

KORMARC 레코드에 대한 FRBR 모델의 적용 실험*

- 국립중앙도서관 서지레코드를 사례로 하여 -

An Application of FRBR Model to KORMARC Records

노 지 현(Jee-Hyun Rho)**

〈 목 차 〉

I. 서론	2. KORMARC의 개체 식별 요소
1. 연구의 배경과 목적	IV. 표본 레코드에 대한 FRBR 모델의 적용
2. 연구의 방법과 대상	1. 표본 레코드에서 드러난 특징
II. FRBR 모델의 적용에 관한 기존 논의	2. 표본 레코드를 이용한 FRBR 모델의 적용
1. 해외에서의 적용 논의	3. 적용 결과에서 드러난 한계와 문제점
2. 국내에서의 적용 논의	V. 결론
III. FRBR 개체와 KORMARC 데이터요소	
1. 제1집단 개체와 MARC21의 식별 요소	

초 록

이 연구에서는 (1) FRBR 모델로의 기계적 변환을 목적으로 현재까지 진행해왔던 핵심 논의들을 살펴보고, (2) 다양한 실험과정을 거쳐 MARC 레코드에 적용하기 위해 개발한 FRBR 알고리즘을 검토한 후, (3) 이러한 자동 알고리즘을 우리의 목록레코드에 적용하고자 할 때 발생하는 문제와 그 문제의 근본 원인을 구체적인 사례를 통해 드러내 보이면서, (4) 향후 FRBR 모델의 능동적 수용을 위해 특히 고려해야 할 사항을 논의해 보고자 하였다. 이를 위해, 이 연구에서는 국립중앙도서관에서 생산한 서지레코드 중에서 '햄릿'에 관한 161건의 표본 레코드를 추출하여 FRBR 모델을 적용해 보는 실험을 수행하였다.

키워드: FRBR 모델, 서지레코드의 기능상의 요건, KORMARC 레코드, FRBR과 KORMARC 데이터요소, FRBR 알고리즘

ABSTRACT

This study intends to make a comprehensive investigation on an application of FRBR model to KORMARC records. To the end, issues associated with FRBR implementation are analyzed, and automatic algorithms to covert bibliographic records into FRBR model are examined thoroughly. On the basis of it, this study implemented four-level hierarchical bibliographic structure of FRBR model using existing KORMARC records. 161 sample records for an experiment were selected from Korea national library. The work chosen as sample records was "Hamlet." Results suggest that considerable difficulties would be encountered in the process of FRBRizing KORMARC records.

Keywords: FRBR Model, Functional Requirements for Bibliographic Records, FRBR Algorithm, FRBRization, Library Catalogs, KORMARC Records

* 이 논문은 부산대학교 자유과제 학술연구비(2년)에 의하여 연구되었음.

** 부산대학교 문헌정보학과 부교수(jrho@pusan.ac.kr)

• 접수일: 2008년 5월 23일 • 최종심사일: 2008년 5월 26일 • 최종심사일: 2008년 6월 21일

I. 서론

1. 연구의 배경과 목적

1998년 IFLA에서 발표한 FRBR 모델은 오늘날 서지세계의 다양하고 복잡한 구조를 반영하고, 지금까지 도서관계에서 추구해 온 목록의 본질적 기능을 달성하는데 유용한 모델이라는 긍정적인 평가를 받고 있다. 그러나 모델 자체가 안고 있는 잠재적 긍정성과는 별도로, FRBR 모델의 실행이 향후 편목업무의 절차와 방식에 미치게 될 파장에 대한 우려도 제기되고 있다. FRBR 모델에서 제시하고 있는 구조는 기존의 방식과 상이하여 업무과정에 상당한 혼란을 초래할 뿐만 아니라 레코드 구축 단계에서 편목사서에 부과되는 업무 부담이 과도할 것으로 예측되고 있기 때문이다. 이렇듯 업무의 혼란과 과중이 우려되고 있는 가운데, 이미 실험 모델의 연구·개발에 착수한 기관에서는 이 모델의 실행을 편목업무를 수행하는 과정에서가 아니라 별도의 '기계적인 작업'을 통해 구현하게 하는 대안을 제시하고 있다. 즉, 도서관에서는 현재의 방식대로 레코드를 생성하되, 차후에 FRBR 모델의 각 개체별로 기계적으로 모으기 위한 방법을 택하는 것이 더욱 효율적이라고 판단하고 있는 것이다. 더불어, 모든 자료를 대상으로 FRBR 모델을 적용하는 것보다 이 모델의 적용을 통해 '이득'이 되는 특정 자료만을 대상으로 모델을 실행할 것을 권고하고 있다.

이에 따라 각국의 서지기관에서는 기존의 MARC 레코드를 FRBR 모델로 자동 변환하기 위한 알고리즘을 개발하고, 자료들 간에 다양한 서지적 관계를 가지고 있어 FRBR 모델을 적용하면 특히 유용할 것으로 판단되는 '문학작품'이나 '음악자료' 등을 대상으로 우선적으로 모델의 구현을 시도하고 있다. 국내에서도 FRBR 모델이 현행 목록체계가 가지는 한계와 문제를 극복하고, 디지털과 네트워크로 대변되는 시대의 특성에 걸맞은 새로운 목록체계라는데 인식을 같이하면서, 최근 들어 모델의 실행에 상당한 관심을 보이고 있다. 그러나 아직까지는 FRBR 모델을 적용했을 때의 효과를 개념적 수준에서 고찰하거나 OCLC나 LC에서 개발한 알고리즘을 이용한 '수정' 알고리즘을 제안하는 연구들이 있을 뿐 서지레코드의 품질검증이 제대로 되어 있지 않고, 특히 기본표목과 통일표제를 적용하지 않고 있는 우리의 목록 환경에서 FRBR 모델을 실제 적용하고자 할 때 봉착할 수 있는 장애와 한계에 대해 분석한 연구는 거의 전무한 상태라 할 수 있다.

이러한 배경에서, 이 연구에서는 우리 도서관에서 생산한 서지레코드를 대상으로 FRBR 모델로의 기계적 변환을 시도할 때 발생할 수 있는 한계와 문제점에 대해 집중적으로 논의해 보고자 한다. 이미 국내 연구진들은 우리 목록의 여건을 감안할 때 기계적인 방법만으로 클러스터링하는 것은 불가능하며, 직접 육안으로 실물 자료와 서지레코드를 일일이 대조하는 과정을 거쳐야 하는 등 다소 까다로운 작업을 수반할 것이라는 예측을 조심스럽게 내놓고 있다. 그러나 이와 관련한 제반 문제를 체계적으로 조사한 사례가 없어, 우리의 서지레코드를 FRBR 모델로 변환하고자 할 때 발

생하는 근본적인 문제에 대해 폭넓고 깊이 있는 논의는 이루어지지 못하고 있다. 이에 이 연구에서는 FRBR 모델의 실행에 있어 선도적인 입장에 있는 주체들이 개발한 FRBR 알고리즘을 참조로 하여 이러한 알고리즘을 KORMARC 레코드에 대입해 보는 과정을 세부적으로 추적하면서, 이 과정에서 발생할 수 있는 실질적인 문제점이나 어려움을 구체적인 사례를 통해 예측해 보고자 한다.

2. 연구의 방법과 대상

이 연구의 목적은 FRBR 모델의 본격적인 실행에 앞서, 이 모델을 우리의 목록 환경에 적용하고자 할 때 나타날 수 있는 제반 문제를 분석해 보는데 있다. 이를 위해, 이 연구에서는 (1) FRBR 모델로의 기계적 변환을 목적으로 현재까지 진행해왔던 핵심 논의들을 살펴보고, (2) 다양한 실험 과정을 거쳐 MARC 레코드에 적용하기 위해 개발한 FRBR 알고리즘을 검토한 후, (3) 이러한 자동 알고리즘을 우리의 목록레코드에 적용하고자 할 때 발생하는 문제와 그 문제의 근본 원인을 구체적인 사례를 통해 드러내 보이면서, (4) 향후 FRBR 모델의 능동적 수용을 위해 특히 고려해야 할 사항을 논의해 보고자 하였다.

구체적으로, 이 연구에서는 국립중앙도서관의 목록레코드를 이용하여 FRBR 모델에서 제시한 제1집단에 대한 군집화를 시도하였으며,¹⁾ 분석대상으로 삼은 표본 레코드는 무작위 추출보다 의도적 방법에 따라 표집하였다. 이 연구에서는 “햄릿”으로 표제 검색한 결과를 표본 레코드로 활용하였는데, 그 이유는 앞서도 언급하였듯이 문학작품의 경우 원본과 다양한 번역본이나 판, 해설이나 비평 등과 같이 서지적 관계를 가진 다수의 저작을 보유하고 있어 FRBR 모델을 적용하였을 경우 특히 유용할 것으로 인식되고 있기 때문이다. 이에 문학작품 중에서도 고전에 해당하는 “햄릿”을 대상으로 하여 기존의 서지레코드를 FRBR 모델로 변환하는 과정을 체계적으로 분석하면, 다양한 유형의 자료를 대상으로 모델의 적용을 계획할 때 발생할 수 있는 어려움을 미리 예측해 볼 수 있을 것이라 판단하였다. 결과적으로, 이 연구에서는 국립중앙도서관에서 생산한 서지레코드(도서, 연속간행물, 시청각자료, 녹음자료 등) 중에서 161건의 표본 레코드를 추출하여 FRBR 모델을 적용해 보는 실험을 수행하였다.²⁾

-
- 1) 국립중앙도서관의 목록레코드를 분석 대상으로 정한 결정적인 이유는 두 가지이다. 첫째, FRBR 모델의 실행과 관련하여 단위도서관에서 개별적으로 실행하는 것보다는 국가도서관이나 주요 서지기관에서 모델을 실행한 후 이 결과를 단위도서관들이 공유하도록 하는 것이 바람직하다는 의견이 현재로서는 지배적이며, 둘째, KERIS와 같은 서지기관에서 구축한 종합목록은 개별 도서관의 상이한 편목관행으로 인해 발생하는 변수가 많으므로 이러한 변수를 통제하기 위해서는 하나의 도서관을 사례로 하여 분석하는 것이 더 적합하다고 판단하였기 때문이다.
 - 2) 이 연구에서는 실물 자료와의 대조과정을 거치지 않고, 현행 서지레코드에 기술된 내용만을 대상으로 하여 FRBR 모델로의 변환을 시도하였다. FRBR 모델을 대규모 서지 DB에 적용하기 위해서는 기존의 서지레코드를 실물 자료와 일일이 대조하는 것이 사실상 불가능하고, 더구나 데이터에 대한 대폭적인 수정이나 보완 작업이 전제되지 않는 한 현행 레코드를 대상으로 모델의 실행을 계획할 수밖에 없다. 이러한 점을 고려하여, 이 연구에서는 가급적이면 현행 서지레코드에 기술된 데이터만을 활용하고자 하였다.

II. FRBR 모델의 적용에 관한 기존 논의

1. 해외에서의 적용 논의

FRBR 모델이 발표된 후 OCLC에서는 이 모델의 실행가능성에 대한 실험을 시작하였다. 그 첫 번째 실험이 WorldCat DB에 있는 레코드를 FRBR 모델에서 제시한 '저작'과 '표현형'으로 군집하는 것이었다. 실험의 결과는 현행 레코드로부터 '저작'을 확인하는 것은 가능하나, '표현형'을 식별하기 위한 단서가 충분하지 않아 현행 레코드의 내용만으로는 표현형에 속하는 중요한 유형이 누락될 수 있으며, 특히 편목 관행이나 출판 당시의 상황에 따라 상이한 표현형이 동일한 표현형으로 간주될 수 있는 것으로 나타났다.³⁾ 이러한 실험 결과에 근거하여, OCLC에서는 현행 레코드로부터 '저작' 계층을 클러스터링하기 위한 알고리즘의 개발에 착수하였다. "FRBR Work-set Algorithm"이라 명명된 이 알고리즘은 '진거레코드'를 활용하여 표목을 자동 생성하고, 이를 '서지레코드'로부터 추출한 데이터 세트와 대조한 후, 동일한 표목을 가질 경우 동일 저작으로 간주하는 방법으로 설계되었다.⁴⁾ 표본 레코드를 추출하여 알고리즘에 적용한 결과, WorldCat에 있는 4,700만 레코드 중에서 상이한 저작으로 간주될 수 있는 것은 3,200만 레코드이며, 이 중 78%에 해당하는 2,500만 레코드는 하나의 구현형만을 가진 '단일저작'인 것으로 밝혀졌다.⁵⁾ 이처럼 OCLC에서는 FRBR 모델의 실행가능성을 타진하는 첫 번째 과정으로서 구현형에 해당하는 현행 레코드로부터 모델의 최상위 계층인 저작을 자동 클러스터하기 위한 알고리즘을 개발하는데 성공하였다. 그러나 표현형을 식별하기 위한 정교한 알고리즘의 개발은 불가능하다고 결론내리고, 모델의 구현 과정에서 표현형은 명시적으로 식별이 가능한 '언어'로 구분하는 수준에 그쳤다.⁶⁾

3) Thomas B. Hickey, Edward T. O'Neill, and Jenny Toves, "Experiments with the IFLA Functional Requirements for Bibliographic Records(FRBR)." *D-Lib Magazine*, Vol.8, No.9(September 2002) ; Edward T. O'Neill, "FRBR: Application of the entity-relationship model to Humphry Clinker." *Library Resources and Technical Services* Vol.46, No.4(October 2002).
<http://www.oclc.org/research/publications/archive/2002/oneill_frbr22.pdf> [cited 2008. 4. 10].

4) Thomas B. Hickey, Jenny Toves, *FRBR Work-Set Algorithm*, OCLC(April 2005).
<http://www.oclc.org/research/software/frbr/frbr_workset_algorithm.pdf> [cited 2008. 3. 20].

5) Rick Bennett, Brian F. Lavoie, Edward T. O'Neill, "The concept of a work in WorldCat : an application of FRBR," *Library Collections, Acquisitions, & Technical Services*, Vol.27(2003), pp.49-50.

6) 이와 관련하여, OCLC의 O'Neill은 저작 "Humphry Clinker"로 군집된 179건의 레코드를 대상으로 표현형으로 간주할 수 있는 모든 유형을 일일이 확인한 결과를 발표하였다. 총 48종의 표현형이 발견되었는데, 그가 여기서 표현형 개체로 분류한 것은 원작(1종), 번역본(8종), 원작의 저자가 아닌 다른 사람이 서문을 쓰거나 편집을 한 경우(15종), 삽화가의 그림이 삽입된 경우(13종), 원작의 저자가 아닌 다른 사람이 서문을 쓰고 삽화가 삽입된 경우(11종) 등이었다(주지하다시피 표현형은 동일 저작에 대한 내용상, 형식상의 차이를 반영한 개념이다. 따라서 원작에 다른 사람이 서문을 쓴 경우나 편집자의 의견이 추가된 경우, 그리고 삽화가의 그림이 수록된 경우에 이들 각각을 '동일 저작에 대한 상이한 표현형'으로 간주하고 있다). OCLC에서는 이 과정을 토대로 표현형을 식별하기 위한 자동 알고리즘을 개발하고자 하였으나 최종적으로 이 작업이 현재로서는 불가능하다고 결론 내렸다. 그 이유는 첫째, 현행 서지레코드에 기술된 데이터만으로 표현형

FRBR 모델의 실행에 큰 관심을 보이고 있는 LC에서도 MARC 레코드에 대한 기능 분석을 실시하고, MARC21을 FRBR로 디스플레이하기 위한 툴("FRBR Display Tool")을 개발하였다. 이 툴은 검색 과정을 통해 화면에 출력된 레코드를 의미 있는 구조로 재배열함으로써 이용자들이 원하는 자료를 보다 편리하게 선택하는데 도움을 주고자 하는 목적에서 개발되었다. 구체적으로, MARC21 형식을 따르는 서지레코드에서 '저자 + 표제' 또는 '표제'를 추출한 후 이들 데이터를 저작을 판별하기 위한 근거로 삼은 다음,⁷⁾ 동일 저작으로 분류된 레코드를 대상으로 다시 '자료유형'과 '언어'로 나누어 차하위 계층인 표현형을 구성하도록 설계하였다. 그리고 동일 표현형에 속하는 구현형 레코드로부터 특정 데이터를 추출한 후 이들 데이터를 저작-표현형-구현형 순으로 계층적으로 제시하도록 하였다. 이처럼 LC에서 개발한 툴은 MARC 레코드를 한꺼번에 FRBR 모델로 변환하기 위한 것이 아니라 OPAC 검색결과를 FRBR 모델에 따라 계층적 구조로 제시함으로써 이용자들이 보다 편리하게 서지DB를 탐색하도록 하는데 의의를 두었다. 그러나 MARC 레코드에 기술된 데이터(특히 표목 부분)의 품질이 전제되지 않는다면 동일 개체로의 클러스터 작업이 불가능하게 되므로, 이 툴은 데이터의 정확성과 일관성이 보장되지 않을 경우에는 제대로 구현할 수 없다는 한계를 가지고 있다.

UCLA의 Martha Lee는 현재의 AACR2와 MARC21에 FRBR의 각 개체를 식별하기 위한 요소가 있는지를 분석하였다.⁸⁾ 그녀는 서지레코드와 전거레코드에서 저작과 표현형을 식별하기 위한 요소를 각각 추출하였는데, OCLC나 LC에서 저작을 식별하기 위한 기준으로 삼았던 '저자'나 '표제' 이외에 '분출표목'이나 '주제명표목' 등도 저작을 식별하는데 중요한 요소가 될 수 있다는 주장을 내놓았다. 그러나 표현형에 대해서는 현행 레코드에 FRBR 모델의 표현형에 해당하는 데이터 요소를 이미 가지고 있음에도 불구하고, 이는 인간의 인지적 과정을 통해서만 식별할 수 있는 것이지 컴퓨터에 의해 기계적으로 인식하는 것은 불가능하다는 사실을 지적하였다. 다만, '통일표제'를 가지고 있는 레코드라면, 통일표제 다음에 나오는 언어나 판(version), 날짜 등을 표현형을 식별하는 중요한 단서로 활용할 수 있을 것이라고 예측하였다.

을 식별하기가 어려워 실물 자료와의 대조가 반드시 필요하며, 둘째, 표현형을 식별할 수 있는 요소가 레코드 전반에 흩어져 있거나 일관되게 적용되지 않고 있어 사람이 직관적으로 판단하는 것은 가능할지언정 이를 일률적으로 알고리즘에 반영하는 것은 어렵다는 점 등이었다. 이에 따라 OCLC에서는 FRBR 모델의 실행에서 '차이'의 정도를 분명히 드러내 놓고 있어 알고리즘의 적용이 비교적 가능한 '다양한 언어로 표현된 번역본'에 대해서만 표현형을 적용할 것을 잠정적으로 결정하였다(Edward T. O'Neill, "FRBR: Functional Requirements for Bibliographic Records Application of the Entity-Relationship Model to Humphry Clinker," *Library Resources & Technical Services* Vol.46, No.4(Oct. 2002), pp.150-159).

7) OCLC와 달리, LC에서 저작을 식별하기 위한 기준 키(key)로서 '서지레코드'를 활용한 이유는 서지레코드의 표목부가 전거레코드를 통해 이미 통제된 상태의 데이터를 수록하고 있기 때문이었다.

8) Martha M. Lee, "FRBRization: a method for turning online public finding lists into online public catalogs," *Information Technology and Libraries*(June 2005), pp.77-95.

2. 국내에서의 적용 논의

국내에서 'KORMARC 레코드를 활용한 FRBR 모델의 자동 구현'에 대해 논의한 것으로는 조재인, 김현희, 김정현 등의 연구가 있다. 먼저, 조재인은 국내에서 구축한 KORMARC 데이터베이스에 LC와 OCLC에서 개발한 FRBR 알고리즘의 적용 가능성을 검토하였다. 이 연구에서는 기본표목과 전거레코드를 정확하게 적용하고 있어 이를 토대로 알고리즘을 개발할 수 있었던 LC나 OCLC와는 달리, 표목의 일관성이 전제되지 않고, 전거통제 과정 없이 관련 자료가 분산되어 있는 우리 목록의 경우 이들 알고리즘에서 제시하는 방법만으로 클러스터링하는 것은 불가능하며, 우리의 경우 보다 정교한 알고리즘의 개발과 동시에 전문가의 육안 식별에 절대적인 시간과 비용을 투자해야 할 것이라고 예측하였다.⁹⁾ 또한, 김현희 등은 LC의 FRBR Display Tool을 참조하여 KORMARC에 적용할 알고리즘을 설계한 후 표본으로 선정한 107건의 음악자료 서지레코드를 대상으로 FRBR 모델의 적용 가능성을 실험적으로 검증하였다.¹⁰⁾ 이 연구 또한 앞서 조재인의 연구와 마찬가지로 KORMARC 레코드에 FRBR 모델을 적용하기 위해서는 전거파일의 활용과 레코드의 일관성 및 정확성이 반드시 전제되어야 함을 지적하고, 특히 음악자료에 대한 표현형을 식별하는 것이 현재의 KORMARC 레코드에 수록된 데이터만으로는 상당히 부족하므로 이러한 한계를 극복하기 위해서는 향후 표현형을 식별하기 위한 적절한 데이터를 추가해야 할 것이라고 제안하였다.

국내에서 처음으로 '저작' 계층으로 서지레코드를 군집하는 작업을 시도한 김정현은 국립중앙도서관에서 표본으로 추출한 1,000건의 KORMARC 레코드를 분석한 결과, 한 저작당 표현형은 평균 1.22개, 그리고 구현형은 평균 1.45개로 나타났으며, 특히 문학작품의 경우 한 저작당 표현형과 구현형의 수가 상대적으로 높은 것으로 확인되었다고 보고하였다.¹¹⁾ 그는 이 연구에서 하나의 저작이 하나의 구현형 밖에 가지지 않는 단일저작의 비율이 전체의 83.1%를, 그리고 복수의 구현형을 가지는 저작의 비율이 16.9%를 차지하고 있음을 밝혀냄으로써 이들 16.9%에 이르는 자료에 FRBR 모델을 적용할 경우 특히 유용할 것이라는 사실을 보여주었다. 또한, FRBR 모델에서 상이한 저작으로 간주하고 있는 후속, 속편, 부록, 보유, 개작, 변형, 모방 등을 모두 동일한 저작의 범주로 간주하여 이를 '동일 저작에 대한 상이한 표현형'으로 취급하는 것이 모델의 유용성을 높일 수 있는 방안임을 실증적으로 보여주었다.

이외에도 FRBR 모델이 개발된 이후 현행 목록레코드에 이 모델을 실행하기 위한 시도가 국

-
- 9) 조재인, "FRBR 알고리즘 분석 및 KORMARC 데이터베이스 적용 방안," 한국문헌정보학회지, 제38권, 제3호(2004. 9), pp.1-21.
10) 김현희·유영준·박서은, "FRBR 모형의 KORMARC 데이터베이스로의 적용 가능성에 대한 실험적 연구: 음악 자료를 중심으로," 한국도서관·정보학회지, 제38권, 제2호(2007. 6), pp.185-202.
11) 김정현, "한국어 서지레코드에 있어 FRBR 모형의 유용성에 관한 연구," 한국문헌정보학회지, 제41권, 제4호(2007. 12), pp.295-314.

내·외에서 다양하게 이루어지고 있다. 여기서 흥미로운 사실은 많은 연구자들이 공통적으로 “모델에서 제시하고 있는 개체에 대한 정의가 상당히 모호하여 이를 실제 상황에 적용하기가 다소 까다로우며,” 특히 “기존의 레코드가 가지고 있는 ‘내용적 불완전성’으로 인해 현행 레코드를 토대로 FRBR 모델을 완벽히 구현하는데 어려움이 있다”고 지적하고 있다는 점이었다. 이러한 이유로 FRBR 모델을 실제 상황에 적용하는 과정에서 저작이나 표현형에 대해 자의적인 해석이 이루어지고, 특히 모델의 구현이 까다로운 표현형은 아예 생략되거나 기계적 식별이 가능한 일부 요소에 한정하여 적용하고 있다는 한계를 보이고 있다. 이에 다음 장에서는 FRBR 모델에서 제시하고 있는 각 개체에 대한 정의를 고찰하고, 기존의 연구들이 현행 레코드(MARC 형식)에서 이들 개체를 식별하기 위해 근거로 삼은 데이터 요소가 무엇인지에 대한 분석을 시도하고자 한다. 분석결과는 이후에 KORMARC 레코드에 FRBR 모델을 적용하기 위한 기초 데이터로 활용될 것이다.

Ⅲ. FRBR 개체와 KORMARC 데이터요소

1. 제1집단 개체와 MARC21의 식별 요소

FRBR 모델은 서지레코드를 이용할 때 이용자가 수행하는 여러 가지 과업과 이 과업을 성공적으로 수행하는데 필요한 서지데이터를 연결하기 위한 목적에서 개발한 개념적 모델이다. 이 모델을 개발하는 과정에서 FRBR 연구진은 이용자들이 특히 관심을 가지는 개체를 세 집단으로 구분하였다: “지적, 예술적 노력의 산물”과 “이러한 산물의 지적, 예술적 내용에 책임을 지거나 제작, 배포, 관리의 책임을 지닌 실체”, “지적, 예술적 산물이 내포하고 있는 주제”가 그것이다. 그리고 이들 세 집단을 형성하는 주요 개체를 체계적으로 분석하여 저작, 표현형, 구현형, 개별자료(이하 제1집단), 개인, 단체(이하 제2집단), 그리고 개념, 대상, 사건, 장소(이하 제3집단)로 정의하였다. 이 가운데 FRBR 모델의 핵심 개념이며, 서지레코드를 대상으로 모델의 구현을 시도하고 있는 서지기관이나 연구자들이 특히 관심을 가지고 있는 것은 제1집단이다.

제1집단의 첫 번째 개체인 저작은 ‘지적으로나 예술적으로 독창성을 지닌 창작물’을 의미한다. 이는 특정한 하나의 유형을 지칭하는 것이 아니라 우리가 추상적으로 인식하고 있는 것, 가령 『햄릿』이라고 했을 때 우리의 머릿속에 떠오르는 추상적 개념이라 할 수 있다. FRBR 모델에서는 똑 같은 『햄릿』이라 하더라도 어떠한 경우는 ‘동일 저작’으로, 또 다른 경우는 ‘상이한 저작’으로 취급하고 있는데, 개정이나 갱신, 축약이나 증보, 언어를 달리한 번역서나 개작, 편곡 등과 같이 내용에 근본적인 변경이 이루어지지 않은 경우는 동일 저작으로, 이와 달리 의역서나 개작서, 변주곡, 초록, 요약문 등과 특정 문학형식이나 예술양식이 다른 양식으로 변경된 경우와 같이 상당한 정도의 내

용적·형식적 변경이 이루어진 경우는 새로운 저작으로 취급하고 있다. 이러한 개념을 실제 레코드에 반영하기 위한 방안으로, 현재 많은 연구자들은 ‘기본표목’을 활용할 것을 권고하고 있다. 도서관의 오랜 관행에 의해 단일 실체에 해당하는 저작이 이미 기본표목에 반영되어 있다는 이유에서이다. 이에 따라 LC와 OCLC 등에서는 저작을 식별하기 위한 근거로서 다음 <표 1>과 같이 기본표목을 활용하는 방법을 기본적으로 채택하고 있다.¹²⁾

<표 1> 저작 식별을 위한 MARC21의 데이터요소

유형	근거	사례
기본표목	100/110/111 ▼a+▼t	100 1\$ ▼aShakespeare, William, ▼d1564-1616. ▼tHamlet.
	130 ▼a+▼h	130 0\$ ▼aGone with the wind(Motion picture). ▼hSound recording.
기본표목 + 통일표제	100/110/111 ▼a +240 ▼a	100 1\$ ▼aShakespeare, William, ▼d1564-1616. 240 10 ▼aAs you like it. ▼lGerman 245 10 ▼aWie es euch gefeällt / ▼dWilliam Shakerspeare
기본표목 + 종합통일표제	100/110/111 ▼a +242/243 ▼a	100 1\$ ▼aShakespeare, William, ▼d1564-1616. 243 10 ▼aWorks. ▼f1983 245 14 ▼aThe Globe illustrated Shakespeare
기본표목 + 표제	100/110/111 ▼a +245/256/247 ▼a	100 1\$ ▼aShakespeare, William, ▼d1564-1616 245 10 ▼aMacbeth / ▼dWilliam Shakerspeare

제1집단의 두 번째 개체인 표현형은 ‘저작을 실현한 것’으로서 저작이 실현될 때마다 그 저작이 가지는 특정한 형식을 반영한 개념이다. 이는 동일 저작에 대한 내용적·형식적 차이를 제시함으로써 이용자들이 자신이 원하는 자료를 선별하도록 하는데 목적이 있다. FRBR 모델에서는 “아무리 사소한 것이라 하더라도” 지적, 예술적 내용에 변화가 발생한 경우 이를 새로운 표현형으로 취급하고 있는데, 앞서 동일 저작으로 취급된 개정이나 갱신, 증보, 번역서, 편곡 등이 동일 저작에 대한 상이한 표현형에 해당하는 사례라 할 수 있다. 그러나 ‘차이’를 인정할 수 있는 범위에 대해 FRBR 모델에서는 “해당 저작의 성격이나 예상되는 이용자의 요구에 따라 달리 적용할 수 있다”고 해석의 여지를 남겨두고 있어, 이를 서지 DB에 실제 적용하는데 상당한 어려움을 안고 있다. 이러한 이유로, 표현형을 식별하기 위한 근거는 ‘명시적으로’ 식별이 가능한 일부 요소에 한정하여 적용하고 있는데, 가령 OCLC에서는 ‘언어’로만, 그리고 LC에서는 ‘언어’, ‘자료유형’, ‘악보형식’,

12) ‘저작’을 식별하기 위한 방안으로 <표 1>과 같은 기준에 따라 ‘저작’을 식별될 수 있는 데이터 요소(진하게 표시된 부분이 저작을 식별하는데 사용되는 기준 데이터임)를 추출하고, 추출한 모든 데이터 요소를 상호 매핑하여 동일한 것으로 판명된 경우 동일 저작으로 분류하는 방식을 일반적으로 채택하고 있다. 다만, 세부적인 적용과 관련해서 LC와 OCLC의 방식이 다소 상이한데, LC에서는 기본표목에 전거레코드에 의해 통제된 데이터가 수록되어 있으므로 ‘서지레코드’를 활용하여 저작을 식별하는 반면, OCLC에서는 저작으로 간주될 수 있는 데이터 요소를 ‘전거레코드’로부터 추출한 다음, 서지레코드로부터 추출한 데이터가 이들 중 하나에 해당할 때 그 저작으로 간주하는 방식을 취하고 있다.

‘지도의 수치데이터’로 표현형을 식별하는 방안을 제안하고 있다.¹³⁾

이렇듯 표현형은 동일 저작의 실현물 간에 내재되어 있는 내용상의 차이를 반영하는 수단으로서 FRBR 모델의 실행에서 특히 중요하게 다루고 있는 부분이라 할 수 있다. 그러나 저작이 ‘기본표목’이라는 장치를 통해 현행 레코드에 일정 정도 반영되어 있다면, 표현형은 지금까지 논의되지 않았던 다소 생소한 개념이면서 동시에 현행 레코드에서 이를 명시적으로 식별할 수 있는 근거가 부족하다는 한계를 가지고 있다. 따라서 표현형 개체에 대한 보다 명확한 개념 정의가 이루어져야 하며, 표현형을 기계적으로 식별하기 위한 장치의 마련 또한 현행 목록규칙과 MARC 형식에 이 개념을 반영한 이후로 유보해야 한다는 의견이 관련 연구자들 사이에서 강하게 대두하고 있다. 그럼에도 불구하고, 비록 한시적이긴 하지만 현 단계에서 표현형을 식별하기 위한 근거로서 다음 <표 2>와 같은 데이터 요소를 제안하고 있다.

<표 2> 표현형 식별을 위한 MARC21의 데이터 요소

유형	근거	사 례
언어	130/240/243 ▼1 008/35-37	130 0♯ ▼aBible. ▼pN.T. ▼pPhilippians. ▼lEnglish.
		100 0♯ ▼aHomer. 240 10 ▼alliad. ▼lGerman
		008/35-37 fre
자료유형	Leader/06	06/c(Notated music)
기타 표현양식	245 ▼d + 700 ▼a	245 10 ▼aThe expedition of Humphry Clinker. With a memoir of the author / ▼cded by Thomas Roscoe and illus. by George Cruikshank
	254 ▼a	254 ♯♯ ▼aComposer's score
	255 ▼a	255 ♯♯ ▼aScale [ca. 1 : 90,000]

제1집단의 첫 번째, 두 번째 개체인 저작과 표현형이 지적, 예술적 노력의 산물에 대한 ‘내용’(content)을 반영한 것이라면, 세 번째, 네 번째 개체인 구현형과 개별자료는 ‘물리적 실체’(carrier)를 반영한 개념이다. FRBR 모델에서는 구현형을 ‘한 저작의 표현형을 물리적으로 구현한 것’으로, 그리고 개별자료를 ‘구현형에 대한 하나의 사례’로 정의하고 있는데, 가령 동일 저작에 속하는 자료가 내용적·형식적인 면에서 동일하지만 제작과정에서 물리적인 형태가 변경되었다면 이를 동일 표현형에 대한 상이한 구현형으로,¹⁴⁾ 그리고 동일 구현형에 속하는 독립된 각 권, 즉 물리적 대상

13) Tom Delsey, *Functional Analysis of the MARC21 Bibliographic and Holdings Format*, Library of Congress, Network Development and MARC Standards Office, 2002.

<<http://www.loc.gov/marc/marc-functional-analysis/home.html>> [cited 2008. 3. 18].

14) 이에 해당하는 것으로는 저작에 대한 내용적·형식적 없이 단순히 종이를 마이크로필름화한 경우나 서체나 폰트 크기의 변경이 이루어진 경우 등이 있다. 이외에 제작과정에서 발행자나 제작자, 배포자 등이 변경된 경우도 새로운 구현형으로 취급된다.

은 동일 구현형에 대한 개별자료로 취급하고 있다. 이러한 구현형과 개별자료는 저작과 표현형을 통해 이용자들이 원하는 자료를 선별한 다음, 그 중에서 이용하기를 원하는 '바로 그 개체'에 접근하거나 입수하고자 할 때 필요한 정보를 제공해 주는데 주된 목적이 있다. 지금까지의 편목 관행이 물리적 실체를 기준으로 그 내용을 기술해 왔다는 점에서 볼 때, 구현형과 개별자료는 현행 레코드에 상당 부분 반영되어 있다고 볼 수 있다. 구현형을 식별하기 위한 요소로는 해당 자료를 식별할 수 있는 모든 요소, 즉 각종 식별기호, 표제, 저자, 판표시, 발행자, 특정자료 종별과 수량, 기타 물리적 특성에 관한 데이터 등을, 그리고 개별자료를 식별하기 위한 요소로는 소장사항과 대출정보 등을 활용하고 있다.

2. KORMARC의 개체 식별 요소

앞서 살펴보았듯이, MARC21에서 FRBR 모델의 최상위 개체인 저작을 식별하기 위해 1차적으로 사용하는 데이터는 기본표목과 통일표제 등이다. 그러나 국내에서 채택하고 있는 '서명주기입방식'은 기본표목을 사용하지 않는 것을 원칙으로 하고 있다. 또한, 무저자명 저작이나 특정 저작의 여러 판이나 번역서에서 상이한 표제를 사용하고 있을 경우 이 모든 저작을 한자리에 모으기 위해 일반적으로 '통일표제'를 적용하지만, 한국목록규칙 제4판에서는 이 개념을 배제하고 있다. 이외에도, 동일 저작을 판별하는데 중요한 근거로 활용되는 '주제명'이나 다수의 저작을 묶어 간행한 합집이나 종합표제를 가진 저작 등에 수록된 개별 저작을 식별하기 위한 '분출표목'이 우리의 목록에서는 일관성 있게 적용되지 않고 있다. 이처럼 우리 목록에서 채택하고 있는 원칙이나 편목관행을 고려해 볼 때, MARC21의 환경을 고려하여 개발한 이들 알고리즘이 우리의 목록레코드에서는 제대로 실행되지 않을 가능성이 상당히 높다고 할 수 있다. 이러한 상황에서, KORMARC을 기반으로 하는 우리의 목록 환경에 적용할 FRBR 알고리즘을 제안한 연구가 최근 발표되어 주목을 받고 있다.

가령, FRBR 모델에 관한 연구를 지속적으로 수행해 온 조재인은 국립중앙도서관을 비롯하여 몇 개의 대학도서관만이 전거레코드를 활용하고 있는 현실을 감안할 때 전거통제된 데이터를 전제로 FRBR의 각 개체를 식별하는 알고리즘의 적용은 불가능하다고 지적하면서, 이에 대한 대안으로 알고리즘의 적용과 동시에 육안 식별 과정을 병행하는 "semi-automated" 방법을 제안하고 있다. 15) 그녀가 '저작'을 식별하기 위해 제안한 방식은 (1) 기본표목으로서의 통일표제(130 필드)가 있을 경우 이를 저작 식별 요소로 활용하고, (2) 통일표제가 없을 경우 서지레코드에 제시되어 있는

15) Jane Cho, "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 Service* Vol.30(2006), pp.202-213.

표제(240 / 245 / 505 / 730 / 740 필드의 첫 번째 표제)와 저자명(100 / 110 / 111 / 700 / 710 / 711 필드)에서 '저자 + 표제' 데이터를 추출한 후 전체 레코드와의 1:1 비교를 통해 동일한 것으로 간주될 경우 동일 저작으로 군집하되, 마지막에는 반드시 전문가가 그 적합 여부를 검증하여 적합하지 않은 레코드를 제거하도록 하는 것이었다. 그녀가 제안한 방식은 전거레코드를 사용하지 않고 있는 우리의 목록 환경에서 충분히 고려해 볼 수 있는 대안이나, 전혀 통제되지 않은 데이터를 대상으로 한 군집화가 얼마나 가능한지 그리고 대규모 서지DB를 대상으로 전문가의 육안식별이 얼마나 실효성을 거둘 수 있을지에 대해서는 향후 보다 심층적인 논의가 필요하다고 할 수 있다.

국립중앙도서관에서도 또한 FRBR 모델의 실행을 위해 Pilot Project의 성격으로 연구를 진행한 바 있다.¹⁶⁾ 이 연구에서는 OCLC의 FRBR 변환 알고리즘을 활용하여, (1) 저자 표목(1XX ▼a, 7XX 10 ▼a)과 표제(245 ▼a, 74X ▼a)를 추출하여 상호 비교를 통해 동일한 것으로 판명될 경우 동일 저작으로 확정하고, (2) 동일 저작으로 판명된 레코드에 대해 언어 정보(008/35-37)의 異同에 따라 표현형으로 구분한 뒤, (3) 이들을 다시 출판사항(260 ▼b), 발행년(008/07-10, 260 ▼c), 매체유형(leader/06-07, 245 ▼h)의 異同에 따라 구현형으로, 그리고 (4) 동일 구현형으로 판명된 레코드 중에서 복분표시(049 ▼c)가 있을 경우 이를 개별자료로 간주하는 'KORMARC 변환 알고리즘'을 제안하였다. 그러나 이 알고리즘에서 저작을 식별하기 위한 1차적 근거로 활용한 저자 표목이나 표제 등이 전혀 통일된 형식으로 기술되어 있지 않은 우리의 현실을 감안한다면, 이 알고리즘 또한 관련 저작들을 실제 얼마나 집중할 수 있을 것인지를 예측하기란 어렵다고 할 수 있다. 이러한 우리 목록 DB의 한계와 문제점을 고려하면서, 다음 장에서는 KORMARC 레코드의 실질적인 분석을 통해 FRBR 모델의 실행을 어렵게 하거나 불가능하게 하는 주된 요인이 무엇인지를 구체적인 사례와 함께 살펴보았다.

IV. 표본 레코드에 대한 FRBR 모델의 적용

1. 표본 레코드에서 드러난 특징

이 장에서는 앞서 살펴본 MARC21의 FRBR 자동 변환 알고리즘과 KORMARC에 적용하기 위해 제안된 '수정' 알고리즘을 참조로 하여, KORMARC 레코드에 이러한 알고리즘의 적용이 실질적으로 가능한지를 직접 살펴보고자 하였다. 이를 위해, 우선 국립중앙도서관에서 표본 레코드를 추출하였는데, 표본 레코드는 표제 키워드를 사용하여 검색한 결과 중에서 KORMARC 형식에 따라 작성된 레코드로 선정하였다. 검색에 사용된 키워드는 '햄릿', '햄리트', 'Hamlet' 세 가지였으

16) 국립중앙도서관, 『NDL 운영 시범사업 최종보고회』, 국립중앙도서관(2007. 12. 20), pp.158-172.

며, 이들 각각의 검색 결과 중에서 상호 중복되는 것과 학위논문¹⁷⁾을 제외한 161건의 레코드를 최종 표본 레코드로 삼았다.¹⁸⁾ 표본 레코드에 대한 분석은 검색결과 화면이 아닌 KORMARC 레코드를 대상으로 이루어졌으며, 이들 표본 레코드를 분석한 결과 국립중앙도서관에서 생산한 KORMARC 레코드는 다음과 같은 특징을 가지고 있음이 밝혀졌다.

첫째, 기본표목의 적용에 관해 통합서지용 KORMARC에서는 “저자주기입 방식의 목록규칙(KCR2, AACR2)에서는 적용하고, 서명주기입 방식의 목록규칙(ISBD, KCR3, KCR4, KORMARC 기술규칙)에서는 원칙적으로 적용하지 않는다”고 명시하고 있다.¹⁹⁾ 그러나 실제 표본 레코드를 분석한 결과, 원칙적으로 기본표목을 적용하지 않아야 함에도 불구하고 ‘대부분’의 레코드는 기본표목(100, 110, 111, 130필드)을 적용하고 있는 것으로 나타났다. 다만, 개인저자의 저작에 해당하는 사례인데도 기본표목을 적용하지 않고 있는 레코드가 일부(29건) 발견되었는데, 이는 모두 ‘레코드의 최종처리일시’가 2006년 이후라는 공통점을 가지고 있었다. 이로써 2005년 이전에 생산한 레코드는 대부분 기본표목을 적용하고 있지만, 2006년 이후에 신규 입력하거나 수정한 레코드는 기본표목 대신 부출표목(700, 710, 711, 730필드)을 적용하고 있음을 알 수 있다. 뿐만 아니라 기본표목을 적용하고 있는 레코드 중에서 4건의 레코드에서 표목의 ‘오류’가 발견되어, 이러한 기본표목의 ‘非일관성’과 ‘부정확성’이 향후 FRBR 모델의 적용에 영향을 미칠 수 있음을 예상할 수 있다.

17) 학위논문을 분석대상에서 제외한 주된 이유는 다음과 같다. 첫째, 학위논문의 경우 대부분 저작 햄릿에 대한 해설이나 분석, 비평 등에 관한 연구물이라 할 수 있는데, 이들은 저작 햄릿에 상당한 정도의 내용적 변경이 일어났기 때문에 FRBR 모델에서 ‘상이한 저작’으로 취급하고 있다. 따라서 ‘동일 저작’으로의 군집화를 시도해 보고자 하는 이 연구의 의도에 다소 벗어난다고 할 수 있다. 둘째, FRBR 모델에서는 상이한 개체로 취급되지만 개체 간에 ‘서지적 관계’를 가지고 있을 때 이들을 특정 관계를 통해 상호 연결하도록 하고 있다. 가령, 이들 학위논문은 ‘햄릿’이라는 공통된 주제를 가지고 있으며, 이들은 이러한 주제관계에 따라 저작 ‘햄릿’과 연결된다. 그러나 국립중앙도서관에서 추출한 학위논문 레코드는 이러한 주제 데이터요소를 거의 가지고 있지 않아 관련 저작과 기계적으로 연결할 수 있는 장치가 없으며, 따라서 개별 학위논문은 모두 ‘단일저작’으로만 취급될 가능성이 농후하다. 이러한 이유로 이 연구에서는 학위논문을 제외하기로 하였다.

18) 구체적인 표본 레코드의 추출 과정은 다음과 같다:

검색 키워드	검색 결과		제외한 레코드	추출된 표본 레코드
햄릿	도서	① 113건	3건(레코드 오류)	110건
	학위논문	65건	65건	0건
	기타	② 14건	0건	14건
Hamlet	도서	③ 82건	50건(①과 중복)	32건
	학위논문	82건	82건	0건
	기타	5건	4건(②와 중복)	1건
햄릿트	도서	9건	5건(① 또는 ③과 중복)	4건
	학위논문	0건	0건	0건
	기타	0건	0건	0건

19) 국립중앙도서관, 『한국문헌자동화목록형식: 통합서지용』, 국립중앙도서관, 2006.

둘째, 국립중앙도서관에서는 내부업무지침에 따라 지난 2000년부터 인명과 단체명에 대한 전거데이터를 적용하고 있다.²⁰⁾ 이러한 전거데이터의 사용은 편목자들에게 여러 형태의 표목 중 어떤 것을 표목으로 선정할 것인가에 대한 고민을 줄여주고, 이용자들에게 자신이 알고 있는 이름뿐만 아니라 다양한 형태의 이름을 하나의 통제된 이름으로 신속하고도 효과적으로 검색하도록 도움을 준다. 그러나 161건의 표본 레코드 중에서 기본표목(100 필드)이 “셰익스피어(Shakespear, William)”로 선정되어 있는 100건의 레코드를 추출하여 분석한 결과, 다음 <표 3>과 같이 무려 11종에 이르는 상이한 표기가 발견되었으며, 이들을 기계적으로 연결하는 장치 또한 마련되어 있지 않다는 점에서 실질적으로 국립중앙도서관에서 생산한 레코드에는 전거데이터가 적용되지 않고 있음을 알 수 있다. 따라서 FRBR 모델의 적용시 기본표목에 ‘동일 저자에 대한 상이한 형태의 표기’를 가진 다수의 레코드가 있다는 사실도 주요 변수로 고려해야 할 것이다.

<표 3> 기본표목(100필드)의 저자명 표기 유형

표기 유형	건 수	표기 유형	건 수
셰익스피어	9	셰익스피어	9
셰익스피어, W.	2	셰익스피어, W.	3
셰익스피어, 윌리엄	25	셰익스피어, 윌리엄	18
셰익스피어, 윌리엄	4	Shakespeare, William	14
셰익스피어, 윌리엄	1	Shakespeare, William, 1564-1616	14
셰익스피어, 윌리엄	1	총 11종	100

* 기본표목 필드(100)를 기준으로 하되, 기본표목을 적용하고 있지 않은 경우에는 부출표목(700) 중 첫 번째 데이터를 기준으로 산정한 것임.

셋째, 위의 사례에서처럼 전거데이터를 적용하지 않는 상황에서 다양한 형태의 이름으로 출판된 여러 저작을 한 번에 검색하도록 하기 위해 국립중앙도서관에서는 異形이나 異名에 해당하는 데이터를 해당 레코드에 함께 수록하고 있는 경우가 적지 않은 것으로 나타났다. 가령, “셰익스피어, 윌리엄”을 기본표목에 기술하였다면, “Shakespeare, William” 또는 “셰익스피어, W” 등을 부출표목이나 로컬표목에 기술하고 있는 것이다. 이처럼 기본표목에 적용된 표기 방식과 상이한 형태의 데이터를 다른 필드에 기술하고 있는 사례가 위 <표 3>에 기술된 100건의 레코드에서 모두 109건 발견되었다.²¹⁾ 그러나 이들 데이터는 다음 <표 4>와 같이 부출표목(700필드)이나 로컬표목(900필드), 심지어는 일반주기사항(500필드)에 분산되어 있고, 이들 필드에 기술되어 있는 데이터의 표기방식 또한 상이한 것으로 나타났다.

20) 이치주, 이재선, “KORMARC를 활용한 저자명 전거통제시스템: 국립중앙도서관의 사례를 중심으로,” 도서관, 제56권, 제1호(2001, 봄호), p.137.

21) 기본표목에 대한 이형 표기가 전혀 발견되지 않은 레코드가 있는가 하면, 일부 레코드에서는 2개 이상의 필드에 중복적으로 기술된 경우도 있었다.

〈표 4〉 기본표목에 대한 이형 표기

적용 필드와 표기 방식	건 수	적용 필드와 표기 방식	건 수
700 ▼aShakespeare	4	900 ▼a셰익스피어, 윌리엄	3
700 ▼aShakespeare, W.	1	900 ▼a셰익스피어	4
700 ▼aShakespeare, William	35	900 ▼a셰익스피어, 윌리엄	9
500 ▼aShakespeare	1	900 ▼aShakespeare, W.	1
500 ▼aShakespeare, W.	1	900 ▼aShakespeare, William	1
500 ▼aShakespeare, William	44	960 ▼a셰익스피어, 윌리엄	1
900 ▼a셰익스피어	1	963 ▼aShakespeare, William	1
900 ▼a셰익스피어, 윌리엄	2	총 15종	109건

이외에도, 번역서인 경우 원표제(246필드)를 기재하고 있는 경우와 그렇지 않은 경우, 다수의 저작을 모아 간행한 합집의 경우 내용주기(505필드)나 분출표목(740표목)을 적용하고 있는 경우와 그렇지 않은 경우 등과 같이 동일한 사례에 대한 적용방식이 일관되지 않은 경우가 상당수 발견되었다. 반면, 표본 레코드 중에서 통일표제(130 또는 730필드)를 적용하고 있는 경우는 161건의 표본 레코드 중에서 단 1건에 불과하였으며, 주제명(6XX필드)도 11건의 레코드에서만 발견되어 대다수의 레코드에 주제명 데이터가 수록되어 있지 않은 것으로 나타났다.

2. 표본 레코드를 이용한 FRBR 모델의 적용

표본 레코드에 FRBR 모델을 적용해 보기 위해 이 연구에서는 스프레드시트(sheet)에 표본 레코드에 수록된 모든 데이터를 옮긴 후 LC나 OCLC 등에서 적용하고 있는 기본 방식에 따라 각 개체로의 군집화를 시도하였다. 즉, 현행 레코드를 '구현형'으로 간주하고 이를 대상으로 (1) '저작'으로의 군집화를 시도한 후, (2) 동일 저작으로 군집된 레코드들을 대상으로 '상이한 표현형'을 구분하였으며, (3) 마지막으로 현행 레코드인 '구현형'으로부터 '개별자료'로 분리하였다.

먼저, 저작 개체로의 군집화는 표본 레코드에서 저자명 데이터와 표제 데이터를 모두 추출한 후 이들을 각각 조합하여 만든 '저자 + 표제' 데이터가 상호 일치할 경우 동일 저작으로 간주하는 방식을 취하였다. 이 때, 저자명 데이터는 기본표목(100필드)과 부출표목(700필드), 로컬표목(900필드)으로부터 각각 추출하였는데, 이는 저자명 데이터가 이들 필드에 흩어져 있고, 이들 각각에 수록된 데이터의 표기방식이 모두 상이하다는 앞서의 분석결과에 근거한 것이었다. 다만, 일반주기(500필드의 ▼a)에 저자명이 기술되어 있는 경우도 상당 수(46건의 레코드) 있었으나 이는 저작 개체로의 군집화 과정에서 부득이 제외할 수밖에 없었다. 일반주기사항 필드에 수록된 다양한 데이터를 인위적인 방식에 따라 저자명과 그 외의 데이터로 분리할 수 없었기 때문이었다. 다음으로, 표제 데이터는 통일표제(240필드의 ▼a)와 본표제(245필드의 ▼a + ▼h)는 물론이고 분출표제(740필드의 ▼a)로

부터 추출하였다. 결과적으로, 100 + 240, 100 + 245, 100 + 740, 700 + 240, 700 + 245, 700 + 740, 900 + 245, 900 + 245, 900 + 740 필드로부터 '저자 + 표제'의 조합이 만들어졌으며, 이들에 대한 상호 매핑을 통해 동일 여부를 가늠하였다.

이러한 과정을 통해 161건의 표본 레코드로부터 만들어진 저작의 수는 총 89종으로, 하나의 저작에 평균 1.8개의 레코드가 포함되는 것으로 나타났다. 또한, 89종으로 분류된 저작 가운데 단지 1개의 레코드만을 가진 '단일저작'(55종 / 61.8%)을 제외하면, 34종(38.2%)의 저작이 2개 이상의 레코드로 구성된 '복합저작'인 것으로 나타났다. 그러나 복합저작으로 간주된 34종의 저작을 최종 검증한 결과, 다음 <표 5>와 같이 실질적으로는 동일 저작으로 취급되어야 함에도 불구하고 단순히 표기 방식의 차이로 인해 상이한 저작으로 간주된 것임을 확인할 수 있었다. 이에 이를 일일이 확인한 후 동일 저작으로 간주될 수 있는 레코드들을 묶어 최종적으로 161건의 표본 레코드로부터 62종의 저작(1저작당 2.6개의 레코드)을 도출할 수 있었다. 이 중 단일저작에 해당하는 55종을 제외하면 106건의 레코드가 불과 7종의 저작(1저작당 15.1개의 레코드)으로 군집되는 결과를 보였다. 반면, 단일저작으로 취급된 자료들 가운데 표목의 선정이나 표목부의 표기 오류로 인해 동일 저작으로 군집되어야 할 대상임에도 불구하고 단일저작으로 취급된 경우가 4건 발견되었다.

<표 5> '상이한 저작'으로 취급된 복수저작의 유형

구 분	저작의 유형	해당 레코드 수
유형 1	(1) 햄릿 / 셰익스피어.	9건
	(2) 햄릿 / 셰익스피어, 윌리엄.	22건
	(3) 햄릿 / 셰익스피어, 윌리엄.	10건
	(4) 햄릿 / 셰익스피어, W.	2건
	(5) 햄릿 / 셰익스피어.	11건
	(6) 햄릿 / 셰익스피어, 윌리엄	3건
	(7) 햄릿 / 셰익스피어, 윌리엄.	20건
	(8) 햄릿 / 셰익스피어, W.	4건
	(9) 햄릿 / Shakespeare.	4건
	(10) 햄릿 / Shakespeare, William.	45건
	(11) 햄릿 / Shakespeare, William, 1564-1616.	12건
	(12) 햄릿 / 셰익스피어, 윌리엄.	2건
	(13) 햄릿 / Shakespeare, William.	2건
	(14) Hamlet / Shakespeare.	2건
	(15) Hamlet / Shakespeare, William.	7건
	(16) Hamlet / Shakespeare, William, 1564-1616.	2건
유형 2	(17) 햄릿 / 램, 차일스.	2건
	(18) 햄릿 / 램, 찰스.	2건
유형 3	(19) 셰익스피어 4대 비극 / 셰익스피어, 윌리엄.	2건
	(20) 셰익스피어 4대 비극 / Shakespeare, William.	2건
	(21) 셰익스피어 4대 비극 / Shakespeare, William, 1564-1616.	5건
유형 4	(22) 로미오와 줄리엣 / 셰익스피어.	2건
	(23) 로미오와 줄리엣 / 셰익스피어, 윌리엄.	6건
	(24) 로미오와 줄리엣 / 셰익스피어.	2건

유형 4	(25) 로미오와 줄리엣 / 셰익스피어, 윌리엄. (26) 로미오와 줄리엣 / Shakespeare, William.	2건 6건
유형 5	(27) 리어왕 / 셰익스피어. (28) 리어왕 / 셰익스피어, 윌리엄. (29) 리어왕 / 셰익스피어, 윌리엄. (30) 리어왕 / Shakespeare, William. (31) 리어왕 / Shakespeare, William, 1564-1616.	3건 2건 2건 2건 2건
유형 6	(32) 맥베스 / 셰익스피어, 윌리엄. (33) 맥베스 / Shakespeare, William.	2건 2건
유형 7	(34) Shakespearean tragedy : lectures on Hamlet, Othello, King Lear, Macbeth / Bradley, A.C.	2건

다음으로, 동일 저작으로 취급된 레코드를 대상으로 이들을 다시 '표현형'으로 구분하는 작업을 수행하였다. 이를 위해, 위의 <표 5>에 제시된 저작의 유형 중에서 [유형 1]에 해당하는 레코드 100건을 별도로 추출하여²²⁾ 이들을 다시 표현형으로 구분하는 작업을 수행하였다. 표현형에 대한 구분은 LC와 OCLC에서 제시한 기준, 즉 '자료유형'과 '언어' 데이터를 기본적으로 활용하였고, 여기에 KORMARC 레코드에 기술되어 있는 데이터 중에서 표현형을 식별하는데 활용할 수 있을 것이라 생각한 '이용대상자 수준'을 덧붙였다. 표현형이 동일 저작에 대한 내용적·형식적 차이를 나타내는 개념이라 할 때, 취학전 아동, 초등학생, 중학생, 고등학생, 성인 등으로 구분하고 있는 '이용대상자 수준'에 관한 정보는 자료의 내용적 특성이나 수준을 충분히 반영한다는 판단에서였다. 표현형을 식별하는데 활용할 자료유형과 언어, 이용대상자 수준에 대한 데이터를 도출한 근거는 다음과 같다.

첫째, '자료유형'에 대한 데이터는 KORMARC의 leader/06에 제시되어 있는 '레코드의 유형'과 각 자료의 유형별로 규정하고 있는 구체적인 '자료의 유형'(008/23 등)에서 각각 추출하였다. 이는 "문자자료(점자자료)", "전자자료(수치정보)"와 같이 자료의 일반적인 유형과 그 유형에 해당하는 구체적인 형태를 동시에 제시하기 위함이었다. 둘째, '언어'에 대한 표현형 식별은 130, 240필드의 ▼1(언어)과, 008/35-37의 3자리 부호, 041필드의 ▼a(본문언어) ▼h(원저작의 언어) 등을 활용할 수 있는데, 앞서의 분석결과에서 130, 240필드의 ▼1(언어)을 적용하고 있는 경우는 단 한건도 발견할 수 없었던 점을 참작하여 여기서는 008/35-37과 041필드에서만 관련 데이터를 수집하였다. 마지막으로 이용대상자 수준에 관한 정보는 008/22에서 추출하였다.

결과는 다음 <표 6>과 같이, '자료유형'과 '이용대상자 수준'의 경우 저작에 대한 표현형이 모두 동일한 것('문자자료'와 '일반이용자용')으로 나타났다.²³⁾ 이에 반해, '언어'에 따라 구분한 결과는

22) 저작은 레코드에 수록되어 있는 저자 데이터(100, 700, 900필드)와 표제 데이터(240, 245, 740필드)의 조합을 통해 도출한 결과를 근거로 하였기 때문에 하나의 레코드로부터 복수의 '저자 + 표제' 데이터가 생성되었다. [유형 1]에 해당하는 레코드는 중복으로 계상된 레코드를 제외하면 총 100건이다.

23) 이러한 결과가 도출된 데에는 두 가지 원인이 있었을 것으로 유추되는데, 첫째는 앞서 [유형 1]에 해당하는 저작

008/35-37과 041필드의 데이터를 적용하였을 때가 상이한 것으로 나타났는데, 그 이유는 다수언어로 된 자료나 번역물인 경우 041필드를 통해 완전한 정보를 제공하여야 함에도 불구하고 실질적으로 이를 기술하지 않은 경우가 많았기 때문이었다.²⁴⁾ 따라서 언어에 대한 식별은 현행 레코드에서 008/35-37에 수록된 데이터에 근거할 수밖에 없었다. 이러한 분석결과에 근거하여, 앞서 [유형 1]의 저작으로 간주된 100건의 레코드는 자료유형과 언어, 이용대상자 수준으로 표현형을 식별한 결과, '문자자료-한국어-일반이용자용'(86건), '문자자료-영어-일반이용자용'(12건), '문자자료-일본어-일반이용자용'(1건), '문자자료-스페인어-일반이용자용'(1건)과 같이 4종의 상이한 표현형으로 구분되었다.

〈표 6〉 동일 저작에 대한 표현형 식별 기준과 결과

구 분	데이터 식별 기준	결 과
(1) 자료유형	Leader/06(레코드의 유형)	문자자료 100건
	도서-008/23(자료의 형태) 전자자료-008/25(전자자료의 유형) 지도자료-008/25(지도자료의 유형) 음악/녹음자료-008/23(자료의 형태) 계속자료-008/23(자료의 형태) 시청각자료-008/29(자료의 형태) 고서-008/23(자료의 형태) 복합자료-008/23(자료의 형태)	해당사항없음 100건
(2) 언어	008/35-37	한국어 86건 영어 12건 일본어 1건 스페인어 1건
	041 ▼a(본문언어) ▼b(요약문언어) ... ▼h(원저작의 언어)	▼a한국어 ▼h영어 46건 ▼a영어 ▼a한국어 7건
(3) 이용대상자 수준	008/22	일반이용자용 100건

한편, 구현형은 현행 레코드에 수록된 내용을 동일 표현형 아래에 나열하는 수준의 작업을, 그리고 개별자료는 각각의 구현형에 대한 복본표시와 소장정보를 제시하는 수준의 작업을 수행하였다. 그러나 다음 〈표 7〉에 제시된 [사례 1]과 같이 두 저작이 실질적으로 '상이한' 구현형에 해당하는 것인지를 KORMARC 레코드만으로는 확인이 불가능한 경우와, [사례 2]와 같이 명백히 개별자

에만 국한하여 적용한 결과로서 이 저작에 해당하는 자료들이 실질적으로 자료유형과 이용대상자 수준이 모두 동일하기 때문일 수 있으며, 둘째는 실제로는 자료유형(008/23)과 이용대상자 수준(008/22)이 상이하지만 이에 대한 데이터 값이 입력되어 있지 않기 때문에 이를 근거로 표현형을 식별하다보니 이와 같은 결과가 도출되었을 수도 있다. 이 연구는 KORMARC 레코드에 수록된 데이터만 기준으로 FRBR 각 개체를 식별하였기 때문에 그 원인을 정확하게 파악할 수 없다.

24) 표현형을 식별하기 위해 선정된 100건의 레코드 중에서 번역서임에도 불구하고 041필드에 해당 데이터를 수록하고 있지 않은 사례는 모두 34건이었다.

료로 취급되어야 함에도 불구하고 레코드 구축과정에서의 오류(중복레코드)로 인해 구현형으로 취급되는 경우가 발견되었다. 이러한 경우, 전자에 대해서는 실물자료와의 대조를 통해 실제 상이한 구현형인지를 일일이 확인해야 하고, 후자에 대해서는 하나의 구현형에 대한 복수의 개별자료로 수정 처리한 후 모델을 실행해야 할 것으로 보인다.

〈표 7〉 동일 표현형에 대한 상이한 구현형의 사례

구분	상이한 구현형으로 간주된 사례	비고
[사례 1]	햄릿, W. 셰익스피어 지음; 권응호 옮김. 해원출판사, 1991. 햄릿, W. 셰익스피어 지음; 권응호 옮김. 해원출판사, 1992.	실제 상이한 구현형인지 실물 자료에 대한 확인 작업 필요
[사례 2]	셰익스피어 4대비극. 지은이: 셰익스피어; 옮긴이: 셰익스피어연구회. 개정판. 아름다운날, 2006. 셰익스피어 4대 비극. 윌리엄 셰익스피어 지음; 셰익스피어연구회 옮김. 개정판. 아름다운날, 2006.	하나의 구현형에 대한 상이한 개별자료로 취급되어야 함

3. 적용 결과에서 드러난 한계와 문제점

FRBR 모델을 실행하게 되면 평면적으로 열거되어 있는 현행 레코드들이 논리적 구조에 따라 입체적으로 연결되고, 이를 통해 이용자들은 그들이 원하는 자료를 보다 논리적으로 탐색할 수 있게 된다. 그러나 문제는 현행 레코드를 FRBR 모델로 어떻게 그리고 얼마나 효율적으로 재구조화할 것인가 하는데 있다. “햄릿”에 관한 161건의 표본 레코드를 대상으로 KORMARC 레코드를 FRBR 모델의 각 개체에 따라 기계적으로 군집하는 과정을 수행해 본 결과, 현행 우리의 목록레코드를 FRBR 모델로 변환하고자 할 때 봉착할 수 있는 한계와 문제로는 다음과 같은 것이 있음이 밝혀졌다.

첫째, FRBR 모델의 제1집단 개체로의 군집화 과정을 저해하는 가장 큰 요인은 非통제된 데이터 때문인 것으로 나타났다. 앞서 표본 레코드의 사례에서도 드러났듯이, 전거통제를 적용하고 있지 않은 우리의 목록 환경을 감안하여 다양한 필드로부터 ‘저자 + 표제’ 데이터를 추출하였음에도 불구하고 이를 토대로 저작 계층으로의 군집화를 수행한 결과는 일일이 육안 확인한 결과와는 상당한 차이가 있었다. 즉, 데이터 상호 비교만을 통해서만 161건의 레코드로부터 총 89종의 저작이 추출되었으나, 이들을 직접 확인한 결과 실제로는 62종의 저작인 것으로 나타났다. 이러한 차이는 단일저작(55종)을 제외한 복합저작의 경우 더욱 확연하게 드러났는데, 2개 이상의 레코드를 가지는 복합저작이 기계적 식별을 통해서만 34종이었던 것이 데이터를 통제된 후에는 불과 7종으로 줄어들었다. 표본 레코드에 대한 분석결과를 토대로 하여, 앞서 LC나 OCLC에서 제시한 알고리즘과 이를 KORMARC에의 적용을 목적으로 ‘수정’한 알고리즘에서 제시한 FRBR 개체 식별 요소보다 더 다양한 필드로부터 데이터를 추출하였음에도 불구하고 이러한 결과가 도출되었다는 사실

은, 통제되지 않은 데이터를 활용하여 FRBR 모델의 각 개체를 식별하는 데에는 분명 한계가 있다는 사실을 입증해 주고 있다.

둘째, FRBR 모델을 구현하기 위해 KORMARC 레코드로부터 데이터를 추출하는 과정에서 일관성 없이 기술된 데이터는 동일 개체로의 집중을 저해하는 요인으로 작용하였다. 기본표목(1XX 필드)을 적용하고 있는 경우와 기본표목 없이 부출표목(700, 710, 711, 730 필드)만을 적용하고 있는 경우는 순차적 방법을 통해 통제한다 하더라도, 유사한 사례임에도 불구하고 부출표목을 가지고 있는 경우와 그렇지 않은 경우, 원표제를 기술하고 있는 경우와 그렇지 않은 경우, 주제명을 가지고 있는 경우와 그렇지 않은 경우 등과 같이 데이터의 기입 패턴이 일관되지 않은 현재의 상황은 이들 데이터를 활용하여 모델을 구현하고자 할 때 근본적인 문제를 야기할 수 있다. 가령, 표본 레코드 중에서 “셰익스피어 4대 비극 / 셰익스피어 지음 ; 김재남 옮김. 하서출판사, 2000”과 “셰익스피어 4대 비극 / 윌리엄 셰익스피어 지음 ; 정홍택 옮김. 소담출판사, 2002”의 경우, 전자는 내용주기(505 필드)에만 ‘햄릿-오델로-리어왕-맥베스’를 기술하고 있는데 반해 후자는 이를 부출표목(740 필드)에 제시하고 있다. 이러한 상황에서 ‘표제’를 식별하기 위한 기준점으로서 부출표목을 적용하게 되면, 전자는 저작 ‘햄릿’에서 누락될 수밖에 없다. 이처럼 모든 레코드에 일관되지 않은 방식으로 적용된 데이터는 FRBR 모델의 적용을 위한 알고리즘을 개발하거나 이를 구현하는 과정에서 커다란 걸림돌로 작용할 수 있다.

셋째, KORMARC 레코드의 정확성, 특히 데이터의 오류나 누락은 FRBR 모델을 실행하는데 치명적인 결함으로 작용하였다. 표본 레코드에서 발견된 명백한 데이터의 오류는 기본표목의 선정에 오류가 있는 경우(예: 번역자를 기본표목으로 기술한 경우)와 원표제를 대등표제로 기재한 경우 등이 있었는데, 기본표목이나 원표제를 활용하여 각 개체를 확인하고자 할 때 이러한 경우 동일 개체로의 군집화를 불가능하게 만드는 요인이 되었다. 또한, 대표적인 데이터의 누락 사례로는 ‘판사항’(250 필드)과 ‘언어부호’(041 필드) 등이 있었는데, 판표시가 누락된 경우에는 구현형과 개별자료 간의 경계를 구분하는데, 그리고 언어부호의 누락은 다수언어로 된 자료나 번역물을 포함한 자료에 대해 완전한 언어 정보를 제공하지 못하기 때문에 이를 근거로 FRBR 모델의 표현형을 식별하는데 한계로 작용하기도 하였다. 이외에도 자료의 유형이나 이용대상자 수준과 같이 FRBR 개체를 식별하는 단서로 사용될 수 있는 데이터의 값이 정확하게 입력되어 있지 않아 이를 기준으로 FRBR 모델을 실행할 경우 각 개체를 구분해 내는데 한계가 있는 것으로 나타났다. 이처럼 데이터의 오류나 누락, 혹은 불완전한 데이터가 수록되어 있는 경우와 같이 KORMARC 레코드 자체가 안고 있는 원초적인 결함은 FRBR 모델의 실행과정에도 그대로 반영되어 결과적으로 모델의 유용성을 저하시키는 직접적인 원인이 될 수 있음을 이 연구의 결과를 통해 확인할 수 있었다.

V. 결 론

FRBR 모델은 국제적 합의에 의해 완성된 것으로 지금까지 목록이 추구해 온 '탐색'과 '집중'의 기능을 강화하고, 복잡한 서지적 관계를 처리하는데 효과적이라는 긍정적인 측면을 분명히 가지고 있다. 그러나 그렇다고 해서 FRBR 모델이 우리의 도서관 환경에서도 반드시 긍정적인 것이라고 확신할 수는 없다. 다양한 변수를 가진 현실 세계에 이 모델을 실행하거나 구현하고자 할 때에 여러 장애와 한계에 봉착할 가능성이 높으며, 무엇보다도 우리 도서관 목록의 제반 여건이 FRBR 모델을 수용할 만한 수준에 있는가를 확신할 수가 없기 때문이다. 이러한 상황에서, 이 연구는 FRBR 모델의 실행에 있어 선도적인 입장에 있는 주체들이 개발한 FRBR 알고리즘을 참조로 하여 이러한 알고리즘을 우리의 목록레코드에도 대입해 볼 수 있는지에 대한 가능성을 타진해 보고, 만일 이 과정이 불가능하거나 어렵다면 그 원인이 무엇인지를 구체적으로 살펴보고자 하였다.

국립중앙도서관에서 추출한 161건의 표본 레코드를 FRBR 모델에서 제시한 각 개체별로 클러스터한 결과, 현재 우리의 목록레코드는 모델의 '제1집단'만을 구현하는 데에도 상당한 제약이 있는 것으로 드러났다. 이에 대한 주요 원인은 명실상부한 국가 '대표' 도서관에서 생산한 레코드임에도 불구하고, 데이터에 대한 통제가 전혀 되어 있지 않고 데이터의 내용이 부정확하거나 불완전할 뿐만 아니라 데이터의 기입 패턴이 일관되지 못하다는 데 있었다. 그러나 겉으로 드러난 이러한 문제 이면에, 애초에 제안된 이들 알고리즘이 우리와는 전혀 다른 서지적 전통 위에서 개발된 것이라는 점에도 주목할 필요가 있다. 즉, '저작을 중심으로 관련 자료를 집중하는 collocation 기능'에 대한 추구가 기본표목이나 통일표제를 중심으로 오랫동안 논의되어 왔고, 이러한 역사를 가진 시스템의 기반 위에서 개발된 것이 이들 알고리즘이기 때문이다. 그러다보니 우리의 목록 환경에 그들의 전통과 가치에 기반하여 개발한 알고리즘을 그대로 적용하고자 하고, 나아가 FRBR 모델의 외형적 모습을 구현하고자 하는 것 자체가 무리였는지 모른다.

지금 FRBR 모델은 여러 전문가들로부터 기존의 평면적인 구조의 서지레코드가 구현하지 못하는 서지레코드들 사이의 서지적 관계를 논리적으로 연결할 수 있어서, 동일한 내용을 가진 자료가 다양한 형태로 구현되고 있는 오늘날의 정보환경에 적합한 혁신적인 모델이라는 긍정적인 평가를 받고 있다. 우리 학계에서도 FRBR 모델을 현재의 서지레코드 구조를 개선하기 위한 유용한 틀로 인식하면서, 이 모델을 우리의 목록체계에 수용하기 위한 방안을 나름대로 모색하는 단계에 있다고 할 수 있다. 그러나 모델을 수용하기에 앞서, 우리의 경우 실무적인 부분을 포함하는 이와 같은 제반 여건에 대해 다시 한 번 진지하게 성찰해 볼 필요가 있다. 지금과 같은 상태에서는 아무리 정교한 알고리즘을 개발해 낸다 하더라도 FRBR 모델을 구현해 내는데 한계가 있을 수밖에 없고, 모델을 구현해 나가는 과정에서 많은 시행착오를 겪거나 난관에 봉착할 가능성이 농후하기 때문이다. 현 단계에서 우리가 반드시 거치고 넘어가야 할 과정은 새롭고 혁신적인 FRBR 모델을 구현해

내는 과정에 대해서만 관심을 가질 것이 아니라 우리의 목록체계를 지금이라도 철저히 점검하고 난 후에 advanced model로서 이 모델을 우리의 실정에서 어떻게 활용할 것인가에 관한 방안을 강구하는 것이다. 부족하지만 이 연구가 FRBR 모델의 수용과 관련하여 우리 도서관계가 직면할 수 있는 제반 문제를 허심탄회하게 논의하고, 우리의 시각에서 이 모델의 범용성을 높일 수 있는 방안을 모색해 볼 수 있는 계기가 되기를 희망한다.

참 고 문 헌

- 국립중앙도서관. “표준 메타DB 모델 및 시맨틱언어 개발 연구.” NDL 운영 시범사업 최종보고회, 국립중앙도서관(2007. 12. 20).
- 김정현. “한국어 서지레코드에 있어 FRBR 모형의 유용성에 관한 연구.” 한국문헌정보학회지, 제41권, 제4호(2007. 12), pp.295-314.
- 김정현. “서지적 관계를 기반으로 한 한국어 도서의 저작유형 분석.” 한국도서관·정보학회지, 제38권, 제3호(2007. 9), pp.183-200.
- 김현희, 유영준, 박서은. “FRBR 모형의 KORMARC 데이터베이스로의 적용 가능성에 대한 실험적 연구.” 한국도서관·정보학회지, 제38권, 제2호(2007. 6), pp.185-202.
- 노지현. “한국의 도서관 환경에서 FRBR 모델의 의미.” 한국도서관·정보학회지, 제38권, 제2호(2007. 6), pp.223-244.
- 조재인. “표현형 계층을 중심으로 한 FRBR 모형 분석 및 목록 체계 수용에 관한 연구.” 한국도서관·정보학회지, 제36권, 제2호(2005. 6), pp.221-239.
- 조재인. “FRBR 알고리즘 분석 및 KORMARC 데이터베이스 적용 방안.” 한국문헌정보학회지, 제38권, 제3호(2004. 9), pp.1-21.
- Bennett, Rick et. al. “The concept of a work in WorldCat: an application of FRBR.” *Library Collections, Acquisitions, & Technical Services* Vol.27(2003), pp.45-59.
- Bowen, Jennifer. “FRBR: Coming Soon to Your Library?.” *Library Resources & Technical Services*, Vol.49, No.3(2005), pp.175-188.
- Cho, Jane. “A study on an 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*, Vol.30(2006), pp.202-213.
- Hickey, Thomas B. etl. al. “Experiments with the IFLA Functional Requirements for Bibliographic Records(FRBR).” *D-Lib Magazine*, Vol.8, No.9(Sep. 2002).

- IFLA Study Group on the Functional Requirements for Bibliographic Records: Final Report, 1997. <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>> [cited 2008. 3. 18] (김태수 역. 『서지레코드의 기능상의 요건: 최종보고서』. 국립중앙도서관).
- LC Network Development and MARC Standards Office. *Displays for Multiple Versions from MARC21 and FRBR: Functional Analysis of the MARC21 Bibliographic and Holdings Formats*. <<http://www.loc.gov/marc/marc-functional-analysis/multiple-versions.html>> [cited 2008. 3. 18].
- LC Network Development and MARC Standards Office. *FRBR Display Tool Version 2.0*. <<http://www.loc.gov/marc/marc-functional-analysis/tool.html>> [cited 2003. 3. 20].
- Martha M. Yee. "FRBRization : a method for turing Online Public Finding Lists into Online Public Catalogs." *Information Technology and Libraries*(June 2005), pp.77-95.
- O'Neill, Edward. "FRBR : application of the entity-relationship model to Humphry Clinker." *Library Resources & Technical Services*, Vol.46, No.4(Oct. 2002). <http://www.oclc.org/research/publications/archive/2002/oneill_frbr22.pdf> [cited 2008. 4. 10].
- OCLC(2005). FRBR Work-Set Algorithm. <http://www.oclc.org/research/software/frbr/frbr_workset_algorithm.pdf> [cited 2008. 3. 20].