

한글 시소러스에서 저자와 저작에 대한 관계 설정과 동형 이의어의 기술

A Study on the Description of Relationships and Homographs in Terms of Creator and Work in the Korean Thesaurus

한 상 길(Sangkil Han)*

최 석 두(Sukdoo Choi)**

목 차

1. 서론	4.3 시소러스에서의 구별 방법
2. 연구 방법	5. 저자명과 저작명의 용어 관계 확장과 한정어 사용 방법의 제안
3. 통합 시소러스의 개념	5.1 저자명
4. 동형 이의어 구별 방법	5.2 저작명
4.1 동형 이의어	6. 결론
4.2 사전 및 데이터베이스에서의 구별 방법	

초 록

개인 저자 집합과 저작 집합 간의 관계를 기술할 때 두 집합 내에서 동형 이의어(즉, 동명이인과 동일 저작명)를 구별하지 못하면 두 집합 간에 전단사(全單射) 관계가 형성되지 않으므로 정확한 정보 검색을 위한 정보로는 사용하기가 어렵다. 실제로 저자명과 저작명을 다루는 도서관 시스템, 문헌, 포털사이트 등에서 동형 이의어를 명확하게 구별하고 있지 않아 색인과 검색 시 다의성에 의한 혼란과 불편을 초래하고 있다. 이에 대한 필요성은 일찍이 전거 데이터의 구축 시 대두된 문제였으나 우리나라에서는 일부 기관이 개별적으로 구축하였을 뿐 국가 차원의 전거 데이터가 없어서 이들의 동형 이의어를 구별하기 위한 기준이 없다. 이에 본 연구자는 개체명 인식을 위한 작업의 일환으로 주제어뿐만 아니라 고유 명사류도 포함되는 한글 통합 시소러스 구축 작업에서 얻은 결과를 바탕으로 저자와 저작의 용어 관계 설정 방법과 두 집합 내에서 그리고 두 집합이 기타 용어와의 관계에서 발생하는 동형 이의어의 기술 방법을 대중 문화 예술 분야를 중심으로 제시하였다.

ABSTRACT

The failure of distinguishing homographs in describing relations of individual authors and relations of authorship(i.e. distinction of persons with the same name or persons of the same literary author's name) will cause difficulties of retrieving exact information. It is because relations of automorphism cannot be formed between the two sets mentioned above. Therefore, it is ultimately necessary to set up the criteria or tools to distinguish homographs in order to retrieve more exact information. In the past, some efforts were made to develop authority data in order to solve the homograph problems by individual libraries, documents and portal sites in Korea. It is well understood that developing authority data by an individual institution was very difficult with no criteria or no rules to clarify the homograph problems at the national level. This study is to develop ways of recognizing individual names including subject words and proper nouns. The results of the study will present methods of distinguishing and describing homographs between individual author sets, and authorship sets particularly focused on the areas of arts and popular culture.

키워드: 한글 시소러스, 저자, 저작, 동형 이의어, 용어 관계, 한정어

Korean Thesaurus, Author, Work, Homograph, Term Relationship, Qualifier

* 대림대학 문헌정보과 교수(skhan@daelim.ac.kr)

** 한성대학교 지식정보학부 교수(sdchoi@hansung.ac.kr)

논문접수일자: 2011년 10월 15일 최초심사일자: 2011년 10월 17일 게재확정일자: 2011년 11월 17일
한국문헌정보학회지, 45(4): 139-155, 2011. [http://dx.doi.org/10.4275/KSLIS.2011.45.4.139]

1. 서론

우리말에서 동형 이의어의 비율은 매우 높다. 이는 가장 많은 표제어를 수록하고 있는 『표준국어대사전』의 전체 표제어 509,076개 중 22.4%에 해당하는 124,254개가 동형 이의어로 나타나고 있는 것으로 알 수 있다. 특히 '사장', '조사' 등과 같이 30개 이상의 동형 이의어를 가지는 단어가 20개나 되고 최대 46개의 동형 이의어를 가지는 단어도 있다(양경용 2010). 동형 이의어의 구별 방법에 대하여 『표준국어대사전』을 위시하여 『금성관 국어대사전』(운평어문연구소), 『우리말큰사전』(한글학회), 『고려대 한국어대사전』(고려대 민족문화연구원) 등 대단위 사전을 조사한 결과, 모두 어께 번호를 사용하고 있었다. 반면, 시소러스에서의 동형 이의어 구별 방법은 대부분 한정어 방법을 적용하고 있었으며, 어께 번호를 사용하는 경우는 전혀 없었다. 오히려 일부 최근에 편집된 소규모 국어사전(예: 『연세 초등국어사전』 등)에 한정어 방법이 적용되고 있었다.

일반 주제어와 마찬가지로 저자와 저작 간의 관계에서도 동명 이인과 동일 저작명이 나타나며 색인과 검색에 많은 장애가 되고 있다. 그러나 일반 주제어의 동형 이의어에 대한 한정어 구별 방법을 고유 명사에 그대로 적용하는 데는 문제가 있다. 인명(여러 가지 이칭 포함)이나 저작명은 끊임없이 생성되어 추가되는 특성을 갖고 있다. 하나의 저작이 매체가 바뀌거나 장르를 넘나들기도 하고, 동일한 저작이 시간 간격을 두고 반복하여 재생산되는 경우도 있다. 따라서 일반 주제어의 구별 방법을 사용하게 되면 새로운 것이 추가될 때마다 기존 인명이

나 저작명 전체의 한정어를 갱신해야 하는 문제가 발생한다.

저자명과 저작명에 관련된 데이터를 구축하기 위해서는 한 번 한정어를 기술함으로써 그 개체를 유일하게 식별될 수 있는 체계가 필요하다. 현재 본 연구자는 개체명 인식을 위한 작업의 일환으로 주제어뿐만 아니라 고유 명사류도 포함되는 용어의 메타데이터로서 통합 시소러스를 구축하고 있다. 이 시소러스에서는 저자명과 저작명 두 집합에서 발생하는 동형 이의어 및 그 이외의 집합과의 관계에서 발생하는 동형 이의어를 구별할 수 있는 한정어 사용 방법을 연구하여 적용하고 있으며, 아울러 저자와 저작의 관계를 정의하고 있다. 본 연구에서는 지금까지 통합 시소러스에서 적용한 원칙과 결과에 대하여 대중 문화 예술 분야를 중심으로 논하고자 한다. 실제로 일정 수 이상의 데이터를 수집하여 구축한 결과는 일반적으로 경우의 수를 망라할 수 없다는 이론적인 측면의 빈 곳을 보완할 수 있을 것이다. 대중 문화 예술 분야로 한정된 것은 전 분야를 대상으로 논하면 너무 광범위하여 논점이 흩어질 수 있다는 점과, 대중 문화 예술 분야가 다른 분야의 데이터에 비해서 비교적 여러 가지 경우를 접할 수 있기 때문이다.

2. 연구 방법

정보의 요구가 정교해짐에 따라 정보 처리에 사용하는 용어는 개체를 식별하고 의미의 표현이 필요하게 되었다. 예컨대, '고양이'가 성이 '고'이고 이름이 '양이'인 사람도 아니고, 성이

‘고’인 여자 아이를 지칭할 때 주격 조사 ‘이’가 붙은 것이 아니고, ‘고양’이 지명이 아니고, 그리고 동물의 다른 ‘고양이’가 아니라 그 ‘고양이’라는 것을 시스템이 이해할 수 있다면 보다 정확한 정보 검색이 가능할 것이다. 이를 위하여 다목적용 용어 메타데이터를 목표로 통합 시소러스를 기획하고 다음 다섯 가지 지침을 정하였다.

- 1) 기본적으로 한글 용어를 대상으로 한다. 다만, 로마자 두문자어를 포함한다.
- 2) 사전을 포함하여 문헌상으로 한 번 이상 사용된 모든 용어를 대상으로 한다.
- 3) 용어는 주제어, 인명(자, 호, 시호, 필명, 본명, 별명, 속명, 법명, 세례명, 유아명, 칭호 등 모든 이명을 포함한다), 법률명, 동물명, 식물명, 기관/단체명, 상품명(모형을 포함한다), 지명(목록 규칙에서 정의하고 있는 지명을 말한다), 기념물명, 작품명(모든 장르의 작품을 포함한다), 사건명, 회의명의 12가지 범주가 포함되며, 문자열의 길이에 관계없이 하나의 용어로 취급한다. 또한 각 범주별로 메타데이터 형식을 만들고 시소러스 표준에서 정의하고 있는 용어 관계는 메타데이터 요소의 일부로 사용한다.
- 4) 대상 용어를 품사별로 보면 명사, 동사, 형용사, 수사, 그리고 주제나 개념을 표현한다면 구, 문도 포함되며, 동일 원칙의 외국어 음차어를 포함한다.
- 5) 이상의 지침에 벗어나지 않으면 방언, 속어, 고어, 잘못 사용되는 용어도 대상이 된다. 다만, 식별 기호, 분류 기호, 한자(우리나라), 대응 외국어 등은 필요한 경우에만

참조 데이터로 이용한다.

이외에 띄어쓰기, 한정어, 용어의 수집, 용어의 선정, 용어의 기술, 우선어(전거; 대표어; 표준어)의 선정, 한자 표기, 외래어 표기, 음차 표기, 외국어 표기 등에 관한 여러 가지 기준을 마련하였다(최석두 외 2006).

이상의 지침에 따라 각 집합의 메타데이터를 처리(입력, 생성, 수정, 삭제, 삽입, 검색, 관리 등)할 수 있을 뿐만 아니라 각 용어가 속해 있는 다양한 집합을 식별할 수 있도록 용어 관리 시스템을 설계하였다. 예컨대, 주제어와 인명의 메타데이터 요소를 보면, 주제어는 용어 관계(NT/BT, RT, USE/UF, 학명, 전후 관계, 각종 외국어 관계, 각종 분류, 일자, 관리 정보 등) 이외에 ‘품사, 정의, 활용형, 용례, 시대, 분류, 이미지, 본문, 링크, 비고’를, 인명은 상기 용어 관계 이외에 ‘정의, 성별, 직군, 생몰년, 본관, 출생지, 직업, 기수 정보, 지위, 소속 기관, 가입 단체, 저작, 수상, 경력, 스승, 제자, 국적, 시대, 학력(초중고대학원), 가족, 지인, 연락처, 분류, 이미지, 링크, 비고’를 가지고 있다. 대부분의 요소가 반복 가능하며 기수 정보, 소속 기관, 학력, 연락처 등은 하위 요소를 가지고 있다. 그리고 모든 요소에 대하여 기술 규칙을 정하고 이 설계에 따라 시스템을 개발하였다.

개발된 시스템을 이용하여 연구용으로 구축하고 있는 통합 시소러스는 약 110만 용어를 수용하고 있으며, 이중 저자(작품이 없는 인명도 포함됨)와 저작 관련 용어 집합은 현재 개인명 17,300명, 기관/단체 16,058개, 문헌 29,207건(물리적인 책의 개념이 아니므로 시, 음악, 그림 등의 한 편도 하나의 개별 문헌으로 다루고

있다), 도합 62,565항목이 포함되어 있다. 이 가운데 대중 예술 분야는 개인명 2,461명, 기관/단체 223개, 문헌 14,895건, 도합 17,579항목이다. 이 데이터의 구축 시 국내의 전거 데이터를 이용할 수 있으면 많은 도움이 되리라 생각되지만 이들은 공개하지 않는 데이터이기 때문에 불가능한 일이었다. 주제어를 포함한 모든 데이터는 직접 구축하였다.

본 연구에서는 통합 시소러스의 구축 실험에서 적용한 두 가지 방법에 대하여 논한다. 우선, 전술한 12가지 분야와 저자명(기관/단체명도 포함) 간에 동형 이의어가 발생하거나, 저자명 간에 동형 이의어가 발생하는 환경에서 시스템과 이용자가 쉽게 동형 이의어를 식별할 수 있는 실증적인 방법이다. 또 하나는 시소러스의 표준 용어 관계에 저자와 저작의 관계를 정의하는 방법과 저자의 역할을 표현할 수 있는 방법이다.

3. 통합 시소러스의 개념

‘통합 시소러스’라는 용어를 사용한 것은 전술한 용어의 범주에 속하는 모든 용어를 통합한다는 의미와 함께 전거 데이터(도서관에서 주로 전거 제어를 위하여 만든 데이터)를 시소러스에 통합한다는 의미를 가지고 있다. 전거 데이터와 시소러스는 형식과 내용이 거의 일치하며 용도는 약간 상이하나 서로 상보 관계에 있기 때문이다. 특히 전거 데이터의 ‘통합’ 개념은 물리적일 수도 있고 가상적일 수도 있으며 단순히 독립적으로 존재하면서 보완 관계일 수도 있다.

개체를 식별하고 의미를 표현하는 방법의 가장 소극적이면서도 효율적인 방법은 다른 개체나 개념과의 관계를 나타내는 방법이다. 전거 데이터를 시소러스에 통합한다는 것은 로컬 정보를 제외한 원래의 전거 데이터에 추가하여 용어 관계를 더 풍부하게 만들고 객관적인 데이터를 만들어 그 용도를 확대하는 일이다. 그것은 전거 데이터에 비하여 시소러스가 용어 혹은 개념 관계의 설정이 매우 자유롭고, 도서관 소장 자료라고 전제하는 전거 데이터는 현실적으로 불가능한 모든 웹 자원을 목록 대상으로 포함시키지 않는 한 망라되지 않는다는 어려움이 있기 때문이다. 예컨대, 인명 ‘박정희(1917-1979:대통령)[朴正熙]’를 찾고 싶은 이용자는 전거 데이터베이스나 인명 사전을 찾으면 되지만, “한국대통령, 10.26사태, 10월 유신, 군인, 정치가” 등의 일반 주제어와 관계를 설정함으로써 이들 용어를 다른 검색의 경로로 이용할 수 있다. 도서관 장서란 물리적인 권책이 주가 되므로 전거 데이터에서는 한계가 있는, 예컨대 ‘야경차[夜警車]’라는 것이 시의 제목이며 ‘김광균(1914-1993:시인)[金光均]’의 시라는 관계를 만들 수 있다. 또한 색인 시에 문장으로 포함되어 있는 “누구를 위하여 종은 울리나”가 형태소 해석기에 의해 하나의 개체명으로 추출되도록 관계를 만들 수 있다.

통합 시소러스에서의 중요한 집합은 명사, 동사, 형용사가 주류를 이루고 있는 주제어이며, 고유 명사는 도서관에서의 용도와 달리 주로 개체명 인식에 사용되기 때문에 통합 시소러스에서의 전거 데이터는 부수적인 정보이다. 저작 없이 등재해야 하는 인명도 있지만, 모든 저자와 저작을 통합 시소러스에 포함시킬 수

있다면 가장 이상적일 것이다. 통합 시소러스의 가장 기본적인 용도는 색인과 검색이지만 다목적용 용어 데이터이다. 즉, 특정 데이터베이스나 분야에 종속되지 않는 소스 데이터이기 때문에 필요에 따라 다양한 용도로 사용할 수 있다. 예컨대, 이용자가 검색어로 인명을 입력하여 '인명, 작품명, 주제어, 지명'이 섞여 10가지의 동형 이의어가 나열되었을 때 이용자는 이 네 가지 그룹을 직관적으로 구별할 수 있을 뿐만 아니라 인명 등 네 가지 그룹 중에서도 자기가 찾는 인명을 쉽게 선택할 수 있도록 할 수 있다.

4. 동형 이의어 구별 방법

4.1 동형 이의어

국어에서는 단어의 철자나 음운 형식이 동일한 경우를 동음어(同音語), 동형어(同形語), 동철어(同綴語) 등 편의에 따라 다양한 말로 표현하고 있다. 여기에 의미의 동일 여부를 덧붙여 동음 이의어(同音異義語), 동형 이의어(同形異義語), 동철 이의어(同綴異義語)라고도 하는데, 이 경우 동형어와 동철어라는 말에는 이미 '이의(異義)'라는 의미를 내포하고 있다고 할 수 있다. 왜냐하면 형태가 같고 의미까지 동일하면서 다른 말이 있을 수 없기 때문이다.

일반적으로 국어에서 동음 이의어는 크게 세 가지 형태로 나타나는데, 첫째, '배(梨)[배], 배(船)[배], 배(腹)[배]'와 같이 철자와 음운 형식이 동일한 경우가 있고, 둘째, '낮[날], 낮

[날], 낮[날]', '학문(學問)[학문], 항문(肛門)[학문]'과 같이 철자가 다르지만 음운 형식이 같은 경우가 있다. 셋째, 이와 비슷한 경우로 '눈(雪)[눈:], 눈(眼)[눈]'과 같이 철자는 동일하지만 음운 형식이 다른 경우가 있는데, 이는 엄밀히 말하면 동음 이의어가 아닌 동형어 또는 동철어라 할 수 있다(유현경 2000). 동형어와 동철어는 음운 형식에 상관없이 철자가 동일한 경우를 일컫는 것으로 위의 첫 번째와 세 번째가 이에 해당된다. 동형어와 동철어는 같은 의미로 사용되는 용어로서 동철어보다 동형어가 좀 더 일반적이라 할 수 있으며, 동형어와 동형 이의어 또한 같은 의미로 사용되는 단어이나 동형 이의어라는 표기가 그 뜻을 보다 분명하게 담고 있다고 할 수 있다.

시소러스나 사전 또는 어휘 데이터베이스 등과 같은 분야에서는 단어의 음운 형식보다 표기 형태가 중요한 정보 처리의 요소가 되고 있어 철자가 동일한 경우를 구별할 필요성이 매우 높다. 따라서 본고에서는 위의 세 가지 유형 가운데 첫 번째와 세 번째의 경우를 대상으로 동형 이의어라는 용어를 사용하되 경우에 따라 동형어라는 용어도 같이 사용하기로 한다.

4.2 사전 및 데이터베이스에서의 구별 방법

동형 이의어를 구별하기 위한 일반적인 방법은 대부분의 국어사전에서 채택하고 있는 어깨 번호를 붙이는 방법이다. 보조 정보로서 한자나 영문을 부기하고 있다. 어깨 번호는 하나의 사전 내에서 어휘를 구별하여 배열의 순서를 정한다는 측면에서는 매우 경제적이고 효율적이다. 그러나 어깨 번호를 붙이는 기준에는 사

용 빈도, 획수, 한자어, 고유어, 품사 등이 있으며 이것은 사전마다 다르다.

대부분의 국어사전에서는 어원이 같으면 동형 이의어로 취급하지 않는다. 다음의 ‘이[齒]’와 ‘종심[終審]’의 경우를 보자.

이[齒]

- 1) 척추 동물의 입 안에 있으며 무엇을 물거나 음식물 씹는 역할을 하는 기관.
- 2) 툇, 툇니바퀴 따위의 뽕족뽕족 내민 부분.
- 3) 기구, 기계 따위의 ...

종심[終審]

- 1) 『교육』논문 심사 때에, 논문 통과 여부를 결정하는 맨 마지막의 심사.
- 2) 『법률』심급 제도에서, 맨 마지막에 해당하는 심급.

정보의 색인과 검색 측면에서는 동물의 이[齒]와 툇니바퀴의 이[齒]를 구별할 필요가 있을 것이다. 시소러스에서 이를 ‘이(잇빨)’와 ‘이(툇니바퀴)’로, ‘종심(법률)’과 ‘종심(논문)’으로 구분한다면, 국어학 측면에서는 하나의 용어이지만 시소러스에서는 동형 이의어가 된다.

박수연(2002)은 동음이의어 구별을 위한 ‘길잡이말(guide words)’ 연구에서 길잡이말의 문제점을 제시하고 외국인을 위한 한국어 학습 사전에서의 동형 이의어 구분을 위한 길잡이말 원칙을 제안하였다. 이는 다음 절에서 논하는 한정어의 사용에 속하는 방법이다.

데이터베이스에서의 인명에 대한 구별은 데이터베이스마다 다르다. 네이버 인물 검색에서는 직업별 구분과 출생일로, 네이트 디렉토리

의 인물 정보에서는 동명 이인을 직업별로 구분하고 다음으로 출생, 신체, 학력, 데뷔, 수상, 경력, 취미 등으로 구분하고 있다. JOINS 인물 정보의 경우는 이름에 한자를 부기하고 소속기관, 직업, 생년, 인맥 관계로 구분하고 있다. 네이트의 예를 보면 다음과 같다.

■ 박상민(영화배우)

출생: 1970년 10월 19일(경상남도 진해)

신체: 177cm, 72kg

학력: 서울예술대학 영화과

■ 박상민(가수)

출생: 1964년 5월 2일

신체: 172cm, 60kg, O형

학력: 도쿄음악대학

4.3 시소러스에서의 구별 방법

대부분의 시소러스가 한정어를 사용하고 있으며 일부를 보면 다음과 같다.

『NASA 시소러스』에서는 두 가지 이상의 의미로 사용되는 경우 혹은 용어 간의 구별을 명료하게 할 필요가 있을 경우에 한정어를 삽입한다. 『주제명요어집』에서는 그대로의 명사로는 의미가 불명료하거나, 또는 여러 주제 분야에서 공통적으로 사용되는 용어는 그 의미를 한정하는 말을 ‘조류(鳥類)’, ‘조류(潮流)’, ‘조류(藻類)’와 같이 원괄호 속에 넣어 부기한다. 또한 『주제명요어집』에서는 지명 아래에서의 주제 구분, 주제의 지리 구분, 시대 구분에서도 괄호 한정어를 사용한다. 『법률시소러스』(정보관리학회 1998)에서는 동형 이의어가 있는 경우, 한글/외국어를 불문하고 한글 한정어로 구

분하고 있으며, 한자가 있는 용어일 경우에는 ‘사법(법률)[私法]’, ‘사법(법원)[司法]’과 같이 한자와는 별도로 한정어를 부기하고 있다. 이때 부기하는 한자는 참조 정보일 뿐 식별 정보는 아니다.

『신문기사 시소러스』에서도 동음이의어를 구별하기 위해 괄호 한정어를 사용하는데, 정보량(색인이나 검색의 빈도수)이 많은 것은 그대로 두고 그 반대의 경우 말미에 식별구를 원괄호에 넣어 부기한다. 괄호를 포함하는 식별구는 색인어의 일부가 된다. 공단의 경우, ‘공단(工團)’, ‘공단(특수법인)(公團(特殊法人))’으로 구분하고 있다. 이 경우 앞의 ‘공단(工團)’은 공업 단지를 의미하며, 이 용어가 뒤의 특수 법인인 공단보다 정보량이 많기 때문에 한정어를 부기하지 않는다. 이때 ‘(工團)’은 한정어가 아니라 한자 부기의 일종이고, ‘(특수법인)’이 한정어이다. 『한국통신 시소러스』는 괄호 한정어를 사용하는 대신에 디스크립터에 영문명을 부기하고 있고, 『과학기술용어 시소러스』는 동음이의어를 구분하기 위해 간장(기관), 간장(식품)과 같이 괄호 한정어를 사용하고 있다. 『리재철의 주제명표목표』(리재철 1961)에서는 동형 이의어를 구분하기 위하여 다양한 방법을 사용하고 있는데, 개념을 구별하기 위하여 외래어 영문 부기, 세부 분야의 한정, 보다 명확한 개념의 부기, 한자의 부기 등의 방법을 사용하고 있다.

이와 같이 시소러스에서의 동형 이의어 구별 방법은 한정어의 사용이 주류를 이루고 있으며, 김태수와 최석두(1977)는 한정어의 사용에 대하여 구체적인 방법을 제시하고 있다. 한정어는 동형 이의어를 구별하는 기능을 갖고 있고

해당 용어와 밀접하게 관련되는 한정어를 제시함으로써 이용자에게 해당 용어의 의미를 직관적으로 압축하여 전달할 수가 있다. 그러나 한정어를 일관성 및 통일성 있게 사용하기가 어렵다. 왜냐하면 개인적인 배경에 따라 구별하는 측면이 달라져 부여하는 한정어를 통일시키는 것이 어렵기 때문이다. 즉, 국어사전에서의 어휘 번호와 마찬가지로 복수의 사전 정보를 함께 다루기는 어렵다. 다만, 한정어 방법은 한정어 자체에 의미를 부여할 수 있기 때문에 그 기술 원칙과 순서를 통제하면 전문 분야별 동형 이의어 식별 기호로서 사용가능성이 높은 방법이라 할 수 있다.

동서양을 불문하고 동명 이인을 식별하기 위한 가장 기본적인 정보로 생몰년을 사용해왔다. 그러나 생몰년만으로 구별할 수 없는 경우도 있으며, 이용자 측면에서는 내가 찾는 사람을 항상 생몰년으로 구별할 수 있는 것은 아니므로 언제나 유일하게 구별할 수 있으면서 이용자가 직관적으로 알 수 있는 보조 정보를 필요로 한다. 또한 새로운 인명이 추가되었을 때 이미 등록된 인명 데이터를 수정하지 않아야 편리하며, 저자명이 작품명이나 주제어 등이 아니라 저자명이라는 것을 알 수 있어야 하며, 이들과의 관계에서 생기는 동형 이의어도 구별되어야 한다. 전술한 시소러스의 한정어 사용은 일반 주제명을 대상으로 한 경우이며 이와 같은 복합적인 관점에서 인명과 저작에 관한 적용례가 없기 때문에 이를 위한 새로운 방안이 필요하다.

5. 저자명과 저작명의 용어 관계 확장과 한정어 사용 방법의 제안

5.1 저자명

5.1.1 저자명의 기술

우선어(전거)로서의 저자명은 개인 저자나 단체 저자를 막론하고 다음과 같은 형식으로 기술한다.

- 이름(생몰년:제1 한정어:제2 한정어:...제 n 한정어)

n의 값은 동형 이의어를 구별할 수 있을 때까지 나열하며, 현재 동형 이의어가 없더라도 제2 한정어까지 기술한다. 동형 이의어가 언제라도 나타날 수 있기 때문이다. 제1 한정어는 생몰년(단체 저자인 경우에는 설립-해산 년도)을, 제2 및 제3 한정어로는 직업, 전공 분야, 시대, 지명, 국명 등(단체 저자인 경우에는 분야, 지명, 국명 등)을 기술한다. 각 구분 기호는 쌍점(:)을 사용한다. 두 가지 이상의 직업을 갖고 있고 어느 것 하나만을 선택할 수 없는 경우에는 모두를 쉼표로 구분하여 기술한다. 직업이나 전공은 통계청의 『한국표준직업분류』(2007)의 구분을 참조한다.

저자명의 이형이, 우선어의 변형인 경우에는 우선어의 형식을 따르고, 그 외에는 다음과 같은 형식으로 기술한다.

- 이명(이명의 종류:제1 한정어:제2 한정어:...제n 한정어)

마찬가지로 n의 값은 동형 이의어가 해소될 때까지 나열한다. 현재 동형 이의어가 없더

라도 제1 한정어까지 기술하며, 제1 한정어로는 우선어를 사용한다. 제2 및 제3 한정어의 경우에는 우선어의 방법을 따른다.

저자명은 한정어를 포함하여 유일한 식별 문자열이 된다. 따라서 한정어 이하의 모든 문자열은 참조 데이터일 뿐이다.

- 예) 남궁원(1934-:영화배우)[南宮遠]
박노식(1930-1995:영화배우)[朴魯植]
김혜자(1941-:탤런트, 영화배우)
[金惠子]
S.E.S(1997-2002:가수)

외국 인명 중에서 서양 인명은 크루즈(1962-:영화배우)[Cruise, Tom]와 같이 성만을 차음하여 기술하고, 원어명은 역순으로 부기한다. 한글의 직순 완전명(예: 탐 크루즈)과 역순 완전명(예: 크루즈, 탐)은 알려져 있는 경우에만 UF로 부기한다. 일본 인명 및 중국 인명은 직순으로 기술한다.

- 예) 크루즈(1962-:영화배우)[Cruise, Tom]
디카프리오(1974-:영화배우)
[DiCaprio, Leonardo Wilhelm]
펠트로(1972-:영화배우)
[Paltrow, Gwyneth]
홍금보(1952-:영화배우)[洪金寶]
기무라 다쿠야(1972-:영화배우)
[木村拓哉]

인명의 이형이 우선어의 변형인 경우에는 우선어의 형식을 따른다.

- 예) 다윈(1809-1882:생물학자)
[Darwin, Charles Robert]
UF 다아윈(1809-1882:생물학자)

[Darwin, Charles Robert]

예) 피아프(1915-1963:영화배우)

[Piaf, Edith]

UF 피아프(1915-1963:영화배우)

[Piaf, Edith]

에디트 피아프(1915-1963:

영화배우) [Piaf, Edith]

에디트 피아프(1915-1963:

영화배우) [Piaf, Edith]

에디트 피아프(1915-1963:

영화배우) [Piaf, Edith]

인명의 이형이 우선어의 변형이 아닌 경우, 즉 호, 자, 시호, 본명, 필명, 예명, 속명 등 인물에 대한 이명은 두 번째 형식을 따른다. 제1 한정어(이름의 우선어)로서 동형 이의어가 해소되지 않는 경우에는 제2 및 제3 한정어를 부기할 수 있다. 우선어에 대하여 보라참조표시를 한다.

예) 허장현(본명:허장강)[許長顯]

USE 허장강(1923-1975:영화배우)

[許長江]

홍경일(본명:남궁원)[洪京日]

USE 남궁원(1934-:영화배우)

[南宮遠]

생몰년 미상인 경우에는 "(?-?:직업 혹은 전공)"으로, 생몰년 중 한쪽의 연도를 모르는 경우에는 "(?-연도:직업 혹은 전공)" 혹은 "(연도-?:직업 혹은 전공)"으로 기술한다.

예) 주상호(?-?:영화배우)[朱相鎬]

성소민(?-1982:영화배우)[成笑民]

5.1.2 인명의 동등 관계

인명의 이형은 동등 관계로 취급한다. 같은 사람에 대한 다른 호칭이므로 등가로 취급하는 것이 당연할 것이다. 이때 동등 관계의 내용을 좀 더 상세하게 세분하기 위하여 패시를 사용한다. 아래 예에서 패시 '본명/예명'은 '선우용녀'가 예명이며 '정용례'가 본명이라는 것을 나타낸다.

예) 선우용녀(1945-:배우)[鮮于龍女]

UF 정용례(본명:선우용녀)

[鄭龍禮]본명/예명

독고성(1929-2004:배우) [獨高成]

UF 전원윤(본명:독고성)[全元潤]

본명/예명

외국인의 이형이 우선어의 변형인 경우에는 모두 동등 관계로 취급한다. 예컨대, 패닝(1994-:영화배우)[Fanning, Hannah Dakota]의 경우, '패닝, 한나 다코타', '다코타 패닝', '한나 다코타 패닝'은 모두 동등 관계가 된다. 이 때 영어명은 역순의 완전명(Fanning, Hannah Dakota)을 기술한다.

예) 패닝(1994-:영화배우)[Fanning,

Hannah Dakota]

UF 패닝, 한나 다코타(1994-:

영화배우)[Fanning, Hannah

Dakota]

UF 다코타 패닝(1994-:영화배우)

[Fanning, Hannah Dakota]

UF 한나 다코타 패닝(1994-:영화배우)

[Fanning, Hannah Dakota]

홍금보(1952-:영화배우)[洪金寶]

UF 홍진바오(1952-:영화배우)

[洪金寶]
 기무라 다쿠야(1972-:영화배우)
 [木村拓哉]
 UF 목촌척재(1972-:영화배우)
 [木村拓哉]

NTI 최강창민(1988-:가수)
 [最強昌珉]

둘 이상의 그룹에 속하는 경우에는 두 개 이상의 계층 관계를 갖는다.

믹키유천(1986-:가수)[秘奇有天]
 BTI 동방신기(2003-:음악그룹)
 [東方神起]
 BTI JYJ(2010-:음악그룹)

5.1.3 인명의 계층 관계

인명은 직업명에 대한 계층 관계의 사례 관계로 설정한다. 직업명에 따라 그 집합의 구성 요소가 대단위가 될 수도 있다. 예컨대, ‘교수’ 아래에는 ‘영화평론가’보다 많은 인명이 속하게 될 것이다. 상하위 관계를 맺음으로써 검색의 효율성을 높일 수 있기 때문이다.

예) 영화 배우(연예인)[映畫俳優]
 NTI 김지영(1937-:영화배우)
 [金志暎]
 NTI 김진규(1923-1998:영화배우)
 [金振奎]
 NTI 김칠성(1918-1973:영화배우)
 [金七星]
 NTI 김태희(1980-:배우)[金泰希]

그룹에 속하는 가수는 그룹명에 대한 계층 관계의 사례 관계로 설정한다.

예) 동방신기(2003-:음악그룹)[東方神起]
 NTI 믹키유천(1986-:가수)
 [秘奇有天]
 NTI 시아준수(1987-:가수)
 [細亞俊秀]
 NTI 영웅재중(1986-:가수)
 [英雄在中]
 NTI 유노윤호(1986-:가수)
 [瑜允浩]

5.1.4 인명의 관련 관계 및 저작 관계

개인이 속한 소속사, 부자 관계 등은 관련 관계로 취급하고 “아들/아버지, 형/제, 스승/제자”와 같은 형식으로 관계 패킷을 부가한다. 대응 관계이므로 대응 용어를 중심으로 보면 그 패킷 관계가 반대로 표시된다.

예) 소녀시대(2007-:음악그룹)[少女時代]
 RT 에스엠 엔터테인먼트[SM Entertainment] 소속사/음악그룹
 허장강(1923-1975:배우)[許長姜]
 RT 허기호(1952-:영화배우)
 [許起虎] 아들/아버지
 RT 허준호(1964-:탤런트)[許峻毫]
 아들/아버지
 허준호(1964-:탤런트)[許峻毫]
 RT 허장강(1923-1975:배우)[許長姜]
 아버지/아들

저자와 저작의 관계는 AUT(HOR)와 TIT(LE)이라는 관계 지시어로 관계를 설정하고 관계 패킷을 부가한다. 관계 패킷은 “영화/주연, 영화/조연, 영화/감독”과 같은 형식으로 기술한다. 이때 반대 관계는 “주연/영화, 출연/영

화, 감독/영화”와 같은 대응 관계가 성립된다.

예) 박노식(1930-1995:영화배우)[朴魯植]

TIT 124군 부대(영화:1970)

[124軍部隊]영화/주연

TIT 갈매기 우는 항구(영화:1964)

[--港口]영화/주연

TIT 검은 불연속선(영화:1963)

[--不連續線]영화/주연

TIT 격퇴(영화:1956)

[擊退]영화/출연

격퇴(영화:1956)[擊退]

AUT 박노식(1930-1995:영화배우)

[朴魯植]출연/영화

5.1.5 인명의 기타 관계

인명의 생몰년은 제1 한정어로 부기되어 있으나 별도의 관계 지시 기호 “YEA”로 분리하여 기술한다. 생몰년, 나이, 시대 등으로 검색하기 위함이다.

예) 버그만(1915-1982:영화배우)[Ingrid

Bergman]

YEA 1915-1982

안성기(1952-:배우)[安聖基]

YEA 1952-

검색의 편의를 위하여 다양한 외국어 관계를 기술한다. 모든 외국 인명은 자국의 문자 세트로 기술되었을 때 언어명을 관계 지시 기호로 사용한다. 즉, 일본 인명을 일본어로 기술하였을 때 ‘JPN’의 관계 지시 기호를 사용한다. 한국 인명을 포함하여 영어권 이외의 인명을 로마자로 표기하였을 때는 로마자 표기라는 의미로 ROM이라는 관계 지시 기호를 사용한다.

예) 김혜자(1941-:탤런트, 영화배우)

[金惠子]

ROM Kim Hae ja

홍금보(1952-:영화배우)[洪金寶]

UF 홍진바오(1952-:영화배우)

[洪金寶]

CHI 洪金宝(1952-:영화배우)

ROM Hóng Jīnbǎo

기무라 다쿠야(1972-:영화배우)

[Kimura Takuya]

UF 목촌척재(1972-:영화배우)

[木村拓哉]

JPN きむらたくや(1972-:영화배우)

[木村拓哉]

JPN 木村拓哉(1972-:영화배우)

ROM Kimura Takuya

5.2 저작명

5.2.1 저작명의 기술

저작물의 표기 방법은 『한국목록규칙』제4판(한국도서관협회 2003) 표제와 책임표시 사항의 기술원칙을 따라 해당 자료의 으뜸 정보원에 기재된 형식 그대로 기재한다. 우선어(전거)로서의 저작이 영화, 드라마, 연극, 노래, 음반 등의 분야이면,

저작명(저작물의 종류:간행연도)

로 기술하며, 문학 작품이나 저서이면,

저작명(저작물의 종류:저자명)

으로 기술한다.

또한 저작명의 이형이 우선어의 변형인 경우에는 이상의 기술 방법을 따르고, 그 외에는 다 음과 같은 형식으로 기술한다.

이형(이형의 종류:우선어)

한정어로 사용하는 저작물의 종류는 저작권 법 제4조(저작물의 예시 등)와 한국학술진흥 재단의 『용어데이터 레지스터리 설계 및 발전 방안 연구』(김이겸, 이재운 2007)에서 제시하 고 있는 기본 속성값을 바탕으로 필요에 따라 확장하여 사용하고 있다. 저작이 영화, 드라마, 연극, 노래, 음반 등의 분야인 경우의 기술 방법 을 보면 다음과 같다.

- 예) 오! 브라더스(영화:2003)
[Oh Brothers]
- 골뱅이@실명제(영화:2001)
[--實名制]
- 소년, 소년을 만나다(영화:2008)
[少年少年--]
- 강철중: 공공의 적 1-1(영화:2008)
[강철중:公共--敵1-1]
- 반올림 #3(드라마:2006)[半--]

참조 정보로서 한자 혹은 영어를 부기한다. 그러나 한자나 영어로 표기할 수 없는 경우에는 대시를 반복하여(--) 부기하고 한자가 있 으나 이를 알 수 없는 경우에는 한글을 그대로 사 용한다. 한자와 영어가 함께 있는 경우는 한자 와 영어를 함께 기술한다.

- 예) 겨울 연가(애니메이션:2009)[--戀歌]
- 고사: 피의 중간고사(영화:2008)
[고死--中間考査]

- 송창식 골든1집(음반:1981)
[宋昌植--一集]
- 종합병원 The Movie
천일동안(영화:2000)
[綜合病院The Movie千日--]
- 퍼펙트 맨(음반:2002)[Perfect Man]

잡제가 있거나 용어의 순서가 명확하지 않는 경우에는 가능한 표현을 모두 UF로 둔다.

- 예) 반올림 #1(드라마:2003-2005)[半--]
UF 성장 드라마:반올림1(드라마:
2003-2005)[成長drama:半--1]
- 종합병원 2(드라마:2008-2009)
[綜合病院]
- UF 종합병원 시즌2(드라마:2008)
[綜合病院--]
- 2009 외인구단(드라마:2009)
[--外人球團]
- UF 외인 구단 2009(드라마:2009)
[外人球團2009]
- 강철중:공공의 적 1-1(영화:2008)
[강철중:公共--敵1-1]
- UF 공공의 적 1-1:강철중(영화:2008)
[公共--敵1-1:강철중]

문학 작품과 영화나 드라마인 경우의 기술 내용을 비교하면 다음과 같다.

- 예) 태백 산맥(소설:김사량)[太白山脈]
- 태백 산맥(소설:조정래)[太白山脈]
- 태백 산맥(소설:전광용)[太白山脈]
- 태백 산맥(희곡:함세덕)[太白山脈]
- 태백 산맥(영화:1994)[太白山脈]
- 인간 시장(소설:김홍신)[人間市場]

인간 시장(드라마:1988)[人間市場]	UF 가문의 영광2:가문의 위기(영화:
인간 시장(드라마:2004)[人間市場]	2005)[家門--榮光2:家門--危機]
인간 시장(영화:1983)[人間市場]	가문의 부활:가문의 영광3(영화:2006)
인간 시장 2(영화:1985)[人間市場2]	[家門-復活:家門-榮光3]
인간 시장 3(영화:1991)[人間市場3]	UF 가문의 영광3:가문의 부활(영화:
2004 인간 시장(드라마:2004)	2006)[家門--榮光3:家門--復活]
[2004人間市場]	

저작명은 한글 표기를 우선으로 하되, 한글 표기가 없는 것은 원어 그대로 사용하고 저작 구분과 연도를 부기한다.

예) TRI-ANGLE(앨범:2004)
Do You Know That?(노래:2008)
Honey 꼬시기(곡:2008)[Honey--]

5.2.2 저작명의 동등 관계

저작명의 원제목이 있을 경우 이를 동등 관계로 취급하여 UF로 처리한다. 이때 아래 예의 '유정무정(원제목: 파문)[有情無情]'과 같이 제 1, 제2 한정어를 기술한다.

예) 파문(영화:1968)[破門]
UF 유정무정(원제목: 파문)
[有情無情]
작은 꿈이 꽃필 때(영화:1972)
UF 불타라 청춘(원제목: 작은꿈이꽃필때)
[--青春]

저작명이 속편에 해당하는 경우, 원제목에서는 대표 제목이 속편의 제목이 되고 원제목이 부제가 되는 경우에는 원제목과 속편의 관계를 동등 관계로 취급한다.

예) 가문의 위기:가문의 영광2(영화:2005)
[家門--危機:家門--榮光2]

5.2.3 저작명의 계층 관계

저작명은 개별 저작이 속하는 주제 용어 하위 관계의 사례 관계로 설정한다.

예) 한국 영화[韓國映畫]

NTI 가면(영화:2007)[假面]
NTI 가면 차선(영화:1992)
[可變車線]
NTI 가요 반세기(영화:1968)
[歌謠半世紀]
NTI 가족 회의(영화:1962)
[家族會議]

동일한 이름으로 계속되는 연작 드라마의 경우는 전체 부분 관계(NTP/BTP)로 취급한다.

예) 베스트 극장(1989-2007:드라마)
[BEST劇場]
NTP 7일간의 선택(드라마:1997)
[七日間--選擇]
NTP 가족 수첩(드라마:1987)
[家族手帖]
NTP 건망증(드라마:2007)[健忘症]
7일간의 선택(드라마:1997)
[七日間--選擇]
BTP 베스트 극장(드라마:1989-2007)
[BEST劇場]

5.2.4 저작명의 관련 관계 및 저자 관계

저작명과 저자와의 관계는 관련 관계로 취급하고 관계 패킷을 부기한다 이때 관계 패킷은 “원작/영화와 같은 형식으로 한다.

- 예) 별들의 고향(영화:1974)[--故郷]
 - YEA 1974
 - BTI 한국 영화[韓國映畵]
 - RT 별들의 고향(소설:최인호)
 - [--故郷]영화/원작
 - AUT 최인호(1945-:소설가)[崔仁浩]
 - ENG The Stars Heavenly Home (movie:1974)

저작명과 저자의 관계는 ‘인명의 관련 관계 및 저자 관계’에서 설명한 바와 같이 TIT(LE)과 AUT(HOR)라는 관계 지시 기호로 관계를 설정하고 관계 패킷을 부기한다.

- 예) 사랑방 손님과 어머니(영화:1961)
 - [舍廊房--]
 - AUT 김진규(1923-1998: 영화배우)
 - [金振奎]주연/영화
 - AUT 김희갑(1922-1993: 영화배우)
 - [金喜甲]조연/영화
 - AUT 도금봉(1930-2009: 영화배우)
 - [都琴峰]조연/영화
 - AUT 신상옥(1926-2006: 영화감독)
 - [申相玉]감독/영화
 - AUT 신영균(1928-: 영화배우)
 - [申榮均]단역/출연

5.2.5 저작명의 기타 관계

저작명에는 저작물이 저작된 연도를 표시한다. 단일 년도에 이어진 것은 연도만 표시하고,

드라마와 같이 2년 연속 이어진 것은 시작과 종료 연도를 관계 기호 “YEA”에도 기술한다.

- 예) 백야행:하얀 어둠속을 걷다(영화:2009)
 - [白夜行:--]
 - YEA 2009
 - 천사의 키스(드라마:1998-1999)
 - [天使--kiss]
 - YEA 1998-1999

저작명의 외국어 관계는 가능한 한 모두 기술한다. 한글 제목의 로마자 표기는 ROM에 기술한다.

- 예) 안개 도시(영화:1988)[--都市]
 - ENG City of Mist(영화:1988)
 - ROM Angaedosi(영화:1988)
 - 마고리엄의 장난감 백화점(영화:2007)
 - [Magorium--百貨店]
 - ENG Mr. Magorium’s Wonder Emporium(영화:2007)

저작물이 중간에 변화가 있는 경우는 이전 용어와 이후 용어는 기관/단체명과 같이 LT(이전 용어), PT(이후 용어)로 정의한다. 이 관계 지시 기호는 주제명, 기관/단체명, 문헌명 등이 변경되었을 때 전체 관계를 참조하기 위한 것이다.

- TV 문학관(드라마:1980-1987) [TV文學館]
 - YEA 1980-1987
 - LT TV 문예 극장(드라마:1991-1992)
 - [TV文藝劇場]
 - YEA 1980-1987

- PT TV 문학관(드라마:1980-1987)
[TV文學館]
- LT 신 TV 문학관(드라마:1996-1997)
[新TV文學館]
- 신 TV 문학관(드라마:1996-1997)
[新TV文學館]
- YEA 1996-1997
- PT TV 문예 극장(드라마:1991-1992)
[TV文藝劇場]
- LT TV 문학관(드라마:1999-2003)
[TV文學館]
- TV 문학관(드라마:1999-2003) [TV文學館]
- YEA 1999-2003
- PT 신 TV 문학관(드라마:1996-1997)
[新TV文學館]
- LT HDTV 문학관(드라마:2005-)
[HDTV文學館]

6. 결 론

정보량의 증가와 정보 검색 기술의 향상에 따라 이용자의 요구는 전문의 빠른 검색과 정확한 검색에 초점이 모아지고 있다. 그러나 실제의 검색 시스템에서 전문의 검색은 속도 면에서 어느 정도 만족할만한 수준이지만 정확한 검색 측면은 원활하게 진행되지 못하고 있다. 그것은 다의성의 해소 기술이 아직 미흡하기 때문이다. 다의성을 해소하는 방법에는 여러 가지가 있으나 가장 근본적으로 잘 구조화된 우리말 용어사전이 필요하다. 어떤 알고리즘을 적용하더라도 구조화 데이터 없이 비구조화 데이터를 처리한다는 것은 거의 불가능하거나 결

과에는 한계가 있다.

본 연구에서는 다양한 목적으로 개체명 인식에 사용할 수 있는 가장 기본적인 데이터, 그 중에서도 우리말 저자와 저작이라는 두 가지 고유 명사 집합 내에서 동형 이의어의 기술 방법과 두 집합의 용어 관계의 설정 방법에 한정하고 대중 문화 예술 분야를 중심으로 제시하였다. 제시된 것은 대중 문화 예술 분야의 개인명 2,461명, 기관/단체 223개, 문헌 14,895건, 도합 17,579 용어에 대한 구축 실험 결과이다. 지금까지 구축된 모든 데이터를 이용하여 내부적으로 두 가지 실험을 실행하였다. 하나는 색인어의 다의성을 식별하지 않고 장서를 색인하여 서비스하는 종합 대학 도서관에 대하여 순수하게 시소러스만을 이용하여 기존 색인 데이터베이스를 재색인함으로써 얼마나 기존 색인어의 다의성을 해소할 수 있는가에 대한 실험이었으며, 또 하나는 시소러스만을 이용하여 전문 데이터를 색인하였을 때 다의성이 어느 정도 해소되는가에 대한 실험이었다. 전자는 시소러스에 등재된 용어의 깊이에 따라, 후자는 전문 데이터의 종류에 따라 해소율은 달라졌으나 두 실험 모두 현실적으로 적용 가능할 정도의 괄목할 만한 객관적인 결과를 보였다 (<http://www.ontorus.net>).

보다 광범위한 다의성의 해소 및 개체명 인식을 위해서는 모든 분야의 용어가 망라적으로 집적될 필요가 있다. 이 작업은 반영구적으로 계속되어야 하며 이를 모두 수작업으로 구축하는 것은 무리라고 보아야 한다. 다만, 시소러스의 자동 구축에 대한 연구와 결과를 보면 어떤 알고리즘을 적용하더라도 인간이 구축한 정확하고 정교한 구조화 데이터가 기본 데이터로서

필요하다는 사실이다. 본 시소러스를 위한 실험에서도 새로운 용어의 자동 추가는 시기 상조이며, 기존 용어 관계를 이용하여 새로운 관계의 자동 생성 실험에서도 역시 보조적인 기능에 머물 수밖에 없었다. 전자와 후자 어느 것도 자동 생성된 결과를 인간이 면밀하게 검토하여 명확하게 안정시키지 않고는 다음 단계로

진행할 수 없었기 때문이다. 그러나 역사를 보면 꼭 필요한 것은 아무리 많은 시간을 들여서라도 만들어 오고 있다는 사실이다. 데이터의 공개와 공유도 그 한 가지 방법이다. 여러 가지 측면에서 인간의 지적 노력을 줄일 수 있는 방법들이 강구되어 보다 빠른 용어 메타데이터의 구축이 오래지 않아 실현되리라 생각된다.

참 고 문 헌

- [1] 김이겸, 이재윤. 2007. 『용어데이터 레지스터리 설계 및 발전방안 연구』. 서울: 한국학술진흥재단.
- [2] 김태수, 최석두. 1997. 동형어의 구별을 위한 한글한정어 사용에 관한 연구. 『정보관리학회지』, 14(1): 107-124.
- [3] 박수연. 2002. 『한국어 학습사전에서 동음이의어 구별을 위한 ‘길잡이말(Guide Words)’ 연구』. 석사학위논문, 연세대학교 교육대학원.
- [4] 양경용. 2010. 『국어사전의 미시정보를 이용한 동형 이의어 구별 방안 연구』. 석사학위논문, 한성대학교 대학원, 문헌정보학과.
- [5] 유현경. 2000. 사전에서의 동형어 구별을 위한 새로운 제안: 구분자(distinguisher)의 사용에 대하여. 『사전편찬학연구』, 10(1): 133-159.
- [6] 이재철. 1961. 『주제명표목표』. 한국도서관협회.
- [7] 최석두 외. 2006. 대단위 우리말 온톨로지 구축을 위한 시소러스 개발. 『정보관리학회지』, 23(4): 233-264.
- [8] 통계청. 2007. 『한국표준직업분류』. 2007년 7월 2일 제6차 개정고시(통계청 고시 2007-3호).
- [9] 한국도서관협회. 2003. 『한국목록규칙』. 제4판. 서울: 한국도서관협회.
- [10] 한국정보관리학회. 1998. 『법률분야 관련어집 최종보고서』. 서울: 법원도서관.
- [11] 온토러스 홈페이지. [online]. [cited 2011.10.10]. <<http://www.ontorus.net>>.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Kim, Lee-Kyum, & Lee, Jae-Yoon. 2007. *A Study on Design and Development of Thesaurus*

- Data Registry*. Seoul: Korea Research Foundation.
- [2] Kim, Tae-Soo, & Choi, Suk-Doo. 1997. "A study on Hangeul qualifier for homographic descriptors." *Journal of the Korean Society for Information Management*, 14(1): 107-124.
- [3] Park, Soo-Yeon. 2002. *A Study on Guide Words(Gil-ja-bi-mal, 낱잡이말) to Distinguish Homonyms in Korean Learner's Dictionary*. M.A. thesis, Graduate School of Education, Yonsei University.
- [4] Yang, Gyeong-Yong. 2010. *A Study on Distinguishing Homographs Using Micro-information of Korean Dictionary*. M.A. thesis, Graduate School of Hansung University.
- [5] Yoo, Hyun Kyung. 2000. "A study on the method for solving the homographic problem in Korean." *Studies in Lexicography*, 10(1): 133-159.
- [6] Lee, Jai Chul. 1961. *Subject Headings*. Seoul: Korean Library Association.
- [7] Choi, Suk-Doo, et al. 2006. "Constructing a grand-scale Korean thesaurus as a base for ontology development." *Journal of the Korean Society for Information Management*, 23(4): 233-264.
- [8] Statistics Korea. 2007. *Korean Standard Classification of Occupations*, Statistics Korea Notice 2007-3.
- [9] Korean Library Association. 2003. *Korean Cataloguing Rules*. 4th ed. Seoul: Korean Library Association.
- [10] Korea Society for Information Management. 1998. *Legal Thesaurus: Final Report*. Seoul: Supreme Court Library of Korea.
- [11] Ontorus. [online]. [cited 2011.10.10]. <<http://www.ontorus.net>>.