# FRBR 저작 개념 기반의 검색시스템 개발 연구

- 존경각 고서서지시스템을 중심으로 -

# Developing a retrieval system based on FRBR work concept

윤재혁, 성균관대학교, seisiel@naver.com Jaehyuk, Yun, Sungkyunkwan University

오늘날의 정보세계에서 정보 자원은 유형이 빠르게 다양화되고 있을 뿐만 아니라 여러 자원과 관계를 맺고 있다. 이러한 상황에서 자원을 평면적으로 보여주는 기존의 MARC 기반 검색시스템은 이용자의 요구를 충족시키기 어렵다. 이에 따라 본 연구의 목적은 서지 자원의 다양성과 관계성을 나타낼 수 있는 FRBR 모델의 저작 개념을 활용하여 새로운 검색시스템을 개발해 보는데 있다. 연구 대상은 성균관대학교 존경각의 고서서지시스템으로 해당 시스템의 서지레코드 중 청구기호는 저작 개념을 반영하고 있다는 특징을 지니고 있다. 본 연구에서는 이러한 점에 주목하여 청구기호를 통해 저작 식별자를 생성하고 각 서지레코드에 할당함으로써 검색결과가 저작을 중심으로 표현되는 실험 시스템을 개발하였다. 저작 단위로 검색결과를 표시하도록 시스템을 구현한 결과 이용자가 키워드 검색을 한 뒤 처음마주하는 검색결과의 수가 눈에 띄게 감소하였다.

### 1. 서 론

빅데이터로 대표되는 오늘날의 정보세계에서는 수많은 양의 데이터가 생성됨에 따라 정보 자원이 급증하고 있으며 자원들의 유형 또한 다양화되고 있다. 이처럼 날이 갈수록 광활해져가는 정보세계로 인해 이용자들은 정보탐색의 어려움을 겪고 있으며 그에 따라 검색시스템의 역할이 중요해지고 있는 상황이다.

국제도서관협회연맹(IFLA)에서 1998년에 발표 한 Functional Requirements for Bibliographic Records (FRBR) 모델 또한 이와 유사한 맥 락에서 등장하였다. 하나의 서지 자원이 언어 나 매체에 따라 다르게 구현되고 파생되면서 자원의 다양성과 관계성을 나타낼 수 있는 목 록의 필요성이 대두된 것이다.

MARC가 서지 자원들에 대한 내용을 평면적으로 보여주는 반면에 FRBR은 저작(work)을 중심으로 서지 자원을 끌어 모아 계층구조로

표현한다는 특징을 지니고 있다. 이와 같은 특징은 이용자가 자신이 원하는 자료뿐만 아니라 관련된 다른 자료, 예를 들어 다른 언어 또는 판(edition)까지 탐색하는 것을 가능하게 한다.

MARC가 중심인 현재의 검색 시스템은 특정 키워드를 포함하고 있는 모든 서지 레코드를 검색 결과로 표시한다. 그 결과 특정 저자의 동일 저작이 언어나 판 사항에 따라 별개의 결 과로 나타나며 이용자는 이를 일일이 살펴보며 분별해야 한다. 그러나 서지 레코드에 FRBR 모델을 적용할 경우 검색 시스템은 저자와 해 당 저자의 대표 저작만을 결과로 나타내며 이 에 따라 이용자는 표제와 저자명만을 확인하게 되므로 자료 탐색의 시간을 단축시킬 수 있다.

MARC 레코드를 FRBR 기반 구조로 변환하는 핵심은 서지 자원을 저작 단위로 끌어 모으기 위해 기존의 서지 자원에 저작 식별자를 부여하는 것이다. 일반적으로 저작 식별자는 알고리즘을 활용하여 저자명과 표제의 대표형

을 기준으로 할당한다. 그러나 성균관대학교 존경각의 고서목록에는 이미 저작 개념이 청 구기호에 반영되어 있으며 이를 활용하면 효 율적으로 MARC 레코드를 FRBR 구조로 변환 하는 것이 가능하다.

본 연구에서는 이러한 점에 주목하여 존경 각 고서 목록에 FRBR 구조를 적용한 실험 시 스템을 개발하는 것이 목적이며 뒤이어 저작 중심의 검색결과 디스플레이가 갖는 특징까지 도 살펴보고자 한다.

## 2. 연구절차 및 데이터 수집

본 연구의 목적은 FRBR 저작 개념 기반의 실험 서지검색 시스템을 구축하여 저작 중심 디스플레이의 유용성을 살펴보는 것이다. 실험 시스템의 설계를 위해 데이터베이스 구축에 필요한 메타데이터 요소 선정 및 데이터베이스 모델링 작업이 선행되었으며 모델링은 Microsoft(MS)의 Visio를 이용하였다.

FRBR 기반 실험 시스템의 특성을 대표적으로 보여주기 위해서는 하나의 저작이 다양한 표현형, 구현형, 개별자료를 포함하면서 다른 저작과의 관계도 가질 수 있는 데이터가 필요하다. 이에 따라 샘플 데이터는 유가(儒家)의 기본 경전으로써 오랜 기간 연구되어 다양한 주석, 해제, 번역이 존재하는 사서오경(四書五經) 중 사서(四書)에 해당하는 "논어", "맹자", "대학", "중용"을 선정하였다.

데이터는 R 프로그래밍을 활용하여 존경각의 고서서지시스템에서 사서(四書)를 서명으로 검색한 결과를 자동으로 수집하였다. 그 결과 논어 68건, 맹자 47건, 대학 67건, 중용 55건의 데이터를 얻었으며 4건의 중복을 제거하고 남은 233건의 데이터를 활용하여 데이터베이스를 개발하였다. 데이터베이스 개발은 MS의 Access를 통해 이루어졌다.

### 3. 실험 시스템 설계

#### 3.1 현행 시스템 분석

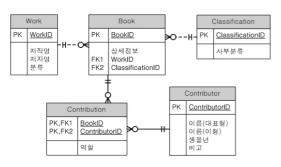
기존 시스템의 초기 화면은 청구기호, 서명, 저자, 판종, 간행년 항목을 통한 상세검색 창 을 제공한다. 검색 결과는 청구기호 순서에 따라 테이블 형태로 목록만 단순히 제공되며 탐 색을 위해 이용자가 클릭할 수 있는 요소는 '서 명'과 '원문보기' 버튼의 두 가지 뿐이다. '서명' 을 클릭하면 상세 서지정보가 나타나며 '원문 보기' 버튼은 원문이 있음을 표시하며 클릭 시 '서명'과 동일하게 상세 서지정보 페이지로 이 어진다.

상세정보 페이지는 단순히 서지정보만을 제 공하며 저자를 기준으로 새롭게 검색하거나 저자 정보를 확인하는 등의 추가적인 탐색 요 소는 제공하고 있지 않다. 이용자가 상호작용 할 수 있는 요소로 '열람 예약', 'KORMARC 조회', '목록' 버튼이 제공되는데 '목록' 버튼 은 정상작동하고 있지 않은 상황이다.

서지 레코드 중 본 연구의 핵심요소인 청구기호는 크게 사부분류코드와 네 자리 숫자로 구성되며 같은 저작은 끝에 알파벳을 붙여 구분한다. 예를 들어 주희(朱熹)의 맹자(孟子)는 'A09E-0001'이라는 청구기호를 부여받았으며 앞의 'A09E'는 사부분류코드이고 뒤의 네 자리 수는 해당 분류에 속하는 레코드의 일련번호이다. 만일 다른 자원이 위의 맹자와 동일한 저작이라 판단되면 'A09E-0001a'와 같이끝에 알파벳을 붙여 구분하므로 사부분류코드와 네 자리 수를 저작 식별자로 활용하는 것이 가능하다.

### 3.2 실험 시스템 설계

본 연구에서는 FRBR 기반의 고서 검색 시 스템을 구축하기 위한 첫 번째 작업으로 구축 에 필요한 메타데이터 항목들을 정리하여 데 이터베이스 모델링 작업을 실시하였으며 그 결과는 <그림 1>과 같다.



<그림 1> 실험 시스템 모델링 ERD

RDA 기반의 검색시스템을 연구한 도슬기 (2014)는 저작 테이블 이외에도 표현형 (expression), 구현형(manifestation) 테이블 까지 별도로 설계하였다. 그러나 본 연구에서 다루는 고서의 경우 다른 수록매체(carrier)가 없으며 표현형과 구현형을 한 데 묶는 것이 효율적이라 판단하여 'Book' 테이블 하나로 구성하였다. 또한 개별자료(item)는 소장정보를 나타내므로 본 연구에서는 제외하였다.

사부분류는 고서를 연구하는 데 있어 매우 중요한 요소이므로 분류를 통한 검색 기능을 그대로 유지하기 위해 테이블을 작성하였다. 이 외에도 저자에 관련된 정보를 추가적으로 제공 하고 추후에 저자 전거로 활용이 가능하도록 기여자(Contributor) 테이블을 새롭게 작성하고 이름과 생몰년, 비고(생애) 등의 요소를 두었다.

## 4. 실험 시스템 구현

시스템 설계에 따라 MS Excel을 이용하여 데이터를 가공하였으며 본 연구에 이용된 데이터는 KORMARC 레코드가 아닌 상세 서지정보 페이지에 표시되는 메타데이터 요소들이다. 기존에 없던 요소인 기여자 관련 정보는

네이버 지식백과에서 검색하여 얻은 내용을 입력하였다. 실험 시스템은 MS Access를 사 용하여 개발하였다.

기본적으로 본 연구는 저작 중심의 검색결과 디스플레이에 초점을 맞추고 있으며 직관적이고 쉬운 검색을 위해 검색기능은 서명과저자명을 이용한 간략검색만 구현하였다. 실험 시스템의 초기 화면은 <그림 2>와 같다.



<그림 2> 실험 시스템 초기 화면

검색결과 디스플레이는 OCLC에서 개발한 Fiction Finder와 호주국립도서관에서 개발한 Trove를 참고하여 '저작 → 표현형+구현형 (판 사항) → 서지레코드'의 순으로 제시되도록 구현하였다. <그림 3>은 '맹자'를 서명으로 검색했을 때 나타나는 첫 번째 결과이며 이용자는 저작 식별자, 저작 이름, 저자, 분류의네 가지 요소를 통해 저작을 구별하게 된다.

고서서지시스템				李統國 号号	학술정보관 App 전망시민한
검색 / 열람신청 사부분류	용어 해설 공지시	참 도움말			
					/ 열람신청     r大學水 <b>후級</b> 開
	검색목록				HOME > B d > 5
흠으로	총 [23]건의 자료	가 있습니다.			
	WorkID		저작명	저자	분류
	A09A-0040	○ 四書旁訓(孟子)		馬夢頑	經集·通義
	A09E-0001	孟子		朱憙	孟子
	A09E-0002	맹자인해(孟子總解		宣組	孟子
	A09E-0003	인해맹자(言解孟子		李範圭	孟子
	A09E-0004	孟子栗谷先生諺解		李珥	孟子
	A09E-0005	孟子集註大全		胡廣	孟子
	A09E-0006	孟子或問精義通攷		宋時烈	孟子
	A09E-0007	增補蘇批孟子		新沙	孟子
	A09E-0008	孟子請演		李在夏	孟子
	A09E-0009	監本孟子		朱憙	孟子
	A09E-0010	孟子大全		胡廣	孟子

<그림 3> 1단계 검색결과

#### 146 제25회 한국정보관리학회 학술대회 논문집

이용자가 첫 번째 화면에서 정보 요구에 맞는 저작을 찾아 클릭하면 저작 정보 페이지로 이동한다(<그림 4> 참조). 해당 페이지에서는 저작의 기본 정보와 함께 '표현형+구현형'정보를 살펴볼 수 있다.

작 정보					딛
WorkID	A09E-000	1			
저작명	孟子				
저자	朱憙				
분류	孟子				
litions of	this w	ork			
litions of	this w	ork <sub>판종</sub>	간사지	간사년	
			간사지 중국	간사년 [清朝年間]刊	
서명		판종			
서명 <u>국</u>		<b>판</b> 종	중국	[清朝年間]刊	
서명 <u>조구</u> 조 <u>구</u>		・   ・   ・   ・   ・   ・   ・   ・   ・   ・	중국 중국	[清朝年間]刊	
서명		世帝           木版本           木版本	증국 증국 한국	[清朝年間]刊 [清朝末期]刊 [朝鮮朝末期]刻[後刷]	

<그림 4> 2단계 검색결과: 저작 정보

두 번째 화면의 '표현형+구현형' 목록에서 자신이 원하는 서지 자원을 발견하여 클릭하 면 서지 상세정보 페이지로 이동하여 개별 레 코드에 대한 확인이 가능하다(<그림 5> 참조).

BookID	BK000070 WorkID		A09E-0001 사=		사부분류	孟丁	
서명	孟子						
	표제	NA		판심제	NA		
판종	木版本		간사지	중국		간사년	[淸朝年間]刊
장정	線裝		권책수	7卷3冊			
광곽	左右雙邊	上下單邊	반곽	19.0cm(総	t) x 12.5cm(横)	계선	有界
항자수	9行17字		주석	註雙行,頭	註		
흑구	NA		어미	NA		크기	25.0cm(縱) x 13.0cm(横
지질	竹紙		기타	NA		청구기호	A09E-0001
주기사항	NA						

<그림 5> 3단계 검색결과: 서지 상세정보

이와 같이 기본적인 자원 탐색은 세 단계에 걸쳐 이루어진다. 그러나 이용자가 저자에 대한 정보도 함께 확인할 수 있도록 저작 정보 와 서지 상세정보 페이지에서 언제든지 저자 의 이름을 클릭하여 저자정보 페이지로 이동 할 수 있게 구현하였다(<그림 6> 참조). 뿐만 아니라 저자의 저작을 클릭하면 저작 정보 페 이지로 이동하게 만듦으로써 자유로운 탐색이 가능하도록 하였다.

저자 정보				닫기
ContributorID	CB000053			
이름(대표형)	朱憙			
이름(이형)	朱熹	주희	주자	
생몰년	1130-1200			
비고	주자로 높여 특	부른다. 남송 시대의	의 이학가(理學家,	지대한 영향을 미친 주희는 흔히 성리학자)로 자는 원회(元晦, 뒤 (晦翕), 둔용이라 했다.
 저작 활동				
저작 활동 <sub>서명</sub>	1	역할		J
		역할		
서당	0			
監本論語	\$	저자		

<그림 6> 저자 정보 페이지

## 5. 실험 시스템 분석

FRBR 저작 개념 기반의 시스템을 실험 제작한 결과 본 연구의 목적인 저작 중심의 검색결과 탐색 및 저자 정보 제공은 상당 부분 달성되었다. 현행 시스템에서 '맹자'를 검색한 결과는 총 47건이며 14건을 제외한 33건이 최소 한 개 이상의 동일한 서명 및 저자명을 가지고 있었다. 이는 이용자로 하여금 검색결과를 살펴보는 데 혼란을 줄 여지가 있다. 그러나 저작 단위로 검색결과를 표시하도록 시스템을 구현한 결과 47건이었던 검색결과가 23건으로 감소하였다. 맹자 이외에도 이용자가 처음 마주하게 되는 검색결과의 수가적게는 21건에서 많게는 37건까지 감소하였다(<표 1> 참조).

〈표 1〉 시스템별 사서(四書) 검색결과 건수

항목	현행 시스템	실험 시스템
논어	68건	31건
맹자	47건	23건
대학	67건	46건
중용	55건	31건
합계	233건*	131건

\* 중복 4건 제외

이외에도 저자 정보 테이블을 별도로 생성 함으로써 저자와 관련된 정보를 열람하거나 저자의 다른 작품들을 검색하는 것이 가능해 졌다. 저자 테이블의 경우 실험 제작된 시스 템에는 반영되지 않았지만 저자 전거 관리에 활용할 수 있는 가능성이 존재한다는 점에서 의미가 있다. 또한 두 번째 검색결과 화면인 저작 정보 페이지에서 동일 저작의 여러 판본 을 비교할 수 있다는 점은 판본과 간사년의 구분이 중요한 고서 연구에 있어 상당한 의미 를 가질 것으로 보인다.

그러나 저작에 대한 기본 정보를 저작 식별 자, 저작 이름, 저자, 분류의 네 가지 요소만 빈약하게 제공하고 있는 실정이다. 따라서 더 많은 정보를 제공할 수 있도록 메타데이터 요 소의 개발이 필요하다. 뿐만 아니라 실험 시스 템에는 관련 있는 저작들 간의 관계가 연결되 어 있지 않은데 이를 개선하여 관련된 다른 저 작을 소개할 수 있다면 고서 및 고문서 연구자 들에게 큰 도움이 될 것으로 여겨진다. 이용자 의 요구에 맞게 검색결과를 필터링(filtering) 할 수 있는 패싯(facet) 기능 또한 추가할 필요 가 있다.

한편 존경각의 청구기호를 기반으로 저작 식별자를 부여하는 것은 저자 전거와 더불어 통일 표제를 관리할 수 있는 테이블을 제작하 는 기반이 될 수 있을 것으로 보인다. OCLC 와 LC를 비롯한 해외 기관에서 개발된 FRBR 알고리즘들은 기본적으로 저자명과 표제의 대 표형을 한 쌍으로 묶은 뒤 저작 식별자를 부 여한다. 이와 같은 알고리즘을 사용하면 저작 의 대표형 표제 또는 저자명과 다른 이형의 표제나 저자명을 동일한 저작으로 판단하지 못하는 경우가 많다. 그러나 존경각의 서지 레코드는 이미 저작들을 구분하여 묶어 놓았 기 때문에 저작 식별자를 중심으로 서지 자원 을 정렬한 이후에 서명을 가져오게 되므로 다 양한 종류의 서명을 한 데 모으는 것이 가능

하다. 본 연구에서 사서(四書)를 검색하여 얻 은 결과에서는 서로 다른 서명을 가진 동일한 저작이 없어 실험 시스템에는 적용하지 못했 지만 이러한 사례들을 확인하고 활용할 수 있 는 방안을 마련할 필요가 있다.

## 6. 결론 및 제언

성균관대학교 존경각의 고서서지 목록데이 터는 품질이 좋을 뿐만 아니라 저작을 중심으 로 공통된 청구기호를 할당받아 FRBR 구조로 변환하는 데 적합한 특징을 가지고 있기 때문 에 데이터 수준에서 저작 식별자를 바로 부여 하는 것이 가능하다. 본 연구에서 이러한 방 식은 사용하여 시스템을 구축해 본 결과, 전 거데이터 및 통일표제를 사용하는 OCLC의 알고리즘이나 키워드 검색의 결과물들을 저자 명과 표제 쌍으로 묶어 저작 식별자를 부여하 는 LC의 알고리즘보다 수월하게 MARC 레코 드를 FRBR 구조에 맞게 변환할 수 있었을 뿐 만 아니라 저작 식별의 정확성이 높은 것으로 나타났다.

현행 시스템은 동일한 저작일지라도 판 사 항을 비롯한 여러 메타데이터 항목에 따라 다 른 자료인 것처럼 검색결과를 표시하기 때문 에 이용자는 검색결과를 일일이 살펴봐야 한 다. 반면에 실험 시스템은 저작을 중심으로 서지 자원들을 묶어서 표시함으로써 이용자가 살펴봐야 하는 검색결과의 수를 줄여 자료 탐 색이 효율적일 뿐만 아니라 한 화면에서 동일 저작의 여러 '표현형+구현형' 자료를 비교해 보는 것도 가능하게 한다. 하지만 본 연구에 서는 존경각 고서서지시스템의 실제 이용자들 이 새로운 시스템에 대해 느끼는 용이성, 효 율성, 만족도 등을 확인할 수 없다. 따라서 FRBR 기반 고서서지시스템에 대한 이용자 평 가 후속 연구가 필요하다.

148 제25회 한국정보관리학회 학술대회 논문집

## 참고문헌

도슬기. (2014). RDA 기반 서지검색시스템에 대한 이용자 평가에 관한 연구. 석사학 위논문. 성균관대학교 일반대학원, 문헌 정보학과.