

# 웹3.0 정보 서비스의 사례와 시사점

■ 박 윤 정\*

개방과 공유, 참여를 통해 이용자를 서로 연결해주던 웹2.0 시대와 달리, 구축되어 있는 막대한 정보 속에서 개인에게 최적화된 정보를 제공해주는 웹3.0 시대의 도래가 임박해 있다. 웹3.0의 핵심은 시맨틱(semantic) 웹으로 메타데이터(metadata)와 온톨로지(ontology) 기술을 기반으로 정보를 서로 연결하여 컴퓨터도 컨텍스트(context)를 이해하게 됨으로써, 검색어의 의미에 맞는 답변을 제공하게 된다. 지금처럼 정보의 양이 방대해지고 파편화될수록(fragmented) 시맨틱 웹의 필요성이 더욱 강조될 것으로 예상된다.

세계 곳곳에서 웹3.0에 대한 논의가 활발해지고 있다. 민간 기업이 관심을 가지고 기술개발을 하는 것은 물론이고, 각국의 정부에서도 공공기관을 중심으로 시범 서비스를 실시하는 추세이다. 본 고에서는 지난 5월 유럽의 최대 시맨틱 웹 컨퍼런스 중 하나인 ESWC에서 발표된 핀란드의 온톨로지 기반 도서관 서비스 ONKI와 영국의 BBC 정보 서비스 사례를 살펴보고, 국내에서 서비스하고 있는 Ontoframe 적용 사례를 통해 시맨틱 웹 기반 서비스가 바람직한 웹3.0 시대로 나아가기 위한 방향을 고찰해 보고자 한다.

## 목 차

- I. 서 론 / 25
- II. 웹3.0의 개념과 특징 / 26
  - 1. 웹2.0의 한계와 웹3.0의 등장 / 26
  - 2. 웹3.0의 키워드: 시맨틱 웹 / 31
- III. 시맨틱 웹 기반 정보 서비스 / 33
  - 1. ONKI / 33
  - 2. BBC / 35
  - 3. Ontoframe 적용 사례 / 37
- IV. 결 론 / 40

## I. 서 론

2010년 미래의 인터넷은 어떤 변화를 겪게 될 것인가? 최근 웹3.0에 대한 논의가 점점 구체화되고 있다. 시맨틱(semantic) 검색, 실시간 웹(real-time web), 클라우드 컴퓨팅(cloud computing), 모바일(mobile) 등과 같이 웹3.0을 설명할 키워

\* 정보통신정책연구원 미래융합연구실 연구원, (02)570-4088, yjpark@kisdi.re.kr

드들도 두각을 드러내고 있다. 그 중에서도 웹3.0의 가장 핵심은 시맨틱 검색으로 웹 3.0의 특징을 형성하는 결정적인 역할을 한다. 일각에서는 웹2.0의 실체조차 명확하지 않을뿐더러, 단지 참여와 공유와 개방이라는 관념적인 성격으로 특징지어진 상황에서 웹3.0에 대한 논의는 시기상조라는 의견이 있다. 또한 웹3.0의 핵심인 시맨틱 웹의 개념이 1998년부터 등장한 만큼 전혀 새로운 개념이 아니라는 의견이 있는 것도 사실이다. 그럼에도 불구하고 웹3.0의 등장이 각광을 받는 이유는 현재 시들해진 웹 2.0에 대한 관심과 웹 산업이 성숙기에 접어들었다는 인식의 확산으로 차별화되는 차세대 웹에 대한 활발한 논의를 재촉하고 있고(이경일, 2008a), 그동안 지나치게 장밋빛으로만 보이던 미래 인터넷 환경의 기술이 최근 수년간 놀라운 기술 발전을 하면서 현실로 다가오고 있기 때문이다.

인터넷을 통한 정보가 급격히 증가하고 이에 대한 통제 능력이 상실되기 시작한 현실에서 정보 조직화에 대한 요구는 점차 커지고 있으며, 요구자 또한 전문가에서 일반인으로 확산되고 있다(정한민 외, 2008). 이러한 상황에서 시맨틱 웹의 발전은 많은 가능성을 제공한다.

본 고에서는 최근 인터넷 산업에서 이슈로 부상하고 있는 웹3.0의 등장배경과 개념에 대해 살펴보고, 웹3.0의 핵심인 시맨틱 웹의 기술이 적용된 서비스 사례를 통해 미래 인터넷 시대에 대비하기 위한 정책 방향의 시사점을 도출해 보고자 한다.

## II. 웹3.0의 개념과 특징

### 1. 웹2.0의 한계와 웹3.0의 등장

2000년 초반, 인터넷의 사용이 확산되면서 누구나 인터넷을 사용할 수 있게 됨에 따라 웹에는 새로운 문화의 지평이 열리게 되었다. 새로운 패러다임이 도래한 것이다. 학자들마다 조금의 차이는 있지만 일반적으로 2000년부터 2010년까지는 웹2.0의 시대가 될 것이라고 전망하고 있다. 웹2.0시대라고 새롭게 규정할 수 있었던 이유는 웹상의 텍스트와 문서가 기존의 공급자 전달 방식이 아니라 문서, 사진, 동영상 등의 멀

미디어를 이용해 사용자 간 개방과 공유와 참여가 이루어질 수 있었기 때문이다. 웹2.0을 대표하는 서비스인 블로그는 누구나 쉽게 인터넷 상에 자신만의 공간을 형성하고 공유할 수 있게 했으며, 2001년 소위 ‘모두의 백과사전’ 위키피디아가 등장하자 소수 엘리트 집단으로부터 정의되어 대중에게 전해지던 지식과 지성은 형성 과정과 관점이 완전히 바뀌게 되었다. 많은 사람들이 힘을 보탬수록 지식이 완벽해진다는 것이다. 2005년 시작된 동영상 서비스 사이트인 유튜브 역시 자유로운 동영상 촬영 및 편집의 사용으로 웹이 범세계적 집단 미디어가 되는 데 기여했다. 웹2.0의 시대는 말 그대로 창작과 소비의 시대이다. 인터넷을 통해 사회 문화와 경제뿐 아니라 정치, 인간관계까지도 일일이 영향을 받고 있는 현재에도, 여전히 우리는 웹2.0 시대에 살고 있으며 다양한 현상을 경험하고, 사회 변화를 거치고 있다.

문제는 이런 현상을 통해 우리가 공유하는 정보는 기하급수적으로 늘어나고, 네트워크는 복잡해지고 있다는 것이다. 세계 언어별로 인터넷 사용자 현황을 보면 2000년부터 2009년까지 채 10년이 되기도 전에 1,000% 이상 사용자가 늘어난 곳도 있다 ([그림 1] 참조).

[그림 1] 언어별 인터넷 사용자 현황

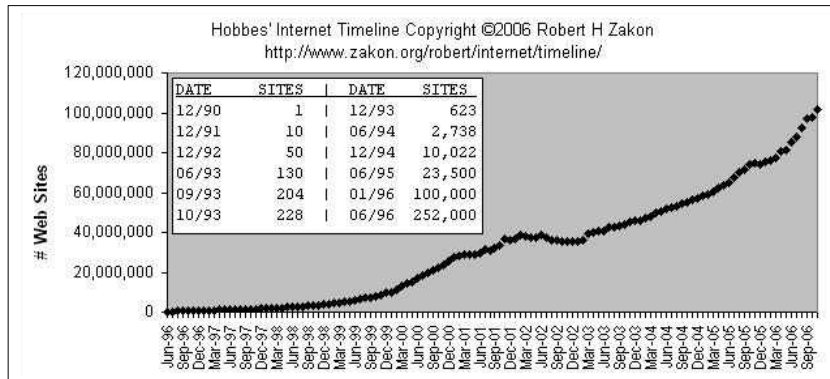
Top Ten Languages Used in the Web ( Number of Internet Users by Language )					
TOP TEN LANGUAGES IN THE INTERNET	Internet Users by Language	Internet Penetration by Language	Growth in Internet ( 2000 - 2009 )	Internet Users % of Total	World Population for this Language (2009 Estimate)
English	478,717,443	37.9 %	237.2 %	28.7 %	1,263,830,976
Chinese	361,364,613	26.3 %	1,018.7 %	21.7 %	1,373,859,774
Spanish	132,963,898	32.3 %	631.3 %	8.0 %	411,631,985
Japanese	94,000,000	74.0 %	99.7 %	5.6 %	127,078,679
French	76,915,917	18.1 %	530.5 %	4.6 %	425,622,855
Portuguese	73,027,400	29.5 %	863.9 %	4.4 %	247,223,493
German	65,243,673	67.7 %	135.5 %	3.9 %	96,389,702
Arabic	49,372,400	17.0 %	1,862.2 %	3.0 %	289,742,641
Russian	38,000,000	27.1 %	1,125.8 %	2.3 %	140,041,247
Korean	37,475,800	52.7 %	96.8 %	2.2 %	71,174,317
TOP 10 LANGUAGES	1,407,081,144	31.6 %	351.5 %	84.3 %	4,446,595,669
Rest of the Languages	261,789,264	11.3 %	430.7 %	15.7 %	2,321,209,539
WORLD TOTAL	1,668,870,408	24.7 %	362.3 %	100.0 %	6,767,805,208

(\*) NOTES: (1) Top Ten Languages Internet Stats were updated for June 30, 2009. (2) Internet Penetration is the ratio between the sum of Internet users speaking a language and the total population estimate that speaks that specific language. (3) The most recent Internet usage information comes from data published by Nielsen Online, International Telecommunications Union, GfK, and other reliable sources. (4) World population information comes from the U.S. Census Bureau. (5) For definitions and navigation help in several languages, see the Site Surfing Guide. (6) Stats may be cited, stating the source and establishing an active link back to Internet World Stats. Copyright © 2009, Miniwatts Marketing Group. All rights reserved worldwide.

자료: internet world state(2009)

사용자가 증가함에 따라 인터넷 사이트와 블로그의 숫자 역시 가파른 증가 추세를 보이는데 조만간 인간이 셀 수 있는 숫자의 범위를 넘어설지 모른다(Bratt(2009)[그림 2] 참조).

[그림 2] Hobbes의 인터넷 타임라인



자료: Bratt(2009)에서 재인용

폭풍처럼 휘몰아치는 정보의 바다에서 어떻게 목적지를 찾아갈 것인가? 원하는 정보를 찾겠다면서 구글의 20번째 페이지까지 검색해본 경우가 종종 있을 것이다. 바로 여기에 웹2.0의 한계가 있다. 정보가 점점 더 방대해지고 파편화되어 통합적인 시각을 갖고 현상을 관찰하기가 쉽지 않다는 것이다. 과거 산업사회에서 정보사회로 진화하는 것이 “지식의 정보화”와 공유를 통해서였다면, 다가오는 지식사회에서는 넘쳐나는 “정보를 지식화”하는 것이 중요해진다(이경일, 2009). 자연스럽게 웹2.0의 한계를 극복하는 웹3.0에 대한 논의가 활발하게 진행되고 있다. 웹2.0 시대의 검색(search)으로 만족할 것이 아니라 웹3.0의 발견(discovery)이 필요한 것이다. 따라서 웹2.0과 웹3.0은 그 태생에서부터 차이가 있다. 웹2.0의 참여와 개방과 공유라는 성격이 데이터와 정보 중심의 상호작용에 의해 비롯되는 사회적, 경제적 조류를 지칭하는 사후적 성격이라면, 웹3.0은 지식과 네트워크 중심의 데이터와 정보를 주문제작(customization)하는 개인화 과정으로 지능화된 웹 환경을 구축하기 위한 기술적 측면을 강조하는 사

전적 조류라고 할 수 있다(김택천, 2009a).

웹3.0이란 용어는 2006년 뉴욕타임즈의 John Markoff 기자가 처음 사용한 이후 논쟁의 중심에 서게 되었다.<sup>1)</sup> 웹3.0은 웹 발전 방향의 흐름을 지칭하는 것일 뿐 아직 명확히 개념화 되어 있지는 않다. 다만 웹3.0을 선도하는 기술들의 특징을 보았을 때, ‘개인화’와 ‘지능화’, ‘상황인식’ 등으로 의견이 수렴되고 있어, 웹의 진화방향을 예측할 수 있다. 웹3.0이란 지능화된 웹이 시맨틱(semantic) 기술을 이용해서 상황인식을 통해 이용자에게 맞춤형 콘텐츠 및 서비스를 제공하는 것이기 때문이다. 즉 웹3.0 시대에는 지능형 웹이 이용자가 원하는 정보, 직관적인 경험을 제공하게 되며, 웹2.0 시대와는 <표 1>과 같은 차이가 있다.

<표 1> 웹2.0과 웹3.0의 비교

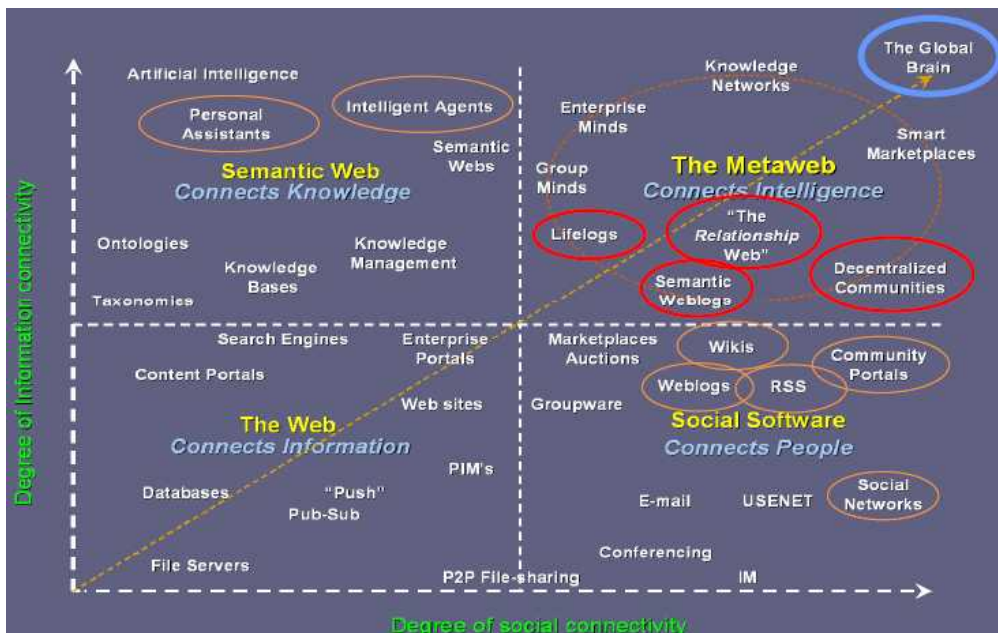
구분	웹2.0	웹3.0
시기	2000~2010	2010~2020
상호작용	read & write	read, write & execute
키워드	참여, 공유, 개방	상황인식(context)
정보 이용자	인간	인간, 기계
정보권력	대형화, 집중화	분산(필요한 정보만 선별)
콘텐츠 이용행태	이용자가 생산, 유통, 소비 모두를 담당	지능화된 웹이 원하는 콘텐츠 제공 → 개인별 맞춤
검색	여러 사이트에 있는 자료 개방 (open API)	이용자 맞춤형 검색
기반 기술	브로드밴드, 서버관리	시맨틱 기술, 상황인식, 클라우드 컴퓨팅
대응 단말	PC와 모바일 일부	PC, 모바일, 시계와 같은 악세서리 등 다양

자료: 김택천(2009b), Bratt(2009), 스트라베이스(2009)를 바탕으로 재구성

1) <http://www.nytimes.com/2006/11/12/business/12web.html>

Nova Spivack(2003)은 정보적 연결성과 사회적 연결성을 기준으로 웹의 발전 방향을 제시한 바 있는데([그림 3] 참조), 소셜 소프트웨어(Social Software)로 대표되는 웹2.0은 참여와 공유와 개방으로 사람들을 연결하여 사회적 연결성(social connectivity)은 높지만, 공유된 정보를 연결하는 정보적 연결성(informational connectivity)은 낮다. 반면 Semantic Web이 주를 이루는 웹3.0 시대는 지식기반 사회에서 시맨틱 웹으로 최적화된 정보를 연결하기 때문에 사회적 연결성은 낮지만 정보적 연결성은 높다. 궁극적으로는 정보와 사람을 모두 긴밀하게 연결하는 Metaweb 혹은 유비쿼터스 웹의 구현이 가능한 시대가 도래할 것이라고 주장했다. 따라서 가장 바람직한 웹의 진화 방향은 정보적 연결성이나 사회적 연결성이 어느 한쪽으로 편중되어 있는 것이 아니라 시맨틱 웹과 소셜 소프트웨어가 동시에 구현되는 환경이라고 하겠다.

[그림 3] Nova Spivack의 Metaweb의 개념



출처: 이선재(2009)에서 재인용

## 2. 웹3.0의 키워드: 시맨틱 웹

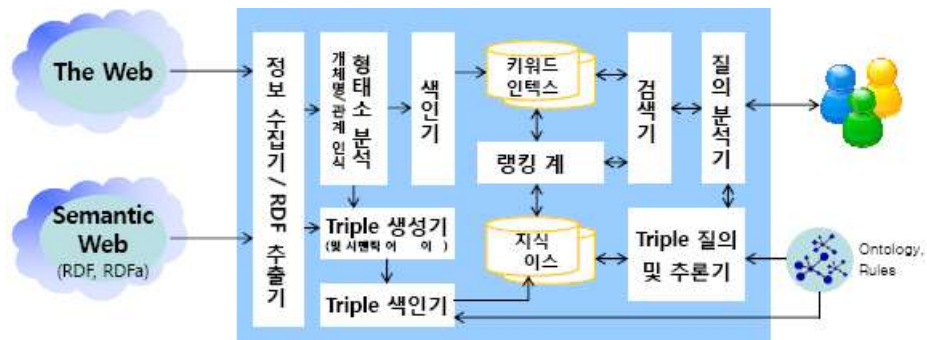
웹3.0에 대한 다양한 키워드가 떠오르고 있지만 가장 핵심적으로 시맨틱 웹이 대두되는 이유는 시맨틱 웹이 파편화되고 방대해진 정보를 서로 연결해 웹2.0의 한계를 가장 잘 극복하기 때문일 것이다. Tim Berner-Lee(1998)는 컴퓨터가 단어, 문장의 뜻을 이해하고 논리적인 추론까지 하는 차세대 웹 기술로 시맨틱 웹을 언급한 바 있다. 시맨틱 웹의 원리는 사람이 이해할 수 있도록 자연어 위주로 되어 있는 현재의 웹문서와 달리, 정보자원들 사이에 연결되어 있는 의미를 컴퓨터가 이해할 수 있는 형태의 언어로 바꾸는 것으로, 컴퓨터가 정보자원의 뜻을 해석하고 기계들끼리 서로 정보를 주고받으면서 자체적으로 필요한 일을 처리하는 것이 가능해 진다(NIDA, 2006). 시맨틱 웹의 주요 역할은 검색 매커니즘의 개선이지만 웹의 정보들을 서로 연결시켜 이용할 수 있게 해준다는 것에 더 큰 잠재성이 있으며, 자동으로 정보를 처리하는 어플리케이션의 진보를 가져올 것이다(Razmerita & Gouarderes, 2005).

시맨틱 웹의 구현을 이해하기 위해서는 기반 기술인 온톨로지(ontology)와 메타데이터(metadata)를 살펴볼 필요가 있다. 온톨로지는 단어와 관계들로 구성된 일종의 사전으로 생각할 수 있으며, 그 속에는 특정 도메인에 관련된 단어들이 계층적으로 표현되어 있고, 추가적으로 이를 확장할 수 있는 추론 규칙이 포함되어 있어 웹 기반의 지식 처리나 응용 프로그램 사이의 지식 공유, 재사용 등이 가능하도록 되어 있다. 온톨로지는 시맨틱 웹 응용의 가장 중심적 개념으로서, 이를 표현하기 위해 스키마와 구문 구조 등을 정의한 언어가 온톨로지 언어(ontology language)이며, 현재 DSML+OIL, OWL, Ontolingun 등이 있다.<sup>2)</sup> 이 온톨로지가 구조화 되어 있는 데이터가 ‘데이터를 위한 데이터’, 즉 ‘메타데이터’이다. 온라인 백과사전 위키피디아의 정의에 의하면 구조화된 정보를 분석, 분류하고 부가적 정보를 추가하기 위해 함께 따라가는 정보라고 할 수 있다.

이러한 기술을 기반으로 웹3.0 시대에는 축적된 개인정보와 검색 성향 등을 분석하

2) KERIS 의미기반검색 서비스 “시스템 소개” 중 발췌

### 〔그림 4〕 시맨틱 웹의 이해



자료: 이경일(2008b)

시맨틱 웹을 중심으로 하는 웹3.0의 현재 발단은 미약하나 이로 인해 발생할 변화의 양상은 사뭇 다양할 것으로 예상된다. 일단 웹2.0 시대에는 집단을 중심으로 한 사회적 참여와 공유를 중시하는 반면, 시맨틱 웹으로 구현되는 웹3.0은 검색과 콘텐츠 이용에 있어 공유한 정보를 기반으로 개인에게 최적화된 지식과 정보가 제공될 것이다. 웹2.0의 참여, 공유, 개방을 통해 ‘보다 빨리’, ‘보다 많은’ 정보를 얻는 것이 목적이었다면 웹3.0은 ‘개인’에 초점을 맞춰 사용자가 원하는 직관적이고 개인적인 정보를 제공하게 된다(김택천, 2009b). 이렇게 개인화된 웹을 이용하면 ‘나는 커서 무엇이 될까?’ 또는 ‘내일은 무엇을 할까?’와 같은 질문에도 검색이 사용자에게 적합한 답변



을 해 줄 수 있게 되며, 실제로 구글의 CEO Eric Schmidt는 이러한 질문에 대답하게 하는 개인화된 검색을 개발하는 것이 구글의 다음 목표라고 밝힌 바 있다.<sup>3)</sup>

검색의 개인화가 실현되면 엔터테인먼트, 여행, 라이프스타일, 전문분야 등 취향과 상황에 따라 차이가 큰 콘텐츠 분야에서 웹3.0이 큰 반향을 일으킬 것으로 예상된다. 다음 장에서는 실제 상기 분야에서 서비스되고 있는 시맨틱 웹 사례를 살펴보도록 하겠다. 사례는 지난 5월 유럽의 시맨틱 웹 컨퍼런스(6th European Semantic Web Conference 2009)에서 발제된 핀란드의 온톨로지 도서관 서비스 ONKI와 영국 국영 방송 BBC의 사례, 국내의 Ontoframe 적용 사례를 통해 살펴보고 시맨틱 웹 기반 서비스가 정보화 시대의 혜택과 공익을 제공하기 위해서 어떻게 미래시대를 준비해야 할 지 함께 고민해 보고자 한다.

### Ⅲ. 시맨틱 웹 기반 정보 서비스

#### 1. ONKI

생활문화 콘텐츠로서 핀란드의 도서관 서비스 ONKI는 다양한 분야의 정보를 도메인 연결을 통해 통합적으로 제시하고 있을 뿐 아니라, 2008년 9월부터 3개월 동안 36,000명의 방문자와 104,000번의 방문을 기록하는 등 이용자들의 참여율도 높아 시맨틱 웹 기반 정보 서비스의 모범 사례라고 할 수 있다.

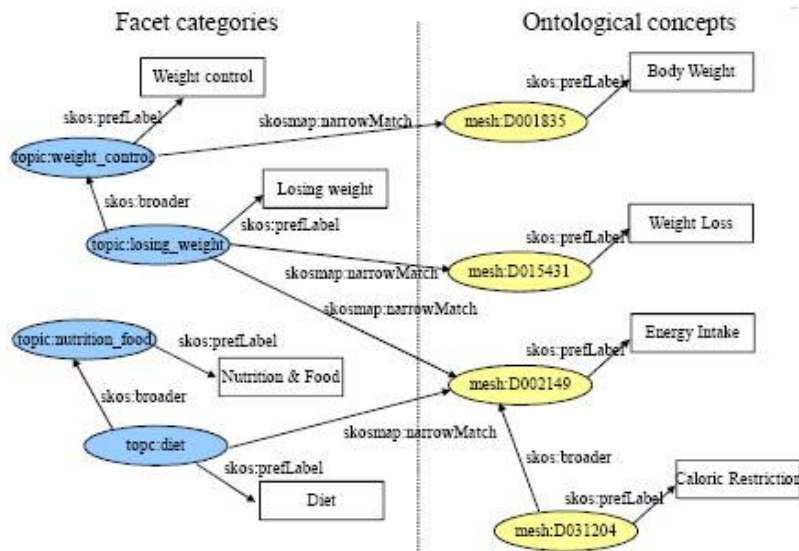
ONKI는 시맨틱 웹 온톨로지 인프라 스트럭처를 국가적인 차원에서 개발하려는 핀란드의 National Semantic Web Ontology project FinnONTO(2003~2010)의 주요 프로젝트 중 하나이다. 최대 규모의 온톨로지 공공 도서관 시스템 ONKI는 이전에 구현된 적이 없는 새로운 서비스로 ONKI의 시맨틱 웹은 온톨로지를 기반으로 웹상의 콘텐츠와 서비스를 메타데이터에 연결해 명확하고 이해하기 쉽도록 제공한다. 현재 약 40여개의 도메인을 가지고 역사, 문화, 자연과학, 의학 정보, 경영 정보, 지리 정보

3) 파이낸셜뉴스(2007. 5. 23)

등의 분야에서 서비스되고 있으며 대부분 무료로 누구나 이용할 수 있는 어플리케이션을 제공하고 있다.

예를 들어, ONKI의 주요 시범 어플리케이션 중 하나인 HealthFinland는 시민들의 건강관리를 촉진시키기 위해 만들어진 포털로 다이어트, 운동, 금연, 건강 생활 등에 관한 콘텐츠를 온톨로지와 메타데이터를 이용하여 하나의 포털에 집합적인 콘텐츠를 제공한다(Hyvonen 외, 2007). 모든 카테고리는 라벨과 위계로 위치가 결정된 하위 개념을 가지고 있다. 검색어가 입력되면 등록된 라벨과 위계에 맞게 콘텐츠를 좁게 일치시켜 제시한다. [그림 5]와 같이 ‘다이어트’에 대해 검색을 하면 온톨로지 맵을 따라 ‘다이어트’에 관한 것 뿐 아니라 관련 개념인 ‘몸무게 조절’, ‘체중 감량’, ‘식이요법’, ‘에너지 섭취’ 등의 하위 개념이 나타나게 된다.

[그림 5] 온톨로지 맵



자료: Hyvonen 외(2007)

ONKI가 제공하는 대부분의 온톨로지는 현존하는 백과사전에 기반하며 만들어졌지만, 사용자가 직접 참여하는 매우 진화한 형태를 보이기도 한다. ONKI Selector는

사용자가 작성한 주석 등의 메타데이터와 해당 개념을 바로 연결할 수 있도록 ‘Fetch button’을 가지고 있다. 이는 시맨틱 웹이 웹2.0의 집단 지성과 결합되는 양상으로, Nova Spivack의 개념에 따르면 이용자를 사회적으로 연결하는 사회적 연결성(social connectivity)과 정보를 연결하는 정보적 연결성(informational connectivity)이 모두 높은, 미래의 웹이 나아가야 하는 바람직한 방향이라 하겠다.

## 2. BBC

개인 취향의 차이가 큰 오락문화 콘텐츠로서 BBC 영국 방송국은 텍스트, 오디오, 비디오를 포함해 엄청난 양의 온라인 콘텐츠를 제공하고 있다. 여느 페이지와 마찬가지로 BBC 역시 도메인에 의해 음식, 음악, 뉴스 등의 마이크로 사이트<sup>4)</sup>로 나뉘져 있었다. 마이크로 사이트의 개발로 인해 BBC는 대중적인 서비스를 생산하고 있지만 분리된 이용의 한계 때문에 그 사이트들이 가진 최대 잠재력을 끌어낼 수는 없었다. 콘텐츠를 서로 연결하는데 실패함으로써 제공된 콘텐츠는 시너지 효과를 창출할 수 없기 때문이다. 사용자 입장에서 BBC가 제공하는 모든 정보를 찾아 이용하기는 힘들다. 예를 들어, 관심 가수에 대한 음악 페이지에서 그 가수가 출연했던 프로그램 페이지로 이동하는 것은 불가능했다.

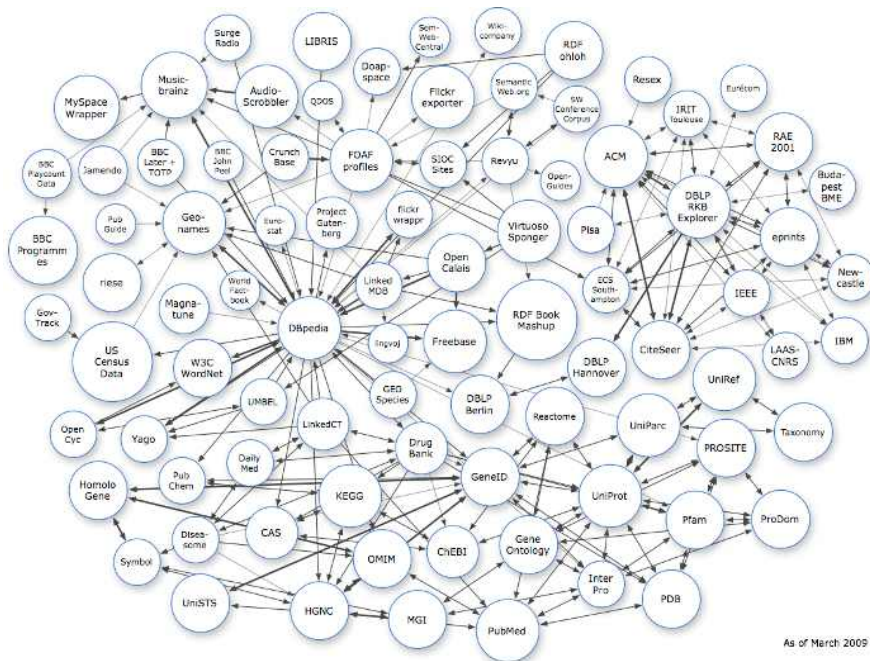
이러한 문제를 해결하기 위해 BBC 프로그램과 BBC 뮤직 사이트는 출연진, 장소, 주제, 프로그램 등 특정한 관계가 있는 것들을 서로 연결하는 서비스를 시도하고자 Wikipedia처럼 데이터를 공동으로 올리고 편집할 수 있는 DBpedia를 적용하였다. BBC 웹사이트 사용자라면 누구나 DBpedia에서 메타데이터를 편집하고 수정할 수 있다. DBpedia는 단순히 연결된 데이터의 URI를 제공하는 것 뿐만 아니라 자동 연결 알고리즘의 핵심인 개념과 개념 간 관계에 따라 구조화된 데이터를 제공해 데이터의

4) 온라인 백과사전 위키피디아에 의하면, 마이크로 사이트는 커다란 웹사이트의 일부로서 독립적으로 선전하는 것이다. 마이크로 사이트는 개별적인 목적들에 맞게 설계되며, 홈페이지를 위한 개별적인 웹주소를 갖는다. 일반적으로, 마이크로 사이트는 같은 웹서버 상에 존재하며, 관련된 대형 사이트의 전반적인 시각 디자인과 로고를 포함한다.

활용성을 높인다. 이렇게 해서 DBpedia에는 BBC 도메인이 점점 모이고 있다([그림 6] 참조).

현재 BBC 뮤직 사이트에서는 BBC 라디오에 출연한 아티스트들에 관한 정보를 메타데이터에 주석으로 연결하여 음악팬들이 미처 알지 못했던 현재 밴드, 과거 소속사, 공연 현장, 아티스트 간 관계 등 추가적인 정보를 얻을 수 있다.<sup>5)</sup> DBpedia 상에서 서로 연결된 정보를 탐색하는 것은 Topic page와 Navigation Badge라는 새로운 인터페이스를 통해서 이용할 수 있다. Topic page에서는 주어진 인물, 장소, 주제에서부터 BBC의 다른 콘텐츠 영역으로 이동할 수 있고 Navigation Badge는 콘텐츠 페이지에서 관련된 통합 페이지로의 링크를 만들어 이동하는 기능을 제공한다.

〔그림 6〕 BBC를 포함한 DBpedia



자료: <http://linkeddata.org>

5) ReadWriteWeb(2009. 1. 21)

거미줄처럼 연결된 데이터들로부터 얼마나 재미있는 서비스가 많이 나오게 될까? 연결된 데이터에서 나올 수 있는 어플리케이션은 화살표 숫자만큼 많을 것이다. 재미를 찾는 욕구란 끝이 없다. 새로운 서비스를 계속 제공은 하되 더 많은 개별 선택, 더 많은 자기표현, 더 많은 자기 정체성 만들기가 가능해야 하고, 중국에는 이런 서비스를 다른 사람들과 함께 즐길 수 있어야 한다(이선재, 2009). 더 많은 선택을 통해 더 많은 자기 정체성 만들기라는 측면에서 오락문화 콘텐츠는 시맨틱 웹 활용도가 매우 높은 콘텐츠이다. BBC의 사례는 웹3.0 시대를 맞아 시맨틱 웹이 새롭게 제공할 수 있는 서비스에 대한 통찰력을 제공해 주고 있다.

### 3. Ontoframe 적용 사례

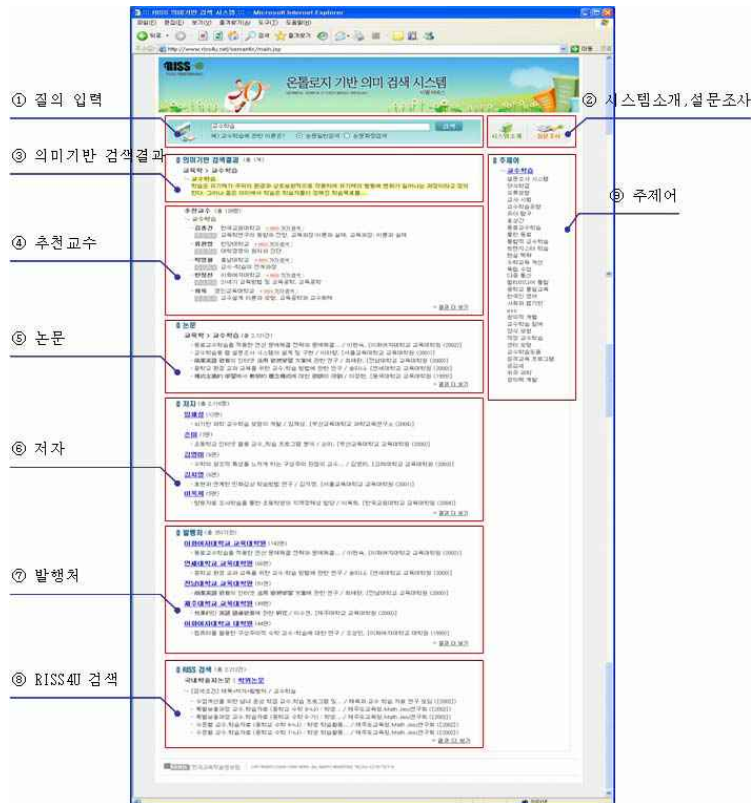
우리나라에서도 시맨틱 웹을 이용한 정보 서비스를 실시하고 있다. 한국교육학술정보원(KERIS)이 제공하는 학술연구정보서비스, 지식경제부 국가표준인증종합정보센터, 법무부가 2007년부터 시맨틱 웹 서비스 프레임워크인 Ontoframe을 적용하고 있다. Ontoframe은 정보를 검색하고 내재된 지식을 발견할 수 있도록 도와주는 시맨틱 서비스 플랫폼으로 한국과학기술정보연구원(KISTI)이 2005년부터 개발해 왔다. 검색어의 유형에 따라 동적으로 검색결과 화면을 구성하는 특허 기술이 적용되어 있다.<sup>6)</sup> 단순 기능을 넘어 R&D 수행 중에 필요한 다양한 고부가가치 정보를 얻을 수 있도록 해주는 지능형 서비스 제공을 목표로, 특히 정확한 정보 제공을 위한 개체 식별 및 관리 기술과 정보 간 연계관계 탐색을 위한 대용량 고속 추론 기술을 핵심 기술을 통해 연구자들의 정보 획득 시간을 획기적으로 단축하고자 노력하고 있다(정한민, 2008).

학술연구정보서비스의 검색은 현재 교육학 분야에 제한적으로 적용되고 있다. 사용자 질의 분석을 통해서 KERIS에서 작성한 교육학 분야의 온톨로지와 관련이 있는 정보를 찾고, 그 정보를 이용하여 KERIS에서 정의한 의미기반 검색결과, 추천교수, 논문, 저자, 발행처, RISS 검색 결과 및 주제어(클러스터링 결과) 정보를 통합하여 보여

6) [http://ontoframe.kr/2008/2008\\_new/main.jsp](http://ontoframe.kr/2008/2008_new/main.jsp)

준다. 물론 단순히 여러 정보를 보여주는 것이 아니라, KERIS 교육학 분야의 온톨로지에서 구축된 의미 정보를 사용자에게 통합하여 보여 준다.<sup>7)</sup> 특히 우측 상단의 온톨로지 분류를 통해 검색하고자 하는 질의어의 의미정보를 포함하고 있는 클래스의 위계정보를 보여주고, 질의어에 대한 의미정보를 보여주는 결과는 매우 유용하다(그림 7) 참조). 예를 들어 “가상 교육”으로 검색을 하면, 기존의 검색에서는 “가상”과 “교육”이 포함된 모든 문서를 보여줬지만, 의미 기반 검색을 이용하면 개념 정의를 비롯해 저명한 추천 교수, 저서 및 논문을 한 번에 볼 수 있다.

[그림 7] 학술연구정보서비스의 의미 검색 시스템

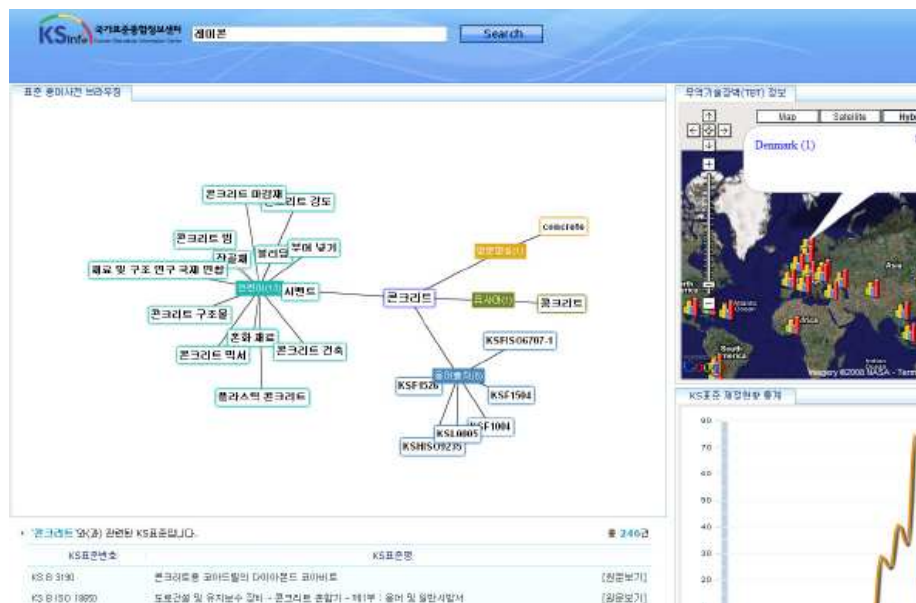


자료: KERIS 의미기반검색 서비스 “시스템 소개”

7) KERIS 의미기반검색 서비스 “시스템 소개”

국가표준인증종합정보센터의 서비스는 각종 제품이나 서비스 제공시 달성해야 하는 기준에 대한 정보를 제공한다. 따라서 콘텐츠 특성상 기업의 부담을 경감하고 국민의 안전 확보에 큰 역할을 한다. 그럼에도 불구하고 전문적이거나 표준 용어가 아닌 경우 적절한 검색결과를 제공하지 못하는 결점을 가지고 있었다. 예를 들어, 레미콘은 레디믹스트콘크리트의 비표준 용어인데 일반 키워드 검색으로는 레미콘의 국가표준을 확인할 수 없었지만 시맨틱 검색이 도입되고 난 후, 대응 표준어와 동일 검색 서비스를 제공하고, 검색 결과를 상황에 따라 유동적으로 재구성하는 방식으로 관련 정보를 집중적으로 제공할 수 있게 되어 검색의 유연성을 확보하였다(정한민, 2008).

[그림 8] 국가표준인증종합정보센터 의미 기반 검색 시스템



자료: 정한민(2008)

우리나라에서 웹3.0 시대를 준비하며 시맨틱 웹의 개발과 적용을 시도하는 것 자체는 매우 바람직한 일이다. 다만 범위가 제한적이고 사용자가 온톨로지 형성에 참여하지 못하는 한계가 있다. 아직까지 웹2.0 서비스에 시맨틱 웹을 적용하는 경우는 활성

화되지 않고 있는 것으로 보인다. 앞의 두 사례와 비교했을 때, ONKI와 BBC가 시맨틱 웹과 인적 커뮤니티를 결합시켜 시너지 효과를 창출하기 위해 노력하는 반면 우리나라는 시맨틱 웹 자체에 집중하는 경향이 있다는 것을 알 수 있다. 현재 시맨틱 웹이 시범서비스로써 공공기관을 중심으로 서비스되고 있기 때문이라고는 하지만, 주도기관이 메타데이터와 온톨로지를 모두 구성한다는 것은 한계가 있고 오류가 발생할 확률이 크므로 이에 대한 대책이 필요하다.

## IV. 결 론

구슬을 서 말 가지고 있어도 꿰어야 보배가 된다고 했다. 웹2.0 시대는 네트워크를 기반으로 정보를 서로 교환하는 참여와 공유, 개방의 문화를 형성해 막대한 정보를 제공하고 있지만 파편화된 정보는 그 진가를 발휘할 수 없다. 웹3.0 시대는 공유된 방대한 정보를 바탕으로, 정보를 서로 연결해 개인화된 정보를 사용자에게 제공하는 한 단계 진화된 서비스를 제공한다.

시맨틱 웹 적용 사례를 살펴본 결과, 우리나라의 시맨틱 웹 서비스와 핀란드, 영국이 제공하는 시맨틱 웹 서비스의 가장 큰 차이점은 메타데이터 상의 온톨로지 형성 과정에 이용자 참여의 여부이다. 핀란드의 ONKI는 온톨로지 만들기를 장려하고, Fatch button을 두어 정보를 서로 연결할 수 있도록 하고 있으며, 영국의 BBC는 DBpedia를 이용해 웹2.0의 집단 지성을 형성하는 방식대로 이용자가 데이터를 수정하고 편집할 수 있도록 참여 가능성을 열어 놓는다. 이는 시맨틱 웹으로 정보 연결성을 높이고, 이용자의 참여를 유도함으로써 사회적 연결성도 동시에 높여 시맨틱 웹 구축과 사용이 선순환하는 효과를 창출하고 있다. 반면 우리나라의 경우 시범적으로 서비스를 시행하는 정부 기관이 해당 사이트에서 시맨틱의 온톨로지를 구성하고 있는데, 이것은 당연히 한계와 오류가 있을 수 밖에 없으며, 이러한 한계를 뛰어넘고 더 먼 인터넷 미래를 준비하기 위해서는 사용자들이 온톨로지 형성에 참여하는 것을 활성화시켜 온톨로지의 형성과 이용이 선순환하도록 만드는 것이 가장 바람직하다.



온톨로지 형성과 이용의 선순환 구조를 만드는 데 가장 큰 문제가 되는 것은 정보를 공유하기를 꺼려하는 우리나라의 문화적 특수성으로 보인다. 웹2.0 시대를 대표하는 위키피디아도 활발한 편집과 의견계재를 통한 정보의 매시 업(mash-up)을 필요로 하지만, 우리나라에서는 저조한 활용성을 보이고 있어 이런 현상이 웹3.0 시대로까지 이어지지 않을지 우려되는 상황이다. 따라서 시맨틱 웹을 이용한 정보 서비스가 원활히 제공되기 위해서는 먼저 시맨틱 웹에 대한 전반적인 인식을 고취시키고, 지식 공유 문화에 대한 인식 전환을 위한 동참 캠페인 실시, 서비스 이용자를 대상으로 온톨로지 형성 기여를 유도하기 위한 정책, 절차, 보상체계 등 대책 마련이 필요하다고 하겠다. 차후의 일이지만 사용자들로 하여금 형성된 온톨로지와 메타데이터의 품질 관리를 위한 작성 기준 제시나 훈련도 꾸준히 실시해야 할 것이다.

웹3.0으로의 전환은 이용자, 사업자, 정부 등 각각의 이해당사자에게 새로운 혜택을 제공할 수 있을 것으로 기대된다. 이용자의 검색 및 콘텐츠 이용에서의 정확도를 높이고 소요시간을 획기적으로 줄이는 한편, 사업자 입장에서는 시맨틱 웹에 기반한 다양한 어플리케이션의 개발로 새로운 시장 개척과 비즈니스 모델 적용이 가능하다. 시범 서비스를 실시하고 기술 개발을 지원하는 정부 입장에서도 개인화된 검색이 다양한 공공정보 콘텐츠를 이용하기 쉽도록 제공할 수 있다는 점에서 공익적인 혜택을 증진시킬 수 있을 것으로 보여, 많은 노력과 관심이 요구된다. 웹3.0 시대가 아직은 미지수로 남아있는 부분이 많다. 이는 오히려 기회가 많다는 뜻이 되기도 한다. 우리나라가 웹3.0 시대에도 디지털 강국의 위상을 확립하고 IT의 발전을 이끌어나가기 위해서는 디지털화된 데이터량 증가에 따른 효율적인 데이터 관리 및 검색 시스템을 마련하고, 정보 관리·유통 체계를 정비해야 한다. 이를 위해 시맨틱 웹의 개발과 더불어 인적 네트워크와의 결합을 장려할 필요가 있다.

## 참고자료

- 김택천(2009a), “웹3.0 시대의 미래 기술의 방향”, 한국인터넷진흥원, issue inside.  
 \_\_\_\_\_(2009b), “웹3.0과 미래”, KT 오픈세미나, 2009. 5.

- 스트라베이스(2009), “web3.0시대의 실체와 인터넷의 미래”, Issue Alert, 2009. 1. 19.
- 이경일(2006), “메타데이터와 온톨로지”, 슬트룩스.
- \_\_\_\_\_(2008a), “시맨틱 웹: 기술을 넘어 서비스 플랫폼으로”, 한국소프트웨어진흥원, 정책리포트.
- \_\_\_\_\_(2008b), “웹 기반 텍스트 마이닝 기술과 사례”, eBiz 연구회 세미나, 2008. 9. 18.
- \_\_\_\_\_(2009), “정보의 지식화와 통찰력 확보”, 시맨틱 검색 기술 세미나, 2009. 6. 9.
- 이선재(2009), “Social Web에서 바라본 Web3.0”, KT 오픈세미나, 2009. 5.
- 정한민·이미경·성원경(2008), “시맨틱 웹2.0 기술 동향”, 정보통신연구진흥원, 포커스.
- 정한민(2008), “시맨틱 웹이 경제, 사회에 미치는 영향”, 정보통신연구진흥원, IT 기획시리즈.
- 한국교육학술정보원, “KERIS 의미기반검색 서비스시스템 소개”, 학술연구정보서비스 공식 홈페이지.
- 한국인터넷진흥원(2006), “시맨틱 웹에 대한 구글과 Tim Berners-Lee의 논쟁”, NIDA issue Tagging.
- Bratt, S(2009), “Future Internet 2020”, W3C.
- Hyvonen, E., Viljanen, K. & Suominen, O.(2007), “HealthFinland-Finnish Health Information on the Semantic Web”, *ISWC 2007*.
- Kobilarov, G., Scott, T., Raimond, Y. & Oliver, S.(2009), “Media Meets Semantic Web- How the BBC Uses DBpedia and Linked Data to Make Connections”, *ESWC2009*, Greece.
- Razmerita & Gouarderes(2005), “Ontology based User Modeling for Personalization of Grid Learning Services”, *the First International Elegi Conference on Advanced Technology for Enhanced Learning*, Napoli.
- Tim Berners-Lee(1998), “Semantic Web Road map”, W3C Design Issues.

Viljanen, K., Tuominen, J. & Hyvonen, E.(2009), “Ontology Libraries for Production Use: The Finnish Ontology Library Service ONKI”, *ESWC2009*, Greece.

《파이낸셜뉴스》, (2007. 5. 23), “구글 개인정보 구축”.

《ReadWriteWeb》, (2009. 1. 21), “BBC’s Semantic Music Project”.

《The New York Times》, (2006. 11. 12), “Entrepreneurs See a Web Guided by Common Sense”.

<http://linkeddata.org>.

[http://ontoframe.kr/2008/2008\\_new/main.jsp](http://ontoframe.kr/2008/2008_new/main.jsp).

[www.internetworldstats.com](http://www.internetworldstats.com).