

# 시맨틱 웹 규칙 언어를 이용한 지능형 상품 정보 검색 에이전트 개발 (Intelligent Product Search Agent based on SWRL)

김우주<sup>1</sup>, 김정명<sup>2</sup>, 최대우<sup>3</sup>

<sup>1</sup> Department of Computer and Industrial Systems Engineering College of Engineering, Yonsei University  
134, Shinchonl-Dong, Soedaemun-Gu, Seoul, 120-749, Korea  
Tel: +82-2-2123-5716, Fax: +82-2-2260-8824, E-mail: kkim@yonsei.ac.kr

<sup>1</sup> 연세대학교 컴퓨터산업공학부  
서울 서대문 신촌동 134, 120-749  
Tel: +82-2-2123-5716, Fax: +82-2-2260-8824, E-mail: wkim@yonsei.ac.kr

## Abstract

We developed Intelligent Product Search Agent based on SWRL, and this agent can search product information with knowledge (facts and rules) on the web, implement price comparison for searched products considering delivery rates.

Existing keyword based product search engines is poor at searching intent products though a user has already perfect knowledge about intent produces. Furthermore if a user has insufficient knowledge, it is impossible to implement search. Also, existing price comparison shopping mall gives users comparison service considering total price (product prices, taxes, delivery rates), this service is valid to single product and has limitations of system expansion and up-dating because of not rule base but programming base.

If there is appropriate knowledge on the Semantic web and this makes product information retrieval possible, above problems can be solved clearly. In this research, we developed Intelligent Product Search Agent based on SWRL that can search product information efficiently by making agent to handle facts and rules by itself.

## Keywords

Intelligent Agents, E-Commerce, E-Business, Semantic Web, Web Service

## 서론

웹 자원이 폭발적으로 증가함에 따라, 정보의 양이 인간이 처리할 수 있는 한계를 초과함으로써, 인간 대신 기계가 웹 자원을 스스로 처리할 수 있도록[1] 웹에 의미를 부여하는 시맨틱 웹 분야 연구에 대한 수요도 기하급수적으로 증가하고 있다.

시맨틱 웹은 그림 1 [2]에서 보듯이 일곱 계층으로

구분되는데, 현재 시맨틱 웹의 추세는 2단계인 발전과 전개 단계(계층 4 - Ontology vocabulary의 활용)이며[3], 한편 다음 단계인 다섯 번째 계층인 규칙에 의미를 더하는 논리 수립에 대한 연구 필요성도 자연스럽게 증가하고 있다[4].

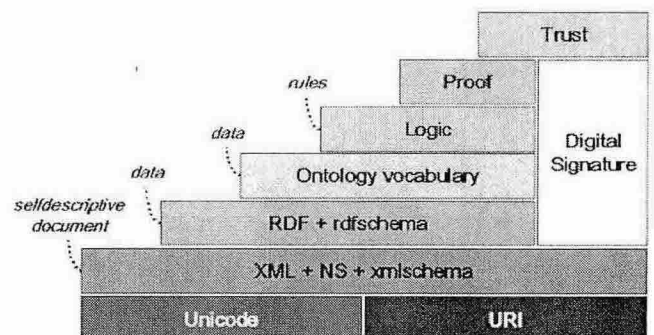


그림 1 - The Semantic Web Layers

본 연구에서는 웹 상에 존재하는 지식을 이용하여 상품 정보를 검색할 수 있고, 비즈니스 규칙에 의해 배송비를 고려한 가격 비교가 가능한 지능형 시맨틱 웹 규칙 언어 기반 상품 정보 검색 에이전트를 개발하고자 한다.

기존의 키워드 기반 상품 검색 엔진의 검색방법은 단어의 결합으로 이루어진 어구 또는 문장을 질의로써 사용하고, 검색 엔진은 유사한 단어들의 조합을 포함하는 상품의 페이지를 검색하였다. 그러나 이러한 전통적인 검색방법은 사용자가 검색하고자 하는 상품에 대한 정확한 지식을 가지고 있다 하더라도 검색이 쉽지 않으며, 주변 지식만을 가지고 검색할 경우 검색이 매우 어렵다. 또한 기존의 비교 쇼핑물은 상품에 대한 상품 가격을 세금과 배송비까지 비교하는 서비스를 사용자에게

제공하지만, 단일 상품에 대한 비교만을 제공하고, 규칙 기반이 아닌 프로그래밍상에서 처리하기 때문에 시스템의 확장 및 갱신이 어렵다.

시맨틱 웹 환경에서 적절한 사실 및 지식이 웹 상에 존재하며, 존재하는 지식을 이용하여 상품 정보를 검색할 수 있고, 또한 그것들을 조합하여 추론을 할 경우 위에서 언급한 문제는 상당부분 해결 될 수 있을 것이다. 본 연구에서는 시맨틱 웹 환경에서 에이전트가 분산되어있는 사실과 규칙을 스스로 처리할 수 있도록 함으로써, 효율적인 검색이 가능한 시맨틱 웹 규칙 언어 기반 지능형 상품 정보 검색 에이전트를 구현하고자 한다.

이제 웹 상의 정보는 인간의 처리 한계를 초과하였기 때문에, 시맨틱 웹 규칙 언어 기반 지능형 상품 정보 검색 에이전트는 기존의 쇼핑 과정을 좀 더 능률적인 과정으로 전환시킴으로써 전자상거래에서 사용자의 만족도 증대에 많은 기여를 할 것으로 기대된다.

## 관련연구

Yahoo!<sup>1</sup>, Google<sup>2</sup> 등과 같은 이미 상업화된 유명한 검색엔진들도, 키워드 기반 검색의 한계 때문에 검색하고자 하는 항목의 키워드를 정확히 모른다면 사용자가 키워드의 키워드를 찾는 지루한 검색 작업부터 시작해야 하는 어려움이 있다.

범용 검색엔진의 한 분야인 상품 정보 검색 엔진으로는 개별 쇼핑몰 사이트에서 운영하는 자체 검색 엔진들과 정보공급자 입장에서의 상품 정보 검색 엔진들로 나누어 볼 수 있으며 정보공급자 입장에서의 대표적 상품 정보 검색 엔진들로는 MySimon<sup>3</sup>, CNET<sup>4</sup>, 그리고 Global Sources<sup>5</sup> 등을 들 수 있다. 이들 상품 정보 검색 엔진들은 보다 정교한 Wrapper 기술이나 해당 쇼핑 사이트와의 계약을 기반으로 상품 정보들을 수집하여 고객에게 검색 서비스를 제공하고 있다. 이들 상품 정보 검색 엔진 및 관련 방법론 역시 앞에서 지적한 바와 같이 검색에 있어 키워드 기반 검색을 기초로 함으로써 키워드 기반 검색이 가지는 단점들을 그대로 가지게 된다.

한편, Yahoo!Shopping<sup>6</sup>, NexTag<sup>7</sup>와 같은 기존의 가격 비교 쇼핑몰은 배송비 및 세금을 고려한 가격비교 서비스를 제공하지만, 단일 상품에 대한 비교만을 제공하고, 규칙 기반이 아닌 프로그래밍상에서

처리하기 때문에 시스템의 확장 및 갱신이 어렵다.

이러한 현재 웹의 한계를 극복하기 위해 지능적인 에이전트의 출현이 요구되었고, 이러한 에이전트가 웹을 이해할 수 있게 만들기 위하여, 시맨틱 웹이라는 분야가 출현하였다. 시맨틱 웹의 주요 연구 분야 중 하나가 온톨로지이며, 온톨로지의 활용에 대하여 크게 두 가지 연구가 있는데 하나는 온톨로지 통합에 대한 연구이고 다른 하나는 온톨로지 검색에 대한 연구이다.

온톨로지 통합 시스템으로 OBSERVER[5]가 있다. OBSERVER의 목적은 다양한 영역에서 Description Logic[6]으로 표현된 이미 존재하는 온톨로지를 통합하는 것이다. 또한 OBSERVER에서 사용자의 질의는 온톨로지 변환 시에 의미보존을 위하여 중간온톨로지 관계에 의해 다시 기술된다.

온톨로지 검색을 위한 시맨틱 웹 검색 엔진으로 Creative Commons License라는 단체가 만든 Swoogle<sup>8</sup>이 있지만 아직 활용이 초기단계이다. Swoogle은 RDF 또는 OWL로 쓰인 시맨틱 웹 문서를 위한 crawler기반의 색인과 검색 시스템이다. Swoogle은 온라인상의 시맨틱 웹 문서를 발견하고, 요약하며, 분석하고, 목록화하며 또한 웹 사용자에게 강력한 질의 인터페이스를 제공한다. 현재 Swoogle은 세가지 서비스를 제공한다. 첫째, Swoogle Search는 시맨틱 웹 문서의 URL과 클래스/프로퍼티 검색을 지원한다. 둘째, Ontology Dictionary는 발견된 시맨틱 웹 문서에 의해 정의된 클래스와 프로퍼티를 목록화한다. 셋째, Swoogle Statics는 발견된 시맨틱 웹 문서에 대한 메타데이터에 기반하여 시맨틱 웹을 기술한다. 그러나 아직 검색자원으로써 RDF 문서 340,000여 건 밖에 목록화되지 않고, 시맨틱 웹에 익숙하지 않은 사용자가 질의를 입력하기에는 어려움이 있다.

이상의 시맨틱 웹 환경을 구축하기 위한 언어로는 RDF[7], RDFS[8], OWL[9]과 같은 언어들이 표준화되었고, 시맨틱 웹에서의 규칙 처리를 위하여 OWL과 RuleML<sup>9</sup>을 결합한 SWRL[10]의 표준화가 진행 중이다. SWRL은 유사 Horn 규칙(Horn-like rules)을 포함시켜 OWL의 공리를 확장함으로써 유사 Horn 규칙을 OWL 지식 베이스와 통합시킬 수 있다.

온톨로지에 대한 질의를 처리하기 위하여 여러 가지 질의 언어가 개발 중이다. RDF를 위한 질의 언어인 SPARQL[11]이 W3C에서 working draft상태이고, RDF/RDFS에 대한 질의가 가능하고 OWL에 대해서도 부분적인 질의가 가능한 RDQL[12]은 W3C에 제안되었으며, OWL 전용의 OWL-QL[13]의 개발은 아직 초기 단계에 있다.

<sup>1</sup> <http://www.yahoo.com>

<sup>2</sup> <http://www.google.com>

<sup>3</sup> <http://www.mysimon.com>

<sup>4</sup> <http://www.search.com>

<sup>5</sup> <http://www.globalsources.com>

<sup>6</sup> <http://shopping.yahoo.com>

<sup>7</sup> <http://www.nextag.com>

<sup>8</sup> <http://swoogle.umbc.edu/>

<sup>9</sup> <http://www.ruleml.org>

## 주요요소

시맨틱 웹 규칙 언어 기반 지능형 상품 검색 에이전트는 *query generator*, *SWRL to Jena rule translator*, *Jena generic purpose rule reasoner*, *ontology search engine*의 4가지 주요 요소로 구성되며, 각 요소의 역할은 다음과 같다.

*query generator*는 사용자가 입력한 키워드를 질의 언어로 변환하는 역할을 한다.

*SWRL to Jena rule translator*는 SWRL 형식의 규칙을 Jena 형식 규칙으로 변환하는 역할을 한다. 현재 Stanford에서 제공하는 온톨로지 편집기인 protege<sup>10</sup>에 내장된 SWRL 편집기는 규칙 편집기능만을 가지고 있고 별도의 추론기능은 제공하지 않는다. 그림 2의 (a)는 a가 b에 대하여 p의 속성을 가지며, b가 c에 대하여 q의 속성을 가질 때, a가 c에 대하여 r의 속성을 가진 규칙을 SWRL형식의 정방향 추론으로 나타낸 것이고, (b)는 이 SWRL 규칙을 Jena형식으로 변환한 결과이다.

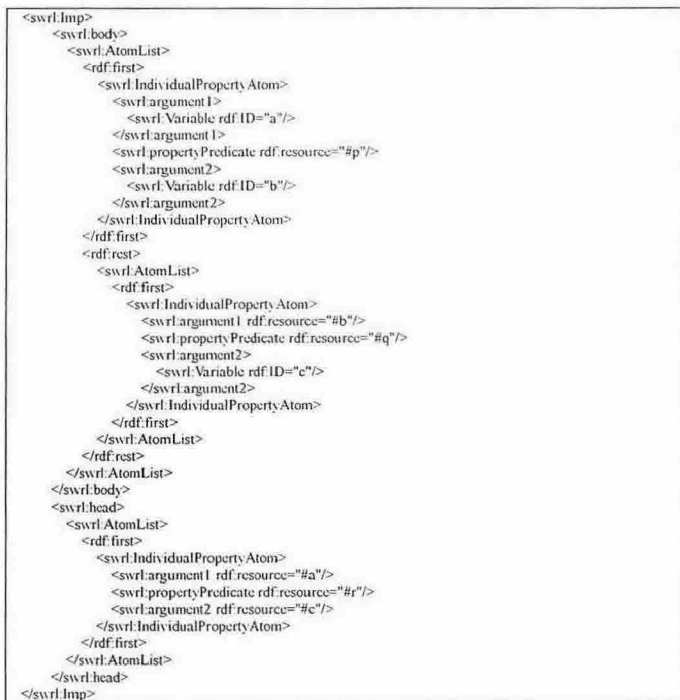


그림 2 (a) Rule in SWRL form

```

|r: (? r concatFirst ?p). (? r concatSecond ?q) -> | rb: (?a ?p ?b) (?b ?q ?c) -> (?a ?r ?c) || ".
  
```

그림 2 (b) Rule in Jena form

SWRL은 정방향 추론만으로 처리될 수 있으며, *Jena generic purpose rule reasoner*는 정방향 추론 및 역방향, 혼합 추론을 모두 처리할 수 있어 SWRL로 표현된 규칙을 수정하면 수용할 수 있다.

*ontology search engine*은 RDQL 기반으로 온톨로지를

검색하는 기능을 가진다. RDQL은 설계 시부터 RDF의 자료 구조를 기반으로 만들어졌기 때문에 RDF의 확장판이라고 할 수 있는 OWL에도 매우 효과적인 질의 언어이다.

우리는 이상의 방법론과 요소를 기반으로 시맨틱 웹 규칙 기반 지능형 상품 정보 검색 에이전트를 구현하였으며 *CDSearch*라고 명명하였다. 다음 장에서는 Elvis Presley를 장인으로 가지는 가수의 음반을 쇼핑몰 별로 비교하여 구매하는 *CDSearch*의 서비스 시나리오를 설명하겠다.

## 서비스 시나리오

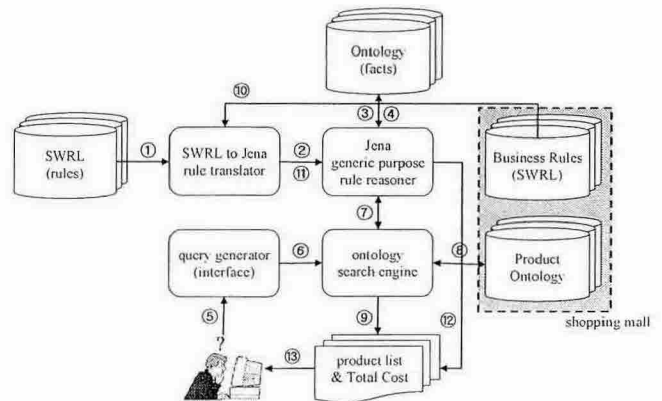


그림 3 - Scenario Architecture

사용자가 Elvis Presley를 장인으로 가진 사람의 음반을 구매하려고 한다. 웹에는 가족관계의 규칙을 나타내는 SWRL 및 사용자가 아직 모르는 사실(Mihael Jackson의 아내가 Lisa Marie Presley이며, Lisa Marie Presley의 아버지가 Elvis Presley 이다)들이 존재한다. *CDSearch*는 이러한 사실(facts)을 기반으로 추론을 실시하여 사용자에게 가수의 이름을 알려주고, 그 가수의 이름으로 CD를 검색하며, 또한 사용자가 결과목록 중에서 CD를 선택하고 희망하는 배송기간을 입력하면 배송비까지 고려한 가격을 쇼핑몰 별로 비교해 사용자에게 보여준다. (그림 3)

①~④는 SWRL로 표현된 규칙이 Jena형식으로 변환되어 온톨로지로 표현된 사실들과 함께 추론된 후, 추론결과가 온톨로지에 갱신되는 과정이고, ⑤~⑧는 상품을 검색하는 과정이며, ⑨~⑬는 배송규칙이 처리되어 최종적으로 사용자에게 보여지는 과정이다.

## 온톨로지 갱신

SWRL로 표현된 규칙은 *Jena generic purpose rule reasoner*에서 그대로 사용될 수 없으므로, *SWRL to Jena rule translator*에서 RDQL로 SWRL의 rule name, property predicate(variable간의 관계를 정의), variable 등의 요소를 RDQL로 추출하여 Jena 형식으로

<sup>10</sup> <http://protege.stanford.edu/>

변환하여 전송하는 단계가 ①, ②이다. 예제에서는 아내(has-Wife)와 아버지(has-Father)관계의 결합으로 표현될 수 있는 장인(father-in-law)에 대한 규칙 처리를 다루려고 한다.

단계 ③에서는 장인관계를 나타내는 규칙들과, 온톨로지에 있는 “Mihael Jackson의 아내가 Lisa Marie Presley이며, Lisa Marie Presley의 아버지가 Elvis Presley 이다”라는 사실들이 **Jena generic purpose rule reasoner**로 전송되어 처리된다. 추론 결과는 Elvis Presley를 장인으로 가지는 사람이 Michael Jackson임을 보여준다.

④는 Michael Jackson이 Elvis Presley를 장인으로 가진다는 새로운 사실이 온톨로지에 갱신되는 과정이다. 기존의 온톨로지가 내부에 있는 사실들과 어떤 웹 상의 규칙에 의해 추론된 새로운 사실들로 갱신되는 모형은 그림 4와 같다.

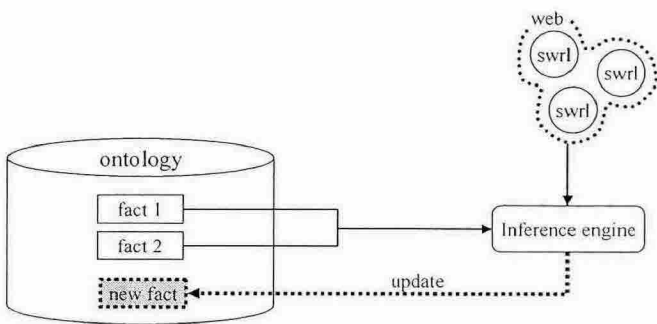


그림 - 4 Ontology Update

상품검색

⑤에서는 사용자가 **query generator**를 통해 찾고자 하는 물품에 대한 주변지식을 입력한다. 이때 Elvis Presley를 장인(father-in-law)로 가지는 가수를 GUI상의 singer 항목에 [hasFatherInLaw Elvis\_Presley]라는 형식으로 입력한다. 자연어 처리도 인공지능 분야에서 주요 분야 중의 하나이지만 본 연구에서는 이러한 자연어 처리에 중점을 두지 않겠다. 그러므로 키워드 처리는 [property value]의 나열로써 극히 간단하게 처리하고자 한다.

⑥에서 **query generator**는 사용자가 입력한 내용을 RDQL 형식으로 변환하여 reasoner에 입력한다. 그림 5의 RDQL은 Elvis Presley를 장인으로 갖는 가수의 음반을 검색하는 내용이다.

```

1) SELECT ?y
2) WHERE (?x, elvis:hasFatherInLaw, "Elvis_Presley")
3)      (?y, shoppingmall:singer, ?x)
4) USING elvis FOR http://www.elvis.com/
5)      shoppingmall FOR http://www.shoppingmall.com/
    
```

그림 5 - RDQL

이러한 질의문을 가지고 ⑦에서 가수가 Michael Jackson으로 도출되고, 이를 기반으로 ⑧에서는 쇼핑몰의 온톨로지를 검색해 Michael Jackson의 음반을 검색한다. 그림에서는 분리되어 있지만 사실 온톨로지(Elvis Presley의 가족관계)와 상품 온톨로지(shopping mall)는 통합되어 있다.

배송규칙 처리

⑨에서 검색결과(음반목록)가 사용자에게 보여지며, 사용자는 음반을 선택하고, 희망 배송일을 입력한다. 예제에서는 배송기간을 2일로 입력하였으며, 배송지별 차이는 고려하지 않았다.

한편, 쇼핑몰은 자체 배달 가능한 쇼핑몰과 자체 배달이 불가능 하여 외부 배달업체를 사용하는 쇼핑몰이 있다. 예제에서는 Amazon.com은 자체 배송 서비스를 보유한 쇼핑몰로, Barnes&Noble는 자체 배송 서비스가 없어 전문 배송 업체인 UPS의 배송 서비스를 사용하는 것으로 구분하였다. ⑨는 Amazon.com과 UPS의 배송 관련 규칙들이 Jena translator로 전송되는 과정이다.

그러나, SWRL을 Jena 규칙 형태로 변환함에 있어 두 가지 한계점이 있다. 첫째, SWRL은 swrl:equal, swrl:greaterThanOrEqual과 같은 내장함수를 가지고 있으나, 현재 SWRL 내장함수를 지원하는 추론엔진이 없다. 한편 Jena에서는 이와 동일한 기능을 하는 equal(?x, ?y), ge(?x, ?y)와 같은 내장함수를 가지고 있다. 그래서 SWRL 규칙 생성시 equal, ge와 같은 property predicate을 새로 정의해서 사용하겠다. 이 Jena 내장함수로 SWRL 규칙을 정의할 때 장점은, 규칙 변환 후 Jena 프레임워크에서 바로 규칙을 사용할 수 있다는 점이다. 둘째, 현재 SWRL editor가 variable의 데이터 타입으로 float type(8.3, 5.05)을 허용하지 않아 배송업체의 SWRL은 편집기를 사용하지 않고 xml형태로 직접 편집하였다. 그림 6는 각 배송업체의 배송기간별 요금을 나타내고, 그림 7은 UPS의 배송기간별 요금 중 배송기간이 2일이면 배송료는 6.48 달러임을 SWRL로 표현한 것이다.

Amazon.com		UPS	
Period	Rates	Period	Rates
1day	\$ 14	1day	\$ 11.98
2day	\$ 8.3	2day	\$ 6.48
3day	\$ 6	3day and more	\$ 2.98
4day and more	\$ 5.05	-	-

그림 6 - Shipments Period and Rates

```

<swrl:Imp rdf:ID="Def-2DayShippingRates">
  <swrl:body>
    <swrl:AtomList>
      <rdf:rest
    
```



```

rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
  <rdf:first>
    <swrl:IndividualPropertyAtom>
      <swrl:propertyPredicate rdf:resource="#equal"/>
      <swrl:argument2>
        <swrl:Variable rdf:resource="#2"/>
      </swrl:argument2>
      <swrl:argument1 rdf:resource="#shippingPeriod"/>
    </swrl:IndividualPropertyAtom>
  </rdf:first>
</swrl:AtomList>
</swrl:body>
<swrl:head>
  <swrl:AtomList>
    <rdf:first>
      <swrl:IndividualPropertyAtom>
        <swrl:argument1 rdf:resource="#shippingRates"/>
        <swrl:propertyPredicate rdf:resource="#equal"/>
        <swrl:argument2 rdf:resource="#8.3"/>
      </swrl:IndividualPropertyAtom>
    </rdf:first>
    <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
  </swrl:AtomList>
</swrl:head>
</swrl:Imp

```

그림 7 - UPS' Shipping Rule in SWRL

⑩, ⑪에서는 이러한 SWRL 규칙들을 **SWRL to Jena rule translator**로 적절한 형식으로 변환하여 다시 **Jena generic purpose rule reasoner**에 전송하는 것을 나타낸다.

⑫, ⑬은 **Jena generic purpose rule reasoner**가 각 쇼핑물의 상품에 대하여 사용자가 입력한 배송기간을 기준으로 배송료를 추론한 후 총 가격(제품가격+배달가격)의 목록을 사용자에게 보여주는 과정을 나타낸다. 그림 8은 사용자에게 최종적으로 보여지는 결과를 나타낸다.

You selected	Shopping Mall	Product Price	Shipping Rates	Total Price
Michael Jackson's Thriller And Shipping Period is 2 days	Barnes&Noble	\$20.49	\$8	\$28.49
	Amazon.com	\$9.99	\$7	\$16.99

그림 8 - Inference Results

## 결 론

본 연구에서는 웹에서 1) SWRL로 표현된 규칙과 OWL로 표현된 사실을 근거로 새로운 사실을 검색하고, 2) SWRL로 표현된 규칙을 이용하여 비즈니스 절차를 실행하는 시맨틱 웹 규칙 언어 기반 검색 에이전트를 구현하여, 전자상거래에 있어 사용자의 만족도를 높이려고 하였다.

표준화가 진행중인 SWRL과, SWRL과 SWRL의 모든 내장함수(built-ins)를 처리할 수 있는 추론 엔진 개발에 대한 연구는 시맨틱 웹을 한 단계 도약시킬 수 있는 중요한 연구이고, 그림 1에서 시맨틱 웹의 최상위 계층인 Proof와 Trust단계의 근간이 될 것이다.

## 참고문헌

[1] Tim Berners-Lee, James Hendler, Ora Lassila. (2001). "The Semantic Web" *Scientific American*. 2001.

[2] Marja-Riitta Koivunen and Eric Miller. (2001). "W3C Semantic Web Activity," *Proceedings of the Semantic Web Kick-off Seminar in Finland*

[3] Eric Miller, (2004). "Semantic Web, Phase 2: Developments and Deployment," *Presented at WWW2004 W3C Track on the Semantic Web*

[4] John A. Miller, Gregory T. Baramidze, Amit P. Sheth, Paul A. Fishwick. (2004). "Investigating Ontologies for Simulation Modeling," *Proceedings of the 37th Annual Simulation Symposium (ANSS'04)*

[5] E. Mena, V. Kashyap, A. Sheth, A. Illarramendi. (1996). "OBSERVER: An Approach for Query Processing in Global Information Systems based on Interoperation across Pre-existing Ontologies

[6] Nardi, D., Brachman, R. J. (2002). "Description Logics Handbook," *Cambridge University Press*

[7] Frank Manola, Eric Miller. (2003). "RDF Primer," *W3C Recommendation*

[8] Dan Brickley, R.V. Guha, Brian McBride. (2004). "RDF Vocabulary Description Language 1.0: RDF Schema," *W3C Recommendation*

[9] Deborah L. McGuinness and Frank van Harmelen eds. (2004). "OWL Web Ontology Language Overview," *W3C Recommendation*

[10] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Groszof, Mike Dean. (2004). "SWRL: A Semantic Web Rule Language Combining OWL and RuleML," *W3C Member Submission*

[11] Eric Prud'hommeaux, Andy Seaborne. (2005). "SPARQL Query Language for RDF," *W3C Working Draft*

[12] Andy Seaborne, (2004). "RDQL - A Query Language for RDF," *W3C Member Submission*

[13] Richard Fikes, Patrick Hayes, and Ian Horrocks. (2003) "OWL-QL - A Language for Deductive Query Answering on the Semantic Web," *Technical report, Knowledge Systems Laboratory, Stanford University, Stanford, CA, KSL-03-14.*