

웹2.0과 시맨틱웹, 그리고 진화의 방향

원광대학교 | 한성국* · 정영식* · 유재규**

1. 서론

팀 버너스리(Tim Berners-Lee)에 의해 1989년 웹 기술이 출현한 이래로, 웹기술은 어느 정보기술보다도 사회발전에 심대한 영향을 주었다. 웹 기반 원격교육, 전자상거래, 전자도서관, 인터넷 포털 등 지식 정보사회의 새로운 모습들이 웹 기술을 기반으로 선보였다. 웹 기술 응용의 급속한 확산과 더불어, 기술적으로도 HTML, DHTML, XHTML 등 개선된 기술이 속속 개발되었고 CSS, 플래시 등 연관 기술들이 숨 가쁘게 선보였다. 웹 기술은 정보 기술의 새로운 지평을 열어 보인 기술로서, 아직도 미완인 채 새로운 기술의 출현을 기다리고 있다.

최근 웹2.0이 인터넷 비즈니스 업계의 화두가 되면서 웹 기술은 또 다른 변신을 기도하고 있다. 구글(Google)이 유튜브(YouTube)를 1억6천만 달러에 인수하면서 더욱 관심이 증폭된 웹2.0은, 2000년 닷컴 버블 붕괴 이후, 인터넷 비즈니스의 황금시장으로 인식되고 있다. 페이스북(Facebook), 세컨드라이프(Second Life), 마이페이스(myFace) 등 웹2.0을 표방하는 벤처 비즈니스 업체가 우후죽순처럼 등장하였고, 사용자 참여와 UCC가 비즈니스 모델의 핵심으로 회자되고 있다. 기업의 이러한 열광적 관심 덕분에, 웹2.0은 단기간 내에 웹 기술의 대표 자리를 점유하게 되었다. 그러나 실체가 없는 비즈니스 모델로써의 웹2.0에 대한 기업의 열광은 하이프 논쟁과 함께 또 다른 닷컴 버블의 징조로 우려의 비판도 고조되고 있다. 웹2.0에 관한 너무 많은 상세 자료와 블루오션 전망 보고서가 출판되어 있어, 웹2.0을 논의하는 것이 진부한 감이 있지만, 웹2.0 열풍을 진지하고 객관적으로 조명해 볼 필요가 있다.

비즈니스 분야에서의 웹2.0에 대한 요란한 관심과는 대조적으로 학계와 연구계에서는 차세대 웹기술의 원형을 지속적으로 탐구하여 왔다. 2001년 팀 버너스리

등이 차세대 웹 기술의 비전으로 시맨틱웹(Semantic Web)을 제시한 이래로, W3C를 중심으로 시맨틱웹 계층구조가 제시되었다. 또한, 온톨로지 기술언어인 RDF/S와 OWL, 질의어인 SPARQL, 규칙언어인 SWRL 등의 표준화가 진행되었다. 시맨틱웹은 기존의 정보 중심 웹이 당면한 제반 문제를 해결하고, 지식 추론이 가능한 지식웹(Knowledge Web)을 구현하여, 환상적인 지식 정보 서비스를 실현할 수 있을 기대감을 고조시켰다. 그러나, 처음의 예측과는 달리, 시맨틱웹의 표준화와 관련 기술의 연구개발 노력에도 불구하고 킬러앱(killer applications) 논쟁과 함께 제2의 인공지능 겨울(AI Winter)이 도래하는 것이 아닌가 하는 우려감도 제기되고 있다. 차세대 웹의 비전으로써 시맨틱웹에 대한 개념 이해와 확실한 방향 정립이 필요한 시점이다.

우리는 지금 서로 다른 관점의 두가지 웹 기술의 혼재 속에서 차세대 웹 기술이 지향해야 할 활로를 탐색하고 있다. 이와 관련하여 웹2.0과 시맨틱웹의 상보적 협력관계나 웹3.0 등에 대한 산발적인 논의가 있어 왔지만 심층적인 접근은 미흡하였다.

본 고에서는 기존 웹 기술을 분석하여, 차세대 웹 기술을 가시화하고 앞으로의 로드맵을 모색해 보고자 한다. 먼저, 제2장에서 웹2.0을 체계적으로 분석하여, 웹 2.0의 실상(참모습)을 파악하고 개념과 특성을 정리 요약하며, 제 3장에서 시맨틱웹의 비전과 관련 기술의 핵심 개념을 파악한다. 제 4장에서 웹2.0과 시맨틱웹의 차이점과 공통점을 도출하여, 상호 보완을 통한 차세대 웹기술의 실체를 조망해 본다. 제 5장에서는 웹기술의 지속적인 연구 개발 체제 확립과 학제적 학문 접근을 위한 웹과학(Web Science)에 대하여 알아보고자 한다. 이를 통해 차세대 웹 기술의 진화 방향을 이해하고, 핵심 기술의 연구 개발에 이해를 돕고자 한다.

2. 웹2.0의 실상

웹2.0이라 하면 UCC나 인맥형성 소셜네트워크(social network)가 생각날 정도로 웹2.0의 원래 개념이 비즈

* 종신회원

** 학생회원

니스적으로 왜곡되고 있다. 비즈니스는 근본 개념보다는 고객을 우선하게 되므로 최종 제품과 서비스에만 관심이 있다. 웹2.0의 실상을 인터넷 정보기술의 측면에서 보다 정확하게 고찰해 본다.

2.1 웹2.0의 개념과 의의

웹2.0의 출현 배경에 대한 참고 자료가 너무 많이 있지만, 인터넷 정보기술의 측면에서 다시 요약, 정리해 보자. 2000년 3월 갑작스러운 닷컴 버블 붕괴로 각광 받던 수많은 인터넷 기업이 줄지에 소멸되었다. 그럼에도 불구하고 인터넷 비즈니스는 아마존, 구글, 이베이 등이 시장을 주도하면서 많은 발전을 이루었다. 이에 따라, 비즈니스 측면에서 살아남은 기업과 소멸한 기업의 차이점을 분석해 볼 필요성이 제기되었다. 그래야, 성공할 수 있는 인터넷 비즈니스 모델을 확실히 정립할 수 있을 것이다. 살아남은 기업들의 핵심 비즈니스 역량을 분석한 결과, 소비자의 행동 양식과 소비 행태에 있어 커다란 변화가 목격되었고, 성공한 인터넷 기업이 지향해 온 기본 원칙이 도출되었다. 소비자의 변화, 즉 인터넷으로 인한 사회 인식의 변화와, 성공한 인터넷 기업의 기본 원칙 등 기존 비즈니스와는 차별화된 인터넷 비즈니스 특성을 본격적으로 논의하기 위하여 “웹2.0 컨퍼런스”(2004년 10월 5일)가 개최된 이래로, 웹2.0은 차세대 인터넷 기술의 대명사로 인식되었다.

웹2.0 출현배경에서 알 수 있는 바와 같이, 웹2.0은 인터넷 비즈니스 기술의 변화를 요약, 정리하여 체계화한 것이다. 이는 마치 동남아에서 각광을 받고 있는 한국 드라마, 한국 대중가요, 김치 등 한국음식에 대한 열광적 관심을 한류로 명명하여 지칭하는 것과 같다. 어린아이가 나이가 들어 20대 성년이 되면, 성숙된 자아의식과 독자적인 인생관을 형성하는 특성을 보이듯이, 웹2.0은 기존의 인터넷 기술의 발전과 변화 특성을 표현하는 개념이다. 즉, 웹2.0은 인터넷 정보 기술의 변화 방향 또는 경향을 지칭하는 정보사회적 개념인 것이다. 웹2.0은 신기술도, 패러다임의 변화도 아니다. 현재, 우리가 직면하고 있고, 앞으로 지향하여야 할 웹 기술의 특징을 개념적으로 대표하는 용어일 뿐이다. 위키나 블로그만을 웹2.0이라고 할 수 없고, 역으로 웹2.0을 위키나 블로그라고 할 수 없다. 웹2.0 이전에 이미 위키와 블로그가 존재하였고, RSS는 시맨틱웹의 RDF 응용일 뿐이며, 매쉬업(mashup)은 컴포넌트지향 개발에서 이미 사용하고 있던 방법이다. 웹2.0의 이러한 특성으로 인하여, 웹2.0의 실체나 기반 기술을 논의하는 것은 무의미하다. 중요한 것

은 웹2.0 논의에서 밝혀 낸 차세대 웹이 지향하여야 할 이념인 것이다.

웹2.0은 인터넷을 기반으로 하는 사회 활동 전반의 변모 양상을 크게 부각시켰고, 변화의 개념과 방향을 명확하게 제시하였다. 웹2.0에서 밝혀낸 정보기술에 대한 사회적 인식 변화와 요구되는 기본 원칙 등에 대한 심도 있는 논의가 필요하다.

2.2 웹2.0의 핵심개념

웹2.0의 실체에 대한 논란속에서 팀 오라일리(Tim O'Reilly)는 웹2.0의 7가지 원칙과 8가지 설계패턴(design pattern)을 제시하였다[1]. 팀 오라일리가 제시한 웹2.0의 개념은 새로운 사실이 아니지만, 차세대 웹 기술을 모색하는 현시점에서 깊이 음미해야 할 중요한 내용을 시사하고 있다. 팀 오라일리의 논점을 다시금 요약하면, 웹2.0은 서로 긴밀하게 연관된 4가지 핵심 개념이 도출된다.

· 개방(Openness)

인터넷은 기본적으로 분산, 개방 환경의 기술이다. 정보와 서비스를 개방하고 공유하여, 정보자원의 가용성과 효용성 증대를 지향하는 열린, 평등공간을 위한 기술이다. 누구든지 참여가 가능하고 상호교류가 보장되며, 원하는 정보자원에 자유스러운 접근이 보장되는 장벽과 차별이 없는 정보공간이다. 차세대 웹은 정보 시스템의 기반 플랫폼으로서 기술과 사회적 제약을 넘어, 웹 공간을 보다 열린 공간으로 확장하여야 할 것이다.

· 공유(Sharing)

웹 공간은 기본적으로 모든 정보자원의 공유를 목적으로 한다. 개방된 정보자원을 상호 공유함으로써, 정보자원의 재사용성과 가용성 등을 제고할 수 있으며, 공유된 정보자원의 창조적 혼합을 통해서 새로운 유용한 정보자원의 생성과 풍부한 사용자 경험을 제공한다. 공유는 네트워크 효과(network effect)를 실현하여 정보자원의 가치를 극대화하는 원동력이기도 하다. 또한 소스 프로그램의 공유(open source)와 데이터베이스의 공유(open database)는 협업을 통해 신속하고 지속적인 기술 개발체제를 형성한다. 차세대 웹에서는 사진, 동영상, 문서 등의 기본 정보자원 뿐만 아니라 서비스를 포함한 다양한 형태의 정보자원에 대한 공유가 이루어져야 할 것이다.

· 참여(Participation)

기존의 웹은 정보제공자가 공급하는 정보를 단순히 소비하는 공급자 중심의 독점화된 일방향 기술이었다.

차세대 웹에서는 공급자로부터 소비자로의 권력이동 (power shift)이 발생하여, 정보자원의 생성, 공급 및 소비의 모든 과정에 사용자가 직접 참여하는 양방향 상호작용 기술로 변모하고 있다. 블로그, UCC 등에서 보는 바와 같이, 사용자의 참여는 새로운 인터넷 문화와 비즈니스의 원동력이 되고 있으며, 위키피디아처럼 놀라운 결과를 가져오고 있다. 열린 인터넷 공간에서 참여가 보장되지 않는 웹 기술의 도태는 가속화될 것이다.

· 협업(Collaboration)

웹은 표준화된 정보자원의 전달 통로나 매체가 아닌 협업 공간으로 진화하고 있다. 함께 연구 토의하고, 함께 설계 개발하고, 함께 조립 제작하는 공동의 작업공간으로 변모하였다. 사용자의 참여와 상호협력을 통해서 새로운 인터넷 문화와 생활 방식을 만들어 가는 거점이 되었다.

웹2.0의 기본 정신(ethos)으로 개방/공유, 참여/협업으로 요약 정리한 논점은, 지금까지의 웹기술 발전을 고려하여, 차세대 웹이 구비하여야 할 핵심특성을 명확하게 밝혀내었다 할 것이다. 웹2.0은 사용자가 참여하는 인간 중심의 웹 기술을 표방하고 있다. 웹2.0에서 지향하는 이러한 차세대 웹의 특징은 몇가지 근거를 기반으로 하고 있다. 집단지성(collective intelligence), 긴꼬리(long tail) 비즈니스 모델, 대규모 데이터의 공유, 서비스 지향(service-oriented), 재사용(reuse)의 효용성 등 웹 기술에서 증명된 현상들이 이러한 방향을 묵시적으로 지지해 주고 있다.

웹2.0이 지향하는 이념, 즉 개방/공유와 참여/협력을 실제로 실현하여 줄 방법 또는 도구로 매쉬업(mashup), 폭소노미(folksonomy), AJAX, 마이크로포맷(Microformat),

Open API, SOAP과 REST 등 여러 기술을 중요하게 여기고 있다. 이외에도 Python, Ruby-on-Rails 등 스크립트 언어도 종종 거론된다. 그러나, 이러한 방법과 도구는 웹2.0을 위해서 새로이 개발된 것이 아니라, 이전부터 사용해온 상용 방법과 도구이다. 다만, 이러한 방법과 도구가 웹상에서 개방/공유와 참여/협력의 시스템을 구현하는데 유용할 뿐이다. 앞서 설명한 바와 같이, 차세대 웹을 위해서 방법과 기술을 구별하여 논의하는 것은 무의미하다. 모두 차세대 웹을 실현하기 위한 요소 기술이기 때문이다.

웹상에서 개방/공유와 참여/협력을 구현한 응용으로서 위키, 블로그, UCC, 소셜네트워크(social network) 등 여러 형태가 있다. 웹2.0에서는 개방/공유와 참여/협력의 이념을 창조적으로 구현할 수 있는 응용 또는 그러한 비즈니스 모델의 발굴이 가장 중요하다. 우수한 웹2.0 사이트들은 모두 독자적인 방법으로 웹2.0이 추구하는 이념을 실현하고 있다[2]. 예를 들어, 사용자의 이야기나 소감을 공유하고 점수를 매겨주는 Digg, 프로젝트 관리와 팀원간 협업을 제공하는 Basecamp, 웹을 플랫폼으로 하여 다양한 정보 서비스를 제공하는 Netvibes 등 모두가 개방/공유 및 참여/협력을 색다르게 실현하고 있다. 웹2.0의 핵심 목표가 개방/공유와 참여/협력의 응용서비스와 비즈니스 모델 개발에 있음을 다시금 확인할 수 있다.

지금까지의 웹 2.0 논의를 요약하면 그림 1과 같다. 웹 2.0은 정보 기술의 발전 동향과 성공한 인터넷 기업의 비즈니스 개념을 결합하여, 차세대 웹이 지향하여 할 이념을 제시하고 있다. 웹 2.0이 표방하는 이념을 고려하면, 현존하는 모든 정보기술이 웹 2.0과 직간접적으로 연관관계가 있다. 다시 말해서, 웹 2.0의

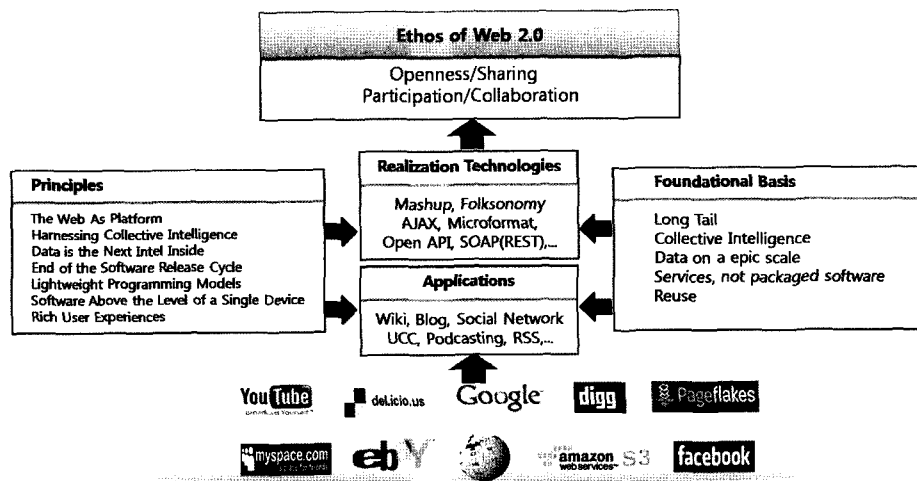


그림 1 웹 2.0의 개념 요약

외연은 모든 정보기술과 사회 분야로 확대될 수 있다. 웹 2.0의 외연은 성격에 따라, 그림 2와 같이 논리적 외연과 물리적 외연으로 나누어 생각할 수 있다. 논리적 외연이란, 웹 2.0이 표방하는 에토스를 중심으로 한 개념의 확대로, 예를들면, Library 2.0, Enterprise 2.0, Government 2.0 등 기존 체계에 웹 2.0의 개념을 접목하는 것이다. 도서관은 전통적으로 공급자인 도서관과 사서에 의해서 모든 정보가 제공되고 관리되어 왔다. 때문에, 사용자가 참여하거나, 함께 정보를 공유할 수 있는 기회가 제공되지 않았다. 이러한 도서관 시스템에 개방/공유 및 참여/협력의 웹 2.0의 개념을 접목한 것을 Library 2.0이다[3]. 웹 2.0이 추구하는 근본 개념은 인간 중심의 정보 서비스이므로, 이처럼 모든 분야에 접목하여 확대가 가능하다. 한편, 물리적 외연은 웹 2.0이 주목하고 있는 사용자 중심의 서비스 실현과 관련된 정보 기술로, RIA (Rich Internet Application), 웹서비스, 서비스 지향구조(service-oriented architecture : SOA), SaaS(Software as a Service) 등과 같은 최근의 정보 기술을 웹 2.0의 외연 기술로 간주할 수 있다. 이들 기술은 웹 2.0의 수직 상승에 편승해서, 정보 서비스 시스템 구축의 필수 기술로 인식되고 있다. 대표적 RIA 기술인 Adobe의 Flex, 마이크로소프트의 SilverLight, Sun의 JavaFX는 웹 2.0의 대표 기술이 되기 위해 치열하게 경쟁하고 있으며, 마이크로소프트의 Windows Live, 구글의 Doc & Spread sheets 등 웹탑(webtop) 기반 RIA 서비스 개발이 가속화 되고 있다. 앞으로의 정보기술은 사용자 중심의 다양하고 편리한 정보 서비스 제공에 초점을 두어야 할 것이다.

3. 시맨틱웹의 이해

시맨틱웹은 어렵고 복잡하며 현재의 웹과는 동떨

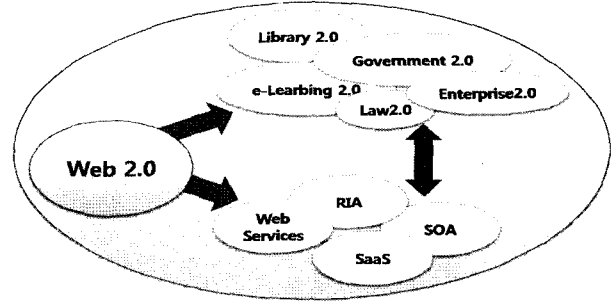


그림 2 웹 2.0의 확장

어져 있는 학계의 연구과제로 생각되는 경우가 많다. 시맨틱웹의 현재 그리고 미래를 중심으로 시맨틱웹의 개념과 구조를 요약 정리한다.

3.1 시맨틱웹의 개념

웹 기술은 정보표현과 전달에 간편한 방법을 제공하여 인터넷이 실생활까지 확산되는 기폭제 역할을 하였다. 그러나, 웹상에 축적된 정보가 방대해 짐에 따라 많은 문제에 봉착하게 되었다. 웹 기술은 축적된 방대한 데이터에 대하여 키워드(keyword)에 의한 정보 접근만을 허용하고 있어, 정보 검색시 무수히 많은 불필요한 정보가 돌출하여 정보 홍수를 가중시키고 있다. 또한, 컴퓨터가 필요한 정보를 추출, 해석, 가공할 수 있는 방법이 없어, 그림3(a) 처럼 모든 정보를 사용자가 직접 개입해서 처리하여야 하는 문제가 있다. 이러한 문제들의 근본원인은 컴퓨터가 정보자원의 의미를 이해하지 못하는데 원인이 있다. 이러한 웹 기술은 팀 버너스리가 초창기에 구상하였던 웹과도 거리가 있다.

2001년 팀 버너스리 등에 의해 웹 기술의 비전으로 시맨틱웹이 제시되었다[4]. 시맨틱웹은 기존 웹을 확장하여 컴퓨터가 이해할 수 있는 잘 정의된 의미를 기반으로 의미적 상호운용성(semantic inteoperability)을

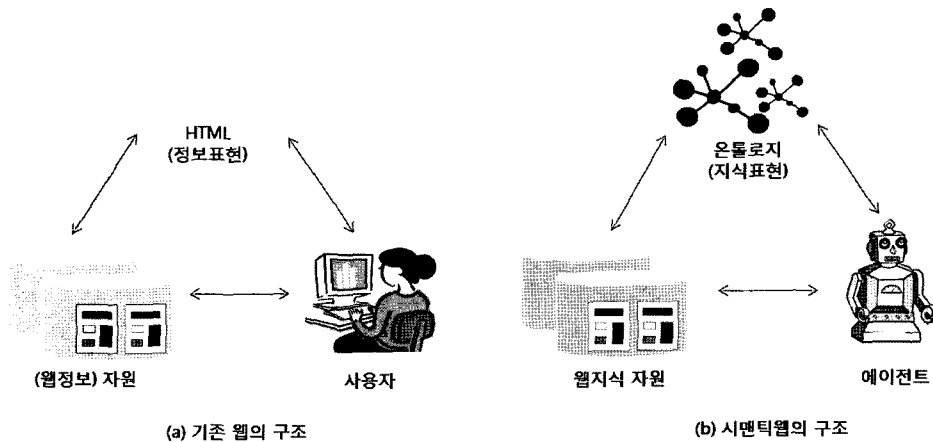


그림 3 기존웹과 시맨틱웹의 구조

실현하여, 다양한 정보자원의 처리 자동화, 데이터의 통합 및 재사용등을 컴퓨터가 스스로 수행하여, 인간과 컴퓨터 간의 효과적인 협력체계를 구축하기 위한 것이다. 즉, 시맨틱웹은 컴퓨터가 웹 정보자원의 의미를 이해하고, 정보의 검색, 추출, 해석, 가공등 제반 처리를 사용자를 대신하여 지능형 에이전트(intelligent agent)가 수행하는 컴퓨터 중심의 기술이다. 일반적으로 시맨틱웹의 개념은 그림 3(b)로 표현할 수 있으며, 3가지 주요 요소로 구성된다.

· 온톨로지(ontology)

온톨로지는 도메인의 공유된 개념화(conceptualization)에 대한 형식적 명세체제로서, 도메인 어휘의 의미 정보를 표현한다. 온톨로지는 일종의 지식 표현(knowledge representation)으로, 컴퓨터는 온톨로지로 표현된 개념을 이해하고 지식처리를 할 수 있다. 추론, 증명등의 처리에 온톨로지의 공리(axiom)와 규칙(rule)이 사용되며, 규칙 표현을 위해서 별도의 규칙언어가 사용된다.

· 의미적으로 주석화된 웹(semantically annotated Web)

온톨로지 주석화된 웹 정보 자원은 일종의 지식 베이스를 형성한다. 시맨틱웹에서는 온톨로지의 의미적 상호 운용성을 기반으로 인터넷의 분산 정보 자원을 의미적으로 통합한 거대한 지식 베이스를 구축할 수 있다.

· 에이전트(agent)

인간(사용자)를 대신하여 정보 자원을 수집, 검색하고 추론하여, 온톨로지를 이용해서 다른 에이전트와 상호 정보 교환 등의 일을 수행하는 지능형 에이전트이다. 지능형 에이전트는 시맨틱웹 기반 응용 서비스의 핵심 요소라 할 수 있다.

시맨틱웹은 정보처리의 실제적 주체인 컴퓨터가 의미를 이해할 수 있도록 하고, 에이전트가 사용자 대신에 업무를 수행하는 지능형 웹 기술이다. 때문에,

맨틱웹에서는 지식 표현의 역할을 하는 온톨로지가 무엇보다도 중요하다. 인공지능 등의 연구에서 알 수 있는 바와 같이, 지식 표현은 도메인의 관점, 시스템 구축의 목적, 지식의 활용 등에 따라 다양한 표현방식이 가능한 난제중의 하나이다. 시맨틱웹에서는 웹 정보자원의 의미를 효과적으로 기술 할 수 있도록, 그림 4처럼 시맨틱웹 계층을 구성하였다.

- Unicode/URI : 웹 정보 자원 서술과 식별을 위한 표준 체계
- XML+NS+XML Schema : 정보구조와 전달 등 정보 표현 형식을 위한 표준 체계
- RDF+RDF Schema : 정보자원의 의미적 연결관계를 정의하고, 메타데이터 수준의 의미 표현을 위한 표준체계
- Ontology : 도메인 온톨로지를 기술하는 표준 체계
- Logic : 온톨로지 기반의 추론을 위한 표준 논리 체계
- Proof, Trust : 논리체계를 이용해서 증명하고 웹 정보자원의 신뢰성을 평가하는 표준 체계

그림4에서 알 수 있는 바와 같은, 시맨틱웹은 웹 정보자원을 대상으로 강력한 의미 표현을 위한 표준 기술 체계를 확립하는데 초점을 두고 있다. 이러한 표준 체계는 웹 정보자원의 의미적 상호 운용성을 실현하는 기반이 된다.

3.2 시맨틱웹 요소 기술및 전망

시맨틱웹에 대한 관심과 연구가 증폭되면서, 관련 기술의 지속적인 연구, 개발이 이루어졌다. 특히 의미적 상호 운용성의 실현을 위한 핵심 기술의 표준화가 W3C를 중심으로 집중적으로 추진되어 왔다. 2004년 1단계 표준화 활동을 정리하고, 시맨틱웹 응용 사례 발굴과 상위 계층 표준화를 위한 2단계 표준화 활동이 진행되고 있다. W3C의 시맨틱웹 표준화 추진 현황[12]을 표 1에 요약하였다.

시맨틱웹 표준화와 연관 기술의 개발로, 초창기에 제시되었던 시맨틱웹의 계층구조가 더욱 정교화되었다. 최근에는 그림 5와 같은 시맨틱웹 계층구조가 새로이 제시되어, 이에 대한 활발한 논의가 진행되고 있다. 그림 4의 초창기 계층구조와 비교해 보면, 몇가지 변화된 점을 확인할 수 있다. 이로부터, 시맨틱웹 실현을 위한 핵심 요소 기술을 새로이 발견할 수 있다.

- URI/IRI : 웹 정보 자원의 식별을 위하여 URI를 사용하여 왔지만, 다양한 정보자원 형태와 프로토콜의 출현으로 통합된 식별체계가 요구되고 있

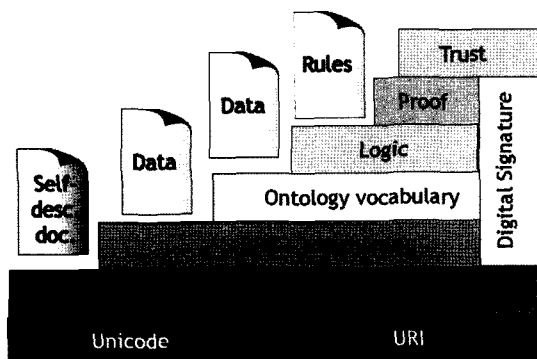


그림 4 시맨틱웹의 계층구조(2001년)

표 1 W3C 시맨틱웹 표준화 추진현황[12]

	Phase 1					Phase 2				
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Semantic Web Interest Group										
Semantic Web CG										
RDF Core WG(종료)										
Web Ontology WG(종료)										
RDF Data Access WG										
Semantic Web Best Practice and Deployment WG(종료)										
Semantic Web Deployment(SWD) WG										
Rule Interchange Format WG										
Semantic Web Health Care and Life Sciences IG										
GRDDL WG										
Semantic Web Education and Outreach(SWEO) IG										

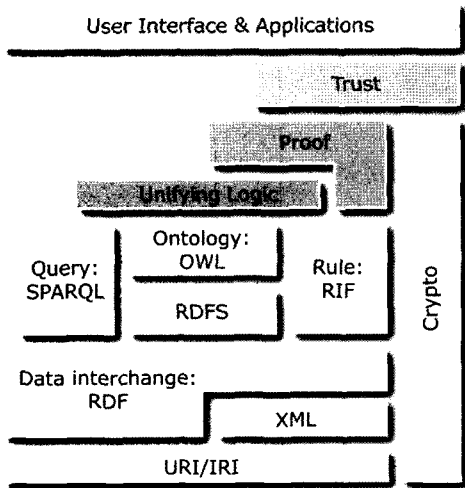


그림 5 시맨틱웹의 계층구조(2006년)

다. 이에 따라 URI를 확장한 IRI(Internationalized Resource Identifier)의 필요성이 증대되고 있다.

- 데이터 상호 교환(data interchange) : 데이터 표현 및 전송 등 상호 교환의 기본 형식으로 XML이 사용되어 왔지만, XML의 기능은 상대적으로 축소된 반면에 RDF 형태가 시맨틱웹의 기본 형식으로 자리잡아 가고 있다. 시맨틱웹에서는 단순 정보 구조화를 위한 XML보다 정보자원의 의미 기술과 관계 형성이 기본이 되기 때문이다.
- 온톨로지, 질의(query)와 규칙(rule)
시맨틱웹의 핵심 계층이 온톨로지, 질의 및 규칙으로 세분화되었다. 온톨로지로 텍소노미(taxonomy)나 시소리스(thesaurus) 등과 같은 단순한 얇은 온톨로지(shallow ontology)를 위한 RDF와, 도메인의 복잡한 깊은 온톨로지(deep ontology)를 위한 OWL로 구분하여 온톨로지 개발 및 구축의 효율

성을 제고하고 있다. 온톨로지로 구현된 시맨틱웹을 검색하는 기본 검색어와 프로토콜로 RDF 그래프 기반의 SPARQL이 표준화 되었다. 한편, 서로 다른 다양한 시스템간 규칙의 공표(publication), 공유, 재사용할 수 있도록 규칙 공통언어(interlingua)로서 규칙 상호 교환 형식(Rule Interchange Format : RIF)의 표준화를 진행하고 있다. 기존에 다양한 규칙언어와 규칙엔진이 널리 사용되고 있기 때문에, 이들의 상호 운용성을 보장하는 공통형식(common format)을 제정하여 시맨틱웹에서 다양한 규칙 응용을 보장하고자 한다.

- 통합 논리(Unifying Logic)
통합 논리는 RDF, RDFS, RIF, OWL, SPARQL 등 시맨틱 웹 요소 기술에 공통으로 적용되는 모델-이론 의미론(model-theoretic semantics)기반의 논리 프레임워크(logic framework)이다. 통합 논리는 규칙 데이터의 공유와 일관된 추론 결과 생성을 위한 것이다.
- 웹 정보자원에 의미 정보 추가
시맨틱웹이 활성화되기 위해서는 온톨로지와 같은 컴퓨터가 이해할 수 있는 의미 정보를 포함하고 있는 웹 페이지가 대량 생산되어야 한다. 온톨로지로 주석화된 웹페이지의 개발이 어려운 만큼, 기존의 웹페이지에 컴퓨터가 이해할 수 있는 의미 정보를 추가가 하는 것이 현실적인 방법이다. 이를 통해서, 인간(사용자)과 컴퓨터가 모두 접근 가능한 웹을 실현할 수 있다. (X)HTML의 class 속성을 활용하여 웹페이지 안에 행사일정이나 연락처 등의 의미 정보를 추가하고 컴퓨터가 이를 알아낼 수 있도록 하는 hCard, hCalendar

등 여러 데이터 형식을 Microformat이라 한다. Microformat은 표준은 아니지만, 커뮤니티를 중심으로 웹 페이지 개발의 설계 지침으로 확산되고 있다. RDFa는 컴퓨터가 처리할 수 있는 RDF 형식의 의미 데이터를 XHTML또는 XML문서에 직접 추가할 수 있는 방법을 제공한다. GRDDL (Gleaning Resource Description form Dialects of Language)는 XML또는 XHTML 문서에 추출, 수집 가능한 데이터를 선언하고, 이들 문서로부터 RDF데이터를 추출, 수집하기 위한 연계 알고리즘을 정의할 수 있는 마크업언어이다. GRDDL은 인간 중심의 기존 웹과 컴퓨터 중심의 시맨틱웹의 간격을 바꾸어 줄 수 있을 것으로 기대되고 있다.

시맨틱웹은 이제 더 이상 연구를 위한 기술이 아니라, 상용화 단계로 발전하고 있다[6,7]. eClassOWL, BioPAX와 같은 다양한 온톨로지, AllegroGraph, Top-Braid와 같은 상용도구가 개발되었고, 노키아 S60 포털, Haper 온라인 매거진, Yahoo!Food등 수 많은 사이트들에서 시맨틱웹 기술이 활용되고 있다. 이미 인

터넷 상에는 10⁷ 시맨틱웹 문서가 존재하는 것으로 알려지고 있다. Garlik, Radar Networks, Joost, Ontology Works 등 수많은 시맨틱웹 기업이 활발한 제품 개발과 비즈니스 활동을 전개하고 있다.

분산된 정보자원을 의미적 상호 운용성으로 통합하여 거대한 지식 베이스를 구축한 시맨틱웹은, 웹서비스에도 의미적 상호운용성을 실현하여, 궁극적으로 그림 6과 같이 시맨틱 웹서비스(Semantic Web Service)로 진화하여 갈 것이다. 온톨로지 언어 OWL 개발과 동시에 서비스 온톨로지인 OWL-S가 개발되었고, W3C의 SWSI나 유럽의 WSMO/L/X 연구등에서 알 수 있는 바와 같이, 시맨틱웹은 지식을 넘어서 지식 서비스로 나아가고 있다[11].

4. 웹2.0과 시맨틱웹의 동거

전장의 개요 설명과 분석으로 알 수 있는 바와 같이, 웹2.0과 시맨틱웹은 접근 방식에서 차이는 있지만, 기존의 웹을 한 차원 진화된 웹으로 발전시키기 위한 노력이다. 이러한 노력의 우열을 가르거나 선호도를 구분하는 것은 무의미하다. 차세대 웹 기술을

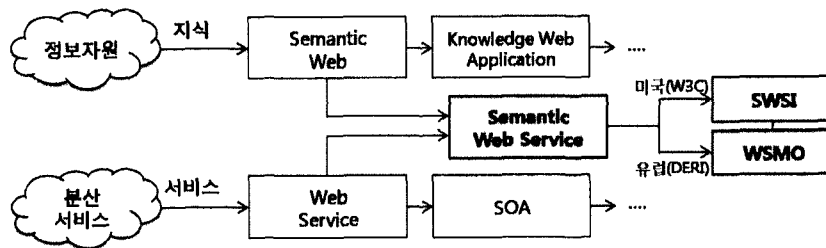


그림 6 시맨틱 웹 서비스

표 2 웹2.0과 시맨틱웹의 상보성 비교

	Web 2.0	Semantic Web
개발자(제안자)	Tim O'Reilly	Tim Berners-Lee
초점	Human-Oriented(Human-to-Human)	Computer-Oriented(Computer-to-Computer)
관심	비즈니스 측면	연구개발 측면
관점	플랫폼으로서의 웹	지식정보자원으로서의 웹
에토스	개방/공유, 참여/협력	개방/공유, 지능화
기반사항	집단지성, Long-tail	Ontological Commitment
핵심요소	Mashup, Open API	온톨로지, 규칙
표준 요소 기술	다양한 일반 정보기술	RDF, RDFS, OWL, RIF, GRDDL, RDFa, ...
응용개발 지원요소	데이터+Open API	온톨로지+추론엔진
응용방향	Rich Internet Application (RIA)	지능형 서비스 에이전트
대표 응용형태	Wiki, Blog, UCC, ...	Semantic Portal, Semantic Desktop, ...
서비스 인식	Software-as-a-Service	Semantic Web Service
사회현상	Library 2.0, Enterprise 2.0, Government 2.0, ...	Semantic Library, Semantic Grid, Semantic Web Services, ...

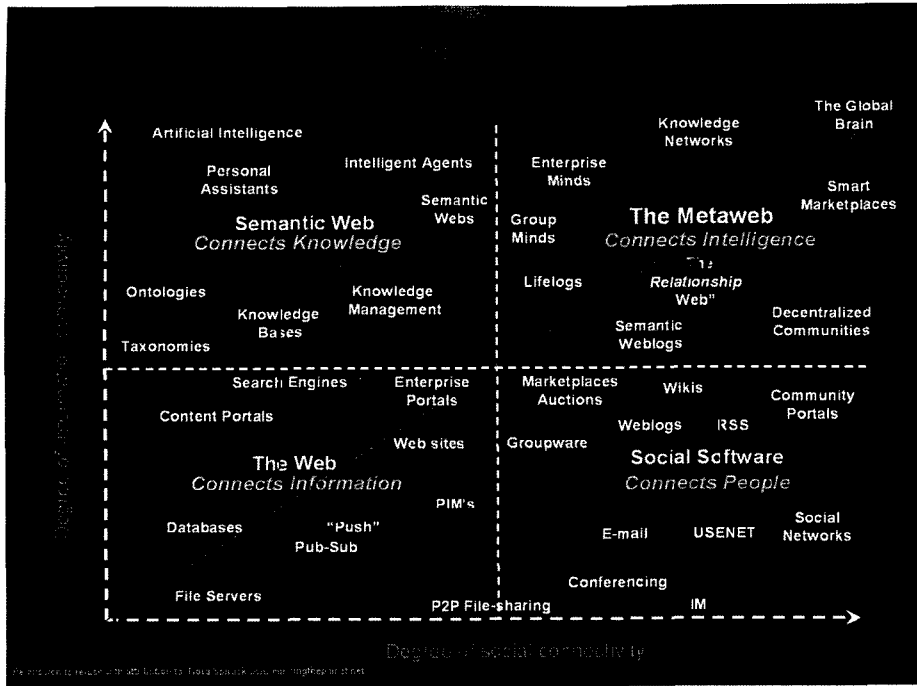


그림 7 차세대 웹 기술의 발전방향

위해 여러 접근 방식이 있다는 것은 각 접근 방식이 간과하거나 고려하지 못했던 점을 상호 보완할 수 있다는 의미이다. 즉, 웹2.0과 시맨틱웹은 경쟁관계가 아니라 상호 보완적 협력 관계이다[8]. 때문에 웹 2.0과 시맨틱웹을 항목별로 비교하는 것은 다소 무리가 따르지만, 각각의 특성을 대조적으로 표2와 같이 요약할 수 있다.

표 2에 나타난 것처럼, 웹2.0과 시맨틱웹은 웹 기술에 대한 고유의 시각과 접근방식을 가지고 진화된 웹기술의 일익을 수행하고 있다. 현재는 서로 독립적이며 자발적으로 발전되어 온 웹2.0과 시맨틱웹은 자신의 영역을 더욱 확대하기 위하여 상존하고 있지만, 궁극적으로는 자연스럽게 통합되어 차세대 웹 기술로 나타날 것이다. 이미, Microformat, RDFa, GRDDL, RSS 등에서 상호 협력이 진행되고 있고, 인간중심의 RIA 인터페이스와 컴퓨터 중심의 시맨틱 비즈니스 프로세스를 기반으로 하는 새로운 엔터프라이즈 시스템의 개발이 가속화되고 있다. 웹2.0과 시맨틱웹은 경계가 분명한 경쟁기술이 아니라, 그림 7의 차세대 웹 기술 발전 추세에서 보는바와 같이 차세대 웹을 향한 진행과정의 기술로 가까운 시일 내에 서로를 포용하게 될 것이다.

5. Web Science의 출현

웹 기술이 출현 이래로, 교육, 경제, 문화, 산업 등 모든 분야에서 커다란 변화가 있었다. 웹 기술이 사

회 변혁의 중추적인 역할을 하고 있음에도 불구하고, 이에 대한 체계적인 접근이나 연구가 아주 미흡하였다. 웹2.0과 시맨틱웹에서 간헐적으로 웹에 대한 구체적인 시각이 제시된 바가 있지만, 이들은 주로 기술적 측면에서 발현된 것이다. 웹과 마케팅, 상거래 등 경제학과의 관계, 인터넷 중독이나 가상공간과 인지체계등에 대한 인지심리학적 접근, 웹 기반 미디어 출현에 의한 사회학적 현상 등등, 웹은 이제 학제간의 관점에서 체계적으로 연구되어야 할 학문의 대상이 되었다. 이럴 때에, 비로서 온전한 차세대 웹 기술을 확실하게 가시화 할 수 있을 것이다. 그림8에서 보는 바와 같이, 웹에는 여러 학문 분야가 직간접으로 혼재하여 있다.

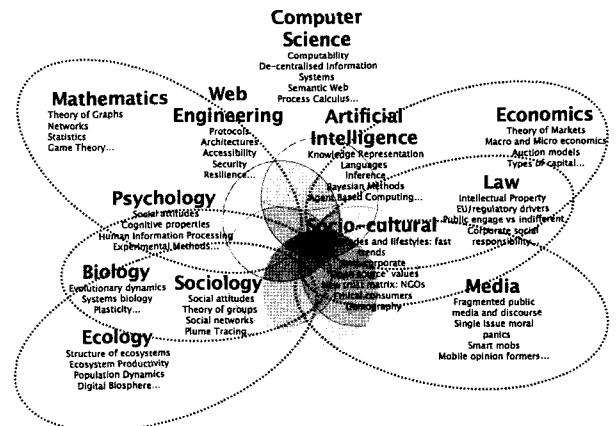


그림 8 학제적인 Web science

웹은 표준화된 정보통신 기술에 의해 생성된 가상 공간이다. 웹 공간에서는 Second Life처럼 현실공간의 인간 상호작용이 그대로 실현된다. 때문에, 웹 공간은 현실세계에 발생하는 제반사항을 포함하고 있으며, 가상세계로서의 특성까지 내포한 복잡한 공간이다. 그럼에도 불구하고, 웹 공간에 대한 다각적 이해와 연구가 체계적으로 수행되지 못하였다. 웹의 지속적인 발전을 위해서는 웹과 관련된 제반 요소를 학문적 관점에서 체계화할 필요가 있다. Web Science가 이러한 목적을 효과적으로 달성하기 위해 태동되고 있다[9,10]. 뿐만 아니라, Web Science에서는 차세대 웹 기술이 갖추어야 할 제반 특성을 제시하고 분석하여, 기술적, 사회적 장애를 극복하고 인류를 위한 강력한 새로운 차세대 웹을 실현하고자 한다. Web Science는 웹 기술과 현상을 이해하고 연구하는 토대 학문의 역할과 차세대 웹 기술 개발의 선도자가 될 것이다.

6. 결론

현대의 일상생활에서 웹이 차지하는 비중이 높아짐에 따라, 차세대 웹 기술에 대한 관심이 고조되고 있다. 웹은 인간과 컴퓨터가 함께 공존하고 생활하는 공간을 조성하여 새로운 지식정보사회 발전의 모태가 되고 있다. 본고에서는 차세대 웹 기술을 견인하고 있는 웹2.0과 시맨틱웹의 개념과 의의를 고찰하였다. 웹2.0은 그동안 웹 기술이 소홀하게 다루어 왔고 잊고 있었던 사용자의 중요성을 재발견하고, 이를 새로이 부각시켜 인간 중심의 웹 기술로 나아가는데 큰 기여를 하고 있다. 시맨틱웹은 컴퓨터가 정보자원의 의미를 이해하고 스스로 처리할 수 있도록 하여 보다 차원 높은 지식 정보 서비스 실현에 노력하고 있다. 웹2.0과 시맨틱웹은 차세대 웹을 주도하기 위한 경쟁자가 아니라, 상호적 협력이 필요한 동반자이다. 웹2.0과 시맨틱웹은 정보자원의 개방과 공유를 통하여, 제약과 장애가 없는 지능화된 지식정보 서비스를 실현하고, 인간과 컴퓨터가 함께 조화되는 협력 체제를 구현하기 위한 기술이다.

웹 기술은 여러 정보 기술을 융합하는 플랫폼으로의 역할을 하고 있기 때문에, 학계와 산업계는 차세대 웹기술을 선도할 수 있도록, 보다 긴밀한 연구 개

발의 공조체제를 구축하여야 할 것이다. 웹기술은 인터넷 기반 기술이기 때문에 글로벌 관점의 연구 협력 체제도 무엇보다 중요하다. 연구실의 제한 공간을 넘어, 연구 자료와 결과를 함께 공유하고, 글로벌 프로젝트에 참여하고 협력하는 차세대 연구 개발자들의 출현을 기대해 본다.

참고문헌

- [1] Tim O'Reilly, What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, Sept., 2005 (<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>)
- [2] Best Web2.0 Software, http://web2.wsj.com/the_best_web_20_software_of_2006.htm
- [3] Maness, J., Library 2.0 Theory: Web 2.0 and Its Implications for Libraries. *Webology*, 3(2) 2006, (<http://www.webology.ir/2006/v3n2/a25.html>)
- [4] Tim Berners-Lee, James Hendler, Ora Lassila The Semantic Web, *Scientific American*, May 2001
- [5] Nigel Shadbolt, Wendy Hall, Tim Berners-Lee, The Semantic Web Revisited, *IEEE Intelligent Systems*, 2007
- [6] Semantic Web Best Practices and Deployment Working Group, <http://www.w3.org/2001/sw/BestPractices/>
- [7] James Hendler, Ora Lassila, SemWeb@5: Current Status and Future Promise of the Semantic Web, *Semantic Technology Conference*, March, 2006
- [8] Anupriya Ankolekar et al, The TwoCultures: Mashing up Web 2.0 and the Semantic Web, *WWW2007*, May, 2007
- [9] Tim Berners-Lee, Wendy Hall, James Hendler, Nigel Shadbolt, Daniel J. Weitzner, *Creating a Science of the Web*, Science, August, 2006
- [10] Tim Berners-Lee et al, A Framework for Web Science, *Foundations and Trends in Web Science*, Vol.1, No.1, 2006
- [11] 전양승 외, 시맨틱 웹 서비스 기술 연구 동향, 정보과학회지, 2006년 4월
- [12] 전종홍 외, 시맨틱웹, *TTA Journal*, No 107, 2006년, 10월



한성국

1979 인하대학교 전자공학과(공학박사)
1984~현재 원광대학교 컴퓨터공학부 교수
1989 University of Pennsylvania 방문교수
2003 DERI 방문교수
2004~현재 대한전자공학회 컴퓨터소사이터티 감사
2005~현재 한국정보과학회 부회장

관심분야: 시맨틱 웹서비스, 온톨로지 공학, 웹서비스, 의료정보, e-Learning

E-mail : skhan@wonkwang.ac.kr



정영식

1993 고려대학교 전산학(박사)
1993~현재 원광대학교 컴퓨터공학부 교수
1997 미시간 주립대학교 전산학과 객원교수
2004 웨인 주립대학교 컴퓨터공학과 객원교수

관심분야: 그리드컴퓨팅, LBS, 분산병렬처리

E-mail : ysjeong@wonkwang.ac.kr



유재규

2007 원광대학교 전기전자 및 정보공학부(공학사)
2007~현재 원광대학교 컴퓨터공학과(석사과정)
관심분야: 시맨틱 웹서비스, 온톨로지 공학, 웹 2.0, 웹서비스

E-mail : jkyoo82@wonkwang.ac.kr

**2nd Korean - Japan Joint Workshop on
Pattern Recognition(KJPR2007)**

- 일 자 : 2007년 10월 25~26일
- 장 소 : 일본
- 내 용 : 논문발표 등
- 주 최 : 컴퓨터비전 및 패턴인식연구회, IEICE PRMU
- 상세안내 : <http://www.image.esys.tsukuba.ac.jp/kjpr2007/>