을 선언한다. 다음의 rdf:RDF 엘리먼트는 </rdf:RDF>로 끝나는 부분까지가 RDF 표현식이란 것을 나타내며 동시에 xmlns를 써서 XML의 namespace를 지정한다. 다시 말해 rdf:로 시작하는 태그들은 URI http://www.w3.org/1999/02/22-rdf-syntax-ns#로 지정된 namespace에, extterms:로 시작하는 태그들은 http://www.example.org/terms/에 속한 것들이라는 것을 나타낸다. 이러한 태그들을 써서 앞의 그래프로 표현한 내용을 4번째부터 6번째 줄에 나타내고 있다. rdf:Description 태그는 자원의 서술이 시작됨을 알리고 서술 대상이 about에 지정한 주어임을 나타낸다. 다음 줄에서는 이 주어의 속성인 extterms:creation-date가 목적어로 표현된 "2003년1월1일"이란 값을 가짐을 나타낸다. RDF는 이렇게 자원-속성-값의 표현을 반복하거나 중첩하여 자원에 대한 서술을 하게 된다. 다음은 RDF 구문의 문법이다.

- [1] RDF ::= [ '<rdf:RDF>' ] description\* [ '</rdf:RDF>' ]
- [2] description ::= '<rdf:Description idAboutAttr?>' propertyElt\* '</rdf:Description>'
- [3] idAboutAttr ::= idAttr | aboutAttr
- [4] aboutAttr ::= 'about="' URI-reference ""

- [5] idAttr ::= 'ID="' IDsymbol ""
- [6] propertyElt ::= '<' propName '>' value '</' propName '>' | '<' propName resourceAttr '>'
- [7] propName ::= QName
- [8] value ::= description | string
- [9] resourceAttr ::= 'resource="' URI-reference ""
- [10] QName ::= [ NSprefix ':' ] name
- [11] URI-reference ::= string, interpreted per[URI]
- [12] IDsymbol ::= (any legal XML name symbol)
- [13] name ::= (any legal XML name symbol)
- [14] NSprefix ::= (any legal XML name space prefix)
- [15] string ::= (any XML text, with "<", ">", and "&" escaped)

RDF에서 제공하는 태그 외에도 사용자가 필요에 따라 새로운 용어를 정의하는 기능이 필요하다. 이러한 용어는 ex:Person과 같은 새로운 대상(thing)이나 ex:age와 같은 새로운 속성(property), 그리고 xsd:integer와 같이 표현 대상이 취할 수 있는 값의 유형(type) 등을 정의하는데 필요하다. 이러한 새로운 용어를 정의하기 위해서 RDF Schema[35]를 사용한다.

### 3.3 DAML+OIL과 OWL

DAML+OIL[6]은 웹 온톨로지 언어로서 DAML (DARPA Agent Markup Language) 프로그램의 DAML-ONT[7]와 주로 유럽에서 개발된 OIL (Ontology Inference Layer)[5]의 결합을 통하여 만들어졌다. OIL의 개발에는 On-To- Knowledge 콘소시움[41]의 구성원이 다수 참여 하기도 했다 [26]. DAML+OIL 웹 온톨로지 언어는 현재 W3C에 의해서 OWL[20][21] 웹 온톨로지 언어로 계승 발전되고 있다. OWL은 DAML+OIL의 네임 스페이스와 속성 클래스 이름 등을 변경하고 RDF 및 RDF Schema의 변화를 수용하였다. 이 장에서는 주로 OWL의 모태인 DAML+OIL을 소개한다.

온톨로지 언어로서 DAML+OIL은 관심영역(domain)의 구조를 서술하기 위한 목적을 갖는다. 이러한 구조는 객체지향적인 방법으로 클래스(class)와 속성(property)을 써서 표현된다. 온톨로지는 클래스와 속성의 성격을 서술한 공리(axiom)의 집합으로 구성된다. 자원(resource)이 특정 DAML+OIL 클래스의 인스턴스(instance)이거나 특정 속성을 갖는 것을 표현할 때는 이러한 목적에 적합한 RDF를 사용한다.

DAML+OIL은 기본적으로 표현력을 강조한 Description Logic이라 할 수 있다[25].

Description Logic과 마찬가지로 DAML+OIL의 클래스들은 URI로 지칭되는 이름이나, 클래스 수식을 만들기 위한 다양한 구성자(constructor)로 표현된 수식(expression)을 나타낸다. 따라서 DAML+OIL의 표현력은 이러한 클래스 구성자와 사용 가능한 공리의 종류에 의해서 결정된다. 다음 표는 DAML+OIL의 클래스 구성자들을 요약한 것이다[26].

| 구성자(constructor) | Description Logic 구문      | 사용 예                      |
|------------------|---------------------------|---------------------------|
| intersectionOf   | $C_1 \cap \dots \cap C_n$ | Human $\cap$ Male         |
| unionOf          | $C_1 \cup \dots \cup C_n$ | Doctor $\cup$ Lawyer      |
| complementOf     | $\neg C$                  | $\neg$ Male               |
| oneOf            | $\{x_1, \dots, x_n\}$     | {john, mary}              |
| toClass          | $\forall P.C$             | $\forall$ hasChild.Doctor |
| hasClass         | $\exists P.C$             | $\exists$ hasChild.Lawyer |
| hasValue         | $\exists P.\{x\}$         | $\exists$ citizenOf.{USA} |
| minCardinalityQ  | $\geq nP.C$               | $\geq 2$ hasChild.Lawyer  |
| maxCardinalityQ  | $\leq nP.C$               | $< 1$ hasChild.Male       |
| cardinalityQ     | $= nP.C$                  | $= 1$ hasParent.Female    |

위의 문장에서 Human  $\cap$  Male을 DAML+OIL으로 표현하면 다음과 같이 표현된다.

```

<daml:Class>
<daml:intersectionOf rdf:parseType="
daml:collection">
<daml:Class rdf:about="#Human"/>
<daml:Class rdf:about="#Male"/>
</daml:intersectionOf>
</daml:Class>
    
```

DAML+OIL의 공리는 클래스나 속성간의 포섭

(subsumption)이나 동치(equivalence)등의 다양한 관계 비롯하여 개별 자원간의 동치 여부와 같은 다양한 성격을 선언하는데 쓰인다.

DAML+OIL에서 사용할 수 있는 공리는 subClassOf, sameClassAs, subPropertyOf, samePropertyAs, disjointWith, sameIndividualAs, differentIndividualFrom, inverserOf, transitiveProperty, uniqueProperty, unambiguousProperty등이 있다.

DAML+OIL은 XML Schema의 모든 데이터형을 지원하여 문자열, 십진수, 실수, 정수 범위 등을 사용할 수 있다. DAML+OIL은 또한 RDFS와 밀접하게 연관되어 있다. DAML+OIL은 기본적으로 Description Logic의 추론 능력과 표현력을 가지고 있기 때문에 기존의 Description Logic의 연구 결과를 활용할 수 있는 장점을 가지고 있다.

### 3.4 다른 웹 온톨로지 언어들

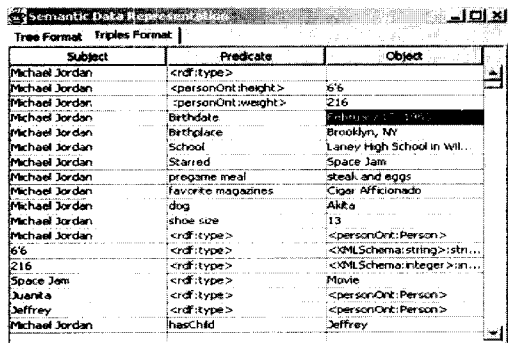
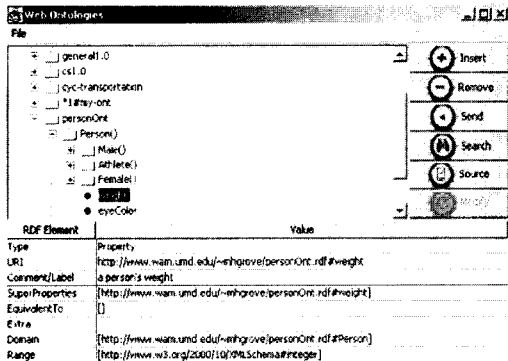
앞서 DAML+OIL과 OWL을 주요 시맨틱 웹 온톨로지 언어로서 소개했으나 지난 수년 간에 개발된 다양한 온톨로지 언어들이 있다. XML 문법을 기반으로 한 Ontology Exchange Language(XOL)[42]와 Simple HTML Ontology Extension(SHOE)[43], Ontology Markup Language(OML)[44] 등이 그것이다. XOL은 초기에 바이오인포매틱스 온톨로지의 교환을 위하여 만들어졌으나 다른 영역에도 활용될 수 있다. SHOE는 HTML을 확장하여 에이전트에 의한 자동처리를 염두에 두고 개발되었으나 DAML+OIL의 개발로 더 이상 활발한 연구가 이루어지고 있지는 않는 상태이다. OML도 SHOE를 기반으로 개발되어 OML Core, Simple OML, Abbreviated OML, Standard OML의 4가지 단계를 제공하나 OML을 위한 저작도구가 따로 존재하지 않는 등 상대적으로 발전 속도가 늦은 편이다. 이들에 대한 비교 자료는 참고 문헌[4]을 참조하기를 권한다.

## 4. 온톨로지의 활용

온톨로지의 설계, 평가, 검사, 수정, 유지보수, 통합, 적용 등의 온톨로지 생명주기 전 과정에서 온톨로지를 효율적으로 개발하고 사용하는 방안을 연구하는 분야를 온톨로지 공학(ontological engineer-

ring)이라 한다[2]. 온톨로지 공학적 측면은 이번 특집호의 다른 논문의 주제이기도 하므로 여기서는 온톨로지의 활용에 대해 단편적으로 소개하고자 한다.

시맨틱 웹 온톨로지는 앞서 언급한 대로 소규모 온톨로지의, 다수에 의한 광범위한 개발을 추구하고 있다. 이를 실현하기 위해 해결할 과제 중의 하나는 온톨로지에 친숙하지 않은 다수의 대중 사용자들이 온톨로지를 자연스럽게 만들어 쓰게 하는 일이다[8][10]. 이를 위해서는 무엇보다 온톨로지를 작성하는 저작 작업이 손쉽게 이루어져야 한다. 온톨로지 저작 도구의 예로 SMORE(Semantic Markup, Ontology and RDF Editor)[9]을 소개한다. SMORE는 Java로 구현되어 WYSIWYG HTML 편집기와 온톨로지를 검색할 수 있는 웹 브라우저 역할을 한다.



또한 그래픽 사용자 인터페이스를 써서 온톨로지를 생성하고 관리할 수 있는 기능을 가지고 있다. 이 프로그램의 소스는 공개되어 있으므로 유사한 프로그램 개발에 도움이 될 것으로 생각된다.

시맨틱 웹 온톨로지는 응용 프로그램에서도 접근이 가능하다. 예를 들어, Jena[45]는 시맨틱 웹 온톨로지의 기반이 되고 있는 RDF를 Java 응용프로그램에서 사용하기 위한 RDF 응용 프로그램 인터페이스(API)를 제공한다. 따라서 RDF API를 이용하면 RDF 그래프를 생성하고 질의하고 저장하는 작업을 Java 프로그램에서 할 수 있게 된다. 특히 온톨로지 질의를 위해서 SQL과 유사한 문법을 가진 RDQL질의 언어를 가지고 있다. Jena는 또한 DAML API도 제공하여 제한적이거나 추론 기능을 포함한 DAML 온톨로지 작업을 가능하게 한다. 온톨로지를 하나의 지식베이스로 간주하여 데이터베이스와 같은 질의를 가능하게 하는 것도 중요한 온톨로지의 활용 분야이다[27]. DQL[11]은 시맨틱 웹을 위한 질의 언어로서 DAML+OIL/OWL로 표현된 지식을 이용하여 에이전트와 에이전트간의 질의 및 응답이 가능하도록 한다.

이러한 온톨로지 저작이나 API에 의한 접근 및 질의 외에도 온톨로지 학습 문제[18], 이질적 온톨로지 간의 변환과 사상 문제, 온톨로지의 검증과 시각화 등의 관리 문제[12], 온톨로지 등록저장(repository), 검색, 질의를 지원하는 온톨로지 서비스 문제[17] 등의 다양한 주제로 논의가 활발하게 이루어지고 있다.

## 5. 결론

온톨로지에 관한 연구는 인공지능 분야의 시작과 함께 지식 표현 분야의 핵심으로 활발히 연구가 이루어져 온 분야이다[46]. 시맨틱 웹의 출현과 더불어 온톨로지의 중요성이 인식되면서 새롭게 주목받는 온톨로지는 새로운 연구 분야가 아닌 새로운 응용 분야로 보는 것이 타당할 것이다. 웹이라는 다수의 사용자가 광범위하게 손쉽게 때와 장소를 가리지 않고 공유할 수 있는 지식의 체계는 인류 문명이 추구해온 기본적인 목표라고 할 수 있다. 바야흐로 시맨틱 웹이라는 지식 표준과 활용 기술이 이러한 꿈을 현실화 시킬 수 있을지 기대되는 바가 크다. 시맨틱 웹의 실현에 따른 기술적, 사회적 파급 효과는 90년대를 휩쓴 웹의 영향보다도 훨씬 지대할 것으로 예측되는 바, 시맨틱 웹의 핵심이라고 할 수 있는 온톨로지에 대한 연구가 활성화되기를 기대한다.

## 참고문헌

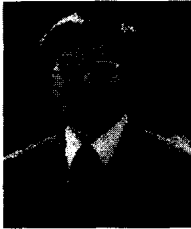
- [ 1 ] Dieter Fensel, James A. Hendler, Henry Lieberman and Wolfgang Wahlster, "Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential." MIT Press, February, 2003.
- [ 2 ] Michael Gruninger and Jintae Lee, "Ontology applications and design: Introduction," Communications of the ACM, vol.45, no.2, February, 2002, pp. 39-41.
- [ 3 ] Fensel, D.; Musen, M.A., "The Semantic Web: A Brain for Humankind." IEEE Intelligent Systems, vol.16, no.2, March/April, 2001, pp.24-25.
- [ 4 ] Asuncion Gomez-Perez and Oscar Corcho, "Ontology Languages for the Semantic Web," IEEE Intelligent Systems, vol.17, no.1, January/February, 2002, pp.54-60.
- [ 5 ] Dieter Fensel, Frank van Harmelen, Ian Horrocks, Debora L. McGuinness, Peter F. Patel Schneider, "OIL: An Ontology Infrastructure for the Semantic Web," IEEE Intelligent Systems, vol.16, no.2, March/April, 2001, pp.38-45.
- [ 6 ] Debora L. McGuinness, Richard Fikes, James Hendler and Lynn Andrea Stein, "DAML-OIL: An Ontology Language for the Semantic Web," IEEE Intelligent Systems, vol.17, no.5, September/October, 2002, pp.72-80.
- [ 7 ] Debora L. McGuinness et al., "DAML-ONT: An Ontology Language for the Semantic Web," <http://www.daml.org/2000/10/daml-ont.html>.
- [ 8 ] Clyde W. Holsapple, K. D. Joshi, "A Collaborative Approach to Ontology Design," Communications of the ACM, vol.45, no.2, February, 2002, pp.42-47.
- [ 9 ] Aditya Kalyanpur, Bijan Parsia, James Hendler, Jennifer Golbeck, "SMORE-Semantic Markup, Ontology, and RDF Editor," <http://www.mindswap.org/papers/SMORE.pdf>.
- [10] Marcelo Tallis, Neil M. Goldman, and Robert M. Balzer, "The Briefing Associate: Easing Authors into the Semantic Web," IEEE Intelligent Systems, vol.17, no.1, January/February, 2002, pp.26-32.
- [11] Richard Fikes, Pat Hayes, and Ian Horrocks, "DQL-A Query Language for the Semantic Web," Knowledge Systems Laboratory, Stanford University, 2002, [ftp://ftp.ksl.stanford.edu/pub/KSL\\_Reports/KSL-02-05.doc](ftp://ftp.ksl.stanford.edu/pub/KSL_Reports/KSL-02-05.doc)
- [12] Deborah L. McGuinness, Richard Fikes, James Rice, and Steve Wilder. "An Environment for Merging and Testing Large Ontologies". In Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR2000), Breckenridge, Colorado, USA, April, 2000.
- [13] Steve Pepper and Graham Moore (Eds.), "XML Topic Maps (XTM) 1.0, TopicMaps.Org Specification," <http://www.topicmaps.org/xtm/1.0/>.
- [14] Jennifer Golbeck, Michael Grove, Bijan Parsia, Aditya Kalyanpur, and James Hendler, "New Tools for the Semantic Web," <http://www.mindswap.org/papers/EKAW02.doc>
- [15] Klein, M., "XML, RDF, and Relatives," IEEE Intelligent Systems, vol.16, no.2, March/April, 2001, pp.26-28.
- [16] James Hendler, "Agents and the Semantic Web," IEEE Intelligent Systems, vol.16, no.2, March/April, 2001, pp.30-37.
- [17] S.A. McIlraith, T.C. Son, and Honglei Zeng, "Semantic Web Services," IEEE Intelligent Systems, vol.16, no.2, March/April, 2001, pp.46-53.
- [18] A. Maedche and S. Staab, "Ontology Learning for the Semantic Web," IEEE Intelligent Systems, vol.16, no.2, March/April, 2001, pp.72-79.
- [19] T. R. Gruber, "A Translation Approach to Portable Ontologies," Knowledge Acquisition, 5(2):199-220, 1993.



- [20] Mike Dean et al. (Eds.), "Web Ontology Language (OWL) Reference Version 1.0," W3C Working Draft 12 November, 2002, <http://www.w3.org/TR/owl-ref/>.
- [21] Michael K. Smith et al., "Web Ontology Language (OWL) Guide Version 1.0," W3C Working Draft 4 November, 2002, <http://www.w3.org/TR/owl-guide/>.
- [22] Peter F. Patel-Schneider et al. (Eds.), "Web Ontology Language (OWL) Abstract Syntax and Semantics," W3C Working Draft 3 February, 2003, <http://www.w3.org/TR/owl-semantics/>.
- [23] Jeff Heflin (Ed.), "Web Ontology Language (OWL) Use Cases and Requirements," W3C Working Draft 3 February, 2003, <http://www.w3.org/TR/webont-req/>.
- [24] Ian Horrocks, "DAML+OIL: a Reasonable Web Ontology Language," Proc. of EDBT 2002, pp. 2-13, Lecture Notes in Computer Science, vol. 2287. Springer Verlag., 2002.
- [25] Ian Horrocks, "DAML+OIL: A Description Logic for the Semantic Web," IEEE Bulletin of the Technical Committee on Data Engineering, vol.25, no.1, pp. 4-9, 2002.
- [26] Peter Patel-Schneider, Ian Horrocks, and Frank van Harmelen, "Reviewing the Design of DAML+OIL: An Ontology Language for the Semantic Web," AAAI'02, 2002.
- [27] Ian Horrocks and Sergio Tessaris, "Querying the Semantic Web: a Formal Approach," Proc of 1st International Semantic Web Conference, Sardinia, Italy, June 9-12, 2002. Lecture Notes in Computer Science, Vol. 2342, Springer-Verlag, 2002.
- [28] Ian Horrocks Reasoning with Expressive Description Logics: Theory and Practice Proceedings of CADE-02, 2002. Springer-Verlag Lecture Notes in Artificial Intelligence, LNAI 2393, pp. 1--15.
- [29] Michael Wooldridge and Nicholas R. Jennings. Intelligent Agents: Theory and Practice. The Knowledge Engineering Review, vol. 10(2) pp. 115-152, 1995.
- [30] CYCR Ontology Guide, <http://www.cyc.com/cyc-2-1/toc.html>
- [31] OpenCyc, <http://www.opencyc.org/>
- [32] WordNet, <http://www.cogsci.princeton.edu/~wn/>
- [33] Extensible Markup Language (XML) 1.0, W3C Recommendation 10-February-1998 <http://www.w3.org/TR/1998/REC-xml-19980210.html>
- [34] Resource Description Framework (RDF), <http://www.w3.org/RDF/>
- [35] RDF Vocabulary Description Language 1.0: RDF Schema, W3C Working Draft 23 January, 2003. <http://www.w3.org/TR/rdf-schema/>
- [36] RDF Primer, W3C Working Draft 23 January 2003, <http://www.w3.org/TR/rdf-primer/>
- [37] DARPA Agent Markup Language (DAML), <http://www.daml.org/>
- [38] Yahoo! Korea, <http://www.yahoo.co.kr/>
- [39] Google 한글, <http://www.google.co.kr/>
- [40] ISO/IEC JTC1/SC34 Web Service, <http://www.ornl.gov/sgml/sc34/>
- [41] On-To-Knowledge, <http://www.ontoknowledge.org/>
- [42] R. Karp, V. Chaudhri, and J. Thomere, "XOL Ontology Exchange Language," <http://www.ai.sri.com/pkarp/xol/>
- [43] SHOE, <http://www.cs.umd.edu/projects/plus/SHOE/>
- [44] Ontology Markup Language, <http://www.ontologos.org/OML/OML%200.3.htm>
- [45] Brian McBrids, "Jena: A Semantic Web Toolkit," IEEE Internet Computing, November/December, 2002, pp.55-59.
- [46] Stuart Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. Prentice-Hall, 1995.

---

**이 재 호**



1985 서울대학교 계산통계학과(이학 학사)  
1987 서울대학교 계산통계학과(이학 석사)  
1997 미시간대학교 컴퓨터공학, 미국 엔지니어(공학 박사)  
1996. 8~1998. 2 수석엔지니어, ORIN CON, 미국 샌디에고 소재  
1998. 3~현재 서울시립대학교 전자전기 컴퓨터공학부 교수  
E mail : jachho@uos.ac.kr

---

**• 제13회 통신정보 합동 학술대회 •**

- 일 자 : 2003년 4월 30일~5월 2일
- 장 소 : 안면도 롯데오션캐슬
- 주 최 : 한국정보과학회 · 한국통신학회 · 대한전자공학회  
한국통신정보보호학회 · 한국정보처리학회
- 문 의 처 : 서울대 이광복 교수 (Tel. 02-880-8415)  
<http://jcci21.or.kr>

**Japan-Korea Joint Workshop on**

**• Algorithms and Computation •**

(워크샵 개최)

- 일 자 : 2003년 7월 3~4일
- 장 소 : 일본 동북대학(Tohoku University, Sendai, Japan)
- 주 최 : 컴퓨터이론연구회
- 상세안내 : <http://www.dais.is.tohoku.ac.jp/waac03/index.html>