

온톨로지 기술 언어

어떤 대상을 디지털 공간에 재현하기 위해 만든 온톨로지는 같은 분야의 지식을 다루는 여러 사람들이 지식의 소통을 위한 공동의 약속으로 사용할 때 큰 효용성을 발휘한다. 예를 들자면, 월드와이드웹의 세계 안에는 수많은 사람들이 개인적으로 작성한 ‘문화유산 답사기’가 존재하는데, 그 글 속에 담긴 객관적 사실 정보들이 앞에서 예시한 바와 같은 ‘문화유산 온톨로지’에 의해 기술될 수 있다면, 그 결과는 월드와이드웹 속에 ‘한국의 문화유산에 관한 시맨틱 웹’이 만들어지는 것이고, 우리는 그것을 통해 더욱 풍성한 문화유산 관련 지식을 얻을 수 있게 될 것이다.

특정 분야의 지식에 관한 온톨로지가 여러 사람들, 또는 다수의 컴퓨터 시스템 사이에서 공유될 수 있기 위해서는 그 설계의 내용이 명시적이고 표준적인 언어로 기술되어야 한다. 앞에서 살펴보았던 RDF(Resource Description Framework)는 가장 기본적인 형태의 온톨로지 기술 언어에 속한다고 할 수 있다. RDF는 단순히 주어와 목적어 사이의 관계를 서술하는 것을 위주로 한다. 그런데 그 관계성이 어떤 대상들 사이에서 유효한지, 그리고 다양한 관계 서술어 사이에는 또 어떤 상관관계가 존재하는지 등도 점검할 필요성이 제기됨에 따라 RDF보다 좀 더 확장된 기능의 온톨로지 기술 언어들이 등장하였다.

학계, 산업계에서 실험적으로, 또는 특정 영역에서 실용적으로 쓰이는 온톨로지 기술 언어는 매우 다양하지만, 월드와이드웹 상에서 운용되는 시맨틱 웹의 구현을 목적으로 하는 온톨로지 기술 언어로는 RDF를 확장시켜 나아간 RDFS(RDF Schema)와 OWL(Web Ontology Language)이 표준적인 약속으로 권장되고 있다.¹⁾ RDF, RDFS, OWL은 별개의 언어가 아니라, 먼저 정해놓은 약속을 기반으로 하면서 더 유용한 수단으로 발전시켜 간 것이라고 할 수 있다.

1) RDF Schema: Resource Description Framework Schema. RDFS, RDF(S), RDF-S, RDF/S 등 다양한 형태의 약어로 지칭되고 있다. RDFS는 RDF 형식으로 표현하고자 하는 대상의 온톨로지를 기술하는 어휘(vocabulary)를 제공한다. RDF Schema는 1998년 W3C에 의해 처음 제안되었으며, 현재 통용되는 권고안은 2004년에 발표되었다.
<https://www.w3.org/TR/rdf-schema/>
OWL(Web Ontology Language)는 RDFS보다 확장된 온톨로지 기술 언어이다. W3C에 의해 2004년에 제안되었으며, 이것을 부분적으로 수정·확장한 OWL2는 2009년에 발표되었다.
<https://www.w3.org/TR/owl2-overview/>

```

xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"

```

OWL 문서에서 기본적으로 참조하는 이름 공간(Name Space)

OWL은 얼마나 복잡하고 정교한 온톨로지를 설계하느냐에 따라 사용하는 어휘의 종류가 달라질 수 있지만, 핵심적인 것은 클래스, 속성, 관계성, 개체를 선언하고, 각각의 개체에 속성을 지정하거나 개체 사이의 관계성을 지정하는데 필요한 어휘들이다.

OWL 어휘	의미
owl:Ontology	온톨로지 선언
owl:Class	클래스 선언
owl:ObjectProperty	관계성 선언
owl:DatatypeProperty	속성 선언
owl:NamedIndividual	개체 선언
rdfs:subClassOf	서브 클래스 선언
rdfs:domain	정의역 지정
rdfs:range	치역 지정
rdf:datatype	데이터 타입 지정

OWL(Web Ontology Language)의 주요 어휘

온톨로지 기술 언어를 사용하면 우리가 앞에서 예시로 다루었던 ‘담양 소쇄원’ 등의 문화유산 관련 정보를 월드와이드웹 상에서 통용되는 시맨틱 데이터로 기술할 수 있다.

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns="http://digerati.aks.ac.kr/terms/CHoKorea#"
  xml:base="http://digerati.aks.ac.kr/terms/CHoKorea"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:dc="http://purl.org/dc/terms/"
  xmlns:CHoKorea="http://digerati.aks.ac.kr/terms/CHoKorea#">

```

```

<!-- 온톨로지 선언-->

<owl:Ontology rdf:about="http://digerati.aks.ac.kr/terms/CHoKorea"/>

<!-- 클래스 선언 예시: '문화유산' 클래스를 선언-->

<owl:Class rdf:about="http://digerati.aks.ac.kr/terms/CHoKorea#문화유산"/>

<!-- 속성 선언 예시: '인물' 클래스의 개체는 '생년' 속성을 갖는다-->

<owl:DatatypeProperty rdf:about="http://digerati.aks.ac.kr/terms/CHoKorea#생년">
  <rdfs:domain rdf:resource="http://digerati.aks.ac.kr/terms/CHoKorea#인물"/>
</owl:DatatypeProperty>

<!-- 관계성 선언 예시:
'isLocatedIn'이라는 관계성은 '문화유산'과 '장소' 클래스 개체 사이에 성립할 수 있다 -->

<owl:ObjectProperty rdf:about="http://digerati.aks.ac.kr/terms/CHoKorea#isLocatedIn">
  <rdfs:domain rdf:resource="http://digerati.aks.ac.kr/terms/CHoKorea#문화유산"/>
  <rdfs:range rdf:resource="http://digerati.aks.ac.kr/terms/CHoKorea#장소"/>
</owl:ObjectProperty>

<!-- 개체 선언 및 속성/관계성 지정 예시:
개체 '양산보'는 ....의 속성을 가지며 개체 '양택지' 등과 ....의 관계성을 갖는다-->

<owl:NamedIndividual rdf:about="http://digerati.aks.ac.kr/terms/CHoKorea#양산보">
  <rdf:type rdf:resource="http://digerati.aks.ac.kr/terms/CHoKorea#인물"/>
  <생년 rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">1503</생년>
  <몰년 rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">1557</몰년>
  <한자 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">梁山甫</한자>
  <hasDescendant rdf:resource="http://digerati.aks.ac.kr/terms/CHoKorea#양택지"/>
  <isDiscipleOf rdf:resource="http://digerati.aks.ac.kr/terms/CHoKorea#조광조"/>
</owl:NamedIndividual>

</rdf:RDF>

```

OWL 어휘를 XML/RDF 문법으로 표현한 웹 온톨로지 기술 문서

이와 같은 웹 온톨로지 기술 문서에서 우리가 의미 있게 보아야 할 것은 대상 세계가 존재하는 모습을 표현하는 모든 데이터가 유일하게 식별될 수 있는 이름을 갖도록 했다는 점이다.

이 온톨로지 전체가 'http://digerati.aks.ac.kr/terms/CHoKorea'라는 고유한 이름을 부여받았으며, RDF 문에서 동사처럼 쓰이는 관계성, 속성 서술어, 그리고 주어와 목적어가 되는 모든 개체가 온톨로지 이름을 접두어로 삼는 고유한 식별자를 갖는다. 실제 세계에서 쓰이는 이름과 용어가 있는데도, 이와 같이 작위적인 유일 식별자를 또 다시 부여하는 이유는 무엇인가? 어떤 이름이 어떤 대상을 가리키는지 분명하게 알 수 있게 하기 위해서이다. 여기서 '분명하게 안다'는 것은 사람의 능력이 아니라 컴퓨터 시스템의 기능을 두고 하는 말이다.