

KORMARC 서지레코드의 FRBR 알고리즘 개발에 관한 연구*

A Study on the Development of FRBR Algorithm for KORMARC Bibliographic Record

김 정 현(Jeong-Hyen Kim)**

이 성 숙(Sung-Sook Lee)***

이 유 정(You-Jeong Lee)****

〈목 차〉

I. 서론	III. KORMARC의 FRBR 알고리즘 개발
1. 연구필요성 및 목적	1. FRBR 알고리즘 설계
2. 연구대상 및 방법	2. FRBR 알고리즘 적용
II. FRBR 알고리즘의 개발 사례 분석	3. FRBR 알고리즘 개발을 위한 시사점
1. FRBR 알고리즘의 논의	IV. 결론
2. FRBR 알고리즘의 비교 분석	

초 록

이 연구의 목적은 KORMARC 서지레코드를 유형별로 분석하여 FRBR을 적용한 검색 알고리즘을 개발하는 데 있다. 이를 위해 OCLC와 LC를 비롯한 국내외 FRBR 구현 알고리즘 개발 사례의 장단점을 비교 분석하여 이 연구의 기반으로 활용하였으며, 국립중앙도서관의 국가서지레코드에서 추출한 KORMARC 실험데이터를 분석하여 FRBR의 4가지 서지적 개체로 식별요소를 추출하였다. 저작세트별로 관련 저작을 군집화하기 ‘저자+표제’와 같은 저작의 전거형 접근점을 작성할 수 있는 알고리즘을 설계하였다. 한편 국가서지레코드를 분석한 결과, KORMARC 서지레코드를 대상으로 FRBR 알고리즘을 적용하기 위해서는 기존의 서지레코드를 정비하고 레코드의 입력 수준을 전반적으로 강화할 것을 제안하였다.

키워드: FRBR, FRBR 알고리즘, KORMARC

ABSTRACT

The purpose of this paper is to development an algorithm for FRBR implementation(functional requirements for bibliographic records), analyzing KORMARC bibliographic records by work types. For this purpose, it was utilized analyzing home and foreign case studies including OCLC and LC's algorithm. Analyzing the experimental data from the Korean National Library's bibliographic records, it was extracted from identifying elements by FRBR's four bibliographic entities. To cluster a related works by work-set, the algorithm was designed to construct the authorized access points as a combination of an author name and a title from the record. I suggested that it should be wholly improved a quality of existing bibliographic records and level of data input to development a FRBR algorithm in Korean libraries.

Keywords: FRBR, FRBR algorithm, KORMARC

* 이 논문은 2014년 국립중앙도서관의 “RDA를 기반으로 한 국가서지데이터의 FRBR 모형 연구” 보고서의 일부를 수정 보완한 것임

** 전남대학교 문헌정보학과 교수 (jhkim@chonnam.ac.kr) (제1저자)

*** 충남대학교 문헌정보학과 부교수 (infolee@cnu.ac.kr) (교신저자)

**** 영남대학교 중앙도서관 사서 (yjlee@yu.ac.kr) (공동저자)

•논문접수: 2015년 2월 14일 •최초심사: 2015년 2월 25일 •게재확정: 2015년 3월 18일

•한국도서관·정보학회지 46(1), 1-23, 2015. [http://dx.doi.org/10.16981/kliss.46.201503.1]

I. 서론

1. 연구필요성 및 목적

1998년 IFLA에서 발표한 FRBR은 자료의 다양성과 목록의 관계성을 수용하는 이용자 지향적 모형으로서 이미 전 세계 목록의 이론적 바탕이 되고 있다. 각국의 서지기관에서는 이러한 추세를 반영하여 기존의 서지레코드를 대상으로 FRBR을 기반으로 한 새로운 구조화를 시도하고 있다. 그런데 FRBR을 적용하여 실제로 목록을 구조화하기 위해서는 FRBR을 적용한 알고리즘 개발이 선행되지 않으면 방대한 서지레코드를 대상으로 관련 저작을 유형별로 군집화 하는 것이 거의 불가능하다.

기존의 서지레코드에 FRBR을 적용하여 관련 서지레코드를 유형별로 군집화 한다는 것은 MARC 등 다른 입력 형식으로 이미 편목되어진 서지레코드를 FRBR 개체로 재분해하는 과정이라고 볼 수 있다. 즉 기존의 서지레코드를 FRBR 모형의 관점으로 분해하여 저작 수준으로 모으고, 모아진 레코드를 표현형, 구현형, 개별자료 수준으로 각각 구분하는 것이다. 이러한 FRBR화 과정은 수작업으로 구현하는 것은 거의 불가능하기 때문에 컴퓨터 알고리즘을 통해 자동화를 통해 이뤄진다. FRBR 모형이 발표되고 서지레코드에 실제적으로 적용하기 위한 과정에서 FRBR화를 위한 알고리즘 개발에 대한 논의가 OCLC나 LC 등을 중심으로 다수 진행되었다.

KORMARC 서지레코드에 있어서도 FRBR과 관련된 많은 연구가 있으나, FRBR 알고리즘을 개발하여 자료검색에 적용하기 위해서는 실제로 프로그램을 개발하여 직접 서지데이터베이스에 적용하여 봄으로서 문제점을 분석한 실제적인 연구가 필요하다. 아울러 FRBR의 저작 유형은 물론, 각종 자료유형별로 KORMARC 서지레코드의 특성을 분석하여 알고리즘에 반영해야 한다.

이 연구의 목적은 KORMARC 서지레코드를 유형별로 분석하여 FRBR을 적용한 검색 알고리즘을 개발한 후, 실제로 검색 과정을 국립중앙도서관의 KORMARC 서지데이터베이스를 통해 직접 실험하여 봄으로써 KORMARC 서지레코드의 FRBR 적용을 위한 기초자료를 제시하는데 있다.

2. 연구대상 및 방법

이 연구의 구체적인 대상과 방법은 아래와 같다.

첫째, 국립중앙도서관 소장 문학류(1910~2013년 발행) 단행본 약 20만 권을 실험대상으

로 한다. 문학류가 다양한 서지적 관계유형을 가장 많이 포함하고 있는 것으로 알려져 있어 이를 실험대상으로 하였다. 단행본으로 한정하였지만 단행본에서 파생된 다양한 유형의 자료도 검색 알고리즘에는 포함될 수 있도록 설계한다.

둘째, KORMARC 서지레코드에 내포된 다양한 서지적 관계 유형을 FRBR 모형에 적용하여 검색에 반영하는 알고리즘을 설계한다.

셋째, FRBR 모형을 적용한 자료 검색 알고리즘 설계는 아래와 같은 순서로 연구를 진행한다.

- ① 알고리즘 설계관련 국내외 사례 분석을 한다. 특히 이 연구와 관련이 높은 OCLC와 LC 등에서 개발한 FRBR 알고리즘을 중점적으로 분석한다.
- ② FRBR 알고리즘 사례 분석을 바탕으로 국립중앙도서관의 서지레코드를 문학 장르와 서지유형별로 분석하여 FRBR 모형을 적용한 알고리즘을 설계한다.
- ③ 국립중앙도서관의 문학류 서지레코드 가운데 가능하면 다양한 저작유형을 반영할 수 있는 사례들을 엄선하여 실험 데이터로 사용한다.
- ④ 알고리즘 설계내용을 바탕으로 유형별로 엄선한 실험 데이터를 직접 적용해봄으로써 FRBR 모형 적용과 관련된 문제점을 분석한다.

Ⅱ. FRBR 알고리즘의 개발 사례 분석

1. FRBR 알고리즘의 논의

FRBR이 발표된 후 기존의 서지레코드를 대상으로 기계적 변환을 통해 FRBR을 적용하려는 논의와 함께 실제로 알고리즘이 개발되기도 하였다. 이 가운데 대표적인 사례로 저작 수준의 개체를 군집화한 OCLC의 FRBR Work-Set Algorithm을 들 수 있으며(Hickey, O'Neill, Toves 2002; 2005; 2009), 이는 원래 O'Neill(2002)이 개발한 것으로 MARC 21 레코드에서 T. Smollett의 저작 "Humphly Clinker"의 분석을 바탕으로 하고 있다. FRBR Work-Set Algorithm의 전체적인 과정은 먼저 전거레코드로부터 표목을 자동 생성하고, 서지레코드에서 저자명과 표제를 추출하여, 전거레코드의 표목과 매칭한 후, 알고리즘에서 정의한 규칙에 의해 일치하는 경우 동일한 Work-Set으로 묶은 다음, 전거레코드에서 추출한 '이름/표제 대표형'을 저작의 대표기로 할당한다. OCLC에서는 소설작품을 대상으로 FictionFinder와 WorldCat 전체 자료를 대상으로 한 WorldCat.org에 이 알고리즘을 적용하여 FRBR화를 시도하고 있다.

Library of Congress(2003; 2009)는 MARCXML과 XSLT를 사용하여 FRBR 적용 검

4 한국도서관정보학회지(제46권 제1호)

색 알고리즘인 FRBR Display Tool Version 2.0을 개발하였다. 이의 작동순서는 서지레코드에서 ‘저자+표제’ 또는 ‘표제’를 추출한 후 이들 데이터를 저작을 판별하기 위한 근거로 삼은 다음, 동일 저작으로 분류된 레코드를 대상으로 다시 ‘자료유형’과 ‘언어’로 나누어 차하위 계층인 표현형을 구성하도록 설계되어 있다.

또한 Carlyle 등(2008)은 OCLC의 알고리즘을 바탕으로 4개의 소설 저작을 바탕으로 식별성능을 분석하였다. 이외에도 OCLC의 알고리즘을 바탕으로 저작의 군집화를 시도한 호주 국립도서관(Rajapatirana 2005), 노르웨이 국립도서관의 BibSYS FRBR conversion tool(Aalberg 2006), CMARC/MARC21간의 알고리즘 매핑을 통해 FRBR OPAC을 시도한 Chinese Koha system(Chang 2011; 2013), JAPAN/MARC 서지레코드를 대상으로 FRBR을 적용한 실험연구로 宮田洋輔(2008), 谷口祥一(2009) 등이 있다.

한편 KORMARC 데이터를 분석대상으로 한 연구는 아래와 같다. 먼저 조재인(2004)은 OCLC와 LC를 중심으로 FRBR 알고리즘의 구성요소 및 기능적 흐름을 분석한 후, KORMARC 데이터베이스의 FRBR 변환가능성을 분석하고 적절한 대안으로 저작 군집화의 기본 구성 및 기능적 흐름을 제시하였다. 또한 2006년도의 연구(Cho 2006)에서도 알고리즘의 적용과 동시에 육안 식별 과정을 병행하는 ‘semi-automated’를 제안하였다.

국립중앙도서관(2007)에서는 국립디지털도서관 운영 시범사업의 연구로 FRBR 모형 기반 온톨로지를 설계하였는데, 한국문학을 대상으로 FRBR 모형 기반 프로토타입을 개발(문학작품, 저자, 장르별, 시대별 구분)하였다. 노지현(2008)은 기존의 FRBR 알고리즘을 검토하고 이러한 알고리즘을 KORMARC 레코드에 대입해 보는 과정을 세부적으로 추적하면서, 국립중앙도서관의 서지레코드(161건)에 FRBR 모형을 적용하기 위한 알고리즘을 논의하고 적용해 보는 실험을 수행하였다.

김현희 등(2007)은 음악자료를 대상으로 KORMARC에서 저작과 표현형을 식별하고 표현형들을 군집시키기 위한 알고리즘을 설계한 후 KERIS의 레코드(107건)를 대상으로 실험 시스템을 적용하였다. 이미화(2008)는 FRBR 저작 및 표현형의 알고리즘을 설계하여 저작 및 표현형의 군집화 정도를 파악하고, 군집화가 낮은 문제점을 목록의 측면에서 분석하였다.

지금까지 살펴본 바와 같이 FRBR 알고리즘의 개발은 OCLC와 LC가 선도하고 있으며, 이를 바탕으로 각국에서 적용 실험이 이루어지고 있다. KORMARC의 경우에도 다양한 연구가 이루어지고 있으나 실제적인 적용에 이르지 않고 실험적인 연구에 머무르고 있다.

2. FRBR 알고리즘의 비교 분석

앞절에서 살펴본 바와 같이 FRBR 알고리즘은 주로 OCLC와 LC를 중심으로 개발되고 있

으며, 이것은 MARC 21의 서지레코드를 대상으로 하고 있어 KORMARC과 유사성이 있다. 따라서 여기서는 OCLC와 LC의 알고리즘 및 KORMARC을 중심으로 분석하고 있는 국내의 FRBR 알고리즘 관련 연구를 중심으로 보다 자세하게 살펴보기로 한다.

우선 OCLC의 알고리즘은 매칭정보로 전거와 서지레코드의 이름과 표제 관련 표목을 사용하는데, 구체적으로는 전거레코드로부터 표목[이름+표제(혹은 이름) 대표형과 변형]를 자동 생성하고, 서지레코드에서 '저자명+표제'를 추출하여, 전거레코드의 표목[이름+표제(혹은 이름) 대표형과 변형] 전거와 서지레코드의 '이름+표제 관련 표목'을 매칭한다. 매칭된 서지레코드의 대표 엔트리로는 전거레코드의 '이름+표제' 또는 '이름 또는 통일표제' 중 적절한 대표형을 저작의 대표키로 할당한다. 반면 LC의 알고리즘은 '이름 또는 표제'로 검색한 결과로 출력되는 서지레코드를 대상으로 FRBR 알고리즘을 적용하며, 매칭정보로 서지레코드의 '저자+표제' 또는 '표제'를 추출한 후 서지레코드끼리 매칭한다. 대표 엔트리로는 서지레코드의 저자, 표제 대표키를 할당한다.

이어서 기존의 FRBR 알고리즘에서 사용된 저작 수준의 대표키 생성과 표현형, 구현형, 개별자료 수준에서의 식별요소를 비교 분석하면 아래와 같다.

첫째, 저작 대표키 조합의 방법에 있어 알고리즘마다 사용하는 표현방법이나 사용하는 표시기호나 식별기호, 우선순위 등의 차이가 있지만, 대체로 아래와 같다. OCLC의 알고리즘은 ① 저자(1XX 필드)가 있는 경우, ② 통일표제 사용, ③ 저자(1XX 필드)가 없는 경우, ④ 표제+OCLC번호의 기타방법 등을 언급하고 있다. LC의 알고리즘은 1XX 필드가 있는 경우와 없는 경우로 구분하고 있으며, OCLC와 달리 1XX 필드 없을 때 7XX 필드를 사용하지 않는다. LC에서는 이름 또는 표제로 검색한 결과로 출력되는 서지레코드를 대상으로 FRBR 알고리즘을 적용하기 때문에 표제만으로 저작 대표키를 사용하고 있다.

둘째, 저작 대표키의 한 부분인 저자키 생성과 관련하여, LC는 1XX 필드만, 나머지 선행연구에서는 7XX 필드도 포함하고 있다. OCLC는 1XX와 7XX 필드가 연결된 880 필드를, 노지현은 900 필드를 포함시키고 있다.

셋째, 저작 대표키의 또 다른 부분인 표제키 생성과 관련하여 LC는 1XX 필드가 있는 경우와 없는 경우로, OCLC는 표제를 단축표제와 완전표제로 구분하여 제시하고 있다. 노르웨이 국립도서관에서는 500, 505 필드를 포함하고 있다. 국내 연구자의 경우 240/245/740 필드의 ▼a를 포함하고 있으며, 조재인은 505과 730 필드, 이미화는 507(분석대상이 USMARC 기반인 H대학을 이므로) 필드를 포함시키고 있고, 김현희는 LC와 유사하게 각 필드의 식별기호를 포함하고 있다. 통일표제에 대해서는 공통으로 130 필드를 언급하고 있다.

넷째, 동일한 저작으로 묶인 서지레코드의 표현형 식별요소는 OCLC는 언어(008 필드/35-37), LC는 언어에 자료유형(리더/06)을 언급하고 있다. 국내의 경우 OCLC나 LC가 언

급한 식별기준이외에 노지현은 언어에 130과 240 필드의 ▼1, 자료유형에는 008 필드/23을 추가하였고, 이용대상자 수준(008 필드/22)을 추가하였다. 이미화는 250 ▼a(숫자만 추출)를 추가하였다. 김현희는 연구대상자료가 음악자료임을 감안하여 546 ▼a와 245 ▼h, 254 ▼a, 511 ▼a, 518 ▼a를 추가하였다.

다섯째, 구현형에 대해 OCLC, 조재인, 이미화의 연구에서는 군집화를 수행하지 않았다. LC에서도 이 수준에서는 군집화를 수행하지 않았다. 다만 구현형 수준에서는 발행년도(008 필드/07-10)로 정렬하며, 디스플레이를 위해 판사항(250 ▼ab), 표제(245 ▼abnp), 저작책임(245 ▼c), 출판사항(260 ▼bcg), 형태사항(300), ISBN(020 ▼acz), ISSN(022 ▼ayz), 출판사번호(028 ▼ab), CODEN(030 ▼az), 복제(533) 등을 제안하고 있다. 개별자료에 대해서는 국립중앙도서관과 노지현의 연구에서만 구현형에 대한 복본표시와 소장자료를 제시하는 수준의 작업을 수행함이 언급되어 있다.

Ⅲ. KORMARC의 FRBR 알고리즘 개발

1. FRBR 알고리즘 설계

가. 설계원칙

FRBR의 구현을 위한 설계 내용은 아래와 같다.

첫째, FRBR은 제1집단의 서지적 개체를 크게 저작, 표현형, 구현형, 개별자료의 4가지로 구분하고 있다. 이 연구에서도 이러한 4가지 개체로 구분하여 알고리즘을 개발하기로 한다.

둘째, FRBR의 핵심은 특정 저작의 모든 관련 저작을 유형별로 제시하는데 있다. 그러기 위해서는 관련 저작을 묶을 수 있는 연결 고리가 필요하며, FRBR을 기반으로 하고 있는 RDA에서는 저작의 전거형 접근점을 사용하고 있다(RDA 6.27). 이 연구에서도 RDA에서와 같이 저작의 전거형 접근점을 작성하여 관련저작을 모을 수 있도록 ‘저자+표제’(저자가 없는 경우 ‘표제’)와 같은 저작의 전거형 접근점을 작성할 수 있는 알고리즘이 설계될 수 있도록 한다.

셋째, 개정, 번역, 비평 등 서지적 관계를 표현하는 필드를 추출하여 알고리즘에 반영한다.

넷째, 표제의 추출을 위해 표제 관련 필드(240, 245, 246, 740) 외에 내용주기 필드(505)에 포함된 표제도 추출하도록 하여 알고리즘을 개발한다.

나. 실험데이터

1) 실험데이터 선정

알고리즘 설계 과정에서 다양한 경우의 수가 포함될 수 있도록 다양한 서지유형과 문학 장르가 포함된 실험데이터를 선정하였으며, 문학 작품의 경우 비교적 널리 알려진 작품을 대상으로 하였다.

문학 이론서의 경우 초판과 개정판 등이 동시에 있는 자료를 선정하였다. 문학 작품의 경우 소설, 시, 희곡의 문학 장르를 포함하여 언어, 출판사항, 자료 유형(도서, 녹음자료, 시청각자료, 장애인대체자료 등)이 다양하며, 이러한 사항이 골고루 포함된 자료를 선정하였다. 한국소설에는 고대소설, 장편소설, 단편소설이 골고루 포함되도록 하였으며, 한국 시의 경우 서지레코드가 많으면서 한국어 외에 다른 언어로 번역된 작품이 포함된 자료를 선정하였다. 외국문학의 경우 희곡, 소설 등 문학 작품의 번역서를 포함하였다. 한국소설과 외국 번역작품의 경우 관련 서지레코드가 적어도 100건 이상인 자료를 선정함으로써 다양한 관련 서지레코드의 사례가 포함될 수 있도록 하였다. 무저자 고전인 『춘향전』을 포함하여 <표 1>과 같이 모두 21건의 실험 데이터를 선정하였다. 선정된 실험데이터는 저자와 표제의 저작세트로 구성하여 MARC 데이터를 추출하여 분석하였으며, 분석 결과는 FRBR 알고리즘을 개발하는데 반영하였다.

<표 1> FRBR 알고리즘의 실험데이터 리스트

저자	표제	장르	저자	표제	장르
문학과문학교육연구소	문학의 이해	문학이론서	허균	홍길동	한국소설 (장편, 단편, 고전, 현대)
김학주	중국문학사		김만중	구운몽	
한국문학편집위원회	한국문학개설		이광수	무정	
조동일	한국문학통사	김동인	감자		
T. Eagleton	문학이론입문	이효석	메밀꽃 필 무렵		
A. Jefferson 외	현대문학이론	황순원	소나기		
윤동주	하늘과 바람과 별과 시	한국시	이문열	우리들의 일그러진 영웅	영미소설 (번역서)
김소월	진달래꽃		W. Shakespeare	로미오와 줄리엣	
고은	만인보		W. Shakespeare	햄릿	
-	춘향전	무저자고전	E. Hemingway	노인과 바다	
			J. Austen	오만과 편견	

2) 실험데이터 분석

(1) 저작세트 추출을 위한 요소 분석

저작은 <표 2>와 같이 ‘저자명+표제’로 구성되며 먼저 저자명을 추출하기 위해 저자명 전거레코드를 먼저 확인하여 저자명을 추출한다. 저자명 전거레코드가 입력되어 있지 않은 경우 서지레코드에서 저자명을 추출한다. 저자의 경우 1XX 필드를 우선 추출하고, 1XX 필드가 없을 경우 7XX 필드를 추출하며 개수에 상관없이 모두 추출한다. <표 2>의 사례와 같이

8 한국도서관정보학회지(제46권 제1호)

100+245, 700+245 ▼a, 700+245 ▼b, 700+740 필드 등에서 저작의 전기형 접근점을 생성하며, 저자 필드와 표제 필드를 대상으로 모든 저작세트가 추출될 때까지 반복한다.

표제의 경우 표제 필드인 240, 245, 246 ▼a를 우선 추출한 다음 740 ▼a를 추출하며, 표제 필드의 개수에 상관없이 모든 표제가 추출될 때까지 반복 실행한다. 표제 추출을 위한 식별기호로는 ▼a뿐만 아니라, ▼n ▼p ▼x를 모두 사용한다. 예를 들면, “24500 ▼a이문열 중단편전집/ ▼d이문열 지음. ▼n4: ▼p우리들의 일그러진 영웅 외”나 “24510 ▼aCielo, viento, estrellas y poes'ia= ▼x하늘과 바람과 별과 詩 / ▼dYun Tong-Ju”에서 “▼p우리들의 일그러진 영웅”과 “▼x하늘과 바람과 별과 詩”를 모두 표제로 추출한다.

합집이나 선집 레코드에서 표제를 추출할 때, 245 필드에 기술된 합집 표제뿐만 아니라 합집에 포함된 개별 표제도 추출한다. 505 필드의 내용주기에 개별 저작을 기술한 경우 개별 저작에 대한 740 필드를 사용하고 부출 표제가 없는 경우 내용주기에 포함된 ▼t 표제를 사용한다.

<표 2> 저작세트 추출의 사례

저작세트	사 례
100+245 ▼a 100+246 ▼a	1001 ▼a이효석, ▼d1907-1942 ▼wKAC200108736 24510 ▼aCuando florece el alforfón / ▼dLee Hyo-seok ; ▼etraducción de Yoon Sun-me, Lee Kang-guk 24619 ▼a메밀꽃필무렵
700+245 ▼a 700+740 ▼a	24500 ▼a한국 대표 단편 소설 13 : ▼b언어의 아름다움과 예술혼 I / ▼d기획·연구: 우한용, ▼e박인기, ▼e정병현, ▼e최병우 ; ▼e역음: 김명환 ... 50500 ▼t돈 -- ▼t메밀꽃 필 무렵 -- ▼t산 / ▼d이효석 ... 소나기 / ▼d황순원 70012 ▼a이효석, ▼d1907-1942 ▼wKAC200108736 740 2 ▼a메밀꽃 필 무렵
700+245 ▼b	24500 ▼a셰익스피어 비극선: ▼b햄릿. 맥베드. 리어왕. 오телло. 로미오와 줄리엣 / ▼d셰익스피어 저; ▼e이태주 [역]; ▼e존 길버트 그림 7001 ▼aShakespeare, William, ▼d1564-1616 ▼wKAC200000024
700+245 ▼a 700+740 ▼a	24520 ▼a(신역) 셰익스피어 전집 / ▼d셰익스피어 著; ▼e韓龍煥 譯 7001 ▼aShakespeare, William, ▼d1564-1616 ▼wKAC200000024 740 2 ▼a햄릿

저자명없이 표제만 있는 경우 <표 3>에서와 같이 표제만으로 저작을 생성하여 군집화한다. 무저자 고전의 경우 <표 4>에서와 같이 130 혹은 730 필드가 입력되어 있는 경우 “저자+표제”의 저작세트 외에 통일표제로 저작을 생성한다.

<표 3> 표제 추출(무저자)의 사례

표제 추출	사 례
245 ▼a	24500 ▼a심청전, ▼a춘향전, ▼a옹고집전 260 ▼a서울: ▼b한국파스퇴르, ▼c2002
	24500 ▼a춘향전 250 ▼a영인본 260 ▼a서울: ▼b문학아카데미, ▼c1990

<표 4> 표제 추출(무저자-통일표제)의 사례

통일표제 추출	사 례
130 ▼a	130 0 ▼a춘향전 24500 ▼a寫眞春香傳/ ▼d朴啓周 著 7001 ▼a박계주
730 ▼a	24500 ▼a열여춘향수절가, ▼a春香傳研究/ ▼d姜慶鎬 編著 7001 ▼a강경호, ▼e편저 730 0 ▼a춘향전

(2) 표현형 식별요소 분석

① 번역

표현형 식별을 위해 <표 5>에서와 같이 008 필드/35-37(언어)과 041 필드(언어부호)를 사용하여 표현형을 구분한다. 즉, 『오만과 편견』은 원서 *Pride and prejudice*에 대한 번역으로, 저작세트 ‘100+245’와 ‘700+246’은 동일한 저작세트 ‘Austen, Jane + Pride and prejudice’를 생성한다.

<표 5> 표현형(번역)의 사례

표현형(번역)	사 례
원서 100 + 245	1001 ▼aAusten, Jane, ▼d1775-1817 ▼wKAC199601201 24510 ▼aPride and prejudice/ ▼dJane Austen ; with a introduction by David Gilson
번역서 700 + 246	008830830s1983 ulk 000 f kor 0411 ▼akor ▼heng 24500 ▼a오만과 편견/ ▼d제인 오스틴 지음; ▼e이상조 옮김 24619 ▼aPride and prejudice 7001 ▼a이상조, ▼e옮김 7001 ▼aAusten, Jane, ▼d1775-1817 ▼wKAC199601201

② 개정

동일 저작에서 판사항은 표현형의 식별요소로 사용한다. <표 6>에서 저작 ‘김소월+진달래꽃’이 개정증보판, 신판, 2판으로 판사항이 각각 다르며, 이는 표현형의 개정으로 사용한다.

<표 6> 표현형(개정)의 사례

표현형(개정)	사 례
250 ▼a	24500 ▼a진달래꽃/▼d김소월 지음
	250 ▼a개정증보판
	260 ▼a서울: ▼b민음社, ▼c1994
	24500 ▼a진달래꽃/▼d김소월 지음
	250 ▼a신판
	260 ▼a서울: ▼b미래사, ▼c2001
	24500 ▼a진달래꽃: ▼b김소월 시집/▼d김소월 [지음]; ▼e김수복 엮음
	250 ▼a2판
	260 ▼a서울: ▼b청개구리, ▼c2006

③ 비평

문학작품에 대한 비평서를 식별하기 위해 008 필드/33(문학형식)에 u(평론)가 입력되어 있는 것은 비평으로 사용한다.

④ 모체-부록

모체레코드와 보유판 및 특별호의 관계를 연결하기 위해 770과 772 필드를 사용한다. 이때 모체와 보유판의 연결은 ISBN이나 제어번호로 연결한다.

⑤ 선행-후속

선행 저록과 후속 저록을 연결하기 위해 780과 785 필드를 사용한다. 이때 선행저록과 후속저록의 연결은 ISBN이나 제어번호를 사용한다.

(3) 구현형 식별요소 분석

구현형은 군집화된 저작과 표현형을 적절한 형식에 맞추어 디스플레이 한다.

① 자료유형

다양한 자료유형을 디스플레이하기 위해 245 필드의 ▼h를 사용한다. 자료유형표시가 없는 것은 단행본이며 자료유형표시가 없는 자료 중 502 필드가 있는 것은 학위논문에 해당한다. 자료유형표시가 있는 것은 ‘녹음자료’, ‘컴퓨터파일’, ‘비디오 녹화자료’ 등 해당 자료유형을 표기한다. 그런데 RDA에서는 자료유형을 보다 세분하여 내용유형, 매체유형, 수록매체유형의 3가지로 구분하여 표시하고 있다. 또한 KORMARC에서도 이를 반영하여 336, 337, 338 필드를 사용하여 표시할 수 있도록 개정되었다. 앞으로는 이 가운데 구현형 식별 요소인 337과 338 필드를 사용하여 자료유형을 디스플레이 하는데 사용할 수 있다.

② 복제

원자료에 대한 복제를 추출하기 위해 534, 580, 776 필드의 유무를 확인하여 원자료와 복제자료를 연결한다.

③ 재판 혹은 증판

250 필드에 재판 혹은 증판으로 입력된 데이터로 동일 저작인 경우 재판의 관계로 연결한다. 250 필드가 입력되어 있지 않지만 저자와 표제, 출판사가 동일한 자료 가운데 출판연도나 페이지가 상이한 경우 쇠의 개념으로 취급하여 구현형의 다른 형태로 표기한다.

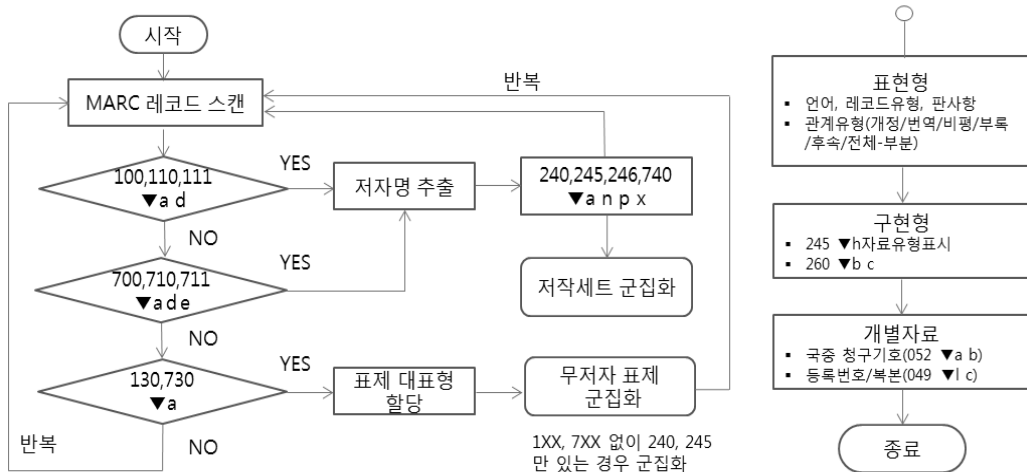
(4) 개별자료 식별요소 분석

개별자료의 사례로는 국립중앙도서관 청구기호 052 필드와 등록번호 및 복본표시 049 필드에 입력된 내용으로 확인한다.

다. 알고리즘 설계

1) 알고리즘 설계 흐름도

FRBR 적용 검색 알고리즘의 흐름을 FRBR의 저작, 표현형, 구현형, 개별자료의 수준별로 살펴보면 <그림 1>과 같다.



<그림 1> FRBR 적용 검색 알고리즘 흐름도

2) 검색키

현재 국립중앙도서관 검색키에는 전체, 제목, 제목(전방), 제목(완전), 저자, 발행자, 청구기호로 구성되어 있으며, 이 연구에서는 전체, 표제, 저자, 발행자를 검색키로 선정하였다.

3) 저작 군집화

저작의 군집화를 위해 저자와 표제의 요소 추출은 <표 7>을 기준으로 하였으며, 표제에 포함된 특수기호(온점, 중간점, 대시 등)는 불용어로 처리하였다.

〈표 7〉 저작(저자와 표제) 추출 기준

저자	100, 110, 111 ▼a ▼d	순서 상관없이 전부 추출
	700, 710, 711 ▼a ▼d ▼e	순서 및 개수 상관없이 전부 추출
표제	240, 245, 246, 740 ▼a ▼n ▼p ▼x	240, 245, 246, 740 필드에서 각 필드 및 필드의 식별기호 개수 상관없이 전부 추출
	505 ▼t	740 ▼a에 표제 부출없이 내용주기에만 기술되어 있을 경우에만 505 필드에서 표제 추출
	130, 730 ▼a	무저자 고전의 경우 저자와 표제를 조합한 저작 외에 통일 표제로 별도의 저작을 군집화

4) 표현형 요소 추출

(1) 자료유형

자료유형에 해당하는 요소를 추출하기 위해 레코드 유형(리더/06)의 값을 사용한다. 일반적인 도서형태는 a(문자자료)로 구분하며 'a'가 아닌 전자자료, 녹음자료, 시청각자료 등은 해당 유형의 값으로 자료유형을 구분한다. 기존의 KORMARC 서지레코드에는 자료유형표시를 245 필드의 ▼h에 기술하고 있으나 여기에는 표현형 요소와 구현형 요소가 혼재되어 있다. RDA에서는 이를 세분하여 표현형 요소인 내용유형, 구현형 요소인 매체유형과 수록매체 유형으로 구분하고 있으며, KORMARC도 이를 반영하여 최근에 수정되어 있다. 즉, 앞으로는 RDA의 내용유형을 구조화한 336 필드를 사용하여 표현형 요소를 추출할 수 있다.

(2) 이용대상자 수준

이용대상자 수준을 구분하기 위해 008 필드/22의 값을 사용하며, 일반이용자(b), 청소년(b, c, d), 아동(j), 특수계층(f)으로 범주화한다.

(3) 문학형식

문학 장르를 표기하기 위해 008 필드/33의 값을 사용하며, 평론(u)의 경우 비평의 관계를 표현하기 위해 사용한다.

(4) 관계유형

개정, 번역, 비평, 부록, 후속, 전체-부분 등과 같은 서지적 관계유형은 대체로 표현형 수준에서 식별한다. 이러한 서지적 관계유형을 식별하기 위한 요소 값은 <표 8>과 같다.

<표 8> 표현형(관계유형) 요소

관계	필드	설명
개정	250 ▽a(판표시)	- '저자, 표제, 출판사' 동일하고 250 필드가 있으면 '개정' 관계로 연결
번역	008/35-37(언어) 041 ▽a(본문언어)	- (저자필드+245)와 (저자필드+24619▽a)를 대조하여 같을 경우 '번역' 관계로 연결 - 번역서의 경우 아래와 같이 입력되어 있어야 함 <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 5px auto;"> 008/35-37 kor 0411 ▽akor ▽hjpn 24500 ▽a이웃집 토토로 24619 ▽aとなりのトトロ </div>
비평	008/33(문학형식) u	- 'u' 입력되어 있는 레코드를 추출하여 저작 비교한 뒤 같을 경우 동일 저작으로 군집화하며, '비평' 관계로 연결
부록	770 ▽t(보유판) 772 ▽t(모체레코드)	- ISBN 혹은 제어번호로 연결
후속	780 ▽a(선행저록) 785 ▽a(후속저록)	- ISBN 혹은 제어번호로 연결
전체-부분	리더/07(서지수준)	- 리더/07(서지수준)로 구분하여 '전체-부분' 관계로 연결 <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 5px auto;"> a 모본에서 분리된 단행자료 성격의 구성요소 b 모본에서 분리된 연속간행자료 성격의 구성요소 c 집서(collection) d 집서의 하위단위 i 갱신자료 m 단행자료/단일자료 s 연속간행물 </div>

5) 구현형 요소 추출

구현형 식별을 위한 추출 요소는 아래와 같이 자료유형, 발행처 및 발행년, 관계유형 등이다.

(1) 자료유형

기존의 KORMARC 서지레코드에는 자료유형표시를 245 필드의 ▽h에 기술하고 있어 이를 구현형 식별 요소로 사용할 수 있지만, 앞으로는 RDA에서 제시하고 있는 내용유형, 매체유형, 수록매체유형의 3가지가운데 매체유형과 수록매체유형을 구현형 식별 요소로 사용할 수 있다. 즉, KORMARC 서지레코드에 RDA를 반영한 337과 338 필드를 사용하여 구현형 요소를 추출할 수 있다.

(2) 발행처 및 발행년

발행처 및 발행년은 260 필드의 ▽b▽c에 기술하고 있어 이를 구현형 식별 요소로 사용한다.

(3) 관계유형

구현형에서 관계유형은 <표 9>의 조건에 따라 데이터를 추출하여 복제, 재판의 관계로 연결한다.

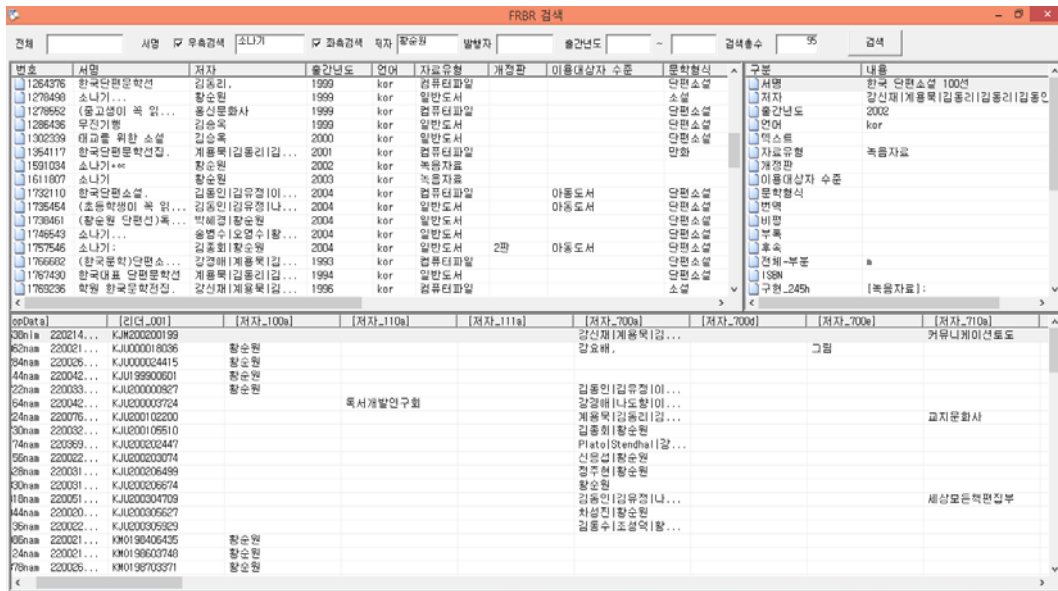
<표 9> 구현형(관계유형) 요소

관계	필드	설명
복제	534 580 776 필드 있는 것	- 534, 580, 776 필드가 있는 것을 추출하여 동일 저작의 '복제' 관계로 연결 - 「장애인대체자료」를 추출하기 위해 [리더/06 'a' + 776 ▼i '시각장애인용'] 으로 입력된 자료를 추출하여 '복제' 관계로 연결
재판	250 ▼a	- '중판', '재판'으로 입력된 것은 '재판' 관계로 연결

2. FRBR 알고리즘 적용

가. 저작세트(소나기+황순원)의 검색 결과

앞서 설계된 FRBR 적용 검색 알고리즘을 구현하여 추출된 레코드의 수는 <그림 2>와 같이 95건이며, 이를 분석한 결과 표현형과 구현형은 <표 10>과 같다. 표현형 자료유형으로 일반도서와 컴퓨터파일이 많았으며, 구현형의 자료유형표시는 텍스트가 대부분이었다. 표현형 개정관계는 2판, 개정판 등 6건의 데이터가 추출되었으며 구현형 재판관계는 중판, 재판이 4건 추출되었다. 문학형식은 단편소설로 입력된 데이터가 가장 많았고 소설, 만화, 문집이 있으며, 미입력 건수가 10건에 이른다.



<그림 2> 저작세트(소나기+황순원)의 추출

〈표 10〉 저작세트(소나기+황순원)의 검색 내용

검색키		표제 “소나기” + 저자명 “황순원”
레코드의 수		95건
표현형	자료유형	일반도서(60건) 녹음자료(3건) 컴퓨터파일(30건) 시청각자료(2건)
	개정관계	6건
	이용대상자	일반이용자(50건) 청소년(17건) 아동(28건)
	문학형식	만화(4건) 단편소설(56건) 소설(21건) 문집(4건) 미입력(10건)
구현형	자료유형표시	텍스트(90건) 녹음자료(3건) 비디오 녹화자료(1건) videorecording(1건)
	재판관계	4건

나. 저작세트(무정+이광수)의 검색 결과

레코드의 수는 88건이 추출되었으며, 분석결과 표현형과 구현형은 <표 11>과 같다. 표현형의 자료유형으로 일반도서와 컴퓨터파일, 전자자료, 녹음자료가 있으며 관계유형으로 개정관계 8건, 비평관계 1건, 부록관계 1건, 전체-부분관계 2건이 추출되었다. 문학형식의 경우 소설이 68건으로 가장 많았으며 미입력이 10건으로 나타났다. 구현형은 텍스트가 85건으로 가장 많았고 재판관계가 4건이었다.

〈표 11〉 저작세트(무정+이광수)의 검색 내용

검색키		표제 “무정” + 저자명 “이광수”
레코드의 수		88건
표현형	자료유형	일반도서(59건) 컴퓨터파일(26건) 전자자료(2건) 녹음자료(1건)
	개정관계	8건
	비평관계	1건 예: (바로잡은)무정 / [이광수 저];김철 校註
	부록관계	1건 예: 빠르게 읽는 현대소설 : 부록. 1-2
	전체-부분관계	2건 예: 代表韓國文學全集.1 - 12 / 金東里;趙演鉉 [공] 責任編輯
	이용대상자	일반이용자(70건) 청소년(7건) 아동(10건) 특수계층(1건)
	문학형식	소설(68건) 만화(1건) 단편소설(5건) 문집(1건) 미입력(10건)
구현형	자료유형표시	텍스트(85건) 전자책(2건) 녹음자료(1건)
	재판관계	4건

다. 저작세트(노인과 바다+Hemingway, Ernest)의 검색 결과

레코드의 수는 117건이 추출되었으며, 분석결과 표현형과 구현형은 <표 12>와 같다. 문학형식이 입력되지 않은 데이터가 18건으로 나타났으며, 『노인과 바다』는 번역서임에도, 표현형 번역관계가 83건에 불과하며 나머지 34건은 원표제 입력이 누락된 것으로 나타났다. 83건 가운데 원저작의 언어를 기입한 건수는 6건에 불과하다. 구현형의 자료유형은 텍스트가

가장 많았으며, 오디오북과 녹음자료도 포함되어 있다.

<표 12> 저작세트(노인과 바다+Hemingway, Ernest)의 검색 내용

검색키		표제 “노인과 바다” + 저자명 “Hemingway, Ernest”
레코드의 수		117건
표현형	자료유형	일반도서(102건), 녹음자료(8건), 시청각자료(1건), 컴퓨터파일(6건)
	개정관계	13건
	번역관계	83건
	이용대상자	일반이용자(83건), 청소년(11건), 아동(23건)
	문학형식	만화(2건), 단편소설(4건), 소설(91건), 문집(2건), 미입력(18건)
구현형	자료유형표시	텍스트(109건), 비디오 녹화자료(1건), 전자자료(1건), 녹음자료(2건), 오디오북(6건)
	복제관계	1건
	재판관계	3건

라. 저작세트(오만과 편견+Austen, Jane)의 검색 결과

레코드의 수는 82건이 추출되었으며, 분석결과 표현형과 구현형은 <표 13>과 같다. 표현형 자료유형에는 일반도서, 컴퓨터파일, 전자자료가 있다. 표현형 번역관계는 61건이며, 나머지 21건은 원표제 입력이 누락되었으며, 61건 가운데 원저작의 언어를 입력한 건수는 7건에 불과하다. 문학 형식의 경우 소설로 입력한 73건을 제외한 9건은 미입력으로 나타났다. 또한 원본자료를 컴퓨터파일로 복제한 구현형 복제관계가 1건 포함되어 있다.

<표 13> 저작세트(노인과 바다+Hemingway, Ernest)의 검색 내용

검색키		표제 “오만과 편견” + 저자명 “Austen, Jane”
레코드의 수		82건
표현형	자료유형	일반도서(77건), 전자자료(1건) 컴퓨터파일(4건)
	개정관계	4건
	번역관계	61건
	이용대상자	일반이용자(71건) 청소년(4건) 아동(7건)
	문학형식	소설(73건) 미입력(9건)
구현형	자료유형표시	텍스트(80건) 컴퓨터파일(1건) 전자책(1건)
	복제관계	1건
	재판관계	1건

마. 저작세트(춘향전)의 검색 결과

『춘향전』의 레코드 수는 <그림 3>과 같이 158건으로 나타났으며, 이 건수는 표제 필드에

“춘향전”으로 입력된 데이터를 추출한 결과이다. 이를 분석한 결과 표현형 자료유형은 <표 14>와 같이 일반도서가 대부분이며, 컴퓨터파일과 녹음자료가 포함된다. 관계유형은 개정 7건, 비평 1건, 전체-부분은 4건이었다. 표현형 이용대상자의 경우 일반이용자가 가장 많았지만 아동용과 청소년용도 많은 건수를 보였다. 문학형식은 소설, 단편소설, 희곡 등의 순이었고 미입력된 데이터도 37건에 이른다. 구현형 재판관계에는 중판, 재판과 영인본이 포함되어 있으며 복제관계는 복제주기 2건이 있다.

번호	서명	저자	출간년도	언어	진공유형	개정판	이용대상...	문학형식	번역	구분	내용
074726	춘향전	김기항	1995	kor	일반도서					서명	내용
693631	춘향전	-	1990	kor	일반도서	영인본				저자	韓國古典文學大系
900059	춘향전	-	1945	kor	일반도서					출간년도	1900
905717	춘향전	구자균	1961	kor	일반도서			소설		언어	kor
912685	(북경)韓國古典文學...	-	1975	kor	컴퓨터파일					학년도	0
912686	韓國古典文學大系	정규래	1975	kor	컴퓨터파일					자료유형	일반도서
913130	韓國古典文學大系	한국고전문학한집...	1980	kor	일반도서					개정판	
913146	韓國古典文學大系	장덕순	1970	kor	컴퓨터파일					이용대상자 수준	
952209	춘향전	김규래	1980	kor	일반도서			소설		문학형식	
956859	춘향전	심성경	1985	kor	일반도서			소설		번역	
1054147	열녀춘향수절가. 1종...	김경호	1990	kor	일반도서	影印...				비판	
1064864	韓國古典文學大系	김규래	1991	kor	컴퓨터파일					부록	
1068398	홍길동전. 1집. 열녀...	여분덕	1991	kor	일반도서			소설		속록	
1076798	조물리는 백학. 1종...	유지건	1949	kor	일반도서			희곡		전체-부분	c
1076126	(原邦)烈女春香傳...	대영달	1951	kor	일반도서			소설		ISBN	
1076127	춘향전	김사영	1952	kor	일반도서					구현_245h	

<그림 3> 저작세트(춘향전)의 추출

<표 14> 저작세트(춘향전)의 검색 내용

검색키		표제 “춘향전”
레코드의 수		158건
표현형	자료유형	일반도서(112건), 컴퓨터파일(44건), 녹음자료(2건)
	개정관계	7건
	비평관계	1건
	전체-부분관계	4건
	이용대상자	일반이용자(80건), 청소년(23건), 아동(55건)
구현형	문학형식	만화(2건), 소설(92건), 단편소설(12건), 희곡(7건), 향가시조가사구비문학(5건), 문집(2건), 미입력(37건)
	자료유형표시	텍스트(155건), 마이크로형태자료(1건), 녹음자료(2건)
	재판관계	2건
	재판관계	10건

3. FRBR 알고리즘 개발을 위한 시사점

앞서 언급한 바와 같이 이 연구에서 알고리즘 설계를 위해 실험데이터를 연구의 목적에 맞게 유형별로 선정해서 분석하였으며, 분석한 결과를 알고리즘 개발에 반영하였다는 점에서 기존 연구와 차이점이 있다고 할 수 있다. 설계된 FRBR 알고리즘의 특징을 살펴보면 아래와 같다.

첫째, 저작은 ‘저자명+표제’로 구성되는데, 저자명 생성과 관련하여 1XX 필드뿐만 아니라 7XX 필드의 ▼a d e를 대상으로, 표제의 경우 240, 245, 246, 740 필드의 ▼a n p x를 대상으로, 통일표제의 경우 130과 730 필드의 ▼a를 대상으로 추출하였다. 합집이나 선집에서 740 필드의 표제부출없이 내용주기에만 기술되어 있는 경우 505 필드도 활용하였다.

둘째, 표현형의 요소로는 자료유형(리더/06), 이용대상자수준(008/22)을 사용하였다. 표현형의 요소뿐만 아니라 표현형의 관계유형 요소도 추출하여 활용하였다. 표현형의 관계유형 요소로 개정(250 ▼a), 번역(008/35-37, 041 ▼a), 비평(008/33 u), 부록(770, 772 ▼t), 후속(780 ▼a), 전체-부분(리더/07)을 추출하였다.

셋째, 구현형의 요소로는 자료유형표시(245 ▼h), 발행처 및 발행년(260 ▼b c)을 사용하였으며, 구현형도 요소뿐만 아니라 관계유형 요소를 추출하여 활용하였다. 구현형의 관계유형 요소로 복제(534, 580, 776), 재판(250 ▼a 중판, 재판)을 추출하였다.

그런데 KORMARC 서지레코드에 있어 FRBR 적용을 위한 알고리즘을 보다 효과적으로 달성하기 위해 몇 가지 사항을 제안하면 아래와 같다.

첫째, FRBR 적용을 위해서는 RDA에서 중요시하는 바와 같이 전거형 접근점의 작성을 강화할 필요가 있다. 인명전거의 중요성은 국내 도서관계에서도 오래전부터 부터 인식하고 나름대로 작성을 하여오고 있었지만, 표제전거는 대부분의 도서관에서 간과되어 왔던 것이 사실이다. 통일표제라는 이름으로 표제전거가 이루어지고 있었지만 실제 도서관 현장에서는 매우 제한적으로 이루어지고 있을 뿐이다. 특히 국내 고전작품의 표제전거를 위한 추가적인 작업이 필요하다고 생각된다.

둘째, 역할어나 연관저록 등의 입력 요소를 강화할 필요가 있다. 지금까지는 번역자, 삽화가, 편역자, 감독자 등과 같은 역할어나 연관저록 등의 요소가 검색에서 그다지 활용되지 않은 관계로 대부분의 도서관에서 실제 목록작업에 있어서도 생략되는 경우가 많았다. 그렇지만 FRBR 및 RDA에서는 자료 식별을 위한 활용을 고려하여 이러한 요소들을 강화하는 경향이 있다. 특히 구현형의 표현이 기계적 알고리즘으로 유연하게 이루어지기 위해서는 요약, 개작, 변형, 비평과 같은 서지적 관계유형도 일반주기보다는 연관저록으로 기술될 수 있도록 확대하는 방안을 고려해 볼 수 있다.

셋째, 저작세트를 구성하여 검색한 결과 특히 구현형 수준에서 검색되어야 할 상당수의 자

료, 예를 들면, 온라인자료(전자책, 원문 등)와 디지털화하여 서비스하는 소장원문자료 등은 실제로는 MODS로 구조화되어 있다. 이 연구의 알고리즘은 MARC 기반이며, MODS를 기반으로 한 레코드에 대해서도 통합할 수 있는 추가적인 연구가 필요하다고 생각된다.

넷째, FRBR의 적용효과가 가장 높다고 기대되는 문학류를 대상으로 알고리즘을 개발하였다. 이를 활용할 시에도 OCLC와 같이 문학에 한정하여 일정기간 시행과정을 거친 후, 효과성이 높을 것으로 예상되는 분야에 대해 단계적으로 검토해야 할 것이다.

다섯째, 국제목록원칙규범의 기본이 되고 있는 이용자 편의성을 달성하기 위해서는 어떠한 요소를 입력하느냐, 얼마만큼 입력했느냐에 달려 있다. FRBR 알고리즘도 제대로 기능을 하기 위해서는 기본적으로 원래의 레코드 제대로 작성되어 있어야 한다. KORMARC 입력 데이터의 수준을 분석한 결과 현행 목록 데이터로는 FRBR을 위한 저작과 표현형을 구분하기 어렵고, 전거나 서지적 관계요소 등이 생략되어 있는 경우가 많다. FRBR를 적용하기 위해 기존의 레코드입력 수준을 전반적으로 강화할 필요가 있다.

IV. 결론

KORMARC 서지레코드의 FRBR 적용 알고리즘을 개발하여 문학류를 대상으로 실험한 결과, 기존의 검색 프로그램에 비해 관련 자료를 보다 망라적이고 체계적으로 제시할 수 있다. 알고리즘의 설계내용과 함께, 연구결과를 요약하면 아래와 같다.

첫째, OCLC의 FRBR Work-Set Algorithm과 LC의 FRBR Display Tool Version 2.0 등을 비롯한 국내외 FRBR 구현 알고리즘 개발 사례의 장단점을 비교 분석하여 이 연구의 기반으로 활용하였다.

둘째, FRBR에서 구분하고 있는 저작 유형에 따라 저작, 표현형, 구현형, 개별자료의 4가지로 구분하여 알고리즘을 개발하였다.

셋째, 국립중앙도서관의 국가서지레코드에서 추출한 실험데이터를 분석하여 저작, 표현형, 구현형, 개별자료의 식별요소를 분석하였다. 저작 식별요소로 ‘저자+표제’, 표현형 식별요소로 번역, 개정, 비평, 모체-부록, 선행-후속, 구현형 식별요소로 자료유형, 복제, 제판 또는 중판, 개별자료 식별요소로 청구기호, 등록번호, 복본기호를 선정하였으며, 여기에 해당하는 KORMARC 요소를 분석하여 제시하였다.

넷째, 저작세트별로 관련 저작을 군집화하기 위해 RDA에서 저작세트를 군집화하기 위한 기본 장치로 적용하고 있는 전거형 접근점을 사용하였으며, ‘저자+표제’와 같은 전거형 접근점을 작성할 수 있는 알고리즘을 설계하였다.

한편 FRBR 알고리즘을 개발하면서 국가서지레코드를 분석한 결과, KORMARC 서지레코드를 대상으로 FRBR 알고리즘을 적용하기 위해서는 기존의 서지레코드를 정비하고 레코드의 입력 수준을 전반적으로 강화할 필요가 있다.

참고문헌

- 국립중앙도서관. 2007. 표준 메타데이터 모델 및 시맨틱 언어 연구. 『국립디지털도서관 운영 시범사업』. 서울: 국립중앙도서관.
- 김정현. 2007. 서지적 관계를 기반으로 한 한국어 도서의 저작유형 분석. 『한국도서관·정보학회지』, 38(3): 183-200.
- 김정현. 2007. 한국어 서지레코드에 있어 FRBR 모형의 유용성에 관한 연구. 『한국문헌정보학회지』, 41(4): 295-314.
- 김현희, 유영준, 박서은. 2007. FRBR 모형의 KORMARC 데이터베이스로의 적용 가능성에 대한 실험적 연구 : 음악자료를 중심으로. 『한국도서관·정보학회지』, 38(2): 185-202.
- 노지현. 2008. KORMARC 레코드에 대한 FRBR 모형의 적용 실험: 국립중앙도서관 서지레코드를 사례로 하여. 『한국도서관·정보학회지』, 39(2): 291-312.
- 이미화. 2008. 『저작 클러스터링 분석을 통한 FRBR의 목록 적용에 관한 연구』. 박사학위논문, 이화여자대학교 대학원 문헌정보학과.
- 이성숙. 2004. 『FRBR 모형의 수용에 관한 연구』. 박사학위논문, 연세대학교 대학원 문헌정보학과.
- 이성숙. 2006. FRBR 모형의 적용전략에 관한 연구 : 해외사례를 중심으로. 『한국문헌정보학회지』, 40(3): 305-331.
- 이유정. 2007. FRBR 실행을 위한 MARC 레코드 생성에 관한 연구. 『한국도서관·정보학회지』, 38(1): 299-314.
- 조재인. 2004. FRBR 알고리즘 분석 및 KORMARC 데이터베이스 적용방안. 『한국문헌정보학회지』, 38(3): 5-21.
- 조재인. 2005. 표현형 계층을 중심으로 한 FRBR 모형분석 및 목록 체계 수용에 관한 연구. 『한국도서관·정보학회지』, 36(2): 221-239.
- 谷口祥一. 2009. FRBR OPAC構築に向けた著作の機械的同定法の検証: JAPAN/MARC書誌レコードによる実験. *Library and Information Science*, 61: 119-151.
- 宮田洋輔. 2008. JAPAN/MARCレコードから自動構築可能, な著作識別子の提案. 『三田図書館情報学会研究大会発表論文集』. 2008: 69-72.

- Aalberg, Trond. 2006. "A Process and Tool for the Conversion of MARC Records to a Normalized FRBR Implementation." *Lecture Notes in Computer Science*, 4312, 283–292.
- Aalberg, Trond, Ole Husby and Frank Berg Haugen 2006. "A Tool for Converting from MARC to FRBR." *Lecture Notes in Computer Science*, 4172: 453–456.
- Carlyle, Allyson, Sara Ranger, Joel Summerlin 2008. "Making the Pieces Fit: little Women, Works, and the Pursuit of Quality." *Cataloging & Classification Quarterly*, 48(1): 35–63.
- Chang, Naicheng and Yuchin Tsai. 2011. "FRBRisation of Koha in the Context of CMARC: a UNIMARC-derived Format." *77th IFLA General Conference and Assembly*. [Conference paper]. 1–15.
- Chang, Naicheng et al. 2013. "Experimenting with implementing FRBR in a Chinese Koha system." *Library Hi Tech News*, 30(10): 10–20.
- Cho, Jane. 2006. "A Study on the Application Method of the Functional Requirements for Bibliographic Records (FRBR) to the Online Public Access Catalog (OPAC) in Korean Libraries," *Library Collections Acquisitions & Technical Services*, 30(3/4): 202–213.
- Delsey, T. 2002. *Functional Analysis of the MARC 21 Bibliographic and Holdings Formats*. <www.loc.gov/marc/marc-functional-analysis/analysis.pdf> [cited 2015. 1. 15].
- Hickey, Thomas B. & Jenny Toves. 2009. *FRBR Work-Set Algorithm: Version 2.0*. Dublin, Ohio: OCLC Research. <<http://www.oclc.org/.../2009-08.pdf>> [cited 2015. 1. 15].
- Hickey, T.B. and E.T. O'Neill. 2005. "FRBRizing OCLC's WorldCat." *Cataloging & Classification Quarterly*. 39(3/4): 239–251.
- Hickey, Thomas B, Edward. T. O'Neill & Jenny Toves. 2002. "Experiments with the IFLA Functional Requirements for Bibliographic Records (FRBR)." *D-Lib Magazine*, 8(9). <<http://www.dlib.org/dlib/september02/hickey/09hickey.html>> [cited 2015. 1. 15].
- Joint Steering Committee for Development of RDA. 2010. *RDA : Resource Description & Access*. Chicago: American Library Association.
- Library of Congress. 2009. *FRBR Display Tool Version 2.0*. <<http://www.loc.gov/ma>

- rc/marc-functional-analysis/tool.html> [cited 2015. 1. 15].
- Library of Congress. 2003. *FRBR Display Tool*. <<http://www.loc.gov/marc/marc-functional-analysis/tool.html>> [cited 2015. 1. 15].
- O'Neill, Edward T. 2002. "FRBR: Functional Requirements for Bibliographic Records Application of the Entity-Relationship Model to Humphry Clinker." *Library Resources & Technical Services*, 46(4): 150-159.
- Rajapatirana, Bemal and Roxanne Missingham. 2005. "The Australian National Bibliographic Database and the Functional Requirements for the Bibliographic Database (FRBR)." *The Australian Library Journal*, 54(1): 31-42.
- Yee, Martha M. 2005. "FRBRization: a Method for Turning Online Public Finding Lists into Online Public Catalogs." *Information Technology and Libraries*, 24(2): 77-95.
- Zhang, Yin & Athena Salaba. 2009. *Implementing FRBR in Libraries: Key Issues and Future Directions*. New York: Neal-Schuman.

국한문 참고문헌의 영문 표기

(English translation / Romanization of reference originally written in Korean)

- Cho, Jane. 2004. "Study on the FRBR Algorithm and Application of KORMARC Database." *Journal of the Korean Society for Library and Information Science*, 38(3): 5-21.
- Cho, Jane. 2005. "The Study on the FRBR Adoption into Cataloging Rule Focused on its Expression Level." *Journal of Korean Library and Information Science Society*, 36(2): 221-239.
- Kim, Hyun-Hee, Yeong-Jun Yoo and Suh-Eun Park. 2007. "An Experimental Study on the FRBR Model Adaptation to KORMARC Database : Focusing on Music Materials." *Journal of Korean Library and Information Science Society*, 38(2): 185-202.
- Kim, Jeong-Hyen. 2007. "An Analysis on the Work Types of Korean Books Based Bibliographical Relationship." *Journal of Korean Library and Information Science Society*, 38(3): 183-200.
- Kim, Jeong-Hyen. 2007. "A Study on the Utility of FRBR Model in Korean Bibliographic Record." *Journal of the Korean Society for Library and*

- Information Science*, 41(4): 295-314.
- Lee, Mihwa. 2008. *A Study of FRBR Implementation to Catalog by Using Work Clustering*. Ph. D. diss., Ewha Womans University, Korea.
- Lee, Sung-Sook. 2004. *A Study on the Adoption of the FRBR Model*. Ph. D. diss., Yonsei University, Korea.
- Lee, Sung-Sook. 2006. "A Study on the Application Strategies of the FRBR Model." *Journal of the Korean Society for Library and Information Science*, 40(3): 305-331.
- Lee, You-Jeong. 2007. "A Study on Creation of MARC Record for FRBR Implementation." *Journal of Korean Library and Information Science Society*, 38(1): 299-314.
- Rho, Jee-Hyun. 2008. "An Application of FRBR Model to KORMARC Records." *Journal of Korean Library and Information Science Society*, 39(2): 291-312.