

# 시맨틱 웹 기반 와인 지식 검색을 위한 웹 서비스 설계

## Framework Design for Wine Knowledge-based Semantic Web Services

전현주, 윤호창\*, 최광웅\*\*

이화여대 정보 과학 대학원, SCCA\*, HCCA\*\*

Jeon Hyun-Joo, Youn Ho-Chang\*, Choi Gwang-Ung\*\*

Ewha Graduate School of Information Science., SCCA\*, HCCA\*\*

### 요약

Well-Being과 관련하여 삶의 질과 관련된 관심이 증가하면서 와인의 관심과 수요가 증가하고 있다. 이와 같은 시기에 와인의 종류 또는 와인인과 어울리는 음식과 같은 여러 가지 지식에 관한 서비스를 온톨로지를 이용하여 사용자가 와인에 관한 정보를 보다 효율적으로 얻도록 하는 것이 필요 된다. 본 연구에서는 와인 온톨로지를 기반으로 시맨틱웹 기술을 활용한 와인 지식 검색 서비스 설계를 제안한다.

### Abstract

As the well-being or quality of life of a population is the common interests, a lot of people are interested in wines. They are willingness to share wine knowledges with other wine experts on the web. Therefore the study for information retrieval system and inference engines are needed to get relevant search results about wine types and suitable wines for given foods. This paper discusses an approach to the architecture of agent-based semantic web services in wine ontology.

## I. 서론

최근에 well-being과 관련하여 삶의 질이 중요시 되면서 와인에 대한 관심과 수요가 증가하고 있다. 인터넷 환경의 발달로 웹의 사용이 일반화됨에 따라 사용자들은 와인에 관한 지식을 웹을 통하여 쉽게 접근하며 습득한 와인 정보를 커뮤니티에서 공유한다.

하지만 현재의 웹 환경은 컴퓨터를 매개로 사람과 사람들간의 정보교환을 위해 발달되어 있어 방대한 양의 정보 추출을 사람에게 의존하고 있기 때문에 원하는 검색결과를 얻는 것이 쉽지 않다. HTML로 이루어진 웹 문서는 부울연산에 의한 구문(syntax) 검색 방법을 사용하고 있기 때문에 유의어, 동음이의어와 같은 문제를 해결하지 못한다. XML은 정보의 정

확한 표현은 가능하지만 의미(semantic) 표현이 불가능하고 다양한 형태의 표현 가능성으로 인해 정보의 애매성을 증가시키고 있다. 지금까지 웹 기술은 사람과 컴퓨터간의 정보교환에 치중하여 컴퓨터와 웹 데이터간의 의미적 해석 및 처리에 문제점을 지니고 있었다. 시맨틱 웹은 이러한 문제를 해결하기 위하여 웹 문서에 의미를 덧붙이고 이를 이용한 의미 추출을 가능하도록 한다.

본 연구에서는 사용자에게 맞는 와인 지식을 제공하기 위해서 데이터베이스를 통하여 의미 있는 지식 표현이 가능하도록 시맨틱 웹 기술을 와인에 적용해보고자 한다. II장에서는 시맨틱웹의 개념과 시맨틱 웹 기술을 이용한 와인 에이전트에 관한 기존의 연구를 살펴본다. III장에서는 시맨틱웹 구현의 핵심 요소로

서 온톨로지에 관해 고찰하고 와인 분류 기준을 제시한다. IV장에서는 시맨틱 웹 기반 와인 검색 시스템 설계 및 구현 도구에 관하여 논의하고 마지막으로 V장에서는 시맨틱 웹 기술을 이용한 와인 지식 검색 서비스 설계에 관하여 제안한다.

## II. 시맨틱 웹

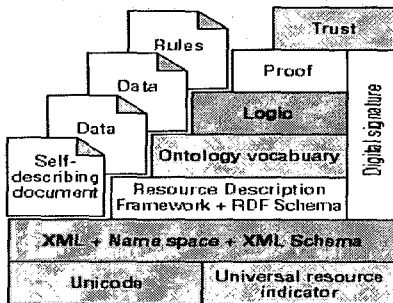
### 2.1 시맨틱 웹 개념

시맨틱웹은 간단히 말하면 '컴퓨터가 정보의 의미를 이해하고 처리할 수 있는 웹'이라 할 수 있다. 시맨틱 웹은 기계가 처리할 수 있는 언어로 표현된 정보 공간이며, 형식적이고 의미적으로 연결된 정보의 웹이다. 형식적이라는 것은 기계가 읽고 처리함으로써 문서의 내용을 자동으로 표현할 수 있음을 의미한다.

시맨틱 웹은 웹 온톨로지를 기반으로 웹상에 존재하는 자원들에 대한 의미 마킹을 하고 검색 에이전트를 이용하여 의미 명세에 대한 추론을 통해 필요로 하는 정보를 보다 정확하게 찾아내는 것을 가능하게 해준다.

### 2.2 시맨틱 웹 계층 구조

월드와이드웹을 처음 도입한 Tim Berners-Lee는 그림 1과 같이 시맨틱 웹의 계층구조를 제안하였다.



▶▶ 그림 1. 시맨틱 웹의 계층 구조

이 계층 구조에서 보면 가장 하위 레벨에서 웹 프로토콜에서 자원을 지칭하기 위한 주소지정(address) 방법인 URI가 뒷받침되고 이를 기반으로 XML과 Namespace, RDF와 RDF스키마, 온톨로지의 순서로 연구가 진행되고 있으며 그 위의 계층인 Logic에 대해서는 인공지능의 추론연구를 밀받침으로 일부 연가가 시작되었다. 또한 보다 더 상위계층인 Proof와 Trust는 시맨틱 웹 정보의 신뢰성과 보안에 대한 내용으로서 아직 개념 정도만 얘기되고 있으며 차후 연구과제로 제시되고 있다.

### 2.3 시맨틱 웹 적용 분야

시맨틱 웹의 응용은 에이전트 기반의 웹 서비스 제공과 같은 유용한 응용 프로그램 개발로 요약될 수 있다. 대표적인 Annotation, authoring, MusicBrainz, ITTalks의 지능형 플랫폼이 요구되는 e-비즈니스 분야, 고객 관리 분야, 바이오 정보 분야, 의료 분야에서 시맨틱 웹을 이용한 응용 서비스 개발에 연구가 이루어지고 있다.

### 2.4 와인 에이전트 선행 연구

와인 에이전트는 주어진 음식과 어울리는 와인을 제공하기 위하여 도메인 온톨로지를 사용하는 데모 시스템이 구현되었다. 와인 에이전트를 사용하여 논리적으로 도메인을 설명하고 최근 시맨틱 웹 기술의 처리, 추론, 증명, 실행을 제공하고 있다. 이 에이전트는 DAML+OIL과 JTP 검증기로 OWL 웹 기반 설명 논리를 조합한다. 결과로서의 지식 베이스는 DQL 질의 언어를 사용하여 원격으로 질의된다. 적절한 페어링은 추론 웹 장치에서 설명되어지고 웹 서비스의 예비 구현을 통해 처리될 수 있다. 이러한 방법론을 위한 원형으로서 제공하기 위해 와인 에이전트는 시맨틱 웹 공리에 의한 추론, 필요되어지는 언어, 설명 요건, 구현과 통합을 고려한 실용적인 이슈와 관련하여 연구되어 왔다.

### III. 와인 온톨로지

#### 3.1 온톨로지 정의

온톨로지라는 말은 희랍어 'ontos(being)'와 'logos(word)'에 기인한다. 이는 원래 철학, 특히 형이상학의 한 분야로, 이 세계에 존재(being)하는 것들의 종류, 그 본성과 관계 등에 대한 연구나 학문을 지칭하는 말이다.

이 용어가 전산학 문헌에서 처음으로 등장한 것은 1967년 S.H. Mealy에 의해서라고 여겨진다. 이후 이 온톨로지는 지난 수십 년간 특히 인공지능 분야에서 지식표현(knowledge representation)에서 사용되고 있다.

Gruber는 "온톨로지란 공유된 개념화의 형식적이고 분명한 명세"로 정의한다. 여기서 개념화란 어떤 목적으로 표현하고자 하는 대상을 추상화하고 단순화시킨 것이다. 형식적이란 규정된 용어들과 그들 사이의 관계를 컴퓨터가 이해할 수 있는 방법으로 표시하는 것이다.

구조적으로 온톨로지는 분류적인 계층구조(taxonomic hierarchies)로 여겨지기도 한다.

철학적으로 규정하기도 어려운 (상위) 범주이지만 인공지능에서 많은 온톨로지 체계에서는 'Thing', 'All' 등의 범주에서 그 하위범주로 개념을 규정해 나가고 있다.

#### 3.2 와인의 종류 및 분류

와인의 종류를 가르는 방법에는 몇 가지 기준이 있다. 와인은 생산지를 기준으로 미국 와인, 이탈리아 와인, 프랑스 와인, 오스트레일리아 인, 독일 와인 등으로 분류된다. 또한 색상, 당도, 바디, 와인의 무게, 용도, 식사에 따라 분류될 수 있다. 표 1은 한국음식에 어울리는 와인을 분류한 표이다. 표에 나오는 와인은 국내에 수입이 된 와인이며 유명한 와인을 중심으로 예를 들었다.

[표 1] 한국음식과 조화를 이루는 와인

음식	추천 와인
Meat(육류)	
Steamed short ribs(갈비찜)	Clos Vougeot Musiny Gros Frere&Soeur
Roast meat(불고기)	Baron Philippe Saint-Emilion
Sweet and sour pork(탕수육)	Gewurztraminer Heimboung
Seafood(해산물)	
Sliced raw fish (생선회)	William FèvreChablis Premier Cru
Fresh-water eel (민물장어)	Le Pin
Fowl(조류)	
Chicken fritters (닭 튀김)	Chateau Bire, Bordeaux Superieur
Roasted Chicken (구운 닭)	Beaujolais Village
Dessert(디저트)	
Cake, Chocolate	Tawny Port(10yrs)
Fruit	Bollinger Special Cuvee Brut

#### 3.3 온톨로지 기술 언어

시맨틱웹에서 형식적이고 의미적인 정보를 표현하기 위해 W3C는 수년에 걸쳐 온톨로지 언어의 개발에 노력해 왔다. W3C에 의해 개발된 RDF와 RDF 스키마는 웹 자원의 메타 데이터를 기술 할 수 있는 표준이지만, 속성의 제약과 클래스의 상속 관계를 표현하는데 제한이 있었다. 최근 W3C의 권고안으로 발표된 OWL(Web Ontology Language)는 웹에서 온톨로지를 기술하기 위해 RDF와 RDF 스키마보다 풍부한 표현력을 지원하기 위한 많은 특성들을 포함하고 있다. OWL은 미국 주도로 개발된 DAML과 유럽 공동체에 의해 개발된 OIL을 기반으로 확장된 언어이며, 두 언어를 통합한 DAML+OIL의 많은 특징을 계승하고 있다. DAML+OIL과 마찬가지로 OWL도 RDF 스키마의 상위 수준에서 어휘의 의미를 정의하고 있으며 RDF의 클래스와 속성을 이용한다.

### IV. 시맨틱웹 기반 와인 검색시스템 구조

#### 4.1 와인 검색을 위한 온톨로지 설계

본 논문에서는 Protege-2000을 사용하여 와인 검색

색을 위한 온톨로지를 설계하였다.

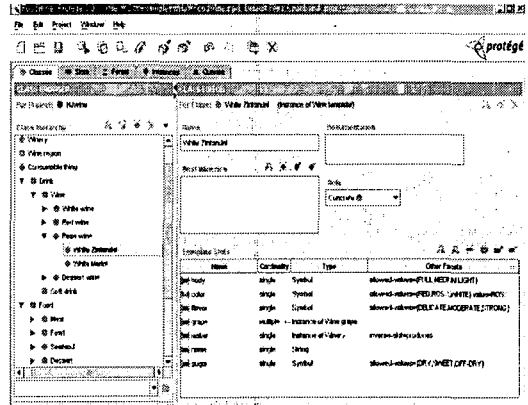
- ☉ Winery
- ☉ Wine region
  - ▶ ☉ USA
  - ▶ ☉ Italia
  - ▶ ☉ France
  - ▶ ☉ Austrailia
  - ▶ ☉ Germany
- ☉ Consumable thing
  - ▼ ☉ Drink
    - ▼ ☉ Wine
      - ▶ ☉ White wine
        - ▶ ☉ Red wine
        - ▼ ☉ Rose wine
          - ☉ White Zinfandel
          - ☉ White Merlot
        - ▶ ☉ Dessert wine
      - ▶ ☉ Soft drink
    - ▼ ☉ Food
      - ▶ ☉ Meat
      - ▶ ☉ Fowl
      - ▶ ☉ Seafood
      - ▶ ☉ Dessert

▶▶ 그림 2. 와인 온톨로지

와인 도메인에서 사용되는 용어들을 정의하고 그들 사이의 관계를 규정하여 모델링을 한다. 이 과정은 개념들의 클래스화, 그 개념들의 관계를 클래스의 계층적인 구조로 정립, 클래스들의 속성과 그 속성에 존재하는 다양성과 제한 요소 정의, 마지막으로 인스턴스들의 생성을 포함한다.

### 4.2 온톨로지 에디터

현재 온톨로지는 수동으로 구축되며 추가, 삭제, 갱신 등에 관련된 온톨로지 관리는 온톨로지 에디터 Protege-2000을 통해 가능하다. 그림 3은 Protege-2000 상에서 와인 온톨로지의 입력화면이다. 좌측은 와인 클래스의 계층적 구조를 표현해 주며 우측은 와인 클래스의 속성을 정의해주고 있다.



▶▶ 그림 3. 와인 온톨로지 입력 화면

### 4.3 온톨로지 기술 언어 OWL

개발된 온톨로지는 OWL을 이용하여 작성될 수 있다. OWL(Web Ontology Language)구문 구조는 키워드 검색의 범위를 넘어서는 기능으로 웹 자원의 의미 기술을 전체 조건으로 하는 것이다. OWL은 웹 온톨로지와 연관된 지식창고(knowledge base), 즉 추론시스템에 축적된 논리적 명제(assertion)들을 정의하는 언어로서 이러한 명제는 하나의 온톨로지에 근거하거나 또는 OWL이 명시하는 방식에 따라 수집하거나 다수의 분산형 온톨로지에 근거할 수 있다. OWL명제는 클래스의 구성원(members)들에 관한 사실과 구성원들 간의 관계를 기술하며, 이러한 명제의 집합으로 이루어진 온톨로지는 구문적으로 정의되지 않은 사실의 논리적 유추를 가능하게 한다. OWL 온톨로지는 클래스(class)와 속성(properties) 및 이에 적용할 수 있는 제약사항(constraints)들을 포함한다.

그림 4는 Protege-2000상에서 입력된 와인 온톨로지를 기술한 OWL 문서의 일부분을 나타내고 있다.

```

</rdfs:subClassOf>
<rdfs:label>Pinot Blanc</rdfs:label>
</owl:Class>
<owl:Class rdf:about="#Red_Bordeaux">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Red_wine"/>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:FunctionalProperty rdf:about="#color"/>
      </owl:onProperty>
      <owl:hasValue rdf:datatype="http://www.w3.org/2001/XMLSchema:string">
        >RED</owl:hasValue>
      </owl:Restriction>
    </rdfs:subClassOf>
  <rdfs:label>Red Bordeaux</rdfs:label>
</owl:Class>
<owl:Class rdf:ID="Winery"/>

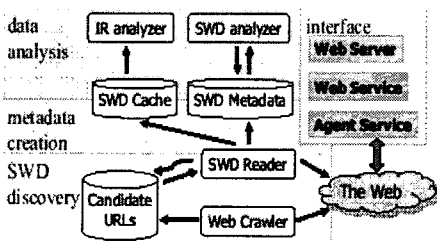
```

▶▶ 그림 4. OWL을 이용해 기술된 온톨로지

#### 4.4 시맨틱 웹 검색엔진 Swoogle

Swoogle은 시맨틱 웹에서 크롤러 기반 인덱싱 검색 시스템(crawler-based indexing and retrieval system)이다. Swoogle은 발견한 문서의 메타데이터를 추출하고 문서 사이의 관계를 계산한다. 발견된 문서는 문서 사이에서의 유사성을 측정하고 관련된 문서를 찾기 위해 N-Gram 또는 URIs를 사용하는 정보 검색 시스템에 의해서 인덱스 된다. Swoogle의 주된 특성은 시맨틱 웹 문서의 중요성을 측정하도록 온톨로지 랭크(ontology rank)를 사용한다는 점이다. Swoogle은 특정 업무에 대한 적절한 온톨로지 발견, 인스턴스 데이터 탐색, 계층적 속성을 보여주는 기능을 포함한다.

그림 5는 Swoogle의 4가지 주요 요소에 대한 아키텍처를 보여준다.

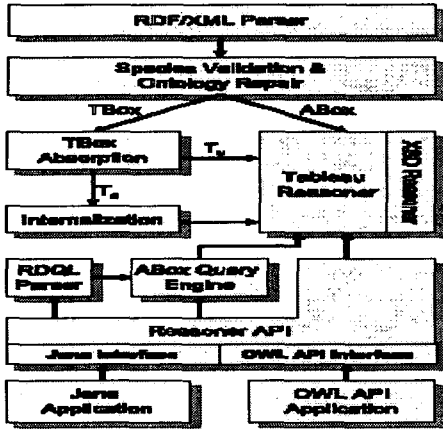


▶▶ 그림 5. Swoogle 아키텍처

SWD discovery은 전 웹에 있는 잠재적 SWDs(Semantic Web datas)을 발견하며 SWDs에 관한 최신 정보를 유지한다. Metadata creation은 SWD 스냅샷을 저장하고 구문(syntax) 레벨과 의미(semantic) 레벨에서의 SWD에 관한 메타데이터를 생성한다. Data analysis에서는 저장된 SWDs와 SWOs, SWDBs, SWDs의 랭크, SWDs의 IR 인덱스의 분류와 같은 분석 레포트를 도출하기 위해 새로 생성한 메타데이터를 사용한다. Interface에서는 시맨틱 웹 커뮤니티에 데이터 서비스를 제공한다.

#### 4.5 Description Logic Pellet Reasoner

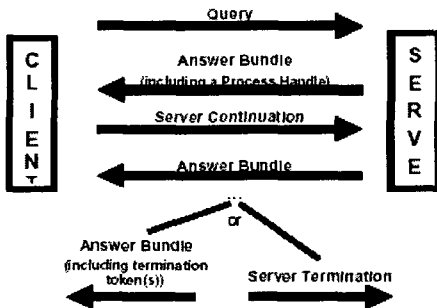
추론 기능은 시맨틱 웹 개발에 중요한 역할을 한다. Pellet은 Description logic(DL)을 위해 개발된 tableaux 알고리즘에 기반을 두어 OWL이 포함된 모든 OWL DL 구조를 지원하고 있다. 그림 6은 Pellet reasoner의 주요 요소를 보여준다. OWL 온톨로지는 RDF 트리플로 파싱되면 Pellet은 트리플이 knowledge base에서 공리로 변환되는 동안 온톨로지 종류를 검증한다. 만약에 온톨로지 레벨이 잃어버린 타입 트리플로 인해 OWL로만 채워지게 되면 온톨로지를 수정하기 위해 heuristics를 사용한다. Pellet은 TBox에 클래스에 관한 공리를 저장하고 ABox 요소의 개체에 관한 공리를 저장한다. TBox 분할, 흡수, 조립식 최적화 구현된다. Tableau reasoner는 표준 standard tableau 규칙을 사용하고 dependency directed backjumping, semantic branching, blocking strategies와 같은 표준적인 다양한 최적화를 포함한다. Pellet은 Java를 이용해 구현된다.



▶▶ 그림 6. Pellet Reasoner 아키텍처

#### 4.6 질의 응답 OWL-QL

OWL Query Language(OWL-QL)은 시맨틱 웹 환경에서 에이전트가 OWL로 표현된 지식베이스를 이용하여 질의응답대화를 수행할 수 있는 질의언어이며 프로토폴이다. 스탠포드 대학에서는 시맨틱 웹 서비스에 OWL-QL을 사용하기 위해 XML 스키마를 정의했다. OWL-QL은 크게 Queries 부분과 Answers부분으로 나눈다. 이 중에서 정보검색 서비스에 맞게 변형할 부분은 Queries부분이며, Answers부분에서는 단순히 사용자 질의에 대한 답을 제공해준다. 질의응답 대화의 전체 구조는 그림 7과 같다.

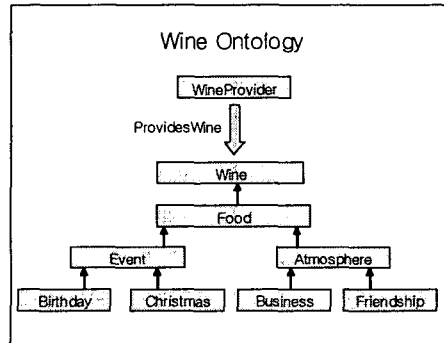


▶▶ 그림 7. OWL-QL Query-Answering

## V. 와인 지식 웹 서비스 설계

### 5.1 시맨틱 웹 시나리오

와인 도메인에서의 시맨틱 웹 시나리오를 소개한다. 기본적인 아이디어는 사용자가 시맨틱 웹에서 음식과 같이 마실 수 있는 와인을 선택하고자 할 때 와인 에이전트가 여러 가지 다양한 와인을 적절하게 추천해 주는 것이다. 와인 에이전트는 특별한 날 기념이나 연회 행사에 따라 정해진 음식과 조화될 수 있는 와인을 정할 수 있게 하며 같이 마시게 와인 메이트와의 비즈니스, 사교 분위기에 따라 와인을 선택할 수 있게 한다. 그림 8에서 보는 바와 같이 에이전트를 위한 시맨틱 웹 구조는 온톨로지에 기반을 두고 있다.



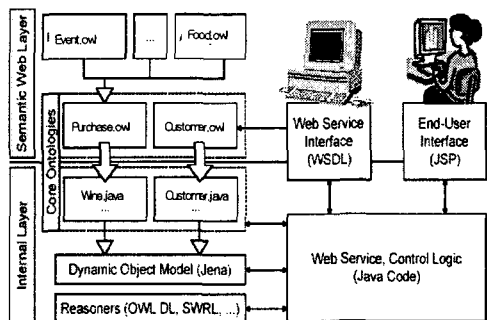
▶▶ 그림 8. 시맨틱 웹 시나리오

와인 온톨로지는 이벤트와 분위기에 관한 메타데이터를 제공할 수 있도록 하고 있다. 와인 제공자는 클래스를 온톨로지로부터 설명할 것이고 웹 사이트의 OWL 파일로 결과 객체를 표현할 것이다. 시맨틱 웹 서비스는 와인을 마시는 환경에 따라 특별하게 되며 가능한 와인을 수집할 수 있도록 크롤러 에이전트를 보내게 된다. 만약에 사용자가 친구 생일에 초콜렛 케이크와 같이 마실 수 있는 와인을 원한다면 에이전트는 적절한 매치가 되는 와인을 찾기 위해 온톨로지 계층 구조를 탐색할 것이다.

### 5.2 시맨틱 웹 애플리케이션 구조

시맨틱 웹 시나리오에서 설명한 온톨로지는 다양한

시맨틱 웹 애플리케이션으로 개발될 수 있다. 그림 9는 소비자에게 특별한 날 음식과 매치되는 와인을 찾을 수 있는 애플리케이션 구조를 제시한다.



▶▶ 그림 9. 시맨틱 웹 애플리케이션 구조

애플리케이션의 기능은 소프트웨어 에이전트가 웹 서비스 인터페이스를 통해 가능하도록 하며 최종 소비자가 웹 브라우저 인터페이스를 통해 이용 가능하도록 한다. 입력과 출력데이터 구조는 OWL 온톨로지를 통해 표시되기 때문에 외부 에이전트도 정확하게 서비스를 사용할 수 있다. 애플리케이션 logic은 Java와 같은 객체 지향 언어로 구현한다.

## VI. 결론

본 논문은 웹 환경에서 시맨틱 웹의 와인 온톨로지를 이용하여 지식 처리를 위한 모델을 설계 하였다. 이는 서양 중심의 와인 문화를 한국 음식과 조화시켜 최근의 온톨로지 언어와 강력한 추론엔진을 사용하여 국내 상황에 맞는 와인 에이전트를 설계하여 지금까지의 연구 결과와는 차별화 된 것이다.

온톨로지 의미 구조와 계층 구조를 활용한 추론을 통해 지식 검색을 설계하여 검색 능력을 향상시키고자 하였다. 향후 과제로는 시나리오에 따른 와인 에이전트를 구현하여 질의 검색 결과를 네이버, 야후와 같은 타사이트와 비교하여 성과를 측정해 보는 것이 필요하다. 그리고 보다 유연한 웹 서비스가 되도록

하기 위해서 다양한 온톨로지가 정의되어야 하며 온톨로지 정의 구조에 따라 자동으로 입출력 인터페이스가 바뀌게 하는 기술이 연구되어져야 한다.

### 참고 문헌

- [1] 하영국, 손주찬, 조영조. "시맨틱 웹 에이전트를 위한 OWL 온톨로지 추론 엔진의 설계 및 구현", 정보처리학회논문지B, 제11권, 제2호, pp.1499~150, 2004.
- [2] Eric I. Hsu, Deborah L. McGuinness. Wine Agent: Semantic Web Testbed Application. Knowledge Systems Laboratory, Stanford University
- [3] 전홍진. "와인과 음식과의 조화에 관한 연구", 한국관광정보학회논문지, 제14권, pp.97~118, 2003.
- [4] 양하영, 유정목, 이만호. "정보검색서비스에 특화된 시맨틱 웹 서비스 환경을 위한 질의언어 설계" 한국정보처리학회논문지, 제11권, 제2호, pp.1543~1546, 2004.
- [5] Richard Fikes, Patrick Hayes, Ian Horrocks, OWL-QL-A Language for Deductive Query Answering on the Semantic Web. Stanford University, Stanford, CA, 2003.
- [6] 강민숙, 조동섭. "Semantic Web을 이용한 Web Service 기반 전자정부 플랫폼" 한국정보과학회논문지, 제30권 제1호, 2003.
- [7] 하상범, 박영택. "온톨로지 기반 추론을 이용한 시맨틱 검색 시스템" 한국정보과학회논문지B, 제 32권, 제3호, pp.0202~0214, 2005.
- [8] Bijan Parsia, Evren Sirin. Pellet: An OWL DL Reasoner, MINDSWAP Research Group, University of Maryland, College Park, MD,
- [9] Li Dong, Tim Finin, Anupam Joshi. Swoogle: A search and Metadata Engine for the Semantic Web. CIKM: Conference on Information and Knowledge Management, pp. 0652~0659, 2004.
- [10] W3C, "OWL Web Ontology Language Guide," <http://www.w3.org/TR/owl-guide/>.
- [11] W3C, "Web Services Architecture," <http://www.w3c.org/TR/ws-arch/>.
- [12] N. Noy. M. Sintek et al., "Creating Semantic Web Contents with Protege 2000". IEEE Intelligent Systems vol. 16 No.2. 2001.
- [13] 오지훈, 최병석, 정영식. "시맨틱 웹 서비스 기술을 위한 수행 온톨로지" 한국정보처리학회 논문지B, 제11권, 제4호, pp.0457~0464, 2004.