

시맨틱웹(RDF) 기반의 e러닝 메타데이터 상호운용성에 대한 고찰

배일환* · 이범진** · 문남미***

1. 서론

e러닝(e-Learning) 영역은 다양한 종류의 학습 기술 제 영역에 대한 표준화를 위하여 활발한 연구가 이뤄지고 있다. 현재까지의 연구의 대부분은 그러한 표준화 규격(예를 들어, IMS 제 규격들, IEEE-LOM, 그리고 ADL의 SCORM 등)을 위한 인코딩 언어로써 XML을 그 근간으로 삼고 있다.

그런데, 최근 몇 년간 W3C(World Wide Web Consortium)[5]는 RDF(Resource Description Framework)라는 메타데이터를 위한 메타 언어(meta-language)를 기반으로 하여, Tim Berners-Lee가 주창한 시맨틱웹(Semantic Web) 버전의 기초를 만들어 내고 있으며[14], 그 결과 시맨틱웹의 빌딩블록의 하나인 온톨로지(ontology)를 위한 언어인 OWL(Web Ontology Language)의 작업초안이 2003년 3월에 발표되기에 이르렀다[13]. 이러한 웹의 발전에 발맞춰 유럽 및 미국의 e러닝 각 진영에서는 시맨틱웹 기술과 e러닝 표준화 기술간의 융합을 조심스레 준비하고 있다. 더블린코어(Dublin Core; DC) 진영은 이미 DC를 위한 RDF 스키마를 발표하였으며, e러닝 메타데이터의 사실상 표준(de facto standard)인 LOM

(Learning Object Metadata) 또한 IEEE, IMS, ARIADNE을 중심으로 하여 이미 LOM을 위한 RDF 스키마를 발표하였다. 더불어 IMS는 Content Package 규격을 RDF로 기술하는 프로젝트를 현재 진행 중이다[2]. 그리고, e러닝 표준이나 규격들 중에서 전세계적으로 시장에서 가장 널리 인정 받고 있는 미 국방부 주도의 ADL SCORM 또한 그 향후 로드맵에서 시맨틱 웹 기술의 접목 또는 전향적 적용 의사를 밝히고 있다[9,10].

이러한 사실들은 미래의 학습 기술과 관련한 중요한 의문점을 불러 일으킨다. 시맨틱웹 기술이, 특히 RDF가 어떤 방식으로 학습 기술 규격들에 적용될 것인가, 즉 RDF가 어떤 측면에서 미래의 메타데이터를 표현할 것인가? 그리고 시맨틱웹 기술이 e러닝을 어떻게 변화시킬 것인가?

이러한 의문들에 대한 해답을 찾기 위해서, 본 논문에서는 시맨틱웹 기술 기반의, 주요하게 RDF 기반의 e러닝 메타데이터 상호운용성 구현에 대하여 논의한다. 본 논문의 구성은 다음과 같다. 2장에서는 시맨틱웹과 RDF 기술에 대하여 기술하고, 3장에서는 e러닝 메타데이터에 대한 기존의 개념들의 오류들에 대한 지점을 확인하고 새롭게 시맨틱메타데이터를 제안하며, 4장에서는 RDF 기술을 적용한 자원기술 방법을, 5장에서는 e러닝에 대한 RDF 기술 적용의 의미성을 논한다. 그리고 6장 결론에서는 향후의 연구과제를 제시한다.

* (사)한국이러닝학회 사무국장
** (사)한국이러닝학회 실장
*** (사)한국이러닝학회 회장

2. 시맨틱웹(Semantic Web)

2.1 시맨틱웹 개요

시맨틱웹은 W3C에 의해서 최근에 시작되었으며, 여러 응용(application)들에 걸쳐서 데이터의 자동적 통합 및 재사용을 위해서 “기계에 의해 사용될 수 있는 방식으로 정의되고 연결된 ‘웹의 데이터’를 만들고자 하는 개념”을 현실화시키고자 추진되고 있다[1]. 이는 기존의 웹 서비스의 한계를 극복하기 위해 1998년 Tim Berners-Lee가 주창한 차세대 웹 기술로, 이를 적용하면 쉽게 인공지능 시스템 개발이 가능하다. 시맨틱(semantic)이란 웹 데이터간의 관계를 의미하며, 이를 사람이 인식하고 이해하는 것이 아닌 기계가 이해할 수 있는 관계를 말한다. 따라서 기존 웹 서비스가 사람이 보고 판단해 이용하는 것이었다면, 시맨틱 웹 기술을 이용하면 컴퓨터가 내용을 스스로 판단해 처리해 주는 인공지능 서비스가 가능해진다[20].

이와 같이 지능형 서비스 즉, 데이터의 혁신적 인고사용성을 위해서 컴퓨터가 사용할 수 있는 정보를 만든다는 것은 메타데이터에 대한 접근법에서 출발한 것이다. 사실상 시맨틱웹의 빌딩블록의 큰 축은 메타데이터(더 발전하여 온톨로지)와 지식표현(Knowledge Representation)이다[18].

시맨틱 웹은 계층화된 구조(layered structure)이다. 구문(syntax)을 전달할 수 있도록 XML이 그 기초를 이루고 있다. RDF는 정보 표현 프레임워크를 제공하고 있다, 그리고 이 계층의 맨 위에서 온톨로지(ontology)가 정보에 대한 지능적 처리를 가능하게 하는 어휘(vocabulary)의 표현을 위해서 반드시 필요한 논리적 구조를 제공한다[20].

미국과 유럽이 시맨틱웹 기술 개발에 박차를

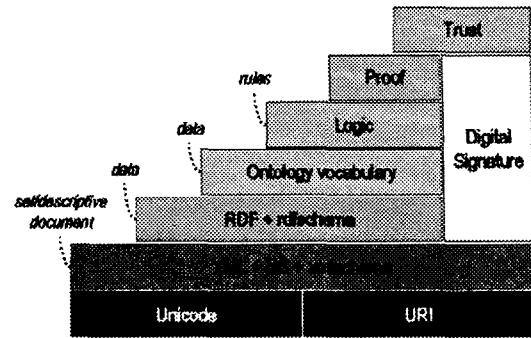


그림. 시맨틱웹 아키텍처 레이어, (20)재인용

가하는 이유는 e비즈니스나 전자상거래 등에 이를 적용하면 인공지능 시스템을 개발할 수 있기 때문이다. 또한 지금까지 큰 관심을 보이지 않던 아시아·태평양 지역에서도 최근 중국과 일본이 준비에 나서고 있는 상황이다[19].

2.2 RDF

RDF는 시맨틱웹 환경의 메타데이터 표현 언어이다. RDF는 문서에 의미를 더하는 방법을 제공한다. 이는 구조화된 메타데이터를 기호화하고(encoding), 교환하고, 재사용하는 것을 가능하게 하는 기반구조이다. 정보는 기계가 이해할 수 있는 RDF 문장의 형태로 저장되어 있다. 검색 엔진, 지능형 에이전트, 정보 브로커, 브라우저 그리고 사람은 RDF로 기술되어 있는 이 의미적 정보(semantic information)를 이해하고 사용할 수 있다. RDF는 구현 독립적이고 XML의 연장이다(구문이 XML로 정의되어 있다).

RDFS(RDF 스키마)와 짝을 이루는 RDF는 필요에 따라 확장할 수 있는 모델링 근원어(modeling primitives)를 제공한다. 기본 클래스의 계층적 구조와 클래스와 객체간의 관계는 RDFS로 표현 가능하다. RDF는 작은 온톨로지 언어로 이해할 수 있을 것이다. RDF를 기반으로 한 웹 온톨로지 언어의 현재 상태는 W3C에서 작업초안 단

제로 발표한 OWL이다.

RDF는 소위 구문(statement)이라고 하는 것으로서 완전히 구성되어 있다. 그리고 RDF 구문은 3개의 요소로써 구성되어 있다:

- 주어(subject)
- 서술어(predicate)
- 목적어(object)

구문은 웹 자원에 대한 것이다, 그래서 주어 및 목적어는 URI들이고, 기계가 읽을 수 있는 식별자들이다. 또한 사용될 수 있는 여러 서술어들을 명확히 하기 위해서, 모든 서술어에 하나의 URI가 부여되어야 한다. 이 경우에, 사용 가능한 표준화된 서술어가 메타데이터의 어휘에 있다, 예를 들어 우리가 사용할 수 있는 것은 DC 메타데이터 중 저작자(creator)인 “<http://purl.org/dc/elements/creator>”이다. 그래서 다음과 같은 트리플 표현이 가능하다:

- 평문 : <http://www.w3c.org/2001/sw/> 문서는 W3C가 만들었다.
- 트리플 : (“<http://www.w3c.org/2001/sw/>”, “<http://purl.org/dc/elements/creator>”, “W3C”)

그리고, RDFS 규격은 용어들간의 관계를 표현하는 기본 어휘를 제공하고 있다: 용어의 인스턴스(instance)를 제시하는 용어(예, <http://www.w3c.org/> is an organization), 다른 용어의 하위 개념인 용어(“a hex-head bolt is a type of machine bolt”) 등등.

이것은 또한 서술어 사용을 제한하는 방법도 제공한다: “is a parent of”는 오직 사람(person)에만 적용된다 등. 이렇게 RDFS에 의해 제공된 어휘를 사용하는 것은, 의미적으로 풍부한 자신만의 어휘를 창조하는 것을 쉽게 해주며 기본적인 온톨

로지 구성을 가능하게 해준다[11].

3. 시맨틱 메타데이터

본 장에서는 시맨틱웹 기술과 e러닝 기술의 접목을 위한 기술 진화에 있어서, 기존의 “메타데이터(Metadata)” 논쟁을 중심으로 당면 e러닝 메타데이터의 문제점을 지적하고 그 대안으로 시맨틱 메타데이터(Semantic Metadata)[8]를 제안한다. 사실상, 시맨틱웹 기술의 기본적인 빌딩블록은 “메타데이터 기술”과 “지식표현(Knowledge Representation) 기술”이다.

2장에서 언급했듯이, 시맨틱웹 기술의 기초가 되고 있는 RDF 기술을 이용한 메타데이터의 표현은 이미 더블린 코어나 IMS를 통해서 일정 정도 완성이 되었다. 물론 이것이 곧바로 “시맨틱 메타데이터(Semantic Metadata)”의 완성을 얘기하는 것은 아니다. 지금도 이에 대한 정교화가 진행되고 있으며, 이는 향후, 현실적으로 봤을 때, OWL의 적용과 그리고 e러닝 진영의 영역 온톨로지(e-learning domain ontology)의 개발과 함께 완성도를 확보할 수 있다.

그렇다면 본 장에서 다루고자 하는 “시맨틱웹 기술을 적용한 시맨틱 메타데이터(또는 차세대 메타데이터)”란 무엇이고, 어떤 기능성을 가지는지 그리고 어떤 의미가 있는지, 시맨틱웹 기술을 적용한 e러닝은 어떤 모습인지에 대해서 ‘간단히’ 기존의 메타데이터 정의에 대한 해석을 통해서 파악해보겠다.

3.1 e러닝 메타데이터의 쟁점

3.1.1 메타데이터의 주관성 vs. 객관성

메타데이터의 가장 널리 알려진 광의의 정의는 “데이터에 대한 데이터(data about data)”[6]라

는 정의이다. 그러나, 이 정의는 많은 오해를 불러 일으키도록, 대부분의 사람들이 서로 다른 의미와 모순될 수도 있는 의미로 이 말을 사용하고 있는 것처럼 보인다. e러닝 응용들이 메타데이터를 구현하는 과정을 보면 이 오해들은 더욱더 구체적이고 현실적인 오해를 불러일으키고 있음이 확인 가능하다. 메타데이터의 “객관성”이라는 오해와 관련해서 정리해보면 다음과 같다.[2]

- 메타데이터는 데이터에 대한 객관적인 (objective) 데이터이다.
- 자원에 대해서 메타데이터는 오직 한 번만 들어진다.
- 메타데이터는 서지 목록의 디지털 버전이다.

첫 번째의 오해와 두 번째 오해는 대부분의 메타데이터 전문 시스템들이 기능적으로 제공하고 있는 오해 중 가장 대표적인 것이다. 먼저, 객관적 메타데이터에 대한 오해를 보면, 초기 메타데이터 출현의 역사적 배경과 관련이 있다. 이것은 제목, 저작자, 식별자 등과 같은 명확한 객관적 정보만을 가진 메타데이터의 상정과 관련되어 있다. 더블린 코어의 대부분의 요소들이 이런 종류임을 확인할 수 있다. 그러나 이러한 오해는 사실상 얼마 지나지 않아서 깨어진다. 더블린 코어에서의 요소 선정 및 기술의 기준을 보면, 객관적 요소들을 위주로 하는 고유성(intrinsicality)을 그 요소 선정 기준의 특성으로 하면서도, 규정된 필수 데이터 요소 이외에 부차적인 내용이나 특성을 기술 요소로 사용할 수 있는 확장성(extensibility)과 응용분야나 표현기법을 규정하지 않는 구문독립성(syntax independence)의 특성을 인정하듯이, e러닝 영역의 메타데이터 정의에서도 객체의 단위의 유형, 교수설계 목적, 평가 그리고 학습 목표 등이 자원을 묘사하기 위해 메타데이터의 요소들로 사용된다. IEEE의 LOM의 요소들을 확인해보

면 이러한 요소들이 강화되었음을 확인할 수 있다.

이렇게 메타데이터에 대한 응용 분야가 확장됨으로써 사실상, 세 번째 오해였던 서지 목록의 디지털 버전으로 한정되었던 메타데이터의 개념도 변경이 된다.

즉, 메타데이터는 자원의 객관적인 정보만으로 한정되는 것이 아니라, 자원(또는 콘텐츠)의 라이프사이클 전체를 촉진하기 위해 메타데이터를 생성하고 서비스하는 것이라는 개념으로 자연스럽게 확장된다. “메타데이터는 콘텐츠의 가치사슬의 각 단계에서 콘텐츠의 가치를 증가”[8]시키게 된 것이다. 메타데이터를 목록 관리 정도만을 수행하는 것으로 한정시키는 것은 이미 그 잠재력을 심각하게 제한하는 상태가 된 것이다. 물론 이는 웹의 급속한 발전과 함께 그리고 메타데이터를 이용하는 각 영역의 성장과 함께 초래된 결과이다. 도서관 정보 네트워크의 대표적인 기관인 UKOLON(UK Office for Library and Information Networking) 또한 메타데이터를 “자원 설명, 탐색, 저작권 관리를 포함한 관리 및 자원의 장기 보존 등의 폭넓은 자원 운영 전반을 지원”[16]한다고 정의하고 있다.

이렇게 객관적 메타데이터라는 오해는 일정정도 해결된 듯 보이지만, 여전히 그 오해로 인해 메타데이터의 진화가 발목 잡혀 있다. 메타데이터 요소에 자원의 응용 맥락에 따라 변경 될 수 있는 다소 주관적인(subjective) 요소들이 포함되었지만, 여전히 메타데이터 전문시스템은 메타데이터를 객관적으로 바라보고 있다. 즉, 여전히 e러닝 또는 메타데이터 관련 시스템은 메타데이터를 자원에 대한 객관화된 정보라고 인정하고 있다. 암묵적으로 메타데이터가 그 자원을 고스란히 설명하고 있다고 인정하고 있다는 것이다. 메타데이터는 자원을 퍼블리싱 할 때 한 번 만들어져서 메타

데이터에 대한 추가적인 수정이나 갱신을 통해서 버전관리만 하면 된다는 식으로 관리되고 있다. 하나의 자원을 설명하는 객관적인 하나의 메타데이터가 존재하고 이 메타데이터를 통해서 자원을 활용한다는 것이다. 사실상 자원에 대한 해석은 하나만이 존재하게 된다.

그러나 현실세계는 이러한 사상과 정면으로 배치된다. 하나의 자원에 대한 해석은 한 가지만 존재하지 않으며, 그 하나의 메타데이터가 그 자원을 온전히 말하는 객관적이고도 권위 있는 정보라고도 할 수 없다. 메타데이터 기술(description)도 사람이 행하는 것이며 그 정보는 주관적이다. 근본적으로 주관적(subjective) 해석이 이뤄지는 것이다. 현실적으로 하나의 자원에 대해서 메타데이터는 여러 개가 존재할 수 있는 것이며 또한 여러 개가 존재해야 한다. 하나의 자원에 대해서 역할(role)별로 다양하게 정보를 정의(메타데이터 생성)할 수 있어야 하며, 또는 자원을 사용할 '의도된 시스템' 별로 서로 다르게 자원에 대한 정보를 정의할 수 있는 것이며, 또는 바라보는 관점(주제 또는 맥락)에 따라 자원에 대한 해석을 달리 할 수 있는 것이다. 심지어는 상반되는 해석이 있을 수도 있다.

하나의 예를 들어보자.[17] 어떤 전쟁 포화 속의 고아를 담은 흑백 사진 한 장이 있다. 이 사진이 '사진학'이라는 맥락에서 교육에 활용될 때와, '전쟁사'라는 맥락에서 교육에 이용될 때는 서로 다른 의미들을 가질 수 있다. 주제나 학습목표도 틀릴 것이며 관련자원에 대한 정보도 틀리다. 전자 는 관련 자원으로 '사진을 찍을 때의 빛의 노출 방법'이나 '필름 현상 방법'을 내용으로 가지고 있는 자원들을 제시하게 된다. 후자는 해당 전쟁과 관련된 여러 가지 사례라든지, 당시의 시대적 배경 등과 관련된 내용을 담고 있는 자원을 관련자원으로 제공하게 된다.

메타데이터의 주관성을 인정하지 않는다면 우리는 하나의 자원이 가지고 있는 다양한 맥락을 잊은 채, 아마도 전문가가 제공하는 하나의 관점, 즉 시스템이 객관적이라고 말하는 하나의 관점에서만 이 자원을 활용하게 될 것이다. 사실 이것은 객관적인 게 아니라 전문가 개인의 주관성이며 아울러 '권위적인 객관', '선언적인 객관'이라 할 수 있다[2].

진정한 객관성은 여러 메타데이터 정보 중에서 전문가의 권장 메타데이터 및 사용자의 사용 빈도 등을 종합한 많은 사용 사례에 관한 데이터의 축적 속에서 그 자원에 대한 가장 근접한, 가장 사용성이 풍부한 메타데이터가 그 지위를 가지게 될 것이다. 그리고 이렇게 '정당화된 객관성'은 정적(static)으로 누군가에 의해 일방적으로 선언 되는 것이 아니라 동적(dynamic)으로 여러 사용자들의 보이지 않는 커뮤니케이션에 의해서 만들어진 것이다. 즉 동의를 생성(consensus building)과정을 통해서 정당화된 객관에 도달하게 된다. 동의를 통해서 권위 있는 정보가 탄생하는 것이다.

3.1.2 메타데이터의 분산성

메타데이터의 객관성과 주관성은 사실 웹(또는 웹자원을) 바라보는 시각과 관련이 있다. 인터넷은 원래 누군가가 누구에게 연결 될 수 있는 (where anyone can connect to anyone) Peer-To-Peer 네트워크로 설계된 것이다. 그리고 이것은 여전히 인터넷이 성공하고 있는 주요한 이유 중의 하나이기도 하다.

그러나 웹은, 팀 버너스리가 최초로 월드와이드웹을 창시했을 때의 애초의 목적을 무너뜨리는 주로 중앙집중적인 정보 관리에 의존하는 Client-Server 기반의 시스템이 지배적으로 개발되었다 [1]. 이것은 앞 절에서 언급한 객관적인 권위와 맥을 같이 하는 것으로써, 이러한 권위를 해체하고

동의(consensus)로써 정의되는 '정당화된 객관성'을 만들기 위한 프로세스의 한 부분이 메타데이터의 주관성을 인정하는 부분과 관련되어 있다.

그리고 이러한 주관적 메타데이터는 웹에 광범위하게 분산되어 있는 메타데이터 전문 시스템에 의해 생성되고, 관리되고, 교류되고, 점차 발전하고, 끊임없이 업데이트와 수정을 받는 정보가 되는 것이다. 기술(description)들간의 경쟁은 장려되고, 그리고 다양한 종류와 계층의 맥락에 특정한 메타데이터가 필요할 때 다른 사람에 의해 언제나 추가될 수 있는 정보의 국제적 네트워크를 가능하게 한다.

보충하자면, 차세대 웹 환경이라는 시맨틱 웹의 기술은 다른 여러 가지 의미들이 있겠지만, 그 중에서도 웹의 분산성을 강력하게 지원한다는 특징이 있다. 이것은 웹 자원에 대한 메타데이터의 주관성과 분산성을 전세계적으로 광범위하게 지원하게 될 것이며, "동의를 만드는 것(consensus building)은 메타데이터 관리의 당연한 부분이 되고 전 세계적으로 메타데이터가 번창하고 서로 번식할 수 있는, 그래서 메타데이터가 점차 발전하고 그리고 새롭고 고정되지 않은 맥락에서 재사용되는 곳, 누구나 참여할 수 있는 곳, 즉 메타데이터 생태계를 가능하게 할 것이다"[2]

3.1.3 메타데이터의 상호운용성

또 다른 메타데이터에 대한 오해는, "메타데이터는 구조화된 구문(syntax) 체계를 가진다"는 것과 관련한 오해이다. 이는 현재 e러닝 진영에서 광범위하게 유행하고 있는 XML 기반의 메타데이터 엔코딩(encoding) 방식과 관련이 깊다. 정확히 말해서 이 오해는 기술적 유행 또는 XML의 한계라고 말할 수 있다. 또한 이는 메타데이터에 대한 여러 규격들의 상호운용성을 해결하는 대안이 숨어 있다. 이 부분에 대한 논의는 시맨틱 웹

기술의 현시기의 현실적인 실체를 가지고 있는 RDF와 XML의 비교를 통해서 파악하고자 한다.

XML 스키마를 사용해서 메타데이터 응용 프로파일을 정의하는 것은, XML 문서의 구문 구조(syntactic structure)를 정의하고 있는 XML 스키마를 통해서, 즉 각각의 메타데이터 응용 프로파일이 "사용하도록 허용된 어떤 스키마"에 대해서 적합성을 가지도록(valid) 정의된다는 의미이다. 이럴 경우는 그 유연성이 심각히 훼손되며 이종 스키마의 메타데이터 간의 상호운용성은 확보될 수가 없다.

각 메타데이터 규격의 어떤 요소명들을 동일한 이름으로 정한다고 해서 상호운용성이 확보되는 것은 아니다. 더더군다나 무수히 많이 존재하는 메타데이터 규격을 전 세계적으로 오직 한 가지만으로 통일해서 사용한다는 것도 있을 수 없는 일이다. 실사, 강력한 조직에 의해서 강력한 규격에 의해서 하나로 통일 된다 하더라도 그 단일한(single) 메타데이터가 여러 영역의 자원을 제대로 다 설명할 수는 없을 것이다. 상호운용성에 대해서 XML 기반의 메타데이터는 대안을 가지고 있지 못하다[11].

RDF는 문서에 의미를 더하는 방법을 제공한다. 구조화된 메타데이터를 엔코딩하고, 교환하고, 재사용하는 것을 가능하게 하는 기반구조이다. RDF는 "구현 독립적"이고 XML의 연장이다. RDF는 웹 상에서 자원을 묘사하기 위한 기술로써, 별개의 메타데이터간의 어의, 구문 및 구조에 대한 공통적인 규칙을 지원하는 기법을 통해 웹 상에 존재하는 정보를 교환하기 위해 W3C에서 제안한 것으로, 메타데이터간의 효율적인 교환 및 상호운용을 목적으로 하고 있다.

상호운용성의 문제에 대한 대안으로서, IMS 메타데이터를 RDF로 기술하는 프로젝트를 주도했던 M.Nilsson의 교훈을 들어보자[11]. 메타데

이터의 RDF 바인딩을 생산한 결과,

- 다른 별개의 표준들과의 상호운용성이 크게 증가되었다. 그 이유는 간단하다. RDF는 매우 다양한 유형의 데이터와 스키마를 위한 단일한 스토리지 모델(single storage model)을 허용한다. 예를 들면, 동일한 데이터베이스에 여러 다른 규격을 따르는 메타데이터를 저장하는 것이 간단하다.
- 기존의 메타데이터 표준들의 재사용은 크게 단순화되었다. 예를 들면, vCard XML 구문을 XML 바인딩에 결합할 지의 여부에 대한 많은 논의가 있어왔었지만, 이것은 이름공간(namespace) 및 XML DTD 문제를 일으킨다. 하지만 vCard RDF 바인딩은 추가적인 노력 없이 가능하다.
- 어떤 어휘(vocabulary)들은 메타데이터 표준들 간에 의미가 정확히 일치하지 않을 수도 있다. 또한 XML 기술로써는 어휘들을 엔코딩하고 분산하는 방법을 제공하지 못한다. RDF에서는 이 문제는 완전히 사라진다. 어휘들의 상호의존성이 표준화된 방식으로 모델링 될 수 있다. 그리고 이것으로 부족하면, 보다 완전해진 온톨로지들로써 더욱 완전해질 수 있다.

RDF와 XML 스키마를 이용해서 자원을 서술하는 것은 근본적으로 틀리다. XML 스키마가 말하고 있는 노드(node)는 어떤 하나의 XML 문서의 구조상의 특정 영역에 있는 노드이다. 그러나 RDF에서의 노드는 문서 자체에 있는 노드가 아니라 URI를 가지고 있는 어떤 자원이다. XML과 XML 스키마로 정형화된 트리 기반 문서로써 표현된 메타데이터는 유연성이 거의 드물며, 메타데이터의 주관적이고 분산적인 또한 여러 종류의 형태로 표현되어 있는 현실에서 그 상호운용성을

담보할 수가 없다.

본 절에서는, 메타데이터의 분산성과 구문독립성, 주관성을 인정하면서도 상호운용성을 확보해야 하는 현실적 필요성과 그 대안으로써 RDF 기술을 이용한 상호운용성 확보의 가능성에 대해서 논하고 있다. 현재 선진 e러닝 진영에서는 그 실험적 시도으로써 RDF를 적용해 보고 있다. 예로, vCard, Dublin Core, Dublin Core Qualifiers, IEEE LOM, SCORM 등에 사용된 여러 다른 포맷들 간의 사용, 검색, 교환에 있어서 봉착한 심각한 문제를 해결하기 위한, "P2P 기반의 교육 메타데이터 교환 네트워크 인 Edutella" 프로젝트를 진행중이다 [12].

3.2 시맨틱 메타데이터 그리고 온톨로지

지금까지 제한적이지만, 메타데이터의 주관성, 분산성, 상호운용성을 주제를 간단하게 다뤘다. 이러한 내용을 바탕으로, 시맨틱 웹 기반의 메타데이터의 특성을 주요하게 개략하면[1],

- 주관적이고 비권위적이다 동일한 자원을 서로 다른 여러 관점으로 설명하는 것을 지원한다.
- 분산되어 있다 누구에 의해서든, 무엇에 관해서든, 어디서든 기술(description)을 지원한다.
- 재사용성과 고이용성을 확대한다 주관성과 분산성을 바탕으로 자원에 대한 효과적인 이용을 촉진한다.
- 상호운용성을 지원한다 다양한 규격의 메타데이터 스키마 간의 상호운용성을 지원할 수 있다.

그렇다면, 엄밀히 말해서, 시맨틱 메타데이터란 무엇인가?

먼저 시맨틱 웹에서 말하고 있는 온톨로지에 대한 개념을 짚어보자.

온톨로지(ontology)란 “어휘들의 의미를 그리고 그런 어휘들간의 상호관계들에 대한 표현”[13]을 말한다. 즉, 온톨로지는 “어떤 영역을 기술하기 위해 사용하는 개념과 그리고 유사한 어휘들을 모아 놓은 어휘집에 관한 규격”[7]이다. 온톨로지는 웹에서 발견되는 이종의, 분산되어 있는, 반(半)구조화된 정보를 기술하기에 안성맞춤이다. 공통의 영역에 대한 공유된 의견을 정의함으로써, 구문이 아니라 의미 교환을 지원하는 방식을 통해서 사람과 기계가 간단히 통신할 수 있도록 한다.

원래 온톨로지(Ontology)는 철학적으로 존재에 관한 이론이다. 학문분야로서의 온톨로지는 존재하는 것들이 어떤 형태인지와 같은 그러한 이론들을 연구하는 것이다. 인공지능과 웹 연구가들은 이 전문어를 흡수하여, 특정 흥미 영역의 정형화되고 공유될 수 있는 개념들을 온톨로지로 기술한 것이다[20].

e러닝에서의 온톨로지의 활용을 보면, 1) 학습자원의 의미를 기술하기 위해, 2) 학습자원의 학습 맥락을 정의하기 위해, 3) 학습코스 내의 학습자원의 구조화(structuring)을 위해 사용될 수 있다.

e러닝에서 공통적으로 공유된 의미의 문제는 학습자원을 제공하기 위한 과정뿐만 아니라, 특정 학습자원에 대한 접근(또는 검색)을 처리하는 과정에서 매우 중요하다. 두 명의 저작자가 같은 주제를 다른 방식으로 표현할 수 있는 위험성이 항상 존재하는 것이다. 키워드를 예로 들면, Agent와 동일 시 될 수 있는 어휘들이 있을 것이다 -“agent, actor, contributor, creator, player, doer, worker, performer” 등. 온톨로지를 이용해서 이러한 개념들의 맵핑(mapping)을 구별하고 또 연관 지을 수 있을 것이며, 또한 ‘Network’라는

주제로부터 ‘Protocol’이라는 주제를 유추해서 관련 자료를 쉽게 제시할 수도 있게 된다. 또한 학습자원에 대한 다양한 맥락하에서의 사용을 촉진할 수 있다. 사용자의 선호도에 근거하여 학습자원에 대해 맥락과 적절히 관련된 것을 검색 가능하게 해줄 것이며, 아울러, 학습자원들 간의 관계를 구조화하는 방법을 통해서 학습코스의 구조를 기술하는 것을 지원해준다.

이러한 온톨로지를 기반으로 하는 시맨틱 메타데이터를 통해서 e러닝은, 학습자원에 대한 의미적 주석이 부여되고, 또 새로운 학습 요구에 부응하기 위해 쉽게 새로운 학습 코스로 조합될 수 있고, 학습자의 취향에 따라서 유용한 학습자원이 쉽게 발견될 수 있고, 이러한 프로세스 뒤에는 온톨로지가 기능함으로써 학습자원에 대한 의미적 질의와 탐색이 가능하도록 지원한다.

부연하자면, 공통적으로 공유된 의미(common-shared-meaning)와 기계가 처리할 수 있는 메타데이터(machine-processable-metadata)를 특성으로 하는 시맨틱웹 기술을 적용한 e러닝이란 빠르고, 즉각적이고, 최적화된 학습을 만족시킬 수 있는 강력한 기능성을 지원하게 된다.

지금까지의 내용을 통해서 시맨틱 메타데이터를 다음과 같이 정의해 볼 수 있다. 시맨틱 메타데이터란, “기계가 이해할 수 있는 콘텐츠를 지원하는 온톨로지 기반의 메타데이터”[17]. 그리고 향후의 e러닝과 시맨틱 웹의 결합에 있어서 그 빌딩블록으로 기여하는 핵심이 바로 메타데이터임을 다시 한 번 강조한다.

3.3 시맨틱 메타데이터 그리고 e러닝

시맨틱 메타데이터의 특성을 언급해보자.

- 주관적이고 비권위적이다 - 동일한 자원을 서로 다른 여러 관점으로 설명하는 것을 지

원한다.

- 분산되어 있다 - 누구에 의해서든, 무엇에 관해서든, 어디서든 기술(description)을 지원한다.
- 재사용성과 고이용성을 확대한다 - 주관성과 분산성을 바탕으로 자원에 대한 효과적인 이용을 촉진한다.
- 상호운용성을 지원한다 - 다양한 규격의 메타데이터 스키마 간의 상호운용성을 지원할 수 있다.
- 확장성이 있다 - 새로운 의미체계를 가진 새로운 어휘를 창조하는 것을 허용한다.
- 유연하다 - 자원들의 예측되지 않은 쓰임을 지원할 수 있다.
- 발전적이다 - 다양한 메타데이터 생태계를 지원한다.
- 개념적이다 - 인간 지식의 진화를 지원한다.

결과적으로, 시맨틱 메타데이터는 단순한 자원에 대한 목록으로서의 기능으로 한정되는 것이 아니라, 인간 지식의 진화를 촉진하는 데에 크다면 기여를 한다.

아울러, 시맨틱 메타데이터를 이용한 e러닝은,

- 학습자원에 대한 의미적 주석이 부여되고,
- 새로운 학습 요구에 부응하기 위해 쉽게 새로운 학습 코스로 조합되어 질 수 있고,
- 학습자의 취향에 따라서 유용한 학습자원이 쉽게 발견될 수 있고,
- 이러한 프로세스 뒤에는 온톨로지가 기능함으로써 학습자원에 대한 의미적 질의와 탐색이 가능하도록 지원한다.

4. RDF를 이용한 e러닝 자원 기술

RDF를 이용한 e러닝 자원 기술(description)이 다른 메타데이터 접근법과 비교되는 가장 기본적인

인 이점(利點)은, RDF를 사용하면 누구나 어떤 식별 가능한 자원에 대한 RDF문을 만들 수 있다는 것이다(메타데이터의 주관성 및 분산성). RDF를 사용하면, 메타데이터를 확장하는 문제점 그리고 이종의 스키마의 여러 다른 포맷(format)의 메타데이터를 조합하는 문제점이 사라진다(메타데이터의 상호운용성 및 확장성), 왜냐하면 RDF는 XML이 가진 구문적 한계를 뛰어넘어 폐쇄형 문서(closed document)를 사용하지 않기 때문이다.

아래에서 RDF로 기술한 e러닝 자원에 대한 메타데이터 사례를 살펴보자[4].

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1"
xmlns:lom_edu="http://www.imsproject.org/rdf/imsmd_educationalv1p2#"
xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#"
... (중략) ...
<dc:creator>
  <vCard:fn>Hong GilDong</vCard:fn>
</dc:creator>
... (중략) ...
<lom_edu:NarrativeText>
  <dc:title>
    <lom:LangString rdf:ID="title">
      <rdf:value rdf:parseType="Resource">
        <rdf:value>e-Learning Metadata</rdf:value>
        <dc:language rdf:resource="#en"/>
        ... (생략) ...
```

이 메타데이터 인스턴스 사례에서 본 바와 같이, RDF로 자원을 기술했을 때, 기존의 XML 기반의 메타데이터 스키마가 가지지 못했던 상호운용성(사례에서는 LOM, DC, vCard의 상호운용을 보여주고 있다)을 확보함을 확인할 수 있다.

RDF 모델과 XML 모델은 근본적으로 서로 다르다. XML 데이터 모델은 텍스트 기반의 마크업

을 지향하는 라벨 트리(text-markup oriented labeled tree)이다. 대조적으로, RDF는 라벨 아크(labeled arcs)를 구성하고 있는 매우 단순한 모델을 가지고 있다. XML로 정형화된 메타데이터를 표현하는 경우 유연성이 거의 드물다. 메타데이터의 성질은 주관적이고, 분산되어 있고 그리고 여러 종류의 형태로 표현된다는 것이다. RDF 모델은 이러한 메타데이터의 원리를 지원할 수 있을 만큼 충분히 유연하며, 여러 메타데이터간의 상호운용성을 근본적으로 보장해준다.

5. RDF를 적용한 e러닝의 새로운 패러다임

RDF를 적용한 시맨틱 메타데이터를 통해서 본, 즉 시맨틱 웹 기술을 사용했을 때의 e러닝의 가능성(possibility)에 대해서 알아보자.

XML 표준이 LMS(Learning Management System)에 대한 입출력 포맷을 명세화함으로써 상호운용성을 가능하게 하는 좋은 기능을 제공하지만(SCORM v.1.2의 규격을 참조), 이러한 XML 표준은 거대한 중앙 집중식의 획일화된 시스템(large, centrally managed, monolithic system)에 유리하다. 메타데이터 및 e러닝을 위한 여러 커뮤니티들에 의해 생산된 규격들(Dublin Core, vCard, LOM, SCORM 등) 모두는 내용적으로는 사실상 뿔레야 뿔 수 없는 복잡한 상호의존성을 가지고 있고 또 그렇게 복잡한 상호의존성은 점점 증가하고 있다. 그러나, 그러한 규격들간의 어휘 및 기술의 재사용은 매우 어려운 작업이라는 것은 당연하다. 근본적으로 상호운용성 문제 때문에 항상 그러하다는 것이다. 이것은 XML 바인딩의 폐쇄성의 결과이다.

그러나, 최근에 생겨난 시맨틱 웹은 e러닝 기술 및 규격의 사용에 있어서 매우 흥미로운 가능성을 제시한다. 필요한 규격을 생산하는 e러닝 커뮤니

티들의 노력들은 여전히 계속될 것이고 또 그러해야 하고, 표준 그룹들간의 상호운용성 논의에서는 최대한의 상호운용성을 위하여 강력하게 지원할 중요한 기술적 플랫폼을 제공할 것이다.

- 학습자가 전세계적으로 분산되어 있는 학습 자원들을 의미적으로 찾거나 사용하는 것을 지원할 것이다. 자원은 웹에 분산되어 있으며, 이들은 공통적으로 동화된 온톨로지(들)에 의해 연결되어 있다. 이것은 관심 있는 주제에 대한 의미적 질의에 의해서 사용자 특성에 맞는 코스로 구성될 수 있을 것이다.
- 어떤 학습 자원에 대한 개인적 주석이 현실적으로 가능한 기술이 될 것이다[15]. 분산 관리 방식하의 많은 출처로부터의 메타데이터 조합을 가능하게 하고, 자원에 대한 교차 주석 처리 그리고 상호재사용성을 높일 것이다.
- 개인화 및 맞춤화 (개인화된 에이전트를 활용한) 사용자는 자신의 필요에 맞춰진 학습 자원을 의미적으로 검색할 수 있으며, 온톨로지는 학습자의 필요와 학습자의 특성을 연결할 것이다[3].
- 정보의 모듈화 덕분에 분산된, 협업형의 저작 및 코스 구성이 더욱 단순하게 수행될 수 있을 것이며 이러한 수행을 촉진하게 될 것이다.
- 활발한 상호교류에 의한 학습자원의 재사용이 중요한 시너지 효과를 현실화할 것이다. “학습자원에 대해서 자원의 횡적 이용성 증가(자원에 대한 여러 맥락에서의 재사용성 증가) 및 자원의 종적 이용성 증가(자원에 대한 개인화 및 맞춤화를 위한 재사용성 향상)를 통해서, 재사용성과 고이용성을 함께 달성할 수 있을 것이다[17].”

시맨틱 웹은, 폐쇄형 시스템에 속박되어 있는 자원들을 자유롭게 함으로써, 학습자원을 위한 웹 기반의 최적화된 환경을 만드는 것을 보장해줄 것이다. 한마디로, 시맨틱 웹의 비전은 e러닝을 위한 중요한 비전이 될 것이다.

e러닝 기술의 장기적 관점에서 보면, 위에서 언급한 바와 같이, 최근 생겨난 시맨틱웹은 학습 기술 규격들의 사용에 대한 매우 많은 그리고 새로운 잠재적 가능성을 보여주고 있음이 사실이다.

RDF 기술로써 학습 기술 규격들을 이용하는 것이 가능함으로써(예, IMS의 Content Packaging 규격의 RDF 바인딩 프로젝트), e러닝에 있어서 폭넓은 응용들이 무궁무진하게 가능할 것이다:

6. 결론

여러 e러닝 메타데이터 기관들에 의해 생산된 규격들(Dublin Core, vCard, LOM, SCORM 등) 모두는 상당한 의존성을 가지고 있고, 또 그만큼 복잡한 상호운용의 필요성이 요구되고 있다.

그러나, 그러한 규격들간의 어휘 및 기술의 재사용 및 상호운용은 XML 기반 하에서는 매우 어려운 작업이라는 것은 당연하다, 그러나 시맨틱웹 기술(본 논문에서는 주요하게 RDF를 검토했다)을 적용할 경우에는 간단히 그 상호운용성을 성공적으로 보장할 수 있음을 확인하였다. 이것은 학습 기술의 채택에 있어서 중요한 의미를 제공한다.

시맨틱웹 시대에 지능적이고 적응적인 e러닝 기술을 확보하기 위해서는 RDF 기술의 적용을 통한 e러닝 메타데이터 상호운용 구현과 함께 다음과 같은 향후 과제들이 함께 이뤄져야 할 것이다.

먼저, 메타데이터 전문 시스템(metadata-aware system)의 개발이 필수적이다.

둘째, 웹 온톨로지 언어(예, OWL) 기반의 e러닝 온톨로지에 대한 연구 및 개발이 시급하다. 더

불어, e러닝 온톨로지와 관련한 응용들(예, 온톨로지 에디터, 추론 및 적합성 검사를 수행하는 온톨로지 엔진 또는 서버, 온톨로지 검색 및 이종의 온톨로지 통합 기술 등)도 함께 연구되어야 한다.

셋째, 시맨틱웹 기술 기반의 분산형 e러닝 리퍼지토리 기술 개발이 이뤄져야 한다.

참고 문헌

- [1] Ljiljana Stojanovic, "eLearning based on the Semantic Web," University of Karlsruhe, 2001.
- [2] Mikael Nilsson, et al., "Semantic Web metadata for e-Learning," Knowledge management research group, 2001.
- [3] Juan M. Santos, et al., "On the Application of the Semantic Web Concepts to Adaptive e-Learning", Universidade de Vigo, 2002.
- [4] IMS, <http://www.imsglobal.org/rdf/>
- [5] W3C, <http://www.w3.org>
- [6] Rachel Heery, "Review of metadata format", 1996.
- [7] Gruber, T., "A translation approach to portable ontology specifications", 1993.
- [8] Amit Sheth, "Content Management, Metadata & Semantic Web", 2001.
- [9] ADL, "Creating The Digital Knowledge Environment", 2002.
- [10] ADL, "Next Generation SCORM", 2002.
- [11] M.Nilsson, "The Semantic Web : How RDF will change learning technology standards", 2001.
- [12] Edutella, <http://edutella.jxta.org>
- [13] W3C, "OWL Web Ontology Language Overview", 2003.
- [14] W3C, <http://www.w3.org/2001/sw/>
- [15] Annotea, <http://www.w3.org/2001/Annotea/>
- [16] UKOLN, <http://www.ukoln.ac.uk/metadata/>
- [17] 배일환, "시맨틱웹 기술과 e러닝 메타데이터에 대한 소론", e-Learning+, 2003.
- [18] 전자상거래표준화통합포럼, "시맨틱웹 기술을

- 적용한 전자상거래 표준 운용체계 연구”, 2001.
- [19] 디지털타임스, “차세대 웹기술 상용화 연구 돌입” 기사, 2002. 09. 30.
- [20] 마이크로소프트웨어, “월드와이드웹에서 시맨틱 웹으로”, 2002. 04.



배 일 환

- 1997년 2월 동아대학교 경영대학 무역학과 졸업(경영학사)
- 1997년~2000년 우인교육 대표
- 2001~2003년 지식정보기술(주) 교육솔루션사업본부 대리
- 2003년 11월 현재 서울정보통신대학원대학교(SIT) 정보통신 석사과정
- 2003년 11월 현재 (사)한국이러닝학회 사무국장
- E-mail : baecilhwan@sit.ac.kr



문 남 미

- 1985년 2월 이화여자대학교 컴퓨터학과 졸업
- 1987년 2월 이화여자대학교 컴퓨터학과 대학원(석사) 졸업
- 1990년 5월 Tulane Uni. 박사과정 수료
- 1998년 2월 이화여자대학교 컴퓨터학과 박사과정 졸업
- 1999년 아주대학교 미디어학과 조교수 대우
- 2000년 3월~2003년 2월 이화여자대학교 인터넷멀티미디어연구소 센터장/ 이화여자대학교 정보통신교육원교육원 부원장 / 조교수
- 2003년 3월~현재 서울정보통신대학원대학교 멀티미디어전공 부교수
- 2003년 9월~현재 (사)한국이러닝학회 회장
- 관심분야 : 디지털 데이터방송 비즈니스 모델, T-Commerce, Mpeg 21 e-Learning 비즈니스 응용



이 범 진

- 1999년 2월 한국외국어대학교 사회과학대학 행정학과 졸업(행정학사)
- 2000년~2001년 ㈜애니빌
- 2001년~2003년 ㈜메디오피아 테크놀로지
- 2003년 11월 현재 서울정보통신대학원대학교(SIT) 정보통신 석사과정
- 2003년 11월 현재 (사)한국이러닝학회 실장