

살아있는 e러닝- 시멘틱 웹기반의 e러닝(1)

학습자가 원하는 학습자원을 컴퓨터가 스스로 찾아내서 학습자에게 전달해주고, 더 나아가 새로운 지식까지 추론해서 제공해 줄 수는 없을까?

살아 있는 e러닝이 되기 위해서는 컴퓨터가 학습자가 원하는 학습자원의 의미를 이해하고, 적시적(Just-in-time)으로 학습자에게 맞는 형태(Just-for-me)로 학습자원을 제공해 주어 학습이 단편적이 아닌 연속적, 통합적으로 이뤄질 수 있는 적응적 학습 환경이 필요하다. 의미의 웹이라 불리 우고 있는 시멘틱 웹(Semantic Web)은 의미적으로 연결돼 있는 학습 정보를 컴퓨터가 의미를 이해해서 학습자가 원하는, 학습자 수준에 맞는 정보를 제공해주고, 더 나아가 지식까지도 추론해서 학습자에게 가장 적합한 형태로 전달해 줄 수 있는 강력한 매카니즘으로 부각되고 있다. 이에 필자는 살아 있는 e러닝이 되기 위해서는 시멘틱 웹과의 통합이 필요하다고 생각해 2회에 걸쳐 시멘틱 웹과, 시멘틱 웹을 e러닝에 어떻게 적용할 것인지에 대해 얘기해 보고자 한다.

글 정의석 디유넷 컨설턴트



우리는 지금 정보의 대홍수 속에서 살고 있다. IT 기술과 오픈 네트워크가 결합되면서 전세계에 흩어져 있던 자료들이 그 물망처럼 얽혀 수 많은 자료들에 접근 할 수 있지만, 오히려 정보의 블랙홀에 빠져 사용자들이 정작 찾고자 하는 정보는 찾을 수 없는 기현상이 발생되고 있다. 1994년 제리 양이 만든 야후(Yahoo)를 필두로 해 알타비스타(Altavista), 라이코스(lycos)와 현재의 구글(Google)에 이르기 까지 검색 엔진은 흩어져 있는 자원들을 원 클릭으로(One-Click) 통합검색 서비스를 제공해주기 위한 노력을 해 왔지만 정보의 대홍수가 범람하고 있는 현재의 웹 구조에서는 사용자의 다양한 검색 니즈를 충족시켜주기에는 역부족이 아닌가 생각된다. 이러한 원인으로서는 지금 까지 모든 자원들은 사람이 이해하기 쉽도록 만들어진 HTML로 기술돼 정보의 외형적인 표현과 위치적 연결에만 중점을 뒀기 때문이다. 이에 컴퓨터는 자원의 의미를 이해하지 못하고 단지 검색 키워드가 들어가는 자원들은 모조리 쏟아내고 이들 자원 중에서 필요한 정보를 골라내는 몫은 사용자들에게 넘겨줘, 사용자가 필요한 정보를 찾는데 많은 시간을 낭비하도록 했다.

지금 검색 엔진을 통해 찾고자 하는 정보를 검색해 보라. 아마도 검색된 결과가 수십만 건에 달해, 당신은 몇 십번의 마우스 클릭을 통해 또 다른 정보의 향해를 떠나야 할 것이다.

복잡한 환경에 맞춰 변화하는 지식 패러다임

여기서 우리의 주 관심사인 e러닝으로 눈을 돌려 보자. 사회가 다양화되고 복잡해짐에 따라 '알'의 의미가 정보를 기억하고 반복하는 패러다임에서 정보를 쉽게 찾아서 이용할 줄 아는 패러다임으로 바뀌게 됐다. 구성주의 이론에 많은 영향을 받은 e러닝은 학습자와 관련된, 학습자에 맞는 학습 자료를 적응적으로 제공해줘, 학습자가 정보를 쉽게 찾아 이용해 스스로 지식을 구성해 나갈 수 있는 적응적 환경이 필요하다. 이를 위해서는 외형적 구조 관점뿐만 아니라, 발견적 학습(heuristic learning)관점에서 학습 자원의 의미적 기술이 컴퓨터가 이해 할 수 있

구분	월드와이드 웹(WWW)	시멘틱 웹
기반	HTML, HTTP, URLs	XML, RDF, Ontology, Agent
시스템	서버-클라이언트	Peer-to-Peer
정보 소비자	사람	사람 및 기계(프로그램)
데이터	외형중심	내용중심, 데이터 관계중심, 메타데이터
표현	하이퍼링크, 스크립트	데이터간의 관계(RDF)
시작 연도	1989년	1999년
추론	없음	인공지능, 에이전트

〈표 1〉 월드와이드 웹(WWW)과 시멘틱 웹의 비교

는 언어로 기술되고, 이를 컴퓨터가 이해해 학습자에게 필요한 정보만을 검색, 지식을 추론해주는 환경이 필요하다.

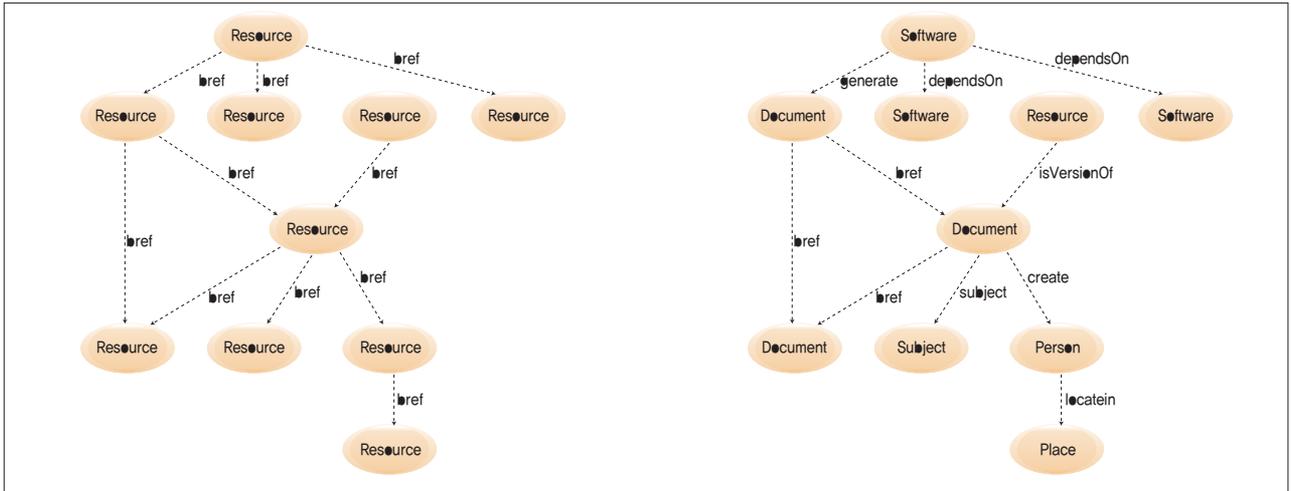
하지만 현재 e러닝에서 전달되는 학습 자료는 교수자가 이미 정해 놓은 한정된 자료가 모든 학습자들에게 친편일률적으로 제공이 되고, 검색 엔진을 통해 전달되는 학습 자료들은 학습자가 필요로 하지 않는 쓰레기 데이터(Garbage Data)들까지도 함께 학습자에게 전달돼 학습자들의 학습 효과를 떨어뜨리고 있다. 이는 웹에 존재하는 학습자원들이 컴퓨터가 의미를 이해할 수 없는 형태로 존재하고, LMS에 탑재돼 있는 학습자원 또한 다른 LMS의 학습자원들과의 의미적 연결 없이 각기 단편적으로 분산돼 있기 때문이다.

최근에는 사실상의 국제 표준인 SCORM(LOM), 국가 표준으로 채택된 KEM으로 학습자원을 기술해 학습자원들 간의 상호운영성(interoperability), 재사용성(reusability)등은 확보됐지만, 이 역시 학습자원들 간의 의미적 연결이 없고, 결정적으로 학습 자원의 의미를 컴퓨터가 이해하지 못하는 구조적 제약 조건을 가지고 있다.

지능을 가진 의미의 웹-시멘틱 웹

시멘틱 웹을 한마디로 정의하자면 '컴퓨터가 정보의 의미를 이해하고 추론을 통해 새로운 의미를 조작할 수 있는 똑똑한 웹'이라고 말할 수 있다.

웹(WWW)을 제안한 팀 버너스-리(Tim Berners-Lee)가 그의 저서 'Weaving the Web'에서 밝힌 '무엇이든지 서로 연결되는 공간'을 지



〈그림 1〉 월드와이드 웹(WWW)과 시멘틱 웹의 비교 (출처 : 멀티미디어 정보공학 연구실 김철순 PT자료)

구상의 컴퓨터 하나하나에 실현시키는 것이라고 말한 웹의 비전이 시멘틱 웹을 통해 비로소 실현 될 수 있다. 시멘틱 웹이 실현되면 검색어 포함 유무에 의한 검색이 아니라 정보의 ‘의미(개념)’를 이용한 검색이 가능해져 사용자가 원하는 정보만을 적응적으로 제공해 줄 수 있다. 기존의 웹은 단순히 자원들의 위치 연결로 구성돼 있고, 시멘틱 웹은 정보들의 의미적 연결구조로, 기존 월드와이드웹(WWW)과는 다음과 같은 차이점이 있다.

〈그림 1〉처럼 기존의 웹은 각 자원들이 하이퍼링크로 위치 연결만 돼 있지만, 시멘틱 웹은 위치적 연결뿐만 아니라, 의미적 연결까지 구성돼 있는 것이다.

이렇게 함으로써 웹 문서에 시멘틱 정보를 붙이고 이를 이용해 컴퓨터가 의미정보를 자동으로 추출할 수 있는 패러다임을 조성하는 것이다. 부수적으로 의미 정보의 자동 추출뿐 아니라 정보의 확장이나 공유 등도 가능하게 될 것이다. 향후 시멘틱 웹 환경 하에서는 정보들 간의 의미적 연결을 통해 새로운 지식을 창출할 수 있고, 더 나아가 정보의 신뢰성을 보장해주는 신뢰의 웹 구현도 가능 할 것이다.

시멘틱 웹의 구조

시멘틱 웹의 구성 핵심은 ▲자원의 성질과 다른 자원들과의 관계를 표현해주는 메타데이터 프레임워크(RDF) ▲각 자원들의 의미적 관계, 추론규칙을 정의한 온톨로지이다. 이들 요소들은 모두 기존의 Unicode를 지원하는 XML을 기반으로 하며, 시멘틱 웹의 계층 구조는 다음과 같다.

위의 각 계층별 기술들은 독립적으로 발전해 왔지만, 시멘틱 웹을 구현하는데 있어서는 하위 계층의 기술들은 상위 계층의 기술들이 구현되는 기반이 된다. 이 계층구조에서 보면 가장 하위 레벨에서 웹 프로토콜에서 자원을 지칭하기 위한 주소지정(addressing) 방법인 URI와 UNICODE가 밑받침되고 이를 기반으로 데이터의 구조를 정의하고 영역을 정의해주는 XML과 Namespace가 그 다음 층을 이루고, 다음 층에는 자원의 의미와 관계를 기술하기 위한 RDF와 RDF 스키마가 있고, 상위에는 자원의 정의와 관계를 정의한 온톨로지와 온톨로지를 지원하는 규칙, 논리, 증명을 위한 기술 요소의 층들로 구성돼 있다. HTML이 웹 문서를 사람이 쉽게 볼 수 있도록 화면에 보여주기(Displaying) 위한

언어였다면, XML은 웹에서의 자원의 기술(Describing)해주기 위한 언어이다. XML에서 사용자가 이해하기 쉽게 태그와 속성을 직접 정의 내려 구조적 정의를 제공해 줄 수는 있지만 정보 자원들 사이의 의미적 관계는 정의 내려주지 못하는 한계점이 있다. 이를 보완하기 위해 RDF를 이용해 웹 자원의 기술을 기계가 해석할 수 있는(machine-understandable) 형식으로 작성하고 이러한 정보를 교환하는 어플리케이션간의 상호운용성을 제공해 줄 수 있다.

RDF는 데이터 모형은 자원에 대해 지정된 속성과 그 값을 표현하기 위한 모델로, 자원(Resource), 속성(Property Type), 문(Statements)의 세 가지 객체유형으로 구성된다.

“jeong eui suk” is the creator of the resource <http://www.dunet.co.kr/goodguy>”의 문장을 3단계 RDF 데이터 모델로 표현하면 다음과 같다.

Subject (Resource)	http://www.dunet.co.kr/goodguy
Predicate (Property)	Creator
Object (literal)	jeong eui suk

〈표 2〉 3단계 RDF 데이터 모델

자원의 정의와 자원과 속성들 간의 관계를 나타내 주기 위해서는 온톨로지(ontology)가 구축돼야 한다. 온톨로지의 역할은 사용되는 어휘(vocabulary)의 공유된 의미를 정형적으로 표현하는 것이다. 일종의 어휘 사전이라고 이해하면 쉽게 이해될 것이다. RDF 스키마가 개체와 속성 사이의 관계성을 정의하고 있는 일종의 온톨로지를 표현하는 언어이고, 현재는 더욱더 어휘 정의표현을 풍부하게 해주는 온톨로지 언어(OIL,DAML+OIL,OWL)에 대한 연구가 W3C를 중심으로 활발한 연구가 진행되고 있다.

학습 환경에서 의미적으로 동일한 개념이 해당 영역, 저자에게 따라 다른 키워드로 표현될 수도 있는데, 이 문제는 온톨로지에서 도메인 어휘를 통합함으로써 해결할 수 있다. 예를 들어, 의미적으로 ‘Agent’와 동일한 어휘들인 ‘agent, actor, contributor, creator, player, doer, worker, performer’ 등의 어휘들을 도메인 온톨로지를 통해 컴퓨터가 동일한 의미로 이해 할 수 있는 것이다. 🌀