

# 온톨로지를 이용한 인터넷 웹 검색에 관한 실험적 연구\*

An Experimental Study on the Internet Web Retrieval Using Ontologies

김현희(Hyun-hee Kim)\*\*, 안태경(Tae-kyoung Ahn)\*\*\*

## 초 록

온톨로지는 웹자원을 지식화함으로써 정보의 효율적 검색, 통합, 재사용을 도모할 수 있는 새로운 기술인 시맨틱 웹의 구현을 위한 가장 핵심적인 요소 기술로 알려지고 있다. 온톨로지는 사람간에 그리고 서로 다른 응용 시스템간에 지식을 공유하고 재이용하는 방법을 제공하는 기술로서 특정 주제에 관한 지식 용어들의 집합으로서 이들 용어뿐만 아니라 용어간의 의미적 연결 관계와 간단한 추론 규칙을 포함한다. 본 연구에서는 인터넷 웹상에서 국제기구에 관한 정보를 체계적으로 관리하고 검색하기 위해서 국제기구 온톨로지를 설계하고 이 온톨로지에 기반 하여 검색 시스템을 구현해 보고, 이 시스템을 20개의 탐색 질문들을 이용하여 기존의 인터넷 검색엔진과 적합성과 탐색 시간이라는 두 가지 요인을 통해서 비교해 보았다. 실험 결과에 의하면 적합성 측정은 온톨로지 기반 시스템은 평균 4.53, 인터넷 검색엔진은 평균 2.51로 온톨로지 기반 시스템의 적합도가 1.80배 높은 것으로 나타났다. 또한 탐색시간은 온톨로지 기반 시스템은 평균 1.96분, 인터넷 검색엔진은 평균 4.74분으로 인터넷 검색엔진이 온톨로지 기반 시스템 보다 2.42배 정도 더 많은 탐색시간이 필요한 것으로 나타났다.

## ABSTRACT

Ontologies are formal theories that are suitable for implementing the semantic web, which is a new technology that attempts to achieve effective retrieval, integration, and reuse of web resources. Ontologies provide a way of sharing and reusing knowledge among people and heterogeneous applications systems. The role of ontologies is that of making explicit specified conceptualizations. In this context, domain and generic ontologies can be shared, reused, and integrated in the analysis and design stage of information and knowledge systems.

This study aims to design an ontology for international organizations, and build an Internet web retrieval system based on the proposed ontology, and finally conduct an experiment to compare the system performance of the proposed system with that of Internet search engines focusing relevance and searching time. This study found that average relevance of ontology-based searching and Internet search engines are 4.53 and 2.51, and average searching time of ontology-based searching and Internet search engines are 1.96 minutes and 4.74 minutes.

키워드: 온톨로지, 시맨틱 웹, 인터넷 웹 검색, 국제 기구, 지식 관리, DAML+OIL, 클래스, 온톨로지 기반 검색시스템, Ontologies, Semantic Web

\* 이 논문은 2002년도 한국학술진흥재단의 지원에 의하여 연구되었음(KRF-2002-005-B20006)

\*\* 명지대학교, 문헌정보학과(kimhh@mju.ac.kr)

\*\*\* 대외경제정책연구원(tkahn@kiep.go.kr)

■ 논문 접수일 : 2003. 2. 25

■ 게재 확정일 : 2003. 3. 11

## 1 서론

### 1.1 연구 목적

문자의 발명에 따라 사람간의 직접적인 대화를 통해서만 가능했던 지식의 전달이 時空을 초월하여 가능하게 되었다. 이 와 더불어 종이의 발명에 의해 문서를 이용한 지식의 전달 계승은 비약적인 확장을 해 오고 있다. 한편 인터넷웹의 출현과 웹문서 증가는 방대한 지식베이스의 구축과 관리라는 새로운 문제를 발생시켰다. 특히 최근에 와서 인터넷 웹이 기하급수적으로 증가하고 있는 추세이기 때문에 이용자가 필요한 정보를 웹으로 검색하는 것이 더욱 어려워지고 있다. 지금까지의 인터넷 웹은 지식표현 능력의 부재로 정확한 정보의 검색이나 내포된 지식의 자동 추론이 어려웠다. 즉 인터넷 검색엔진은 키워드 검색의 한계, 동의어 및 동음이의어의 구별 불가, 일반적인 용어나 구절에 대한 정확한 이해 부족, 함축적인 정보를 알아내지 못하는 등의 많은 문제점들을 지니고 있다.

웹자원을 지식화함으로써 정보의 효율적 검색, 통합, 재사용을 도모할 수 있는 새로운 기술인 시맨틱웹(semantic web)이 현재 인터넷 검색의 어려움을 많이 감소시켜줄 것으로 기대하고 있다. 시맨틱 웹은 인터넷웹상에 존재하는 데이터에 의미를 부여하며 이러한 의미들을 기계가 이해할 수 있도록 만드는 것이다. 온톨로

지는 이러한 시맨틱 웹의 구현을 위한 가장 핵심적인 요소 기술로 알려지고 있다. 온톨로지는 지식 시스템에 이용되는 암묵적인 정보를 명시하는 역할을 갖고 있으며 사람간에 그리고 서로 다른 응용 시스템간에 지식을 공유하고 재이용하는 방법을 제공하는 기술로서 인공지능이나 웹 관련 연구에서는 기본적으로 개념 규정을 말하며 일반적으로 용어와 용어들간의 관계를 표현하는 컴퓨터 판독이 가능한 공식 규정을 의미한다. 시맨틱웹에서의 온톨로지는 특정 주제에 관한 지식 용어들의 집합으로서 이들 용어뿐만 아니라 용어들간의 의미적 연결 관계와 간단한 추론 규칙을 포함한다.

본 연구의 목적은 인터넷에서 경제 분야 국제기구에 관한 정보를 검색하기 위해서 국제기구 온톨로지를 설계하고 이 온톨로지에 기반 하여 검색 시스템을 구현해 보고 실제 탐색 질문들을 이용하여 구현한 검색 시스템을 기존의 인터넷 전문검색엔진과 적합성과 탐색 시간이라는 두 가지 요인을 통해서 비교해 봄으로써 좀 더 효율적인 인터넷 검색 방안을 모색하는데 있다.

### 1.2 연구 방법과 제한점

인터넷웹 검색에 활용할 온톨로지를 설계하기 전에 어떤 분야의 온톨로지를 만들 것인지를 결정해야 했다. 경제 관련 연구소들에서 IMF, World Bank 등 국제기구의 웹사이트에 있는 출판물, 인물, 프로젝트, 조직 및 국제기구 멤버에 대한

자료를 빈번히 검색하고 있으나 웹사이트의 구조가 매우 복잡하여 쉽게 최종 이용자가 자료를 찾을 수 없다는 점이 제안되어 경제 분야 국제기구에 대한 온톨로지를 구성하기로 결정하였다.<sup>1)</sup>

온톨로지 구성은 다섯 단계 즉 온톨로지 목적과 범위 결정 단계, 지식 수집과 개념화 단계, 온톨로지 통합 단계, 개념 기술과 공식적 설명 단계 및 평가와 문서화 단계를 통해서 출판물(Publication), 인물(Person), 프로젝트(Project), 조직(Organization), 멤버(Member)로 구성된 온톨로지를 설계하였다(Uschold and Gruninger 1996). 특히 기존의 온톨로지를 많이 참조하였고, 분석 대상인 11개 국제기구의 웹사이트들과 이용자들의 질문들을 정밀하게 분석하고 이러한 분석 결과를 바탕으로 경제학자, 경제분야 전문 사서 및 시스템 개발자가 개념들을 수집하여 분류하였다.

온톨로지를 설계한 후 이를 웹 온톨로지 언어인 DAML+OIL로 기술하여 클래스와 클래스, 클래스와 서브클래스 그리고 클래스와 속성간의 관계를 명확히 규명하였다. DAML+OIL로 기술한 온톨로지를 기반으로 검색시스템을 구현하기 위해서 출판물, 인물, 프로젝트, 멤버, 조직의 다섯 개 데이터베이스 파일을 구성하였다. 이 시스템은 전문사서는 물론 최종 이용자를 대상으로 하며 온톨로지 기반

탐색과 전문검색엔진 탐색의 두 개 서브시스템으로 구성된다. 시스템 구현환경은 하드웨어는 Windows NT 서버를 사용하고 온톨로지는 DAML+OIL과 MS ACCESS로 기술하였다. 프로그램은 비주얼 스크립트 언어로 작성한 ASP(Active Server Page)를 이용하였다.

구현한 시스템의 만족도를 조사하기 위해서 이 시스템을 기존 국제기구의 웹사이트인 전문 검색엔진과 비교하기로 결정하였다. 실제 이용자들로부터 의뢰된 질문들을 바탕으로 하여 재구성된 20개의 임의의 질문을 준비하고 10명의 이용자로 하여금 본 연구에서 구축한 온톨로지 기반 시스템과 국제 기구 전문 검색엔진, 일반 검색엔진 등을 각각 검색해 보게 한 후 적합성과 탐색 시간을 비교해 보았다.

## 2 이론적 배경

### 2.1 시맨틱웹과 온톨로지

시맨틱웹은 웹상의 정보에 컴퓨터가 쉽게 해석할 수 있는 잘 정의된 의미(semantics)를 부여하여 사람과 컴퓨터간의 협동 작업을 원활하게 하기 위해서 제안되었다. 현재 웹상의 문서는 사람이 읽고 해석하기에 편리하도록 설계되어 있기 때문에 컴퓨터가 문서의 내용을 읽어 그 내용에 적합하게 처리하기에는 어려움이

1) 11개의 경제 분야 국제 기구 IMF, World Bank, IFC, OECD, EBRD, ADB, WTO, APEC, UNCTAD, UNDP, UNIDO를 대상으로 하였다.

많다. 시맨틱웹에서는 자연어 위주의 기존 웹 문서와 달리 컴퓨터가 해석하기 쉽도록 의미를 부여한 계층을 가지고 있기 때문에 자동화된 프로그램(에이전트)이나 정교한 검색 엔진들이 부여된 의미를 이용하여 높은 수준의 자동화와 지능화를 이룰 수 있게 된다(전자상거래 표준화 통합포럼 2001). 웹에 의미를 부여한다는 것은 이용자 인터페이스를 위한 자연어 처리 기능을 부가하고자 하는 것이 아니라 컴퓨터가 처리하기 용이하게 하고자 선언적인 추가 정보를 부여하는 것을 뜻한다. 이와 같이 시맨틱웹은 웹자원을 지식화함으로써 정보의 효율적 검색, 통합, 재사용을 도모할 수 있는 새로운 기술이다(오삼균 2002).

온톨로지는 지식의 구조를 컴퓨터가 처리할 수 있는 형태로 표현하는 기술로서 시맨틱웹의 구현을 위한 가장 핵심적인 요소기술이다. 온톨로지는 원래 철학에서 존재의 본질과 유형에 관한 이론을 의미하며 인공지능이나 웹 관련 연구에서는 기본적으로 개념 규정을 말하며 일반적으로 용어와 용어들간의 관계를 표현하는 컴퓨터 판독이 가능한 공식 규정을 의미한다. 시맨틱웹에서의 온톨로지는 특정 주제에 관한 지식 용어들의 집합으로서 이들 용어뿐만 아니라 용어들간의 의미적 연결 관계와 간단한 추론 규칙을 포함한다. 시맨틱웹에서는 몇몇의 복잡한 대규모 온톨로지 대신에 다수의 소규모 온톨로지와 이들 간의 상호 협조로 이루어진

웹을 추구한다.

## 2.2 지식관리에서 온톨로지의 필요성

지식 시스템에 이용되는 암묵적인 정보를 명시하는 역할을 갖으며 사람 간에 그리고 서로 다른 응용 시스템간에 지식을 공유하고 재이용하는 방법을 제공하는 온톨로지가 암묵지를 강조하고 있는 지식 관리 분야에 활용되고 있다(古崎晃司 2001). 여기서는 지식 관리 중에서 인터넷 웹 정보 검색에서 온톨로지의 필요성을 살펴보고자 한다. 일반적인 정보검색 기법은 이용 가능한 정보의 엔코딩(예, 고정된 분류 코드)이나 간단한 쉘文 분석에 의존한다. 이와 같은 접근 방식은 몇 가지 문제점이 있다. 첫째는 이 두 방법은 검색 대상인 정보원의 어휘와 완전히 일치하지 않을 수 있는 입력된 이용자의 어휘에 전적으로 의존한다는 것이다. 두 번째는 특정 엔코딩은 질문의 재현율을 급격하게 감소시킬 수 있다는 점이다. 왜냐하면 조금 다른 형식으로 엔코딩된 적합 정보가 매칭되지 않기 때문이다. 또한 쉘文 분석은 단어의 의미가 명확하지 않기 때문에 정확률을 감소시키기도 한다.

더 나아가 기존의 인터넷 웹 검색엔진들은 주로 키워드 기반 탐색 방법을 채택하고 있는데 이 방법은 빠르고 유용하기는 하나 다음과 같은 문제점들을 가지고 있다(Hyvönen 2001). 첫째는 문헌에 나타난 키워드가 그 문헌이 적합하다는 것

을 반드시 의미하지 않는다. 예를 들어서 “This page is not about politics”라는 구절을 갖는 웹페이지는 키워드 “politics”를 이용할 때 이 페이지가 검색 결과에 포함될 수 있다. 두 번째는 검색엔진은 동의어간을 구분할 수 없으며 또한 동음이의어도 구분하지 못한다. 세 번째는 검색엔진은 일반적인 용어나 구절을 이해하지 못한다. 예를 들어서 “AI 프로그래밍 기법”에 대해서 논의하는 웹페이지를 검색할 때 아마도 미리 각 AI 프로그래밍 기법의 명칭을 나열해야 할지 모른다. 네 번째는 적합성의 문제인데 질문에 대한 문헌의 적합성을 평가하는 것이 쉽지 않는 일이다. 일만 건의 검색 결과가 나왔을 때 적합도 순서에 따라 그 결과가 정렬되지 않는 한 자료를 이용하는데 어려움이 있을 것이다. 다섯 번째는 검색엔진들은 함축적인 정보는 찾아내지 못한다. 명확한 키워드를 포함하고 있을 때만 웹페이지들이 검색될 수 있다. 예를 들어서, 아인슈타인의 연구에 대해서 흥미가 있을 때 아인슈타인을 언급하지 않는 한 그의 상대성 이론을 기술하는 웹페이지가 검색되지 않는다. 마지막으로 분포된 관련 정보를 모으기가 어렵다. 예를 들어서 Y 회사와 비즈니스 관련을 맺고 있는 X라는 회사에 대한 모든 정보를 찾을 때, 이러한 정보들은 관련된 회사들의 웹페이지들 상에서 여러 방법으로 분포되어 있을 것

이다.

이러한 환경에서는 컴퓨터에 관하여 충분히 지식이 있는 일부 사람은 인터넷의 정보를 활용할 수 있지만 그 밖의 사람은 인터넷의 정보를 활용할 수 없게 되어 디지털 디바이드<sup>2)</sup>가 확대되는 우려를 갖게 하고 있다. 이러한 문제점을 완화시키려면, 서로 다른 개발자에 의해 만들어진 특정 분야의 인터넷 웹사이트의 공통된 구성 요소들을 정리하여 공유할 수 있는 절차가 필요한데 이를 가능케 하는 것 중 하나가 온톨로지(Ontology)이다. 온톨로지란 특정 분야에서 사용되는 표준 어휘들의 모음이라고 정의될 수 있는데 의미론적 수준에서 웹정보를 검색할 수 있는 기초를 제공한다. 더 나아가 탐색자의 경험적 지식을 분석하여 이를 온톨로지를 활용한 검색 절차에 이용할 수 있다. 실제로 WordNet(Fellbaum 1998)과 같은 약식의 온톨로지 이용은 탐색 과정에서 동의어들을 포함시킴으로써 질문의 재현율을 향상시켰다. 또한 개념 그래피크와 같은 좀더 공식적인 표현의 이용은 재현율과 정확률을 양쪽의 향상을 가져다 주었다(Sowa 1999). 즉 계승 관계에 대한 추론으로 재현율을 향상 시켰고, 구조를 매칭시킴으로써 정확률을 향상시켰다. 결론적으로 온톨로지는 정보원의 기술 어휘와 질문의 어휘와의 차이를 완화시키기 때문에 재현율은 물론 정확률도 향상시키는 것으로 나타났다.

2) 컴퓨터를 통해 정보를 손쉽게 획득할 수 있는 사람과 그렇지 못한 사람간의 격차에 의해 재화가 분배되는 사회의 양분화 현상을 말한다.

### 2.3 웹 온톨로지 언어

DAML은 RDF와 RDF 스키마를 기반으로 하여 RDF 스키마에서 구현되지 않는 서로 동일한 의미의 요소, 역관계, 합집합 등의 메타데이터의 중요한 관계를 지원하는 기능을 첨가한 웹 온톨로지 언어이다. DAML이 나오기 이전에 이미 OIL이라고 하는 온톨로지를 정의하기 위한 언어가 개발되어 있었다. 그러나 OIL만으로 온톨로지를 표현하기에 부족하여 이를 보완할 새로운 언어로 이 두 언어를 결합한 DAML+OIL이라는 언어가 탄생하게 되었다. DAML+OIL의 문법뿐만 아니라 사양 역시 RDF 문서로 작성된다. DAML+OIL은 네임스페이스 선언, 클래스 정의, 속성 정의, 속성 제한 부여 등으로 구성된다. 웹 온톨로지 언어의 가장 최근 동향으로 제시되고 있는 Web Ontology Language(OWL)는 DAML+OIL의 취약점으로 지적되어 온 용어 의미 혼동을 보완한 웹 온톨로지 언어로 DAML+OIL과는 완벽한 수준의 호환성을 유지한다(오삼균 2002). 의미의 일관성을 위해 DAML+OIL의 여러 클래스와 속성명을 재정의 하였고 DAML+OIL이 부족한 대칭적 속성(Symmetric Property)을 첨가하였다.

### 2.4 선행 연구

정보 관리에 온톨로지를 구현한 국내의 연구를 살펴보면 다음과 같다. SHOE

(Simple HTML Ontology Extensions) 프로젝트는 온톨로지 메타데이터에 많은 영향을 끼친 초기 연구 중 하나로 문헌의 내용을 기술하기 위해서 저자에 의해서 사용된 온톨로지들로 간단한 추론 규칙을 정의하여 이용할 수 있다(Luke et al. 1996; Heflin et al. 1998). SHOE 프로젝트의 데이터베이스 구성을 살펴보면 온톨로지 메타데이터가 그것과 관련된 문헌의 내용에 첨가된다(Kampa 2002). SHOE 수집기(crawler)는 지식을 추출하기 위해서 문헌을 분석하고 탐색 엔진을 이용하여 이용자는 수집된 지식을 의미적으로 탐색할 수 있다. 이 프로젝트는 Semantic Web Agents 프로젝트에 의해 계속되고 있다(<http://www.mindswap.org/>). Ontobroker(Fensel et al. 1998)는 SHOE와 같이 웹 문서를 분석하기 위해서 온톨로지를 이용하고 온톨로지 기반의 질문 처리 절차를 갖고 있다. 확장된 HTML 문법이 이용자들이 문서를 온톨로지 구조로 마크업하기 위하여 제안되었다. 이와 같이 마크업된 문서는 Ontobroker 수집기(crawler)에 의해서 처리된다. 이 시스템의 인터페이스는 Ontobroker에 의해서 수집된 지식을 이용자들이 탐색할 수 있도록 하며 문헌의 내용을 이해하기 위해서 그 지식을 이용하게 한다. 이 프로젝트는 KAON(Karlsruhe Ontology) 프로젝트에 의해 계속되고 있다(Handschuh et al. 2001).

SHOE 프로젝트의 핵심 연구원인 Heflin

(2001)은 SHOE 연구를 발전시켜 메릴랜드 대학의 컴퓨터 공학과에 대한 온톨로지를 설계한 다음 이 온톨로지를 기반으로 하여 메릴랜드 대학의 컴퓨터 공학과의 웹페이지를 분석하여 데이터베이스로 구축하여 검색 시스템을 구현하였다. 이 연구의 목적은 인터넷 웹사이트에 있는 다양한 사이트들을 의미지도(semantic map)를 통해서 효율적으로 검색하는 데 있다. 이 연구는 현재 일반 인터넷 검색엔진을 활용할 경우, 예를 들어서 “메릴랜드주에 있는 공립 대학교의 교수들이 쓴 온톨로지에 관한 문헌들을 찾고 싶다”라는 질문이 있는 경우 대학 홈페이지, 교수 리스트 홈페이지 및 각 교수의 출판 정보 페이지들이 필요한데 이와 같은 질문은 현재의 검색엔진으로는 어려움이 있다는 가정 하에 시작되었다.

古崎晃司(2001)는 석유정제 플랜트 온톨로지를 개발하였다. 이 연구에서 강조된 것은 다른 기존의 온톨로지에서 이용되지 않은 역할 개념을 온톨로지 구성에 활용한 점이다. 역할 개념에 따라 어떤 고등학교에 근무하고 있는 남성은 고등학교에서는 “교사”, 가정에서는 “부친”, 부부사이에서는 “남편”이라는 콘텐츠의 변화에 따라서 다른 개념으로 인식되어 진다. 이 역할 개념을 활용하여 기존의 온톨로지 기술 환경에서 다루지 못하였던 개념적인 성질을 역할 개념을 중심으로 하여 통일적인 구조에서 명시적으로 다룰 수 있게 되었다.

Vasconcelos(2001)는 지식 집약적 기업체의 그룹 메모리 시스템(group memory system)과 같은 기업 정보 시스템에 온톨로지를 어떻게 활용할 수 있는지를 기술하고 있다. 이 연구에서는 product, information 그리고 competence 온톨로지를 갖고 있는 그룹 메모리 시스템(GMS)을 설계, 구현 및 평가한다. 세 가지 온톨로지는 각 그 기능을 갖고 있는데 먼저 product 온톨로지는 조직에 의해서 이용되는 생산품, 정보 기술에 대한 특정한 지식을 기술하는 온톨로지이며, information 온톨로지는 그룹 멤버간의 정보원 관리에 초점을 맞추었다. 마지막으로 competence 온톨로지는 기업체의 특정 비즈니스 기능을 수행하기 위해서 작업장에서 필요한 지식과 기술을 기술한다.

국내의 경우는 尹洪珪(2001)는 온톨로지 개념을 활용하여 지리 정보시스템을 구현하였다. 특히 지리 정보시스템에서 사용하는 어휘는 일반인들이 이해하기에 어려운 부분이 많기 때문에 온톨로지를 활용한 지리 정보시스템은 최종 이용자들이 효율적으로 이용할 수 있다고 기술하고 있다. 崔武羅(2001)는 자동차 산업의 제품 데이터 중 부품 데이터의 효율적인 검색을 위해서 각 부품간의 관계나 용어에 대해 온톨로지를 적용하고 웹 기반의 부품 정보검색시스템을 구현하였다.

### 3. 온톨로지 구성

#### 3.1 설계 과정

온톨로지 개발을 위한 여러 연구들이 제안되었으나 가장 널리 이용되고 있는 방법은 온톨로지 목적과 범위, 지식 수집과 개념화, 통합, 이행, 평가의 다섯 단계 과정이다(Uschold and Gruninger 1996). 다음은 이 다섯 단계에 의해서 국제기구 온톨로지를 설계하는 과정을 설명한다.

##### 3.1.1 온톨로지 목적과 범위 결정 단계

이 단계는 온톨로지의 목적과 범위를 정의하는 단계이다. 이는 이용, 이용자, 요구된 형식의 정도, 온톨로지의 범위에 대한 비공식적인 기술을 포함한다. 엔티티, 관련 용어, 속성 및 관계에 대한 비공식 기술을 포함하여 도메인 용어를 정의하기 위한 최초의 접근이 이 단계에서 기술된다. 이 단계의 결과물은 자연어로 기술한 온톨로지 설명문서가 될 것이다. 본 온톨로지는 경제 분야 국제기구를 대상으로 하였으며 이용자는 경제 관련 학자, 연구원, 사서 등을 대상으로 하였다. 이 온톨로지의 목적은 이용자들이 좀 더 쉽게 원하는 웹자료를 검색할 수 있도록 하는데 있다.

##### 3.1.2 지식 수집과 개념화 단계

이 단계는 주어진 도메인으로부터 지식

을 수집하는 단계이다. 이는 도메인 텍스트 분석, 전문가 인터뷰, 설문지 응용과 같은 전통적인 방법을 적용한다. 지식 수집 과정은 도메인 개념화 즉 도메인 엔티티 또는 개념, 인스턴스, 관계, 속성을 확인하여 표현하고 이들을 도메인 용어와 연결시키는 과정을 이끌어 낸다. 각 용어와 용어간의 관계는 비공식 언어로 표현된다. 이 단계는 도메인 전문가에 의해서 사용된 용어를 정의하려는 의도이다. 이는 도메인 관계자에 의해서 사용된 모든 다른 용어들의 포괄적인 정의, 그것의 의미 및 관계를 이끌어내기 위해서 도메인 공학자와 도메인 전문가간의 반복되는 작업이 된다.

본 연구에서는 도메인으로부터 지식을 수집하기 위해서 IMF, World Bank 등 11개의 국제기구 웹사이트와 이용자들이 자주 탐색하는 질문들을 분석하였고 이 분석 결과를 바탕으로 경제 분야의 연구자와 專門 사서 그리고 시스템 개발자와의 지속적인 작업을 통해서 도메인 용어 리스트를 작성하고 그것들의 연관된 의미를 규정한 다음 도메인 개념들을 카테고리화하여 분류하였다.

##### 3.1.3 온톨로지 통합 단계

기존의 온톨로지를 통합하거나 재이용하는 단계이다. 이 단계는 다른 온톨로지를 재사용함으로써 새로 개발할 온톨로지의 구성 과정을 빠르게 한다. 온톨로지간의 일률성을 유지시키기 위해서 다른 온



〈표 1〉 참조한 온톨로지

|   |
|---|
| Organisation ontology [DAML Crawler( <a href="http://www.daml.org/crawler">http://www.daml.org/crawler</a> ), <a href="http://orlando.drc.com/daml/Ontology/Organization/3.1.1/Organization-ont.daml">http://orlando.drc.com/daml/Ontology/Organization/3.1.1/Organization-ont.daml</a> ] |
| Publications ontology (DAML Ontology Library, <a href="http://www.daml.org/ontologies/4">http://www.daml.org/ontologies/4</a> )   |
| Academic department ontology (DAML Ontology Library, <a href="http://www.daml.org/ontologies/64">http://www.daml.org/ontologies/64</a> )  |
| National Security Orgnization ontology (DAML Ontology Library, <a href="http://www.daml.org/ontologies/293">http://www.daml.org/ontologies/293</a> )  |
| Organisation ontology (DAML Ontology Library, <a href="http://www.daml.org/ontologies/184">http://www.daml.org/ontologies/184</a> )   |
| People ontology (DAML Ontology Library, <a href="http://www.daml.org/ontologies/61">http://www.daml.org/ontologies/61</a> )   |

톨로지의 정의들도 재사용될 수 있다. 기존의 온톨로지들을 분석하여 이용할 수 있는 클래스와 속성을 추출하여 앞 단계에서 구성한 분류된 도메인 개념과 통합하였다. 엄밀히 말하자면 이 단계는 앞의 지식 수집과 개념화 단계와 동시에 진행되었다. 다음 〈표 1〉은 참조한 온톨로지의 출처를 기술하고 있다.

### 3.1.4 이행: 개념 기술과 공식적 설명 단계

이 단계는 온톨로지를 Ontologua와 같은 공식적인 표현 언어를 이용하여 표현할 수 있음을 보여준다. 여기서 각 용어의 공식화와 온톨로지에 의해서 이용된 제한점들이 기술된다. 이와 같은 용어들은 클래스, 관계, 기능, 인스턴스 등을 통해서 표현된다. 본 연구에서는 온톨로지를 DAML+OIL 언어를 사용하여 표현하였다.

### 3.1.5 평가와 문서화 단계

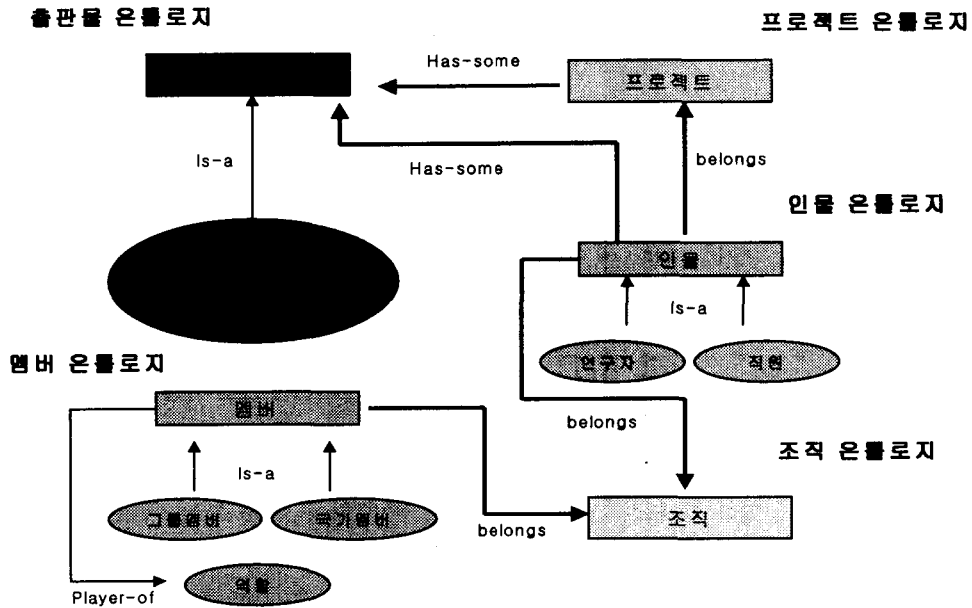
온톨로지를 개발할 때 사용되는 중요한 기준은 명확함, 일관성, 완전성, 확장성

등이다. 온톨로지 확인과 검증은 불완전성, 불일치성, 중복성을 체크하는 일련의 지침을 통해서 이루어진다. 온톨로지에 있는 용어들은 분명하게 정의되어야 하고 전문 용어들은 일관성이 있어야 하고 용어간의 관계는 논리적으로 일관성이 있어야 한다. 본 연구에서는 온톨로지의 평가를 위해서 온톨로지 기반 정보 시스템을 구성하고 실제 질문을 갖고 테스트하여 사용자들의 만족도를 조사해 보았다.

## 3.2 온톨로지 표현

### 3.2.1 개요

다음은 위의 다섯 단계에 의해서 구성된 온톨로지를 좀 더 자세히 설명한다. 구성된 온톨로지는 출판물 온톨로지(Publication Ontology), 인물 온톨로지(Person Ontology), 프로젝트 온톨로지(Project Ontology), 멤버 온톨로지(Member Ontology), 조직 온톨로지(Organization Ontology)의 다섯 개로 구성된다. 〈그림 1〉에서 굵은 선은



〈그림 1〉 온톨로지 구성

온톨로지간의 관계를 기술하고 있다. 프로젝트 온톨로지와 인물 온톨로지는 프로젝트 출판물과 연구자(또는 직원)의 출판물로 출판물 온톨로지와 연결되고 있다. 조직 온톨로지는 조직의 구성 요소인 인적 자원으로 인물 온톨로지와 연결되며 멤버 온톨로지는 조직 온톨로지와 멤버구성원으로 연결된다.

출판물, 인물 및 멤버 온톨로지에 사용된 "is-a" 관계는 개념의 일반화-상세화의 관계를 기술할 때 이용된다. 예를 들어서 "인물"과 "연구자" 사이의 "연구자 is-a 인물"이라는 관계가 성립될 때 "연구자"는 "인물"을 상세화한 개념이고, 반대로 "인물"은 "연구자"를 일반화한 개념이다. 이때 일반화된 개념을 상위 개념,

상세화된 개념을 하위개념으로 부른다. 온톨로지에서는 "is-a" 관계를 이용하여 개념 즉 클래스를 계층적으로 체계화한다. 또한 프로젝트가 출판물을 가지고 있을 수 있다는 것을 "has-some"으로 표현했고 "belongs"를 이용하여 인물과 멤버가 조직에 속해 있다는 것을 기술하였다.

1) 출판물 온톨로지

출판물 온톨로지에는 루트에 Publication 클래스가 있고 이 클래스는 Serial, Article, UnofficialPublication 등 11개의 서브 클래스로 구성되어 있다(〈표 2〉 참조). Serial 서브 클래스는 다시 JournalArticle, Outlook 등의 다섯 개의 서브클래스를 갖고 있다. Publication 클래스에는 pubCode

〈표 2〉 출판물 온톨로지

|  |
|--|
| Publication (pubCode, pubAuthor, pubTitle, pubLan, pubYear, pubResearch_topic,<br>pubResearch_country __region)<br>Serial (pubFrequency)<br>JournalArticle (journalTitle)<br>Statistics (code)<br>WorkingPaper (seriesStat)<br>DiscussionPaper (seriesStat)<br>Outlook ()<br>Article ()<br>Conference (conferenceTitle)<br>Program<br>Paper<br>TechnicalReport ()<br>Book ()<br>Map ()<br>Manual ()<br>News (speaker, pubMonth, pubDay)<br>Video (speaker)<br>Specification ()<br>UnofficialPublication () |
|--|

를 포함하여 일곱 개의 속성을 갖고 있고 이 속성은 Publication 밑에 있는 모든 하위 클래스에 그대로 계승된다. 반면에 Serial 서브 클래스의 속성인 pubFrequency는 Serial의 서브 클래스들에만 적용되며 journalTitle과 seriesStat 속성들은 JournalArticle과 WorkingPaper/DiscussionPaper 서브클래스에만 각각 적용된다. 출판물 온톨로지를 이용하여 각 국제기구

에서 생산한 출판물을 유형별로 탐색해 볼 수 있다.

2) 인물 온톨로지

인물 온톨로지에는 루트에 Person 클래스가 있고 이 클래스는 Researcher와 Staff의 2개의 서브 클래스로 구성되어 있다(〈표 3〉 참조). 인물 온톨로지는 출판물 온톨로지과 같이 상위 클래스의 속성들이

〈표 3〉 인물 온톨로지

|   |
|---|
| Person (name, email, homepage, officephone, sex, address, department, position,<br>associated_with_project, nationality)<br>Researcher (expertise_topic, expertise_country_or_area)<br>Staff () |
|---|

〈표 4〉 프로젝트 온톨로지

|  |
|--|
| Project (title, starting_year, ending_year, subject_topic,<br>subject_country_or_area, funding_sponsors, budget,<br>proPublications, has_members, team_title,<br>is_conducted_by, contact_point) |
|--|

〈표 5〉 조직 온톨로지

|  |
|--|
| Organization (mission, name, org_type, has_groupmember, has_countrymember,<br>affiliatedOrganization, community, address, telephoneNumber, email,<br>subOrganization)<br>OrganizationType () |
|--|

〈표 6〉 멤버 온톨로지

|  |
|--|
| Member (org_name, member_code, member_name, has_role, has_URL, joining_year)<br>GroupMember ()<br>CountryMember () |
|--|

하위 클래스에 그대로 계승된다. 인물 온톨로지를 이용하여 관심 분야(주제 또는 나라/지역)별, 국적별, 부서별 등으로 연구자들을 탐색해 볼 수 있고 직원들도 찾아 볼 수 있다.

### 3) 프로젝트 온톨로지

프로젝트 온톨로지에는 루트에 Project 클래스가 있고 이 클래스는 title, contact\_point 등 12개의 속성을 갖고 있다. 프로젝트 온톨로지를 이용하여 각 국제기구의 프로젝트를 분야(주제 또는 나라/지역)별, 수행시기별, 책임자별 등으로 탐색해 볼 수 있고 프로젝트의 출판물도 탐색해 볼 수 있다.

### 4) 조직 온톨로지

조직 온톨로지에는 Organization과 OrganizationType 클래스가 있고 Organization 클래스에는 mission, name 등 11개의 속성을 갖고 있다. 조직 온톨로지를 이용하여 국제 기구들을 유형별로 탐색해 볼 수 있으며 그 외에 각 국제기구의 서브 조직, 커뮤니티, 협력 기구에 대한 정보를 탐색해 볼 수 있다.

### 5) 멤버 온톨로지

멤버 온톨로지는 루트에 Member 클래스가 있고 이 클래스는 GroupMember와 CountryMember라는 두 개의 서브 클래스로 구성되어 있다. Member 클래스는 6개의 속성을 갖고 있다. 멤버 온톨로지를

〈표 7〉 네임스페이스 선언 부분

```

<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
xmlns:xsd="http://www.w3.org/2000/10/XMLSchema#"
xmlns:org="http://orlando.drc.com/daml/ontology/Organization/3.1.1/Organization-ont#">
    
```

이용하여 각 국제기구 멤버의 역할, 가입  
년도 등을 탐색해 볼 수 있다.

### 3.2.2 DAML+OIL로 표현한 온톨로지

위의 각 온톨로지를 클래스로 표현하여  
하나의 온톨로지 구성 후 DAML+  
OIL 언어로 표현하였다. 가장 최신 웹 온  
톨로지 언어인 OWL을 사용하지 않는 이  
유는 개발 당시 OWL은 작업 초안  
(working draft) 상태였기 때문에 DAML+  
OIL 언어로 표현한 후 두 언어간에 호환  
이 가능하기 때문에 나중에 OWL 언어로  
바꿀 계획을 세웠다. 전체 구성은 맨 앞  
부분에 네임스페이스를 설정한 다음 클래스  
와 속성 등을 정의하여 기술한다. 다음  
은 네임스페이스를 설정하는 부분을 포함  
하여 Publication 클래스와 Organization  
클래스의 일부를 소개한다.

#### 1) 네임스페이스 선언

DAML+OIL은 RDF로 작성되어 있다.  
즉, DAML+OIL 마크업은 RDF 마크업  
의 특별한 종류라고 생각하면 된다. RDF  
는 XML로 작성되어 있으며, XML은 역

시 XML 네임스페이스와 URI로 작성된  
다. 다음에서 “rdf:” 라는 접두어가 붙은  
것은 앞으로 이를 “http://www.w3.org/  
1999/02/22-rdf-syntax-ns#” 네임스페이  
스를 참조하겠다는 것을 의미한다.

#### 2) 클래스와 속성의 정의

Publication 클래스는 매릴랜드 대학  
(http://www.cs.umd.edu/projects/plus/S  
HOE/)에서 작성한 온톨로지인 “base1.0.  
daml”에서 사용한 “SHOEEntity” 클래스  
의 서브 클래스로 정의하였다. 속성은 이  
진 관계로서 두 개의 아이টে를 관계지어  
연결하는 역할을 한다. 속성은 객체를 데  
이터유형 값에 연결시키는 “데이터형 속  
성”과 객체간의 관계를 기술하는 “객체형  
속성”으로 나뉜다(오삼균 2002). 전자  
는 “daml:DataTypeProperty”에 해당하며  
후자는 “daml:ObjectProperty”에 해당한  
다. Publication 클래스의 속성인 Publication-  
Author는 “데이터형 속성”으로 정의되고  
있는데 PublicationAuthor 값은 XML 스  
키마의 표준 위치를 참조함으로써 XML  
스키마에서 정의한 스트링을 이용할 수  
있도록 하였다. Organization 클래스의 속

〈표 8〉 클래스와 속성 정의

```

<!-- ***** PUBLICATION CLASS*****-->
<daml:Class rdf:ID="Publication">
  <daml:label>publication</daml:label>
  <daml:subClassOf rdf:resource="http://
    www.cs.umd.edu/projects/plus/DAML/onts /base1.0.daml #SHOEEntity"/>
  <daml:comment>information about publication </daml:comment>
</daml:Class>
<!-- ***** PUBLICATION DATATYPE PROPERTY*****-->
<daml:DatatypeProperty rdf:ID="PublicationAuthor">
  <daml:domain rdf:resource="#publication"/>
  <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema #string"/>
  <daml:comment>publication's author </daml:comment>
</daml:DatatypeProperty>

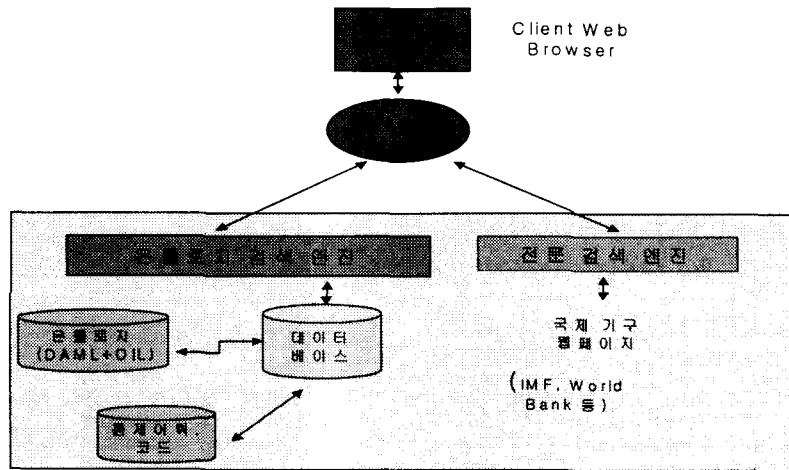
<!--*****ORGANIZATION OBJECT PROPERTY*****-->
<daml:ObjectProperty rdf:ID="Org__type">
  <daml:domain rdf:resource="#Organization"/>
  <daml:range rdf:resource="#OrganizationType"/>
</daml:ObjectProperty>

<!--*****ORGANIZATIONTYPE CLASS(ENUMERATION)*****-->
<daml:Class rdf:ID="OrganizationType">
  <daml:oneOf rdf:parseType="daml:collection">
    <daml:Thing rdf:ID="Finance__org"/>
    <daml:Thing rdf:ID="Economic__org"/>
    <daml:Thing rdf:ID="Trade__org"/>
    <daml:Thing rdf:ID="Industrial__development__org"/>
  </daml:oneOf>
</daml:Cass>

```

성인 “Org\_\_type”은 “객체형 속성”으로 정의되고 있다. Org\_\_type 속성의 도메인 (domain)과 범위(range)가 각각 Organization, OrganizationType이라고 선언하여 두 클래스(객체)간의 관계를 기술하고 있으며 이 두 클래스는 Org\_\_type이라는 속성에 의해서 연결된다. 끝으로 Organi-

zationType 클래스를 살펴 보면 oneOf 엘리먼트를 이용하여 OrganizationType 클래스의 값을 열거된 리스트에서 선택할 수 있도록 하였다.



<그림 2> 시스템 구성도

## 4. 온톨로지 기반 검색 시스템

### 4.1 시스템 구성

#### 4.1.1 개요

시스템은 웹기반으로 구축하였고 전체 구성은 온톨로지 검색 서브시스템과 인터넷 전문엔진 서브시스템으로 나뉜다(〈그림 2〉 참조). 온톨로지 검색 서브시스템에는 세 가지 종류의 데이터베이스 즉 DAML+OIL로 기술된 온톨로지, 이 온톨로지를 기반으로 구축한 출판물, 인물, 프로젝트, 멤버, 조직의 다섯 개의 액세스 데이터베이스 파일 그리고 액세스 데이터베이스 파일에서 사용된 통제어와 코드 데이터가 있고 이러한 데이터베이스를 저장하고 검색하는 프로그램인 온톨로지 검색

엔진이 있다. 위의 각 온톨로지를 클래스로 표현하여 하나의 온톨로지를 구성한 후 DAML+OIL 언어로 기술하였다. 전문 검색엔진 서브시스템은 온톨로지 검색 서브시스템과 검색 효율성을 비교하기 위해서 포함시켰는데 이 서브시스템을 통해서 표본으로 선정한 11개의 국제 기구의 웹사이트로 연결된다. 시스템 구현환경은 하드웨어는 Windows NT 서버를 사용하고 온톨로지는 DAML+OIL과 MS ACCESS로 기술하였다. 프로그램은 비주얼 스크립트 언어로 작성한 ASP(Active Server Page)를 이용하였다.

#### 4.1.2 데이터베이스 구조

##### 1) DBMS 데이터

액세스 DBMS에 저장된 데이터베이스는 총 다섯 개로 1,384개의 레코드로 구

성된다. 다음은 각 데이터베이스에 대해서 자세히 설명한다.

(1) 출판물 데이터베이스

11개의 국제 기구들의 웹페이지에 올려진 출판물을 분석하여 595개의 레코드로 구성된 출판물 데이터베이스를 구축하였다. 출판물 데이터베이스는 <표 9>와 같이 21개의 필드로 구성되어 있다. 이러한 작업은 모두 수작업으로 진행되었는데 이 과정을 자동화 또는 반자동화하는 절차가 필요한 것 같다. 필드값 아래에는 IMF 웹페이지에 올려진 하나의 워킹페이퍼를

코당한 값을 보여 주고 있다. 21개의 필드는 출판물의 모든 유형의 문헌에 해당되는 속성값으로 워킹페이퍼의 경우는 학술지명, 컨퍼런스명, 발표자, 발간월, 발간일, 통계코드는 해당 사항이 없으므로 공란으로 비워 두었다.

(2) 인물 데이터베이스

11개의 국제 기구들의 웹페이지에 올려진 연구자와 직원에 관한 정보를 분석하여 120개의 레코드로 구성된 인물 데이터베이스를 구축하였다. 인물 데이터베이스

<표 9> 출판물 레코드 구조와 실제 입력된 값

| 필드명           | 필드값                                  | 필드명              | 필드값                               | 필드명          | 필드값  |
|---------------|--------------------------------------|------------------|-----------------------------------|--------------|--|
| ID            | 레코드번호                                | conf_title       | 컨퍼런스명                             | speaker      | 발표자  |
|               | 18                                   |                  |                                   |              |  |
| org_type      | 기구 유형                                | date_year        | 발간년                               | n_month      | 발간월  |
|               | monetaryfund                         |                  |                                   |              |  |
| org_name      | 기구 명칭                                | lang             | 언어                                | n_day        | 발간일  |
|               | IMF                                  |                  |                                   |              |  |
| docu_form     | 문헌 유형                                | research_topic1  | 주제-토픽(대)                          | code         | 통계 코드  |
|               | serial/worp                          |                  | Economics                         |              |  |
| author        | 저자                                   | research_topic2  | 주제-토픽(중)                          | frequency    | 출판 빈도  |
|               | Leigh, Danie                         |                  | Monetary and Financial Issues     |              | 부정기적   |
| title         | 표제                                   | research_topic3  | 주제-토픽(소)                          | series_title | 총서명  |
|               | Exchange Rate Pass-Through in Turkey |                  | Monetary and Exchange Rate Policy |              |  |
| journal_title | 학술지명                                 | research_country | 주제-나라/지역                          | URL          | 원문 및 자료 연결 주소                                    |
|               |                                      |                  | Turkey                            |              | www.imf.org/external/pubs/ft/wp/2002/wp02204.pdf |



〈표 10〉 인물 레코드 구조와 실제 입력된 값

| 필드명         | 필드값             | 필드명         | 필드값   | 필드명                  | 필드값   |
|-------------|-----------------|-------------|---|----------------------|---|
| ID          | 레코드번호           | email       | 이메일   | expertise<br>topic   | 전문 분야-토픽  |
|             | 23              |             |   |                      | 개발  |
| org_name    | 기구 명칭           | homepage    | 홈페이지  | expertise<br>country | 전문 분야-나라  |
|             | World Bank      |             |   |                      | 세계 일반   |
| status      | 신분 구분           | officephone | 사무실 전화  | department           | 부서  |
|             | researcher      |             | (202) 473-1000                                  |                      | Senior Management   |
| name        | 이름              | sex         | 남녀 구분   | position             | 위치  |
|             | Nankani, Gobind |             | 남자  |                      | Vice President and Network Head, Poverty Reduction and Economic Management                                  |
| nationality | 국적              | address     | 주소  | URL                  | 상세 정보 링크  |
|             | 가나, 아프리카        |             | 1818 H Street, N.W. Washington, DC 20433 U.S.A. |                      | <a href="http://web.worldbank.org/WBSITE/EXTERNAL/.....">http://web.worldbank.org/WBSITE/EXTERNAL/.....</a> |

는 〈표 10〉과 같이 15개의 필드로 구성되어 있다. 필드값 아래에는 World Bank 웹페이지에 올려진 한 연구자에 대한 정보를 보여 주고 있다. 해당 정보가 없는 것은 공란으로 비워두었다.

(3) 프로젝트 데이터베이스

11개의 국제 기구들의 웹페이지에 올려진 프로젝트에 관한 정보를 분석하여 272개의 레코드로 구성된 프로젝트 데이터베이스를 구축하였다. 프로젝트 데이터베이스는 〈표 11〉과 같이 17개의 필드로 구성되어 있다. 필드값 아래에는 World Bank 웹페이지에 올려진 한 프로젝트에 대한 정보를 보여 주고 있다. 해당 정보가 없

는 것은 공란으로 비워두었다.

(4) 멤버 데이터베이스

11개의 국제 기구들의 웹페이지에 올려진 멤버에 관한 정보를 분석하여 376개의 레코드로 구성된 멤버 데이터베이스를 구축하였다. 멤버 데이터베이스는 〈표 12〉와 같이 7개의 필드로 구성되어 있다. 필드값 아래에는 OECD 기구의 한 멤버인 MEXICO에 대한 정보를 보여 주고 있다.

(5) 조직 데이터베이스

11개의 국제 기구들의 웹페이지에 올려진 각 기구의 조직에 관한 정보를 분석하여 11개의 레코드로 구성된 조직 데이터

〈표 11〉 프로젝트 레코드 구조와 실제 입력된 값

| 필드명           | 필드값                             | 필드명             | 필드값   | 필드명           | 필드값                             |
|---------------|---------------------------------|-----------------|---|---------------|---------------------------------|
| ID            | 레코드번호                           | budget          | 예산  | Pub_author    | 프로젝트 폴관물 (저자)                   |
|               | 79                              |                 | US \$ million<br>15   |               |                                 |
| org_name      | 기구 명칭                           | members         | 구성원   | Pub_title     | 프로젝트 출판물 (표제)                   |
|               | World Bank                      |                 |   |               | Environmental and social<br>... |
| starting_year | 시작 연도                           | team_title      | 팀 타이틀   | Pub_URL       | 프로젝트 출판물 (URL)                  |
| 2002          | 2002                            |                 |   |               | www-wds.worldbank.org/..        |
| ending_year   | 종료 연도                           | subject_topic   | 주제-토픽<br>개발   | contact_point | 연락 정보                           |
| 2004          | 2004                            |                 |   |               |                                 |
| sponsor       | 지원 기관                           | subject_country | 주제-나라   | conductor     | 총 책임자                           |
|               |                                 |                 | 아프카니스탄  |               |                                 |
| title         | 표제                              | URL             | 상세 정보 링크  |               |                                 |
|               | Afghanistan -<br>Emergency .... |                 | http://www4.worldbank.org/sprojects/Project.asp?pid<br>=P077896 |               |                                 |

〈표 12〉 멤버 레코드 구조와 실제 입력된 값

| 필드명      | 필드값   | 필드명         | 필드값    | 필드명         | 필드값   |
|----------|---|-------------|--------|-------------|-------|
| ID       | 레코드번호   | member_name | 멤버명    | member_code | 멤버 코드 |
|          | 1   |             | MEXICO |             | bcl   |
| org_name | 기구 명칭   | r_year      | 가입 연도  | role        | 멤버 역할 |
|          | OECD  |             | 1994   |             | 일반 멤버 |
| URL      | 상세 정보 링크  |             |        |             |       |
|          | http://www.oecd.org/oecd/pages/document/displaywithoutnav/0,3376,EN-do<br>cument-notheme-1-no-no-9464-0,00.html |             |        |             |       |

베이스를 구축하였다. 조직 데이터베이스는 〈표 13〉과 같이 12개의 필드로 구성되어 있다. 필드값 아래에는 World Bank의 조직 전반에 대한 정보를 보여 주고 있다. 해당 정보가 없는 것은 공란으로 비워두었다.

2) DAML+OIL로 기술된 온톨로지

출판물, 인물, 프로젝트, 멤버, 조직의 다섯개의 온톨로지를 클래스로 표현하여 하나의 온톨로지 구성 후 DAML+OIL 언어로 표현하였다. 전체 구성은 맨 앞부분에 네임스페이스를 설정한 다음 클래스와 속성 등을 정의하여 기술한다<sup>3)</sup>.

3) 좀 더 자세한 내용은 3.2 온톨로지 표현에서 참조하고 전체 DAML+OIL 파일은 부록 1을 참조

<표 13> 조직 레코드 구조와 실제 입력된 값

| 필드명      | 필드값                        | 필드명     | 필드값  | 필드명      | 필드값  |
|----------|----------------------------|---------|------|----------|--|
| ID       | 레코드번호                      | com_org | 위원회  | t_number | 전화 번호  |
|          | 1                          |         |      |          | (202) 473-1000   |
| org_name | 기구 명칭                      | aff_org | 협력기구 | email    | 이메일  |
|          | World Bank                 |         |      |          | msilvero@worldbank.org.  |
| org_type | 기구 유형                      | mission | 사명   | members  | 회원국  |
|          | monetaryfund               |         |      |          | Our dream is a world free of poverty To fight overtly .....        |
| sub_org  | 하위 국제기구                    | address | 주소   | URL      | 홈페이지   |
|          | IBRD,IFC, MIGA, ICSID, IDA |         |      |          | Headquarters -General Inquiries The World Bank 1818 H Street, N.W. |

3) 코드와 통제어휘 데이터

출판물, 프로젝트 및 인물 데이터베이스의 주제-토픽, 주제-나라/지역은 통제어휘를 사용하였다. 주제-토픽은 대,중,소의 세 단계로 구분되며 총 252개의 주제어로 구성되며 주제-나라/지역은 132개의 단어로 구성된다. 출판물 데이터베이스에 사용된 언어 코드는 9개의 단어로 구성되며 끝으로 인물 데이터베이스에 사용된 국적 코드는 14개의 단어를 사용한다.

4.2 시스템 구현

4.2.1 개요

시스템의 기능을 소개하기 위해서 두 가지 검색 사례를 온톨로지 선택, 속성값의 입력 및 검색 결과를 중심으로 설명한다.

4.2.2 정보 검색

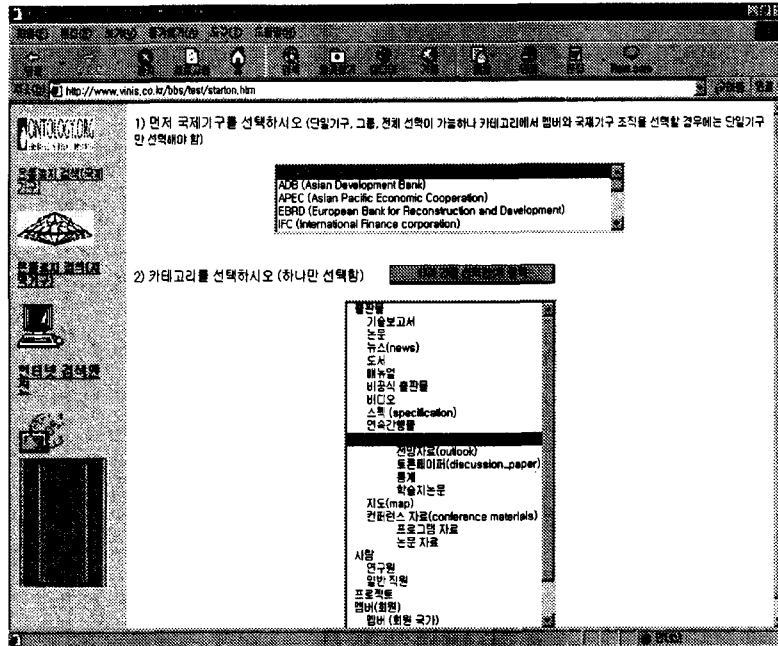
1) 출판물 검색

(1) 온톨로지 선택

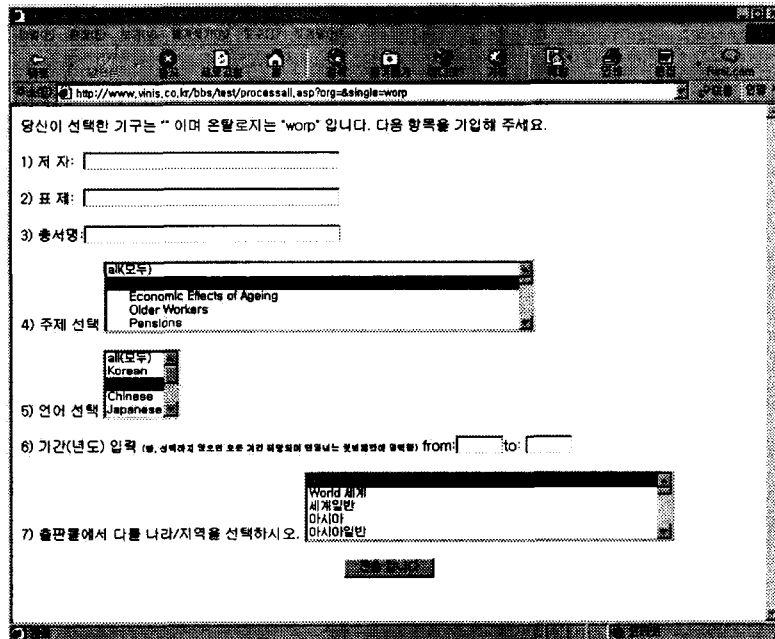
시스템의 초기 화면에서 검색할 국제기구와 온톨로지(카테고리)를 선택하는데 "국제기구 워킹페이퍼 중 고령화사회"에 관련 자료를 찾기 위해서 국제기구 전체와 온톨로지는 연속간행물의 워킹페이퍼를 선정하였다(<그림 3> 참조).

(2) 속성값 입력

시스템의 초기화면에서 선택한 온톨로지에 따라 속성값 입력 화면이 디스플레이 되는데 워킹페이퍼를 선택한 경우 <그림 4>에서 처럼 저자, 표제, 총서명 등의



<그림 3> 은톨로지 선택



<그림 4> 은톨로지의 속성값 입력

7개의 속성이 출력된다. 저자, 표제, 총서명, 기간은 입력하지 않고 주제 선택에서는 “고령화 사회”, 언어 선택에서는 “english”, 출판물에서 다룬 나라/지역 선택에서는 “all”을 선택하였다. 이는 연도, 특정 저자나 표제와 관련없이 “고령화 사회”를 다루는 모든 문헌들을 검색하게 된다.

(3) 검색 결과

속성값을 입력하면 해당 자료의 간략 화면이 출력된다. <그림 5>는 간략 화면에서 문헌번호 “469”를 선택한 후 출력된 상세 화면이다. 이 결과에서 원문 링크를

누르면 해당 문헌의 원문인 PDF 파일에 접근한다.

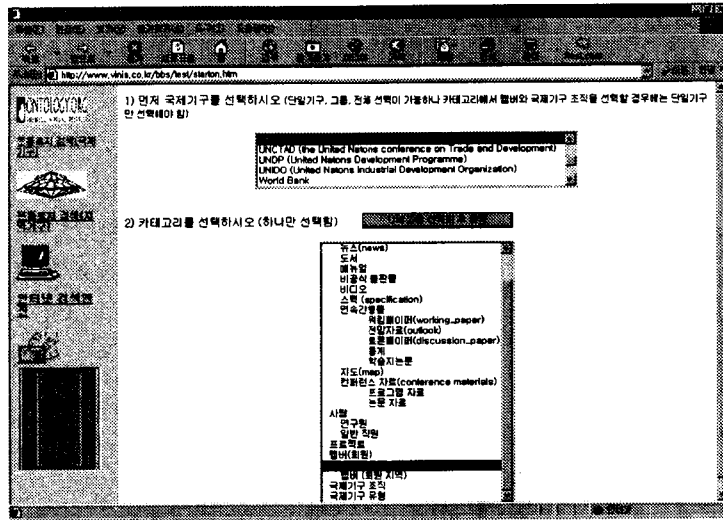
2) 멤버(회원 국가) 검색

(1) 온톨로지 선택

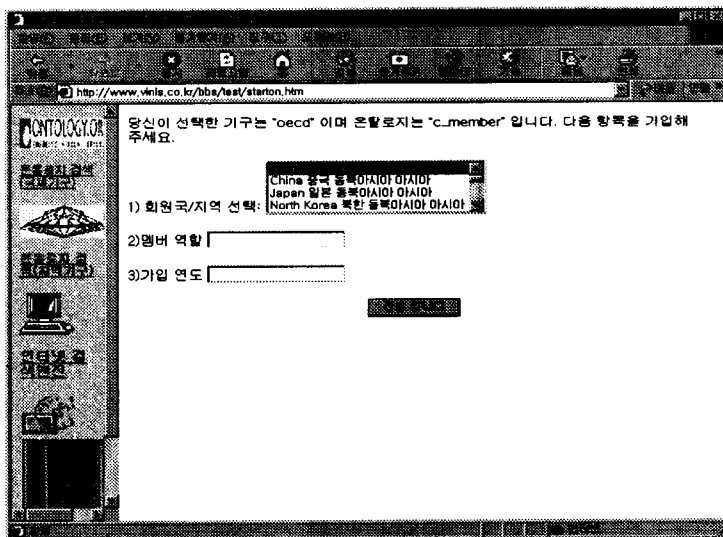
시스템의 초기 화면에서 검색할 국제기구와 온톨로지(카테고리)를 선택하는데 “OECD 회원 국가와 가입 년도”에 관련된 자료를 찾기 위해서 국제기구 OECD와 온톨로지로는 멤버(회원 국가)를 선정하였다(<그림 6> 참조).

|         |   |
|---------|---|
| 번호      | 353   |
| 출판물 코드  | pub   |
| 기구명     | IMF   |
| 표제      | Can Inheritances Alleviate the Fiscal Burden of an Aging Population?  |
| 저자      | Lueth, Erik   |
| 스피커     | :   |
| 년도      | 2001  |
| 통계 코드   | :   |
| 뉴스 발간월  | :   |
| 뉴스 발간일  | :   |
| 문헌유형(1) | serial  |
| 문헌유형(2) | worp  |
| 학술지명    | :   |
| 총서명     | Working Paper No. 01/97   |
| 회의명     | :   |
| 주제(1)   | 1   |
| 언어      | eng   |
| 국가/지역   | w0  |
| 원문 링크   | <a href="http://www.imf.org/external/pubs/ft/wp/2001/wp0197.pdf">http://www.imf.org/external/pubs/ft/wp/2001/wp0197.pdf</a> |

<그림 5> 검색 결과(상세 화면)



<그림 6> 멤버 온톨로지 선택화면

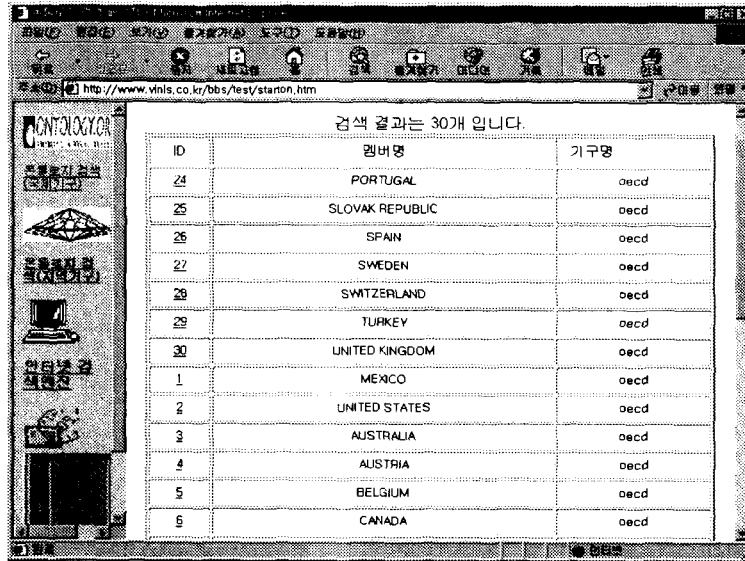


<그림 7> 멤버 온톨로지 속성값 입력

(2) 속성값 입력

시스템의 초기화면에서 선택한 온톨로지에 따라 속성값 입력 화면이 디스플레이 되는데 멤버(회원 국가)를 선택한 경

우 <그림 7>과 같이 회원국/지역, 멤버 역할, 가입연도의 3개의 속성이 출력된다. 회원국/지역 항목에 "all"을 선택하고 멤버 역할과 가입연도는 입력하지 않고 전



<그림 8> 검색결과 (간략 화면)

| 레코드 내용      |   |
|-------------|---|
| 번호          | 24  |
| 국제 기구명      | oecd  |
| 멤버명         | PORTUGAL  |
| 가입 연도       | 1961  |
| 역할          | ;   |
| 자세한 정보 (링크) | <a href="http://www.oecd.org/oecd/pages/document/displaywithoutnav/0,3376,EN-document-notheme-1-no-no-9464-0,00.html">http://www.oecd.org/oecd/pages/document/displaywithoutnav/0,3376,EN-document-notheme-1-no-no-9464-0,00.html</a> |

<그림 9> 검색결과 (상세화면)

송 버튼을 선택하였다. 이로서 멤버 역할과 가입 연도와는 관련없이 "OECD"의 모든 회원국에 대한 레코드들을 검색하였다.

### 3) 검색 결과

속성값을 입력하면 <그림 8>과 같은 간

략 화면이 출력된다. 이 결과에서 레코드 번호 "24"를 선택하면 <그림 9>와 같이 회원 국가인 PORTUGAL에 대한 상세 화면이 출력된다. 자세한 정보(링크)를 누르면 회원국들에 대한 정보를 기술한 웹사이트에 링크된다.

### 4.3 시스템 평가

구현한 시스템의 만족도를 조사하기 위해서 이 시스템의 검색 기능을 국제기구의 웹사이트인 전문 검색엔진(또는 일반 검색엔진)의 검색 기능과 적합성과 탐색 시간이라는 두 가지 요인을 중심으로 비교, 분석하였다. 다음은 실험 질문, 실험 방법과 절차 그리고 실험 결과에 대해서 기술한다.

#### 4.3.1 실험 질문

실험 질문은 실제 경제학 관련 연구기관 연구자들이 요구한 질문들을 재구성하여 사용하였으며 사용된 20개의 질문을 정보 유형별로 구분하면 다음과 같다.

#### 1) 출판물 관련 정보

##### (1) 일반 논문 관련 정보

- Q5) IMF, OECD 등에서 2000이후 발표된 고령화 관련 자료를 조사하고자 합니다.
- Q7) 세계의 해외직접투자(FDI)에 관한 동향/데이터를 조사하고자 합니다.
- Q12) 중남미 경제에 관한 working paper 정보를 입수하고자 합니다.
- Q13) 저자 Cerra, Valerie의 중국 무역에 관한 자료를 입수하고자 합니다.
- Q17) 2002년 이후 발표된 중국 경제에 관한 논문자료를 조사하고자 합니다.

##### (2) 통계 관련 정보

- Q3) 무역통계자료 중 SITC코드로 되어 있는 자료를 입수하고자 합니다.
- Q14) 최근(2002년 이후) 경제, 금융 관련 국제기구의 세계경제 전망 자료를 입수하고자 합니다.

##### (3) 컨퍼런스 관련 정보

- Q6) 아시아에 관한 컨퍼런스 자료 또는 컨퍼런스 프로그램을 조사하고자 합니다.

##### (4) 멀티미디어(비디오 이미지) 관련 정보

- Q8) 세계경제에 관한 토론 비디오 자료를 입수하고자 합니다.
- Q9) 세계의 인구 지도를 찾고자 합니다.
- Q19) 보고서 작성에 동아시아 지도를 활용하고자 합니다.

##### (5) 뉴스 관련 정보

- Q18) 국제기구에서 발표한 최근 아르헨티나 경제 관련 뉴스를 조사하고자 합니다.

#### 2) 인물 관련 정보

- Q1) 아시아 지역 연구자 명단과 연락처를 조사하고자 합니다.
- Q10) 국제기구의 라틴아메리카 전문가를 조사하고자 합니다.
- Q16) 국제기구의 전문가 중 한국 국적을 가지고 있는 사람을 조사하고자 합니다.



자 합니다.

3) 프로젝트 관련 정보

- Q2) 동아시아 지역에 어떠한 산업 프로젝트가 수행되고 있는지에 관하여 조사하고자 합니다.
- Q4) 국제기구에서 수행하는 대 북한 프로젝트와 그 규모(금액)에 관하여 조사하고자 합니다.
- Q11) 전후 아프카니스탄에 관련된 경제, 산업 프로젝트에 관하여 조사하고자 합니다.
- Q15) 베트남 개발에 관한 프로젝트는 어떠한 것이 있는지 프로젝트 명을 알려 주십시오.

4) 멤버(회원국) 관련 정보

- Q20) OECD 회원 국가와 가입 년도를 알고자 합니다.

4.3.2 실험 방법과 절차

사서 7명과 연구원 3명으로 구성된 10명의 탐색자들이 20개의 실험 질문을 가지고 본 연구에서 구축한 온톨로지 기반 시스템과 국제 기구 전문 검색엔진(또는 일반 검색엔진)을 각각 검색한 후 적합성<sup>4)</sup>과 탐색 시간을 비교해 보았다. 여기에서 전문 검색엔진과 일반 검색엔진을 자유롭게 사용하도록 하였는데, 이는 실제

로 이용자들이 탐색과정에서 어떠한 탐색 도구를 사용하여 검색을 하는지 알아보기 위함이다. 질문에 대한 적합성은 객관성을 유지하기 위하여 최종 검색건수가 15건을 넘는 경우는 15건만을 선정한 후 각 건수에 적합도를 부여하게 한 후 평균값을 사용하였다. 또한 최종 검색건수가 15건 미만인 경우에도 각 건수에 적합도를 부여하게 한 후 평균값을 사용하였다. 질문에 대한 적합성 판정은 탐색자가 탐색한 결과물에 대하여 판정 한 후 경제분야 전문 사서가 검토하여 최종적으로 정리하였다.

4.3.3 실험 결과 분석

4.3.3.1 검색효율 측정

검색효율의 측정은 전통적인 방법으로는 검색결과와 정도율과 재현율로서 효율을 측정하는 것이 바람직하다. 그러나 현재 인터넷 검색엔진의 경우 재현율의 측정이 불가능하고 동시에 온톨로지 기반 시스템에 적용하는 것도 적절하지 않다고 판단되어 본 연구에서는 두 검색시스템을 적합성과 탐색 시간을 기준으로 측정하였다.

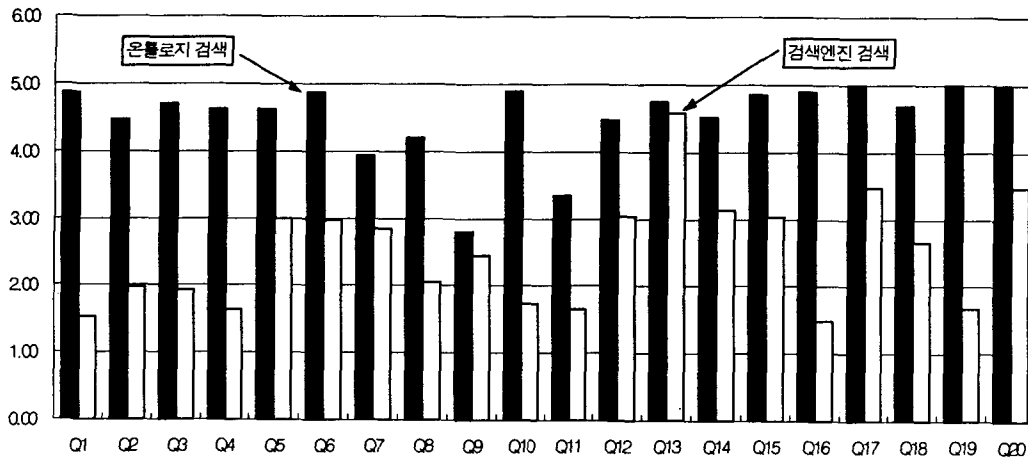
1) 적합성 측정

적합성 측정은 탐색자가 탐색한 결과물

---

4) 적합성 : 최종검색결과에 대한 적합도로 구간척도를 사용 하였다.  
 0: (미해결), 1: (매우 비적합), 2: (비적합)  
 3: (보통), 4: (비교적 적합), 5: (적합)

〈표 14〉 온톨로지 검색과 검색엔진 검색의 적합성 비교



에 대하여 1차적으로 측정하게 한 후 경제분야 전문 사서가 잘못된 부분을 수정하여 최종적으로 정리하였다. 〈표 14〉에 나타난 바와 같이 온톨로지 기반 시스템은 평균 4.53, 검색엔진은 평균 2.51로 온톨로지 기반 시스템의 적합도가 1.80배 높은 것으로 나타났다.

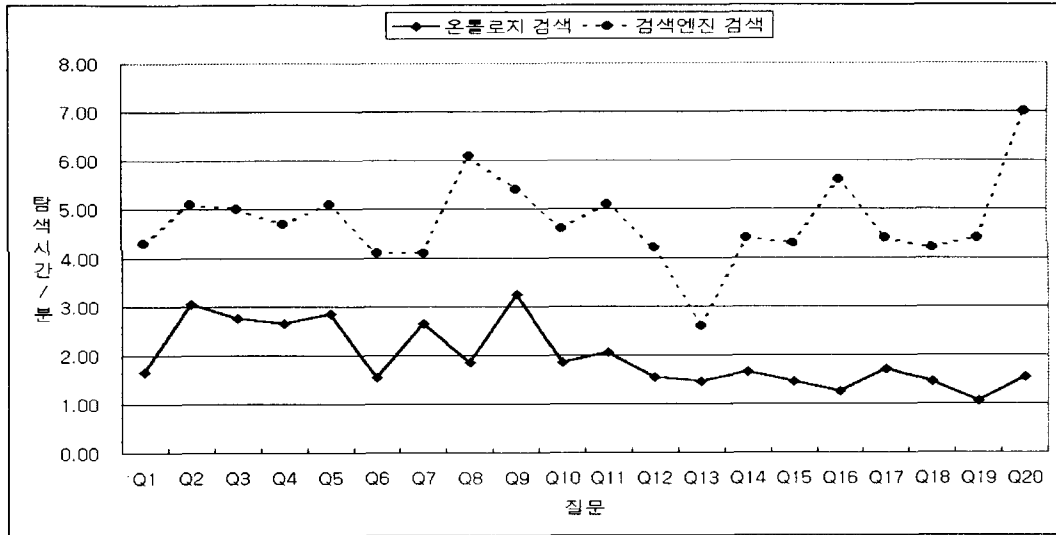
적합성 측정을 정보 유형별로 살펴보면 출판물 정보 중 일반 논문(Q5, Q7, Q12, Q13, Q17)의 경우 온톨로지 기반 시스템은 평균 4.61, 검색엔진은 평균 3.42로 나타나 온톨로지 기반 시스템의 적합도가 1.35배 높게 나타났다. 통계(Q3, Q14)의 경우 온톨로지 기반 시스템은 평균 4.61, 검색엔진은 평균 2.53으로 나타나 온톨로지 기반 시스템의 적합도가 1.82배 높게 나타났다. 컨퍼런스(Q6)의 경우 온톨로지 기반 시스템은 4.87, 검색엔진은 2.97로 나타나 온톨로지 기반 시스템의 적합도가

1.64배 높게 나타났다.

멀티미디어(비디오, 이미지)(Q8, Q9, Q19)의 경우 온톨로지 기반 시스템은 평균 4.00, 검색엔진은 평균 2.06로 나타나 온톨로지 기반 시스템의 적합도가 1.95배 높게 나타났다. 뉴스(Q18)의 경우 온톨로지 기반 시스템은 4.83, 검색엔진은 2.49로 나타나 온톨로지 기반 시스템의 적합도가 1.94배 높게 나타났다. 출판물 정보 중 멀티미디어가 1.95배로 가장 큰 차이를 보였으며 일반논문이 1.35배로 가장 작은 차이를 나타내고 있다.

인물 정보를 다루는 질문(Q1, Q10, Q16)의 경우 온톨로지 기반 시스템은 평균 4.89, 검색엔진은 평균 1.58로 나타나 온톨로지 기반 시스템의 적합도가 3.09배 높게 나타났다. 프로젝트 정보를 다루는 질문(Q2, Q4, Q11, Q15)의 경우 온톨로지 기반 시스템은 평균 4.33, 검색엔진은

〈표 15〉 온톨로지 검색과 검색엔진 검색의 시간비교



평균 2.07로 나타나 온톨로지 기반 시스템의 적합도가 2.09배 높게 나타났다. 멤버(회원국) 정보를 다루는 질문(Q20)에서는 온톨로지 기반 시스템은 5.00, 검색엔진은 3.56로 나타나 온톨로지 기반 시스템의 적합도가 1.40배 높게 나타났다.

온톨로지 검색에서 Q9, Q11의 적합도의 평균이 각각 2.81, 3.35로 낮은 결과를 나타내고 있는데, Q9는 지도정보를 구하는 질문으로 원문에 접근하였을 때 지도 이미지가 너무 늦게 구동되어 확인을 명확하게 하지 못한 결과이다. 또한 Q11의 경우에는 인물정보를 구하는 질문으로 내용 중 라틴아메리카 항목이 온톨로지 검색에서는 중남미로 설정이 되어 있어 탐색자들이 혼란을 가져와 명확하게 설정하지 못한 이유이다.

## 2) 탐색 시간

온톨로지 기반 시스템과 인터넷 검색엔진의 탐색 시간에 대한 자세한 정보가 〈표 15〉에 기술되었다. 10명의 탐색자가 20개의 질문 모두를 검색하는데 소요된 시간은 온톨로지 기반 시스템은 평균 1.96분, 검색엔진은 평균 4.74분으로 검색엔진은 온톨로지 기반 시스템 보다 2.42배의 탐색시간을 사용하고 있다.

### 4.3.3.2 논의

적합성 측정은 온톨로지 기반 시스템이 평균 1.80배 차이로 인터넷 검색엔진 보다 높은 것으로 나타났다. 정보 유형별로 살펴 보면 인물관련 정보가 3.09배의 차이를 보이며 가장 차이가 크게 나타났고 그 다음에는 컨퍼런스 관련 정보가 2.09

배의 차이를 보였고 멀티미디어 관련 정보와 뉴스 정보가 각각 1.95배, 1.94배로 나타나 평균값 보다 높게 나타났다. 한편 일반논문이 1.35배로 가장 작은 차이를 나타내고 있다. 이러한 결과를 보면 현재 검색엔진 등 탐색 도구들에서 많이 다루고 있지 않은 정보원의 경우 더 큰 차이를 보이고 있는 것으로 나타났다. 따라서 온톨로지 기반 시스템이 인터넷 검색엔진의 부족함 점들을 보완해 줄 수 있음을 보여주고 있다. 탐색 시간은 온톨로지 기반 시스템은 평균 1.96분, 검색엔진은 평균 4.74분으로 검색엔진은 온톨로지 기반 시스템 보다 2.42배의 검색시간을 사용하고 있다. 이러한 차이가 생겨난 것은 온톨로지 기반 시스템에서는 주어진 카테고리(클래스)와 설정된 항목(속성)을 이용한 탐색 경로를 활용함으로써 인터넷 검색 엔진에 비해 탐색 시간을 많이 줄일 수 있었기 때문이다.

검색 결과를 살펴본 결과 특징으로 나타난 점은 첫째, 정보 탐색 방법에서 탐색자들은 찾고자 하는 정보가 여러 사이트에 흩어져 있는 한 주제에 관한 탐색에서는 각 국제 기구의 전문 검색엔진을 일일이 체크하기 보다는 일반 검색엔진을 주로 사용한다는 점이 두드러지게 나타났다. 둘째, 실험에 참가한 탐색자들은 세 명의 연구원(S6, S7, S8)과 일곱 명의 전문 사서로 구성하였는데 이 중 연구원들은 탐색속도가 온톨로지 기반 시스템에서는 평균 3.50분(전체 평균 1.96분), 검색

엔진에서는 평균 8.23분(전체 평균 4.74분)으로 전체 평균 보다 훨씬 많은 탐색 시간을 사용한 것으로 나타났다. 이는 전문 사서들은 서지 정보를 주로 다루며 서지정보로서 정보 확인을 하는 탐색 행태를 가지고 있는 반면, 연구원들은 사서들과 달리 보통 최종 문헌을 읽고 확인하는 등의 절차를 거치기 때문에 검색 속도가 지연된 것으로 판단된다.

## 5 결론

최근 인터넷웹을 효율적으로 관리하고 이용하기 위하여 시맨틱웹과 온톨로지에 대한 연구가 활발히 진행되고 있다. 본 연구에서는 인터넷웹에서 경제 분야 국제기구에 관한 정보를 검색하기 위해서 국제기구 온톨로지를 설계하고 이 온톨로지에 기반하여 검색 시스템을 구현해 보았다. 구현된 시스템의 성능을 평가하기 위해서 실제 탐색 질문들을 이용하여 이 시스템과 인터넷 검색엔진에서 적합성과 탐색 시간이라는 두 가지 요인을 중심으로 비교, 분석해 보았다.

분석 결과, 적합성 측정은 온톨로지 기반 시스템이 평균 1.80배 차이로 인터넷 검색엔진 보다 높은 것으로 나타났다. 정보 유형별로 살펴 보면 인물 정보가 3.09배의 차이를 보이며 가장 차이가 크게 나타났다. 그 다음에는 컨퍼런스 정보가 2.09배의 차이를 보였고 일반논문은 1.35

배로 가장 작은 차이를 나타내었다. 이러한 결과를 보면 현재 검색엔진과 같은 탐색 도구들이 많이 다루고 있지 않은 정보원의 경우 더 큰 차이를 보이고 있는 것으로 나타났다. 따라서 온톨로지 기반 시스템이 인터넷 검색엔진의 부족한 점들을 보완해 줄 수 있음을 보여주고 있다. 탐색 시간은 온톨로지 기반 시스템은 평균 1.96분, 검색엔진은 평균 4.74분으로 검색엔진은 온톨로지 기반 시스템 보다 2.42배의 검색시간을 사용하고 있다. 이와 같은 차이를 보이는 것은 온톨로지 기반 시스템에서는 주어진 카테고리화 설정된 속성을 이용한 탐색 경로를 활용함으로써 좀 더 체계적인 검색을 할 수 있었기 때문이다.

온톨로지를 활용한 인터넷웹 정보 검색은 일차원적인 메타데이터를 이용하여 검색하는 것 보다 개념 간의 계층적 관계, 각 개념에 맞는 속성들을 활용할 수 있기 때문에 재현율과 정확율을 모두 높일 수 있는 방법으로 앞으로 많은 연구가 수행되어야 할 것이다. 다만 온톨로지에 맞춰 인터넷웹 페이지를 수작업으로 코딩하는 것은 너무 많은 시간과 노력이 요구되므로 이러한 과정을 자동화 또는 반자동화하는 방안이 요구된다. 특정 분야에 대한 온톨로지를 여러 분야의 사람들이 공동작업을 통해서 설계하고 이 온톨로지가 앞으로 생겨날 그 분야의 웹사이트에 대한 기준이 된다면 온톨로지는 정보 검색뿐만 아니라 시맨틱 웹 구성에도 유용하

게 이용될 수 있을 것이다. 현재는 온톨로지를 설계하고 그 온톨로지에 맞춰 인터넷웹 사이트를 코딩하는 것이 많은 시간과 노력이 들기 때문에, 많은 사람들에 의하여 자주 사용되며 중요한 정보를 갖고 있는 분야를 선정하여 온톨로지를 만드는 것이 바람직해 보인다.

## 〈참고문헌〉

- Borghoff, Uwe and Pareschi, Remo. 1998. *Information technology for knowledge management*. Berlin Heidelberg: Springer-Verlag.
- Deccaroni, Luigi. 2001. *Ontowedss-an ontology-based environmental decision-support system for the management of wastewater treatment plants*. Ph.D. diss., University Politecnica de Catalunya.
- Choo, Chun Wei. 2002. *Information management for the intelligent organization: the art of scanning the environment*. Medford: Information Today, Inc.
- Ding, Ying. 2001. "A review of ontologies with the semantic web in view." *Journal of Information Science*, 27(6): 377-384.
- Ding, Ying., Foo, Schubert. 2002. "Ontology

- research and development. Part 1: a review of ontology generation." *Journal of Information Science*, 28(2): 123-136.
- \_\_\_\_\_. 2002. "Ontology research and development. Part 2: a review of ontology mapping and evolving." *Journal of Information Science*, 28(5): 375-388.
- Fellbaum, C. 1998. *WordNet: an electronic lexical database, language, speech, and communication series*. Cambridge, Mass: MIT press.
- Fensel, Dieter. 2001. *Ontologies: a silver bullet for knowledge management and electronic commerce*. Berlin: Springer-Verlag.
- Fensel, Dieter et al. 1998. Ontobroker: or how to enable intelligent access to the WWW. In *Proceedings of 11th workshop on knowledge acquisition, modeling and management*. Berlin: Springer-Verlag.
- Gandon, Fabien. 2002. *Distributed Artificial Intelligence and Knowledge Management: ontologies and multi-agent systems for a corporate Semantic Web*. Ph.D. diss., INRIA and University of Nice-Sophia Antipolis.
- Guarino, N. 1998. *Formal ontology in information systems*. Oxford: IOS Press.
- Handschuh et al. 2001. *KAON-The Karlsruhe ONtology and Semantic Web Tool Suite*. Tech. rept. University of Karlsruhe.
- Heflin, Jeffrey Douglas et al. 1998. Reading between the lines: Using SHOE to discover implicit knowledge from the web. In *Proceedings of the AAAI-98 workshop on AI and information integration*. [Menlo Park: AAAI Press].
- Heflin, Jeffrey Douglas. 2001. *Towards the semantic web: knowledge representation in a dynamic, distributed environment*. Ph.D. diss., University of Maryland.
- Hjelm, Johan. 2001. *Creating the semantic web with RDF*. New York: John Wiley & Sons, Inc.
- Horrocks, Ian., Hendler, James. 2002. *The semantic Web-ISWC 2002 : first international semantic web conference sardinia, Italy, June 9-12, 2002 proceedings*. Berlin: Springer.
- Hyvönen, Eero. 2002. *The Semantic Web: the new internet of meanings in Semantic Web Kick-Off in Finland : vision, technologies, research and applications*. Hel-

- sinki : Helsinki Institute for Information Technology.
- Jian Qin. 2001. "Building and ontology for Chinese historical resources." In *Global digital library development in the New Millennium: fertile ground for distributed cross-disciplinary collaboration* edited by Ching-chih Chen. Beijing: Tsinghua University Press.
- Kaminski, Piotr. 2002. *Integrating information on the Semantic Web using partially ordered multi hypersets*. M.A. thesis, University of Victoria.
- Kampa, Simon Robert. 2002. *Who are the expert? e-scholars in the Semantic Web*. Ph.D. diss., University of Southampton.
- Kerschberg, Larry. 2000. "Conceptual models and Architectures for advanced information systems." *Applied Intelligence*, 13: 149-164.
- Luke et al. 1996. Ontology-based knowledge discovery on the world wide web In *Proceedings of the AAAI-98 workshop on Internet-based Information Systems*. [Menlo Park: AAAI press].
- Maedche, Alexander D. 2002. *Ontology learning for the semantic web*. London: Kluwer Academic Publishers.
- Mars, Nicolaas J. I. 1995. *Towards very large knowledge bases*. Oxford: IOS Press.
- McDonald, Craig. 2000. *Analysing and coordinating the conceptual structures of the information professions(a domain oriented approach)*. Ph.D. diss., University of New South Wales.
- Mizoguchi, Riichiro et al., 2002. Hozo: An Environment for Building/Using Ontologies Based on a Fundamental Consideration of "Role" and "Relationship." In *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, 13th International Conference, EKAW 2002 Sigüenza, Spain, October 1-4, 2002 Proceedings*. Berlin: Springer.
- Sowa, J. F. 1999. *Knowledge representation: logical, philosophical, and computational foundations*. London: Boors/Cole.
- Tamma, Valentina A.M. 2001. *An ontology model supporting multiple ontologies for knowledge sharing*. Ph.D. diss., University of Liverpool.

- Thuraisingham, Bhavani. 2002. *XML database and the semantic web*. London: CRC Press.
- Uschold, M. and Gruninger, M. 1996 "Ontologies: Principles, methods and applications", *The Knowledge Engineering Review*, 11(2): 93-136.
- Vasconcelos, Jose Angelo Braga de. 2001. *An ontology-driven organisational memory for managing group competencies*. Ph.D. diss., University of York.
- Weinstein, Peter. 1999. *Integrating ontological metadata: algorithms that predict semantic compatibility*. Ph.D. diss., Computer Science and Engineering in The University of Michigan.
- 古崎晃司. 2001. 『ロール・関係概念理論に基づくオントロジ-記述環境の開発と利用に関する研究』. 博士論文, 大阪大学 大学院, 工学研究科.
- 古崎晃司 等. 2002. 「ロール」および「関係」概念に関する基礎的考察に基づくオントロジ-記述環境の開発. 『人工知能学会論文誌』, 17 (3)B: 196-208.
- \_\_\_\_\_. 2002. オントロジ-構築・利用環境「法造」の開発と利用: 実規模プラントのオントロジを嚆として. 『人工知能学会論文誌』, 17(4)G: 407-419.
- 溝口理一郎 等. 1998. 『大規模知識ベースに関する調査研究報告書: オントロジ-工学に関する調査研究』. 東京: 日本情報処理開発協会.
- 尹洪珪. 2001. 『온톨로지 개념을 이용한 지리정보시스템 데이터의 검색』. 석사학위 논문, 인하대학교 대학원, 자동화공학과.
- 오삼균. 2002. 시맨틱웹 기술과 활용방안. 『情報管理學會誌』, 19(4): 297-319.
- 전자상거래 표준화 통합포럼. 2001. 『시맨틱 웹 기술을 적용한 전자상거래 표준 운용체계 연구』. [서울]: 동포럼.
- 崔武羅. 2001. 『분산환경에서 온톨로지 개념을 이용한 부품정보 검색』. 석사학위 논문, 인하대학교 대학원, 자동화공학과.



<부 록 1> DAML+OIL로 표현한 온톨로지

<!--네임스페이스 선언 -->

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/r
df-schema#"
xmlns:daml="http://www.daml.org/2001/03/daml+oil#" xmlns:xsd="http://www.w3.org/2000/10/XML
Schema#"
xmlns:base = "http://www.cs.umd.edu/projects/plus/DAML/onts/base1.0.daml#"
xmlns:dc="http://orlando.drc.com/daml/ontology/DC/3.1/dces-ont#" xmlns:ves="http://orlando.drc.co
m/daml/ontology/VES/3.1/drc-ves-ont#" xmlns:org="http://orlando.drc.com/daml/ontology/Organizati
on/3.1.1/Organization-ont#">
```

<!--International Organization Ontology -->

```
<daml:Ontology rdf:about="">
  <dc:title>International Organization </dc:title>
  <dc:location> http://www.vinis.co.kr/ontology/orgont1.0.daml</dc:location>
  <ves:version>1.0</ves:version>
  <ves:email>kimhh@mju.ac.kr</ves:email>
<dc:creator>HHKim</dc:creator>
  <ves:status>published</ves:status>
  <ves:statusDate>2002-11-12</ves:statusDate>
  <daml:imports rdf:resource="http://www.daml.org/2001/03/daml+oil"/>
  <daml:imports rdf:resource="http://orlando.drc.com/daml/Ontology/daml-extension/3.1
/daml-ext-ont"/>
  <daml:imports rdf:resource="http://orlando.drc.com/daml/ontology/UniversalProperty/
3.1/UniversalProperty-ont"/>
  <daml:comment>This is an ontology for an international organizational structure.</
daml:comment>
</daml:Ontology>
```

**<!-- Publication Class -->**

```
<daml:Class rdf:ID="Publication">
    <daml:label>publication</daml:label>
    <daml:subClassOf rdf:resource="http://www.cs.umd.edu/projects/plus/DAML/onts/bas
e1.0.daml#SHOEEntity" />
</daml:Class>
```

**<!-- Publication Properties -->**

```
<daml:DatatypeProperty rdf:ID="pubCode">
    <daml:domain rdf:resource="#publication"/>
    <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
    <daml:comment>publication's formal code </daml:comment>
</daml:DatatypeProperty>
```

```
<daml:DatatypeProperty rdf:ID="pubAuthor">
    <daml:domain rdf:resource="#publication"/>
    <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
    <daml:comment>publication's author </daml:comment>
</daml:DatatypeProperty>
```

.....생략.....

**<!-- Serial Class -->**

```
<daml:Class rdf:ID="Serial">
    <daml:label>serial</daml:label>
    <daml:subClassOf rdf:resource="#publication" />
</daml:Class>
```

**<!-- Serial Properties-->**

```
<daml:DatatypeProperty rdf:ID="pubFrequency">
  <daml:domain rdf:resource="#publication"/>
  <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
  <daml:comment>publication's frequency </daml:comment>
</daml:DatatypeProperty>
```

**<!-- journalArticle Class-->**

```
<daml:Class rdf:ID="journalArticle">
  <daml:label> journalArticle</daml:label>
  <daml:subClassOf rdf:resource="#publication" />
  <daml:subClassOf rdf:resource="#serial" />
</daml:Class>
```

.....생략.....

**<!-- Article Class-->**

```
<daml:Class rdf:ID="article">
  <daml:label>article</daml:label>
  <daml:subClassOf rdf:resource="#publication" />
</daml:Class>
```

.....생략.....

**<!-- UnofficialPublication Class -->**

```
<daml:Class rdf:ID=" UnofficialPublication >  
    <daml:label> UnofficialPublication </daml:label>  
    <daml:subClassOf rdf:resource="#publication" />  
</daml:Class>
```

**<!-- Person Class -->**

```
<daml:Class rdf:ID="Person">  
    <daml:label>person</daml:label>  
    <daml:subClassOf rdf:resource="http:// www.cs.umd.edu/projects/plus/DAML/onts/bas  
e1.0.daml#SHOEEntity" />  
</daml:Class>
```

**<!-- Researcher Class -->**

```
<daml:Class rdf:ID=" Researcher">  
    <daml:label> researcher </daml:label>  
    <daml:subClassOf rdf:resource="#Person" />  
</daml:Class>
```

**<!-- Staff Class -->**

```
</daml:Class><daml:Class rdf:ID="Staff">  
    <daml:label> Staff </daml:label>  
    <daml:subClassOf rdf:resource="#Person" />  
</daml:Class>
```

**<!-- Person Properties -->**

```
<daml:DatatypeProperty rdf:ID="name">
  <daml:domain rdf:resource="#Person"/>
  <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
</daml:DatatypeProperty>
```

.....생략.....

**<!-- Project Class -->**

```
<daml:Class rdf:ID="Project">
  <daml:label>project</daml:label>
  <daml:subClassOf rdf:resource="http://www.cs.umd.edu/projects/plus/DAML/onts/base
  1.0.daml#SHOEEntity" />
</daml:Class>
```

**<!--Project Properties -->**

```
<daml:DatatypeProperty rdf:ID="title">
  <daml:domain rdf:resource="#Project"/>
  <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
</daml:DatatypeProperty>
```

```
<daml:ObjectProperty rdf:ID="has_members">
  <daml:domain rdf:resource="#Project"/>
  <daml:range rdf:resource="#CountryMember"/>
</daml:ObjectProperty>
```

.....생략.....

**<!-- Member Class -->**

```
<daml:Class rdf:ID="Member">
  <daml:label>Member</daml:label>
  <daml:subClassOf rdf:resource="http://www.cs.umd.edu/projects/plus/DAML/onts/bas
e1.0.daml#SHOEEntity" />
  <daml:comment>A country or a group who has membership in an organization.</d
aml:comment>
</daml:Class>
  <daml:Class rdf:ID="CountryMember">
    <daml:label>Country Member</daml:label>
    <daml:subClassOf rdf:resource="#Member"/>
    <daml:comment>A country who has membership in an organization.</daml:comment>
  </daml:Class>
  <daml:Class rdf:ID="GroupMember">
    <daml:label>Group Member</daml:label>
    <daml:subClassOf rdf:resource="#Member"/>
    <daml:comment>A group who has membership in an organization.</daml:comment>
  </daml:Class>
```

**<!-- Member Properties -->**

```
<daml:DatatypeProperty rdf:ID="member_name">
  <daml:domain rdf:resource="#Member"/>
  <daml:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
  <daml:comment>member's name in an organization.</daml:comment>
</daml:DatatypeProperty>
```

.....생략.....

**<!-- Organization Class -->**

```
<daml:Class rdf:ID="Organization">
  <daml:comment>An organization is something made up of units with varied functions that
  contribute to the whole and to collective functions.</daml:comment>
</daml:Class>
```

**<!-- Organization Properties -->**

```
<daml:ObjectProperty rdf:ID="Org_type">
  <daml:domain rdf:resource="#Organization"/>
  <daml:range rdf:resource="#OrganizationType"/>
</daml:ObjectProperty>
```

**<!-- OrganizationType Class (Enumeration) -->**

```
<daml:Class rdf:ID="OrganizationType">
  <daml:oneOf rdf:parseType="daml:collection">
    <daml:Thing rdf:ID="Finance_org"/>
    <daml:Thing rdf:ID="Economic_org"/>
    <daml:Thing rdf:ID="Trade_org"/>
    <daml:Thing rdf:ID="Industrial_development_org"/>
  </daml:oneOf>
</daml:Class>
```

..... 생략 .....

```
</rdf:RDF>
```