 UNIVERSITY OF SEOUL

2002년 한국지능정보시스템학회
춘계학술대회
부산 BEXCO

시맨틱 웹(Semantic Web)의 개요

이재호
서울시립대학교 전자전기컴퓨터공학부
jaeho@uos.ac.kr
2002. 5. 10.

춘계 KIISS-2002 1


2

웹과 시맨틱웹

<p>■ 웹(Web)</p> <ul style="list-style-type: none">▶ 1989년, Tim Berners-Lee▶ 1991년 일반 공개▶ URIs, HTML, HTTP	<p>■ 시맨틱 웹(Semantic Web)</p> <ul style="list-style-type: none">▶ 1999년, Tim Berners-Lee (W3C)▶ 내용 중심의 정보 표현과 해석▶ 기계(agent)에 의한 정보 추출, 통합, 가공, 추론▶ XML, RDF, Ontology, 에이전트
--	--

■ **World Wide Web Consortium (W3C)**

- ▶ 1994년 Tim Berners-Lee, MIT, CERN
- ▶ XML, XML Namespace
- ▶ DTD, XML Schema, DOM(API), XPath, XPointer, XSLT, XML Query
- ▶ XLink, XHTML, SVG, SMIL, MathML, XForms

춘계 KIISS-2002  UNIVERSITY OF SEOUL
서울시립대학교

개요

■ 인간 뿐만 아니라 기계도 정보를 이해할 수 있도록 하는 기술

- ▶ 현재 대부분의 웹 정보는 인간의 소비를 전제로 개발
- ▶ 기계가 인식하여 처리하기 위해서는 가공되지 않은 원 데이터로 본에 후 재처리 필요

■ 웹 상에 존재하는 데이터를 컴퓨터 판독이 용이하게 표현

■ 데이터간의 연관 관계를 체계적으로 정의

■ 자동화되고 통합된 데이터의 공유를 제공

입형과 내용

	입형	내용
입상 예	Fax	Computer File
	PDF	Power Point
	JPEG	SVG
생산지	Marketing	Domain Expert
소비지	Human	Machine
표현	Procedural (how)	Declarative(what)
기술	WWW	Semantic Web
표현식	Table, Fixed Tags	XML(hierarchical) RDF(relational)

관련 사이트: 일반

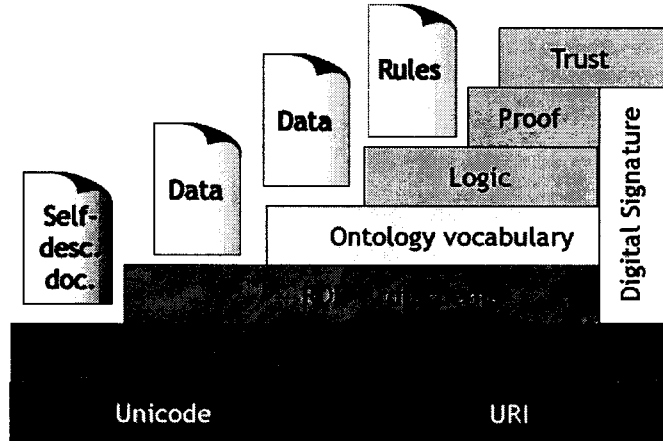
- SemanticWeb.Org (The Semantic Web Community Portal)
 - ▶ <http://www.semanticweb.org/>
- World Wide Web Consortium (W3C)
 - ▶ <http://www.w3.org/>
- Resource Description Framework (RDF)
 - ▶ <http://www.w3.org/RDF/>
- Web-Ontology (WebOnt) Working Group
 - ▶ <http://www.w3.org/2001/sw/WebOnt/>
- Extensible Markup Language (XML)
 - ▶ <http://www.w3.org/XML/>

관련 학술대회

- 1st International Semantic Web Conference (ISWC2002)
 - ▶ <http://iswc.semanticweb.org/>
- Workshop on "Web Semantics" (WEBS)
 - ▶ <http://www.dexa.org/WEBScallforpaper.html>
- International Conference on Web Intelligence
 - ▶ <http://kis.maebashi-it.ac.jp/wi01/>
- Semantic Web Conference
 - ▶ <http://www.net.intap.or.jp/INTAP/english/what-new/seminarOTO29.html>
- SEMWEB@KR2002
 - ▶ <http://profs.sci.univr.it/~cristani/semweb/>
- International Conference on Electronic Commerce
 - ▶ <http://icec.net>
- 웹 기반 기술 표준 컨퍼런스
 - ▶ <http://www.nca.or.kr>
- The Pacific Asian Conference on Intelligent Systems
 - ▶ <http://www.pais2001.org>

연구배경: 시맨틱 웹의 구조

(Berners-Lee의 XML 2000 학술대회 발표자료)



시맨틱 웹(Semantic Web)의 요소기술

■ 데이터:

▶ XML

■ 자원

▶ RDF, RDFS

■ 지식

▶ 온톨로지

■ 추론

▶ 인공지능

▶ 에이전트

XML (eXtensible Markup Language)

- 컴퓨터 판독(computer-readable)이 용이한 문서 작성 규정
- 마크업(markup)
 - ▶ HTML과는 달리 XML에서는 사용자가 필요한 태그(tag)를 직접 만들어 웹 문서에 사용
- 메타언어(meta language)의 기능
 - ▶ 언어를 표현하는 언어 역할
- Syntactic
 - ▶ 임의의 문서 구조를 지정할 수 있는 반면에 문서 구조가 갖는 의미에 대해서는 정의하지 않음

관련사이트: XML

- XML 1.0, W3C Recommendation, 10-February-1998
 - ▶ <http://www.w3.org/TR/1998/REC-xml-19980210.html#elemdecls>
- XML Data, Layman, et. al.
 - ▶ <http://www.w3.org/Submission/1998/01/>
- XML Schema Part 1: Structures
 - ▶ <http://www.w3.org/TR/2000/WD-xmlschema-1-20000225/>
- XML Schema Part 2: Datatypes
 - ▶ <http://www.w3.org/TR/2000/WD-xmlschema-2-20000225/>
- XML Linking Language (XLink),
 - ▶ <http://www.w3.org/TR/WD-xml-link>
- Namespaces in XML
 - ▶ <http://www.w3.org/TR/1999/REC-xml-names-19990114>

RDF (Resource Description Framework)

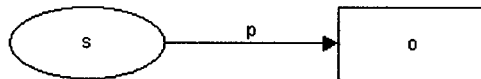
- **Semantic: 문서의 의미를 지정하는 역할**
- **XML로 표현**
- **의미 부여를 위해서는 데이터 뿐만 아니라 데이터에 대한 추론을 규정하는 규칙(rule)을 표현할 수 있는 언어**
- **세 개의 정보를 지닌 쌍들을 정의**
 - ▶ 일반 문장의 주어, 동사(시술어), 목적어에 해당
 - ▶ 사람이나 일 문서 등 특정 대상(object)이 특정 속성(attribute)에 대하여 특정 값(value)을 가지고 있는 것을 표현한다.
 - ▶ 주어에 해당하는 대상과 목적어에 해당하는 값, 그리고 동사에 해당하는 속성 이름 모두 URI로 지정
 - ▶ DAML+OIL (DARPA Agent Markup Language + Ontology Inference Layer) 언어와 같은 확장된 기능의 의미 부여 언어 등의 연구 진행

RDF 관련자료

- **RDF 모델 및 문법 규격**
 - ▶ <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>
- **What is RDF?**
 - ▶ <http://www.xml.com/pub/a/2001/01/24/rdf.html>
- **RDF Applications with Prolog**
 - ▶ <http://www.xml.com/pub/a/2001/07/25/prologrdf.html>
- **Using the Jena API to Process RDF**
 - ▶ Jena is a freely-available Java API for processing RDF.
 - ▶ <http://www.xml.com/pub/a/2001/05/23/jena.html>

RDF 데이터 모델

- 자원(Resources) 집합
- 문자(Literals) 집합
- 특성(Properties): 자원의 부분 집합
- 문장(Statements) 집합
 - ▶ 문장 각각의 요소(element)는 { pred, sub, obj }의 triple로 구성
 - ▶ pred : 특성의 멤버
 - ▶ sub : 자원의 멤버



RDF 문장의 예

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:s="http://purl.org/dc/elements/1.1/">
<rdf:RDF>
  <rdf:Description about="http://www.ecif.or.kr">
    <s:Creator rdf:resource="http://www.ecif.or.kr/staffId/85740"
      s:Name="webmaster" s:Email="webmaster@ecif.or.kr" />
  </rdf:Description>
</rdf:RDF>
  
```

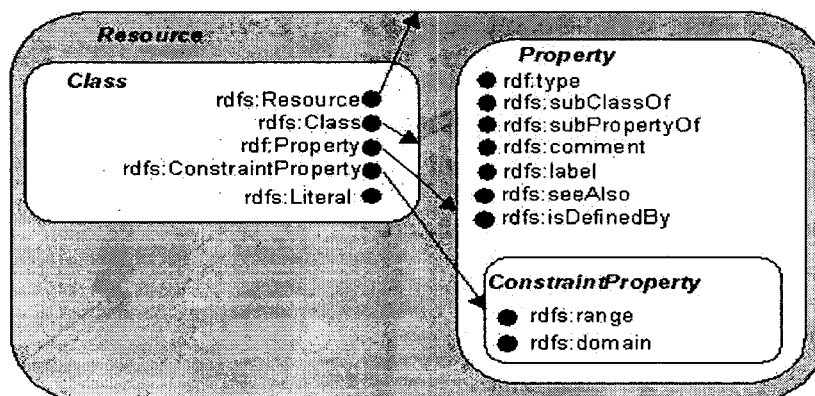
RDFS (RDF Schema)

■ RDF 스키마 명세 1.0

▶ <http://www.w3.org/TR/2000/CR-rdf-schema-20000327/>

- RDF 데이터 모델은 특성과 값을 명명함으로써 자원간의 상호 관계성을 기술
- RDF 스키마는 이러한 특성들을 선언하고 특성들과 다른 자원간의 관계 정의를 위한 메커니즘을 제공
- 스키마 지정 언어 (Schema Specification Language)
- RDF 스키마는 RDF 데이터 모델로 표현

RDFS: 클래스와 자원



온톨로지(Ontology) : Philosophy

- 존재론
- A systematic account of existence
- Epistemology is about knowledge and knowing

온톨로지(Ontology) : Knowledge Sharing

- T. R. Gruber. A translation approach to portable ontologies. Knowledge Acquisition, 5(2):199-220, 1993. http://ksl-web.stanford.edu/KSL_Abstracts/KSL-92-71.html.
- An ontology is an explicit specification of a conceptualization.
- An ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents.
- An ontological commitment is an agreement to use a vocabulary (i.e., ask queries and make assertions) in a way that is consistent (but not complete) with respect to the theory specified by an ontology.
- We build agents that commit to ontologies. We design ontologies so we can share knowledge with and among these agents.

온톨로지의 역할 (PAIS 2001, Vladan Devedzic)

■ Intelligent Services

- ▶ Information brokers, search agents, information filters, intelligent information integration, knowledge management

■ Higher Level of Interoperability

- ▶ Semantic interoperability

■ Explicit Inference Layer

- ▶ Higher level of abstraction

■ Semantic Markup

- ▶ Machine understandable

관련 사이트: 온톨로지

- DARPA Agent Markup Language (DAML)
 - ▶ <http://www.daml.org/>
- OpenCyc
 - ▶ <http://www.opencyc.org/>
- CYC® Ontology Guide
 - ▶ <http://www.cyc.com/cyc-2-1/toc.html>
- Web-Ontology (WebOnt) Working Group
 - ▶ <http://www.w3.org/2001/sw/WebOnt/>
- KSL Interactive Ontology Server
 - ▶ <http://www-ksi-svc.stanford.edu:5915/>
- SHOE (Simple HTML Ontology Extensions for agents)
 - ▶ <http://www.cs.umd.edu/projects/plus/SHOE/>
- Ontology.Org
 - ▶ <http://www.ontology.org/>

관련 사이트: 온톨로지

- DARPA Agent Markup Language (DAML)
 - ▶ <http://www.daml.org/>
- OpenCyc
 - ▶ <http://www.opencyc.org/>
- CYC® Ontology Guide
 - ▶ <http://www.cyc.com/cyc-2-1/toc.html>
- Web-Ontology (WebOnt) Working Group
 - ▶ <http://www.w3.org/2001/sw/WebOnt/>
- KSL Interactive Ontology Server
 - ▶ <http://www.ksl-svc.stanford.edu:5915/>
- SHOE (Simple HTML Ontology Extensions for agents)
 - ▶ <http://www.cs.umd.edu/projects/plus/SHOE/>
- Ontology.Org
 - ▶ <http://www.ontology.org/>

DAML (Darpa Agent Markup Language)

- XML-based semantic language that ties the information on a page to machine readable semantics
- <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:daml="http://www.daml.org/2000/12/daml+oil#"
 - xmlns="http://www.daml.org/2000/12/daml+oil-ex#" >
 - <daml:ObjectProperty rdf:ID="hasFather">
 - <rdfs:subPropertyOf rdf:resource="#hasParent"/>
 - <rdfs:range rdf:resource="#Male"/>
 - </daml ObjectProperty>
- 용 용:
 - ▶ ISI People page (<http://www.isi.edu/>)

DAML 관련자료

■ DARPA Agent Markup Language (DAML)

▶ <http://www.daml.org/>

■ Introduction to DAML: Part I

▶ <http://www.xml.com/pub/a/2002/01/30/daml1.html>

■ Introduction to DAML: Part II

▶ <http://www.xml.com/pub/a/2002/03/13/daml.html>

온톨로지 관련 연구

■ Topic Maps: www.topicmaps.com

▶ “GPS of the information universe”

▶ ISO standard for describing knowledge structures and associating them with information resources

• <http://www.gca.org/papers/xml/europe2000/papers/sTI-01.html>

• Topics, Associations, and Occurrences (TAO)

▶ Topic map sites:

<http://www.ontopia.net/topicmaps/sites.html>

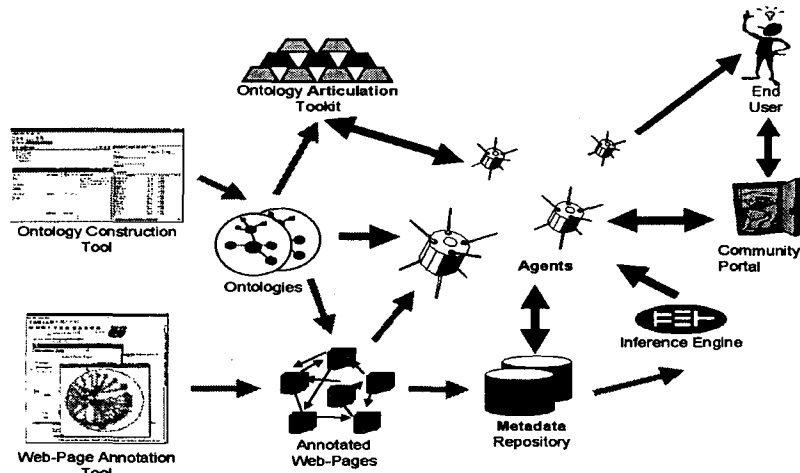
온톨로지 관련 연구

- **WordNet:** <http://www.cogsci.princeton.edu/~wn/>
 - ▶ An Electronic Lexical Database for the English Language
 - ▶ English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept
 - ▶ Different relations link the synonym sets (synsets)
 - ▶ Wordnet RDF Representation: <http://www.semanticweb.org/library/>
 - ▶ 한국어: <http://dewey.yonsei.ac.kr/memexlee/doc/LeeKim1999.htm>
- **GCIDE_XML:** <http://www.ibiblio.org/webster/>
 - ▶ GNU version of The Collaborative International Dictionary of English

온톨로지 관련 도구

- Cerebra™: 상용 Windows/Linux, CORBA 기반
- COHSE: Conceptual Open Hypermedia Service Environment)
- HealthCyberMap: 의학정보 map
- iMorph InfoMinder : Website 변경내용 추적
- KAON: RDF 기반의 온톨로지 관련 도구
- OntoEdit: 온톨로지 모델의 설계, 적용을 위한 개발환경
- Protégé-2000: 온톨로지 에디터 내장(Text/RDF/JDBC), Cerebra Client
- SiLRI: Java로 작성된 추론엔진
- SWAG : Semantic Web Agreement Group (SW 용어 Dictionary DB)
- OILed
- Ontoprise
- Flogic
- SymOntos

에이전트 (<http://www.semanticweb.org/> 자료)



응용: 전자상거래

- 웹 상에 존재하는 방대한 정보를 활용한 전자상거래 프레임워크의 필요성
- 시맨틱 웹은 W3C의 메타 데이터(Metadata) 사업의 후속으로 진행되고 있으며 W3C의 XML, RDF, 온톨로지(Ontology) 사업을 기반으로 웹 상에서 이루어지는 전자카탈로그, 전자상거래, 비즈니스 모델 등의 모든 활동에 공통적으로 적용되는 표준안이 필요
- 시맨틱 웹의 활성화는 전자상거래를 위해 부문별로 이루어진 표준화 노력을 통합하여 운용하기 위한 실용적 표준이 필요
- 시맨틱 웹의 온톨로지와 RDF로 표현된 데이터 간의 관계 정보는 이러한 중앙등록부와 같은 역할을 수행하여 웹 상의 모든 정적인 성격의 정보와 동적인 성격의 서비스를 하나의 데이터베이스처럼 운용할 수 있는 이상을 실현하는 기교가 될 것으로 기대

관련 표준

- The Foundations for Intelligent Physical Agents (FIPA) : <http://www.fipa.org/>
- Knowledge Interchange Format (KIF) : <http://www-ksl.stanford.edu/knowledge-sharing/kif/index.html>
- PICS: Platform for Internet Content Selection: <http://www.w3.org/PICS/>
- PICS Rating Vocabularies in XML/RDF, Brickley and Swick, 2000: <http://www.w3.org/TR/2000/NOTE-rdf-pics-20000327>
- Dublin Core Metadata Initiative : <http://purl.org/dc/>
- The Cambridge Communique, W3C NOTE, 7 October 1999: <http://www.w3.org/TR/1999/NOTE-schema-arch-19991007>
- Unified Modeling Language (UML) : <http://www.rational.com/uml/resources>
- DAML+OIL: <http://www.daml.org/2001/03/daml+oil-index.html>

응용분야

- Multi-agent Collaboration and Interoperation
- Education
- Modeling (Knowledge Engineering)
- E-Commerce
- Search Engines (Semantic Search Engine)
- Bioinformatics

결론

■ 비전

- ▶ 내용 중심 표현과 메타데이터 표현
- ▶ 개념 중심 작업 지원
- ▶ 새로운 개념(단어) 정의 및 단어 간의 관계 표현
- ▶ 에이전트 기반 정보 추출 및 가공
- ▶ 연산, 논리적 추론, 증명 지원
- ▶ 대규모 분산 소형 온톨로지

■ 인공지능, 에이전트, 시맨틱 웹