



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

碩 士 學 位 論 文

시맨틱 웹 기반의 연구자정보
검색 인터페이스 설계에 관한 연구



漢城大學校 大學院

文獻情報學科

文獻情報學專攻

朴 美 香

碩 士 學 位 論 文
指導教授 徐 恩 璟

시맨틱 웹 기반의 연구자정보
검색 인터페이스 설계에 관한 연구

Designing the Researcher Information Retrieval
Interface Based Semantic Web



2008年 12月 日

HANSUNG
UNIVERSITY

漢城大學校 大學院

文獻情報學科

文獻情報學專攻

朴 美 香

碩 士 學 位 論 文

指導教授 徐 恩 璟

시맨틱 웹 기반의 연구자정보
검색 인터페이스 설계에 관한 연구

Designing the Researcher Information Retrieval
Interface Based Semantic Web

위 論文을 圖書館學 碩士學位論文으로 提出함

2008年 12月 日

漢城大學校 大學院

文獻情報學科

文獻情報學專攻

朴 美 香

朴美香의 圖書館學碩士學位論文을 認准함

2008年 12月 日



HANSUNG
UNIVERSITY

審査委員長 (인)

審査委員 (인)

審査委員 (인)

목 차

국문초록	v
------------	---

I. 서론

1. 연구의 필요성 및 목적	1
2. 연구의 방법 및 범위	3

II. 이론적 배경

1. 시맨틱 웹을 위한 온톨로지 기법	5
1.1 온톨로지 정의	6
1.2 온톨로지 기술 언어	8
2. 선행연구	12
2.1 온톨로지 기법을 이용한 연구	12
2.2 인물 및 연구자정보 검색에 관한 연구	15

III. 연구자정보 온톨로지 분석

1. 연구자정보 기술요소 분석	19
1.1 인물정보 DB 조사	19
1.2 인물정보 DB 기술요소 비교	28
1.3 연구자정보 기술요소 선정	31
2. 연구자정보 온톨로지 설계	32
2.1 클래스 이름 및 클래스 계층구조 정의	34
2.2 속성의 선정	37
2.3 속성의 목록 정의	40
2.4 속성의 제약 조건 정의	43

2.5 관계추론을 위한 공리 정의	47
IV. 연구자정보 검색 인터페이스 설계	
1. 인명 검색	52
1.1 인적정보	53
1.2 연구업적정보	54
1.3 소속정보	55
1.4 강의정보	56
1.5 관련인물	57
2. 주제어 검색	57
2.1 학술논문	58
2.2 저·역서	59
2.3 프로젝트	59
2.4 연구자군	60
2.5 강의명	60
3. 소속 검색	61
4. 상세 검색	64
V. 결론	65
참고문헌	68
ABSTRACT	73

표 목 차

<표 1> 조인스닷컴 기술요소 항목	21
<표 2> 조선닷컴 기술요소 항목	22
<표 3> 로마켓 기술요소 항목	23
<표 4> KRI 기술요소 항목	24
<표 5> NTIS 기술요소 항목	25
<표 6> SESTAT 기술요소 항목	26
<표 7> ReaD 기술요소	27
<표 8> COS Scholar Universe 기술요소 항목	27
<표 9> 인물정보 검색시스템의 기술요소	28
<표 10> 각 시스템별 기술요소 비교	30
<표 11> 범주별 추출된 공통 기술요소	32
<표 12> 연구자정보 클래스 사전	36
<표 13> 연구자정보 클래스 관계 사전	38
<표 14> 연구자정보 속성 목록	40
<표 15> OWL로 기술된 인적정보 결과 일부	45
<표 16> 연구자정보 공리 목록서	49
<표 17> 연구자정보 관계 테이블	50

그 림 목 차

<그림 1> 웹 자원 간의 연결 관계	6
<그림 2> 온톨로지의 분류	7
<그림 3> 토픽 공간과 리소스 공간	11
<그림 4> EOE 방법론 프로세스	33

<그림 5> 연구자정보 개념적 설계 과정	34
<그림 6> protégé에서 표현된 계층화된 클래스들	37
<그림 7> protégé에서 personallInfo의 속성 제약 화면	44
<그림 8> 연구자정보 클래스간의 관계와 추론 정의	48
<그림 9> 연구자정보 검색시스템 로그인 화면	51
<그림 10> 인명 검색 관계정의	52
<그림 11> 연구자 이름을 통한 인적정보 브라우징 결과 화면 ...	53
<그림 12> 연구자의 연구업적정보 브라우징 결과 화면	54
<그림 13> 연구자의 강의정보 브라우징 결과 화면	56
<그림 14> 주제 검색 관계정의	57
<그림 15> 주제어 검색을 이용한 학술논문 브라우징 결과	58
<그림 16> 주제어검색을 이용한 강의정보 검색결과	61
<그림 17> 소속검색 관계정의	62
<그림 18> 연구자군 브라우징 결과 화면	63
<그림 19> 상세검색 화면	64

국 문 초 록

인터넷이 지속적으로 발달함에 따라 이용자들은 검색엔진의 도움을 받아 다양한 정보를 탐색하게 되었고, ‘정보화시대’라는 말에 걸맞게 넘쳐나는 수많은 정보 중에서 사용자가 원하는 정보를 정확하게 찾는 것이 가장 중요한 문제로 대두되고 있다. 최근에는 이러한 문제의 대안으로 시맨틱 웹 기술이 정보검색 시스템에 응용되고 있다. 시맨틱 웹은 검색환경이 단순히 정보를 검색하고 검색된 정보를 나열하여 주는 것이 아니라, 정보의 적합성과 관련성을 컴퓨터가 판단하여 제공해주는 웹 환경을 말한다.

본 연구는 여러 분야에서 연구를 진행하고 있는 연구자에게 연구의 근간이 되는 관련 주제의 연구자정보를 체계적이며 포괄적으로 제공하는 검색 인터페이스를 설계하는 것이 목적이므로 시맨틱 웹 기술을 이용해 보다 동적이며 의미연관성이 있고, 쉽게 정보에 접근할 수 있는 검색 인터페이스를 제안하고자 한다.

이를 위해 먼저 연구자정보의 속성과 이에 대응하는 클래스를 추출하기 위하여 국내외 인물 및 연구자정보 데이터베이스의 기술요소들을 조사하여 비교하여 가장 필수적이며 범용적인 기술요소를 선정하였다. 이러한 즉 도메인에 대한 요구분석 결과를 토대로 하여 연구자정보 온톨로지 분석을 EOE 방법론에 따라 단계적으로 실행하였다. 선정된 연구자정보 기술요소들을 범주화하여 클래스를 결정하고, 각 클래스의 정의와 관계를 정의한 다음, 각 클래스에 속하는 속성들을 추출하고 속성간의 계층적 관계, 핵심클래스와 행위클래스간의 계층적 관계를 정의하였다. 각 속성에 대한 데이터 타입, 범위, 차수 등을 정의하였고 속성간의 제약조건을 정의하였으며, 마지막으로 관계추론을 실행시키기 위한 공리목록서와 관계테이블을 정의하였다. 온톨로지를 기술하는 언어로는 OWL을 이용하였고 protégé 프로그램을 이용하여 클래스를 계층화하고 제약조건을 정의하였다.

이러한 연구를 통하여 다음과 같은 연구결과를 얻었다.

첫째, 연구자정보에 대한 요구분석을 한 결과 연구자정보에는 인적정보,

연구업적정보, 직무정보, 소속정보가 필요하다는 것을 알 수 있었다.

둘째, 핵심클래스는 행위클래스를 가지고 각 클래스간의 관계를 정의하는 것이 검색에 도움을 줄 수 있는 것이라고 판명하였다.

셋째, 온톨로지 분석을 통해 추론된 관계를 바탕으로 검색 인터페이스를 설계하였다.

위 연구결과를 토대로 본 연구에서 제시한 연구자정보 검색 인터페이스를 통해 ‘사회관계정보’가 추출될 수 있음을 알 수 있었고 연구 분야로서 밀접한 다른 연구자 그룹, 같은 주제를 강의하는 또는 연구하는 연구자 그룹, 소속된 기관에 대한 기관소개 및 임원진 정보 등을 즉각적으로 브라우징 할 수 있었다. 또한, 기존의 웹 정보원도 관계정의를 수행함에 따라 검색결과를 보다 다양하게 브라우징 할 수 있었고, 유연한 데이터 검색이 가능하였다. 이처럼 구조화되고 계층화된 데이터들을 관계정의를 통해 추론이 가능하다면 이용자에게 다양하고 동적인 검색서비스를 제공할 수 있을 것이고, 이러한 시맨틱 웹을 이용한 관계기반의 웹 검색은 유연한 검색 환경을 구축하는데 이바지 할 수 있을 것이다.



I. 서론

1. 연구 필요성 및 목적

정보검색은 1990년대 이전만 해도 도서관에서 서지정보를 통한 문헌정보 검색이거나 상용데이터베이스를 이용한 색인초록 검색이 대부분을 차지하였으나 그 후, 원문 데이터베이스가 구축됨에 따라 전문검색이 가능한 정보검색 시스템이 보편화되기 시작하였다. 최근 정보검색 환경이 웹 환경 기반으로 전환됨에 따라 이용자들은 보다 편리한 방법으로 원하는 데이터베이스를 선택할 수 있게 되었고, 지능적인 검색엔진의 도움을 받아 다양한 형태의 웹정보 자원을 쉽게 탐색하거나 브라우징할 수 있게 되었다.

웹 환경에서의 정보검색은 구조화된 디렉토리나 링크로 연결된 많은 접근점을 기반으로 하는 브라우징 탐색과 이용자가 입력한 단어의 매칭을 기반으로 하는 키워드 탐색이 주를 이루고 있다. 그러나 이러한 기존의 검색 방식은 이용자가 질의 시 제시한 정보요구의 의미와 범위를 컴퓨터가 인지하지 못함에 따라 이용자는 찾고자 하는 정보를 발견할 때까지 정보 질의와 검색해야하는 정보원들을 변경해야 하는 문제점을 갖고 있다¹⁾. 따라서 이용자의 정보요구에 적합하고 신뢰성 있는 정보만이 검색될 수 있는 환경이 요구되고 있다. 즉, 적합한 정보를 보다 효율적으로 얻을 수 있는 방법에 대한 논의가 끊임없이 제기되고 있는 상황이다.

최근에는 유의미한 정보검색을 추구하며 관련 문헌들을 보다 적합하게 제공할 수 있는 시맨틱 웹 기술이 웹 환경 기반의 정보검색 시스템에 응용되고 있다. 시맨틱 웹은 웹 검색환경이 단순히 정보원을 검색하고 검색된 정보원을 나열하여 주는 것이 아니라, 이용자가 원하는 정보를 구조적으로 그룹화해주고, 정보의 적합성과 관련성을 컴퓨터가 자동 처리해주어 이용자의 정보검색 목적을 성취할 수 있도록 해주는 웹 환경을 말한다²⁾.

1) Heting, Chu. 2006. 『디지털 시대의 정보문헌과 검색』. 장혜란 옮김. 서울: 한국도서관협회, pp. 159-167.

특히 정보검색 기능에 이용되는 시맨틱 웹은 이용자에게 적절한 정보를 제공하는 것뿐 만 아니라 최종적으로는 컴퓨터에 의한 추론기능을 이용하여 이용자가 필요로 하는 정보를 자동적으로 수집하는 에이전트의 구축이 가능하게 된다³⁾.

본 연구는 이러한 특성을 가지는 시맨틱 웹을 기반으로 하여 연구자정보 검색 인터페이스를 설계하고자 한다. 최근 연구자에 대한 정확한 정보와 상호연관성에 대한 정보요구가 많아지게 되자 다양한 웹사이트에서 연구자정보를 제공하고 있다. 즉, 포털 정보 서비스업체는 일차적으로 인물 프로파일은 물론이고 과거 인쇄물 형태에서 제공할 수 없었던 인물 관련 정보를 제공하고 있으며 정부 기관에서 개발한 연구자정보서비스시스템들은 연구자에 대한 개별 연구업적이나 프로젝트에 대한 검색 및 링크를 목적으로 관련 정보 및 연구자의 폭넓은 브라우징이 가능하도록 하고 있다⁴⁾. 그러나 이러한 인물 또는 연구자정보 검색시스템은 단편적 탐색만이 가능하고 그 탐색결과도 포괄적이지 못하고⁵⁾, 데이터갱신도 수작업으로 이루어짐에 따라 최신 정보검색이 잘 이루어지지 못하는 한계점을 가진다. 이러한 한계점을 극복하고 여러 분야에서 연구를 진행하고 있는 연구자에게 연구의 근간이 되는 관련 주제의 연구자정보를 체계적이며 포괄적으로 제공할 수 있는 시맨틱 웹 기반의 검색 인터페이스를 설계하는 것이 본 연구의 목적이다.

연구자정보 검색 인터페이스를 시맨틱 웹 기반으로 구축함으로써 연구자정보에 대한 정확한 사실정보를 검색할 수 있을 뿐만 아니라 이를 기반으로 사회적 관계정보 또는 연구 관계정보 등이 추론되어 검색될 수 있을 것이다. 또한 관계 정의에 따라 연계된 사이트로 인하여 자동적으로 연구자정보가 갱신될 수 있다. 이러한 특징을 지닌 시맨틱 웹의 분석을 기반으로 한 검색 인터페이스는 이용자에게 동적이며 특정성 높은 검색은 물론

2) 현승순. 2006. 차세대 지능형 시맨틱 웹 & 온톨로지. 『주간기술동향』, 정보통신연구진흥원. 1265: 25.

3) 최석두, 김이겸. 2008. 『온톨로지 알고리즘 II』. 서울: 한울. p. 106.

4) 유수현 외, 2007. 과학기술분야 해외 인력정보서비스 평가. 『한국정보관리학회 학술대회 논문집』, 14: 105.

5) 박윤희. 2008. 『온톨로지 인스턴스 생성을 위한 인물정보 추출』. 석사학위논문, 경북대학교 대학원.

다각적인 브라우징을 할 수 있게 하며 궁극적으로 이용자의 검색만족도를 높일 수 있을 것이다. 또한 향후 도메인 중심의 시맨틱 웹 기반 검색시스템 구축 시 본 연구가 도움이 될 수 있을 것이라 본다.

2. 연구의 범위와 방법

본 연구는 웹상에 존재하는 자원들에 대해 그 관계를 기술하고 검색 에이전트를 이용하여 의미를 추론하게 함으로써 필요로 하는 정보를 보다 정확하게 그리고 포괄적으로 찾아내주는 기능을 하는 시맨틱 웹을 기반의 연구자정보 검색시스템을 구축 시 전제조건으로 필요하는 연구자정보에 대한 온톨로지 분석과 연구자정보 검색 인터페이스를 설계하였다.

이를 위하여 본 연구가 수행한 방법을 다음과 같다. 첫째, 시맨틱 웹과 온톨로지에 대한 개념과 그 특징을 알아보고 온톨로지를 이용하여 수행된 다양한 선행연구들과 국내와 국외의 연구자 정보 검색시스템에 대하여 고찰하였다.

둘째, 시맨틱 웹이 구현될 수 있도록 연구자정보 도메인에 대한 온톨로지를 설계하였다. 이를 위해 먼저 연구자정보의 속성과 이에 대응하는 클래스를 추출하기 위하여 국내외 인물 및 연구자정보 데이터베이스의 기술요소들을 조사하여 비교하여 가장 필수적이며 범용적인 기술요소를 선정하였다.

셋째, 이러한 즉 도메인에 대한 요구분석 결과를 토대로 하여 연구자정보 온톨로지 분석을 EOE 방법론⁶⁾에 따라 단계적으로 실행하였다. 따라서 본 연구는 온톨로지 분석을 다음과 같이 다섯 단계로 나누어 단계적으로 수행하였다: 1) 선정된 연구자정보 기술요소들을 범주화하여 클래스를 결정하고, 각 클래스의 정의와 관계를 정의하였다; 2) 각 클래스에 속하는 속성들을 추출하고 속성간의 계층적 관계, 핵심클래스와 행위클래스간의

6) 한국전산원. 2004. 『웹 온톨로지 개발 지침 연구』. 서울: 한국전산원.

계층적 관계를 정의하였다; 3) 각 속성에 대한 데이터타입, 범위, 차수 등을 정의하였고; 4) 속성간의 제약조건을 정의하였다; 5) 마지막으로 관계 추론을 실행시키기 위한 공리목록서와 관계테이블을 정의하였다. 온톨로지를 기술하는 언어로는 OWL을 이용하였고 protégé 프로그램을 이용하여 클래스를 계층화하고 제약조건을 정의하였다.

넷째, 구조화된 온톨로지를 기반으로 연구자정보를 검색할 수 있는 검색 인터페이스를 제안하였다. 본 연구는 인명, 주제어, 소속기관명을 이용하여 연구자와 관련되는 모든 정보를 브라우징할 수 있는 인터페이스와 정의된 속성의 조합을 이용하여 특정 정보만을 검색하게 하는 인터페이스를 제안하였다.

본 연구는 다음과 같은 한계점을 가진다. 첫째, 연구자정보 도메인에 대한 요구분석을 실제 연구자에게 직접적으로 얻는 의견을 토대로 수행한 것이 아니라 기존의 인물정보/연구자정보 제공 시스템이 연구자의 요구분석을 토대로 구축되었을 것이라는 전제하에 연구자정보 제공 시스템에서 제공하는 기술요소를 분석하는 것으로 요구분석을 대신하였다.

둘째, 연구자정보 검색화면과 검색결과 화면은 검색 인터페이스 방안을 제안하기 위하여 실험적으로 구축되었다. 따라서 이미지의 출처, 검색 속도, 시각적 명료성, 브라우징 속도에 대한 문제점을 논할 수 없다.

셋째, 본 연구에서는 연구자정보 데이터를 구축하지 않았으므로 검색 인터페이스에 나타난 검색 결과는 임의적으로 보여진 것이다.

II. 이론적 배경

1. 시맨틱 웹을 위한 온톨로지 기법

시맨틱 웹은 "컴퓨터가 정보의 의미를 이해하고 의미를 조작할 수 있는 웹"이라고 할 수 있다⁷⁾. 시맨틱 웹은 기존의 웹과 완전히 구별되는 것이 아니라, 현재 웹을 확장하여 웹에 올라오는 정보에 잘 정의된 의미를 부여하고 이를 통해 컴퓨터와 사람이 협동적으로 작업을 수행할 수 있도록 하는 패러다임을 제공하는 것이다⁸⁾.

시맨틱 웹의 목적은 웹상의 정보에 잘 정의된 의미를 부여하여 컴퓨터가 쉽게 정보자원의 의미를 해석할 수 있도록 함으로써 정보의 검색·통합·해석이 자동적으로 이루어지고 효율화시키는 것이다⁹⁾. 특히 시맨틱 웹은 웹상에 존재하는 자원들에 대해 그 관계를 기술하고 검색 에이전트를 이용하여 의미를 추론하게 함으로써 필요로 하는 정보를 보다 정확하게 찾아내주는 기능과 더불어 현재 웹에서 실행되고 있는 디렉토리 검색 또는 키워드 매칭을 통한 검색으로는 동일한 의미를 갖는 다른 정보나 이용자가 원하는 정확한 의미를 갖는 정보를 검색할 수 없는 한계점을 극복할 수 있게 해준다. <그림 1>은 현재의 웹과 시맨틱 웹 자원간의 연결의 차이점을 보여주고 있다¹⁰⁾. 즉 현재의 웹을 하이퍼링크로 자원간의 연결만 제공해 준다면 시맨틱 웹을 자원간의 관계의 의미를 기술해줌으로써 보다 적합한 자원만을 검색할 수 있거나 자동추론으로 인하여 관련된 자원이 검색될 수 있다.

따라서 시맨틱 웹은 기계 스스로 개념화 과정과 추론 과정을 통해 사람이 원하는 정보를 찾아주는 환경을 말한다. 이때 이러한 과정을 자동적으

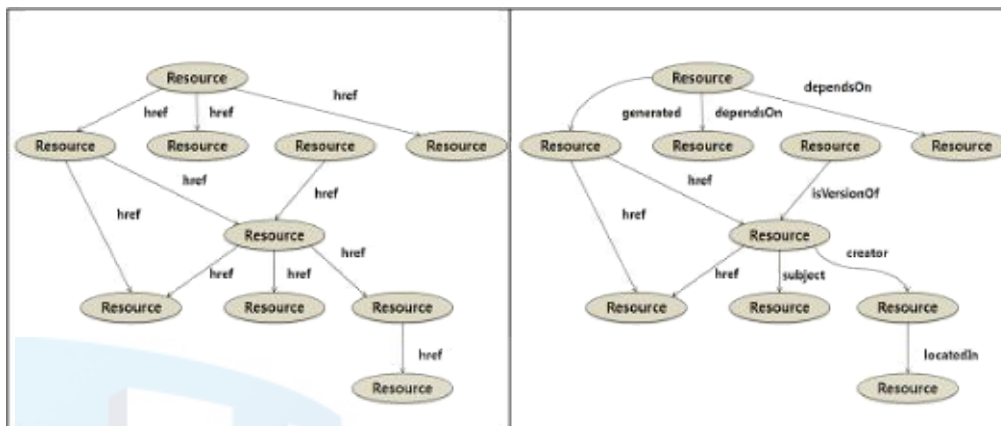
7) 현승순. 2006. 차세대 지능형 시맨틱 웹 & 온톨로지. 『주간기술동향』, 정보통신연구진흥원. 1265: 25.

8) 최중민. 2003. 시맨틱 웹의 개요와 연구동향. 『정보과학회지』, 21(3): 4.

9) 김수경, 안기홍. 2005. 시맨틱 웹 기반의 비교구매 에이전트를 위한 동적 웹 온톨로지에 대한 연구. 『한국지능정보시스템학회논문지』, 11(2): 31-45

10) 상계서.

로 처리하기 위해 실세계의 지식을 기계가 처리가능 하도록 표현하는 것이 중요하다. 이런 표현이 가능하도록 해주는 것이 온톨로지이다. 온톨로지 기법이란 특정한 주제 영역을 표현하는 데이터 모델로서 일반적으로 개념과 개념 사이의 관계를 기술해주는 것을 말한다. 즉, 종속관계, 부분-전체관계 등의 다양한 관계를 정형화된 어휘로 기술하도록 하여 개념간의 관계를 추론할 수 있도록 하고 보다 논리적이며 의미적 행위 또는 의사결정이 자동적으로 처리가능하게 해준다.



a. 현재의 웹

b. 시맨틱 웹

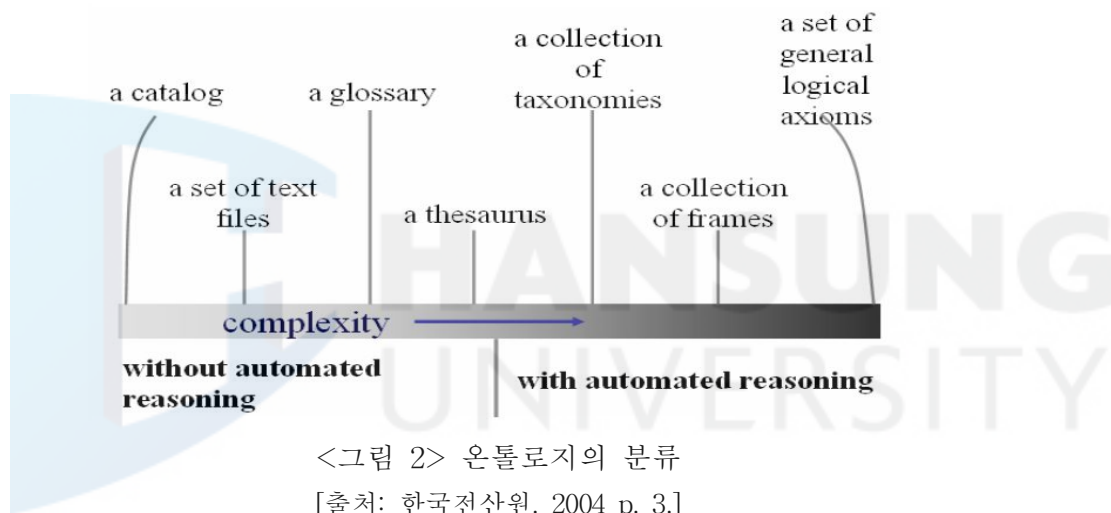
<그림 1> 웹 자원 간의 연결 관계

[출처: 김수경, 안기홍, 2005, p. 308.]

1.1 온톨로지 정의

사람들은 일반적으로 세상에 있는 각각의 사물이나 사건을 경험하면서 이들 속에 있는 특징을 파악하고 이를 언어로 개념화 시킨다. 온톨로지는 컴퓨터가 이러한 특징을 사람과 같이 이해할 수 있도록 개념을 관계정의로 기술해주는 것으로 궁극적으로 컴퓨터가 인간과 같은 인식을 할 수 있게 하는 기법이라 할 수 있다. 온톨로지는 1967년 S. H. Mealy에 의해 전산학 분야에 도입되어 사용되었으며, 최근에는 인공지능분야에서 시맨틱 웹을 표현하는 중요한 언어로 각광받고 있다¹¹⁾. 본래 온톨로지라는 용어는

철학의 “존재론”에서 온 것으로 고대 그리스에서 존재론은 자연법칙이나 사물의 형태에 대해서 존재론적 관점으로 각각의 문제를 다루었으며, 그 문제에 대한 해결책으로 공용어로 기술하고, 구성 개체들 간의 관계를 상세히 표현하는 것에서 시작되었다¹²⁾. 그러나 현재 정보학 관점에서의 온톨로지는 “특정 분야를 개념화하기 위하여 명시적으로 정형화한 명세서”라고 말할 수 있고, 이러한 그루버의 정의는 비교적 널리 받아들여지고 있다¹³⁾. 여전히 명세화(specification)와 개념화(conceptualization)에 대한 명확한 정의가 내려지고 있지는 않지만 어떠한 실체(reality)를 기술하기 위한 특정한 어휘의 집합과 그리고 의도된 의미를 가지고 있고 어휘간의 위계적인 구조, 명시적인 가정, 기타 다른 관계나 제약을 표현하는 공리를 정의해주는 것이 온톨로지라고 보았다¹⁴⁾.



<그림 2>는 온톨로지의 분류를 나타낸 것으로 ‘온톨로지’로 분류되

- 11) 서 휘. 2006. 온톨로지 자동구축을 위한 OWL의 어휘와 구문 사용방법에 대한 이론적 연구. 『한국도서관·정보학회지』, 37(2): 192.
- 12) 최원중. 2004. 『온톨로지 계층 정보를 이용한 웹 서비스 발견 알고리즘』. 석사학위 논문, 한양대학교 대학원.
- 13) T. R. Gruber. 1993. "A translation approach to portable ontology specification.", *Knowledge Acquisition*, 5: 199-220.
- 14) N. Guarino. 1998. "Formal ontology and information systems, In N. Guarino(ed.).", *Formal Ontology in Information Systems*. ISO Press.

는 많은 것들이 매우 넓은 스펙트럼 내에 병존하고 있다¹⁵⁾.

최근에 온톨로지는 효율적인 정보통합과 정보 검색 그리고 지식 관리 등의 기능이 필요한 전자학습, 디지털 도서관, 지능형 에이전트 그리고 정보검색 시스템 등의 분야에서 필요한 정보에 대하여 연계적인 추출이 가능한 지능적인 추론 등에 활용되고 있다¹⁶⁾. 특히 정보검색에서의 온톨로지는 개념을 명확히 정의해줌으로써 탐색 시 개념의 모호성을 없애주고 용어적 혼동을 줄여줄 수 있다. 또한, 속성으로 기술된 사항들이 색인화가 됨에 따라 포괄적이고 정확하게 검색이 가능하고, 관계정의와 공리정의를 기반으로 하여 관련된 다른 지식은 물론 추론적 정보까지도 검색이 가능하게 해주는 장점을 가진다. 다만 아직까지도 이러한 기능을 수행하기 위해서는 복잡한 절차에 따른 분석과 기술을 요구한다는 단점을 가진다.

1.2 온톨로지 기술 언어

1) RDF/OWL

현재 시맨틱 웹의 표준은 데이터 층과 온톨로지 층 두 개의 층으로 이루어져 있다. 이중 RDF(Resource Description Framework)는 데이터 층이며, 기계가독형 정보의 통합과 교환을 위한 표현방식을 제공한다. RDF는 W3C에서 제정한 표준으로서 자원, 속성, 속성 값 세부분으로 자원을 기술하는 메타데이터이다. 즉 자원의 특성을 표현하는 속성을 통해서 자원과 속성값 사이의 관계를 나타내준다. 그러나 RDF는 그 자체만으로는 속성에 대해 설명할 방법을 제공하지 않을 뿐 아니라, 속성간의 관계나 다른 자원간의 관계를 표현할 방법 또한 제공하지 않는 단점을 가진다. RDF에서 자원들은 URI로 식별될 수 있다. 어떤 자원들은 무명 자원으로 식별이 불가능하다. 속성도 자원으로 구분되어서 URI를 통하여 식별될 수 있다. 속성값은 자원이거나 값을 가지는 문자열로 정의될 수 있다.

15) 한국전산원. 2004. 전계서.

16) 서 휘. 2006. p. 193.

다른 층인 온톨로지 층은 간단한 RDF 스키마 언어와 보다 상세한 웹 온톨로지 언어인 OWL(Web Ontology Language)을 지원한다. OWL은 시맨틱 웹의 실현을 위해, 데이터 표현 단계인 RDF를 보조하기 개발되었으며, RDF 스키마가 제공하는 기본적인 기능을 넘어서, 기계가 자원에 대한 추론을 가능하게 한다. OWL은 기술될 수 있는 형식과 어휘 면에서 다른 언어 즉, XML이라든지 RDF보다 풍부하며 의미표현 수단을 더 많이 제공함으로써 웹상에서 컴퓨터가 해석할 수 있는 정보를 작성하는데 있어 이들 언어보다 뛰어난 성능을 보인다. OWL은 분류 법칙을 이용한 검색 기능 확장이 가능하며 따라서 인터넷 포털 등에서 응용될 수 있다. 또한 멀티미디어 내용에 기반을 둔 검색, 데이터와 문서의 자동 텍사노미 생성, 사용자의 선택사항 및 관심사항 표현이 가능한 지능형 에이전트, 웹서비스 발견 및 구성, 저작권 및 접근 관리 등에 응용될 수 있다. OWL은 RDF 스키마를 확장하여 RDF 스키마 클래스간의 복잡한 관계나 클래스 및 속성에 대한 정교한 제한 사항을 표현할 수 있다. 예를 들면, 클래스의 속성 수나 타입을 제한할 수 있고, 특정 속성을 포함하는 정보가 특정 클래스에 속함을 추론할 수 있으며, 특정 클래스의 모든 멤버가 특정 속성을 지니고 있음을 결정하거나, 클래스간의 상호관계 또한 표현할 수 있다¹⁷⁾.

RDF/OWL의 장점으로서는 정보통합 및 공유 어휘를 재사용할 수 있게 지원하며, 구조화되지 않은 데이터를 처리할 수 있다. 또한 데이터 모델링과 구문체계가 분리되어 있으며 확장이 용이하고, 세밀한 구문체계에 기초하여 추론과 분류가 가능하다. 그리고 그래프 구조 모델링을 통해 보다 융통성 있는 표현이 가능하다. 그러나 이러한 장점에도 불구하고, RDF/OWL의 단점으로는 문서검증 능력 및 성능이 떨어지며, 학습하는데 많은 시간이 소요된다. 또한 자원에 부여된 여러 특성을 연계시키는데 한계가 있다¹⁸⁾.

17) Kanzaki Masahide. 2008. 『시맨틱 웹을 위한 RDF/OWL 입문』. 황석형, 양해술 옮김. 서울: 홍릉과학출판사.

18) 서 휘. 2006. 전제서. p. 194.

2) TOPIC MAPS

토픽맵은 ISO가 제정한 표준으로, 정보 자원간의 상호 연관성을 도출, 연결 및 조직하여 효과적인 지식 구조를 기술할 수 있도록 하는데 목적이 있다. 토픽맵은 대용량의 구조화되지 않은 정보를 효율적으로 검색하기 위한 해결 방안으로 제안되었다¹⁹⁾. 토픽맵은 초기에 서적의 인덱스, 용어, 시소러스, 목차 등의 기술을 위한 도구로 개발되었지만, 온톨로지 개념이 활성화되면서 그 응용 영역이 크게 확대되었다. 토픽맵은 토픽 중심의 도메인 의미정보 모델링을 지원하는 온톨로지 언어로, 도메인의 지식정보를 온톨로지 형식으로 모델링하여 지식정보의 공유와 상호 교환을 실현해 주는 국제 표준 언어라고 할 수 있다²⁰⁾.

토픽맵은 토픽(Topic), 어커런스(Occurrence), 연관관계(Association)라는 핵심적인 세 가지 개념으로 구성되어 있다. 우선 토픽은 웹 페이지, 책, 개인, 사건 혹은 어떤 추상적 객체와 같은 사물일 수 있다. 토픽의 주된 목적은 무엇에 관한 개념을 구체화하기 위한 것이고 이는 어떤 추상적인 개념을 토픽맵이라는 지식구조로 표현하고 식별할 수 있는 객체로 만들기 위한 것이다. 어커런스는 각각의 토픽은 자신이 참조하는 하나 이상의 실제적인 지식 항목과 연결될 수 있다. 이러한 연결을 어커런스라고 한다. 어커런스는 여러 가지 형태를 가질 수 있는데 문서 파일, 이미지 파일, 비디오 파일, 데이터베이스 내의 특정 레코드 등으로 나타날 수 있다. 연관관계는 토픽맵에서 둘 혹은 그 이상의 토픽 사이의 관계를 선언할 수 있다. 토픽간의 관계가 연관관계로 표현됨과 동시에 그 관계가 어떤 것인지는 연관관계 타입을 통해 나타내며, 이 관계에서 각 토픽이 어떤 역할을 하는지 역할(Role)을 통해 기술된다²¹⁾.

토픽맵은 정보 자원의 구성, 추출, 네비게이션을 다양하게 제공하는 새로운 패러다임이며, 정보와 지식 관리를 위해 최적화된 표현 형식을 제공

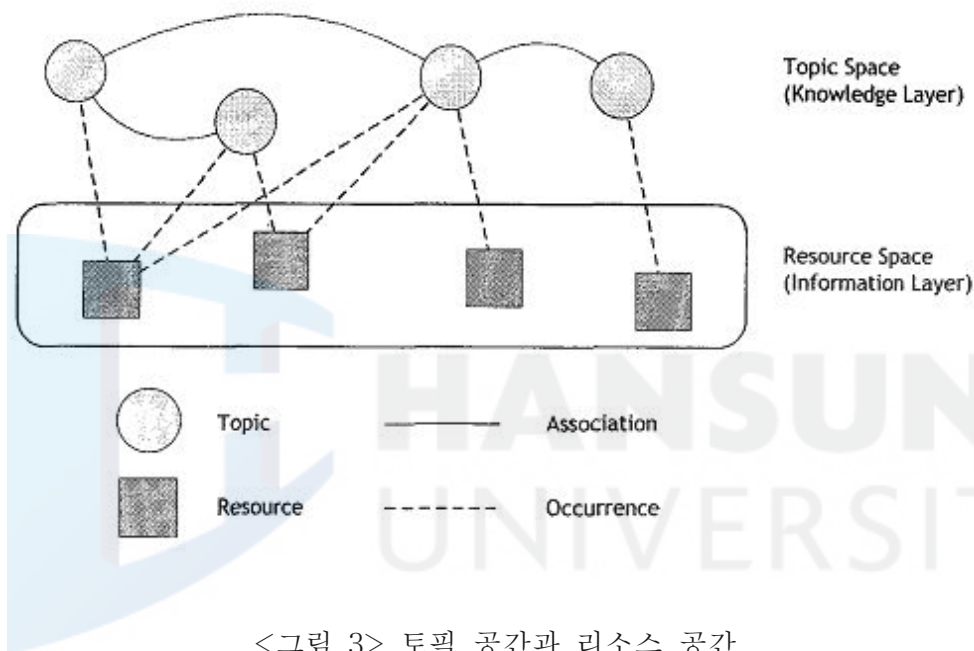
19) 한국전산원. 2004. 전계서.

20) 이현실, 한성국. 2005. Topic Maps를 이용한 MARC데이터의 FRBR모델 구현에 관한 연구. 『한국정보관리학회지』, 22(3): 289-306.

21) 박희영. 2006. 『더블린코어 메타데이터를 이용한 서지적 관계 구현에 관한 연구』. 석사학위논문, 연세대학교 대학원.

한다. 또한, 주제를 초점으로 관련된 자원을 융통성 있게 연계시킬 수 있는 효율적 방법을 제공한다²²⁾.

<그림 3>와 같이 토픽맵은 적합한 정보를 어커런스를 통하여 지시해준다. 어커런스는 데이터베이스에서 정보 오브젝트가 여러 위치들을 연결시켜주고 관계를 검색 할 수 있다. 토픽맵은 불분명한 개념들을 명확하게 대상과 연결시켜 주기 때문에 향후 지식 표현 스키마나 실세계의 대상들에 대한 명확한 설계가 가능하다²³⁾.



<그림 3> 토픽 공간과 리소스 공간

[출처: 문정수, 2006. p. 17.]

22) 문정수, 2006. 『토픽맵 기반의 인터넷 포털의 서비스 분류 및 연관체계 관리 방법에 관한 연구』. 석사 학위 논문, 한국과학기술원.

23) 상계서, p. 17.

2. 선행연구

2.1 온톨로지 기법을 이용한 연구

온톨로지 연구는 현재 다양한 분야에서 여러 방법론을 가지고 진행되고 있다. 주요 연구 분야로는 언어학 관점에서의 분류 및 의미 관계 연구, 인공지능 영역에서의 지식 표상 연구, 자연과학 영역에서의 용어와 개념들의 계층관계 연구, 전자 상거래 환경에서의 정보의 계층화 연구, 시맨틱 웹을 위한 공학적 온톨로지 구축 연구 등을 들 수 있다²⁴⁾. 또한, 최근에는 정보 검색 향상을 위해 온톨로지를 활용한 연구가 많이 진행되고 있다.

정현숙과 신영숙(2008)은 온톨로지 기반의 강의계획서 관리 시스템을 설계 하였다. 온라인상에 제공되는 다양한 유형의 강의계획서 양식을 분석하고 이를 통해 표준 모델로 활용 가능한 강의계획서 온톨로지의 클래스, 속성, 관계, 제약조건 및 학습 개념 온톨로지를 설계하였다²⁵⁾.

박선경(2008)은 기존에 제안된 대표적인 온톨로지 구축방법 중 구축과정이 비교적 상세한 Methonotology, OTK 구축방법을 바탕으로 한국문헌정보학회지에 게재된 문헌정보학 논문들의 표제를 대상으로 온톨로지를 작성해보고 그 과정을 비교분석하였다. 비교분석한 결과를 바탕으로 문헌정보학 논문을 위한 온톨로지 구축방법을 크게 다섯 단계로 제시하였다²⁶⁾.

멀티미디어 검색의 효율을 높이기 위하여 김수경과 안기홍(2007)은 온톨로지와 웹 온톨로지 관련 연구들을 분석하여 시맨틱 웹을 위한 웹 온톨로지의 특징들을 파악하고, 기존 온톨로지 구축과 웹 온톨로지 구축 기법을 제안한다. 제안된 방법에 따라 서술 논리와 SWRL의 표현 방법을 이용한 공리 규칙을 기반으로 온톨로지를 모델링하여 추론 기반의 웹 온톨로지를

24) 남지순, 2008. 정보검색엔진을 위한 <인물>관련 온톨로지 구축연구 사례들에 대한 소개 및 논의. 『프랑스어문교육집』, 27: 201-229.

25) 정현숙, 신영숙. 2008. 온톨로지 기반 강의계획서 관리 시스템 설계 및 구현. 『한국정보기술학회지』, 6(5): 108-114.

26) 박선경. 2008. 『문헌정보학 논문을위한 온톨로지 구축방법에 관한 연구』. 석사학위논문, 숙명여자대학교 대학원.

구축하였으며, 구축된 온톨로지의 검증을 위한 온톨로지 추론 실험을 통해 온톨로지 성능을 검증하였다. 또한 구축된 웹 온톨로지를 기반으로 하는 시맨틱 웹 응용의 예로 지능형 이미지 검색 시스템을 실험 시스템으로 구현하였다²⁷⁾.

박창섭(2007)은 의미적 연관성을 온톨로지 기반으로 구축하였다. 기존의 멀티미디어 검색 방법은 멀티미디어 콘텐츠에 포함된 의미 개념들이 연관성을 효과적으로 이용하지 못하고 콘텐츠에 대한 메타데이터 검색에 의존함에 따라 제한적인 검색 결과만을 제공하였기 때문에 도메인 온톨로지를 활용하여 멀티미디어 콘텐츠의 의미적 연관성에 기반을 둔 멀티미디어 검색 방법 및 시스템 구조를 제안하였다. 이용자가 검색어를 직접적으로 포함하지 않지만 의미적으로 연관성이 있는 개념들을 온톨로지에서 검색하고 이들에 대한 랭킹을 결정하기 위한 클래스 연관도 척도를 정의하고, 이를 효율적으로 계산하기 위한 알고리즘을 제시하였다. 제안한 검색 방법은 온톨로지에 기술된 개념들 간의 의미적 연관성을 바탕으로 검색어를 직접적으로 포함하지 않지만 연관성을 가지고 있는 다양한 멀티미디어 콘텐츠를 검색결과로 제공할 수 있으며, 검색된 콘텐츠를 온톨로지로부터 추출된 의미적 연관성에 따라 분류하여 보여줄 수 있다. 또한 온톨로지로부터 추출된 검색어 집합을 통해 검색 엔진에서 해당 메타데이터를 검색하기 때문에 온톨로지와 메타데이터의 독립성을 보장할 수 있었다²⁸⁾.

이정희와 김희섭(2007)은 국·공립 대학의 전자기록관리시스템이 보다 적합한 전자기록물을 검색할 수 있는 기반 환경을 조성하기 위하여 온톨로지 기반의 검색 시스템을 설계 및 구현하였고 그 성능을 기존의 키워드 기반 검색시스템과 비교해 보았다. 이 실험에 사용된 키워드 기반 검색시스템은 국립 한국해양대학교에서 설치·운영 중인 전자기록관리시스템을 사용하였고, 온톨로지 기반 검색시스템은 OntoStudio 1.4를 사용하여 자체 설계 및 구현을 하였다. 재현율과 정확율에 의한 성능을 비교해본 결과 온

27) 김수경, 안기홍. 2007. 지능형 이미지 검색시스템을 위한 추론기반 웹 온톨로지 구축. 『한국정보관리학회지』, 24(3): 119-147.

28) 박창섭. 2007. 의미적 연관성을 이용한 멀티미디어 정보검색. 『인터넷정보학회논문지』, 8(5): 67-79.

톨로지 기반의 검색시스템이 키워드 기반의 검색시스템보다 높은 성능을 보였다. 또한, 장문과 단문의 질의집단 비교평가에서도 온톨로지 기반의 검색 성능이 키워드 기반의 검색시스템 보다 뛰어난 것으로 나타났다²⁹⁾.

김학래와 김흥기(2007)는 개인용 전자문서를 검색하는 시스템을 구축하기 위하여 온톨로지 기술을 활용하였다. 전자 문서의 관리와 검색을 위해 개발한 ONTALK 시스템을 소개하였다. 문서를 메타데이터로 관리하면 실체문서의 전체 내용을 파악하거나 검색하지 않고도 내용을 이해할 수 있기 때문에 문서관리나 검색의 효율성을 높여 줄 수 있다. ONTALK는 온톨로지 기반으로 메타데이터를 기술함으로써 문서 검색 시 추론을 통한 정확성을 높일 수 있었다. 개발된 ONTALK의 특징으로는 온톨로지를 기반으로 메타데이터를 표현함으로써 기술된 데이터 간의 관계를 추론할 수 있는 방법을 제공하였고, 다양한 검색 형태를 제공함으로써 사용자가 원하는 수준의 검색 결과를 제공하였다. 또한, 사용자가 패턴 매칭을 통해 검색 결과를 얻을 수 있을 뿐만 아니라 추론 엔진을 통해 암묵적으로 존재하는 자원간의 관계를 검색할 수 있었다³⁰⁾.

차승준 연구팀(2007)은 전통약재정보의 효율적 활용을 위한 온톨로지 시스템을 개발하였다. 한의학에서 사용되는 수많은 약재정보들에 대해서 효율적인 검색을 위해 시맨틱 웹을 사용하였다. 한의학 연구원에서 현재 쓰이는 약재 데이터는 총 16개의 속성을 가지고 있다. 각각의 클래스, 속성, 개체를 정의하고, 기존에 구축된 온톨로지와 그것들의 관계성을 바탕으로 약재 데이터 및 속성들의 값을 추가, 수정, 삭제가 가능하도록 온톨로지 관리시스템을 구현하였다³¹⁾.

Srinivasan과 Huang(2005)은 온톨로지 기법을 이용하여 디지털 박물관을 설계하였다. 문화유산에 대한 관심이 높아지면서 보존의 문제뿐만 아니라, 관리적 측면에서도 더욱 높은 기술이 요구되고 있는데 온톨로지 기법

29) 이정희, 김희섭. 2007. 대학 전자기록물을 위한 온톨로지 기반 검색시스템 설계 및 구현. 『정보관리학회지』, 24(3): 343-362.

30) 김학래, 김흥기. 2007. 시맨틱 웹/온톨로지 기술을 이용한 개인용 전자문서검색시스템. 『한국전자거래학회지』, 12(1): 135-149.

31) 차승준 외. 2007. 전통 약재 정보의 효율적인 활용을 위한 온톨로지 시스템 개발. 『한국정보과학회 학술발표 논문집』, 34(2): 195-199.

을 기반으로 큐레이터와 일반 관람자 사이에 상호작용 및 추론과 관련해 어떠한 관계내역을 추적할 수 있을 것으로 내다보았다. 디지털 박물관 설계에 있어 메타 데이터를 유동적 온톨로지 사용의 배후 핵심 아이디어를 표현하는데 사용할 수 있다고 보았다³²⁾.

Weinberger(2008) 연구팀은 온톨로지에 기반한 조직적인 메모리의 평가를 연구하였다. 조직 메모리(OM) 조직 혁신과 경쟁 우위를 위해 노력하고 구현하기 위해 온톨로지를 이용하였고, 도메인 즉, 스키마를 통해 온톨로지화를 하였다. 최상위 레벨에는 구조용 메모리를 설정하였고, 학습관리 지식과 조직의 활성화를 유도하였다³³⁾.

Fonseca(2007)³⁴⁾는 정보과학 연구에 있어서 온톨로지의 이중 역할에 대해 연구하였다. 온톨로지는 정보과학에서, 개념적 모델과 개념적 스키마 검증 모두에 대해 제안되는 모델이라 밝히고 있다.

Yildiz와 Miksch(2007)는 온톨로지 탐색 정보 시스템에 대해 연구하였다. 온톨로지는 여러 응용분야에서의 활용이 증가되고 있으며, 정보 시스템 내에서 보다 유연한 통합을 위한 요구사항들을 가능하게 한다고 하였다. 따라서 온톨로지는 정보시스템의 성능을 향상시키는데 필수적으로 요구되는 기술이라고 설명하였다³⁵⁾.

2.2 인물 및 연구자정보 검색에 관한 연구

또한 주 연구 목적이 온톨로지를 이용한 연구자정보 검색서비스는 아니지만, 인물관련 온톨로지를 논의한 연구는 다음과 같다.

32) R. Srinivasan and J. Huang. 2005. "Fluid ontologies for digital museums." *International Journal on Digital Libraries*, 5(3): 193-204.

33) H. Weinberger, D. Te'eni, and A. J. Frank. 2008. "Ontology-based evaluation of organizational memory." *Journal of the American Society for Information Science and Technology*, 59(9): 1454-1468.

34) F. Fonseca. 2007. "The double role of ontologies in information science research." *Journal of the American Society for Information Science and Technology*, 58(6): 786-793.

35) B. Yildiz and S. Miksch. 2007. "ontoX - A method for ontology-driven information extraction." *ICCSA*, 3: 660-673.

김현희와 안태경(2003)은 인터넷 웹상에서 국제기구에 관한 정보를 체계적으로 관리하고 검색하기 위해서 국제기구 온톨로지를 설계하고 이 온톨로지에 기반하여 검색 시스템을 구현하였다. 또한, 구현된 검색 시스템을 20개의 탐색 질문들을 이용하여 기존의 인터넷 검색엔진과 적합성과 탐색 시간이라는 두가지 요인을 통해서 비교해 보았다. 실험 결과에 의하면 적합성면에서 온톨로지 기반 시스템의 적합도가 더 높은 결과를 나타내었다. 또한, 탐색 시간 역시 온톨로지 기반 시스템이 인터넷 검색엔진에 비해 보다 더 적은 탐색시간이 걸린 것으로 나타났다³⁶⁾.

권범중과 곽승진(2006)은 과학기술분야 인력정보관리시스템들의 데이터를 기반으로 실제적으로 적용이 가능한 전문가 선정방법을 제안하기 위해 먼저 국내외 주요 과학기술분야 인력정보관리시스템의 운영현황을 개괄한 후, 각 시스템의 데이터 구조를 분석하였다. 향후, 분석된 데이터구조를 통한 가중치 적용요소 추출, 객관적 근거를 가진 평가기준의 제시, 알고리즘 적용을 통한 결과도출 및 평가 등에 대한 추가적인 연구가 필요하다고 하였다³⁷⁾.

한용진 연구팀(2007)은 인물데이터베이스 검색의 문제점을 해결하고자 온톨로지 기술을 응용하였다. 즉 잘 정리된 형태로 인물정보를 집약한 인물데이터베이스 기반의 검색은 하나의 데이터베이스 내에서는 검색이 효율적으로 이루어질 수 있으나 여러 내용이 흩어져서 저장된 경우 필요한 내용을 찾고 재구성하는 많은 시간과 노력이 필요하다고 보았다. 이런 문제점을 해결하기 위하여 IT-People Event Ontology(ITPEO)스키마를 구축하여 서로 다른 소스로부터 추출한 인물정보를 하나의 온톨로지로 인스턴스화하였다. 연구 결과로, 데이터베이스를 기반으로 한 검색은 신뢰할 수 있는 정보를 제공해 주지만, 변하는 사건정보나 부분적인 개념관계에 대해 유연하게 대처하지 못한다. 따라서 온톨로지를 이용할 경우, 질의어를 의미적으로 처리함으로써 질의 요구를 반영한 검색을 할 수 있었다. 또

36) 김현희, 안태경. 2003. 온톨로지를 이용한 인터넷 웹 검색에 관한 실험적 연구. 『한국정보관리학회지』, 20(1): 417-455.

37) 권범중, 곽승진. 2006. 과학기술분야 인력정보관리시스템 데이터기반 전문가 선정 방법 연구. 『한국정보관리학회 학술대회 논문집』, 13: 223-228.

한 웹 혹은 데이터베이스의 다양한 정보를 자동으로 온톨로지화함으로써 인물과 관련된 변화하는 사건 정보도 개념 단위로 처리할 수 있다고 보았다³⁸⁾.

박윤희(2008)는 ITPEO를 효율적으로 구축하기 위하여 자연어 처리 기반의 인물 정보 추출을 가능하게 하는 시스템을 설계하였다. 즉 자동 온톨로지 인스턴스 생성하기 위한 자연어 처리기반 인물 정보 추출 시스템을 제안하였다. 제안된 시스템은 구 단위의 정보 추출, 인물 대용어 처리 및 절 단위의 정보 추출의 세 가지 모듈로 구성하였고 실험은 텍스트로부터 추출 가능한 여러 가지 의미적 관계들 중에서도 IT-피플 이벤트 온톨로지를 위한 인물 정보만을 그 추출 대상으로 실험하였다. 추출된 인물의 직위 정보는 현재의 인물의 직위 정보뿐만이 아니라 변동된 직위까지 정보를 담고 있다. 따라서 인물의 과거 변동된 정보의 히스토리까지 추출이 가능하다는 장점이 있었다³⁹⁾.

최근 남지순(2008)은 정보검색 엔진을 위한 <인물>관련 온톨로지 구축 연구 사례들에 대하여 소개하였다. 그는 외국의 인물 온톨로지 연구들은 다음과 같은 세 가지 유형으로 분류될 수 있다고 하였다. 첫째는 개인의 인적 사항 및 신상에 관련된 정보를 담고 있는 ‘개인정보 온톨로지’ 유형이며, 둘째는 개인의 사회적 관계망을 보이는 ‘사회관계’ 온톨로지 유형, 그리고 셋째는 ‘유생물 개체로서의 인간’에 대한 온톨로지 유형이다. 이와 같은 분류에 따른 대표적 프로젝트는 다음과 같다⁴⁰⁾.

첫째, 미국의 메릴랜드 대학교의 컴퓨터 공학과에서 제안한 SHOE(Simple HTML Ontology Extensions)을 들 수 있다. 이 시스템은 포괄적인 개인 정보 온톨로지를 구성하였다. 제안된 인물정보 클래스로는 활동영역, 연락영역, 고용영역, 진행영역, 사회그룹영역, 업무영역으로 나누고 그 밑에 상세한 속성값을 제안하고 있다⁴¹⁾. 둘째, 유럽연합의

38) 한용진 외. 2008. 이벤트 온톨로지 기반의 의미 정보 검색. 『정보과학회논문지』, 14(1): 96-100.

39) 박윤희. 2008. 『온톨로지 인스턴스 생성을 위한 인물정보 추출』. 석사학위논문, 경북대학교 대학원.

40) 남지순, 2008. 전게서.

41) J. Heflin, 2000, <Personal Ontology> in SHOE.
<<http://www.cs.umd.edu/projects/plus/SHOE/onts/personal1.0.html>>. [cited. 2008. 7. 3].

JRC(Joint Research Centre)는 테러리즘과 관련된 인물들의 사회적 관계에 대한 정보를 지식베이스화 하기 위해 구축한 온톨로지인 SNO(SocialNetwork Ontology)를 구축하였다. 이를 위하여 People, Event, Organization, Location의 하위 온톨로지를 구축하고 각각에 적합한 클래스를 제안하였다⁴²⁾. 셋째, CIA World Fact Book에서는 SUMO(Suggested Upper Merged Ontology)에서 제안하는 각 나라의 민족과 국민에 대한 개념 온톨로지인 People 온톨로지를 사용하였다. 이 온톨로지에서 사용한 상위 클래스는 인구, 연령 구조, 인구 증가율, 출생률, 사망률, 순이동율, 성 비율, 유아사망률, 수명, 총 출산을 등이 있다⁴³⁾.

또한, 국내 온톨로지 구축 사례에 대해 살펴보면 한국과학기술정보연구원에서 최근 발표한 OntoFrame⁴⁴⁾이 있다. OntoFrame은 시맨틱 웹 기술을 바탕으로 연구자가 원하는 정보를 쉽고 정확하게 제공하는 학술정보 분석 및 융합 서비스를 제공하기 위한 목적을 가지고 있다. OntoFrame은 세 가지 온톨로지를 기반으로 지식을 구축하고 있는데 하나는 "국가 과학 기술 R&D 기반정보 온톨로지"이고 나머지들은 시맨틱 포털 서비스를 위한 "응용 온톨로지"와 "개인화 온톨로지"이다. 기반정보 온톨로지는 연구자와 연구기관, 논문이나 저서, 특허, 보고서 등의 성과물, 성과물의 주제, 기관이 위치한 지역 등의 학술 연구정보를 지식으로 모델링 하고 있다. 연구 분야의 다른 온톨로지와 달리, 이 온톨로지는 연구자와 연구자의 현재 소속기관 사이의 관계뿐만 아니라 특정 성과물을 저작할 당시의 소속기관과의 관계도 모델링 하고 있다. OntoFrame은 여러 서비스를 제공하고 있다. 학술연구토픽의 경향 분석, 토픽의 주요 연구자 파악, 토픽의 주요 연구기관 파악, 연구자 간의 연구자 네트워크, 토픽의 지역별 연구 분포 분석, 학술연구정보의 통계 분석 등을 제공한다⁴⁵⁾.

42) P. O. Wennerberg. 2005. "Ontology Based Knowledge Discovery in Social Networks." Final Report JRC. <<http://langtech.jrc.it>>. [cited. 2008. 7. 3].

43) The Suggested Upper Merged Ontology (SUMO). <<http://www.ontologyportal.org/>>. [cited. 2008. 6. 24].

44) 김평 외. 2008. OntoFrame 기반 학술정보 분석 서비스. 『정보과학회논문지: 소프트웨어 및 응용』, 35(7): 431-441.

45) 이미경, 정한민, 성원경. 2008. 웹 서비스 기반의 추론 서비스에 관한 연구. 『한국정보과학회 학술발표 논문집』, 35(2): 200-203.

Ⅲ. 연구자정보 온톨로지 분석

본 장에서는 연구자정보를 다양한 각도로 기술하여 이용자가 연구자에 대해 사실정보뿐만 아니라 사실정보를 기반으로 한 의미 있는 연관성 정보를 제공하는 검색 인터페이스를 설계하기 위하여 연구자정보의 특징을 표현하는데 필요한 기술요소 즉 속성을 선정하고 속성간의 관계를 정의하였다.

1. 연구자정보 기술요소 분석

현재 제공되고 있는 연구자정보 시스템이 어떤 정보를 제공하고 있는지를 알아보기 위하여 다양한 인물 또는 연구자정보 제공 시스템을 선정하여 각각의 서비스의 특징과 제공내용을 일차적으로 조사하였고 연구자정보 제공 시스템들이 기술하고 있는 각각 기술요소들을 기반으로 하여 본 연구에서 사용하는 연구자정보 속성의 클래스를 산출하였다.

1.1 인물정보 DB 조사

인물정보는 연구자정보, 인력정보, 인물정보 등 다양한 명칭으로 기관별 독자적으로 관리·운영되고 있다. 인물정보는 크게 연구지원기관 및 대학에서 구축한 연구자중심의 인력 DB와 언론기관의 정치·사회 유명인 중심의 인력DB로 구분될 수 있다. 전자의 경우 학력, 경력, 연구실적, 학회활동 등 개인의 연구역량을 나타내는 정보로 이루어지고, 후자는 가족관계, 취미, 등의 신변적 정보를 제공하고 있다⁴⁶⁾.

본 연구에서는 현재 웹 기반 인물정보원 중 수록인물의 범위가 포괄적

46) 지정규. 2003. 국가학술연구인력 데이터베이스 구축에 관한 연구. 『한국컴퓨터정보학회지』, 10(1): 16-23.

인 조선닷컴 인물정보, 동아닷컴 인물정보, 연합뉴스 인물정보, 조인스 인물정보⁴⁷⁾ 중 자체적으로 데이터베이스를 구축하고, 그 중 30만건의 넘는 데이터량을 보유하고 있는 조인스⁴⁸⁾, 조선일보⁴⁹⁾와 주제별 전문인사 인명정보를 제공하는 로마켓⁵⁰⁾을 조사대상으로 선정하였다. 그리고 국내 연구인적정보시스템인 한국학술진흥재단의 한국연구업적통합정보시스템⁵¹⁾과의 한국과학기술정보연구원의 NTIS⁵²⁾, 미국의 연구인적정보서비스인 SESTAT⁵³⁾와, 일본의 연구인력 정보서비스인 ReaD⁵⁴⁾, ProQuest사의 전문연구자 정보 데이터베이스인 COS Scholar Universe⁵⁵⁾도 추가적으로 선정하였다. 이에 따라 8개의 시스템에 대한 연혁 및 소개, 서비스의 특징, 검색 인터페이스의 특징 그리고 기술요소 항목 등을 조사하였다.

1) 조인스닷컴

조인스닷컴은 1985년부터 인물정보 데이터베이스 구축을 시작하여 30만여명의 대한민국 각 분야 주요 인사들의 프로필 정보를 보유하고 있다. 또한 뉴스인물, 주요기업의 인사이동, 직업별, 출신학교별, 출신지역별, 국회의원 등의 카테고리를 두어 원하는 인물정보에 보다 쉽게 접근할 수 있도록 하였다. 조인스닷컴의 인물정보 검색 서비스의 주요 특징은 인맥정보를 상세히 알려준다는 것이며, 특히 졸업동기를 링크해주거나 관련된 인맥을 시각화하여 각 그룹별 연계도를 보여주고 있다. 따라서 관련 인물들을 연속적으로 탐색할 수 있는 장점을 가진다.

조인스닷컴의 ‘인물정보’는 기본정보로는 이름, 한문성명, 영문성명, 생년

47) 박준식, 박경호. 2001. 한국 웹 기반 인물정보원의 분석적 연구. 『한국문헌정보학회지』, 35(1): 187-203.

48) <<http://people.joins.com/>>.

49) <<http://www.chosun.com/national/people/>>.

50) <http://www.lawmarket.co.kr/people/home/people_home.asp>.

51) <<http://www.kri.go.kr/index.jsp>>.

52) <<http://hurims.ntis.go.kr/index.jsp>>.

53) <<http://sestat.nsf.gov/SESTAT/NewHomeLogin.html>>.

54) <<http://read.jst.go.jp/>>.

55) <http://csaweb115v.csa.com/ids70/quick_search.php?SID=ho1bkqgdmi4ct17s0ta8momd74>.

월일, 나이, 소속, 성별, 직업, 전문분야, 직장주소, 직장전화/팩스, 이메일, 학력, 경력-기관경력/기타, 활동-상훈, 활동-저서, 가입단체, 지인, 자녀수, 가족사항 등의 정보를 제공하고 있다(표 1 참조).

<표 1> 조인스닷컴 기술요소 항목

시스템명	분류	기술요소
조인스닷컴	인적정보	조인스 ID, 이름, 한자명, 영문명, 별칭, 생년월일, 주민등록번호, 성별, 본관, 전문분야, 직장주소, 직장전화, 팩스, 자택주소, 자택전화, 휴대폰, 이메일, 홈페이지, 본적, 원적, 출생지, 결혼기념일, 자녀수, 흡연량, 주량, 기호품, 신장, 체중, 취미, 종교, 시력, 특기, 병역사항, 혈액형, 가훈, 생활신조, 자격/면허, 작품활동, 학력사항, 경력사항, 상훈
	연구정보	저서, 논문
	소속정보	소속기관, 부서, 직위, 담당업무, 가입단체
	사회정보	지인
	기타	자녀수, 가족사항

2) 조선닷컴

조선닷컴의 인물검색 서비스 역시 30만 여명의 각계각층의 주요인사 프로필을 제공하는 인물정보서비스이다. 이름으로 검색했을 경우, 동명의 인물리스트가 제공이 되며 사진, 소속, 생년월일 등을 간략하게 제공하여 원하는 인물을 리스트 상에서 선택할 수 있게 하였다. 제공되고 있는 기술요소는 성명, 한자성명, 영문성명, 생년월일, 성별, 출생지, 국적, 직업, 전문분야, 직장주소, 직장전화/팩스, 자택주소, 자택전화/팩스, 휴대폰, 이메일, 홈페이지, 학력, 경력, 자격면허, 상훈, 저서, 논문, 소속기관, 소속부서, 직위, 소속단체, 교우/지인, 가족사항, 자녀수 등으로 상세한 정보를 제공해주는 특징을 가진다(표 2 참조).

<표 2> 조선닷컴 기술요소 항목

시스템명	분류	기술요소
조선닷컴	인적정보	한글명, 한자명, 영문명, 주민등록번호, 생년월일, 아호, 띠, 성별, 본명, 종교, 별칭, 출생지, 국적, 본관, 혈액형, 본적, 원적, 전문분야, 직장주소, 직장전화/팩스, 주요직업, 자택주소, 자택전화/팩스, 휴대폰, 이메일, 홈페이지, 병역, 복무기간, 최종계급, 신장/체중, 시력, 주량, 흡연량, 취미, 특기, 결혼일, 결혼형태, 자격/면허, 존경인물, 좌우명, 가훈, 경영관, 감명깊은 책, 선호음식, 감명깊은 영화, 애창곡, 즐겨하는 운동, 좋아하는 연예인, 애장품, 좋아하는 술, 좋아하는 차, 좋아하는 색깔, 좋아하는 꽃, 건강관리, 외국어 구사, 스트레스 해소법, 학력사항, 경력사항, 상훈, 작품/기록,
	연구업적정보	저서, 논문
	소속정보	소속기관, 부서/직위, 담당업무, 소속단체, 가입정당
	사회정보	교우/지인
	기타	가족사항, 자녀수

3) 로마켓

법률 포털사이트인 로마켓은 변호사 등의 법률관련 인물이나 사회 주요 이슈 인물의 출신학교, 주요경력 등에 관한 정보를 제공하고 있다.

로마켓은 성명, 한문성명, 영문성명, 생년월일, 출생지, 성별, 전문분야, 직업, 직장주소, 직장전화/팩스, 이메일, 홈페이지, 학력, 유학국가, 경력, 시험/자격증, 수상, 논문, 저서, 프로젝트, 지적재산권, 직장, 소속학·협회, 친한인물/동료, 가족사항, 사법연수원기록 등의 정보를 제공하고 있다(표 3 참조). 또한, 가까운 인물 찾기 서비스를 통해 출생지, 고등학교, 대학교, 대학원, 해외대학교 등을 통해 인맥관계 파악이 가능하다.

<표 3> 로마켓 기술요소 항목

시스템명	분류	기술요소
로마켓	인적정보	이름, 한자명, 영문명, 생년월일, 생년, 출생지, 성별, 호(별명), 본관, 본적, 취미, 특기, 좌우명, 가훈, 가족, 병역, 종교, 존경하는 인물, 혈액형, 신장/체중, 감명깊은 책, 감명깊은 영화, 좋아하는 음식, 좋아하는 노래, 주량, 흡연량, 전문/집중분야, 나의 동영상, 나의 소개, 직업, 직장, 연락처, 초등학교, 중학교, 고등학교, 대학교/대학원, 최종학위, 전공분야, 유학, 경력, 시험/자격증, 수상, 예술발표, 영화, 드라마, 앨범, 뮤직비디오, 기타활동
	연구업적정보	논문, 저서, 프로젝트, 지적재산권
	소속정보	직장, 소속학·협회
	사회정보	친한인물/동료
	기타	가족사항, 사법연수원기록

4) 한국학술진흥재단

한국학술진흥재단의 KRI(Korean Researcher Information, 한국연구업적통합정보)시스템은 대학 및 기관 연구자들의 연구업적정보를 국가차원에서 공유 및 활용하기 위해 구축한 한국연구업적통합정보를 제공하는 정보시스템이다. KRI의 목적은 한국학술진흥재단이 연구자의 연구업적정보를 체계적이며 종합적으로 관리하고 검증된 연구업적을 기반으로 학술연구지원 사업 과제 선정에 공정하고 투명하게 하는데 있다. 또한, KRI가 구축됨에 따라 연구자들은 자신들의 업적 및 이력관리에 소요되는 시간과 번거로움을 축소시킬 수 있으며 연구업적 정보를 검색하여 동일한 연구를 피할 수 있는 장점을 가진다.

KRI의 연구자정보 기술요소는 성명, 한문성명, 영문성명, 생년월일, 성별, 연구자등록번호, 국적, 연구실 전화/팩스, 핸드폰, 자택전화, 이메일, 홈페이지, 직장주소, 소속 기관, 분교/캠퍼스, 대학원, 단과대학, 학과(부), 세부전공, 임용일, 퇴직일, 타기관 정보제공, 정보공개여부, 메일수신여부 등의 기본정보와 연구분야, 경력사항, 취득학위, 논문실적, 연구비, 저역서,

지식재산, 전시작품, 학술활동, 수상사항, 자격사항, 기술이전 등의 정보를 제공하고 있다(표 4 참조).

<표 4> KRI 기술요소 항목

시스템명	분류	기술요소
한국학술 진흥재단 (KRI)	인적정보	성명, 한문성명, 영문성명, 생년월일, 성별, 연구자 등록번호, 국적, 연구실 전화/팩스, 핸드폰, 자택전화, 이메일, 홈페이지, 직장주소, 연구분야, 경력사항, 취득학위, 수상사항, 자격사항
	연구정보	논문실적, 연구비, 저역서, 지식재산, 전시작품, 학술활동, 기술이전
	소속정보	소속 기관, 세부전공, 임용일, 퇴직일
	기타	타기관 정보제공, 정보공개여부, 메일수신여부

5) 국가R&D참여인력정보서비스

한국과학기술정보연구원의 국가R&D참여인력정보서비스는 NTIS 서비스와의 연계되어 국가연구개발사업에 참여하는 연구자에 대한 상세정보를 통합·제공하는 서비스이다. 연구자가 국가연구개발사업에 참여하기 위해 여러 연구관리전문기관에 이중등록해야 하는 불편함을 해소하도록 하였고, 각 연구자의 연구내용과 국가연구개발사업 등을 검색할 수 있다. 주요 특징으로는, 해당분야의 연구내용이나 전문가 파악 및 특정조건에 적합한 인력정보 검색 서비스를 제공하고 있으며, 과학기술표준분류별 전문분야, 전공별 등의 디렉토리 탐색 서비스가 가능하다. 또한, 이공계인력 육성·활용 관련 정책입안자에게 정책수립 지원을 위한 통계정보 서비스가 제공되고 있다. 주요 특징으로는 평가위원 후보 추천서비스를 제공, 평가위원 선정의 객관성, 전문성, 공정성을 높일 수 있다는 점이다.

국가R&D참여인력정보서비스는 기본정보로는 성명, 한문성명, 영문성명, 생년월일, 국적, 홈페이지, 전문분야 등이 있고, 그 외 상세정보는 소속기관사항, 최종학력사항, 과제정보, 기사검색, 웹정보보기, 연구보고서, 관련

논문 등이 제공되고 있다(표 5 참조).

<표 5> NTIS 기술요소 항목

시스템명	분류	기술요소
국가R&D 참여인력정보 서비스 (NTIS)	인적정보	성명, 한문성명, 영문성명, 생년월일, 국적, 홈페이지, 전문분야, 최종학력사항
	연구정보	과제정보, 관련논문, 연구보고서
	소속정보	소속기관사항
	기타	웹정보보기, 기사검색

6) 미국의 SESTAT

미국 국립과학재단(National Science Foundation)은 1950년대 설립된 연방정부기관으로, 과학기술자의 직업 범주를 크게 5개로 분류하여 컴퓨터과학 및 수학자, 생명과학자, 물리학자, 사회과학자, 공학자 등으로 구분하고 있다. NSF는 국가연구위원회(National Research Council)의 권고를 통해 미국 내 과학기술자를 파악할 수 있는 기본적인 데이터베이스로 SESTAT(Scientists and Engineers Statistical Data System)을 운영, 관리하고 있다⁵⁶⁾. 과학기술관련 분야 대학졸업자·최근의 대학졸업자·박사 학위 취득자 등 3개 영역의 자료를 통합해 만든 데이터베이스인 SESTAT는 미국의 과학·기술인력 수급에 대한 정책분석 및 연구를 위한 자료 제공하고 있는 포괄적이고 통합적인 미국의 과학기술인력 통계시스템이다.

SESTAT를 통해 과학·기술자 고용시장에 관한 정보를 알 수 있다. 과학·기술 취업자에 대해 주직업과 급여, 은퇴여부, 고용기관의 유형 및 고용형태, 직무상 관리 대상의 인원 등 사항을 파악할 수 있는 것이다. 아울러 일과 최종 학력과의 관계, 전형적인 일의 형태, 직무상 요청되거나 소

56) 권범중, 박승진. 2006. 과학기술분야 인력정보관리시스템 데이터기반 전문가 선정 방법 연구. 『한국정보관리학회 학술대회 논문집』, 13: 224.

지하고 있는 면허 및 자격증, 미국 정부의 연구지원, 부업 등에 관한 내용도 알 수 있다(표 6 참조).

<표 6> SESTAT 기술요소 항목

시스템명	분류	기술요소
SESTAT	인적정보	시민권, 나이, 인종/민족, 성별, 장애상태, 출생국, 취업자와 미취업자로 구분된 직무에 관한 사항에 대한 정보, 학사 및 가장 최근의 학위, 고등학교 졸업시기, 박사후 과정 course, 연구분야, 고용자의 지원
	연구정보	특허, 저서
	소속정보	전문가 협회 및 학회 가입 여부
	기타	결혼여부, 배우자의 고용상태, 자녀여부, 자녀의 나이, 부모의 교육수준

7) 일본 JST의 ReaD

일본과학기술진흥사업단(Japan Science and Technology Corporation)은 과학기술정보의 수집·정리 및 제공을 주된 업무로 하는 일본과학기술정보센터(Japan Information Center of Science and Technology)와 신기술의 위탁개발이나 창조과학기술사업, 기초연구사업 및 연구교류사업을 주된 업무로 하는 신기술사업단(Research Development Corporation of Japan)이 통합한 법인이다. JST는 일본내 대학·공공 연구기관등에 관한 기관정보, 연구자정보, 연구 과제정보, 연구 자원 정보를 망라적으로 수집·제공하는 연구개발지원 종합 디렉토리(Directory Database of Research and Development Activities)를 운영하고 있다. ReaD는 연구과제명, 연구개요, 연구성과 등의 내용이 포함된 약 200,000명의 연구자 정보데이터를 수집·관리하고 있다⁵⁷⁾.

ReaD는 연구자정보를 크게 기본정보와 상세정보로 구분하고 있는데, 주로 해당인력의 연구분야, 연구실적 등 연구전반에 걸친 부분이 주가 되고 기본적으로 성명, 소속기관명, 부서명, 직명, 연락처, 연구분야, 연구분야

57) 권범중, 박승진. 2006. 전계서. pp. 225.

키워드, 현재연구과제, 홈페이지 등의 정보를 제공하고, 상세정보로는 연구 실적, 출신 학교, 전공, 취득 학위, R&D 연구분야, 연구경력, 소속학과 등이 있다(표 7 참조). 또한, 연구자가 재직중인 기관의 정보를 볼 수 있고, 또한 그 기관에 속한 다른 연구자의 리스트도 볼 수 있다.

<표 7> ReaD 기술요소

시스템명	분류	기술요소
ReaD	인적정보	개정일, 성명, 소속기관명, 부서명, 직명, 연락처, 홈페이지, 출신학교, 전공, 취득학위
	연구정보	연구분야, 연구분야 키워드, 현재연구과제, 연구실적(저서, 발표논문), 연구실적(현지조사 등), R&D 연구분야, 연구경력
	소속정보	소속학과

8) COS Scholar Universe

ProQuest에 합병된 CSA의 COS Scholar Universe는 전문 연구자 정보 데이터베이스로, 전 주제 분야 연구자정보와 연구 공동체정보를 제공하고 있다. 200개가 넘는 학술 주제 분야에 대해 미국, 영국, 캐나다, 호주 등 영어권 8개 국가에서 활동하고 있는 100만 명 이상의 연구자들에 대한 정보를 제공하고 있다. 기술요소로서 이름, 이메일, 웹페이지주소, 이력서 링크, 소속기관, 최종학위, 관심 연구분야, 수상내역, 저작물, 공저자 등을 기본정보로 제공하고 있으며, 저자가 소속된 기관의 멤버 리스트도 확인할 수 있다(표 8 참조).

<표 8> COS Scholar Universe 기술요소 항목

시스템명	분류	기술요소
COS Scholar Universe	인적정보	이름, 이메일, 웹페이지주소, 이력서 링크, 최종학위, 관심분야
	연구정보	저작물, 공저자
	소속정보	소속기관

1.2 인물정보 DB 기술요소 비교

본 연구에서는 인물정보 또는 연구자정보를 제공해주는 8개의 국내외 시스템이 제공하는 기술요소를 비교하였다(표9 참조). 단, 웹 기반 인물정보 DB는 인물에 대한 사적인 정보도 기술요소로 채택하고 있으나 본 연구에서는 ‘연구자’에 대한 정보를 기술하는 것이므로 몇몇 기술요소는 비교대상에서 제외시켰다⁵⁸⁾.

<표 9> 인물정보 검색시스템의 기술요소

	시스템명	기술요소
웹기반 인물정보 검색서비스	조인스닷컴	성명, 한자성명, 영문성명, 별칭, 생년월일, 성별, 소속기관, 부서, 직위, 담당업무, 직업, 전문분야, 직장주소, 직장전화, 팩스, 자택주소, 자택전화, 휴대폰, 이메일, 홈페이지, 자격면허, 가입단체, 학력, 경력, 상훈, 저서, 논문, 가족사항, 지인
	조선닷컴	성명, 한자성명, 영문이름, 생년월일, 성별, 출생지, 국적, 직업, 전문분야, 직장주소, 직장전화/팩스, 자택주소, 자택전화/팩스, 휴대폰, 이메일, 홈페이지, 학력, 경력, 자격면허, 상훈, 저서, 논문, 소속기관, 소속부서, 직위, 소속단체, 교우/지인, 가족사항, 자녀수
	로마켓	성명, 한문성명, 영문성명, 생년월일, 출생지, 성별, 전문분야, 직업, 직장, 직장주소, 직장전화/팩스, 이메일, 홈페이지, 학력, 유학국가, 경력, 시험/자격증, 사법연수원, 수상, 논문, 저서, 소속학·협회, 프로젝트, 지적재산권, 가족사항, 친한인물/동료
국내 연구자 정보서비스 기관	KRI	성명, 한문성명, 영문성명, 생년월일, 성별, 연구자 등록번호, 국적, 연구실 전화/팩스, 핸드폰, 자택전화, 이메일, 홈페이지주소, 주소, 연구분야, 경력사항, 취득학위, 수상사항, 자격사항, 논문실적, 연구비, 저역서, 지식재산, 전시작품, 학술활동, 기술이전, 소속 기관, 세무전공, 임용일, 퇴직일, 타기관 정보제공, 정보공개여부, 메일수신여부
		성명, 한문성명, 영문성명, 생년월일, 국적, 홈페이지,

58) 제외한 기술요소는 다음과 같다: 별칭, 본관, 본적, 원적, 결혼기념일, 흡연량, 주량, 기호품, 신장, 체중, 취미, 종교, 시력, 특기, 병역사항, 혈액형, 가훈, 생활신조, 예술작품활동(영화/드라마/뮤직비디오), 주민등록번호, 가입정당, 존경인물, 경영관, 감명깊은 책, 감명깊은 영화, 선호음식, 즐겨하는 운동, 좋아하는 연예인, 애장품, 좋아하는 술, 좋아하는 차, 좋아하는 색깔, 좋아하는 꽃, 건강관리, 외국어 구사, 스트레스 해소, 좋아하는 노래, 나의 동영상, 나의 소개, 기타활동

	NTIS	전문분야, 최종학력사항, 과제정보, 관련논문, 연구보고서, 소속기관사항, 웹정보보기, 기사검색
국외 연구자 정보서비스 기관	SESTAT	시민권, 나이, 인종/민족, 성별, 장애상태, 출생국, 취업자와 미취업자로 구분된 직무에 관한 사항에 대한 정보, 학사 및 가장 최근의 학위, 고등학교 졸업시기, 박사후 college course, 연구분야, 고용자의 지원분야, 특허, 저서, 전문가 협회 및 학회 가입 여부, 결혼여부, 배우자의 고용상태, 자녀여부, 자녀의 나이, 부모의 교육수준
	ReaD	개정일, 성명, 소속기관명, 부서명, 직명, 연락처, 홈페이지, 출신학교, 전공, 취득학위, 연구분야, 연구분야 키워드, 현재연구과제, 연구실적(저서, 발표논문), 연구실적(현지조사 등), R&D 연구분야, 연구경력, 소속학과
	COS Scholar Universe	이름, 이메일, 웹페이지주소, 이력서 링크, 소속기관, 최종학위, 관심분야, 저작물, 수상내역, 공저자, 소속학·협회, 소속기관 멤버리스트, 웹콘텐츠

<표 9>는 각 검색 시스템이 제공하고 있는 기술요소를 나열한 것이고, <표 10>은 각 시스템이 제공하는 기술요소를 크게 다섯 가지 영역으로 분류한 다음, 각각의 기술요소에 대한 정보제공 여부를 각 시스템별로 표시한 것이다. 8개의 시스템은 전체적으로 50개의 기술요소들을 제공하고 있는 것으로 판별되었다. 개인정보에 관한 기술요소는 31개, 연구업적정보로는 12개, 사회관계정보에는 2개, 그리고 링크정보에 관한 2개의 기술요소와 정보공개에 관한 3개의 기술요소가 이에 속한다. 예상했듯이 인물정보 검색시스템에서 제공되는 기술요소는 인적정보 중심의 즉, 성명, 출생정보, 연락정보 그리고 인물의 학력·경력·소속정보를 중점적으로 제공하고 있다. 특히 학력·경력·소속정보를 근간으로 하는 인맥관계 정보제공에 그 중심을 두고 있는 것으로 나타났다.

그러나 연구자 관련 정보 검색시스템인 그 외의 5개의 시스템은 연구자의 학력·경력·소속정보 뿐만 아니라 연구자의 연구 산출물에 대한 상세정보와 연구분야 및 관심사에 대한 정보를 제공하고 있다. 이 중 SESTAT 시스템은 상세한 가족관련사항과 고용정보를 기술하도록 하여 단순한 인력정보 시스템이라기보다는 사회적 통계정보를 제공하는 기능도 갖추고 있다. 또한 ReaD와 KRI는 매우 유사한 속성을 가지고 있으며 연구분야와

연구경력, 연구물 정보가 상세하게 제공될 수 있도록 하였다.

<표 10> 각 시스템별 기술요소 비교

	제공업체 속성값	조인스	조선 닷컴	로마켓	KRI	NTIS	SES TAT	ReaD	COS
개인 정보	성명	○	○	○	○	○	○	○	○
	한문성명	○	○	○	○	○		○	
	영문성명	○	○	○	○	○	○	○	○
	생년월일	○	○	○	○	○			
	나이						○		
	성별	○	○	○	○		○	○	
	출생지	○	○	○					○
	가족사항	○	○	○			○		
	시민권/국적		○		○	○	○		
	인종/민족						○		
	장애상태						○		
	소속	○	○	○	○	○	○		
	임용일				○				
	퇴직일				○				
	직업	○	○	○	○			○	
	전문분야	○	○	○	○	○			○
	취업상태						○		
	부서/직위	○	○	○					
	가입단체	○	○					○	
	소속학·협회			○			○		○
	학력	○	○	○	○	○	○	○	
	시험/자격증	○	○	○	○				
	경력	○	○	○	○		○	○	
	자택주소	○	○						
	자택전화	○	○		○				
	직장주소	○	○	○	○		○	○	
	전화번호/팩스	○	○	○	○		○	○	
	핸드폰번호	○	○		○				
	이메일	○	○	○	○				○
	개인홈페이지	○	○	○	○	○			○
	연구자 등록번호				○				
연구 업적 정보	수상	○	○	○	○				
	논문	○		○	○	○		○	○
	연구보고서					○		○	○
	연구분야 키워드							○	
	현재 연구과제							○	
	저·역서	○	○	○	○	○	○	○	
	연구비지원기관				○				
	프로젝트			○					

사회 정보 링크 정보 공 개 정보	지식재산			○	○				
	전시작품				○				
	학술활동				○				
	기술이전				○				
	교우/지인	○	○	○					
	소속기관멤버정보							○	○
	웹정보 보기					○			○
	기사 검색					○			
	정보공개 여부				○				
	타기관 정보제공				○				
공 개 정보	메일 수신여부				○				

1.3 연구자정보 기술요소 추출

연구자에 대해 다양한 정보와 특징을 기술할 수 있도록 해주는 기술요소들을 찾아내는 작업은 온톨로지 설계에 중요한 기반 작업이라고 볼 수 있다. 이에 따라 본 장에서는 앞에서 기술요소를 비교분석한 자료를 토대로 하여 핵심적 공통 기술요소를 추출한 후, 더불어 연구자정보가 지니는 특성을 기반으로 하여 연구자정보를 기술하는데 적합한 클래스 확정하였다.

웹 인물정보 시스템과 국내의 학술연구자정보 시스템 서비스에 제공되고 있는 기술요소들 중 4개 이상의 시스템에서 공통적으로 나타난 기술요소들을 연구자정보 속성으로 추출하였다. 추출된 기술요소⁵⁹⁾들은 성명, 한문성명, 영문성명, 생년월일(나이), 성별, 출생지, 국적, 가족사항, 소속, 가입단체(소속학·협회), 면허/자격증, 경력, 직장주소, 직장전화/팩스, 이메일, 홈페이지, 수상, 논문, 저·역서 등이었다. 선정된 기술요소들 중 공통된 속성을 그룹화한 결과, 즉 연구자의 개인적 정보를 알려주는 인적정보, 연구성과물을 제시해주는 연구업적정보, 연구자의 연구성과물 이외의 활동을 알려주는 직무정보, 연구자가 소속되어 있는 다양한 기관과 학·협회를 알려주는 소속정보로 나눌 수 있다.

59) 비슷한 성격의 기술요소들은 통합하였다: 생년월일(나이), 가입단체(소속학·협회).

<표 11> 범주별 추출된 공통 기술요소

범주	항목
인적정보	성명, 한문성명, 영문성명, 생년월일, 성별, 출생지, 국적, 전문 분야, 학력, 직장주소, 직장전화번호/팩스, 이메일, 홈페이지, 경력, 수상, 면허/자격증, 가족사항
연구업적정보	논문, 저·역서
직무정보	직업
소속정보	소속, 가입단체(소속학·협회)

<표 11>은 이와 같이 4개의 범주에 앞서 8개 기관에서 공통적으로 나타난 기술요소를 할당한 결과이다. 인적정보에 대한 기술요소는 세분화되어 있는 반면, 연구업적정보, 직무정보, 소속정보에 대한 기술요소는 그다지 많지 않다. 연구를 수행하는 연구자에게 도움을 줄 수 있는 연구자정보 데이터베이스를 구축하기 위해서는 인적정보 이외의 정보를 상세하게 기술할 수 있는 항목개발이 필요하고 또한, 기술요소간의 계층 및 연관관계를 구체화할 필요가 있다.

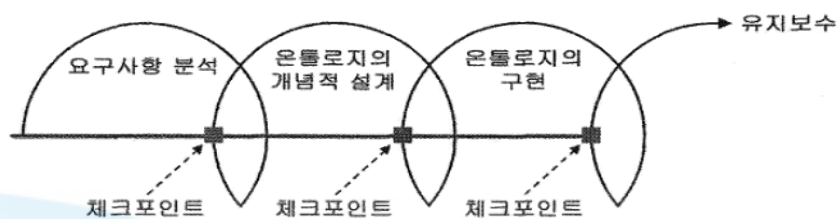
2. 연구자정보 온톨로지 설계

시맨틱 웹 기반으로 연구자정보를 검색하기 위하여 먼저 웹 도메인 온톨로지를 구축하여야 한다., 복잡한 온톨로지를 체계적으로 구축하기 위하여 선행되는 것이 바로 개념적 설계⁶⁰⁾이며 본 장에서는 이와 같은 개념적 설계를 EOE(Evolving Ontology Engineering) 방법에 따라 수행하였다.

EOE 방법론은 온톨로지를 점진적으로 구축하게 하는 방법으로 먼저 요

60) N. F. Noy and D. L. McGuinness. 2002. 『Ontology Development 101 : A Guide to Creating Your First Ontology』, Stanford University press.

구분석을 실시하여 이용자가 원하는 정보를 제공할 수 있도록 도메인에 대한 기술요소와 특징을 파악하도록 한 다음, 도메인에 대한 온톨로지를 개념적으로 설계하고 그 마지막으로 온톨로지를 구현하도록 하고 있다⁶¹⁾. 이와 같은 EOE 방법론은 각 프로세스가 진행되는 동안 ‘문서화’와 ‘지식 획득’ 작업이 지속적으로 진행되도록 하고 있으며 각 단계가 마무리되기 전에 점검을 통해 미흡한 부분에 대해 수정하도록 하고 있다. 이러한 EOE 방법론은 각 계속적인 점검과 단계별 진행으로 신뢰성이 있는 것으로 인지되고 있다⁶²⁾. <그림 4>는 이러한 세 가지 단계적 프로세스를 설명하고 있다.



<그림 4> EOE 방법론 프로세스

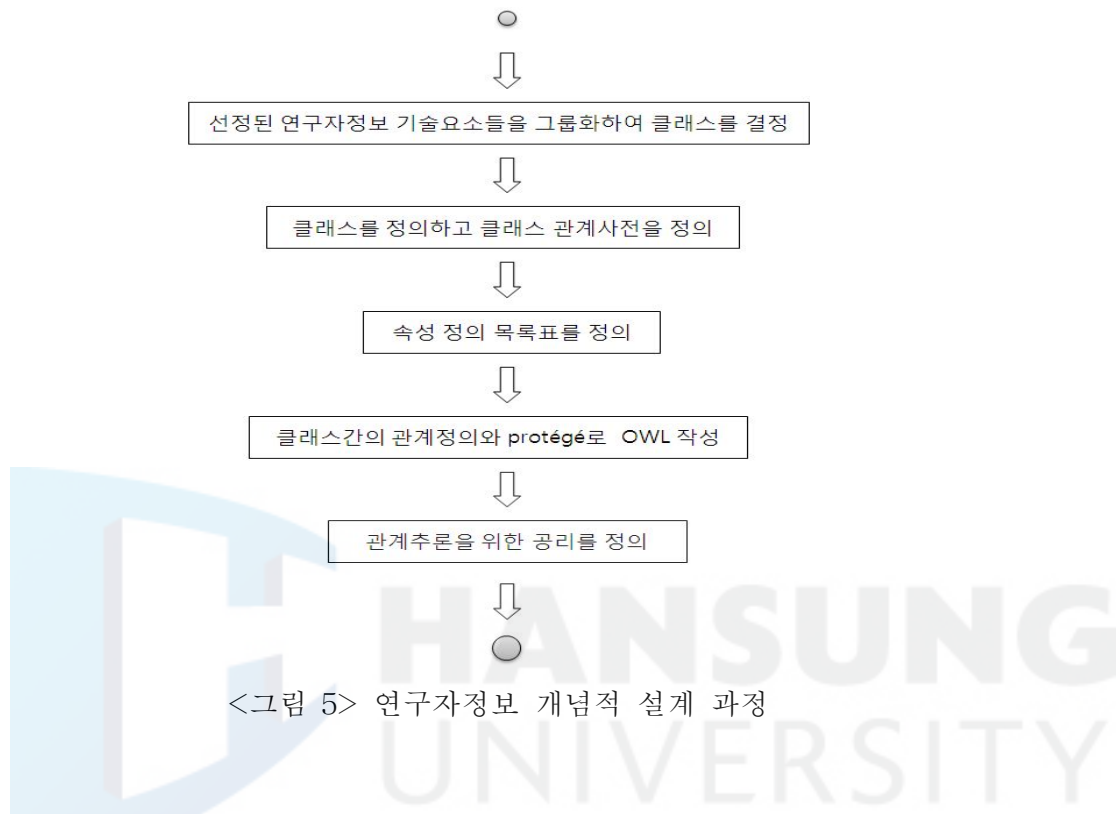
[출처: 강래구. 2007. p. 23.]

본 연구에서는 요구분석 단계는 연구자정보에 대한 실질적 요구분석을 수행하기 보다는 연구자/인물에 대한 정보를 제공하는 시스템을 분석하여 도메인에 대한 요구를 분석하였다. 다음, 연구자정보에 대한 개념적 설계는 한국전산원의 웹 온톨로지 개발지침 연구에서 제시한 단계를 기본으로 하여 진행하였다. 즉, 첫 번째 단계로 선정된 연구자정보 기술요소들을 비슷한 속성으로 그룹지어 클래스를 결정하였고 두 번째 단계로 클래스를 정의하고 관계사전을 만든 다음, 세 번째로 속성 정의 목록표를 정의하였다. 네 번째로 계층화된 클래스간의 관계를 새롭게 정의하여 검색 인터페이스에서 추론된 정보를 브라우징할 수 있도록 하였다. 이때 정의된 클레

61) 한국전산원. 2004. 전계서.

62) 강래구. 2007. 『시맨틱 웹 환경에서 온톨로지 기반의 지능형 상품 검색 시스템 설계 및 구현』.. 석사학위논문, 조선대 대학원 p. 23.

스와 속성들은 실제 OWL 편집기인 protégé를 이용하여 OWL언어로 작성하였다. 마지막으로 관계추론을 위한 공리를 정의하였다(그림 5 참조). 이로써 연구자정보의 온톨로지를 개념적으로 설계하였고 다음의 각 장은 단계별로 분석된 결과이다.



<그림 5> 연구자정보 개념적 설계 과정

2.1 클래스 이름 및 클래스의 계층구조 정의

클래스의 이름을 정의하는 것은 온톨로지 모델링에 있어 중요한 일이다. EOE 방법론에서는 클래스 이름을 정의하고 체계적인 이름 관리를 위해 클래스 사전을 이용할 것을 제시하고 있다. 클래스 사전은 명사 분석 요약서나 기존의 온톨로지에서 추출한 클래스의 이름을 정의한 문서이다. 일반적으로 클래스의 이름을 정의하고 계층의 구조를 정의하는 데 있어서 고려해야 할 규칙은 다음과 같다⁶³⁾.

- R.1.1 유사어는 통합하여 하나의 클래스로 정의한다.
- R.1.2 클래스 이름을 명확하게 정의한다.
- R.1.3 클래스 이름에 대한 다국어 표현을 고려한다.
- R.1.4 단수형을 복수형의 IS-A 관계로 정의하지 않는다.
- R.1.5 형제 관계는 같은 단계로 정의한다.
- R.1.6 하나의 클래스에 서브 클래스가 12개 이상일 경우엔 새로운 부모 클래스를 만든다.
- R.1.7 다중 상속을 고려한다.
- R.1.8 복잡한 관계에 있는 클래스들을 새로운 클래스로 표현하고, 이 클래스와 기존 클래스의 관계를 재정의한다.

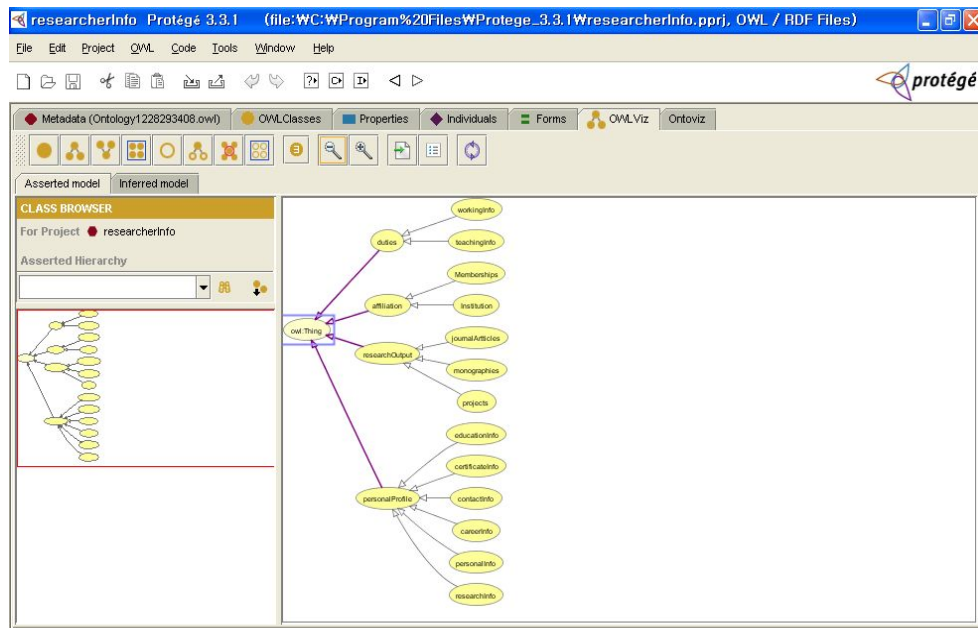
본 연구는 첫 번째 단계로 연구자정보를 기술하는데 필요한 클래스를 앞 장에서 추출한 4개의 범주로 대체 하였고, 핵심 클래스에 해당되는 행위<하위> 클래스를 선정하였다. 핵심클래스는 최상위 클래스 즉 이용자정보 아래 주요 기능을 구분해주는 기능을 하고, 행위 클래스는 실세계에서 계속적으로 생성되고 활용되는 영역을 지칭하는 것이다. <표 12>은 최상위 클래스에서 파생된 상위·하위 클래스간의 관계와 각 클래스의 이름과 설명을 정의한 것이다. 연구자정보 클래스 사전은 각 클래스의 이름, 영문명, 클래스의 설명과 비교란을 두어 보다 상세한 정보를 알 수 있도록 하였다.

본 연구에서는 클래스 및 속성간의 도식화된 계층구조를 얻기 위해 스탠포드 대학에서 개발한 Protégé 3.3.1 온톨로지 에디터를 사용하였다. 맨 왼쪽이 최상위 클래스(owl:thing)이고, 그 하위에 핵심 클래스로서 인적정보, 학력정보, 연구업적정보, 직무정보, 소속정보의 클래스가 있다. 각 핵심 클래스에는 각각의 행위 클래스가 하위에 위치하게 된다. <그림 6>은 각 클래스를 정의하고 그에 따른 클래스의 계층화 모습을 보여준다. 이에 따라 클래스 사전에서 정의된 18개의 클래스를 다시 계층화하여 상위 클래스 4개와 하위 클래스 14개로 구조화하였고 각각의 하위 클래스에 앞서 추출된 기술요소에 대응하는 속성을 선정하였다.

63) 한국전산원. 2004. 전게서.

<표 12> 연구자정보 클래스 사전

클래스명	영문명	설명	비고
강의정보	teachingInfo	연구자가 현재 강의하고 있는 강의의 강의명, 강의소개, 참고문헌, 강의 계획서 등의 정보를 담고 있음.	행위클래스
개인정보	personalInfo	연구자의 이름, 나이, 성별, 생년월일, 출생지 등의 정보를 담고 있음.	행위클래스
경력정보	careerInfo	연구자의 강의경력과 직장경력 등의 정보를 담고 있음.	행위클래스
소속기관	institution	연구자가 소속된 기관의 이름, 부서명, 기관종류, 직위 등의 정보를 담고 있음.	행위클래스
소속정보	affiliation	연구자가 소속된 기관의 정보나, 소속된 학·협회 등의 정보를 담고 있음.	핵심클래스
소속학·협회	memberships	연구자가 소속된 학·협회의 학·협회명, 학·협회상대, 활동기간, 연락처 등의 정보를 담고 있음.	행위클래스
업무정보	workingInfo	연구자의 현재 주요업무1, 주요업무2, 직무종류, 관련업무 분야 등의 정보를 담고 있음.	행위클래스
연구정보	researchInfo	연구자의 주요 연구분야를 담고 있음.	행위클래스
연구업적정보	researchOutput	연구자의 학술논문, 연구보고서, 프로젝트 등의 연구업적물 정보를 담고 있음.	핵심클래스
연구자정보	researcherInfo	연구자의 인적, 연구업적물, 직무, 소속기관 등 연구자의 모든 속성을 가지고 있음.	최상위클래스
연락정보	contactInfo	연구자의 자택주소, 자택전화, 직장주소, 직장전화, 직장팩스, 홈페이지, 이메일 등의 정보를 담고 있음.	행위클래스
인적정보	personalProfile	연구자의 이름, 나이, 성별과 같은 개인적 정보와 연구업적물, 경력, 자격등의 훈·포상 정보를 담고 있음.	핵심클래스
저·역서	monographies	연구자가 저술한 연구보고서나 저역서 등의 단행본 정보를 담고 있음.	행위클래스
직무정보	duties	연구자의 강의정보나 업무정보를 담고 있음.	핵심클래스
프로젝트	projects	연구자가 수행한 프로젝트의 성과물에 대한 정보를 담고 있음.	행위클래스
학력정보	educationInfo	연구자가 졸업한 고등학교, 대학교, 대학원의 정보를 담고 있음.	행위클래스
학술논문	journalArticle	연구자의 국내외 학술잡지에 게재된 학술논문 정보를 담고 있음.	행위클래스
훈·포상정보	certificateInfo	연구자의 자격증, 수상정보 등을 담고 있음.	행위클래스



<그림 6> protégé에서 표현된 계층화된 클래스들

2.2 속성의 선정

핵심클래스와 행위클래스가 확정되어지면 이에 대한 속성을 지정해주어야 한다.: 속성 정의 규칙

R.2.1 해당 도메인에서 일반적으로 통용되지 않는 것은 속성으로 정의하지 않는다.

R.2.2 수퍼 클래스에 공통적인 속성은 서브 클래스보다 수퍼 클래스의 정의한다.

R.2.3 속성값이 더 이상 분해할 수 없는 특징일 경우 속성으로 정의한다.

R.2.4 속성값이 분해할 수 있다면 클래스로 정의하고, 이를 속성으로 제약한다.

이와 같은 규칙에 따라 행위클래스에 속하는 속성을 선정하였다(표 13 참조).

<표 13> 연구자정보 클래스 관계 사전

최상위 클래스	핵심 클래스	행위 클래스	속성명	영문 속성명
연구자정보 researcherInfo	인적정보 personalProfile	개인정보 personalInfo	성명	name
			영문성명	engName
			한문성명	hanName
			성별	gender
			생년월일	birthDate
			출생지	birthPlace
			국적	nationality
			가족사항	familyRelations
		학력정보 educationInfo	고등학교	highSchool
			대학교/학과	universityBA
			대학원/전공분야/석사	graduateSchoolMS
			대학원/전공분야/박사	graduateSchoolPhD
		연락정보 contactInfo	주택 주소	homeAddress
			직장 주소	officeAddress
			주택 전화	homePhone
			직장 전화	officePhone
			휴대폰	cellPhone
			이메일	email
			홈페이지 주소	homepage
		연구정보 researchInfo	주요연구분야1	primaryTopic
			주요연구분야2	secondaryTopic
			현재진행중연구	onGoingResearch
		경력정보 careerInfo	학교경력	academicCareer
			직장경력	professionalCareer
		훈·포상정보 certificateInfo	자격명	certificateName
			자격취득일	certificateDate
			수상명	awardName
			수상일	awardDate
			수여기관	endowInstitution
	연구업적 정보 researchOutput	학술논문 journalArticles	논문한글제목	korTitle_A
			논문영문제목	engTitle_A
			잡지명	journalName
			발행년도	pubYear_A
			권	volume
			호	number
			주제 키워드	subjectKeyword_A

			CODEN	CODEN
			DOI	DOI
	저·역서 monographies	저·역서 monographies	지역서 한글서명	korTitle_M
			지역서 영문서명	engTitle_M
			발행장소	pubPlace_M
			발행처	pubName_M
			발행년도	pubYear_M
			주제 키워드	subjectKeyword_M
			ISBN	ISBN
	프로젝트 projects	프로젝트 projects	프로젝트 한글명	korTitle_P
			프로젝트 영문명	engTitle_P
			발행기관	organization_P
			공동연구자	co-workers
			연구비 지원 기관	fundedBy
			발행년도	pubYear_P
			수행기간	duration
			주제 키워드	subjectKeywords_P
직무정보 duties	강의정보 teachingInfo	강의정보 teachingInfo	강의코드	courseID
			강의명	courseName
			강의소개	courseDescription
			강의참고문헌	courseReference
			강의수강 학점	courseUnit
			강의스케줄	courseSchedule
	직무정보 workingInfo	직무정보 workingInfo	주요업무1	primaryWork
			주요업무2	secondaryWork
			직무분야	workingType
			관련직무분야	relatedWorkingArea
소속정보 affiliation	소속기관 institution	소속기관 institution	기관명	institutionName
			기관소개	institutionDescription
			부서/학과명	departmentName
			기관종류	institutionType
			직위	position
	소속 학·협회 memberships	소속 학·협회 memberships	학·협회명	membershipName
			학·협회소개	membershipDescription
			학·협회직위	membershipSatus
			활동기간	duration_M
			연락처	contact

2.3 속성의 목록 정의

다음으로 연구자정보의 속성으로 지정된 항목에 대한 속성형식, 값형식 범위 등을 지정하는 속성목록을 만들었다. 이에 대한 결과는 <표 14>이다.

<표 14> 연구자정보 속성 목록

속성명	영문명	설명	속성형식	값형식	범위	차수	소속클래스
성명	name	연구자의 이름	인스턴스	String	XSD Type	(1, n)	personalprofile
영문성명	engName	연구자의 영문이름	인스턴스	String	XSD Type	(1, n)	personalprofile
한문성명	hanName	연구자의 한문이름	인스턴스	String	XSD Type	(1, n)	personalprofile
성별	gender	연구자의 성별	인스턴스	String	XSD Type	(1, 1)	personalprofile
생년월일	birthDate	연구자의 생년월일	인스턴스	YYYY-MM-DD		(1, 1)	personalprofile
출생지	birthPlace	연구자의 출생지	인스턴스	String	XSD Type	(1, 1)	personalprofile
국적	nationality	연구자의 국적	인스턴스	String	XSD Type	(1, 1)	personalprofile
가족사항	familyRelations	연구자의 가족관계	인스턴스	String	XSD Type	(null, n)	personalprofile
고등학교	highSchool	졸업한 고등학교	인스턴스	String	XSD Type	(1, 1)	educationInfo
대학교	universityBA	졸업한 대학교	인스턴스	String	XSD Type	(1, n)	educationInfo
대학원/석사	graduateSchoolMS	졸업한 대학원과 석사전공명	인스턴스	String	XSD Type	(null, n)	educationInfo
대학원/박사	graduateSchoolPhD	졸업한 대학원과 박사전공명	인스턴스	String	XSD Type	(null, n)	educationInfo
주택주소	homeAddress	연구자의 주택주소	인스턴스	String	XSD Type	(1, 2)	contactInfo
직장주소	officeAddress	연구자의 직장주소	인스턴스	String	XSD Type	(1, 2)	contactInfo
주택전화	homePhone	연구자의 주택전화번호	인스턴스	Number		(1, n)	contactInfo
직장전화	officePhone	연구자의 직장전화번호	인스턴스	Number		(1, n)	contactInfo
휴대폰	cellPhone	연구자의 휴대폰번호	인스턴스	Number		(1, n)	contactInfo
이메일	email	연구자의 이메일주소	인스턴스	String	XSD Type	(1, n)	contactInfo
홈페이지 주소	homepage	연구자의 홈페이지주소	인스턴스	String	XSD Type	(1, n)	contactInfo
주요연구 분야1	primaryTopic	연구자의 주요연구분야1	인스턴스	String	XSD Type	(1, n)	researchInfo
주요연구 분야2	secondaryTopic	연구자의 주요연구분야2	인스턴스	String	XSD Type	(1, n)	researchInfo

현재진행 중연구	onGoingResearch	연구자가 현재 진행중인 연구	인스턴스	String	XSD Type	(1, 1)	researchInfo
강의경력	academicCareer	연구자의 강의경력	인스턴스	String	XSD Type	(1, n)	careerInfo
직무경력	professionalCareer	연구자의 직무경력	인스턴스	String	XSD Type	(1, n)	careerInfo
자격명	certificateName	취득한 자격증의 이름	인스턴스	String	XSD Type	(1, n)	certificateInfo
자격취득 일	certificateDate	자격을 취득한 날짜	인스턴스	YYYY-MM-DD		(1, n)	certificateInfo
수상명	awardName	수상한 상의 이름	인스턴스	String	XSD Type	(1, n)	certificateInfo
수상일	awardDate	상을 수여받은 날짜	인스턴스	YYYY-MM-DD		(1, n)	certificateInfo
수여기관	awardInstitution	상을 수여받은 기관의 이름	인스턴스	String	XSD Type	(1, n)	certificateInfo
논문한글 서명	korTitle_A	학술논문의 한글제목	인스턴스	String	XSD Type	(null, n)	journalArticles
논문영문 서명	engTitle_A	학술논문 영문제목	인스턴스	String	XSD Type	(null, n)	journalArticles
잡지명	journalName	학술논문이 게재된 잡지의 이름	인스턴스	String	XSD Type	(null, n)	journalArticles
발행년도	pubYear_A	학술논문이 발행된 년도	인스턴스	YYYY-MM-DD		(null, n)	journalArticles
권	volume	학술논문이 게재된 잡지의 권	인스턴스	Number		(null, n)	journalArticles
호	number	학술논문이 게재된 지널의 호	인스턴스	Number		(null, n)	journalArticles
주제 키워드	subjectKeywords_A	학술논문의 주제 키워드	인스턴스	String		(null, n)	journalArticles
CODEN	CODEN	과학과 기술분야의 특정 연속간행물을 식별할 수 있도록 6자리의 문자로 부여되는 부호	인스턴스	Number		(null, n)	journalArticles
DOI	DOI	DOI (digital object identifier) 디지털 객체 식별자	인스턴스	Number		(null, n)	journalArticles
저역서 한글서명	korTitles_M	연구자가 저술한 저서나 역서의 한글서명	인스턴스	String	XSD Type	(null, n)	monographies
저역서 영문서명	engTitles_M	연구자가 저술한 저서나 역서의 영문서명	인스턴스	String	XSD Type	(null, n)	monographies

지역서 발행장소	pubPlace	지역서를 발행한 장소	인스턴스	String	XSD Type	(null, n)	monographies
지역서 발행처	pubName	지역서를 발행한 기관	인스턴스	String	XSD Type	(null, n)	monographies
지역서 발행년도	pubYear_M	지역서를 발행한 년도	인스턴스	YYYY-MM-DD		(null, n)	monographies
지역서 주제 키워드	subjectKeywords_M	지역서의 주제 키워드	인스턴스	String	XSD Type	(null, n)	monographies
ISBN	ISBN	지역서의 국제표준도서번호	인스턴스	Number		(null, n)	monographies
프로젝트 한글명	korTitle_P	프로젝트성파물의 한글명	인스턴스	String	XSD Type	(null, n)	projects
프로젝트 영문명	engTitle_P	프로젝트성파물의 영문명	인스턴스	String	XSD Type	(null, n)	projects
프로젝트 발행기관	organization	프로젝트성파물을 발행한 기관이름	인스턴스	String	XSD Type	(null, n)	projects
프로젝트 공동연구자	co-workers	연구자와 공동으로 참여한 프로젝트 공동연구자	인스턴스	String	XSD Type	(null, n)	projects
프로젝트 연구비지 원기관	fundedBy	연구자의 프로젝트 연구비를 지원하는 기관이름	인스턴스	String	XSD Type	(null, n)	projects
프로젝트 발행년도	pubYear_P	연구자의 프로젝트 발행년도	인스턴스	YYYY-MM-DD		(null, n)	projects
수행기간	duration_P	연구자가 프로젝트 프로젝트에 참여한 기간	인스턴스	YYYY-MM-DD*YYYY-MM-DD		(null, n)	projects
프로젝트 주제 키워드	subjectKeywords_P	프로젝트성파물의 주제 키워드	인스턴스	String	XSD Type	(null, n)	projects
강의코드	courseID	연구자가 현재 강의하고 있는 강의의 코드	인스턴스	Number		(null, n)	teachingInfo
강의명	courseName	현재 강의하고 있는 강의의 이름	인스턴스	String	XSD Type	(null, n)	teachingInfo
강의소개	courseDescription	현재 강의하고 있는 강의 소개글	인스턴스	String	XSD Type	(null, n)	teachingInfo
강의참고 문헌	courseReference	현재 강의하고 있는 강의의 참고문헌	인스턴스	String	XSD Type	(null, n)	teachingInfo
강의수강 학점	couseUnit	현재 강의하고 있는 강의의 학점	인스턴스	Number		(null, n)	teachingInfo

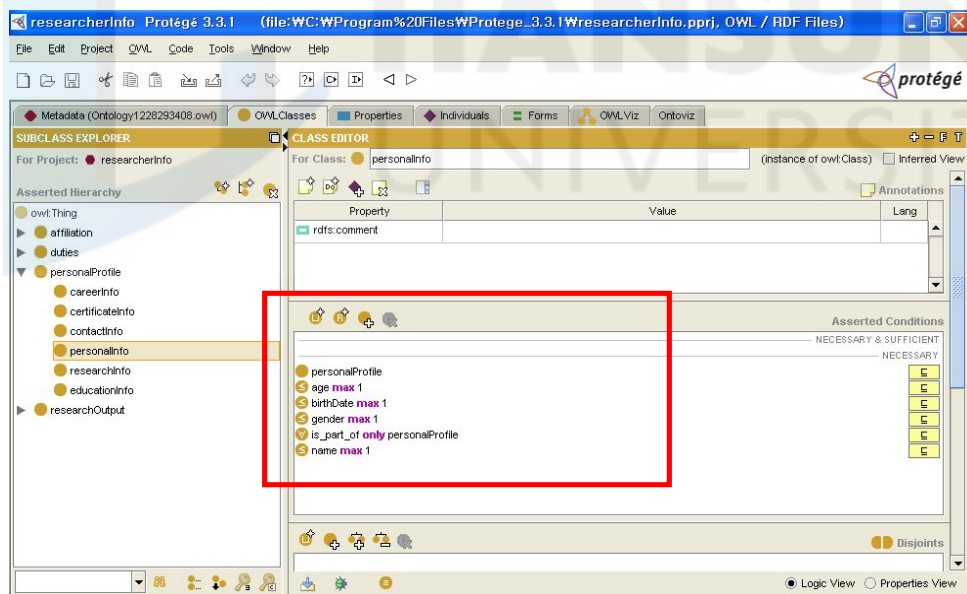
강의스케줄	couseSchedule	현재 강의하고 있는 강의의 스케줄	인스턴스	String	XSD Type	(null, n)	teachingInfo
주요업무1	primaryWork	연구자의 주요업무1	인스턴스	String	XSD Type	(null, n)	workingInfo
주요업무2	secondaryWork	연구자의 주요업무2	인스턴스	String	XSD Type	(null, n)	workingInfo
직무분야	workingType	현재하고 있는 일의 직무분야	인스턴스	String	XSD Type	(null, n)	workingInfo
관련직무 분야	relatedWorkingArea	현재하고 있는 일의 관련된 직무 분야	인스턴스	String	XSD Type	(1, n)	workingInfo
소속기관명	institutionName	연구자가 소속된 소속기관이름	인스턴스	String	XSD Type	(1, n)	institution
소속부서명	departmentName	소속된 부서이름	인스턴스	String	XSD Type	(1, n)	institution
소속기관 종류	institutionType	소속된 기관종류	인스턴스	String	XSD Type	(1, n)	institution
직위	position	현재 직위	인스턴스	String	XSD Type	(1, n)	institution
학·협회명	mebershipName	소속된 학·협회의 이름	인스턴스	String	XSD Type	(null, n)	memberships
학·협회직위	membershipsStatus	소속된 학·협회에서의 직위	인스턴스	String	XSD Type	(null, n)	memberships
활동기간	duration	소속된 학·협회에서의 활동기간	인스턴스	YYYY-MM-DD/YYYY-MM-DD		(null, n)	memberships
연락처	contact	소속된 학·협회의 연락처	인스턴스	Number		(null, n)	memberships

2.4 속성의 제약조건 정의

속성을 제약하는 것은 온톨로지를 구축할 때 발생하는 오류를 최소화시키며 향후 추론규칙을 구성하는데 그 기본이 된다. 아래 내용은 EOE가 제시하는 제약조건 기술 규칙으로 일반적으로 열두 가지 영역을 제한하도록 하고 있다.

- R.3.1 owl:allValuesFrom 클래스의 모든 인스턴스에 대해 속성의 모든 값이 제약을 만족하는 것을 나타낸다.
- R.3.2 owl:someValuesFrom 적어도 하나의 값은 주어진 클래스에 선언되어야 하는 것을 말한다.
- R.3.3 wl:hasValues 속성의 값을 특정한 개체로 제한할 때 사용한다.
- R.3.4 owl:maxCardinality 속성값의 최대 출현 회수
- R.3.5 owl:minCardinality 속성값의 최소 출현 회수
- R.3.6 owl:cardinality 지정된 회수만 출현 가능
- R.3.7 owl:equivalentProperty 2개의 속성이 서로 서브속성일 때 사용한다.
- R.3.8 owl:inverseOf 2개의 속성이 반대관계일 때 나타낸다.
- R.3.9 owl:FunctionalProperty 역함수형 속성값의 출현횟수가 0 또는 1이라는 제약
- R.3.10 owl:InverseFuctionalProperty 역함수형 속성값의 출현횟수가 1이라는 제약
- R.3.11 owl:TransitiveProperty 논리적으로 추이적인 성질을 나타낸다.
- R.3.12 owl:SymmetricProperty 논리적으로 대칭적인 성질을 나타낸다.

위의 속성이 지니는 제약조건 요소와 속성 목록에서 기술된 성질을 이용하여, 각 속성의 상세한 정의를 기술하였다.



<그림 7> protégé에서 personalInfo의 속성 제약 화면

<그림 7>은 protégé 3.3.1을 이용하여 personalInfo 클래스의 속성을 기술하는 화면이고 <표 15>는 인적정보를 OWL로 기술한 결과의 일부를 보여주고 있다.

<표 15> OWL로 기술된 인적정보 결과 일부

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.owl-ontologies.com/Ontology1228293408.owl#"
  xml:base="http://www.owl-ontologies.com/Ontology1228293408.owl"
  xmlns:p1="http://www.owl-ontologies.com/assert.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#">
  <owl:Ontology rdf:about=""/>
  <owl:DatatypeProperty rdf:ID="age">
    <rdf:type
rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
    <rdfs:domain rdf:resource="#personalInfo"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="birthPlace">
    <rdfs:domain rdf:resource="#personalInfo"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="birthDate">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#date"/>
  </owl:DatatypeProperty>
  <owl:Class rdf:ID="careerInfo">
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#is_part_of"/>
        <owl:allValuesFrom rdf:resource="#personalProfile"/>
      </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf rdf:resource="#personalProfile"/>
  </owl:Class>
  <owl:DatatypeProperty rdf:ID="cellPhone">
```



```

    <rdfs:domain rdf:resource="#contactInfo"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:Class rdf:ID="certificateInfo">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#is_part_of"/>
      <owl:allValuesFrom rdf:resource="#personalProfile"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="#personalProfile"/>
</owl:Class>
<owl:Class rdf:ID="contactInfo">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#is_part_of"/>
      <owl:allValuesFrom rdf:resource="#personalProfile"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="#personalProfile"/>
</owl:Class>
<owl:DatatypeProperty rdf:ID="email">
  <rdfs:domain rdf:resource="#contactInfo"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="gender">
  <rdf:type
rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/>
  <rdfs:domain rdf:resource="#personalInfo"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="homeAddress">
  <rdfs:domain rdf:resource="#contactInfo"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="homepage">
  <rdfs:domain rdf:resource="#contactInfo"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>

```

```

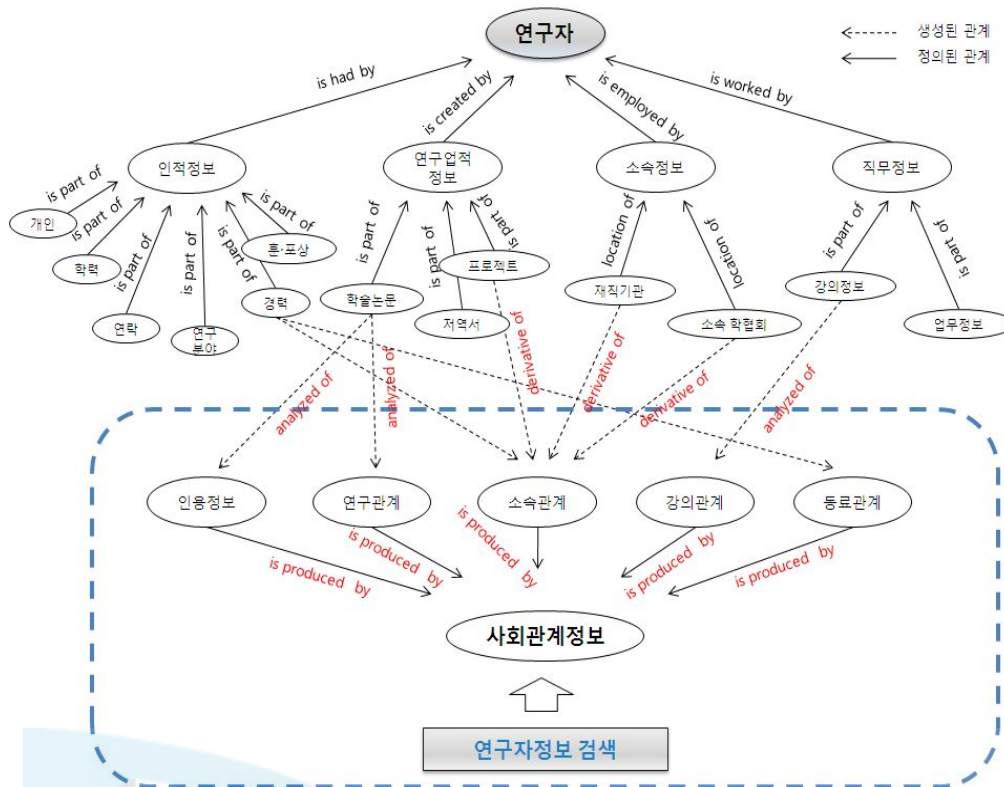
<owl:DatatypeProperty rdf:ID="homePhone">
  <rdfs:domain rdf:resource="#contactInfo"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>

```

2.5 관계추론을 위한 공리 정의

온톨로지 분석은 관계추론을 가능하게 해준다. 특히 시맨틱 웹기반일 경우 기술된 속성 이외에 관련된 웹사이트에 관계정보를 추가하여 관련된 웹정보들이 검색될 수 있고 또한 정보 갱신이 자동적으로 이루어질 수 있다. 다음의 그림 8은 구축된 속성정보를 기반으로 생성될 수 있는 정보와 추론과정을 통하여 생성될 수 있는 정보에 대한 개념도이다.

<그림 8>은 본 연구에서 추론될 수 있는 정보를 정의한 공리를 바탕으로 한 관계정의 모델이다. 가장 상위에 연구자 클래스가 있고, 그 하위에 각각 인적정보, 연구정보, 소속정보, 직무정보 4개의 클래스가 위치하고 있다. 인적정보 클래스는 연구자정보의 하위 클래스로서 'is had by'라는 관계를 통해 정의되고, 연구정보 클래스는 연구자에 의해 'is created by'라는 관계를 통해 정의된다. 마찬가지로 소속정보는 연구자에 의해 'is employed by' 되며, 직무정보는 'is worked by'로 정의된다. 이 네 개의 핵심 클래스들은 각각의 하위(행위) 클래스를 가지고 있으며, 각 클래스의 성질에 따라 'is part of', 'location of' 등으로 정의된다. 각 행위 클래스들 간의 공리를 통해 연구관계, 소속관계, 강의관계, 동료관계 등을 추론할 수 있게 된다. 또한, 구축된 연구자정보를 기반으로 하여 각 연구자의 연구업적물에 대한 인용정보도 생성될 수는 있으나 본 연구에서는 이에 대한 추론은 제외하였다.



<그림 8> 연구자정보 클래스간의 관계와 추론 정의

이와 같은 관계추론을 구현시키기 위해서는 먼저 클래스간의 관계를 공리로 정의내려야 한다. 공리(Axiom)란 선언적이고 동시에 엄밀하게 표현된 지식으로 그 옳음을 증명하지 않고도 받아들일 수 있는 지식으로 논리학에서 공리는 추론의 전제가 되는 지식을 나타낸다. 온톨로지 구축 시 공리는 어휘/개념의 의미 정의를 엄밀하게 표현할 수 있도록 해주며, 온톨로지에 포함된 어휘/개념을 사용하여 질문에 대해 추론할 수 있도록 해주는 역할을 한다⁶⁴⁾.

예를 들어 클래스를 구성하는 개체들을 모두 열거하여 클래스를 정의할 때 사용할 수 있으며, 클래스의 구성 개체들은 정확히 열거된 개체들의 집

64) 한국전산원. 2004. 전계서.

합과 일치한다. 만약, 계절을 클래스로 정의할 때, 개체들 즉, 봄, 여름, 가을, 겨울을 나열하여 계절(Season)클래스를 정의할 수 있다. 이때 계절 클래스의 어떤 속성이 allValuesFrom 제약의 공역으로 선언되어 있다면, 추론엔진을 이용하여 그 속성의 maxCardinality가 4임을 유추할 수 있다.

본 연구에서 추론될 수 있는 정보를 정의하기 위해 다음과 같은 공리 목록서를 작성하였다(표 16 참조)와 클래스와 인스턴스간의 관계(표 17 참조)를 작성하였다.

<표 16> 연구자정보 공리 목록서

추론	구분	설명
연구관계	공리명	연구관계 정보
	내용	학술논문의 주제를 분석하여 같은 주제를 연구하는 연구자들의 관계를 보여준다.
	개념	creating
	관계	연구업적정보, 연구자 개인정보, 원문 데이터베이스
소속관계	공리명	소속관계 정보
	내용	연구자의 학력, 재직기관, 소속학·협회, 프로젝트 참여 정보를 내용으로 하여 같은 소속을 가진 연구자들의 관계를 보여준다.
	개념	creating
	관계	학력정보, 개인정보, 연구정보(프로젝트), 소속정보, 대학 및 기관 사이트, 학회 사이트, 협회 사이트
강의관계	공리명	강의관계 정보
	내용	연구자가 강의하는 과목명을 분석하여 같은 주제의 강의를 하고 있는 연구자들간의 관계를 보여준다.
	개념	creating
	관계	강의정보, 개인정보, 학과 소개 사이트
동료관계	공리명	동료관계 정보
	내용	연구자의 소속학교정보와 출신학교 정보를 분석하여 연구자의 동료관계를 보여준다.
	개념	creating
	관계	경력정보, 개인정보

<표 17> 연구자정보 관계 테이블

Relation name	Source concept	Source cardinality (Max)	Target concept	Inverse relation
is created by	researchOutput	N	researcherInfo	create
is employed by	affiliation	N	researcherInfo	employe
is had by	personalprofile	N	researcherInfo	has a
is part of	workingInfo	N	duties	has a
is part of	journalArticles	N	researchOutput	has a
is part of	monographies	N	researchOutput	has a
is part of	projects	N	researchOutput	has a
is part of	teachingInfo	N	duties	has a
is part of	educationInfo	N	personalProfile	has a
is part of	careerInfo	N	personalProfile	has a
is part of	certificateInfo	N	personalProfile	has a
is part of	cantactInfo	N	personalProfile	has a
is part of	researchInfo	N	personalProfile	has a
is part of	personalInfo	N	personalProfile	has a
is worked by	duties	N	researcherInfo	work
location of	institution	N	affiliation	has a
location of	memberships	N	affiliation	has a

IV. 연구자정보의 검색 인터페이스 설계

3장에서는 설계된 연구자정보 온톨로지를 기반으로 하여 연구자의 이름, 주제어/키워드, 소속기관 및 학·협회명으로 검색하여 연구자에 대한 사실 정보 뿐만 아니라 이를 기반으로 추론될 수 있는 연구관계, 소속관계, 강의관계, 동료관계 등의 사회관계 정보를 브라우징할 수 있으며, 각각의 속성을 조합하여 원하는 정보만 검색할 수 있는 검색 인터페이스를 설계하였다.

본 연구에서의 웹 기반 연구자정보 검색 시스템을 IS RIRS(Information Science Researcher Information Retrieval System)이라 명명하고자 한다. <그림 9>는 연구자정보 검색의 첫 화면이다. 연구자정보 검색 시스템을 개인정보의 보안을 위하여 접근권한과 이용권한을 통제하도록 하였다.

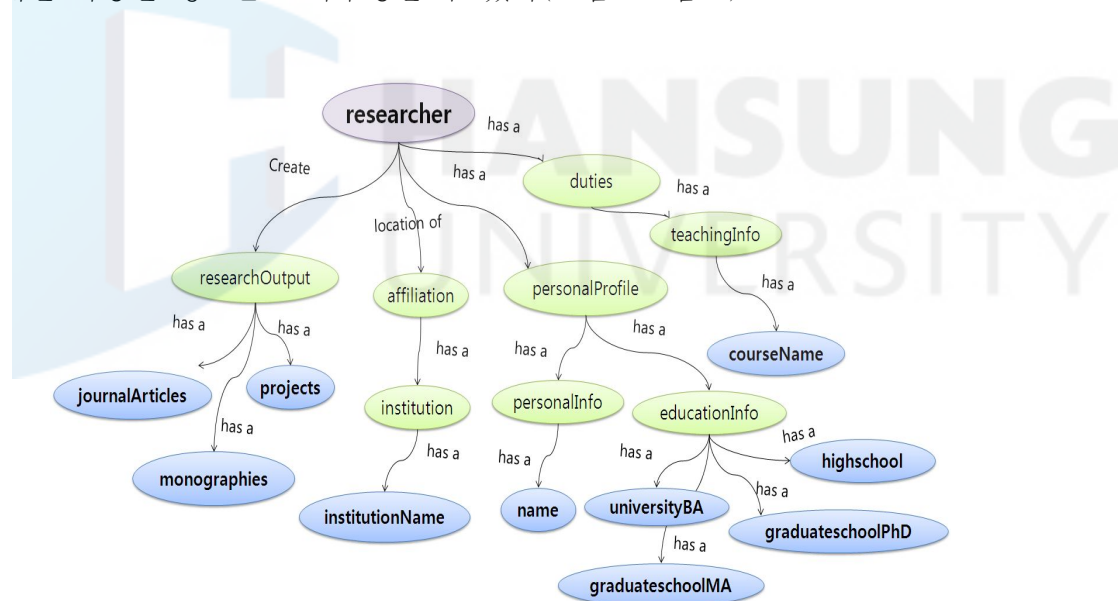


<그림 9> 연구자정보 검색시스템 로그인 화면

즉, 연구자의 개인 신상정보를 검색하기 위해서는 기본적으로 실명인증 및 공인인증서를 통한 로그인 절차가 필요하다. 따라서, 검색시스템의 첫 화면은 로그인 작업과 이에 대한 안내사항을 알려주고 직접 회원가입을 할 수 있는 기능을 가진다.

1. 인명 검색

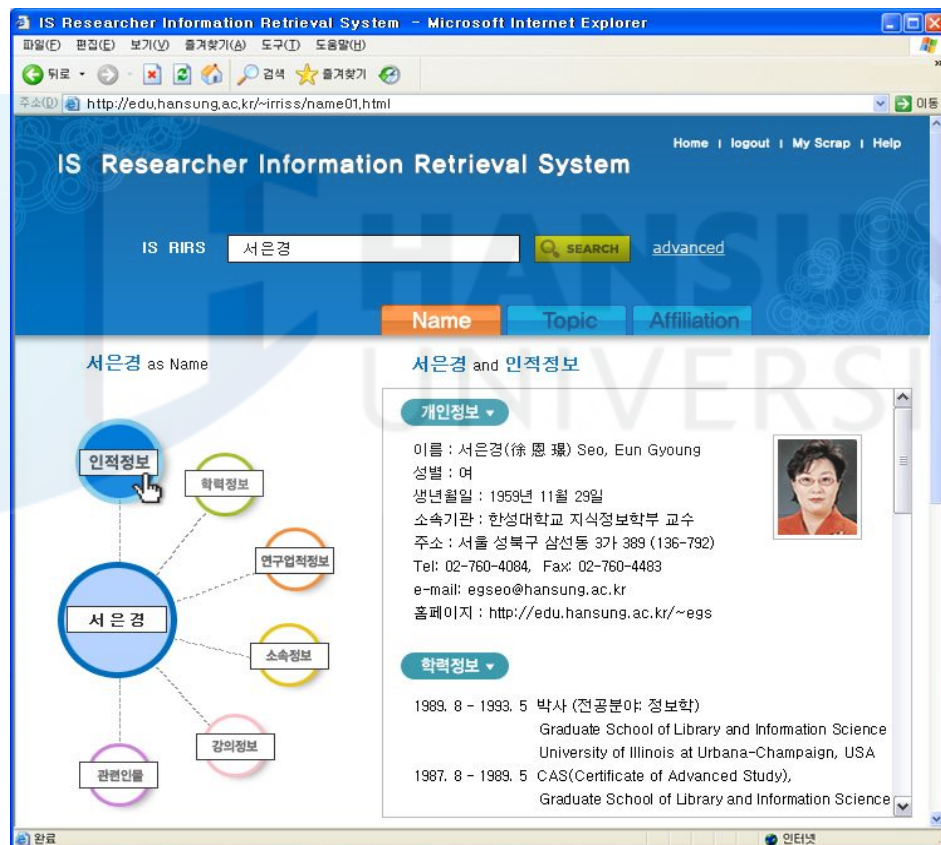
인명 검색은 <그림 10>의 관계정의를 토대로 하여 연구자의 이름을 검색 접근점으로 이용해 연구자에 대한 기본적인 개인정보, 연락처, 소속 정보, 강의정보, 관련인물정보 등을 한눈에 브라우징할 수 있도록 설계하였다. 연구자정보 검색시스템이 제공하는 3개의 접근점 중에서 'Name'으로 선택한 후 연구자명을 입력하게 되면 연구자의 인적정보, 학력정보, 연구업적정보, 소속정보, 강의정보, 동료정보의 세부 접근점에 따라 연구자에 대한 다양한 정보를 브라우징할 수 있다(그림 10 참조).



<그림 10> 인명 검색 관계정의

1.1 인적정보

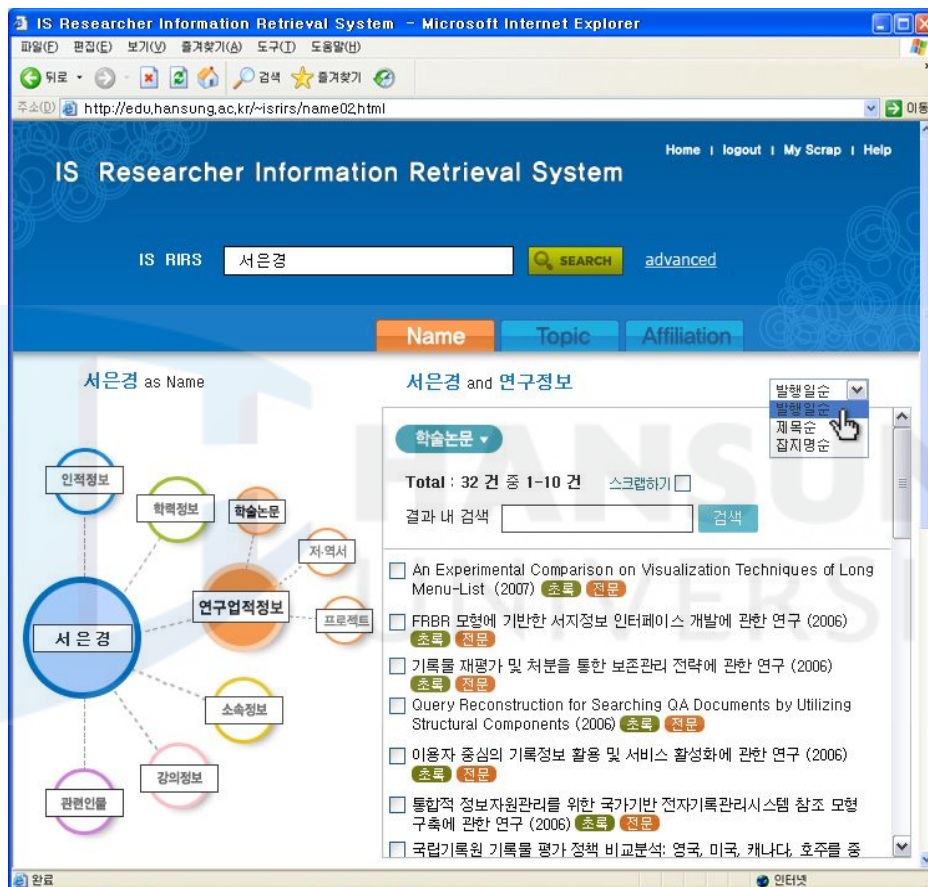
인적정보는 연구자의 이름을 검색 접근점으로 이용해 연구자에 대한 기본적인 개인정보, 학력정보, 연락처정보, 연구업적정보, 경력정보, 훈·포상 정보 등을 한눈에 브라우징할 수 있도록 설계하였다. <그림 11>과 같이 입력창에 ‘서은경’이라는 연구자의 이름을 입력하고 이중 인적정보를 선택하게 되면 해당 연구자의 이름, 성별, 생년월일, 출생지 등과 같은 개인정보와 자택주소, 직장주소, 전화번호, 휴대폰, 이메일, 홈페이지주소와 같은 연락처, 학력, 경력, 수상경력, 자격증 등의 정보를 브라우징할 수 있으며, 연구자가 각 정보의 공개여부를 선택할 수 있도록 설계하였기 때문에 비공개를 원할 경우 브라우징 검색 결과에는 나타나지 않는다.



<그림 11> 연구자 이름을 통한 인적정보 브라우징 결과 화면

1.2 연구업적정보

연구업적정보는 연구자정보의 하위 클래스 중 연구업적정보 클래스를 통해 구현될 수 있다. 연구자의 업적물인 학술논문, 연구보고서, 프로젝트 등의 정보를 메인화면의 오른쪽 프레임에서 브라우징할 수 있도록 설계하였다(그림 12 참조).



<그림 12> 연구자의 연구업적정보 브라우징 결과 화면

왼쪽 프레임의 연구자정보의 하위클래스 중 연구업적정보 원형이미지를 선택하면, 연구업적정보의 하위클래스로 연구자의 학술논문, 연구보고서,

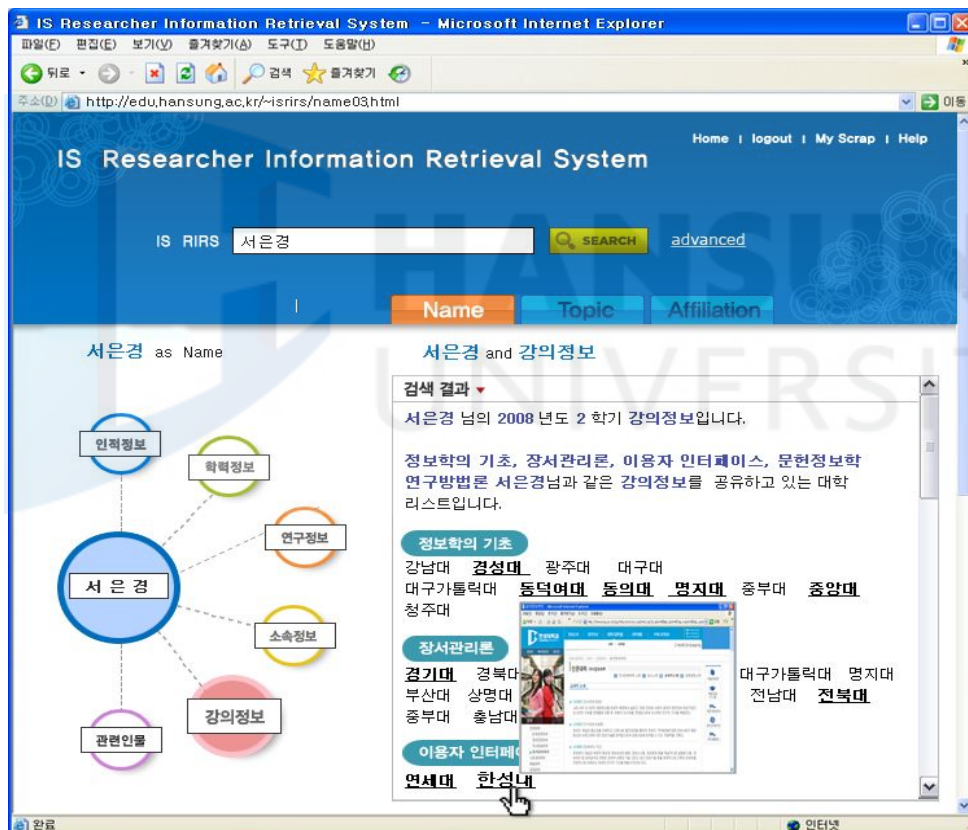
프로젝트 등이 제공되며 이중 이용자가 학술논문을 클릭했을 경우, 오른쪽 프레임에는 연구자의 업적물 중 학술논문이 나타나게 된다. 따라서 이용자는 연구자의 이름으로 검색하면 연구자의 모든 정보가 한꺼번에 나타나는 것이 아니라 왼쪽 선택 프레임에서 계층적으로 표현된 연구자의 속성을 보고 원하는 정보를 선택하여 오른쪽 프레임에서 브라우징할 수 있게 된다. 시스템은 검색결과 첫 번째 화면에는 연구업적 제목만을 제공하도록 하여 손쉽게 브라우징할 수 있게 하였고 보다 상세한 정보를 알기 위해서는 초록이나 목차 버튼을 클릭하도록 하였다. 또한 전문이 브라우징이 가능한 경우 전문 버튼이 활성화되고 원문데이터베이스의 링크를 통하여 제공해주도록 한다. 또한, 검색된 제목 앞에 체크박스를 두어 원하는 연구정보를 My Scrap에 저장할 수 있도록 하였고 검색결과가 학술논문인 경우에는 제목, 잡지명, 발행일 순으로 정렬이 가능하게 하였고, 단행본인 경우에는 제목명, 출판사명과 발행년도, 프로젝트일 경우는 제목명, 발행기관, 연구비지원기관 그리고 발행년도 순으로 정렬할 수 있게 하였다.

1.3 소속정보

소속정보 브라우징은 연구자정보의 하위 클래스 중 소속정보 클래스를 통해 구현될 수 있다. 연구자의 소속정보 클래스 중 기관 클래스와 소속학·협회 클래스의 속성을 이용해 연구자가 현재 재직 중인 기관과 학·협회에 대한 정보를 브라우징할 수 있다. 왼쪽 연구자 정보 선택 프레임에서 소속정보를 클릭하였을 경우, 하위클래스인 기관정보와 소속학·협회정보를 선택할 수 있는 메뉴가 나타나게 된다. 기관정보를 선택하면, 현재 연구자가 재직 중인 기관의 이름, 소속부서, 소속기관의 종류, 직위 등의 정보를 볼 수 있고 마찬가지로 소속학·협회 정보를 선택하면 학·협회명, 학·협회직위, 활동기간, 연락처 등의 정보를 볼 수 있다.

1.4 강의정보

강의정보 브라우징은 연구자정보의 하위 클래스 중 직무정보>강의정보 클래스를 통해 구현될 수 있다. 연구자가 대학교원인 경우, 직무클래스 중 강의정보 클래스의 속성을 이용해 연구자가 현재 강의 중인 교과목에 대한 정보를 브라우징할 수 있게 제안하였다. 왼쪽 연구자 속성 선택 프레임에서 강의정보를 보고 싶을 경우, 해당 메뉴를 클릭하면 현재 해당학기에 개설된 강의리스트가 나오고, 같은 강의를 진행하고 있는 해당학교의 리스트를 볼 수 있다. 동일한 강의정보를 제공하고 있는 학교의 웹페이지를 링크시켰으며, 하이퍼링크 표시를 이용하여 해당학교의 링크페이지 유무를 한눈에 구별할 수 있도록 표현하였다(그림 13 참조).



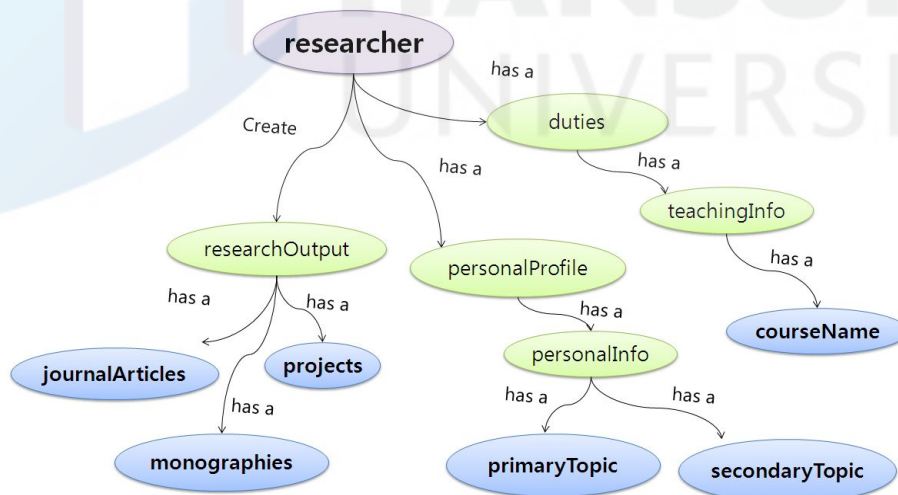
<그림 13> 연구자의 강의정보 브라우징 결과 화면

1.5 관련인물

관련인물 브라우징은 연구자정보의 인적정보 클래스와 소속정보 클래스를 통해 제공될 수 있다. 연구자의 학력정보를 통해 출신학교 정보를 찾아 동기 선·후배 등의 인물을 볼 수 있다. 또한 재직기관을 통해 현재 재직 중인 기관에 같이 근무하는 인물들을 볼 수 있도록 제안하였다. 마찬가지로 소속학·협회를 현재의 같은 학·협회에서 활동 중인 인물들의 리스트를 보여주고 마우스를 인물의 이름으로 이동하면 연구자의 이름, 소속 등의 간략한 정보가 임시팝업창 형태로 나타나게 하였다. 각 인물의 이름을 선택하였을 경우, 다시 해당 인물로 탐색할 수 있도록 링크하였다.

2. 주제어 검색

주제어 검색은 Topic 접근점을 활성화한 다음 연구자의 연구물의 주제 키워드나 연구분야 키워드를 이용해 연구자의 연구저작물과 동일한 연구분야의 연구자 리스트를 브라우징할 수 있도록 제안하였다(그림 14 참조).

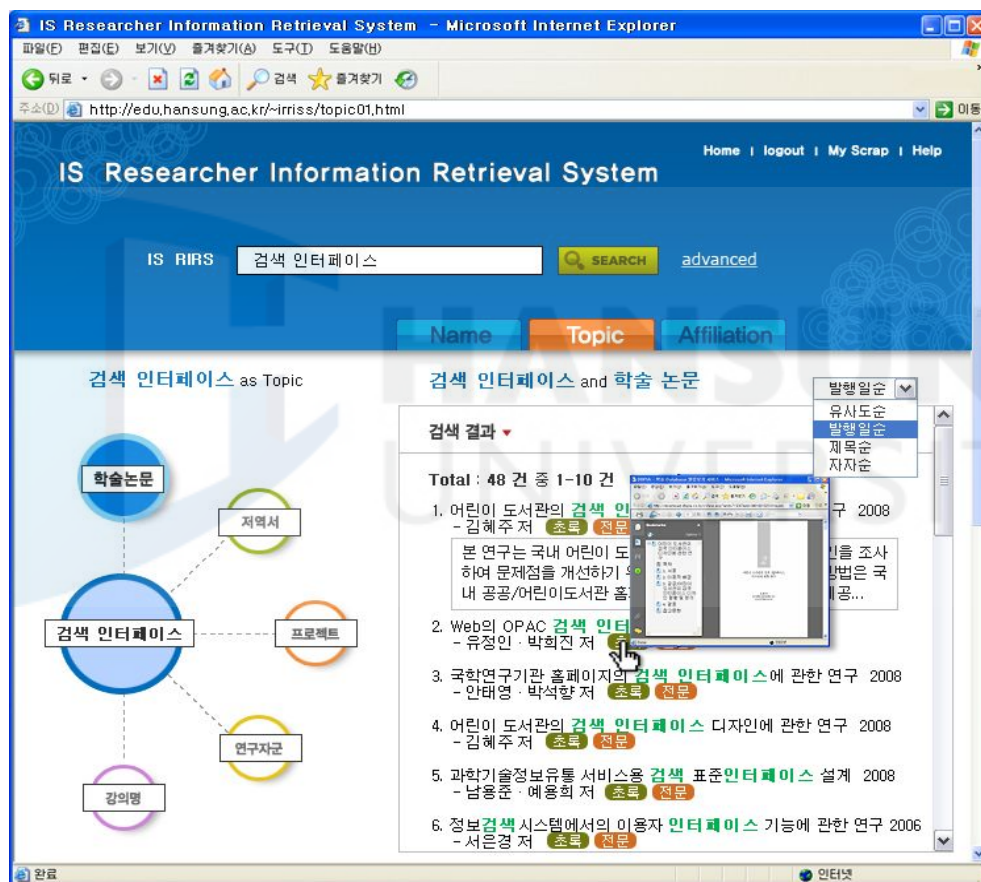


<그림 14> 주제 검색 관계정의

2.1 학술논문

<그림 15>는 주제어 검색의 예로써 검색어로 ‘검색 인터페이스’를 입력하고 학술정보를 선택한 결과를 보여주는 화면이다.

검색된 논문은 초록과 전문링크 기능을 가지며, 10개씩 브라우징할 수 있도록 설계되었다. 또한, 해당 논문을 기본적으로 유사도순으로 정렬하였으며, 이용자가 발행일순, 제목순, 저자순으로 재정렬할 수 있게 표현하였다. 학술논문명의 앞에 체크박스를 두어 원하는 논문정보를 My Scrap에 저장할 수 있다(그림 15 참조).



<그림 15> 주제어 검색을 이용한 학술논문 브라우징 결과

2.2 저·역서

저·역서 브라우징은 이용자가 입력한 키워드인 ‘검색 인터페이스’가 가질 수 있는 속성정보 중 저·역서를 선택하였을 경우, 오른쪽 프레임에 IR RIRS에 구축되어 있는 저·역서 중 ‘검색 인터페이스’를 키워드로 가지고 있는 자료들을 볼 수 있다.

검색된 저·역서는 초록과 전문링크 기능을 가지며, 10개씩 브라우징할 수 있도록 제안하였고 해당 저·역서를 기본적으로 유사도순으로 정렬하였으며 이용자가 발행일순, 제목순, 저자순으로 재정렬할 수 있게 설계하였다. 학술논문과 마찬가지로 서명 앞에 체크박스를 두어 원하는 저·역서정보를 My Scrap에 저장할 수 있다.

2.3 프로젝트

프로젝트 브라우징은 이용자가 입력한 키워드인 ‘검색 인터페이스’가 가질 수 있는 속성정보 중 프로젝트를 클릭하였을 경우, 오른쪽 프레임에 IR RIRS에 구축되어 있는 프로젝트 발간물 중 ‘검색 인터페이스’를 키워드로 가지고 있는 자료들을 볼 수 있다.

검색된 프로젝트 발간물은 초록과 전문링크 기능을 가지며, 10개씩 브라우징할 수 있도록 제안하였다. 또한, 해당 발간물을 기본적으로 유사도순으로 정렬하였으며 이용자가 발행일순, 제목순, 저자순으로 재정렬할 수 있다. 학술논문과 마찬가지로 서명 앞에 체크박스를 두어 원하는 프로젝트 발간물을 My Scrap에 저장할 수 있다.

2.4 연구자군

연구자군 브라우징은 이용자가 입력한 키워드인 ‘검색 인터페이스’가 가질 수 있는 속성정보 중 연구자군을 클릭했을 경우, 오른쪽 프레임에 IR RIRS에 구축되어 연구자들 중에 ‘검색 인터페이스’를 연구분야로 가지고 있는 연구자들을 브라우징할 수 있다. 검색된 연구자들은 이름, 소속, 강의명을 가지고 있고, 이름순, 소속기관명순 등으로 재정렬할 수 있으며 연구자의 이름을 클릭하면, 해당연구자의 인적정보를 볼 수 있도록 링크하였다.

2.5 강의명

주제어 검색에서 ‘검색 인터페이스’로 검색한 후, ‘강의명’이라는 속성을 선택하였을 경우, 연구자가 현재 강의하고 있는 강의과목명을 검색 접근점으로 이용해 IR RIRS에 구축되어 있는 모든 연구자중 ‘검색 인터페이스’를 강의명 및 키워드로 가지고 있는 강의를 브라우징할 수 있다. 또한 현재 학기에 진행하고 있는 연구자의 이름과 소속, 강의명, 강의소개, 강의계획서, 강의 참고문헌 등의 결과물을 제공하고 있다(그림 16 참조). 그리고 검색된 결과물을 강의자순, 강의명순, 학교순으로 재정렬할 수 있도록 하였다.



<그림 16> 주제어검색을 이용한 강의정보 검색결과

3. 소속 검색

소속 검색은 affiliation을 활성화한 다음 연구자가 재직 중인 기관명과 현재 소속중인 학·협회명을 이용해 해당 기관의 기관소개, 발간물, 소속연구자군, 임원진 정보 등의 결과물을 브라우징할 수 있도록 제안하였다(그림 17 참조). 따라서 상단의 세 개의 검색접근 메뉴 중 기관을 선택하고 검색창에 원하는 기관명을 입력할 수 있도록 설계하였다. 기관소개, 발간물, 임원진 정보의 경우 웹에 있는 해당정보들을 불러와 브라우징 결과로 보여줄 수 있기 때문에 관계추론 그림에서는 제외하였고, 연구자군의 경우 <그림 17>처럼 인적정보 클래스와 소속정보 클래스를 통해 구현될 수 있다.



<그림 17> 소속검색 관계정의

<그림 18>은 소속검색중 학회명을 검색한 예로써 검색어로 ‘한국정보관리학회’를 입력하였을 경우 브라우징하는 과정을 나타낸 화면이다. ‘한국정보관리학회’로 검색하였을 경우 하단 왼쪽 프레임에 ‘한국정보관리학회’가 가지고 있는 속성 중 기관소개, 소속연구자, 임원진 정보를 볼 수 있게 하였고, 그 중 연구자군을 선택하였을 경우, 오른쪽 내용 프레임에 현재 한국정보관리학회에 소속되어 있는 연구자들의 이름, 소속, 직위, 협회직위 리스트가 출력되도록 하였다. 정렬박스를 통해 연구자의 이름순, 소속기관 순으로 재정렬 할 수 있도록 하였고, 연구자 중 IR RIRS에 구축되어 있는 인물은 이름을 선택하였을 경우, 해당 인물의 이름, 소속, 생년월일, 이메일, 연락처 등의 간략한 개인정보를 볼 수 있다.

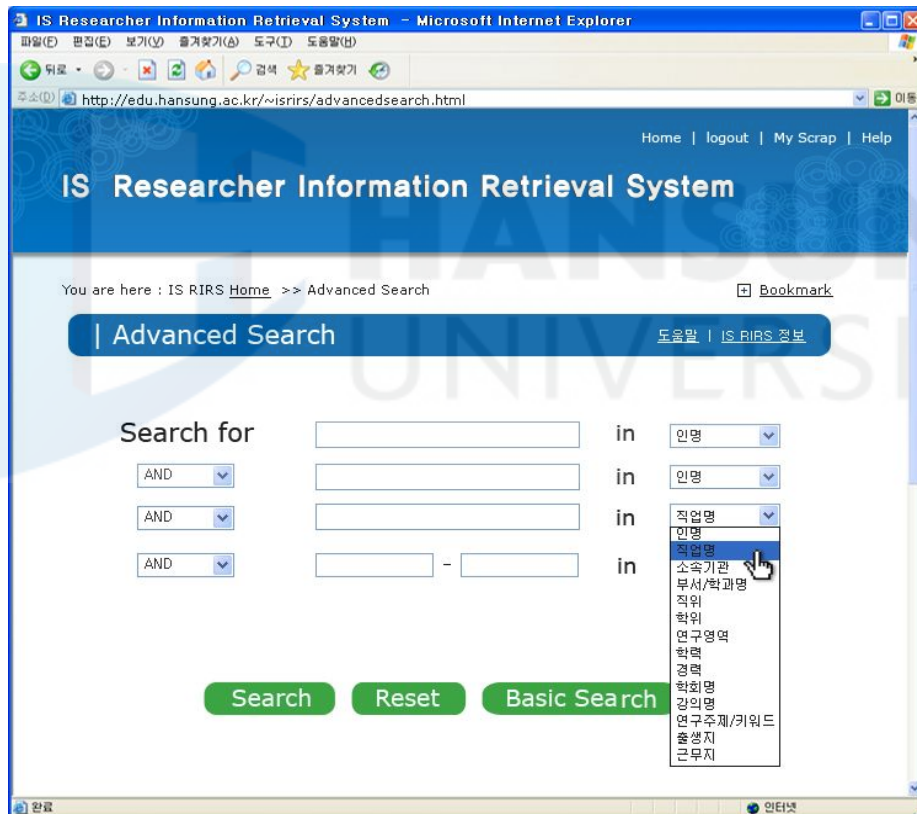


<그림 18> 연구자군 브라우징 결과 화면

또한, 소속기관명으로도 검색이 가능하다. 예를 들어 ‘한성대’로 검색하였을 경우 하단 왼쪽 프레임에 ‘한성대’가 가지고 있는 하위정보 중 한성대에 대한 소개인 기관소개, 한성대에서 발행한 발간물, 한성대 소속연구자, 한성대 임원진 정보를 볼 수 있게 하였다. 오른쪽 내용 프레임에 현재 한성대학교에 재직 중인 연구자들의 이름, 소속, 직위, 협회 직위 리스트가 출력되도록 설계하였다. 정렬박스를 통해 연구자의 이름순, 출신학교순으로 재정렬 할 수 있고, 연구자 중 IR RIRS에 구축되어 있는 인물은 이름을 클릭하면, 해당 인물의 이름, 소속, 생년월일, 이메일, 연락처 등의 간략한 개인정보를 볼 수 있다.

4. 상세 검색 인터페이스

상세검색 인터페이스는 앞서 브라우저를 통한 검색이 아닌 속성간의 조합을 이용하여 보다 특징적인 정보를 검색할 수 있도록 해준다. 즉, 인명, 직업, 소속기관, 부서/학과명, 직위, 학위, 연구영역, 학력, 경력, 학회명, 강의명, 연구주제/키워드, 출생지, 근무지를 자유롭게 조합하여 원하는 연구자정보를 검색할 수 있게 하였다. 또한, 출생년도, 재직기간, 활동기간, 교육기간을 입력하여 검색결과를 제한할 수 있게 하였다. 상세 검색 결과 역시 My scrap 기능에서 이용자가 보관하고 싶은 정보를 스크랩할 수 있다. 또한, Bookmark 기능과 도움말 기능을 제공하고 있다<그림 19 참조>.



<그림 19> 상세검색 화면

V. 결론

현재 ‘정보화시대’라는 말에 걸맞게 하루에도 수많은 정보가 생성되어 이용자가 활용하고 있으나, 지나치게 많은 정보가 생성되고 검색됨에 따라 정보선택에 어려움을 겪는 경우가 많다. 따라서 보다 효율적이고 적합한 정보를 어떻게 얻을 것인가에 대한 논의가 끊임없이 제기되고 있는 상황이다. 최근에는 이러한 문제의 대안으로 시맨틱 웹 기술이 정보검색 시스템에 응용되고 있다. 시맨틱 웹은 검색환경이 단순히 정보를 검색하고 검색된 정보를 나열하여 주는 것이 아니라, 이용자가 원하는 정보를 구조적으로 그룹화해주고, 정보의 적합성과 관련성을 컴퓨터가 판단하여 제공해주는 즉, 이용자의 정보검색 목적을 성취할 수 있도록 해주는 웹 환경을 말한다.

따라서, 본 연구는 여러 분야에서 연구를 진행하고 있는 연구자에게 연구의 근간이 되는 관련 주제의 연구자정보를 보다 체계적이며 포괄적으로 제공할 수 있도록 하기 위해 시맨틱 웹 기반의 연구자정보 검색 인터페이스를 설계하고자 하였다. 이에 따라 먼저 시맨틱 웹과 온톨로지에 대한 개념과 그 특징을 알아보았고 시맨틱 웹이 구현될 수 있도록 연구자정보 도메인에 대한 온톨로지를 설계하였다. 이를 위해 먼저 국내외 인물 및 연구자정보 데이터베이스의 기술요소들을 조사하여 비교하여 가장 필수적이며 범용적인 기술요소를 분석하였고 이러한 도메인에 대한 요구분석 결과를 토대로 하여 연구자정보 온톨로지 분석을 EOE 방법론⁶⁵⁾에 따라 단계적으로 실행하였다. 마지막으로 구조화된 온톨로지를 기반으로 연구자정보를 검색할 수 있는 검색 인터페이스를 제안하였다. 특히 본 연구는 인명, 주제어, 소속기관명을 이용하여 연구자와 관련되는 모든 정보를 브라우징할 수 있는 인터페이스와 정의된 속성의 조합을 이용하여 특정 정보만을 검색하게 하는 인터페이스를 제안하였다.

65) 한국전산원. 2004. 전계서.

이러한 연구를 통하여 다음과 같은 연구결과를 얻었다:

첫째, 연구자정보에 대한 요구분석을 한 결과 연구자정보에는 개개인의 학력, 경력 그리고 개인적 정보를 기술해주는 인적정보, 연구자의 연구성과물을 기술할 수 있는 연구업적정보, 연구자가 현재 수행하고 있는 업무를 설명해주는 직무정보, 마지막으로 연구자가 현재 또는 과거에 소속한 기관이나 학·협회에 대한 정보를 기술하는 소속정보가 필요하다는 것을 알 수 있었다.

둘째, 핵심클래스는 행위클래스를 가지고 각 클래스간의 관계를 정의하는 것이 검색에 도움을 줄 수 있는 것이라고 판명하였다. 이에 따라 인적정보 핵심클래스는 개인정보, 학력정보, 연락정보, 연구정보, 경력정보, 훈·포상정보의 6개의 행위클래스를 명명하였고 이에 대한 29개의 속성들을 선정하였다. 다음으로 연구업적정보 핵심클래스에는 학술논문, 저역서, 프로젝트 행위클래스를 두었고 이에 대한 24개의 속성을 선정하였다, 직무정보 핵심클래스에는 강의정보와 직무정보라는 2개의 행위클래스를 두고 10개의 속성을 선정하였다. 소속정보 핵심클래스도 소속기관과 소속학·협회라는 2개의 행위클래스를 두고 이에 대한 10개의 속성을 선정하였다.

셋째, 온톨로지 분석을 통해 추론된 관계를 바탕으로 검색 인터페이스를 설계하였다. 검색 접근점을 인명, 주제어, 소속으로 설정하여 기존의 키워드 기반의 검색결과와 차별화된 검색 인터페이스를 제안하였다. 우선, 인명 검색을 통해 연구자의 인적정보, 연구정보, 강의정보, 소속정보, 관련인물 브라우징이 가능하도록 표현하였다. 주제어 검색은 키워드를 입력하면 이용자가 입력한 키워드에 대해 학술논문, 저역서, 프로젝트, 연구자군, 강의명 등의 연관된 연결고리를 통해 연속적인 정보탐색이 가능하다. 소속 검색의 경우, 이용자가 기관명을 입력하였을 때 해당 기관에 대한 관련정보가 나타나게 된다. 기관소개, 발간물, 연구자군, 임원진 정보 등을 브라우징할 수 있다.

따라서 본 연구는 다음과 같은 의의를 갖는다. 첫째, 연구자정보 도메인에 대한 온톨로지 분석을 통해 상세 속성을 기술한다면 ‘사회관계정보’가 추출될 수 있음을 알 수 있었다. 즉 얼마나 다른 연구자에 대하여 인용될

수 있는지를 알 수 있는 인용정보, 연구 분야로서 밀접한 다른 연구자 그룹, 같은 주제를 강의하는 또는 연구하는 연구자 그룹, 소속된 기관에 대한 기관소개 및 임원진 정보를 즉각적이고 직관적으로 브라우징할 수 있었다.

둘째, 기술된 속성 이외에 기존의 웹정보원도 관계정의를 수행함에 따라 검색결과를 보다 다양하게 브라우징할 수 있었다. 정보자원의 위치는 연속적으로 변화하기 때문에 기존의 키워드 방식의 검색은 잘못된 주소를 결과값으로 제시하는 경우가 종종 있다. 그러나 시맨틱 웹 기반의 온톨로지 기법으로 데이터베이스를 구축한다면, DOI를 이용해 학술논문을 검색하는 것처럼 보다 동적인 데이터 검색이 가능하게 될 것이다.

이와 같이 모든 분야를 온톨로지 기법으로 구축하는 것은 상당한 많은 시간과 노력이 소요된다. 그러나 연구자정보와 같은 우선적으로 많이 이용되고 적합한 분야를 온톨로지화해서 점차 추진해나간다면, 보다 효율적으로 검색할 수 있을 것이다. 즉, 구조화되고 계층화된 데이터들을 관계정의를 통해 추론이 가능하다면 이용자에게 다양하고 동적인 검색서비스를 제공할 수 있을 것이다.



참고문헌

[국내 문헌]

- 권범중, 곽승진. 2006. 과학기술분야 인력정보관리시스템 데이터기반 전문가 선정 방법 연구. 『제13회 한국정보관리학회 학술대회 논문집』, 13: 223-228.
- 김수경, 안기홍. 2005. 시맨틱 웹 기반의 비교구매 에이전트를 위한 동적 웹 온톨로지에 대한 연구. 『한국지능정보시스템학회논문지』, 11(2): 31-45.
- 김수경, 안기홍. 2007. 지능형 이미지 검색시스템을 위한 추론기반 웹 온톨로지 구축. 『한국정보관리학회지』, 24(3): 119-147.
- 김은주. 2003. 시맨틱 웹. 『TTA 저널/정보통신표준화소식』, 한국정보통신기술협회. 87: 134-140.
- 김평, 이승우, 강인수, 정한민, 이정연, 성원경. 2008. OntoFrame 기반 학술정보 분석 서비스. 『정보과학회논문지: 소프트웨어 및 응용』, 35(7): 431-441.
- 김학래, 김흥기. 2007. 시맨틱 웹/온톨로지 기술을 이용한 개인용 전자문서 검색시스템. 『한국전자거래학회지』, 12(1): 135-149.
- 김현희, 안태경. 2003. 온톨로지를 이용한 인터넷웹검색에 관한 실험적 연구. 『한국정보관리학회지』, 20(1): 417-455.
- 남지순. 2008. 정보검색엔진을 위한 <인물>관련 온톨로지 구축연구 사례들에 대한 소개 및 논의. 『프랑스어문교육집』, 27: 201-229.
- 문정수. 2006. 『토픽맵 기반의 인터넷 포털의 서비스 분류 및 연관체계 관리 방법에 관한 연구』, 석사 학위 논문. 한국과학기술원.
- 박선경. 2008. 『문헌정보학 논문을위한 온톨로지 구축방법에 관한 연구』. 석사학위논문, 숙명여자대학교 대학원.
- 박윤희. 2008. 『온톨로지 인스턴스 생성을 위한 인물정보 추출』. 석사학위논문, 경북대학교 대학원.
- 박준식, 박경호. 2001. 한국 웹 기반 인물정보원의 분석적 연구. 『한국문헌

- 정보학회지』, 35(1): 187-203.
- 박창섭. 2007. 의미적 연관성을 이용한 멀티미디어 정보검색. 『인터넷정보학회논문지』, 8(5): 67-79.
- 박희영. 2006. 『더블린코어 메타데이터를 이용한 서지적 관계 구현에 관한 연구』. 석사학위논문, 연세대학교 대학원.
- 서은경. 1999. 정보검색 시스템에서의 이용자 인터페이스 기능에 대한 분석적 고찰. 『정보관리학회지』, 16(4): 125-158.
- 서 휘. 2006. 온톨로지 자동구축을 위한 OWL의 어휘와 구문 사용방법에 대한 이론적 연구. 『한국도서관·정보학회지』, 37(2): 193.
- 유수현, 최호남, 최희운. 2007. 과학기술분야 해외 인력정보서비스 평가. 『제 14회 한국정보관리학회 학술대회 논문집』, 14: 105.
- 이미경, 정한민, 성원경. 2008. 웹 서비스 기반의 추론 서비스에 관한 연구. 『한국정보과학회 학술발표 논문집』, 35(2): 200-203.
- 이원경. 2008. 『한국화 검색 인터페이스 설계에 관한 연구』. 석사학위논문, 한성대학교 대학원.
- 이정희, 김희섭. 2007. 대학 전자기록물을 위한 온톨로지 기반 검색시스템 설계 및 구현. 『정보관리학회지』, 24(3): 343-362.
- 정현숙, 신영숙. 2008. 온톨로지 기반 강의계획서 관리 시스템 설계 및 구현. 『한국정보기술학회지』, 6(5): 108-114.
- 지정규. 2003. 국가학술연구인력 데이터베이스 구축에 관한 연구. 『한국컴퓨터정보학회지』, 10(1): 16-23.
- 차승준, 유정연, 엄동명, 고유미, 이규철. 2007. 전통 약재 정보의 효율적인 활용을 위한 온톨로지 시스템 개발. 『한국정보과학회 학술발표 논문집』, 34(2): 195-199.
- 최석두, 한상길. 2008. 『온톨로지 알고리즘 I』. 서울: 한울.
- 최석두, 김이겸. 2008. 『온톨로지 알고리즘 II』. 서울: 한울.
- 최원중. 2004. 『온톨로지 계층 정보를 이용한 웹 서비스 발견 알고리즘』. 석사학위 논문, 한양대학교 대학원.
- 최중민. 2003. 시맨틱 웹의 개요와 연구동향. 『정보과학회지』, 21(3): 4.

- 한국전산원. 2004. 『웹 온톨로지 개발 지침 연구』. 서울: 한국전산원.
- 한용진, 박세영, 이영화, 김권양. 2008. 이벤트 온톨로지 기반의 의미 정보 검색. 『정보과학회논문지』, 14(1): 96-100.
- 현승순. 2006. 차세대 지능형 시맨틱 웹 & 온톨로지. 『주간기술동향』, 정보통신연구진흥원. 1265: 25.
- Chu, Heting. 2006. 『디지털 시대의 정보문헌과 검색』. 장혜란 옮김. 서울: 한국도서관협회.
- Kanzaki Masahide. 2008. 『시맨틱 웹을 위한 RDF/OWL 입문』. 황석형, 양해술 옮김. 서울: 홍릉과학출판사.

[국외문헌]

- Bechhofer, S., L. Carr, C. Goble, and W. Hall. 2001. "Conceptual open hypermedia: the semantic web" *In Proceedings of the WWW 2001, Semantic Web Workshop*, Hongkong, pp. 44-50.
- Decker, S., M. Erdmann, D. Fensel, and R. Studer. 1999. "Ontobroker: Ontology based access to distributed and semi-structured information." *Proceedings of IFIP TC2/WG2.6 8th Working Conference on Database Semantics(DS-8)*, pp. 351-369.
- Fonseca, F. 2007. "The double role of ontologies in information science research." *Journal of the American Society for Information Science and Technology*, 58(6): 786-793.
- Gruber, T. R. 1993. "A translation approach to portable ontology Specification", *Knowledge Acquisition*, 5: 199-220.
- Guarino, N. 1998. "Formal ontology and information systems, In N. Guarino(ed.).", *Formal Ontology in Information Systems*. ISO Press.
- Noy, N. F. and D. L. McGuinness. 2002. 『Ontology development 101 : A guide to greating your first ontology』 Stanford University press.
- Srinivasan, R. and J. Huang. 2005. "Fluid ontologies for digital

museums." *International Journal on Digital Libraries*, 5(3): 193-204.

Weinberger, H., D. Te'eni, and A. J. Frank. 2008. "Ontology-based evaluation of organizational memory." *Journal of the American Society for Information Science and Technology*, 59(9): 1454-1468.

Yildiz. B. and S. Miksch. 2007. "ontoX-A Method for ontology-driven Information Extraction." *ICCSA*, 3: 660-673.

[웹사이트]

로마켓 법률 및 주요인물 검색 서비스. [cited. 2008. 9. 30].

<<http://www.lawmarket.co.kr/>>.

미국의 과학재단 홈페이지 SESTAT. [cited. 2008. 9. 30].

<<http://sestat.nsf.gov/SESTAT/NewHomeLogin.html>>.

일본의 문부과학성 소속 연구자정보 검색 서비스. [cited. 2008. 9. 30]

<<http://read.jst.go.jp/>>.

조선닷컴 검색플러스 인물검색 서비스. [cited. 2008. 9. 30].

<http://image.chosun.com/sv/people/200710/people_db_main.html>.

조인스닷컴 인물정보 검색서비스. [cited. 2008. 9. 30].

<<http://people.joins.com/>>.

한국과학기술정보연구원 국가R&D참여인력정보서비스. [cited. 2008. 9. 30].

<<http://hurims.ntis.go.kr/index.jsp>>.

한국학술진흥재단 KRI 시스템. [cited. 2008. 9. 30].

<<http://www.kri.go.kr/index.jsp>>.

COS Scholar Universe 검색 서비스. [cited. 2008. 11. 4].

<http://csaweb115v.csa.com/ids70/quick_search.php?SID=ho1bkqgdmi4ct17s0ta8momd74>.

Heflin, J. 2000. <Personal Ontology> in SHOE. [cited. 2008. 7. 3].

<<http://www.cs.umd.edu/projects/plus/SHOE/onts/personal1.0.html>>.

The Suggested Upper Merged Ontology (SUMO). [cited. 2008. 6. 24].
<<http://www.ontologyportal.org/>>.

Wennerberg, P. O. 2005. "Ontology based knowledge discovery in social networks." Final Report JRC. [cited. 2008. 7. 3].
<<http://langtech.jrc.it>>.



ABSTRACT

Designing the Researcher Information Retrieval Interface Based Semantic Web

Park, Mi Hyang

Major in Library & Information Science

Dept. of Library & Information Science

Graduate School, Hansung University

As the continuous development of the internet search engine, many people are navigating a lot of information by using the internet search engine. However, it is one of the most important things to find exactly the information which many users want among heavily overflowing information, as you know, in the information age. Recently, as an alternative to this problem, however, there has applied to the Semantic Web Technology in the information search system. The Semantic Web Technology means not only that a web search environment can find the needed information and arrange it, but also that the computer can think of the adequacy and relevance of the searched information.

This study for many researchers who study in diverse areas is suggesting a search interface which can easily approach the needed information. In other words, for many researchers, there is a purpose to design this search interface, which systematically and comprehensively can offer the information of the researchers studying in same areas. Therefore, this study is suggesting the search interface that can approach the needed information more dynamically, meaningful connected and easily. To understand these things, above all, this study

will find out the Semantic Web Technology and a Ontology. Then, this study will find out a lot of former researches using the Ontology and the information search systems of the domestic and foreign researchers.

Frist, this study is designing the Ontology about the information of the researchers, and to do this, this study is choosing the needed properties in order to describe the information of the researchers, and then defining the structural relationship and meaning among properties according to the steps presented by the EOE methodology. Second, this study is selecting an OWL language that defines the Ontology, and describing a protégé program by using it. Third, this study is suggesting the search interface which can search for the information of the researchers by using the structured Ontology-based information.

Through the search interface of the information of the many researchers suggested in this study, I could know that 'the information in the social relationship' can be extracted. Also, I could immediately browse the information of the other researchers' groups which are studying similar topics, or which are studying or lecturing the same topics, and could find the introduction of the institution and executives belonging to many institutions. In addition, I could browse search results more diversely by performing the definition of the relationship about the best source of searched information, and could search for the correct ones among many data. As a result, the relationship-based web search using the Semantic Web will be able to contribute to design a flexible search environment.