

# 시맨틱 기술의 경제적 파급효과 분석

## An Economic Ripple Effect Analysis of Semantic Technology

박 성 옥\* · 한 선 화\*\* · 성 원 경\*\*\*

Sung-Uk Park · Sun-Hwa Hahn · Won-Kyung Sung

### 차 례

- |             |                |
|-------------|----------------|
| 1. 서 론      | 4. 경제적 파급효과 분석 |
| 2. 시맨틱 웹 기술 | 5. 결 론         |
| 3. 시맨틱 웹 시장 | · 참고문헌         |

### 초 록

국가 과학기술 경쟁력 향상 및 창조적 부의 원천 확보를 위해서는 지능화된 시맨틱 기술을 통한 정확한 정보 제공과 고품질의 정보 제공이 필요하다. 시맨틱 기술은 정보를 의미적으로 연계 및 통합하고 원하는 정보를 보다 쉽게 얻을 수 있는 기술이며, IT전반에 걸쳐 파급효과가 매우 큰 핵심기술로서 다양한 산업분야에서의 기술적 경쟁력을 확보하고 이를 성장동력화 하고 있다. 이에 본 논문은 시맨틱 기술이 국내 차세대 웹 시장을 통해 경제 각 부문에 미치는 효과를 한국은행(2009)의 산업연관분석을 이용하여 생산유발효과, 부가가치유발효과, 수입유발효과, 취업자유발효과로 나누어 경제적 파급효과를 분석하여 시맨틱 기술에 대한 경제적 타당성을 검증하고자 한다.

### 키 워 드

시맨틱 기술, 시맨틱 웹 시장, 생산유발효과, 부가가치유발효과, 수입유발효과, 취업자유발효과

- 
- \* 한국과학기술정보연구원 정책연구실 선임연구원  
(Senior Researcher, Dept. of Policy Research, Korea Institute of Science and Technology Information, supark@kisti.re.kr)
  - \*\* 한국과학기술정보연구원 정책연구실 책임연구원  
(Principal Researcher, Dept. of Policy Research, Korea Institute of Science and Technology Information, shhahn@kisti.re.kr)
  - \*\*\* 한국과학기술정보연구원 정보기술연구실 책임연구원  
(Principal Researcher, Dept. of Information Technology Research, Korea Institute of Science and Technology Information, wksung@kisti.re.kr)
- 논문접수일자: 2010년 4월 22일
  - 최종심사일자: 2010년 7월 13일
  - 게재확정일자: 2010년 7월 13일

## ABSTRACT

In order to enhance national science technology and to secure the origin of creative wealth, there is a need to offer accurate and high quality information with semantic technology. Semantic technology enables to significantly unify information and obtain the wanted information in an easier way. Also Semantic technology is a key technology of the next generation having great ripple effect on the whole of IT. In this paper, We analyze an economic ripple effects in terms of production inducement effect, added value inducement effect, import inducement effect, labor inducement effect of Semantic technology using an input-out analysis of the bank of Korea(2009). We also examine an economic propriety of Semantic technology.

## KEYWORDS

Semantic Technology, Semantic Web Market, Production Inducement Effect, Added Value Inducement Effect, Import Inducement Effect, Labor Inducement Effect

## 1. 서론

### 1.1 연구배경 및 목적

팀 버너스리에 의해 제창된 현재의 웹은 사람이 읽고, 쓰고, 실행할 수 있도록 발전되어 왔다. 인터넷의 발전과 웹 기반 산업의 성장은 폭발적인 정보량의 증가를 가져왔고, 이제는 수많은 정보들에서 어떻게 정확한 정보를 찾고, 활용할 수 있는가가 관건이 되고 있다. 폭발적인 정보의 생성, 공유와 유통은 웹을 놀라운 산업 플랫폼으로 만들었으나, 웹이 가지고 있는 구조적 잠재성에 비해 기존 웹의 정보 표현체제는 웹 기반 산업의 성장과 발전의 한계로 작용하고 있다. 특히, HTML에 기반한 정

보 표현 체제는 사람만이 읽고 쓰고 처리할 수 있기 때문에 방대한 정보를 효과적으로 자동화하는 것이 거의 불가능한 상태에 이르렀다.

팀 버너스리는 시맨틱 웹이 현재 웹과 동떨어진 것이 아니라 의미적으로 잘 정의된 정보를 사용하여 사람과 컴퓨터 간의 협력적인 일을 잘 할 수 있도록 하는 현재 웹의 확장이라고 정의한다. 시맨틱 웹이 현재의 웹과 동떨어져 있지 않다는 것은 현재의 웹이 사용하고 있는 URI 체계나 XML, namespace와 같은 개념과 시스템을 그대로 사용하고 있기 때문이다. 현재 웹이 가진 이러한 요소들에 데이터를 트리플로 표현해 내고, 그 위에 온톨로지, 로직 등의 요소가 덧붙여져 새로운 기술 구조를 가지게 되었다.

세계적으로는 미국의 W3C, DARPA, MIT 를 비롯하여 유럽의 DERI, STI 등이 표준화 및 연구를 주도하고 있고, IBM, Oracle 등의 대기업과 Ontoprise, Franz 등의 전문 벤처 기업들이 산업화 및 시장을 주도하고 있다. 국내에서도 정부 차원의 다양한 연구개발 프로젝트들이 KISTI, ETRI, KAIST를 중심으로 진행되고 있으며, 산업적으로는 KT와 솔트룩스가 상용 솔루션과 서비스를 내놓고 있다. 인터넷이 지난 10년 동안 우리 생활을 변화시킨 가장 획기적인 기술이었다면 시맨틱 웹 기술은 앞으로 10년의 우리 생활을 가장 크게 변화시킬 기술 일 것이다. 현재의 시맨틱 웹 기술과 산업은 인터넷의 초기 상황과 비슷하다.

그동안 국내에서는 시맨틱 기술, 시맨틱 시장, 시맨틱 웹 동향 등 서비스별로 단편적으로 조사 및 연구가 진행되어 왔다. 그리고 그에 대한 성과 및 효과도 합리적이고 객관적인 기법을 적용하기 보다는 일차적인 업무 성과 위주로 확인하고 있어 관련 사업에 대한 투자효과를 제대로 확인을 하지 못하고 있다. 이에 본 논문에서는 시맨틱 기술이 국내 차세대 웹 시장을 통해 우리나라 경제 각 부문에 미치는 효과를 한국은행(2009)의 산업연관분석을 이용하여 생산유발효과, 부가가치유발효과, 수입유발효과, 취업자유발효과로 나누어 경제적 파급효과를 분석함으로써 시맨틱 기술개발에 대한 경제적 타당성을 검증하고 이를 바탕으로 관련 사업에 대한 투자확대의 논리를 제공할 수 있다는 점에서 본 논문은 의미가 있다.

## 1.2 연구내용 및 방법

기존의 정보서비스에서는 정보량이 급증함에 따라 ‘부의 미래’에서 앨빈토플러가 말한 ‘무용지식(obsoledge)’으로 인한 전체 지식정보 체계의 비효율성이 가속화되고 있다. 국가과학기술 경쟁력 향상 및 창조적 부의 원천 확보를 위해서 지능화된 시맨틱 기술개발을 통한 정확한 정보제공과 고품질의 정보 제공이 필요하게 된다. 이에 시맨틱 기술이 국내 차세대 웹 시장을 통해 경제 각 부문에 미치는 효과를 한국은행(2009)의 산업연관분석을 이용하여 생산유발효과, 부가가치유발효과, 수입유발효과, 취업자유발효과로 나누어 경제적 파급효과를 국민경제적 차원에서 측정하여 향후 관련 연구사업에 대한 예산 배분 등 미래의 투자 계획 및 정책수립 등의 기초자료로 활용하고자 한다.

시맨틱 기술의 경제적 파급효과를 분석하기 위해 한국은행(2009)의 투입-산출표인 산업연관표에 의거하여 분류함으로써 시맨틱 기술이 국내 산업부문 중 어느 부문에 속해 있는지 파악할 것이다. 그리고 시맨틱 기술에 대한 파급효과를 이론적 차원에서 생산유발효과, 부가가치유발효과, 수입유발효과, 취업자유발효과로 나누어 분석하고 시맨틱 기술에 대한 각각의 파급효과를 한국은행의 투입-산출모형에 입각하여 분석할 것이다.

본 논문의 제2장에서는 시맨틱 기술이 무엇을 의미하고 그 기술에 속한 세부 기술들은 무엇들이 있는지 개략적으로 설명한다. 제3장에

서는 시맨틱 기술에 의한 차세대 웹시장의 세계 및 국내시장의 현황을 살펴보고, 제4장에서 는 경제적 파급효과 분석 방법론에 대해 설명 하고 본 연구에서 사용하고 있는 산업연관분석 모형에 대해 정리한다. 제5장에서는 본 논문의 결론 및 연구의 한계에 대해 정리해 본다.

## 2. 시맨틱 웹 기술

### 2.1 지능형 에이전트 기술

시맨틱 웹의 출현은 지능형 에이전트를 위 한 새로운 공간을 제공하기 위한 것이다. 웹이 사람을 위한 사이버 공간을 제공하였다면 시맨틱 웹은 지능형 에이전트를 위한 공간을 제 공하여 보다 사람이 편리하게 에이전트로부터 정보를 찾거나 서비스를 받을 수 있게 하는 것이다. 지능형 에이전트란 사전적인 의미로 ‘지 능적 대리인’ 또는 ‘지능을 가진 도우미’라고 정의 내릴 수 있다. 현재 인터넷에는 사용가능 한 정보가 무제한으로 증가하고 있는 이유로 인해서 적절한 정보이용에는 오히려 방해가 되고 있다. 이런 면에서 지능형 에이전트는 무 제한 정보를 이용자의 요구나 기호에 맞도록 가공한 자료를 제공하며, 필요한 정보에 좀 더 빠르게 접근할 수 있도록 도와주는 역할을 한 다는 점에서 의의를 가지고 있다. 지능형 에이 전트의 특성으로는 자율성(Autonomy), 대화 능력(Communication Ability), 협동성(Ca-

capacity for Cooperation), 판단력(Capacity for Reasoning), 적응력(Adaptive Behavior), 신뢰성(Trustworthiness) 등이 있다.

현재, 웹에 있는 HTML 문서는 사람이 보 고 이해하는 데는 불편함이 없도록 설계되어 있으나 소프트웨어 에이전트는 HTML 문서를 이해 할 방법이 없다. 예를 들면 우리가 일본 영화를 볼 때 일어를 모르면 영화의 내용을 이 해할 수 없는 것과 같이 소프트웨어 에이전트 도 HTML 문서를 보면 그 내용을 이해할 수 없다. 번역된 자막을 보면 우리가 일본 영화의 내용을 이해할 수 있는 것과 유사하게 소프트 웨어 에이전트는 시맨틱 웹 환경에서 HTML 문서의 메타데이터를 보면 HTML 문서의 내 용을 파악하게 된다. 그러나 일본어 대본이 가 지고 있는 의미와 자막에 번역된 의미의 차이 는 매우 크다. 번역된 자막은 영화를 이해하는 데 필요한 최소한의 정보만을 가지고 있을 수 밖에 없다. 따라서 소프트웨어 에이전트가 이 해할 수 있는 단순한 언어를 이용하여 HTML 문서의 내용을 메타데이터의 형태로 표현하면 된다. 즉, HTML 문서가 있는 웹 공간은 사람 을 위한 사이버공간이고, HTML 문서의 메타 데이터가 있는 웹 공간은 소프트웨어 에이전 트를 위한 사이버공간이라고 할 수 있다. 이때 메타데이터는 HTML 문서의 내용을 표현하고 있어야 하는데 사람이 생각한 모든 내용을 표 현한다는 것이 불가능하므로 사람의 생각 중 에서 소프트웨어 에이전트에게 가장 필요한 부분만을 개념화하여 표현하는 방식을 취하게

된다. 이와 같이 시맨틱 웹에서는 두 개의 공간을 준비하여 지능형 에이전트가 활동할 수 있는 구조를 제공하게 된다.

시맨틱 웹 환경에서는 기존의 웹 공간과 이 웹에 있는 정보를 표현하는 메타 공간으로 구성된다. 이 메타공간에 있는 정보는 웹 공간에 있는 HTML 문서의 의미를 표현하고 있으며 지능형 에이전트가 메타 공간에 있는 정보를 이해할 수 있도록 설계되어 있다. 메타공간에 표현되는 언어는 한정된 어휘를 가지고 있고 일정한 규칙을 가지도록 설계되었으므로 소프트웨어 에이전트는 메타공간에 표현되는 언어를 이해할 수 있다.

## 2.2 온톨로지 기술 개발

지능형 에이전트가 동작할 수 있도록 구조적인 공간을 만들어 주기 위해서는 사람이 만든 HTML 문서의 내용을 메타데이터의 형식으로 표현하여야 한다. 이와 같은 메타데이터 공간을 지능형 에이전트가 활용하게 되므로 메타데이터를 표현하는 것은 매우 중요하다. 사람은 머릿속에 있는 생각과 개념을 이용하여 HTML 문서를 표현하였으므로 이와 같은 HTML 문서가 가지는 의미를 소프트웨어 에이전트에게 표현하기 위해서는 사람이 가지고 있었던 개념이나 생각을 표현하는 것이 필요

하다. 이와 같은 시맨틱 개념을 이용하여 소프트웨어 에이전트를 위한 메타데이터를 표현하게 되는데 사람이 생각한 시맨틱 개념을 소프트웨어 에이전트가 이해할 수 있도록 표현한 것을 온톨로지<sup>1)</sup>라고 한다.

온톨로지는 두가지 중요한 의미를 가지고 있다. 첫째는 메타데이터를 표현하기 위한 해당분야의 개념이라고 할 수 있다. 즉 앞에서 예를 든 영화가 공상과학 영화라면 자막에는 과학 용어가 나오게 될 것이고 우리는 그 과학 용어에 대한 개념을 알고 있을 때 그 자막을 이해할 수 있다. 반대로 그 영화가 일본 역사 물이라면 자막에 나오는 역사 사건에 대한 개념이 있을 때 그 영화의 내용을 이해할 수 있다. 이와 같이 자막과 같은 메타데이터가 있더라도 그 자막에 있는 어휘에 대한 개념을 가지고 있을 때 그 내용을 알 수 있게 되고 온톨로지는 메타데이터에 있는 어휘에 대한 개념이라고 할 수 있다.

둘째는 이와 같은 온톨로지는 공유된다는 점이다. 즉 온톨로지를 설계하는 것은 시맨틱 웹 프로그래머이고 이 프로그래머는 온톨로지를 설계할 때 일정한 규칙과 잘 정의된 방식에 따라서 설계하여 이 내용을 소프트웨어 에이전트에게 알려주게 된다. 이와 같은 과정을 거쳐서 온톨로지는 사람도 이해하고 소프트웨어 에이전트도 이해하는 특징을 가지게 된다.

1) 위키백과에 의하면 온톨로지란 사람들이 세상에 대하여 보고 듣고 느끼고 생각하는 것에 대하여 서로 간의 토론을 통하여 합의를 이룬 바를 개념적이고 컴퓨터에서 다룰 수 있는 형태로 표현한 모델로, 개념의 타입이나 사용상의 제약조건들을 명시적으로 정의한 기술이다.

## 2.3 메타데이터 표현 기술

시맨틱 웹이 성공하기 위해서는 메타데이터가 필수적이다. 시맨틱 웹이 출현한 이유는 지능형 에이전트가 웹의 정보를 이해할 수 있는 공간을 구축하기 위함인데 이와 같은 공간이 결국은 메타데이터이다. 메타데이터는 일반적으로 데이터에 관한 데이터로서 정보자원의 속성을 기술하는 데이터를 의미한다. 즉, 메타데이터란 실제로 저장하고자 하는 데이터(예: 비디오, 오디오, 텍스트 등)는 아니지만 이 데이터와 직접적으로 혹은 간접적으로 연관된 정보를 제공하는 데이터를 말한다.

웹에서는 HTML 문서와 같은 한 종류의 정보만 존재하지만 시맨틱 웹에서는 HTML 문서가 가지고 있는 의미 있는 정보를 지능형 에이전트가 이해할 수 있도록 메타데이터를 필요로 한다. 메타데이터는 웹상의 문서나 정보의 의미를 표현하는데 그 문서의 주요 개념, 저자, 저작 연월일 등이다. 어떻게 보면 매우 간단한 정보같지만 기존의 웹상에서는 이와 같은 정보가 명시적으로 표현되지 않았기 때문에 소프트웨어가 처리하는데 매우 불편했고 따라서 소프트웨어의 지능이 떨어지게 되었다.

## 3. 시맨틱 웹 시장

### 3.1 시맨틱 웹 시장 동향

시맨틱 웹 시장은 현재 매우 초기 단계에

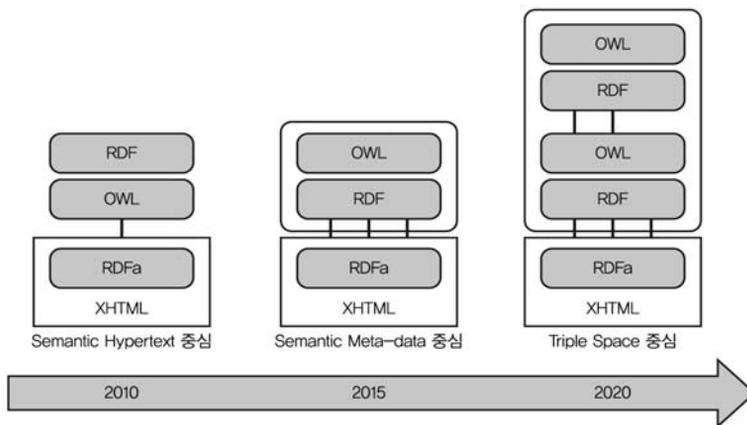
있다. 현재의 실정은 1980년대 말의 웹 시장과 유사한 상황이다. 온톨로지 표현 언어는 표준화가 된 상태이지만, 메타데이터를 자유롭게 저작할 수 있는 저작 도구가 보편화되어 있지 않아서 일반대중이 활용하는 데는 많은 어려움이 있다. 그러나 초기의 웹 시장이 그러하듯이 전문적인 분야에서는 매우 활발하게 시장을 형성하고 있다. 미국과 EU에서는 W3C, DARPA, DERI 등의 연구소와 IBM, HP, Google 등의 회사에서 LEAD(Live Early Adoption and Development)와 같은 방식으로 킬러 어플리케이션을 발굴하는 작업을 수행하고 있다. 현재 시맨틱 웹 시장은 온톨로지나 메타데이터를 적극 활용하는 분야나 활용할 수밖에 없는 분야에서 시장을 형성하고 있다. 전자의 경우는 블로그 시장을 예로 들 수 있는데, 블로그 시장에 있는 모든 사용자들은 자신의 정보를 광고하는 데 열심인 그룹들이다. 이 그룹에서는 자신의 정보를 널리 알리기 위해서 온톨로지 기반의 메타데이터를 적극 활용하는 분위기이므로 시맨틱 웹의 킬러 어플리케이션이 생길 수 있는 가능성이 매우 크다. 이런 이유 때문에 MIT 인공 지능 연구실과 IBM T.J. Watson 연구소, 그리고 EU의 SWAD-E(Semantic Web Advanced Development-Europe) 등에서 연구가 진행되어 왔다. 두 번째 시장은 엔터프라이즈 소프트웨어 시장이다. 이 시장은 지식 공유를 매우 중시하는 분야이므로 온톨로지를 이용한 메타데이터를 저작한다는 불편함을 극복하고서라도

지능형 소프트웨어를 생산하려는 의지가 매우 높은 시장이다. 따라서 향후 ERP(Enterprise Resource Planning), CRM(Customer Relationship Management), SCM(Supply Chain Management) 시장 등도 점진적으로 온톨로지 기반의 지능형 소프트웨어로 발전할 것이라는 예측이 많이 나오고 있다.

가트너 그룹의 경우 시맨틱 웹 기술이 Hype Cycle의 기술 과장(Inflation) 단계를 넘어 이제 상용화 단계로 넘어가고 있다고 진단했다. 이경일(2008)의 분석에 의하면 <그림 1>처럼 이미 시작된 시맨틱 웹의 초기 단계는 RDFa와 같은 형태로 HTML 문서에 내장된 Semantic Hypertext가 폭 넓게 쓰일 것이며, 2010년 경에는 웹에서 RDFa와 RDF가 매우 일반화되어 사용될 것으로 전망하고 있다. 또 한 궁극적 시맨틱 웹 비전의 달성은 10년에서 15년이 걸릴 것으로 예측하고 있다.

현재 웹시장은 서비스를 포함하는 시장으로

확산되고 있다. 즉, 과거의 웹에는 문서나 정지영상과 같은 정적인 콘텐츠를 포함하고 있었지만 근래에 들어 웹에서 서비스를 처리할 수 있는 응용 소프트웨어가 출현하고 있다. 현재 WSDL(Web Service Definition Language), UDDI(Universal Description Discovery and Integration), SOAP(Simple Object Access Protocol)와 같은 방식을 이용하여 웹에 있는 서비스를 표현하고, 저장하고, 검색하는 웹 서비스 시장이 활발히 만들어지고 있다. 시맨틱 웹 서비스 분야에서는 이와 같은 웹 서비스에 온톨로지를 이용하여 서비스에 대한 메타데이터를 표현함으로써 지능형 에이전트가 사람을 대신하여 작업을 수행할 수 있는 시장을 개척하는 시도가 이루어지고 있다. 궁극적으로 미래 사회에서는 현재 사람이 하는 모든 웹상이 트랙잭션을 지능형 에이전트가 대신하게 될 것이므로 시맨틱 웹 서비스 기술은 매우 필수적일 것으로 예측을 하고 있다.

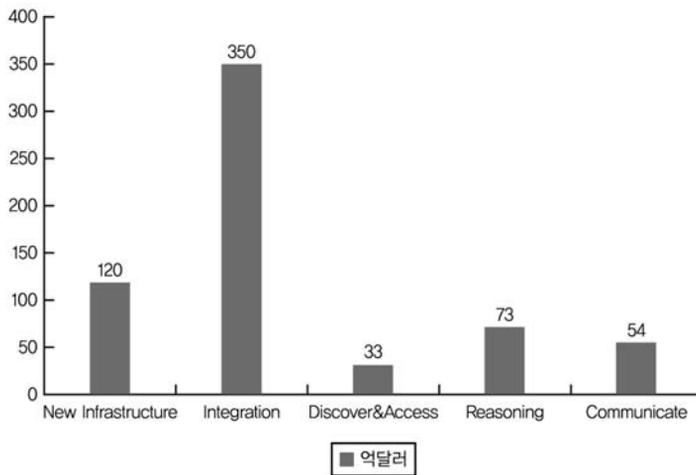


<그림 1> 시맨틱 웹의 진화(이경일 2008)

시맨틱 웹과 관련한 시장의 출현은 필수적인 것으로 간주되고 있는데 박영택(2008)의 분석에 의하면 온톨로지 기반의 기술이 향후 핵심 기술로 자리 잡을 것이며 2010년 이후에는 인공지능의 지식표현 기술을 활용하는 강력한 온톨로지가 응용 프로젝트의 80% 이상

을 차지할 것으로 예상하고 있다. 또한, Top-Quadrant사를 인용하여 IT 분석가와 벤더들의 연구 결과를 바탕으로 2010년의 시맨틱 웹 시장을 <그림 2>와 같이 예상하고 있다.

또한, 정보통신산업진흥원(2010)을 재구성하면 세계 차세대 웹 시장규모는 2008년 248



<그림 2> 2010년의 시맨틱 웹 시장(박영택 2008)

<표 1> 세계 차세대 웹 시장 규모

(단위: 백만 달러)

구분	2008년	2009년	2010년	2011년	2012년	CAGR	
웹 2.0	포털 및 응용미들웨어	9,500	10,000	10,526	11,080	11,662	5.3%
	콘텐츠연동	5,341	5,842	6,450	7,120	7,861	10.1%
	소계	14,841	15,842	16,976	18,200	19,523	7.1%
시맨틱 웹 및 온톨로지 (시맨틱기술)	시맨틱 인프라, 개발	978	1,305	1,740	2,121	3,095	33.4%
	지식자동화 및 업무활용	2,329	2,856	3,500	4,191	5,261	22.6%
	소계	3,307	4,161	5,240	6,512	8,356	26.1%
모바일 및 유비쿼터스 웹	4,586	7,375	11,179	15,250	19,149	42.9%	
웹 검색	2,219	2,455	2,699	2,966	3,255	10.1%	
합계	24,863	29,831	36,094	42,728	50,283	19.3%	

※ 웹 2.0 기술을 이용한 연계서비스인 포털 및 응용 미들웨어

자료: 정보통신산업진흥원(2010) 재구성

억 달러에서 2012년 503억 달러로 연평균성장률(CAGR)은 19.3% 증가할 것이고 특히 시맨틱 웹 및 온톨로지인 시맨틱 기술의 시장규모는 2008년 33억 달러에서 2012년 84억 달러로 26.1% 증가할 것으로 예상할 수 있다 (<표 1> 참조).

국내 시맨틱 웹 기술과 관련된 시장규모는 <표 2>에서 보듯이 정보통신산업진흥원(2010)을 재구성하면 2010년 2983억원을 예상하고 있으며, 2012년에는 19.1% 증가하여 5028억 원을 전망할 수 있다.

### 3.2 시맨틱 웹 기술 개발 현황

국내 시맨틱 웹 기술과 관련된 시장규모는 <표 2>에서 보듯이 긍정적인 시장 전망에도 불구하고 기술 개발의 어려움과 그에 따른 중장기 투자의 부담으로 인하여, 민간 부문보다는 국가 연구개발 사업을 중심으로 관련 기술에 대한 연구들이 수행되고 있는 상황이다.

현재 주요 국내 연구로는 방송통신위원회의 선도기반기술개발 사업에서 추진 중인 “차세대 웹을 위한 시맨틱 서비스 에이전트 기술개발”과 “국가 IT 코어 온톨로지 구축”이 있는데 대체적으로 초기 연구 단계로서 핵심 요소 기술과 자원을 개발하고 제한된 분야를 대상으로 한 시험적 응용을 시도하고 있는 상태에 머물고 있다.

시맨틱 기술의 실용화 가능성은 2006년도에 한국과학기술정보연구원에 의해 KISTI-STAR와 OntoFrame®이 개발·공개되고 2007년도에는 한국학술진흥재단의 연구정보 분석 서비스와 심사자 추천 서비스, 2008년도 및 2009년도에는 법무부의 지능형 입법지원시스템과 기술표준원의 온톨로지 시범서비스에 성공적으로 적용되고 활용되면서 가시화되고 있다.

BI(Business Intelligence), CRM 등과 같은 엔터프라이즈 분야에서의 요구가 급속히 확대되고 있는 실정이지만 실질적인 상용화 모델은 아직까지도 주로 정형자료를 바탕으로

<표 2> 국내 차세대 웹 시장 규모

(단위: 억원)

구분	2008년	2009년	2010년	2011년	2012년	CAGR
웹2.0	1,484	1,584	1,698	1,820	1,952	7.1%
시맨틱 웹 및 온톨로지 (시맨틱기술)	331	416	524	651	836	26.1%
모바일 및 유비쿼터스 웹	458	737	1,179	1,525	1,914	43.0%
웹 검색	222	246	270	297	326	10.1%
합계	2,495	2,983	3,671	4,293	5,028	19.1%

자료: 정보통신산업진흥원(2010) 재구성

한 데이터 마이닝 수준에 머물고 있는 실정이다. 이는 시장 협소의 차원에서가 아니라 상당히 많은 사용자 요구사항에도 불구하고 국내 고유 기술의 부족으로 나타나는 현상으로 볼 수 있다. 비정형 자료에 대한 진정한 의미의 텍스트 마이닝 소프트웨어는 아직까지 국내에서 상용화되지 않았으며, 정보검색 및 자연어처리 전문 업체들이 초보적인 수준의 텍스트 마이닝 기능들을 개발하고 있는 상황이다. 이러한 상황에서 BI 및 정보분석 분야에서 세계 최고 수준의 기업들인 SAS, SPSS, BO, Cognos 등과 같은 외국업체들이 국내 시장을 잠식하고 있는 상황에 있으며 논문, 특허, 기술문서 등과 같은 전문정보 대상 텍스트 마이닝 및 분석 시도는 거의 이루어지고 있지 않고 있다.

현재 시맨틱 웹 기술 관련 세계 시장 규모는 <표 1>에서 보듯이 폭발적 증가 추세를 보이고 있으며 이에 따라 시장 선점을 위한 세계 각국의 R&D 투자도 지속적으로 증가하고 있다. 이와 같은 시장 추세에 발맞추어 최근의 시맨틱 기술 관련 연구들은 대용량 정보처리를 위한 기술 개발에 집중되고 있어, 미국을 중심으로는 대용량 지식베이스를 대상으로 한 저장소, 추론 엔진 등의 개발을 위한 연구들이 활발히 진행되고 있으며 특히, 영국은 Univ. of Southampton을 중심으로 분산 환경 내의 다중·이형 온톨로지 연계 기술을 개발하고 있다. 기술 선진국들은 관련 국제 표준을 선도함과 아울러 시맨틱 기술과 관련된 기반 연구를 완료하고 2000년도부터는 시맨틱 기술을

다양한 환경에 적용하기 위한 제품들을 상용화하여 발표하고 있다. 2000년도 중반부터는 영국의 AKT(Advanced Knowledge Technologies)와 EU의 SEKT(Semantic Knowledge Technologies) 등과 같이 대규모 정보유통 서비스의 효율화를 위하여 정보유통 기술과 시맨틱 기술 간의 융합을 시도하는 연구들이 유럽을 중심으로 시작되었다.

정보서비스 기술과 시맨틱 기술 간의 융합은 현재 미국을 중심으로 본격적으로 진행되고 있어 한국과학기술정보연구원(2009)에 의하면 IBM, BEA, CISCO, Vivisimo, Cerebra 등 약 200여개 이상의 업체가 시맨틱 기술과 관련된 제품을 공급하고 있으며 정부기관, 금융, 제조, 물류, 교통, 통신, 에너지 등 다양한 분야에서 100여 개 이상의 시맨틱 기술 기반의 정보서비스 개발 사례가 보고되었다. 선진국은 텍스트 마이닝에 대한 중요성을 일찍부터 인식하였으며 국가적으로 대내적인 연구지원사업을 추진하고 있다. 이에 그러한 결과로 현재는 기술수준이 이미 연구단계를 넘어서서 실용화 단계에 접어든 실정이다. 또한 같은 보고서에서 Gartner 자료를 인용하여, 2007년 30개 국가의 1,400명의 CIO를 대상으로 한 조사에서 BI가 2006년에 이어 2007년에도 기업정보화 투자 우선순위 1위로 기록되었으며 이는 기업들이 넘쳐나는 다양한 형태의 정보를 대상으로 정확한 분석 및 요약 그리고 예측을 가장 최우선 경영과제로 여기고 있음을 나타내고 있다.

## 4. 경제적 파급효과 분석

### 4.1 산업연관분석의 의의 및 구조

한국은행(2009)에 의하면 한 국가 경제에서 각 산업들은 생산활동을 위해 상호 간에 재화와 서비스를 구입하고 판매하는 과정을 통해 직접 또는 간접적으로 서로 관계를 맺게 되는데 산업연관표는 일정기간(보통 1년) 동안의 이러한 산업간 거래관계를 일정한 원칙에 따라 행렬식으로 기록한 통계표이다. 산업연관분석 또는 투입산출분석은 산업연관표를 바탕으로 하여 산업간 상호연관계를 수량적으로 분석한 방법인 것이다. 산업연관분석은 최종 수요가 유발하는 생산, 고용, 소득 등 각종의

파급효과를 산업부문별로 구분하여 분석할 수 있기 때문에 경제정책의 수립, 정책효과의 측정 등에 활용되고 있다. 다만 한국은행 산업연관표를 통한 생산파급효과 등은 시맨틱 기술에 대한 R&D 투자 등의 직접적인 효과를 고려하지 못하는 점에서 과소평가의 우려가 있지만 가격의 변화를 고려하지 못하는 정태분석이라는 점에서 효과를 과대평가할 우려가 상존하고 있어서 과소 또는 과대평가의 상대적 크기에 대해 단정적인 평가를 내리기는 어려운 점이 존재하고 있다.

한국은행의 산업연관표를 이용하여 국내 각 산업에 대한 시맨틱 기술에 대한 차세대 웹 시장의 생산, 부가가치 등의 파급효과를 추정하기 위해 시맨틱 기술이 속한 '컴퓨터관련서비

〈표 3〉 산업연관표 기본 구조

내생부문	중 간 수 요						최종수요	수입(공제)	총산출액	
	1	2	·	·	·	n				
중 간 투 입	1	$x_{11}$	$x_{12}$	·	·	·	$x_{1n}$	$Y_1$	$M_1$	$X_1$
	2	$x_{21}$	$x_{22}$	·	·	·	$x_{2n}$	$Y_2$	$M_2$	$X_2$
	·	·	·	·	·	·	·	·	·	·
	·	·	·	·	·	·	·	·	·	·
	n	$x_{n1}$	$x_{n2}$	·	·	·	$x_{nn}$	$Y_n$	$M_n$	$X_n$
부가가치	$V_1$	$V_2$	·	·	·	$V_n$	외생부문			
총투입액	$X_1$	$X_2$	·	·	·	$X_n$				

주) 1. 총투입액 = 중간투입+부가가치 (투입구조)

총산출액 = 중간수요+최종수요-수입 (배분구조)

총투입액 = 총산출액

2. 세로방향: 각 산업부문이 해당상품 생산을 위하여 지출한 생산비용의 구성, 즉 투입구조를 설명함  
가로방향: 각 산업부문의 생산물이 어떤 부문에 중간 수요 또는 최종수요 형태로 얼마나 판매되었는가 하는 배분구조를 의미

자료: 한국은행(2009)

스'에 속하는 행과 열을 모두 '0'으로 대체하여 만들고 파급효과 추정을 위한 각종 계수도 이를 통해 산출하여 유발효과를 얻을 수 있다.

## 4.2 각종 계수의 산출 및 산업연관효과

한국은행(2009)과 신용보증기금(2009)에 의하면 각 산업부문이 해당 부문의 재화나 서비스 생산에 사용하기 위하여 다른 부문으로부터 구입한 원재료 등의 중간 투입액을 총투입액으로 나누어 산출한 계수를 투입계수라 하고 투입계수를 산업연관표의 내생부문과 같은 모양으로 배열한 행렬이 투입계수표라 설명하고 있다. 시맨틱 기술에 의한 국내 차세대 웹 시장에 대한 경제적 파급효과를 산출하기 위해서는 수입된 재화 또는 서비스로 인한 산출효과를 배제해야 하므로 수입거래 금액을 제외한 국산 투입계수는 식(1)과 같다.

ij 산업 간의 국산투입계수

$$a_{ij} = \frac{X_{ij} - M_{ij}}{X_j} \quad \text{식(1)}$$

생산유발계수는 식(1)을 이용하여 시맨틱 기술이 속한 '컴퓨터관련서비스' 부문을 외생 변수화 한 후에 식(2)를 사용하여 생산유발계수를 산출한다.

$$\text{생산유발계수} = A_s^d (I - A^d)^{-1} \quad \text{식(2)}$$

- $A_s^d$  : 시맨틱기술분야(금융 및 보험)의 국산투입계수 행벡터

- $I$  : 1로 이루어진 대각 행렬(diagonal matrix)
- $A^d$  : 국산투입계수( $a_{ij}$ ) 행렬

부가가치계수는 총 산출에서 부가가치가 차지하는 비중을 나타내며, 산업연관표에서 각 산업의 부가가치 합계를 총 산출로 나누어 식(3)처럼 구한다.

$$i \text{ 산업의 부가가치계수 } v_i = \frac{V_i}{X_i} \quad \text{식(3)}$$

노동계수란 일정기간 동안 생산활동에 투입된 노동량을 총산출액으로 나눈 계수로 한 단위의 생산에 직접 필요한 노동량을 의미하며, 노동량에 자영업주 및 무급가족종사자를 포함하느냐의 여부에 따라 취업자계수와 고용계수로 구분한다.

$$\begin{aligned} \text{취업계수 } l_w &= \frac{L_w}{X}, \\ \text{고용계수 } l_e &= \frac{L_e}{X} \end{aligned} \quad \text{식(4)}$$

- $L_w$  : 취업자수,  $L_e$  : 피용자수,  $X$  : 총 산출액

노동유발계수는 어느 산업의 생산물을 한 단위 생산하는데 직접 필요한 노동량 뿐 아니라 생산파급과정에서 간접적으로 필요한 노동량까지 포함하고 있으며 식(5)로 나타내며, 본문에서는 취업자유발계수를 이용하여 취업자유발효과를 분석한다.

$$\begin{aligned} \text{취업자유발계수} &= \hat{l}_w (I - A^d)^{-1}, \\ \text{고용유발계수} &= \hat{l}_e (I - A^d)^{-1} \end{aligned} \quad \text{식(5)}$$

- $\hat{l}$  : 노동계수의 대각 행렬,  $I$  : 단위 행렬,  $A^d$  : 국산투입계수 행렬

산업연관분석을 적용한 국내 기존연구로는 박성욱(2008)은 국가 지식정보시스템 개발에 대해 경제적 파급효과를 분석하고, 박성욱 등(2008)은 정보보호산업의 경제적 파급효과를 분석하였다. 김도환(2007)은 정보통신산업에서의 경제적 기여도를 파악하고 정보통신산업의 특정 세부산업이 국민경제에 미치는 영향력을 분석하고 있다. 또한, 김수현 등(2006)은 IT기기 수출이 국내 산업에 미치는 경제적 파급효과를 분석하고 있는 등 최근에 관련 문헌상에 산업연관분석이 경제적 파급효과의 분석도구로 사용되고 있다. 산업연관분석은 경제부문 간의 재화와 서비스의 흐름이 비교적 안정적이라는 점을 활용하여 경제체계의 모습

을 보다 자세하게 통계적으로 분석함으로써 경제현상에 대한 설명을 보다 구체적으로 해주는 역할을 하고 있다. 특히 산업연관분석은 한 나라의 경제정책수립 및 효과분석과 관련된 분야에 많이 이용되고 있어 본 논문에서 분석하고자 하는 시맨틱 기술에 대한 국내 차세대 웹 시장의 경제적 파급효과를 분석하는데 적절하며 타당하다고 볼 수 있다.

우선, 산업연관분석을 적용하기 위해서는 본 논문에의 시맨틱 기술은 통합소분류의 ‘컴퓨터관련서비스’에 적용시켰다.

본 논문에서는 식(1)~식(5)를 통하여 생산유발계수, 부가가치유발계수, 수입유발계수, 취업자유발계를 산출하여 <표 5>에 정리를 하였으며 각각의 유발계수는 시맨틱 기술에 대한 최종수요가 한 단위 증가하는 경우 전 산업에서 직·간접적으로 일어나는 유발효과를 설명하고 있다.

<표 5>를 살펴보면 시맨틱 기술에 대한 부

<표 4> 시맨틱 기술의 산업연관표 해당부문

통합중분류(78부문)	통합소분류(168부문)	기본분류(402부문)
67. 사업관련 전문서비스	153. 컴퓨터관련서비스	366. 소프트웨어개발공급 367. 컴퓨터관련서비스

자료: 한국은행(2009)

<표 5> 시맨틱 기술의 유발계수

통합소분류	생산유발계수	부가가치유발계수	수입유발계수	취업자유발계수 (명/10억원 당)
컴퓨터관련서비스(153)	1,713,937	0,827,544	0,172,456	15,1628
전체 산업 평균	1,940	0,722	0,278	10,4

자료: 한국은행(2009)

가가치유발계수는 0.827544로 전체 산업평균인 0.722보다 크게 산정됨으로써 부가가치를 유발하는 계수가 큰 기술이고 취업유발계수도 10억원 당 15.1628명으로 전체 산업평균인 10.4명에 비해 취업을 유발하는 계수가 큰 기술임을 알 수 있다. 이제 <표 5>에서 구한 시맨틱 기술의 유발계수를 바탕으로 유발효과를 계산하기 위해 우선 생산유발효과는 시맨틱 기술에 의한 국내 차세대 웹 시장액에 산업연관표를 통해 산출한 생산유발계수를 곱하여 추정하고, 부가가치유발효과는 시맨틱 기술을 통해 얻을 수 있는 국민경제적인 순가치이며 생산유발효과 중 부가가치 귀속부분을 부가가치계수를 통해 추정할 수 있다. 시맨틱 기술로 인한 국내 차세대 웹 시장이 증대함에 따라 취업자의 증가로 이어지는 관점에서 취업자유발효과를 추정할 수 있으며 이는 생산유발효과에 취업자유발계수를 곱하여 산출할 수 있다.

<표 6>의 결과처럼 2008년~2012년까지의

시맨틱 기술에 의한 생산유발효과는 2008년 4276억원에서 2012년 86186억원으로 예상되어 총 31657억원, 부가가치유발효과는 2008년에는 3539원에서 2012년에는 7132억원으로 예상되어 총 26198억원, 수입유발효과는 2008년에는 737억원에서 2012년에는 1486억원으로 총 5459억원의 경제적 파급효과가 발생함을 알 수 있다. 또한, 취업자유발효과는 2008년에 6,484명에서 2012년에는 1만3,067명으로 예상되어 총 4만8,000명의 취업자유발이 예상된다.

## 5. 결론

정휘웅 외(2010)에 의하면 폭증하는 정보를 효과적으로 관리하고 활용하기 위해서는 W3C가 제시한 시맨틱 기술을 활용하여 지식의 생명주기를 효과적으로 관리하고 이를 기

<표 6> 시맨틱 기술의 경제적 파급효과

(단위: 억원/명)

구분	2008년	2009년	2010년	2011년	2012년	총 효과
생산유발효과	4,276	5,113	6,292	7,358	8,618	31,657
부가가치유발효과	3,539	4,231	5,207	6,089	7,132	26,198
수입유발효과	737	882	1,085	1,269	1,486	5,459
취업자유발효과	6,484	7,752	9,540	11,157	13,067	48,000

반으로 지식을 적시에 탐지하고 융·복합하여 제공함으로써 국가 R&D 생산성 제고를 크게 기여할 수 있다고 설명하고 있다. 시맨틱 기술은 IT 전반에 걸쳐 경제적 파급효과가 매우 큰 차세대 핵심기술로서 다양한 산업분야에서의 기술적 경쟁력을 확보하고 이를 성장동력화하기 위해 국가차원의 원천 및 응용 기술 보유가 필요한 부분이다.

본 논문에서는 시맨틱 기술에 대한 국내 시맨틱 웹 시장에서의 경제적 파급효과를 분석하기 위해 한국은행(2009)의 산업연관표를 사용하였으며 2008년~2012년까지의 국내 시맨틱 웹 시장을 통해 산업연관분석을 수행하였다. 그 결과로 전체 산업 평균에 비해 부가가치유발효과 및 취업유발효과가 크다는 점에서 시맨틱 기술이 미치는 경제적 파급효과가 타 산업보다 매우 크다는 것을 확인하였다. 하지만, 본 논문에서 활용된 한국은행의 산업연관표를 통한 경제적 파급효과 분석은 시맨틱 기술에 대한 R&D 투자 등의 직접적인 효과를 고려하지 못하는 점에서 과소평가의 우려가 있지만 가격의 변화를 고려하지 못하는 정태 분석이라는 점에서 효과를 과대평가 할 우려가 상존하고 있어서 과소 또는 과대평가의 상대적 크기에 대해 단정적인 평가를 내리기는 어려운 점이 존재하고 있다. 또한 분석대상기간인 2008년~2012년까지 산업의 구조가 동일한 가정을 취하고 있다는 점을 약점으로 지적할 수 있다.

마지막으로, 본 논문을 통해 살펴본 시맨틱

기술은 정보를 의미적으로 연계·통합하고 원하는 정보를 보다 쉽게 얻을 수 있게 하기 위한 기술로서 과학기술 하부구조 고도화를 위한 과학기술 정보유통 체제 구현에 크게 기여할 수 있을 것으로 기대되며, DB 모델링, 정보검색, 기업의사결정(Business Intelligence), 지식관리 등 다양한 산업분야에 적용되어 각종 지식정보를 보다 정확하고 편리하게 구축·관리·분석·제공할 수 있는 지능형 제품 및 서비스 개발을 직접적으로 유도할 수 있을 것으로 기대된다. 또한, 시맨틱 기술 기반의 차세대 정보서비스 기술은 원하는 정보를 보다 쉽게 얻을 수 있는 지식정보 서비스 구축을 가능하게 함으로써 지식-정보-지능화 사회 실현에 크게 기여할 수 있기를 기대한다.

## 참고문헌

- 김도환. 2007. 산업연관분석에 의한 정보통신산업의 경제적 파급효과. 『한국경영과학회지』, 32(3): 81-96.
- 김수현, 강희일, 홍승표, 정해식. 2006. IT 기기 수출의 경제적 파급효과. 『기술혁신학회지』, 9(2): 279-303.
- 박성욱. 2008. 국가 지식정보시스템 개발의 경제적 효과분석-한국과학기술정보원(KISTI)의 연구활동을 중심으로. 『정보관리연구』, 39(1): 73-94.
- 박성욱, 이상호. 2008. 한국의 정보보호산업과 경제적 파급효과-산업연관분석을 이용

- 하여-. 『산업혁신연구』. 24(2): 1-22.
- 박영택. 2008. 『차세대 웹 기술: 시맨틱 웹』.  
[인용 2010.02.18].  
〈<http://itisfun.tistory.com/181>〉.
- 신용보증기금 성과평가부. 2009. 『2008년도 신용보증 성과 및 국민경제기여도 분석』.  
서울: 신용보증기금.
- 이경일. 2008. 시맨틱 웹: 기술을 넘어 서비스 플랫폼으로. 『SW Insight 정책리포트』,  
32-47.
- 정휘웅, 김경선, 정한민. 2010. 시맨틱 검색 기술  
동향. 『주간기술동향』, 1431호: 14-27.
- 정보통신산업진흥원. 2010. 『정보통신 중점기술 표준화 로드맵 Ver. 2010』. 대전: 정보통신산업진흥원.
- 한국과학기술정보연구원. 2009. 『차세대 정보유통 핵심기술 연구·개발』. 대전: 한국과학기술정보연구원.
- 한국은행. 2009. 『2007년 산업연관표』. 서울: 한국은행.
- 한국은행. 2009. 『산업연관분석 해설』. 서울: 한국은행.