Designing Researcher Information Retrieval Interface based on Ontological Analysis*

온톨로지 기반의 연구자정보 검색 인터페이스 설계

Eun-Gyoung Seo**
Mi-Hyang Park***

ABSTRACT

Recently, semantic search techniques which are based on information space as consisting of non-ambiguous, non-redundant, formal pieces of ontological knowledge have been developed so that users do exploit large knowledge bases. The purpose of the study is to design more user-friendly and smarter retrieval interface based on ontological analysis, which can provide more precise information by reducing semantic ambiguity or more rich linked information based on well-defined relationships. Therefore, this study, first of all, focuses on ontological analysis on researcher information as selecting descriptive elements, defining classes and properties of descriptive elements, and identifying relationships between the properties and their restriction between relationships. Next, the study designs the prototypical retrieval interface based on ontology-based representation, which supports to semantic searching and browsing regarding researchers as a full-fledged domain. On the proposed retrieval interface, users can search various facts for researcher information such as research outputs or the personal information, or carrier history and browse the social connection of the researchers such as researcher group that is lecturing or researching on the same subject or involving in the same intellectual communication.

초 록

최근 특정 도메인의 개념에 대한 정확한 정의, 계층적 관계나 추론규칙을 정형화된 어휘로 기술된 온톨로지 지식을 기반하는 시맨틱 탐색기법이 정보검색시스템에 응용됨에 따라 이용자는 보다 쉽게 지식기반을 항해할 수 있다. 본 연구의 목적은 의미적 애매성을 줄여주고, 의미적 관계성을 제시해 줌으로서 이용자가 보다 정확하고 쉽게 정보를 검색할 수 있게 해주는 온톨로지 기반의 정보검색 인터페이스를 개발하는 것이다. 이를 위해서 본 연구는 연구자정보 기술요소들의 클래스의 계층관계, 데이터의 계층관계 및 속성, 관계 연계성을 정의하였고 이를 기반으로 하여 구조화된 온톨로지를 기반으로 기술된 연구자정보를 검색할 수 있는 검색 인터페이스를 제안하였다. 이용자는 연구자들의 연구결과물, 개인적 데이터, 학력 또는 경력과 같은 개인정보를 검색할 수 있을 뿐만 아니라, 연구자의 사회적 연결 즉 같은 주제의 연구나 강의를 수행하는 연구자들 또는 같은 분야의 지적 커뮤니케이션에 속하는 연구자들을 브라우징 할 수 있다.

Keywords: researcher information, retrieval interface, ontology, ontological analysis, semantic search technique
연구자정보, 정보검색, 온톨로지, 온톨로지 분석, 시맨틱 탐색기법

---

* This research was supported by Hansung University in 2008.
** Professor, Division of Knowledge & Information, Hansung University(egseo@hansung.ac.kr)
*** Graduate School of Hansung University(redmuriann@nate.com)
■ Received : 18 May 2009 ■ Revised : 22 May 2009 ■ Accepted : 3 June 2009
■ Journal of the Korean Society for Information Management, 26(2): 173-194, 2009.
〔DOI:10.3743/KOSIM. 2009.26.2.173〕

1. Introduction

As the environment for information retrieval is shifting to the intelligent web, users can search or browse the web information resources through the more convenient methods such as the way-finding retrieval based on space metaphor, the collaborative retrieval based on human information interaction(HII), and the ambient retrieval based on pull and push(Morbille 2005). Accordingly, many different techniques such as cognitive map, information visualization, filtering, pull and push, information interaction by collaborating, and the search interface based on information scent are applied to information retrieval. However, it still happens that a computer cannot semantically recognize the contents and the context of information needs which a user presents as a search term during searching, so that the user should change the query statements or browse the other information resources repeatedly until the wanted information is found(Chu 2006).

Recently, semantic search techniques which are based on information space as consisting of non-ambiguous, non-redundant, formal pieces of ontological knowledge have been developed so that users do exploit large knowledge bases(Vallet et al. 2005). The use of ontologies to overcome the limitations of key-word based search has been put forward as one of the motivations of the semantic web. Ontology, an explicit specification of conceptualizations(Gruber 1993), is a description of things, of the relationships between things and the properties of things, and of inference rules in a way that computers can understand. That is, the ontology is a mesh or a vision of information linked up in such a way as to be easily processed by computers within the specific domain.

Moreover, ontology improves knowledge management and retrieval ability in huge information space. Ontologies used within the information systems can make domain knowledge explicit so that meaningful information or more relevant related documents be retrieved efficiently. Ontology approach has brought improvements over keyword-based search through e.g. query expansion based on class hierarchies and rules on relationships or multi-level searching and browsing(Yildis & Miksch 2005). It allows provide more precise information by reducing semantic ambiguity of queries or/ and results, and browse multifacetedly and precisely linked information by providing well-defined and inferred relationships.

This study is performing ontological analysis on researcher information in order to develop more user-friendly and smarter researcher information retrieval interface of than that of the existing researcher information systems. Recently the people's information retrieval systems supported by internet portal of information service provider(ISP) allow search and browse not only the detailed personal profile of the first level, but also the related people's information which we could not get from the printed material in the past, by using various methods. The people's information sites have been visited by who want to know multifaceted in-

formation about the person such as academic and career profile, personal information, related acquaintances. These sites are getting popular and utilized increasingly by many searchers, but users are not satisfied with these sites because of obsolescence information, primitive link, and unpredictable browse(Park 2008). Government institutes, also, have developed the researcher information service systems which are designed for easy retrieval of individual research achievements and project of researcher, but it does not provide inclusive browsing on relationships among the researchers or various related information regards to researchers.

In this paper we propose a retrieval interface based on ontology-based representation, which supports to semantic searching and browsing in a full-fledged domain such as researchers' information. The purpose of this paper is to develop a retrieval interface for researcher information based on ontology analysis, that is able not only to find out the research outputs and the personal/carrier history of one researcher, but also to browse the social connection of researchers, e.g. researcher group that is lecturing or researching on the same subject, group that is related in social activities, and group that is involved in the same intellectual communication. In order to accomplish this purpose, this study, first of all, focuses on ontological analysis on researcher information as selecting descriptive elements which could explain general researcher information, defining classes and properties of descriptive elements, and identifying relationships between the properties and their restriction between relationships. Next, the study designs the prototypical retrieval interface for researcher information with multiple access points. Specially, the study is exploring a design space of possible functionalities and interfaces in order to search and browse not only 'the precise factual information' but also 'the social relation information.' Therefore, researchers easily navigate from one researcher to others based on semantic link and dig out hidden information points for retrieval, and free from keyword searching.

2. Literature Review

In recent years, ontology research has been explored in terms of understating of what ontologies are. Fonseca(2007) examined that the role of ontologies in information science research. He indicated that ontology, in philosophy, is the basic description of things in the world, but in information science, refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality. He clarified a differentiation between ontologies of information systems and ontologies for information systems. Srinivasan and Huang(2005) discussed the concept of "fluid ontologies," a novel, dynamic structure for organizing and browsing knowledge.

Besides, there are several studies about ontology-driven information systems, which are developing various techniques for retrieval systems.

Yildiz and Miksch(2005) examined what the requirements of ontologies are that have to be reconciled in order to enhance their smooth integration within the information systems and later they proposed an extraction method that utilizes the content and predefined semantics of ontologies formulated in the OWL to perform the information extraction task(Yildiz and Miksch 2007). Vallet and his colleagues(2005) proposed a model for the exploitation of ontology-based knowledge bases to improve search over large document repositories. They specially developed an ontology-based scheme for the semi-automatic annotation of documents and semantic search techniques which combined with keyword-based search. Kim and Ahn(2007) proposed a suitable a method of building web ontology for characteristics of semantic web by comparing the existing ontology construction and inference-based web ontology construction. They established a web ontology-based intelligence image retrieval system, evaluated performance of the web ontology built for the study, and suggested an example of implementation of a semantic web application and utilization. Park(2007) proposed a multimedia retrieval system which is exploiting semantic relevancy of multimedia contents based on a domain ontology, which provides a wide range of multimedia contents having semantic relevancy to the input keyword, and which displays the results categorized by the semantic meaning and relevancy to the keyword derived from the ontology. Lee and Kim(2007) built an ontology-based retrieval system for the electronic records of universities and compared its performance with the existing keyword-based retrieval system. Kim and Kim (2007) developed a document management and retrieval tool which is named Ontalk based on a semi-automatic metadata generator and an ontology-based search engine.

In addition, there are several ontology studies enhancing usability within a certain domain. Suh and Yoo(2008) developed hotel ontology using currently available semantic web technologies such as RDF, OWL and SWRL and to show how it could be used to help travelers find hotels of their interest. Specially they analyzed available hotel-related ontologies, investigated typical terms which are used when searching for hotels in the Q&A communities, and developed Semantic Hotel Search System(SHSS). Chung and Shin(2008) designed a syllabus ontology for solving the problem of current text-based irregular syllabuses, which defined the standard structure of syllabus and semantic relationships between entities like courses, syllabuses, teachers, texts, and so on. They are found that a syllabus management system using the syllabus ontology can provide improved search facility to find semantically relevant courses and learning materials.

There are some researches which utilized ontology methods for developing researcher or research information retrieval system. Nam(2008) introduced People Ontology Systems such as JRC-SNO (Social Network Ontology) built by the European Union to structure information on human networks of terrorists; DERI-SWPP(People's Semantic Web Portal) to share information from participants in

Web Portal; SHOE-Personal Ontology presented by the University of Maryland on the general information of the individual; SOUPA-PersonOntology designed to represent the intelligent agent in Ubiquotos; FOAF-PeopleOntology proposed for the Machine-readable Web pages on individuals and companies; KW-PersonOntology built by DERI group for R & D Project Development; SUMO-HumanOntology proposed by Standard Upper Ontology Working Group in IEEE. Han and his colleagues(2008) developed IT-People Event Ontology by using people information extracted from web portals, which could provide constant information and time-temporal information on people. It were found that the system could provide the well-organized information which is suitable for users' demand. Kim et al. (2008) proposed the design methodology of ontology and service system for academic information based on OntoFrame as a service framework which includes ontology, reasoning engine, and triple store. OntoFrame served automatic recommendation of reviewers based on subject of project, reviewer's major, expertness of reviewer, relationship between applicant and reviewer and also served the analysis of researchers' accomplishments based on books, articles, patents, reports and work of art.

3. Developing Researcher Information Ontology

As Noy and McGuinness(2002) mentioned, an ontology is a formal explicit description of concepts in a domain of discourse, properties of each concept describing various features and attributes of the concept, and restrictions on properties. Therefore, the study, first, identifies features and users' needs of researcher information, by analyzing systems which provide multifaceted information about researchers or persons. On the basis of this result, the study is to determine classes representing concepts in the ontology, to arrange the classes in a taxonomic hierarchy, to define properties of classes and allowed values for these properties, and to describe the relationships among properties.

3.1 Analyzing Descriptive Elements of Researcher Information

The databases of people information are examined to identify what kinds of information are provided. Generally, people information falls into two broad categories: one is the researcher oriented man power databases which are developed by research institutes or universities and the other political/social figure oriented manpower databases which are developed by mass media. The former is made of information about personal research capability such as educational background, carrier history, research achievement, and academic activities. The latter provides personal information such as education, affiliation, personal hobbies or interests, and family or friend relationship(Jee 2003). The study examines information description elements which explain each researcher or person

in the people information systems such as 1) JOINS People Information and 2) People Chosun, which have an inclusive range of recorded persons and more than 300,000 items of data, 3) lawMarket's People of Korea which provides professional people information, 4) KRF's Korean Researcher Information System, 5) NTIS's National R&D Manpower Information Service, 6) SESTAT (Scientists and Engineers Statistical Data System) of America, 7) ReaD(Directory Database of Research and Development Activities) of Japan, and 8) ProQuest's COS Scholar Universe.

It is found that these eight systems provide about 50 descriptive elements as follows: 31 kinds of personal information elements, 12 kinds of research achievement elements, 2 kinds of social relationship elements, 2 kinds of linking information, and 3 kinds of open resources. As expected, people information systems(Joins, Chosun, and lawMarket) provide chiefly personal information such as name, birth data, address and phone number, educational background, carrier history, and affiliated organization. Especially, they lay an emphasis on the human networking information based on educational background, carrier history, and affiliated organization. However, the other systems which are researcher information systems provide not only personal information, but also detailed information about research product, research area, and interested topic, etc. ReaD and KRI have very similar elements and provide detailed information about research area, research carrier, and research product. SESTAT provides detailed information about family relation and employment information so that it could inform the social statistics to us(See <Table 1>).

The common elements which more than four systems provide are as follows: name, birth date, gender, birthplace, family relation, nationality, affiliation, occupation, specialty, education, license/certificate of qualification, carrier history, address and phone/fax number, email address, homepage, award records, academic papers, and publication. These all elements fall into 4 groups as follows: personal information about researcher's personal data, research achievement information about research products, job in duty information about researcher's other activities besides research product, and affiliation information about researcher's affiliated organization and academic association.

3.2 Defining the Classes and the Class Hierarchy

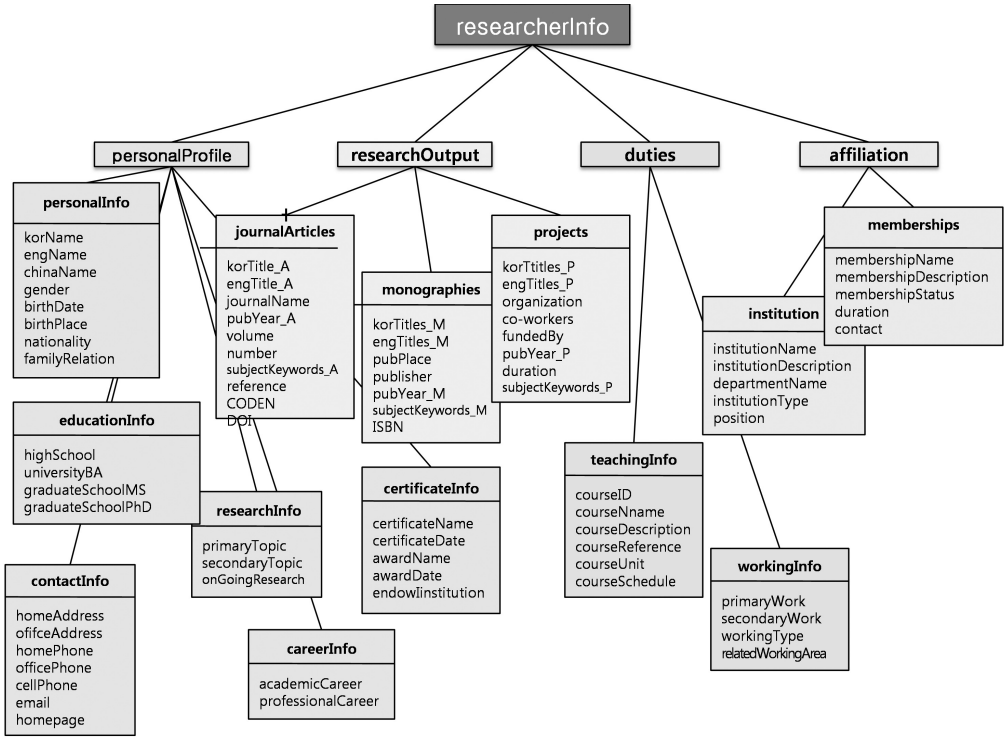
This is the step to determine the scope of the researcher ontology and to enumerate important elements in the researcher ontology. First of all, the four kinds of descriptive elements which are explained the previous chapter are selected as the main concepts of researcher ontology. Therefore, the four concepts describing personal information of researchers, research achievement information, duty and activity information, and affiliation information figure into the ontology and are represented as main classes.

Next step is to develop the class hierarchy and

〈Table 1〉 The Comparison of Descriptive Elements among 8 People Information Systems

	Provider / Attribute	Joins	Chosun	Law Market	KRI	NTIS	SESTAT	ReaD	COS
Personal Information	Name	o	o	o	o	o	o	o	o
	Name_China	o	o	o	o	o		o	
	Name_English	o	o	o	o	o	o	o	o
	Birth Date	o	o	o	o	o			
	Age						o		
	Gender	o	o	o	o		o	o	
	Birth Place	o	o	o					o
	Family	o	o	o			o		
	Nationality		o		o	o	o		
	Race						o		
	Disability						o		
	Department	o	o	o	o	o	o		
	Appointment Date				o				
	Retirement Date				o				
	Occupation	o	o	o	o			o	
	Specialty	o	o	o	o	o			o
	Employment Status						o		
	Position	o	o	o					
	Join Group	o	o					o	
	Association			o			o		o
	Education	o	o	o	o	o	o	o	
	Certification	o	o	o	o				
	Career	o	o	o	o		o	o	
	Home Address	o	o						
	Home Phone	o	o		o				
	Office Address	o	o	o	o		o	o	
	Office Phone	o	o	o	o		o	o	
	Mobile Phone	o	o		o				
	E-mail	o	o	o	o				o
	Homepage	o	o	o	o	o			o
	Researchers No.				o				
Research Information	Award	o	o	o	o				
	Thesis	o		o	o	o		o	o
	Research Report					o		o	o
	Keyword							o	
	Current Projects							o	
	Monography	o	o	o	o	o	o	o	
	Funding Agency				o				
	Project			o					
	IP			o	o				
	Exhibition Works				o				
	Academic Activities				o				
	Transfer				o				
Social Information	Friends	o	o	o					
	Membership Info.							o	o
Link Information	View web source					o			o
	Article Search					o			
Public information	Disclosed info.				o				
	Providing info.				o				
	Receive e-mail				o				

determine subclasses. The study follows the a bottom-up development process which starts with the most specific classes, the leaves of hierarchy, with subsequent grouping of the classes into more general concepts. Therefore, the study enumerates all attributes as concerned with the coverage of the domain, the usages of ontology, the types of questions answered in the system, and actual users of the ontology. And then study lists more than 80 descriptive elements and selects about 70 attributes. Next, the selected attributes are categorized by the four classes and the attributes which are belonged to each class also are grouped by common value. Finally the study determines the classification and names 13 subclasses. The class of Personal Profile has 6 subclasses such as personal information, educational information, contact information, research information, carrier information, and certificate information. The class of Research Output has 3 subclasses such as journal articles, monographies, and projects. The class of Duties has 2 subclasses such as teaching information and working information. The class of Affiliation also has 2 subclasses such as institution and memberships. The classes alone will not provide enough information to answer the competency questions. Next step is to determine which descriptive entity is belonged to subclasses and to give themselves



〈Figure 1〉 The Class Hierarchy and Its Properties

their own appropriate name. The < Figure 1> is the result of the class hierarchy and its properties.

3.3 Defining the Properties of Classes

Properties have different facets describing the value types, allowed values, the number of the values, and other features of the values(Noy and McGuiness 2002). Properties can be used to state relationships between individuals or from individuals to data values. The study determines whether the entity is best described as an atomic type or would be better represented with a class, which in turn could be decomposed into subcategories and identifies what are significant entities and relationships associated with other entity.

These formal features and relationship are described by OWL Lite language in which are special identifiers that are used to provide information concerning properties and their values. The study defines, first, the intrinsic features of properties using by the two types of properties. If the property is an atomic property, it is specified to a datatype Property. If this is not an atomic property but other concepts, then it is specified to a objectProperty. An object property is similar to a data type property, but the difference is that the range for the object property is a class instead of a data type(McGuiness and Harmelen 2004). Also the study determines the permitted cardinality for the range of the property and characteristics such as inverseOf, transitive Property, semmetricProverty, functionalProperty, and inverseFunctionalProperty.

A property also represents the connection between a class and something that the class has or affects. Therefore, a property in semantic relations(that is, relationship) is represented as a verb such as hasArticles, isEmployedBy, createID. etc. The <Figure 2> shows the properties of Personal Information described using by Protégé 3.3.1.

```
〈?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"/〉
```

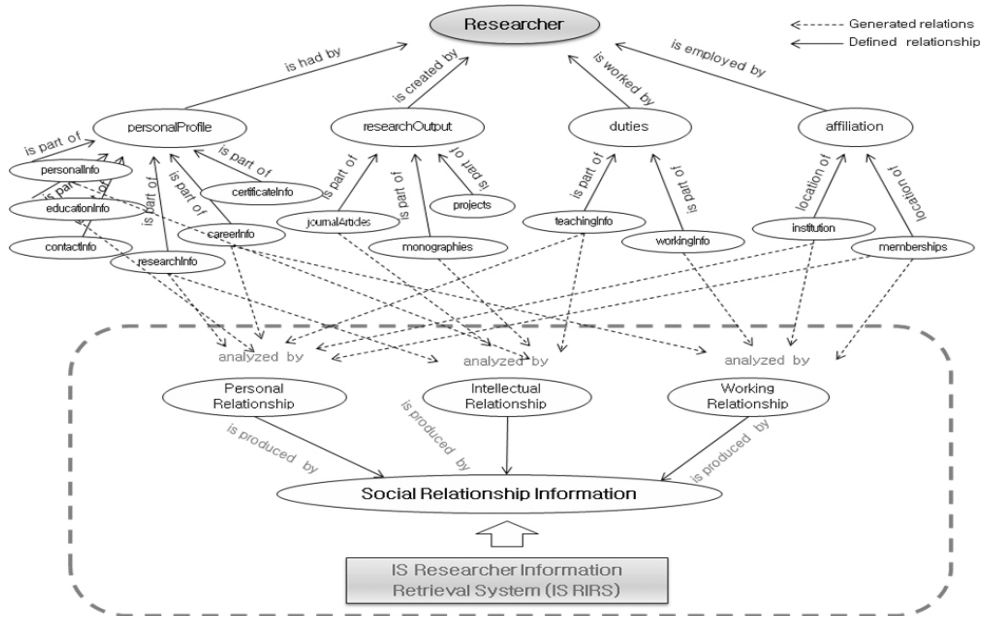
```
    〈/owl:Class〉
    〈owl:DatatypeProperty rdf:ID="email"〉
        〈rdfs:domain rdf:resource="#contactInfo"/〉
        〈rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/〉
    〈/owl:DatatypeProperty〉
    〈owl:DatatypeProperty rdf:ID="gender"〉
        〈rdf:type rdf:resource="http://www.w3.org/2002/07/owl#FunctionalProperty"/〉
        〈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〉
```

〈Figure 2〉 The Properties of Personal Information Described using by Protégé 3.3.1.

3.4 Developing Relationships

Ontological analysis makes relational inference possible. Researcher information provides not only multifaceted information about one's own self, but also intellectual or social relationships with other researchers. In other words, we can find some semantic relationships between researchers in the field of research area, educational/carrier history, academic publication, affiliation and etc. The <Figure 3> is a conceptual map which explains the semantic relationship found in researcher information ontology, personal networking, intellectual networing, and working networking.

The Personal Relationship provides social relationships of authors or human networking in terms of research subject, colleagues and course lectures. The Intellectual Relationship provides information about related authors, related publications, and references based on the researcher's academic papers and publications. The Working Relationship can be grasped from the information about position, title, affiliated organization, academic association, and participated project. The <Table 2> is a simple definition of semantic relationship. This research tries to analyze researcher information ontology



〈Figure 3〉 A Conceptual Map of Properties Relationship

〈Table 2〉 Definition of Semantic Relationship

Semantic Relationship	Relationship	Description
Personal Relationship	Research Relationship	to show the whole research relation between researchers by analyzing researcher's research area and specialty in major
		PrimaryTopic, secondaryTopic, onGoingTopic, korName
	Colleagues Relationship	To show the peer or colleagues relation by analyzing information about researcher's affiliation, attended school, membership
		highSchool, universityBA, graduateSchoolMS, graduateSchoolPhD, membershipName, institutionName, professionalCareer, academicCareer korName
	Teaching Relationship	To show the relation between researchers who are teaching the same subject, by analyzing the researcher's teaching subject
		courseName, courseDescription, primaryWork, secondaryWork, korName
Intellectual Relationship	Relationship of Authors	to show the relation between researchers who are working on the same subject, by analyzing the subject of academic paper and publication
		korTitle_A, korTitle_M, korTitle_P, korName,
	Relationship of Publication	to show the papers which have similar topic with the author's published paper
		subjectkeyword_A, subjectkeyword_M, subjectkeyword_P, korTitle_A, korTitle_M, korTitle_P, korName
	Relationship of Citation	to show the citation information about author or publication
		korName, korTitle_A, reference_A
Working Relationship	Relationship of Positions	to show the relation between researchers who perform the same job, by analyzing position and title
		primaryWork, secondaryWork, workingType, position, membershipStatus
	Relationship of Affiliation	to show the relation between researchers who belong to the same organization, by analyzing affiliated organization and academic associationt
		institutionName, membershipName, professionalCareer

on the level of lightweight ontology without definition of axiom and inference function.

4. Developing Researcher Information Retrieval Interface

On the basis of researcher information ontology designed in Chapter 3, the study constructs two separate retrieval interfaces for the simple retrieval interface which allows for users to directly search and browse researcher information and the advanced retrieval interface which provides a number of option to formulate complex queries. In other words, the one is developed to get the factual information using by three access points such as researcher's name, topic/keyword, and affiliation' name and the other is developed not only to search researchers using by Boolean logic and field limitation, but also to browse the human networking, intellectual networking, working networking. For construction retrieval interfaces, the study utilizes My SQL, J2EE 6.0 and XML.

The IS Researcher Information Retrieval System (RIRS) meets requirements for securing personal privacy with access control and usage restriction. The system is started with login process through authentication of real name and official certificate process. The first screen of this system provides login boxes and general information for joining memberships. After finishing login process, the system leads to the next screen which allows users to search researcher information based on name of a researcher, topic or subject keyword, the name of affiliated organization.

4.1 Name Retrieval

According to ontology modeling, the search interface for Name is designed in order to access the five kinds of information using by 35 properties from 12 subclasses. That is, users browse personal information, research information, affiliation information, teaching information, related people information about one researcher.

If users input the name of researcher in search box and activate "Name" mode, they find 5 kinds of information concerning that researcher. Personal Information provides the researcher's basic personal information such as gender, birth date, birthplace, home/office address, phone number, mobile phone number, email/homepage address, educational/carrier history, award record, and certificate record. As the researchers joined in our RIRS decide whether the each information is open to the public or not, all of personal information cannot be provided.

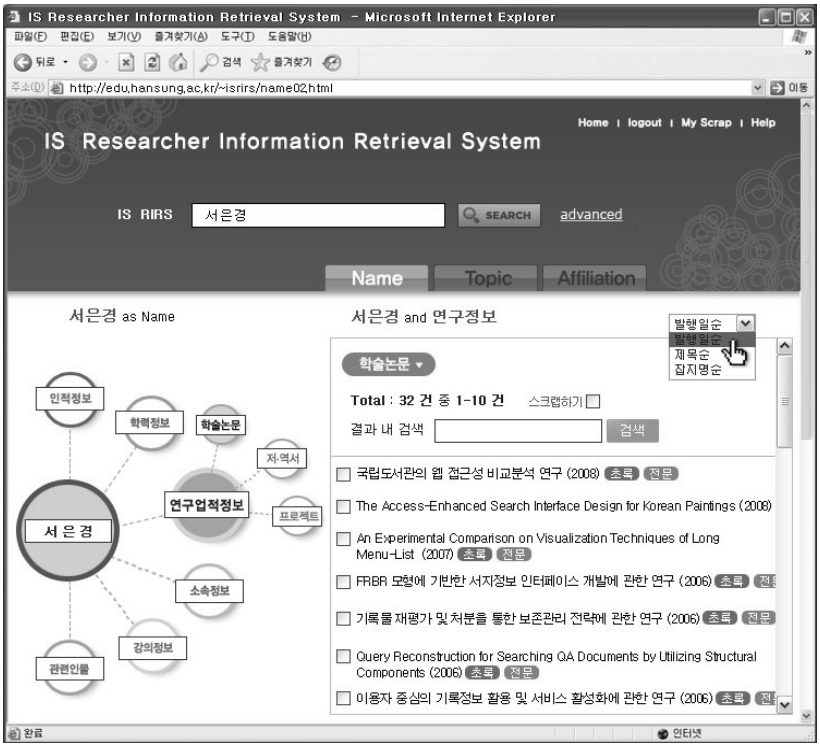
Research Information is provided separately according to the three types of research outputs such as journal articles, monographies, and project reports. If 'Journal Articles' button is selected, the list of article titles which that researcher has published in academic journals until now is displayed. The results, that is, article titles are sorted by alphabet title or journal name, or by date of publication. If users want to review abstract or full text, they

just activate a 'Abstract' button or 'Full Text' button. By using the 'My Scrap' function, users can select and save the retrieved articles which they want to keep. Also, the system provides 'Search within the Results' function which can narrow down the first search results(See <figure 4>). 'Monographies' and 'Project' can be searched and browsed with the same manner.

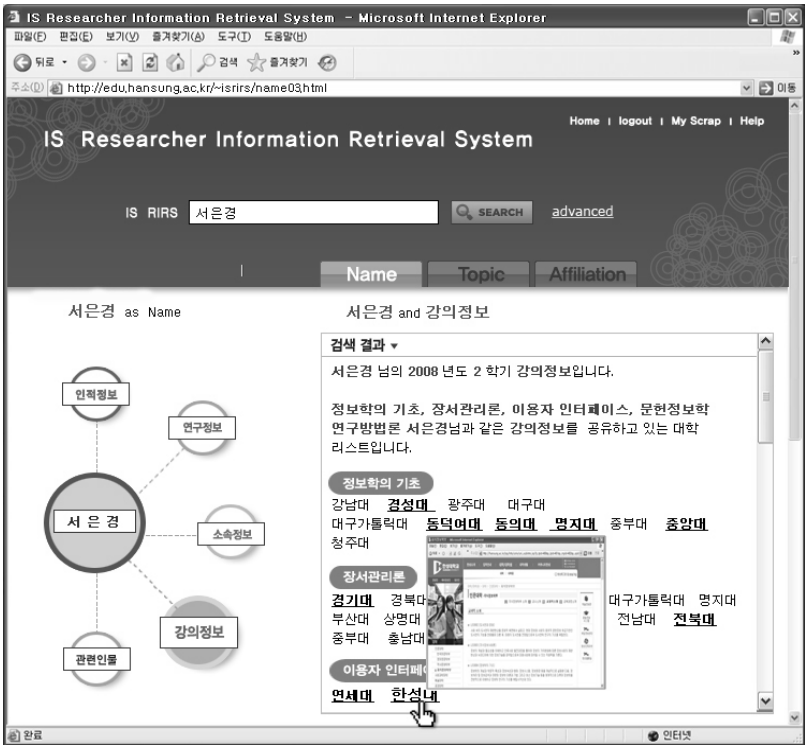
Affiliation Information provides two types of information, that is, 'Institution Information' and 'Membership Information' which a researcher belongs. Institution Information is covered by institution name, department name, Institution type, and position and Membership Information is covered by membership name, membership status, participation duration, and contact point.

Teaching Information lists the course titles which the researcher is opening during the present semester. Below, under the course title, the system lists universities alphabetically where the same course is open. So users identify at a glance which university offers the course or how many universities offer the course. Moreover, if the course introduction of that university is open to public in the website, the system links to the web page (See <Figure 5>).

Lastly, Related People Information could be created on the basis of personal information and



〈Figure 4〉 Search Interface for Research Information in NAME Mode



〈Figure 5〉 Search Interface for Teaching Information in NAME Mode

affiliation information. Through this mode, users can find colleagues or seniors/juniors who graduated from the same university or graduate school using by the educational background information. Besides, users browse people who are working at the same institution or organization. If a user puts the mouse on the name in the result box, the brief information like researcher's name and affiliation appears on temporary pop-up window.

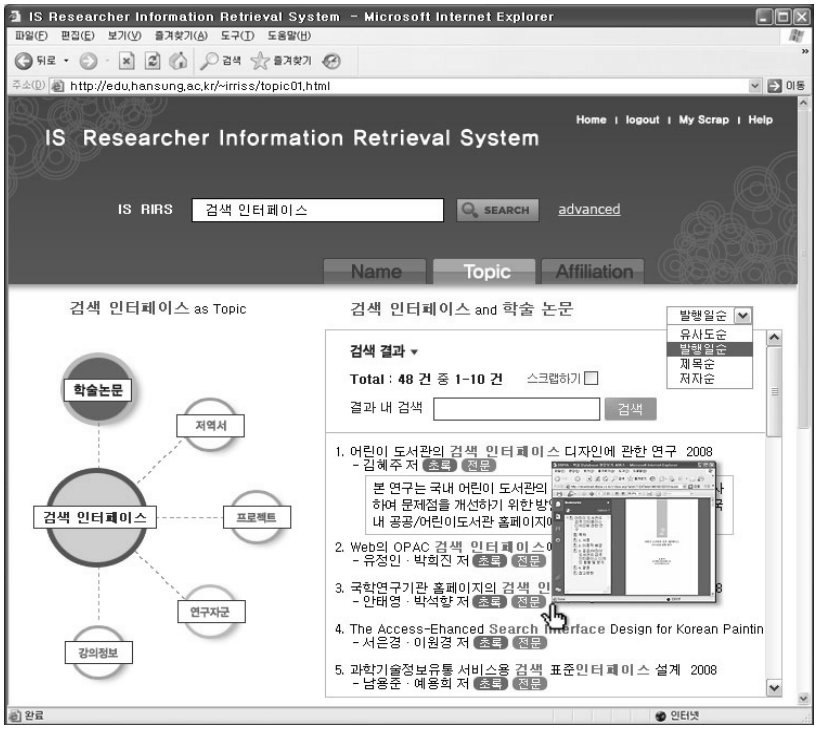
4.2 Topic Keyword Retrieval

The search interface for Topic is designed in order to access the five kinds of information using by 20 properties from 7 subclasses. Therefore, users browse journal articles, monographies, projects, research group, and teaching information. Especially, the system could identify the relation between researchers who are working on the same subject by analyzing the subject of academic papers and publications, and the relation of publications which have similar topic with the author's published paper.

If "Topic" mode is activated, users can search researchers' research outputs, researcher group who have researched on the same subject, and the list of courses which cover with the same subject,

by using topic keyword. If users input subject keyword in the search box and select the 'Journal Articles', the system, first, displays title, author(s) and publication year of journal articles which play on that subject. Users can also select the sorting key such as relevance, date of publication, article title, or author. If a user puts the mouse on the article title in the result box, the full information of publication appears on temporary pop-up window. Also, if users want to review abstract or full text, they just activate a 'Abstract' button or 'Full Text' button. By using the 'My Scrap' function, users can select and save the retrieved articles which they want to keep(See <Figure 6>). With the same manner, users can search 'Monographies' and 'Project' which treat that subject and browse them.

Research Group shows the researchers working on the same subject. That is, researchers whose research fields match with the keyword could be searched. The results are sorted alphabetically by the name of researchers or the name of affiliation which the researcher serves. If a user puts the mouse on the name in the result box, the brief information like researcher's name, affiliation, and lecture titles appears on temporary pop-up window. If a user click on the name, personal information dataset such a gender, birth date, birthplace, home/office



〈Figure 6〉 Search Interface for Journal Articles in TOPIC Mode

〈Figure 7〉 Search Interface for Teaching Information in NAME Mode

address, phone number, mobile phone number, email/homepage address, educational/carrier history, award record, and certificate record also appears on temporary pop-up window.

If a user inputs a search term as “retrieval interface” and activates Teaching Information, the system displays the course titles which contain a search term, the name of lecturer, affiliation of lecturer, and the syllabus which includes course description, course schedule, references etc. The results are sorted by the name of course, course title, or name of institution which offers the course (See <Figure 7>).

4.3 Affiliation Retrieval

The search interface for Affiliation is designed in order to access the three kinds of information using by 20 properties from 3 subclasses. Therefore, users browse Introduction of affiliation, its Publication and Research Group. According to ontology modeling, the system could show the relation between researchers who perform the same job, by analyzing position and title and show the relation between researchers who belong to the same organization, by analyzing affiliated organization and academic association.

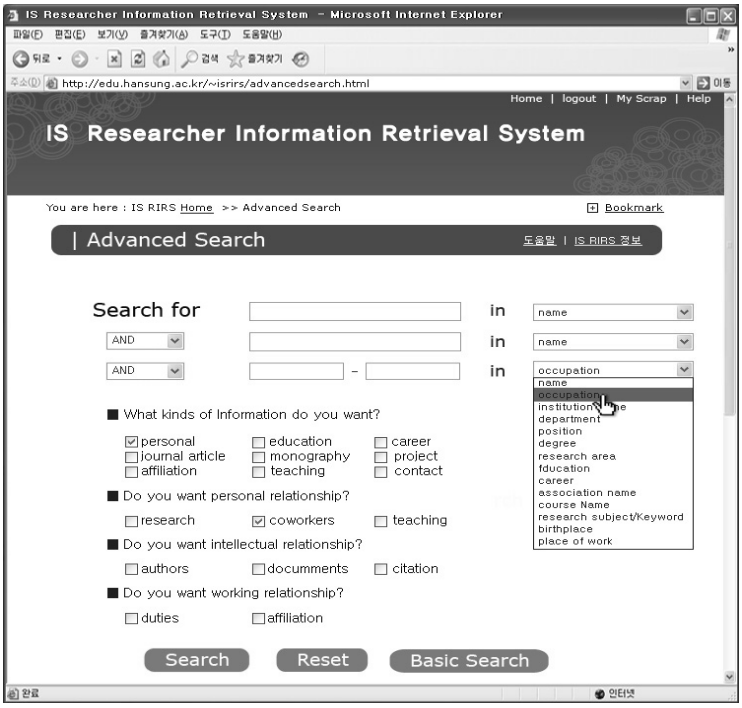
If “Affiliation” mode is activated, users can

〈Figure 8〉 Search Interface for Research Group in AFFILIATION Mode

browse researchers who are in employ of a certain institution or/and association(See <Figure 8>). The information about affiliation of researchers could be get from three subclasses such as Career Information, Institution and Memberships. If a user inputs a name of organization which would be a institution, an association, or an academic society, he/she could browse Introduction, Publications, or Researcher Group of an organization: 'Introduction' informs about the history, features and board members of the organization; 'Publication' lists the various types of publication which the organization has published; 'Researcher Group' displays the name list of researchers who are in employ of that organization with the name and position of other organization which that researcher also belongs to.

4.4 Advanced Retrieval Interface

The Advanced Retrieval Interface provides a number of options to formulate complex queries. Here, users can select a field to search and enter search term(s) with specific Boolean search conditions. The interface specially provides a combo box for 14 kinds of fields such as name, occupation, institution, department, position, degree, research area, education history, career history, asso-



〈Figure 9〉 Advanced Search Screen

ciation name, lecture name, research subject keyword, birthplace, working place. Also it is possible to specify the retrieval result by input of birth date, period in education, period in employment, working duration. So, users limit easily the search results(See <Figure 9>).

Underneath the retrieval formula box, there are the four questions. The first question allows users to specify the types of information about the researcher such as personal, education, career, journal articles, monography, project, affiliation, teaching, and contact. For example, if a user wants to know only researcher's personal data, he/she checks the box of 'personal.' And the second to fourth questions ask whether users want social relationship information such as personal relationship, intellectual relationship, or working relationship. The interface allows also users to specify the range of relationship.

5. Conclusion

Gruber's view that an ontology is an explicit specification of an abstract, simplified view of a world and specifies both the concepts inherent in this view and their interrelationship(Gruber 1995). A typical reason for constructing an ontology is to give a common language for sharing and reusing knowledge about phenomena in the world of interest

and to analyze and make explicit domain knowledge(Noy and McGuinness 2002). Therefore, in the emergent era of knowledge-based systems, the importance of ontology is growing more and more. Ontology which is used to model and reason about today information systems at the conceptual level can play a major role in many of information systems. Ontologies are one step forward in our endeavor to create better models and are also methods and tools to help the identification of the basic things in the world or the basic constructs of information systems(Fonseca 2007). Especially, the concept of ontologies are brought to the information system as part of the search for an answer to some problems in conceptual modeling(Yildiz and Miksch 2005).

The objective of the study is to design more user-friendly and smarter retrieval interface based on ontological analysis, which can provide more precise information by reducing semantic ambiguity or more rich linked information based on well-defined relationships. Therefore, we propose a retrieval interface based on ontology-based representation, which supports to semantic searching and browsing regarding researchers as a full-fledged domain. Particularly, we dessign the retrieval interface which directly retrieves various facts for researcher information such as research outputs or the personal information, or carrier history and browses the social connection of the researcher such as researcher groups that is lecturing or researching on the same subject or involving in the same intellectual communication.

This study, first of all, focuses on ontological analysis on researcher information as selecting descriptive elements, defining classes and properties of descriptive elements, and identifying relationships between the properties and their restriction between relationships. Next, the study designs the prototypical retrieval interface for researcher information with multiple access points.

The main contribution of our work is as follows: first, this study is to do an in-depth analysis of researchers as domain knowledge and to draw up a declarative specification of researchers. This analysis is valuable when both attempting to develop other ontologies and to elaborate these ontologies later by other program. Therefore, this ontology can then be used as a basis for some application in a suite of people-searching tools. Second, it is found that the ontological analysis about researchers allows infer social relationship of researchers such as personal networking, intellectual networking, and working networking. Though intuitive browsing, users, therefore, easily identify the colleagues or alumni of a researcher, researcher group who are closely related in the specific research field, researchers who is teaching or researching the same subject, researchers enrolled as the same member, and etc. Third, the study dessigns the retrieval interface which allows users searching researchers and browsing its results within one screen. Also the researcher information retrieval system proposed by the study attempts to import the various web resources if the resources are open for the public. Therefore, users could browse various kinds of information pertaining to

a certain researcher and grasp the relationship to others in the search interface.

The creation of ontologies for semantic applications is a difficult and time-consuming task because it usually requires the knowledge of domain experts, the skills of ontology engineers, and the analysis of demands and resources. However, it is no wonder that well defined ontology based knowledge representation should be built for semantic search improvements and for contextual or conceptual interrelationships. Much work remains to be defined axiom and inference rules for identifying update relationships and to be designed retrieval interface based on visualization techniques for improving ambient findability.

References

Chu, Heting. 2006. *Information Representation and Retrieval In the Digital Age. New Jersey: Information Today,* Inc.

Chung, Hyun Sook and Young Sook Shin. 2008. "Design and implementation of a syllabus management system based on ontology." *Journal of Korean Institute of Information Technology,* 6(5): 108-114.

Fonseca, Frederico. 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. "Atransaction approach to portable ontologies." *Knowledge Acquisition,* 5(2): 199-220.

Han, Yong-Jin, Se-Young Park, Young-Hwa Lee, and Kweon-Yang Kim. 2008. "Semantic search based on event ontology." *Journal of Electrical Engineering and Information Science,* 14(1): 96-100.

Jee, Jeong Gyu. 2003. "A study on the development of national researcher information database." *Journal of Korean Society for Computer Information,* 10(1): 16-26.

Kim, Hak-Lae and Hong-Gee Kim. 2007. "Personal electronic document retrieval system using semantic web/ontology technologies." *Journal of the Society for e-Business Studies* 12(1): 135-149.

Kim, Pyung, Seung-Woo Lee, In-Su Kang, Han-Min Jung, Jung-Yeoun Lee, Won-Kyung Sung. 2008. "The academic information analysis service using OntoFrame - recommendation of reviewers and analysis of researchers' accomplishments." *Journal of Electrical Engineering and Information Science,* 35(7): 431-441.

Kim, Su-Kyoung and Kee-Hong Ahn. 2007. "An implementation of inference-based web ontology for intelligent image retrieval

system." *Journal of the Korean Society for information Management,* 24(3): 119-147.

Lee, Jung Hee and Hee Sop Kim. 2007. "A design and implementation of ontology-based retrieval system for the electronic records of universities." *Journal of the Korean Society for information Management,* 24(3): 343-362.

McGuinness, Deborah L. and Frank Harmelen. 2004. "OWL web ontology language: Overview." <http://www.w3.org/TR/2004/REC-owl-features-20040210>.

Morbille, Peter. 2005. *Ambient Findability*. O'Reilly & Associates, Inc.

NAM, Jee-Sun. 2008. "Etude sur les ontologies des informations de l'individu construites pour des moteurs de recherché." *Journal of French Education,* 27: 201-229.

Noy, N.F. and D.L. McGuinness. 2002. "Ontology development 101: a Guide to grating your first ontology." <http://protege.stanfaord.edu/publications/ontology-development/ontology-noy-mcauiness.html>.

Park, Chang Sup. 2007. "Multimedia information retrieval using semantic relevancy." *Journal of Korean Society for Internet Information,* 8(5): 67-79.

Park, Yoon Hee. 2008. *Extracting the Information about Individuals to Build the Ontology Instance."* Master's thesis, Kyungpook National University.

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

Suh, Yong moo and Dong hee Yoo. 2008. "An ontology-based hotel search system using semantic web technologies." *Journal of Society for e-Business Studies,* 13(4): 71-92.

Vallet, David, Miriam Fernandez, and Pablo Castells, 2005. "An ontology-based information retrieval model." <http://www.acemedia.org/aceMedia/files/document/wp7/2005/eswc05-uam.pdf/>.

Yildiz, Burcu and Silvia Miksch. 2005. "Ontology-Driven Information systems: challenges and requirements." <http://public.tuwien.ac.at/files/pub-inf_4601.pdf>.

Yildiz, Burcu and Silvia Miksch. 2007. "ontoX - a method for ontology-driven information extraction." *ICCSA,* 3: 660-673.