

# 국립중앙도서관 주제명표목표의 고품질화 방안에 관한 연구\*

## A Study on Improvements for High Quality in National Library of Korea Subject Headings List

최 윤 경 (Yoon Kyung Choi)\*\*

정 연 경 (Yeon-Kyoung Chung)\*\*\*

### 목 차

- |                   |                   |
|-------------------|-------------------|
| 1. 서 론            | 4. 주제명표목표 고품질화 방안 |
| 2. 주제명표목표의 개념과 사례 | 5. 결 론            |
| 3. 주제명 데이터 분석 결과  |                   |

### 초 록

본 연구의 목적은 2002년에 개발된 『국립중앙도서관 주제명표목표』의 용어 품질을 검토하여 개선 방안을 제안하는 것이다. 이를 위해 문헌 연구와 국내외 국가도서관의 주제명표목표 사례 조사, 국립중앙도서관 주제명표목표의 주제명 및 관계 데이터 분석을 수행하였다. 본 연구 결과, 주제명 기술의 비일관성, 불분명하고 불필요한 관계 설정, 계층 관계 적용의 비일관성, 최신성 부족, 분류기호의 부정확한 기입, 신규 주제명의 중복 신청, 비우선어 및 미등록어 부여와 같은 문제점이 나타났다. 주제명표목표의 문제점을 중심으로, 주제명 정비, 서지 데이터 정비, 주제명표목 검토위원회 운영, 주제명 지침 보완을 고품질화 방안으로 제시하였다.

### ABSTRACT

The purpose of this study was to propose the improvements for high quality of *National Library of Korea (NLK) Subject Headings List* which was developed as a subject access tool in 2002. For this study, literature review, case study for subject headings lists of national libraries, and data analysis of headings and relationships were performed. Several problems were found as follows: inconsistency of subject headings descriptions, unclear and unnecessary relationships among headings, inconsistency of hierarchical relationships, lack of currency, incorrectness of classification numbers, duplication of newly requested terms, and assignment of non-preferred and unregistered terms in bibliographic records. Focusing on these problems, modifications of subject headings and bibliographic records, installation and operation of a review committee on subject headings, and supplementation of a manual were suggested to improve NLK Subject Headings List.

키워드: 주제명표목, 국립중앙도서관, 통제어, 주제 접근, 용어 품질

Subject Heading, National Library of Korea, Controlled Vocabulary, Subject Access, Terminology Quality

\* 본 연구는 2013년도 국립중앙도서관의 지원을 받아 연구되었음.

\*\* 이화여자대학교 일반대학원 문헌정보학과 박사과정(yunee20@gmail.com) (제1저자)

\*\*\* 이화여자대학교 사회과학대학 문헌정보학과 교수(ykchung@ewha.ac.kr) (교신저자)

논문접수일자: 2014년 1월 9일 최초심사일자: 2014년 2월 4일 게재확정일자: 2014년 2월 10일

한국문헌정보학회지, 48(1): 75-95, 2014. [http://dx.doi.org/10.4275/KSLIS.2014.48.1.075]

## 1. 서론

### 1.1 연구의 배경 및 목적

오늘날 대부분의 이용자들은 주제 검색 대신 키워드 검색으로 정보원에 접근하는 경향이 있는데, 키워드 검색은 동의어, 이형, 복수형, 특수문자, 축약어, 숫자 등 자연어 요소들이 검색 결과의 정확률과 재현율을 저해시킨다. 그래서 단순 검색에서는 키워드 검색이 성공적일 수 있으나, 좀 더 정교하고 집중된 탐색을 제공하면서 검색 결과의 정확률과 재현율을 높이려면 전문적인 주제 접근이 필요하다.

도서관에서는 이용자의 주제 접근을 도와주기 위하여 주제명표목표, 분류체계, 시소러스 등 다양한 도구들을 개발·유지해오고 있다. 이 중 주제명표목표는 전통적으로 도서관의 소장 자료에 대한 색인과 검색에 활용되어 왔으며, 이용자의 주제 기반 정보 탐색을 지원하는 핵심 도구가 되어왔다. 주제명표목표의 표목으로 선정되는 단어나 구는 미리 통일된 원칙에 따라 선정하고, 표목을 부여하기 때문에 주제명표목표는 통제어 체계 기반이라 간주한다.

최근 문헌정보학 분야에서 고비용의 주제 접근 도구에 대한 필요성과 가치에 대해 의문을 제기하기도 하지만, 여전히 통제어는 동의어와 이형어, 연관어를 제공함으로써 검색 효율성과 효과성을 증진시키는 역할을 한다(Chan and O'Neil 2010). 따라서 주제명표목표는 도서관 소장 자료는 물론, 웹 자원으로까지 적용이 가능하고, 일단 구축된 주제명은 상호 공유가 가능하므로, 양질의 주제명표목표는 단순히 도서

관 소장 자료를 넘어 국가 지식 정보의 활용을 위한 강력한 주제 접근 도구로 활용될 수 있다. 그러나 주제명표목표는 기본적으로 구축 및 유지에 많은 비용이 소요되기 때문에, 개별 도서관보다는 국가도서관이 주축이 되어 개발되어야 한다.

우리나라에서도 이러한 필요성에 의거하여, 국립중앙도서관은 1992년 작성된 '주제명 검색 요어집'을 바탕으로 2002년에 전 주제 분야를 대상으로 한 시소러스형 후조합 색인언어 구조로 『국립중앙도서관 주제명표목표』를 개발하여 표준화된 접근점을 제공하고자 하였다. 그러나 학문의 분화와 융합에 따른 새로운 용어의 등장과 기존 개념의 세분 및 변화 등으로 기존 주제명의 갱신 요구가 생겼는데 2007년 이후 용어의 추가 및 수정이 거의 이루어지지 않았다. 또한 링크드 데이터의 구축 및 활용에 대한 논의가 본격화되면서 전 세계적으로 주제명표목표의 개방과 공유 요구도 증대되고 있다. 따라서 국립중앙도서관은 양질의 주제명표목표를 구축·유지함으로써 국내 도서관의 주제 접근 인프라를 지원하고, 글로벌 정보 환경에 적극적으로 대응할 수 있어야 한다. 그러므로 본 연구의 목적은 국립중앙도서관 주제명표목표의 용어 품질을 전면적으로 검토하여, 실질적인 개선 방안을 제안하는 것이다.

### 1.2 연구 방법

본 연구는 국립중앙도서관 주제명표목표의 용어 고품질화 방안을 제안하고자 문헌 연구와 사례 조사, 실사 분석, 전문가 자문회의를 수행하였다. 첫째, 문헌연구에서는 주제명표목표의 역

할과 기능에 대해 고찰하였다. 둘째, 사례 조사에서는 국립중앙도서관 주제명표목표와 외국 국가도서관의 주제명표목표를 분석하였다. 외국 사례로는 미국 의회도서관 주제명표목표(Library of Congress Subject Headings List, LCSH), 일본 국립국회도서관건명표목표(NDLSH, 國立國會図書館件名標目表), 2개를 선정하여 주제명표목표의 구조 및 특징, 구축범위 및 수록건수, 관리방식 등을 조사하였다. 셋째, 실사 분석에서는 국립중앙도서관 주제명표목표의 용어 및 관계 데이터를 입수하여 주제명 구축 및 이용 현황과 용어 정보 입력 현황, 표본 데이터 용어 품질을 조사하여, 문제점을 도출하였다. 주제명 구축·이용 현황에서 주제명 데이터 구축 및 이용 건수, 비활용어 현황 등을 조사하고, 용어 정보 입력 현황에서는 항목별 입력 통계치를 분석하여, 핵심 입력 항목과 저빈도 입력 항목을 도출하였다. 마지막으로 표본 데이터 품질 조사에서는 활용 빈도를 기준으로 고빈도 용어 1,124개와 저빈도 용어 285개, 총 1,409개를 대상으로 주제명과 그 관계를 분석하고, 1회 이상 서지 데이터에 부여된 용어 45,830개를 대상으로 비우선어와 미등록 용어 실태를 조사하였다. 그리고 2005년부터 2013년 7월까지 신규 신청된 용어 3,607건의 특성도 함께 파악하였다. 넷째, 전문가 자문회의에서는 국립중앙도서관 주제명표목표 고품질화에 관한 검토 의견을 수렴하였다. 전문가는 국내 주제명표목표 및 시소러스 구축·관리 실무자와 학계 전문가 3명으로 구성하였고, 자문회의에서는 국립중앙도서관 주제명표목표의 문제점을 공유하고, 품질 개선 방안에 대한 의견을 수렴하였다. 이상의 연구 내용을 종합하여 국립중앙도서관 주제명표목표의

품질 개선 방안을 제시하였다.

### 1.3 선행연구

국내외로 주제명표목표에 관한 연구는 오랫동안 다양한 분야에서 이루어져왔다. 국내의 경우, 급변하는 정보환경에서 주제어의 역할이 더욱 중요해지고 있는 상황 하에 기존 주제명표목표에서 나타난 다양한 문제점을 분석하고 기본 구조를 변형하여 제안하는 연구가 수행되었다(한상완 등 1995; 김태수 1997; 황면 등 2002). 그리고 전 세계적으로 가장 이용률이 높은 LCSH를 대상으로 본질적인 특성과 구조에 대한 연구가 이루어졌다(김주성, 김태수 1995; 윤혜영 2003; 유영준 2009). 이들은 주로 LCSH의 표목 구조와 시소러스의 용어 구조와 색인 구조를 비교 분석하거나, LCSH가 가진 구조적 문제점을 제기하였다. 이외에도 LCSH나 다른 국가의 주제명표목표를 대상으로 한국 관련 주제명을 비교·분석하거나 변천사를 제시한 연구가 이루어졌다(윤정옥 2001; 전창호 2004; 남태우 등 2008; 김정현, 문지현 2009; 정연경 2010; 김정현 2010; 2011; 2013).

그러나 국립중앙도서관 주제명표목표에 대해서는 2002년 구축 이후 연구가 진행되지 않다가 FRSAD(Functional Requirements of Subject Authority Data)를 FAST의 패킷과 연계하여 국립중앙도서관 주제명표목표에 시범 적용한 결과, 주제명표목을 하위구분하고 기존의 의미관계를 더욱 명확히 할 수 있는 가능성을 제시한 연구가 수행되었다(박지영 2011).

외국의 경우, 미국 국가도서관의 주제명표목

이자 전 세계적으로 활용되고 있는 LCSH에 대한 연구가 온라인 목록에서 주제 접근 도구로서 주제명 표목과 LCSH의 효용성에 관해서 먼저 이루어졌다(Frost 1985; Carlyle 1989; Fernandez 1991; Vizine-Goetz and Drabenstott 1991). 2000년 이후부터는 LCSH의 가치와 중요성을 재조명하면서 새로운 정보 환경에 적합한 구조를 모색하고, 문제점과 개선방안을 제시하는 연구들이 수행되었다(Chan and Hodges 2000; Mann 2000; Wool 2000; Fischer 2005; Anderson and Hoffmann 2006; Yi and Chan 2010; Walsh 2011). 이러한 연구는 대부분 LCSH가 현대의 정보검색 기술과 방식에 맞는 구조로 개선되어야 할 필요를 주장한 것이었다. 이와 같은 맥락에서 LCSH의 단점을 극복하고자 OCLC의 주도로 시작된 LCSH를 패싯화한 FAST에 관한 연구가 많이 진행되었다(Dean 2004; Jin 2008; Chan and O'Neill 2010).

이상의 국내외 선행 연구를 보면 외국은 LCSH를 중심으로 문제점을 제기하고 이에 대한 개선 방안을 모색하는 연구들이 지속적으로 이루어졌다. 또한 LCSH의 패싯화된 구조인 FAST의 개발 및 적용에 관한 연구가 최근까지 지속적으로 발표되고 있다. 그리고 통제어휘 도구 뿐 아니라 저자나 이용자가 부여하는 주제어의 상호 관계, 주제어 검색의 효율성 증진을 위한 시스템과 인터페이스 등 주제 접근에 대한 다양한 연구가 수행되고 있다. 그런데 국내의 경우는 주제명표목표 개발에 관한 연구가 최근 거의 진행되지 않았으며, 국립중앙도서관 주제명표목표가 구축된 후로는 이에 대한 직접적인 연구가 거의 이루어지지 않았다.

## 2. 주제명표목표의 개념과 사례

### 2.1 주제명표목표의 개념과 구조

#### 2.1.1 정의와 역할

『문헌정보학용어사전』에 따르면, 주제명표목이란 주제명목록에서 저작의 주제를 표현하는 표목이 되는 단어나 구를 지칭한다. 그래서 주제명은 일반적인 주제어 이외에도 자료의 주제로 다뤄질 수 있는 개인명, 단체명, 회의명, 통일서명, 지명까지 포괄한다. 주제명표목표는 사서가 통제어를 생성·부여함으로써 저자가 사용한 '주제어'와 이용자가 검색하는 '주제어'를 연결시키는 일종의 매개체 역할을 한다. 즉, 사서의 '색인' 작업과 이용자의 '검색' 과정에서 용어 통제 역할을 수행하며, 부가적으로 주제명 간의 의미 관계를 통해 다양한 용어정보를 제공한다(국립중앙도서관 2002, 87-89). 또한 동일한 개념을 표현하는 동의어나 유사어를 함께 모아서 동일 주제에 관한 자료를 모아주는 기능을 한다. 기본적으로 통제어의 사용 목적은 이용자 검색어가 가리키는 주제와 관련된 자료를 많이 전달하고, 부적절한 검색 결과를 최소화시키는 것이다. 또한 동일한 용어가 다른 주제 분야나 맥락에서 다양한 의미로 사용되는 경우 한정어라 부르는 괄호 안에 용어로 맥락화시켜서 가능한 모호함을 배제시킬 수 있다. 즉, 주제명표목표를 활용한 주제 접근은 높은 재현율을 가진다는 점에서 부적절한 자료를 최소화시키는 정확률도 잘 달성할 수 있다.

#### 2.1.2 표목의 구조와 선정

주제명표목은 보통 하나 이상의 단어로 이루어

어지는 명사류로, 단일어 표목과 복합 표목으로 구분된다. 단일어 표목은 단일 개념을, 전조합에 의한 복합 표목은 2개 이상의 독립 개념을 접속 사나, 세목, 도치 등의 방법으로 조합되어 표현한다. 전통적인 주제명표목표는 주표목과 세목, 참조로 구성되는데, 주표목은 세목 없이 기본개념을 표현하며 성격에 따라 주제 표목, 형식 표목, 고유명사 등이 있다. 세목은 중심 주제를 세분하는 기능으로, 주표목의 형태와 관점을 표현하며 성격에 따라 주제 세목, 형식 세목, 지리 세목, 시대 세목 등이 있다. 주표목에서 동형어의 어를 구분하거나 불분명한 용어의 의미를 표현하기 위해 한정어를 사용한다. 그리고 주제명표목을 상호 연결하기 위한 장치로 전통적인 주제명표목표에서는 보라 참조(See Reference)와 도보라 참조(See Also Reference)가 있다. 보라 참조는 표목으로 채택되지 않은 주제명에서 채택된 주제명표목으로 연결하는 것으로 주로 동의어와 유사어 간을 연결하는데 사용되며 도보라 참조는 특정표목과 관련된 다른 표목 간을 연결하기 위한 것으로 주제의 상호관계와 종속관계를 표현하는데 사용된다(최달현, 이창수 2010, 247-258).

한편, 주제명표목표는 어떠한 용어를 표목으로 구성하느냐에 따라 효용성과 가치가 결정되기 때문에 표목의 선정이 매우 중요하다(최달현, 이창수 2010, 244-246; Chan and O'Neill 2010, 7-9). 주제명표목의 선정 기준으로는 첫째, 이용자의 요구와 최신성을 반영하며, 둘째, 동의어와 유사어를 통제할 수 있는 용어로 통일되고, 주제명표목표 내에서 고유한 표목이 마련되어야 하고, 셋째, 표현 형식이 통일되어야 한다는 것이다. 이러한 주제명표목표의 효용성을 높이

고, 용어 선정과 유지·관리에 일관성을 기하기 위해서 대부분의 주제명표목표는 주제명표목의 선정과 기술에 관한 세부 규칙과 규정을 명문화하여 정의하고 있다.

## 2.2 국내외 주제명표목표 사례

### 2.2.1 한국의 국립중앙도서관 주제명표목표

#### (1) 표목 구조

국립중앙도서관 주제명표목표는 단일 디스크립터, 즉 세목 없이 기본 개념을 표현하는 주표목으로만 이루어진다. 그리고 주표목은 단일어와 복합어를 모두 포괄하고, 한정어는 동음이의어를 식별하기 위해 사용되며 원괄호 안에 기입하고, 용이한 식별을 위해 한자를 각괄호 안에 기술한다(국립중앙도서관 2002, 34-37). 참조는 동의어 관계(USE/UF), 계층관계(BT/NT), 연관관계(RT) 등 시소러스의 기본 관계를 중심으로 계층 관계를 속성에 따라 속(屬, generic), 사례(instance), 부분(partial)으로 세분하고 각각 BTG/NTG, BTI/NTI, BTP/NTP로 정의하였다. 용어 간의 관계성을 식별하는 참조 이외에도 우리나라 실정을 감안하여 북한어(NK)와 북한어에 대응되는 국어(SK)에 관한 참조, 분류기호(KDC, DDC)나 범위주기(SN)와 같은 용어의 부가 정보에 관한 참조도 있다. 특히 영어(ENG), 스페인어(ESP), FRA(불어), 독일어(GER) 등 외국어 대응 관계가 매우 다양하게 마련되었다(국립중앙도서관 2002, 6-8).

#### (2) 구축범위 및 수록건수

국립중앙도서관 주제명표목표는 2002년 개발된 이래, 일반 주제명을 중심으로 일부 개인명,

단체명, 지명 등이 구축되었다. 일반 주제명을 기준으로 560,563건이 수록되었으며, 현재는 국내 단행본에만 주제명을 부여하고 있다(국립중앙도서관 2009: 2013, 11). 주제명표목표에 대한 인쇄본은 발행되지 않고 국립중앙도서관 디브리리 검색에서 활용할 수 있다.

### (3) 관리방식

국립중앙도서관 주제명표목표는 2002년도에 외부 용역을 통해 약 20만건을 구축한 이후, 2006년 약 30만건을 추가되었으나, 그 이후 유지 및 관리가 제대로 이루어지지 않았다. 관련 지침으로는 여섯 장으로 이루어진 간략한 『주제명표목표 매뉴얼』이 있으며, 7개 부문인 부여대상, 사용 tag, 부여범위, 주요 용어 관계지시 기호, 주제명 신규용어 부여 제외대상, 주제명 부여 방법, 주제명요청기술 규칙(요약)을 다루고 있다(국립중앙도서관 2009).

## 2.2.2 미국의 LCSH

### (1) 표목 구조

LCSH의 표목은 주표목과 세목, 참조, 기타로 구성된다(정연경 2010). 첫째, 주표목은 하나 이상의 단어로 이루어지고, 두 단어 이상의 표목은 '형용사 + 명사', 접속사(and) 연결 혹은 전치사구로, 도치 표목인 경우는 언어나 국적 등을 기술하는 형용사로 시작하는 단어이다. 복합어가 표목으로 지정되기도 하는데 2개 표목이 매우 유사하여 저작에서 함께 자주 다루어지는 경우이다. 한정어는 동음이의어나 용어의 문맥상에 고유한 의미를 명확히 설정하기 위해 원괄호 안에 기입한다.

둘째, 세목은 주제의 특수한 측면들을 표현

하는 것으로, 표목의 구체성을 높이는데 사용되었으며 긴 줄표(-)로 식별한다. 성격에 따라 주제 세목, 형식 세목, 시대 세목, 지리 세목이 있으며, 특이하게 자유부가세목(free-floating subdivision)은 분류표의 조기표와 유사한 기능을 하며, 일부 주제와 형식 세목을 표준 집합으로 구성하여 가능한 표목에 모두 조합할 수 있다.

셋째, 참조는 동의어 관계(USE/UF), 계층관계(BT/NT), 연관관계(RT) 등 시소러스의 기본 관계를 일부 적용하고 도보라(See Also, SA) 참조도 허용한다. 용어 간의 관계성을 식별하는 참조 이외에도 분류기호나 범위주기와 같은 용어의 부가 정보에 관한 참조도 있다. 분류 기호에는 미의회도서관분류기호를 부여하며, 각괄호([ ]) 안에 기술한다. 현재 87,500건의 주제명에 분류기호가 부여되어있으며 범위주기는 목록 레코드에 주제명을 일관성 있게 부여하기 위하여 주제의 범주를 명시하고, 관련 표목들 간의 차이점을 기술하며, 표목의 의미가 도서관 목록에서 사용이 제한되는 경우를 설명하는데 사용된다(Library of Congress 2012).

### (2) 구축범위 및 수록건수

LCSH는 1914년 초판이 발행되어, 2013년에 인쇄본이 35판까지 발행되었는데 여기에는 332,500건의 주제명을 수록하고 있으며, 2012년의 34판에 비해 약 5,000건의 신규 주제명이 추가되었다. 제 34판 기준, 329,000건의 주제명을 수록하고 있으며, 755건의 USE 참조 등을 포함하고 구축 범위 및 수록 건수는 인명(23,200건), 단체명(8,800건), 회의명(10건), 통일서명(480건), 지리명(58,300건), 일반 주제명(234,000건)이었

다(Library of Congress 2012).

### (3) 관리방식

LCSH는 미의회도서관의 협동목록프로그램(PCC, Program for Cooperative Cataloging)의 일환으로 SACO(Subject Authority Cooperative Program)가 주로 새로운 표목 제안 및 개정에 참여하고 있으며, 분야별로 12개의 주제그룹(funnel)을 구성하고 있다. SACO 참여 도서관은 신규 주제명의 제안과 기존 주제명의 변경에 핵심적인 역할을 하며, SACO Proposal system을 이용해 신청할 수 있다. SACO에 참여하지 않는 도서관도 신규 주제명 신청이 가능하며, 표목명, UF, BT, RT, 해당 표목을 판단할 수 있는 참고 정보를 신청서에 기입해야 한다. SACO 참여자가 제안서를 제출하면 위원회가 검토하여 최종 결정하며 신규 및 변경 표목에 대한 정보는 주간으로 서비스하는 『Library of Congress Subject Headings Weekly List』에 반영된다(Library of Congress 2012).

관련 지침으로는 Subject Heading Manual(SHM, 2008)과 SACO Participant's Manual(2007)이 있다. SHM의 내용 구성을 살펴보면, H0000은 LCSH의 개략적 소개와 특성, H0100 표목의 생성과 부여, 주제 전거 레코드의 변경 및 삭제에 관한 사항, H0200 신규 주제명의 제안, 표목의 형식에 관한 사항, H0300은 한정어, 세목, 참조에 관한 사항, H0400은 범위주기, 인명 및 단체명에 관한 사항, H0600은 세목에 관한 사항, H0700~H1000은 지리명과 자유부가세목에 관한 사항, H1100~H2200 주제별 또는 자료유형별 적용 사항 등으로 매우 상세하다.

## 2.2.3 일본의 NDLSH

### (1) 표목 구조

NDLSH의 표목은 주표목과 세목, 참조, 기타 정보로 구성된다. 주표목은 하나 이상의 단일어와 복합어, 병렬어, 구형식 등으로 다양하고, 띄어 쓰지 않는다. 한정어는 의미가 모호하거나, 동일한 용어를 여러 분야에서 사용하는 경우에 주제명의 사용 범위를 명확히 하기 위해 원괄호 안에 기입한다. 이외에도 예술 분야의 작품 부기, 사건 이름, 전쟁 이름 등의 연대 부기, 기타 참고(범위 주기로의 안내)로 지시하는 경우에 사용된다. 세목은 지시가 있는 경우, 해당 주제명을 세목으로 다른 주제명에 결합하여 사용할 수 있으며 이중불입표(--로 식별한다. 세목의 성격에 따라 주표목과 세목에 모두 사용할 수 있는 것과 세목으로만 사용할 수 있는 것으로 구분하고, 세목의 내용에 따라 지명 세목, 지명 결합 세목, 시대 세목, 주제 세목, 형식 세목, 특수 세목으로 총 6개 유형이 있다(김정현 2013).

참조는 보라 참조와 도보라 참조로 크게 나누어, 보라 참조는 시소러스의 USE-UF 관계와 유사한 것으로, 동일한 개념에 대한 다른 표기 형태, 개념에 차이는 있지만, 구분하기가 어렵거나 구분 필요성이 매우 적은 것, 약자, 약어 중 널리 알려진 것, BSH(기본 주제명표목), LCSH 표목을 기입하였다. 도보라 참조의 경우, 2004년부터 계층관계(BT/NT)와 연관관계(RT)를 입력하였다.

부가 정보로는 국립국회도서관분류(NDLC)와 일본십진분류(NDC) 신정 9판에 해당하는 분류기호를 부여하였다. 범위주기도 다양하게 사용하였는데, 주요 역할은 (1) 해당 주제명의 사용 범위 제한, (2) 유사한 주제명과 구분을 위

한 설명, (3) 주 표목 및 세목 구분 지시, (4) 주제명 적용 시 유의해야 할 사항 등이다. 또한 관련 링크에는 해당 주제명에 상응하는 LCSH와 일본도서관협회가 간행한 BSH의 ID를 기재하고, 출처에는 표목 생성의 근거가 되는 저작의 제목이나 표목 형태 확정에 참고한 근거 자료 등을 기재하였다(國立國會圖書館 2013).

### (2) 구축범위 및 수록건수

NDSLH는 1964년부터 구축되었으며 2012년 말 기준, 일반 주제명 구축 건수는 102,242건(세목 포함)으로 전체 전거 데이터 1,111,018건의 약 9.2%를 차지하였다. 주표목은 개인명이나 단체명과 같은 고유명 표목은 수록하지 않는 것이 원칙이나, 일반 주제명과 더불어 개인명, 가족명, 단체명, 지명, 통일서명 등도 일부 인정하였다. 1949년부터 현재까지 국립 국회 도서관 소장 도서 및 서양서에 적용하고 있다(김정현 2013: 國立國會圖書館 2013).

### (3) 관리방식

NDSLH는 국립국회도서관 수집서지부에서 주제명 신설 및 수정, 삭제, 부여 등 관리 전반을 담당하며 외부기관과의 협력 구축은 하지 않는다. 관련 지침으로는 『국립국회도서관명명작업지침서』가 2010년 초판, 2013년 7월에 개정판이 발행되었다. 온라인으로 이용 가능하며 수록 내용은 일반 주제명과 더불어 개인명, 단체명, 통일서명 등 고유명 전거에 관한 지침으로 209 페이지에 달한다. 내용상 크게 두 개 부분으로 구분되며, 전반부는 주제명 신설 및 부여에 관한 원칙과 기준을, 후반부는 주제별 작업 지침을 수록한다. 구체적으로는 전반부에 주제명 부여

원칙, 주표목에 적용되는 공통 지침, 주제명 유형별(개인명, 가족명, 단체명, 지명, 통일서명, 일반 주제명) 신설 및 부여 기준과 특성, 세목 등을, 후반부에 NDC 기준으로 0 종류부터 9 문항까지 주제별로 적용 지침을 제시하고 있다(國立國會圖書館 2013).

## 3. 주제명 데이터 분석 결과

### 3.1 분석 내용 및 대상

국립중앙도서관 주제명표목표의 품질을 조사하기 위해 2013년 7월까지 입력된 주제명 및 관계 데이터를 수집·분석하였다. 분석 내용은 국립중앙도서관 주제명표목표의 전체 현황을 비롯하여 주제명의 속성 및 관계, 비활용 주제명과 신규 신청 주제명의 특성, 비우선어 및 미등록 용어 입력 실태를 다음과 같이 파악하였다. 첫째, 전체 현황은 전체 수록 용어 수와 유형, 관계 수와 유형을 분석하였다. 둘째, 주제명의 속성 및 관계에서는 서지 데이터 650 \$a에 100회 이상 부여된 주제명 1,124개, 관계 데이터 30,997건을 분석하였다. 셋째, 비활용 주제명은 현재까지 한 번도 부여된 적이 없는 주제명 중에서 주제명 285개를 추출하고, 이에 대한 관계 데이터 1,144건을 분석하였다. 넷째, 신규 신청 주제명은 2005년 이후 신규 신청된 용어 데이터 3,607건을 분석하였다. 다섯째, 비우선어와 미등록 용어 입력 실태는 서지 데이터 650 \$a에 1회 이상 부여된 주제명 45,830개를 중심으로 파악하였다.

### 3.2 전체 현황

#### 3.2.1 용어

국립중앙도서관 주제명 데이터베이스에 구축된 용어는 총 560,563개로 우선어(272,623건, 48.6%), 외국어(외국어 대응 코드에 입력된 용어, 173,498건, 31.0%), 비우선어(66,237건, 11.8%), 분류어(KDC, DDC에 입력된 용어, 49,205건, 8.78%) 순이었다. 다음으로 우선어 272,623건에 대하여 2003년부터 2013년 7월까지 서지 서지 데이터에 부여된 이용 건수를 기준으로 1회 이상 활용된 주제명이 45,830개(16.81%), 비활용 주제명은 232,390개(85.24%)이었으며 이 중 100건 이상의 서지 데이터에 활용된 주제명은 1,124개(0.41%)에 불과하였다.

#### 3.2.2 관계

국립중앙도서관 주제명표목표 데이터베이스가 수록한 관계 유형은 97개로 전체 용어 관계 수는 1,600,639건이었다. 관계 유형별로는 시소러스 기본 관계(RT, BT, NT)의 분포가 가장 높았고, 세부적으로는 RT(관련어), BT(상위 개념어), NT(하위 개념어), KEN(영어에 대응하는 한글 용어), ENG(영어), KDC(KDC 분류 4판), DDC(듀이십진분류 21판), USE(~을 사용하라), UF(~대신 사용하라), WLL(용어 수준), UNS(UNSPSC 상품분류표) 순이었다. 그러나 E.G(용례), JAP(일본어), KTH(태국어에 대응하는 한글용어), ND(미판정어), POS(품사), SLO(슬로바키아어), THA(타이어) 관계는 한 번도 부여된 적이 없었다.

### 3.3 주제명의 속성 및 관계

#### 3.3.1 주제명 속성

2003년부터 2013년까지 서지 데이터 650 \$a에 100회 이상 부여된 주제명 1,124개를 중심으로 주제명의 속성을 분석한 결과 단일 개념어(906, 80.6%)를 비롯하여 2개 이상의 단어로 이루어진 복합어(218개, 19.4%)가 많았으며, 복합어의 경우 일반 주제성 용어에 지리, 연대, 사건, 형식 및 장르 등 다양한 속성을 가진 용어들이 결합되었다. 예를 들어 한국 요리는 '요리(주제적 속성)'와 '한국(지리적 속성)', 한국 현대 시는 '시(주제적 속성, 형식 및 장르적 속성)' + '한국(지리적 속성)' + '현대(연대적 속성)'와 같은 속성으로 조합된 형태였다.

#### 3.3.2 관계 설정

2003년부터 2013년까지 서지 데이터 650 \$a에 100회 이상 부여된 주제명 1,124개의 관계 데이터 30,997건을 분석한 결과, 관계 유형은 총 28개로 주로 기본 관계였다. 구체적으로는 RT(13,885건, 1,015개 주제명), NT(9,786건, 711개 주제명), BT(1,175건, 84개 주제명), DDC(1,248건, 995개 주제명), KDC(1,244건, 956개 주제명), ENG(1,011건, 651개 주제명) 순이었다. 10개 미만 관계를 가진 주제명(465개)이 가장 많았으며, 10개 이상 30개 미만(357개), 30개 이상 50개 미만(123개) 순이었고, 이 중 가장 많은 관계를 가진 주제명은 '교육[教育]'으로 390개의 관계가 설정되었다.

##### (1) BT 관계

834개 주제명에 1,175건의 BT 관계가 있었으

며, 주제명 1개 당 평균 1.41건이었다. BT는 학문적 상위 개념으로 설정된 경우가 많았으며 복합어인 경우, 제 2요소를 BT로 설정하였다. 예를 들어 '기독교 교육 교재'는 '교육 교재'를, '교통 정책'은 '정책'을 BT로 설정하였다. 1개 BT를 가진 주제명(604개)이 가장 많았으며, 2개(161개), 3개(46개), 4개(11개) 순으로 3개 이상 BT를 가진 주제명은 69개 8.27%, 최다 9개 BT를 가진 주제명은 '물(액체)[水]'이었다.

(2) NT 관계

670개 주제명에 9,786건의 NT 관계가 있었으며, 주제명 1개 당 평균 14.61건이었다. NT는 주로 국가/지역별, 시대, 인종 등으로 세분되어 하위 개념어를 구성하였고, 기독교, 문학, 경제학 등 학문명에 관한 주제명은 NT가 매우 상세하게 전개되었다. 5개 미만의 NT를 가진 주제명(282개)이 가장 많았으며, 5개 이상 10개 미만은 139개, 100개 이상의 NT를 가진 주제명은 11개, 최다 199개 NT를 가진 주제명은 '교육[教育]'이었다.

(3) RT 관계

1,015개 주제명에 RT 관계가 13,885건이 포함되었으며, 주제명 1개 당 평균 13.70건이었다. RT는 주로 해당 주제명이 포함된 용어, 관련된 정책, 인물, 법률(특정 법률명 포함), 책(집), 기관, 시설, 직업, 이론 및 개념, 원인, 사건, 시험, 학문, 운동(사상), 시대 등이었다. 5개 미만의 RT를 가진 주제명이 451개로 가장 많았으며, 5개 이상 10개 미만은 201개였다. 100개 이상의 RT를 가진 주제명은 15개였는데, '정치(권력)[政治]', '경제(경제학)[經濟]', '교육[教育]'

과 같은 주제어였다.

### 3.4 비활용 주제명

2003년부터 2013년까지 서지 데이터 650 \$a에 한 번도 부여된 적 없는 주제명 285개와 관계 데이터 1,144건을 분석한 결과, 용어 측면에서는 '가자미 식혜', '감압 복사기', '고체 수축' 등 전문적이거나 특수한 주제명이 많았다. 관계 유형은 총 20개로 외국어 대응 코드가 조금 적었으며, 비활용 주제명 1개당 평균 관계 수는 4.01건이었다.

한편, 전체 주제명표목 중 비활용 주제명이 80% 이상에 달하고 있으나, 이는 현재 단행본만을 대상으로 주제명표목을 부여하므로 특수성이 높은 주제명의 활용성이 낮은 수준에 머물고 있기 때문이다. 앞으로 주제명표목 부여 대상이 전문 학술 논문 등 보다 광범위한 자원으로 확대 적용될 때에는 비활용 주제명의 활용성이 증가할 가능성이 높다. 더 나아가 국립중앙도서관 주제명표목표를 개방·공유한다면 이러한 비활용 주제명은 주제 전문도서관에서 보다 높은 활용성을 보이게 될 것이다. 그러므로 비활용 주제명의 잠재적 가치를 인식하고 활용 가능성을 열어 두어야 할 것이다.

### 3.5 신규 주제명

신규 주제명 요청 용어 3,607건(2013년 7월 31일 기준) 중 시스템 반출 오류로 인한 중복 레코드를 제거한 후, 총 3,136건을 분석하였다. 분석 결과, 연도별 신규 주제명 요청 용어 수는 2007년 1,294건으로 가장 많았고, 최근 3년 간

요청 건수는 평균 75.7건이었다. 신규 신청 주제명의 사례를 유형화하면 첫째, 2002년 월드컵과 같은 특정 사건이나 행사명, 둘째 학술 용어와 이론에 대한 용어로 Ajax 기법(전산), DNA 칩, NIE 교육, 복잡계, 산화질소합성효소 등이 해당되었다. 셋째, 직업 또는 직종, 자격증, 자격 시험에 관련된 용어로 ERP 정보 관리자, 법학 적성 시험, 병원 코디네이터 등이 있었다. 넷째, 시사 상식과 정치·경제·문화 분야의 용어로 K-POP, 건강 도시, 고령 사회, 소셜 네트워크, 블루오션 등이 해당되었다. 다섯째, KORMARC 통합서지용과 같은 표준 관련 용어, 여섯째, 가족 관계 등록법, 국가 균형 발전, 납세자 보호관 제도, 녹색 성장 등 정부 정책이나 법령(법률)에 관한 용어였다. 일곱째, 개성공업지구와 같은 지명과 여덟 번째, 교육 에세이, 건강 에세이, 산악 문학 등 새로운 장르 문학에 관한 용어였다.

### 3.6 문제점 및 개선방안

국립중앙도서관 주제명 및 관계 데이터 분석 결과, 다음과 같은 문제점을 파악하였다.

#### 3.6.1 주제명 기술의 비일관성

국립중앙도서관 주제명표목표 매뉴얼에서는 개념 단위로 띄어 쓰기를 권고하였으나 '현대시', '현대 소설', '독일 시'와 같이 비일관적으로 적용된 경우가 있었다. 국립중앙도서관 주제명표목표의 현 구조가 후조합식 색인 언어 기반이고, 세목을 사용하지 않으므로, 개념 구분을 위해 띄어쓰기를 유지할 필요가 있다. 따라서 개념 단위로 띄어 쓰지 않은 주제명은 모두 수정해야 할 것이다.

#### 3.6.2 불분명한 관계 설정

BT(상위개념어)와 BTI(BT 사례 관계), NT(하위개념어)와 NTI(NT 사례관계) 구분이 모호하였다. 예를 들어 '일본어 능력 시험'과 '언어 능력 시험' 간에는 BT로 맺어져 있지만, '일본어 어휘'와 '어휘' 간에는 BTI로 설정되어 있어 관계 설정에 일관성이 떨어졌다. 따라서 BT와 BTI, NT와 NTI를 통합하여 BT와 NT만 사용하거나 일관성 있게 세분 관계를 설정할 필요가 있다.

또한 상식적 수준에서 설정 관계를 이해하기 어려운 경우나 잘못 설정된 관계는 소급해서 수정해야 한다. 계층적 맥락에서 관계 설정이 잘못된 경우도 있었는데, 이럴 경우 한정어를 사용하여 주제명의 의미를 구체화하고 맥락에 따른 주제명을 구분하여 입력해야 한다. 예를 들어 '직무 적성 검사'와 '공직 적성 평가'는 '적성 검사'의 하위개념어가 맞으나, 그 상위 개념어인 '교육 평가'에 모두 속한다고 볼 수 없다. 이럴 경우, 한정어를 사용해서 구분하고, 각각 적합한 BT와 NT를 설정할 필요가 있다.

#### 3.6.3 계층 관계의 비일관성

유사한 의미의 용어가 다르게 취급된 경우도 있었는데, 예를 들어 '일본어 어휘'는 BT로 '어휘'를 설정한 반면 '일본어 회화'는 BT가 없는 최상위어, '영어 회화'는 BT로 '회화(언어)'로 설정되어 있었다. 이런 경우, 일관성을 유지할 수 있도록 소급하여 관계를 수정해야 하며, 근본적으로 주제명 시스템의 검색 기능을 강화하여 주제명의 관계 설정 시 유사한 용어의 관계를 참조할 수 있도록 해야 한다.

### 3.6.4 다중 BT로 인한 용어 의미의 불명확성

상이한 의미를 가진 2개 이상의 용어를 BT로 설정하여 용어 의미가 불명확해지는 사례가 많았다. 예를 들어 '존엄사', '자연사'는 '사망'의 하위 개념이 맞으나, '사망'의 상위 개념인 '업무상 재해'에 속한다고 볼 수 없었다. 또한 2개 이상의 BT를 설정하는 경우도 있었는데, 예를 들어 '한의학'의 BT는 '동방의학'과 '동양의학'인데, 이 두 개 주제명 간에 '동방의학'의 BT가 '동양의학'으로 다시 계층 관계가 설정되어 있었다.

한편, 의미상 차이를 구분하지 않고 NT를 설정하여, 상이한 성격의 NT가 섞여 있는 경우도 발생하였다. 예를 들어 '전쟁'의 NT에 '전투'와 '취업전쟁'과 같이 상이한 의미를 가진 주제명이 함께 설정되었다. 전쟁의 사전적 정의는 '국가와 국가, 또는 교전 단체 사이에 무력을 사용하여 싸움'과 '극심한 경쟁이나 혼란 또는 어떤 문제에 대한 아주 적극적인 대응을 비유적으로 이르는 말'을 모두 포함한다. 이런 경우에는 한정어를 사용해서 주제명의 의미상 차이를 구분하고 관계를 재설정해야 한다.

### 3.6.5 용어 의미의 모호성

주제명 의미 자체를 명확하게 파악하기 어려운 경우도 있었는데, 예를 들어 '교육학'과 '교육과학'은 해당 분야 전문가가 아니면 명확한 의미의 차이를 알기 어려웠다. 주제명은 사서뿐만 아니라 일반 이용자가 활용할 수 있는 정보 접근 도구이므로, 이용자가 쉽게 주제명을 식별하고 의미를 파악할 수 있도록 범위 주기를 활용하여 용어 의미를 구체화시킬 필요가 있다.

### 3.6.6 최신성 부족

최신 주제명의 반영이 전반적으로 미흡하였는데, 일례로 '신재생 에너지'의 일종인 '재생 에너지'나 '태양열 에너지'에 관한 단행본이 많이 출간되었음에도 불구하고 현재 국립중앙도서관 주제명표목표에는 '신재생 에너지'의 하위 개념어는 구축되지 않았다. 그러므로 주제명의 최신성을 보장할 수 있도록 최신 용어를 추가하고, 기존 관계는 수정해야 한다. 또한 신규 주제명에 대한 신속한 검토 및 등록 처리를 병행해야만 주제명의 최신성을 확보할 수 있다.

### 3.6.7 불필요한 관계 생성

불필요한 계층 관계를 설정하여 주제명의 구조를 복잡하게 만드는 경우도 있었다. 예를 들어 '음악가'의 BT는 '예술가' 이외에도 '공연 예술 전문가', '지리 정보 시스템'의 BT는 '정보 시스템' 이외에 '컴퓨터 응용'이라는 모호한 관계가 설정되어 있었다. 이를 해결하려면 주제명 간의 명확한 계층 관계가 성립되는 경우만 설정하고, BT를 3개 이하로 권고하는 등 주제명 업무지침을 구체화해야 한다.

### 3.6.8 분류기호의 부정확한 기입

분류기호가 부정확하게 기입되거나, 잘못된 기호가 기재되어 있는 경우가 있었다. 예를 들어 '연애 소설'은 DDC 분류기호가 800, '영재 교육'은 270.07로 설정되어 있었다. 정확한 분류기호를 입력해야만 주제명과 분류체계 간의 연계를 최대한 활용할 수 있기 때문에 분류기호의 오류는 수정해야 할 것이다. 그리고 동서의 경우, 주제명 표목에는 KDC 제4판 기호를 부여하고 있는데, 국립중앙도서관에서 현재 사용하는

KDC 제5판 기호로 갱신해야 할 것이다.

### 3.6.9 신규 주제명의 중복 신청

신규 주제명의 경우, 신청 단계에서 한자와 한정어를 요청자에 따라 기입하는 방식이 달라 동일 용어가 중복 신청되는 경우가 있었다. 예를 들어 다문화(多文化)를 지칭하는 주제명으로 '다문화', '다문화(多文化)', '다문화[多文化]', 총 3가지 유형으로 신청하였다. 이 경우, 지침 상에 신규 주제명의 입력 규칙을 명시하고, 시스템에서도 용어, 한자, 한정어를 구분하여 입력하도록 수정해야 한다.

### 3.6.10 비우선어 및 미등록어 부여

100회 이상 부여된 1,124개 주제명 중에서 비우선어가 16건 발견되었다. 이는 국립중앙도서관 주제명표목표가 2002년에 구축된 이후, 2006년 추가·수정되는 과정에서 우선어 변경에 따른 비우선어 및 이전 우선어의 처리가 미흡하였기 때문이다. 이용자 검색은 서지 데이터에 부여된 주제명을 검색 대상으로 삼기 때문에, 검색 효율성이 떨어질 수밖에 없다. 또한 판타지 소설, 환상 소설과 같이 최신성 측면에서 우선어보다 오히려 비우선어가 더 적합한 경우도 있으므로 우선어와 비우선어 간의 타당성을 함께 검토할 필요가 있다.

한편, 서지 데이터에 1회 이상 사용된 주제명 45,830개 중에서 주제명표목표에 등록되지 않은 954개 용어가 부여된 서지 데이터도 무려 7,674건이나 발견되었다. 예를 들어 우선어로 등록된 '한자(글자)[漢字]' 이외에도 '한자(문자)[漢字]', '한자[漢字]' 등 동일한 의미이나 다른 형태의 주제명이 서지 데이터에 입력되었다. 서지 데이터에 입력된 미등록 용어의 유형을 분석

한 결과 한정어, 특수문자, 대소문자 표기, 띄어쓰기 등으로 인하여 동일한 의미를 가진 여러 형태의 용어가 서지 데이터에 부여되었다. 이런 경우 주제명표목표에는 한 개만 등록되고 나머지는 미등록되어, 서지 데이터 650 \$a에 입력된 중복 주제명 중, 주제명표목표에 미등록된 용어 추출하여 수정해야 한다.

## 4. 주제명표목표 고품질화 방안

### 4.1 주제명 정비

주제명표목표의 고품질화는 근본적으로 양질의 주제명 데이터 구축과 지속적인 관리가 반드시 전제되어야 한다. 현재 주제명이 수록한 정보에는 용어(표목)와 관계, 그리고 범위 주기, 분류기호와 같은 부가적인 정보로 구성되기 때문에, 주제명의 정비 범위는 위 세 가지 모두가 포함되어야 한다.

#### 4.1.1 용어

용어의 고품질화는 우선어의 타당성을 검증하면서, 현재 기술 규칙에 위배된 용어를 수정하는 과정을 포함한다. 첫째, 우선어가 개념의 대표어로 적합한지를 검토하여, 적합하지 않은 주제명은 새로운 우선어로 변경하고, 이전 우선어는 비우선어(UF)로 처리해야 한다. 이 때 타당성의 기준은 우선어가 업무지침의 대상 용어로 적합한지와 단행본이나 잡지기사 제목, 법률명 등으로 현용되고 있는 용어인지 즉, 최신성이 확보되었는지를 중심으로 설정되어야 한다. 그래서 LCSH와 NDLSH의 경우, 우선어의 형

식을 결정할 때 다양한 참고정보원과 현용 출판물 등을 반영하도록 권고하고 있다.

둘째, 띄어쓰기가 개념 단위로 되어 있지 않거나, 한정어 등 기술 규칙에 위배되는 용어 표기를 수정해야 한다. 특히 전문가 자문회의에서 국립중앙도서관 주제명표목표의 현 구조가 후조합식 색인 언어 기반이고, 세목을 사용하지 않으므로, 개념 구분을 위해 띄어쓰기를 유지할 필요가 있다고 권고하였으므로, 원칙에 어긋나는 용어를 수정 입력해야 한다.

#### 4.1.2 관계

관계의 고품질화는 잘못 설정된 관계에 대한 수정 및 삭제, 계층 및 연관 관계에 대한 검토 및 변경, 관계 단순화의 결정 및 적용 과정을 포함해야 한다.

첫째, 잘못 설정된 관계에 대한 수정 및 삭제는 우선어를 중심으로 직관적으로 잘못 설정된 관계는 검토하여 수정·삭제하고, 유사한 성격 또는 의미를 가진 주제명의 계층 관계가 비일관적으로 취급되었다면 이를 수정해야 한다(<표 1> 참조).

둘째, 계층 및 연관 관계에 대한 검토 및 변경에 있어서는 먼저, 불필요한 관계가 설정된 경우, 검토하여 수정해야 한다. 참조 관계는 용어의 계층 구조와 의미적 맥락을 파악하는데 도움을 주지만, 불필요한 관계는 주제명 구조만 복

잡해지고, 이용자 검색에서도 도움이 되지 않는다. LCSH와 NDLSH의 경우, 참조 관계에 대한 구체적인 지침을 명시하고, 주제명 간의 명확한 계층 및 연관 관계가 성립되는 경우에만 설정하도록 제안하고 있다.

다음으로 BT가 없는 최상위어를 검토하여 가장 가까운 상위 개념어에 BT 관계를 부여해야 한다. LCSH의 경우, 학문이나 주제 범위가 매우 광범위한 경우가 아니라면 모든 주제어는 1개의 BT만 모두 설정하도록 권고하고 있다. 국립중앙도서관 주제명표목표에서는 100회 이상 부여된 주제명 1,124개 중 최상위어가 189개나 되므로, 최상위어로 설정하는 것이 적합한지 검증 절차가 필요하다.

한편, 여러 개의 BT가 설정된 주제명은 용어의 의미가 불분명해질 수 있으므로 필요한 경우 한정어를 이용해 주제명을 구분하고, 가장 가까운 상위 개념어를 BT로 설정해야 한다. LCSH에서는 BT를 3개 이상 부여하지 않으면서, 가장 가까운 개념어에 설정하도록 권고하고 있다. 계층 관계에서 하위 개념은 상위 개념의 속성을 원칙적으로 상속받기 때문에, 하위 개념이 반드시 상위 개념에 속하도록 기존 관계를 수정해야 한다. 유사한 맥락에서 부모 혹은 형제 관계에 속한 2개 이상의 BT를 설정한 경우는 계층적으로 가장 가까운 상위 주제명을 BT로 설정해야 한다(<표 2> 참조).

<표 1> 주제명의 비일관적인 관계 설정에 대한 개선(안)

(예)	일본어 어휘	BT	어휘	⇒	일본어 어휘	BT	어휘
	일본어 회화	BT			일본어 회화	BT	회화
	영어 회화	BT	회화(언어)		영어 회화	BT	회화(언어)
	영문법	BT			영문법	BT	문법

〈표 2〉 다중 BT 설정에 대한 개선(안)

(예)	사망	BT	업무상 재해 생사(삶과죽음) 사중	⇒	사망(재해)	BT	업무상 재해
		NT	재해사 존엄사 자연사		사망	BT	생사(삶과죽음) 사중
	한의학	BT	동방의학	⇒	한의학	BT	동방의학
	동방의학	BT	동양의학		동방의학	BT	동양의학

마지막으로 상이한 의미를 가진 주제명이 NT로 설정된 경우는 한정어를 사용하여 주제명의 의미를 구체화하고 맥락에 따른 주제명을 구분하여 입력해야 한다. 여기에는 주제명 자체가 여러 가지 다른 의미를 가지는 경우와 주제나 학문 분야별로 주제명의 개념적 범주가 다른 경우가 있다. 두 경우 모두 한정어를 추가하고, 범위 주기에 주제명의 정의나 적용 주제 범위 등을 기술해야 한다.

#### 4.1.3 부가 정보

부가 정보의 고품질화는 풍부한 범위 주기의 작성과 분류 기호의 수정 및 추가 과정도 수반해야 한다. 첫째, 범위 주기는 주제명에 대한 정의뿐만 아니라 주제명의 사용 범위를 제한하거나, 유사한 주제명 간의 구분하거나, 기타 주제명 사용상의 유의해야 할 사항 등을 기재할 때도 사용해야 한다. 범위 주기를 기재하는 것이 번거롭고 시간과 노력이 많이 드는 과정이나 주제명의 일관성 있는 부여를 위해서는 반드시 필요하다. 둘째, 분류 기호의 수정 및 추가에 있어서는 우선 해당 주제어의 개념 보다 상위 개념에 해당하는 분류기호를 기재하였거나 잘못된

기호를 기입한 경우 이를 수정해야 한다. 그리고 더 나아가 동서의 경우 현재 국립중앙도서관에서 사용하는 KDC 제5판에 해당하는 분류기호를 추가적으로 입력해야 할 것이다.

#### 4.2 서지 데이터 정비

국립중앙도서관 주제명표목표의 용어 품질 분석 결과, 서지 데이터 상에 비우선어가 부여되어 있거나 주제명 DB에는 수록되지 않은 미등록 용어가 부여된 것이 발견되었다. 특히 미등록 용어는 한정어, 특수문자, 대소문자 표기, 띄어쓰기 등 기술 방식의 차이로 발생하는 경우가 많았다. 아무리 주제명을 고품질화하더라도 이용자 검색은 서지 데이터에 입력된 주제명을 대상으로 수행하기 때문에, 서지 데이터를 정련하지 않는다면 검색 효율성이 떨어질 수밖에 없다. 자문회의 결과, 비우선어를 우선어와 함께 검색 엔진 차원에서 확장 검색하는 대안적인 방안이 제안되었으나, 그 전에 잘못 입력된 서지 데이터가 정비되어야 한다.

비우선어와 미등록 용어에 대한 자동 및 수동의 처리 방안은 첫째, 비우선어의 경우 주제

명 DB에 등록된 비우선어와 서지 데이터 650 \$a에 입력된 주제명 간에 매핑을 수행하고, 매칭된 서지 데이터에 대해 해당하는 우선어를 입력하도록 한다. 둘째, 미등록 용어의 경우는 주제명 DB에 등록된 모든 용어와 서지 데이터 650 \$a에 입력된 주제명 간의 매핑을 수행하여, 비매칭된 서지 데이터를 수작업으로 검수하고, 적절한 주제명을 입력한다. 이외에도 앞서 제시한 우선어가 변경되거나, 한정어가 추가되어 주제명의 형식이 변경된 경우, 반드시 관련된 서지 데이터의 변경 작업이 병행되어야 한다.

#### 4.3 주제명표목 검토위원회 운영

주제명표목표의 안정적인 유지와 관리는 사서 개인이 전담하기에는 그 역할과 책임이 매우 크다. 특히 신규 주제명 선정이나 기존 주제명의 변경 및 삭제, 주제명 부여 대상 자료의 설정, 관계 및 부가 정보 항목의 관리는 '국가서지'라는 거시적이고, 장기적인 안목으로 좀 더 체계적이고 조직적인 방식으로 결정되어야만 한다. LCSH의 경우, 주제명 관리를 위해 주제전거협력프로그램인 SACO와 내·외부 전문가로 구성된 위원회가 설치·운영되고 있다. 이에 본 연구는 주제명표목 검토위원회를 설치·운영하는 방안을 다음과 같이 제안하였다.

첫째, 구성 및 운영 측면에서 검토위원회에는 국가서지와 소속 목록 담당 관리자 및 실무자를 비롯하여 정보조직 분야 외부 전문가 등으로 구성할 필요가 있다. 특히 주제명은 실제 문헌을 반영해야 하기 때문에 실무자의 적극적인 참여가 필수적이다. 검토 회의는 정기적으로 개최하

되, 신규 및 변경 요청 주제명이 많은 경우, 또는 긴급하게 처리가 필요한 경우 수시로 회의를 운영해야 할 것이다.

둘째, 논의 내용으로는 신규 요청된 주제명의 생성이나 기존 주제명의 수정 및 삭제, 요청 주제명 이외에 추가 또는 변경이 필요하다고 판단되는 주제명에 대한 생성·수정·삭제에 관한 사항 등이다. 이외에도 현재는 주제명 부여 대상 자료가 단행본에 국한되어 있으나 향후 학위 논문이나 학술지 논문에도 부여범위를 확장해야 한다면 위원회에서 적용 가능성과 효용성을 사전에 검토하여 결정해야 할 것이다. 또한 BTI, BTP, NTI, NTP 등 세분화된 관계를 통합하고 단순화하거나 새로운 관계를 추가하고자 할 경우에도 위원회가 충분히 검토할 필요가 있다. 이외에도 주제명의 부가 정보 항목 등 주제명표목 관리에 필요한 제반 사항이나 업무지침의 내용 등도 논의 대상이 될 수 있다.

#### 4.4 주제명 지침 보완

국립중앙도서관은 이미 2009년부터 총 6페이지로 구성된 『주제명표목표 매뉴얼』을 보유하고 있으며, 7개 부문인 부여대상, 사용 tag, 부여 범위, 주요 용어 관계지시 기호, 주제명 신규용어 부여 제외대상, 주제명 부여 방법, 주제명요청기술규칙(요약)으로 구성되었다. 그러나 수록 내용이 매우 간단하고 압축적이라 실제 신규 주제명 구축이나 주제명 관계 설정에 참고하기 어려운 상황이다.

NDLSH와 LCSH의 지침에서는 주제명 생성과 부여에 관한 일반 원칙과 학문 분야와 자료 유형별로 세부 원칙이 자세하게 마련된 반면,

국립중앙도서관 주제명표목표의 경우 원칙보다는 시스템을 이용해 주제명을 부여하는 절차를 위주로 기술되었다. 물론 LCSH는 1914년부터 NDLSH는 1964년부터 주제명을 구축·관리해 왔고, 그 간의 경험을 바탕으로 지침도 지속적으로 보완해왔기 때문에 국립중앙도서관의 『주제명표목표 매뉴얼』의 내용은 미흡할 수밖에 없다. 그러나 국가 지식정보에 '국립중앙도서관 주제명표목표'를 일관성 있게 부여하게 하고, 체계적인 주제명 생성과 관리를 도모하려면 업무 지침 상에 신규 주제명의 생성과 선정, 기존 주제명의 수정 및 삭제, 한정어의 활용 방법, 참조관계 설정 기준 등 기본적인 원칙들을 보강할 필요가 있다.

## 5. 결 론

오늘날 인터넷에 접속해 막대한 양의 정보를 쉽게 이용할 수 있게 되면서 정보의 양은 50년 전에 대충이 이용할 수 있는 것보다 10배 이상으로 많아졌다. 이제 사람들은 대량 정보의 저장고를 접하면서 정보를 검색하고 관리하는 능력에 도전을 받고 있으며 이들이 적시에 최적의 정보를 찾을 수 있도록 도서관은 정보를 조직하고 서비스해야 한다.

이러한 맥락에서 본 연구는 국립중앙도서관이 주제 접근 도구로 2002년에 개발한 국립중앙

도서관 주제명표목표를 전 세계에 개방하고 공유하기 위한 첫 번째 단계로 국립중앙도서관 주제명표목표의 용어 품질을 검토하여 개선 방안을 마련하였다. 이를 위해 문헌연구와 한국, 미국, 일본의 국가도서관의 주제명표목표 사례 조사, 그리고 국립중앙도서관 주제명표목표의 용어 및 관계 데이터를 분석하였다. 이를 바탕으로 국립중앙도서관 주제명표목표 용어 품질 개선 방안을 크게 4가지 부문, 주제명 정비, 서지데이터 정비, 주제명표목 검토위원회 운영, 주제명 지침 보완으로 제시하였다.

국립중앙도서관은 장·단기적으로 주제명표목표를 고품질화하는 전략을 수립하고, 이를 통해 우수한 주제명 데이터를 개방해서 주제명 데이터를 가진 외부 기관과 공유 할 수 있는 전략을 주도적으로 추진해야 할 것이다. 이를 위해 무엇보다도 외부 유관 기관의 이해당사자들과 기술적, 행정적 측면의 협력을 우선적으로 모색할 필요가 있으며, 향후 주제명 시스템에서 가능한 모든 잠재적인 융통성, 정보 기술의 잠재적 가능성, 자원 조달 전략과 유관 기관과의 파트너십을 최대한 강구해야 할 것이다. 그래서 국립중앙도서관은 정보와 지식이 국가 발전의 가장 중요한 자원이며 정보자원의 조직과 접근이 매우 중요한 기반임을 확고히 하고 이용자 친화적인 주제명표목표로 고도화되어야만 할 것이다.

## 참 고 문 헌

- [1] 國立國會圖書館. 2013. 『國立國會圖書館件名作業指針』. [online]  
〈<http://www.ndl.go.jp/jp/library/data/kenmeimanual.pdf>〉
- [2] 국립중앙도서관 편. 2002. 『국립중앙도서관 주제명표목표 개발 최종보고서』. 서울: 국립중앙도서관
- [3] 국립중앙도서관. 2009. 『주제명표목표 매뉴얼』. [미간행자료]. 서울: 국립중앙도서관
- [4] 국립중앙도서관. 2013. 『국립중앙도서관 주제명표목표 개방을 위한 고품질화 연구』. 서울: 국립중앙도서관
- [5] 김정현, 문지현. 2009. LCSH 한국관련 주제명표목의 변천과정에 관한 연구. 『한국도서관·정보학회지』, 40(3): 49-68.
- [6] 김정현. 2010. 미국의회도서관 주제명표목표의 한·중·일 관련 주제명표목의 변천과정 비교 분석. 『한국도서관·정보학회지』, 41(2): 147-169.
- [7] 김정현. 2011. 주요 국가의 국가도서관 목록에 나타난 한국관련 주제명 비교 분석. 『정보관리학회지』, 28(3): 141-162.
- [8] 김정현. 2013. 주요 국가의 주제명표목표에 나타난 한국의 국경관련 주제명 비교 분석. 『한국도서관·정보학회지』, 44(2): 217-239.
- [9] 김주성, 김태수. 1995. LC주제명표목표와 시소러스의 색인어 구조 비교연구. 『한국정보관리학회 학술대회 논문집』, 2: 111-114.
- [10] 김태수. 1997. 주제명표목의 구조에 관한 연구. 『정보관리학회지』, 14(1): 77-105.
- [11] 남태우, 전말숙, 정연순, 장로사. 2008. 자료조직 측면에서 독도표기 문제에 관한 연구. 『한국문헌정보학회지』, 42(4): 291-310.
- [12] 박지영. 2011. 서지레코드의 주제전거를 위한 FRSAD 모형 연구. 『정보관리연구』, 42(1): 113-135.
- [13] 유영준. 2009. 신학술지의 색인어에 관한 연구. 『한국도서관·정보학회지』, 40(3): 137-156.
- [14] 윤정옥. 2001. 미국의회도서관 주제명표목표(LCSH)의 한국과 일본 관련 주제표목의 비교 연구. 『도서관』, 56(1): 30-55.
- [15] 윤혜영. 2003. LC Subject Headings 색인어 구조 분석. 『인문사회과학논문집(중부대학교)』, 7(2): 133-149.
- [16] 전창호. 2004. DDC에서의 한국, 조선 및 동해의 표기 문제. 『도서관문화』, 45(9): 68-73.
- [17] 정연경. 2010. 정보의 접근점으로서 미국의회도서관 주제명 표목에 관한 연구. 『한국문헌정보학회지』, 44(2): 51-72.
- [18] 최달현, 이창수. 2010. 『정보자료의 분류와 주제명』. 서울: 한국도서관협회.
- [19] 한상완, 김태수, 최석두, 이은철, 남태우. 1995. 주제명 검색요어집 개발에 관한 연구. 『도서관』, 50(1): 3-23.

- [20] 황면, 이재선, 한정원, 한인선, 류희경, 송지혜, 김승정. 2002. 시소러스를 도입한 주제명표목표에 관한 연구. 『도서관연구회논문집』, 2002: 123-138.
- [21] Anderson, James and Hofmann, Melissa. 2006. "A fully faceted syntax for Library of Congress Subject Headings." *Cataloging & Classification Quarterly*, 43(1): 7-38.
- [22] Carlyle, A. 1989. "Matching LCSH and user vocabulary in the library catalog." *Cataloging & Classification Quarterly*, 10: 37-63.
- [23] Chan, Lois Mai and O'Neill, Edward T. 2010. *FAST, Faceted Application of Subject Terminology: principles and applications*. Santa Barbara, Calif: Libraries Unlimited, 2010.
- [24] Chan, Lois Mai and Hodges, Thedora. 2000. "Entering the millennium: a new century for LCSH." *Cataloging & Classification Quarterly*, 29(1): 225-234.
- [25] Dean, Rebecca J. 2004. "FAST: development of simplified headings for metadata." *Cataloging & Classification Quarterly*, 39:1-2, 331-352.
- [26] Fernandez, Cheryl Wise. 1991. "Semantic relationships between title phrases and LCSH." *Cataloging & Classification Quarterly*, 13(1): 51-77.
- [27] Fischer, Karen. 2005. "Critical view of LCSH, 1990-2001: the third bibliographic essay." *Cataloging & Classification Quarterly*, 41(1): 63-109.
- [28] Frost, Carolyn O. 1985. Student and faculty subject searching in a university online public catalog. [online] <<http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED264872>>
- [29] Jin, Qiang. 2008. "Is FAST the right direction for a new system of subject cataloging and metadata?" *Cataloging & Classification Quarterly*, 45(3): 91-10.
- [30] Library of Congress. 2012. Introduction to Library of Congress Subject Headings. [online] <<http://www.loc.gov/aba/publications/FreeLCSH/lcshintro.pdf>>
- [31] Mann, Thomas. 2000. "Teaching Library of Congress Subject Headings." *Cataloging & Classification Quarterly*, 29(1): 117-126.
- [32] Taylor, Arlene G. and Jourdre, Daniel N. 2009. *The Organization of Information*. Westport, CT: Libraries Unlimited.
- [33] Vizine-Goetz, Diane and Drabenstott, Karen Markey. 1991. "Computer and manual analysis of subject terms entered by online catalog users." *Proceedings of the ASIS Annual Meeting*, 28: 156-161.
- [34] Walsh, John. 2011. "The use of Library of Congress Subject Headings in digital collections." *Library Review*, 60(4): 328-343.
- [35] Wool, Gregory. 2000. "Filing and precoordination: how subject headings are displayed in

- online catalogs and why it matters.” *Cataloging & Classification Quarterly*, 29(1-2): 91-106.
- [36] Yi, Kwan and Chan, Lois Mai, 2010. “Revisiting the syntactical and structural analysis of Library of Congress Subject Headings for the digital environment.” *Journal of the American Society for Information Science and Technology*, 61(4): 677-687.

• 국문 참고자료의 영어 표기

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

- [1] National Diet Library of Japan, 2013. *Subject heading manual in National Diet Library of Japan*. [online] <<http://www.ndl.go.jp/jp/library/data/kenmeimanual.pdf>>
- [2] National Library of Korea, 2002. *A study on development of National Library Subject Headings*. Seoul: National Library of Korea.
- [3] National Library of Korea, 2009. *Subject Heading List manual*. [unpublished]. Seoul: National Library of Korea.
- [4] National Library of Korea 2013. *A study on high quality for opening National Library of Korea Subject Heading List*. Seoul: National Library of Korea.
- [5] Kim, Jeong-Hyen and Moon, Ji-Hyun, 2009. “A study on the development of topical headings related to Korea in LCSH.” *Journal of Korean Library and Information Science Society*, 40(3): 49-68.
- [6] Kim, Jeong-Hyen, 2010. “A comparative study of subject headings related to Korea, China, and Japan in the LCSH.” *Journal of Korean Library and Information Science Society*, 41(2): 147-169.
- [7] Kim, Jeong-Hyen, 2011. “A comparative analysis of subject headings related to Korea in the National Library OPACs of major countries.” *Journal of the Korean Society for Information Management*, 28(3): 141-162.
- [8] Kim, Jeong-Hyen, 2013. “A comparative analysis of subject headings related to Korean border in the subject headings of major countries.” *Journal of Korean Library and Information Science Society*, 44(2): 217-239.
- [9] Kim, Joo-Sung and Kim, Tae-Soo, 1995. “A comparative study on the structures of indexing languages between LC Subject Headings and thesaurus.” *Proceeding of Korean Society for Information Management*, 2: 111-114.
- [10] Kim, Tae-Soo, 1997. “A study on the structure of subject headings.” *Journal of the Korean Society for Information Management*, 14(1): 77-105.

- [11] Nam, Tae-Woo, Chun, Mal-Suk, Chung, Yon-Soon and Chang, Ro-sa. 2008. "A study of the Dokdo notation problem in terms of library materials organization." *Journal of the Korean Library and Information Science Society*, 42(4): 291-310.
- [12] Park, Zi-young. 2011. "A study on the functional requirement of subject authority data for subject authority control." *Journal of information management*, 42(1): 113-135.
- [13] Yoo, Yeong-Jun. 2009. "Analytical study on the index terms in the theological journals." *Journal of Korean Library and Information Science Society*, 40(3): 137-156.
- [14] Yoon, Cheong-Ok. 2001. "A comparative study on subject headings related to Korea and Japan in Library of Congress Subject Headings." *The Library*, 56(1): 30-55.
- [15] Yoon, Hye-Young. 2003. "Analytical study on the structure of indexing terms in LC Subject Headings." *Journal of Humanities & Social Science*, 7(2): 133-149.
- [16] Jeon, Chang-Ho. 2004. "Problems of romanizing Korea, Chosun and East Sea in DDC." *KLA journal*, 45(9): 68-73.
- [17] Chung, Yeon-Kyoung. 2010. "A study on the Library of Congress's Subject Headings as an information access point." *Journal of the Korean Library and Information Science Society*, 44(2): 51-72.
- [18] Choi, Dal-Hyun and Lee, Chang Soo. 2010. *Classification and subject heading of information materials*. Seoul: Korean Library Association.
- [19] Han, Sang-Wan, Kim, Tae-Soo, Choi, Lee, Eun Chul and Nam, Tae-Woo. 1995. "A study on the development of search keywords for subject headings." *The Library*, 50(1): 3-23.
- [20] Hwang Myun, Lee, Jae-Sun, Han, In-Seon, Ryu, Hee-Kyoung, Song, Ji-Hye and Kim, Seung-Jung. 2002. "A study on the subject heading List adopted thesauri." *Doseogwanyeonguhuinmunjib*, 2002: 123-138.