FRBR 모형 기반 서지검색시스템의 검색 효율성 평가 연구

An Experimental Study on the Retrieval Efficiency of the FRBR Based Bibliographic Retrieval System

김 현 희(Hyun-Hee Kim)*

〈 목 차 〉

Ⅰ. 서 론
Ⅱ. 이론적 배경
1. FRBR 데이터 모형
2. 선행 연구
3. 연구 가설
Ⅲ. 실험 시스템 설계와 구축
1. 데이터베이스 표본 선정
2. OPAC 기반 시스템
3. FRBR 기반 시스템
Ⅳ. FRBR 기반 시스템의 검색효율성 평가
1.실험 설계
2.데이터 분석과 가설 검증
3.분석결과에 따른 시스템간 상호비교
Ⅴ. 결 론

초 록

본 연구에서는 FRBR 모형 기반 서지검색시스템의 검색 효율성을 평가하기 위해서 두 개의 실험 시스템을 구축하였다. 즉, FRBR화 알고리즘을 적용한 FRBR 기반 시스템과 일반 OPAC 기반 시스템을 387건의 KORMARC 기반 음악 자료를 데이터베이스로 활용하여 구성하였다. 그런 다음, 6개의 연구 가설을 세우고 28명의 피조사자와 표현형 질문 6개와 구현형 질문 6개로 구성된 총 12개의 검색 질문을 이용하여 이 두 시스템을 재현율, 정확률 및 검색 시간을 측정 기준으로 하여 비교, 분석하였다. 중다변량분석과 t-검증을 이용하여 실험 결과를 분석한 결과, FRBR 시스템이 질문의 유형을 가리지 않고 평균 재현율이 OPAC 시스템 보다 단연 앞서는 것으로 나타났고, 구현형 질문의 경우 OPAC 기반 시스템이 FRBR 기반 실험시스템 보다 평균 정확률이 조금 높았고 평균 검색 시간이 더 짧게 나타났다. 현 검색 환경에 FRBR 기반 시스템을 적용시키기 위해서는 우선 서지적 연관성이 높은 분야에 활용될 수 있으며 이러한 분야에 적용될 때 OPAC 시스템을 대체하기 보다는 이를 보완하면서 FRBR 시스템의 특성을 살려 질의 유형이 표현형인 경우 또는 높은 재현율을 요하는 경우 특히 효율적으로 적용될 수 있음을 제안하였다.

주제어: FRBR모형, FRBR화, 검색 효율성, 서지검색시스템, 재현율, 정확률, 검색시간, 음악자료, 구현형 질문, 표현형 질문

ABSTRACT

This study examines the retrieval efficiency of the FRBR-based bibliographic retrieval system. To do this, we built two experimental retrieval systems(a FRBR-based system constructed through FRBRizing algorithms and an OPAC-based retrieval system) using 387 music materials coded in a KORMARC format. Next, we set up six hypotheses and compared these two systems in terms of recall, precision, and retrieval time using 28 participants and a questionnaire with 12 queries. The results show that the average recall value of the FRBR-based system is higher than that of the OPAC system regardless of query types and the average precision and retrieval time values of manifestation queries of the OPAC system is more efficient that those of the FRBR-based system. This study results can be used to customize digital library interfaces as well as to improve the retrieval efficiency of the bibliographic retrieval system.

Key Words: FRBR Model, OPAC, KORMARC, FRBRizing, Retrieval Efficiency Recall, Precision, Retrieval Time

* 명지대학교 인문대학 문헌정보학과 교수(kimhh@mju.ac.kr)
• 접수일: 2007년 8월 24일 • 최초심사일: 2007년 8월 3일 • 최종심사일: 2007년 9월 21일

I. 서 론

전통적인 도서관은 목록의 기술 부분을 강조해 오고 있지만, 네트워크 환경에서의 디지털 도서관은 자료들이 접근되고, 링크되고 그리고 배열되는 방법에 더 관심을 두고 있다. 최근 이용자 중심의 디지털 도서관 시스템의 항해 구조를 설계하는 방법 중 하나로 FRBR 모형이 많은 관심을 받고 있다. Brenne(2004)은 FRBR 모형 기반 정보 시스템이 음악 자료의 축적과 검색에 적합한지를 살펴보기 위해서 전통적인 도서관 정보시스템과 비교 분석하였고, 그 연구 결과는 FRBR 모형이 전통적인 도서관 시스템에서 함축적으로 존재하는 여러 관계들을 좀 더 분명히 들어냄으로써 목록의 항해 기능을 개선시켰다고 보고하였다. 또한, Cooper(2003)는 FRBR 모형을 적용한 음악 자원을 제공하는 정보시스템인 Variations2를 구축하여 탐색 인터페이스의 사용성과 탐색 결과의 기능을 평가하였는데, 피조사자들은 Variations2를 전통적인 도서관 시스템 보다 훨씬 더 선호하였다. 이는 Variations2가 사용하기가 쉽고 레코드에 대한 자세한 정보를 제공하기 때문이라고 보고 하였다.

이러한 연구들은 이용자들이 다양한 서지적 관계를 파악하여 관련 자료들을 체계적으로 그룹핑하여 보여주는 FRBR 기반 시스템을 기존 OPAC 시스템 보다 더 선호한다고 보고하고 있다. 앞의 두 연구들이 음악 분야를 주제 카테고리로 정한 이유는 음악 자료가 다양한 파생관계를 갖고 있기 때문이다. 예를 들어서, 한 저작과 그 저작을 편곡한 여러 저작들이 있을 수 있고, 한 저작이 여러 연주 단체, 연주자 및 지휘자에 의해서 연주될 수 있으며, 더 나아가 특정 연주는 다양한 매체에 수록될 수 있고, 한 매체에 여러 저작들이 포함될 수 있다.

본 연구에서는 자료의 주제 분야, 검색 질문의 유형 또는 어떤 유형의 검색 결과를 원하느냐에 따라서 정보시스템별로 유용성과 검색 효율성이 차이가 난다고 가정하였다. 즉, FRBR 기반 시스템은 음악, 문학, 프로젝트 자료 및 방송 영상 자료와 같이 자료간의 서지적 연관성이 높은 분야에 더 적합하며, FRBR 기반 시스템이 관련 자료를 구조화하여 보여 주고 저자와 표제의 전거통제를 전제로 하기 때문에 질문의 유형에 관계없이 재현율에서 OPAC 기반 시스템의 성능을 앞설 것으로 가정하였다. 또한 정확률과 검색 시간의 경우, 구체적으로 서지정보가 알려진 자료를 찾는 구현형 질문의 경우는 OPAC 시스템이 FRBR 시스템 보다 효율성이 높고, 표현형 질문인 경우는 FRBR 시스템이 OPAC 시스템 보다 효율성이 높을 것으로 가정하였다.[1] 본 연구의 목적은 이러한 가정들을 기초로 하여 387건의 음악 자료, 12개의 질문 및 28명의 피조사자를 활용하여 두 시스템의 검색 효율성을 비교, 분석한 후 효율적인 정보검색시스템을 구현하는데 있다.

---

1) "모짜르트의 피아노 협주곡 20번을 *Vladimir Ashkenazy*가 연주한 공연/연주 자료를 찾습니다"와 같이 특정 저작(들)의 특정 연주(또는 공연, 악보 등)로 표현된 자료(들)을 찾는 질문을 표현형 질문으로 구분하고, "Edwin F. Kalmus사에서 발간한 모짜르트의 교향곡 제41번에 대한 악보를 찾습니다"와 같이 구체적으로 서지정보(예, 출판 사항 등)가 알려진 자료를 찾는 질문을 구현형으로 하였다.

Ⅱ. 이론적 배경

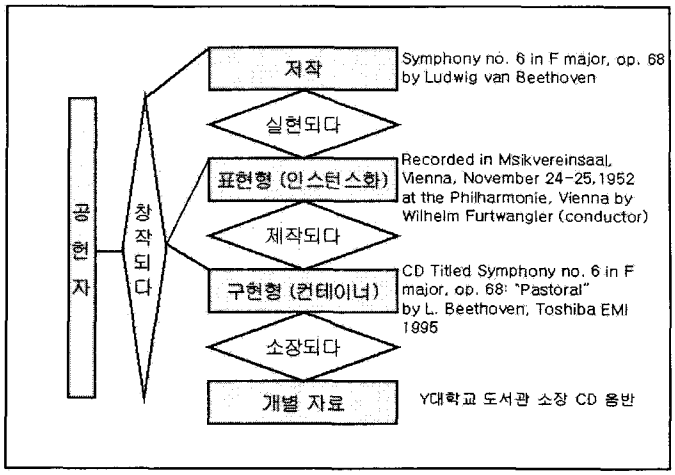
1. FRBR 데이터 모형

FRBR은 1998년 IFLA에서 제시한 이용자가 필요로 하는 데이터요소를 어떻게 조직하느냐에 중점을 두고, 다양한 서지관계를 통해 개체를 연결하여 저작을 집중시키고 항해기능을 지원하여 이용자의 검색을 돕는 참조모형이다.[2] FRBR 개념 모형에서는 서지정보개체를 3개 집단으로 정의하고 각 집단의 개체를 제1집단은 저작, 표현형, 구현형 및 개별자료, 제2집단은 개인과 단체, 제3집단은 개념, 대상, 사건 및 장소로 정의하였다. FRBR 제1집단 개체의 구체적인 의미는 다음과 같다. 저작은 지적 예술적 창작물을 지칭하는 추상적인 개념이며 이를 문자로 기록하거나, 음향, 이미지, 동작 등의 형식으로 실현한 것이 표현형이다. 또한 한 저작의 표현형을 물리적인 매체로 제작한 것이 구현형이며 하나의 구현형에 대한 구체적인 사례를 개별자료라고 한다.[3]

본 연구에서 사용한 FRBR 데이터 모형은 음악 자료를 위한 Variation2의 데이터 모형을 기본으로 하여 재구성한 것이다(그림 1 참조). Variation2의 데이터 모형은 FRBR 모형을 기본으로 하고 있으나, 표현형 대신에 인스턴스화(instantiation), 구현형 대신에 컨테이너(container)를 사용하고 있다. 즉, 인스턴스화는 공연, 연주, 악보 등으로 표현된 것으로 FRBR 표현형과 유사하며, 컨테이너는 개별적인 물리적인 매체(예, CD)를 의미하는 것으로 구현형의 의미와 유사하다. Variation2 모형에서는 디지털화된 음악 자료만을 다루기 때문에 개별자료 대신 디지털화된 사운드 파일과 같이 컨테이너를 디지털 방식으로 표현하는 미디어 객체(media object)를 사용하고 있다. 본 연구에서 사용한 데이터 모형에서는 디지털화된 자료 뿐만 아니라 인쇄본 자료도 다루어야 하기 때문에 미디어 객체 대신 개별자료를 사용한다. 〈그림 1〉은 베토벤 교향곡 제6번에 대한 저작을 FRBR 모형으로 표현한 것으로 교향곡의 표현형을 구분하기 위한 서지요소로 연주단체와 지휘자를 사용하였고, 구현형으로는 이 저작이 수록된 1995년에 발간한 CD 자료를 예로 들었다.

2) 이성숙, FRBR 모형의 수용에 관한 연구(박사학위논문, 연세대학교 대학원 문헌정보학과, 2005), p.15.

3) IFLA, *Functional Requirements for Bibliographic Records: final report*, UBCIM Publications, New Series Vol.19(Munich: K. G. Saur, 1998), pp.82-87.



〈그림 1〉 FRBR 데이터 모형

2. 선행연구

FRBR 모형 기반의 정보검색시스템을 구현하여 검색 효율성과 시스템 사용성 평가를 수행한 연구를 살펴보면 다음과 같다.

처음에 언급하는 두 연구는 음악자료의 검색에 FRBR 모형을 적용한 경우로 첫째, Brenne[4]은 FRBR 모형 기반 정보 시스템이 음악 자료의 축적과 검색에 적합한지를 살펴보기 위해서 전통적인 도서관 정보시스템과 비교 분석하였다. 연구 방법은 문헌 연구 및 음악 분야 사서와의 면담을 이용하였다. 구축한 실험 시스템에는 존 레논과 폴 매가트니의 "Yesterday"에 대한 서지 레코드를 표본 자료로 활용하였다. 그는 FRBR 모형이 하나의 자료가 여러 개의 저작들로 구성되어 있다 하더라도 하나의 서지 레코드가 하나의 자료에 대해서 구성되는 현재 서지 레코드의 문제점을 해결할 수 있다고 제안하였다. 즉, FRBR 모형을 적용하면 구분된 저작 레코드들이 각 저작을 위해서 만들어지며 이는 관심이 없는 레코드들을 걸러내는 기능을 높여 줄 것이라고 하였다. 이 연구 결과는 FRBR 모형이 전통적인 도서관 서지정보검색 시스템에서 함축적으로 존재하는 여러 관계들을 좀 더 분명히 들어냄으로써 목록을 항해하는 기능을 개선시켰다고 보고하였다. 둘째, Cooper[5]는 인디아나대학의 디지털음악 도서관 프로젝트로 FRBR 모형을 적용한 음악 자원을 제공하는 정보시스템인 Variations2를 구축하여 탐색 인터페이스의 사용성과 탐색 결과의 기능을 평가하였다. 연구 방법은 대학원 수준의 음악 과목을 이수하고 있는 7명의 학생을 대상으로 하였다. 평가

4) Marte Brenne, *Storage and Retrieval of Musical Documents in a FRBR-based library catalogue*(Master Thesis, Oslo University College, 2004).

5) Michelle Cooper. Variations2: IU Digital Music Library Version 2.1 Usability Test Report: Searching(2003). [cited 2007.6.1] 〈http://www.dml.indiana.edu/pdf/V2SearchUsabilityReport12-03.pdf〉

결과, Variations2의 탐색 결과에 대한 만족도는 보통으로 나왔다. 그러나 피조사자들은 Variations2를 전통적인 도서관 시스템인 IUCAT 보다 훨씬 더 선호하였는데 이는 시스템을 사용하기가 쉽고 레코드에 대한 자세한 정보를 제공하기 때문이라고 하였다.

다음 세 연구는 음악 분야 이외의 분야에 FRBR 모형을 적용한 연구들이다. 정진규[6]는 방송영상자료의 효율적인 메타데이터 시스템을 구축하기 위하여 FRBR 기반의 서지구조 모형을 개발하고 실험시스템을 구축하여 검색 효율성과 유용성의 측면에서 이용자 평가를 실시하였다. 30명의 피실험자들을 선정하여 FRBR 기반 실험 시스템을 일반시스템과 비교한 결과 검색효율성 측면에서 검색 질의 수는 차이가 없었고, 검색에 소요된 접근 단계 수는 FRBR 기반 시스템이 더 많은 것으로 조사되었다. 유용성 측면에서는 화면구성의 복잡 정도에 대한 항목을 제외하고는 모두 FRBR 기반 시스템의 만족도가 더 높은 것으로 조사되었다.

그리고 김소형[7]은 영화정보의 체계적인 관리와 효율적인 이용을 위하여 전문가집단의 평가를 통해 메타데이터 요소를 개발하고 FRBR 기반의 시스템을 구현하여 구성된 질의에 따른 검색결과를 현행시스템과 비교분석하였다. 시스템간의 유용성을 평가하기 위해서 일반 이용자를 대상으로 만족도를 수행하는 것 대신에 연구자가 직접 비교분석하였다. 분석 결과 연관저작의 서지적 관계 표현 측면에서는 형식적 관계와 내용적 관계 표현에 모두 효과적이었으며, ISAN과 수퍼 저작 개념을 도입하여 연관저작의 집중기능이 강화되었다. 또한 이성숙[8]은 저작과 표현형의 관계유형을 재정의 한 M-FRBR 모형을 개발하고, 이를 한국문학작품에 적용하여 그 활용성을 검토하였다. 그 결과 저작과 표현형간의 경계의 모호함이 개선되었고, 목록의 집중기능 강화, 복수버전의 효과적 처리, 개선된 인터페이스를 통한 항해기능이 강화되어 이용자의 서지적 관계검색에 유용한 것으로 확인되었다.

이러한 선행 연구들은 기존 OPAC 시스템과 비교하여 FRBR 모형 기반 시스템이 목록의 집중기능을 강화시켜 목록을 항해하는 기능을 개선시켰다고 보고하고 있다. 본 연구에서는 FRBR 기반 시스템과 OPAC 시스템이 질문의 유형에 따라서 검색 효율성이 다를 것으로 가정하였다. 이러한 가정은 다음과 같은 6개의 연구 가설들로 구성된다.

3. 연구 가설

가. 연구 가설 1: 표현형 질문의 경우 FRBR 기반 실험시스템이 일반 시스템(OPAC 기반 시스템) 보다 평균 재현율이 더 높게 나타날 것이다.

---

6) 정진규, 방송영상자료의 FRBR기반 서지구조모형에 관한 연구(박사학위논문, 성균관대학교 대학원 문헌정보학과, 2006).

7) 김소형, 영화정보를 위한 FRBR 모형 메타데이터 시스템 구현에 관한 연구(박사학위논문, 성균관대학교 대학원 문헌정보학과, 2006).

8) 이성숙, FRBR 모형의 수용에 관한 연구(박사학위논문, 연세대학교 대학원 문헌정보학과, 2005).

나. 연구 가설 2: 구현형 질문의 경우 FRBR 기반 실험시스템이 일반 시스템(OPAC 기반 시스템) 보다 평균 재현율이 더 높게 나타날 것이다.

다. 연구 가설 3: 표현형 질문의 경우 FRBR 기반 실험시스템이 일반 시스템(OPAC 기반 시스템) 보다 평균 정확률이 더 높게 나타날 것이다.

라. 연구 가설 4: 구현형 질문의 경우 일반 시스템(OPAC 기반 시스템)이 FRBR 기반 실험시스템 보다 평균 정확률이 더 높게 나타날 것이다.

마. 연구 가설 5: 표현형 질문의 경우 FRBR 기반 실험시스템이 일반 시스템(OPAC 기반 시스템) 보다 평균 검색 시간(초)이 더 짧을 것이다.

바. 연구 가설 6: 구현형 질문의 경우 일반 시스템(OPAC 기반 시스템)이 FRBR 기반 실험시스템 보다 평균 검색 시간(초)이 더 짧을 것이다.

Ⅲ. 실험 시스템 설계와 구축

여섯 개의 연구 가설을 검증하기 위해서 표본으로 사용할 데이터베이스를 선정하고 두 개의 실험 시스템 즉, FRBR 기반 시스템과 OPAC 기반 시스템을 구현하였다.

1. 데이터베이스 표본 선정

FRBR화 알고리즘 적용을 위해서 모짜르트와 베토벤의 교향곡, 피아노 협주곡 및 피아노 소나타의 각 한 저작을 선정하여 총 6개의 저작을 표본 저작으로 선정하였다. KERIS 종합목록을 사용하여 저자명과 표제명 입력 및 결과내 검색 기능을 활용하여 자료를 검색한 후 중복된 자료 또는 관련 없는 자료를 제외하여 총 387건의 자료를 검색하여 이용하였다(표 1 참조).

〈표 1〉 데이터베이스 표본

모짜르트		베토벤	
저 작	건 수	저 작	건 수
(1) Symphony No. 41 in C Major K.551 "Jupiter"	93	(1) Symphony No. 6 in F Major, op. 68 "Pastoral"	54
(2) Piano Concerto No. 20 in d Minor, K.466	53	(2) Piano Concerto No.5 in Eb major Op.73. "Emperor"	98
(3) Piano Sonata No.11 A major K.331 "Turkish March"	25	(3) Piano Sonata No.23 in f minor op.57 "Appassionata"	64
소 계	171	소 계	216
전 체 합 계			387

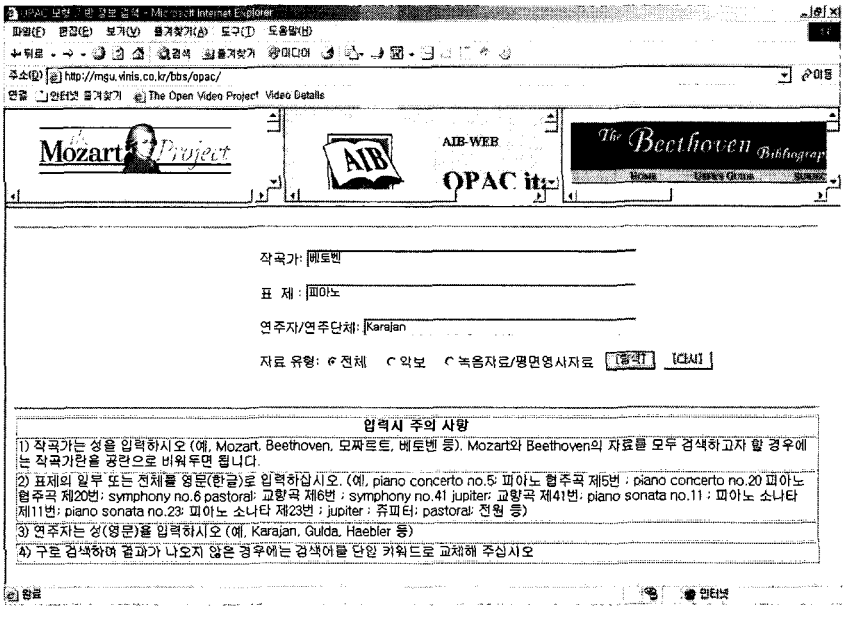
2. OPAC 기반 시스템

가. 콘텐츠 구축

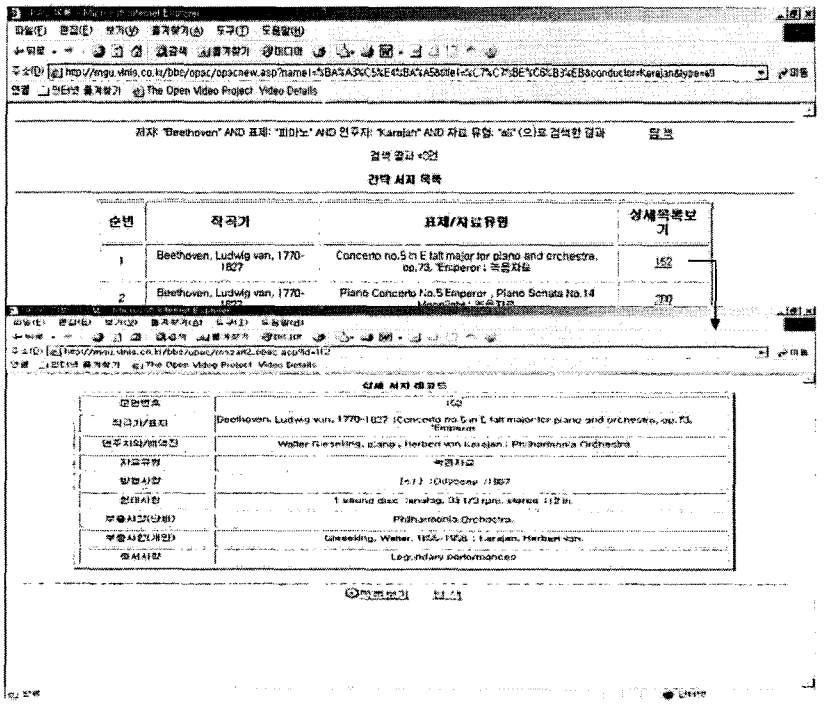
KERIS에서 다운 받은 387건의 KORMARC 자료를 이용하여 OPAC 시스템의 콘텐츠를 구축하였다. 즉, 다운로드한 387건의 자료에 대한 리더(06), 245, 260, 300, 653 태그 등 총 13개의 태그값을 액세스 데이터베이스의 13개의 필드에 입력하였다. 음악자료를 기술하기 위해서 대부분의 도서관이 비도서 KORMARC를 사용하고 있었지만 일부 도서관에서는 단행본 KORMARC를 사용하고 있었다. 대학에 따라 특정 태그값을 입력하지 않은 경우에는 필드를 공란으로 남겼다.

나. 인터페이스

OPAC 기반 시스템의 인터페이스는 〈그림 6〉에서처럼 작곡가, 표제 및 연주자/연주단체로 검색할 수 있으며 자료 유형을 선택하여 자료를 유형별로 제한할 수 있는 기능을 추가하였다(그림 2 참조). 작곡가는 전거 파일을 통해서 다양한 표현들(예, Mozart, 모짜르트, 모차르트 등)을 통제하였으나, 표제는 표제에서 사용한 키워드 또는 구절만으로 검색될 수 있도록 하여 전거 통제를 하지 않았다. 〈그림 2〉 초기 화면에서 작곡가는 "베토벤", 표제는 "피아노", 연주자/연주단체는 "Karajan" 그리고 자료 유형은 "전체"를 선택하면, 〈그림 3〉에서 처럼 검색 결과의 간략화면과 상세화면을 확인해 볼 수 있다.



〈그림 2〉 OPAC 기반 시스템의 초기 화면



〈그림 3〉 OPAC 기반 시스템의 간략 서지 목록 및 상세 서지 레코드

3. FRBR 기반 시스템

가. 콘텐츠 구축

FRBR 기반 시스템의 콘텐츠 구성은 OPAC 기반 시스템을 위해서 구축한 기본 데이터베이스에 표현형 필드를 추가하여 총 14개 필드로 구성하였다. 표현형 필드에는 "연주/공연: Berliner Philharmoniker(연주단체)-Herbert von Karajan(지휘자)", "악보-한국어"와 같은 데이터 값이 수록되어 있다. 갱신된 기본 데이터베이스 외에 FRBR화를 위한 데이터베이스를 추가로 구축하였다.

FRBR화를 위한 데이터베이스 구축을 위해서 FRBR화 알고리즘을 기본 데이터베이스에 적용하였다.[9][10] FRBR화를 위한 데이터베이스는 저작(작곡가와 표제), 표현형, 문헌유형 및 관련문헌 건수의 4개의 핵심 필드로 구성되며 이 데이터베이스는 표현형 필드를 통해서 기본 데이터베이스와 연결된다. 즉, FRBR화를 위한 데이터베이스를 통해서 저작과 표현형이 결정되는데 이때 저작은 전거통제된다. 저작과 표현형이 선택된 후 이것과 연결된 레코드들에 대한 상세 정보(구현형

9) 본 연구에서 사용된 알고리즘은 『김현희, 유영준 & 박서은 저. "FRBR 모형의 KORMARC 데이터베이스로의 적용 가능성에 대한 실험적 연구: 음악자료를 중심으로," 한국도서관·정보학회지, 제38권, 제1호(2007.3), pp.299-314』에서 사용된 알고리즘을 재구성한 것이다.

10) 알고리즘을 이용하여 데이터베이스를 FRBR화 하는 상세한 과정은 부록 1을 참조하기 바람.

정보)는 기본 데이터베이스를 이용하여 확인한다. 사용된 FRBR화 알고리즘은 현재 대부분의 도서관이 사용하고 있는 비도서용 KORMARC에 기반한 것으로 자세히 설명하면 다음과 같다.

(1) 제 1단계: 저작 확인 단계

(가) 저작을 저자명과 표제를 활용하여 확인하는 단계로 첫째, "100abcd(또는 110abcd, 111acdn)"를 순서대로 체크하여 저자명을 확인한다. 1XX 태그를 사용하지 않은 경우에는 "700abcd(또는 710abcd, 711acdn)"를 순서대로 체크하여 저자명을 확인한다. 둘째, "240adkmnp$r(또는 245$anp)"를 순서대로 확인하여 표제를 확인한다.

(나) 저작을 표제를 활용하여 확인하는 단계로 1XX와 7XX 태그를 사용하지 않은 서지 레코드에 적용하는 단계로 첫째, "130akmnpr"를 순서대로 확인한다. 둘째, "240adkmnp$r(또는 245$anp)"를 순서대로 체크한다.

(2) 제 2단계: 그룹화된 저작을 표현형으로 구분하는 단계

악보와 공연/연주로 구분한다. 악보는 악보의 종류와 해당 악보가 어떤 언어를 이용하여 표현했는지를 하나의 표현형으로 기술할 수 있다. 공연/연주인 경우에는 하나의 표현형으로 기술하기 위해서는 연주자, 지휘자, 연주 단체는 물론 연주 장소와 일자를 결합하여 하나의 표현형으로 기술할 수 있을 것이다.

(가) 악보: Leader/06을 이용하여 악보의 종류를 확인한 다음 이를 다시 008/35-37을 사용하여 언어별로 구분한다. 즉, '필사악보-한국어판', '필사악보-독일어판' 등으로 구분한다. 008/35-37 정보가 없는 경우에는 041$a를 이용하여 언어를 확인하였다.

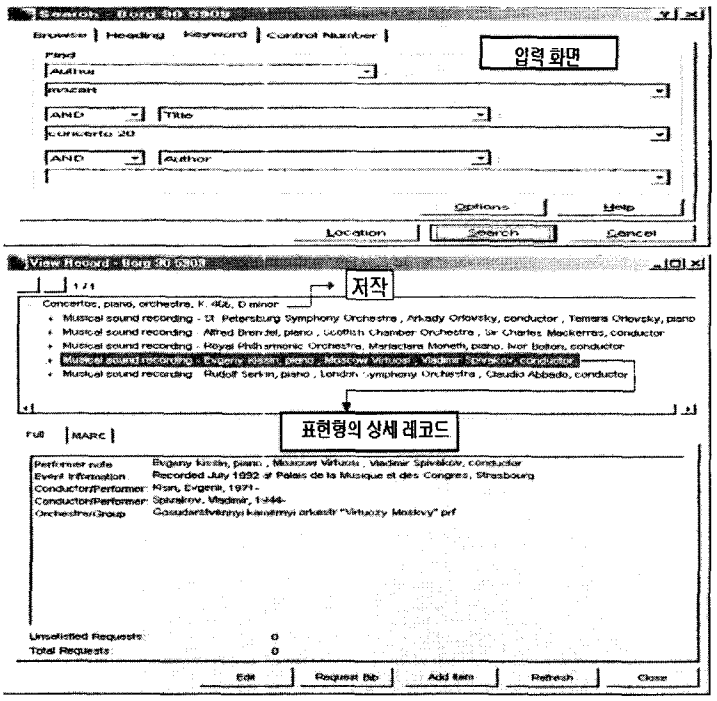
(나) 공연/연주: 'Leader/06'을 활용하여 확인하여 그 값이 평면영사자료 또는 녹음자료(음악)이면 공연/연주 자료로 확인한다. 공연/연주로 확인되면 700 adem(또는 710abem)를 이용하여 동일한 정보를 확인하여 연주자, 연주단체 및 지휘자에 대한 정보를 확인한다. 7XX태그를 사용하여 연주정보를 확인할 수 없는 경우에는 511$a와 518$a 태그를 이용하여 연주자와 배역진 그리고 촬영/녹음 일시와 장소를 확인한다.

(3) 제 3단계: 각 표현형에 속한 구현형들을 "260bc$g"과 "300$abcef$g"를 순서대로 확인하여 그룹핑하는 단계이다.

나. 인터페이스

실험 시스템의 인터페이스 설계를 하기 전에 VTLS사에서 개발한 VIRTUA를 표본 시스템으

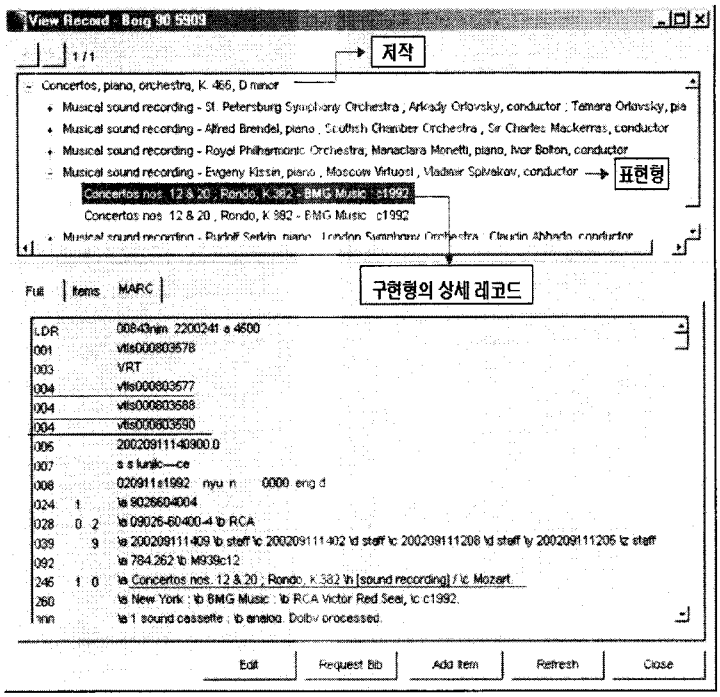
로 정하고 인터페이스의 구성을 자세히 살펴보기로 한다.[11][12] 〈그림 4〉에서 위쪽 화면은 입력 화면으로 작곡가와 표제를 입력할 수 있도록 설계되어 있다. 〈그림 4〉의 아래 화면은 검색된 저작의 한 표현형(Evgeny Kissin, piano; Moscow Virtuosi(Orchestra); Vladimir Sprivakov, conductor)에 대한 상세 레코드를 보여 주고 있다. 〈그림 5〉는 〈그림 4〉에서 선택한 표현형에 속한 특정 구현형에 대한 상세 레코드를 보여주고 있다. 구현형 상세 레코드는 MARC21 형식과 관련된 서지 레코드를 위한 004 통제 번호로 기술되어 있다. 즉, 세 개의 004태그들이 기술되어 있는데 이는 해당 레코드가 세 개의 다른 표현형 레코드들과 연관되어 있음을 나타내며, 더 나아가 245태그에서 확인할 수 있듯이 해당 구현형 레코드는 하나 이상의 저작을 갖고 있는데 이는 모차르트의 "Concerto no. 12"와 "Concerto no. 20 the Rondo, k.382"이다.



〈그림 4〉 VIRTUA 시스템의 화면 1

11) Mimno, D., Crane, G., and Jones, A. "Hierarchical Catalog Records Implementing a FRBR Catalog." *D-Lib Magazine*. Vol.11, No.10(October 2005), [cited 2007. 5. 1] 〈http://www.dlib.org/dlib/october05/crane/10crane.html〉

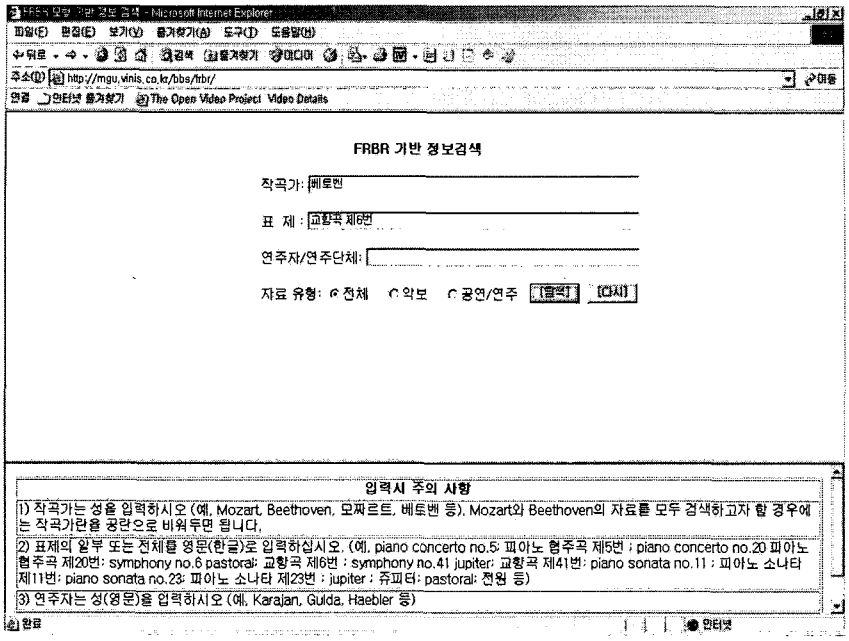
12) Virtua. Visionary Technology in Library Services Inc. [cited 2007. 4. 21] 〈http://www.vtls.com/Products/virtua.shtml〉.



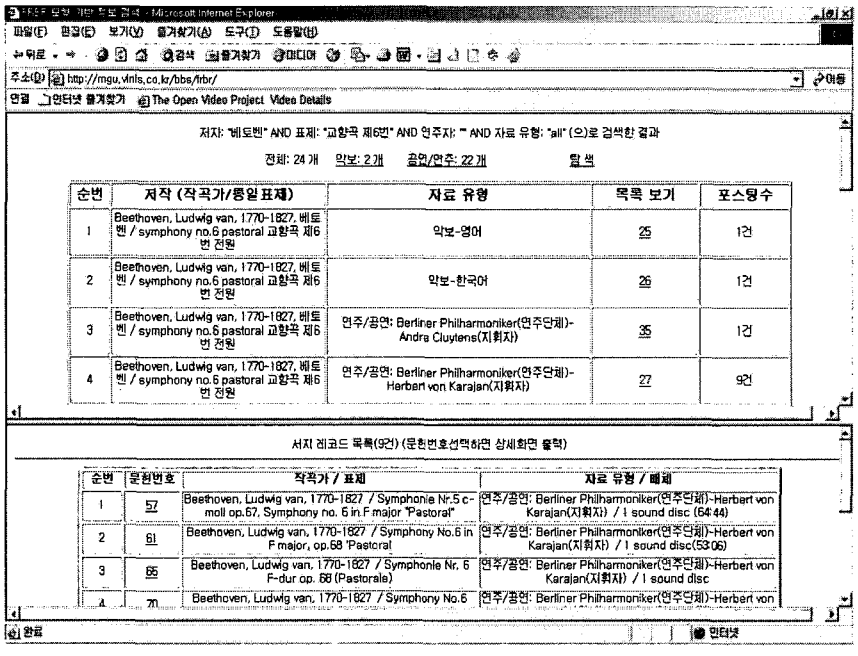
〈그림 5〉 VIRTUA 시스템의 화면 2

앞에서 표본 시스템으로 기술한 VIRTUS의 인터페이스를 참고하여 실험 시스템의 검색 인터페이스를 설계한 후 도서관 사서 2명과 음악 분야 전공자 1명을 대상으로 인터페이스를 평가하게 한 다음 최종적으로 〈그림 6〉과 같은 검색 초기 화면을 구성하였다. 작곡가, 표제 및 연주자/연주단체로 검색할 수 있으며 자료 유형을 선택하여 자료를 유형별로 제한할 수 있는 기능을 추가하였다.

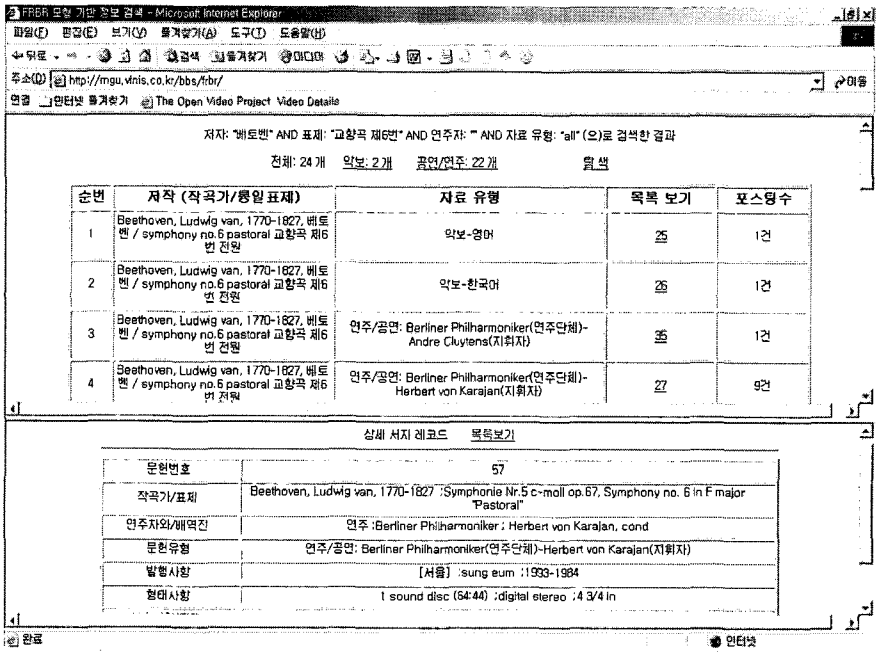
예를 들어서, 작곡가는 "베토벤", 표제는 "교향곡 제6번" 그리고 자료 유형은 "전체"를 선택하면 〈그림 7〉의 화면 위쪽과 같은 검색 결과의 간략화면을 보여준다. 〈그림 7〉에서 처럼 공연/연주 "22건", 악보 "2건"이 검색되었으며, 이 중에서 "연주/공연: Berliner Philharmoniker(연주단체)-Herbert von Karajan(지휘자)"를 선택하면 〈그림 7〉의 화면 아래쪽에 선택한 표현형에 대한 구현형 레코드들의 간략 정보가 배열되며, 간략 레코드 중 하나를 선택하면 〈그림 8〉의 화면 아래쪽에 선택된 레코드의 상세화면이 배열된다. 본 실험 시스템의 인터페이스는 하나의 화면에서 특정 저작의 표현형과 구현형에 대한 간략 및 상세 정보를 확인할 수 있도록 하였고, 처음 선택하지 않은 표현형을 다시 선택하여 해당 표현형과 구현형에 대한 정보를 계속 브라우징 할 수 있어서 화면을 이동하지 않고도 관련 자료들을 저작, 표현형 및 구현형의 3단계 계층적 구조에서 확인해 볼 수 있다.



〈그림 6〉 FRBR 기반 시스템의 초기화면



〈그림 7〉 FRBR 기반 시스템의 간략화면



〈그림 8〉 FRBR 기반 시스템의 상세화면

Ⅳ. FRBR 기반 시스템의 검색효율성 평가

1. 실험 설계

가. 설문지

설문지는 표현형과 구현형을 위한 각각 6개의 질문을 구성하여 총 12개의 질문을 구성하였다.[13] 12개 질문에 대한 검색시간(초)과 적합 문헌이라고 판단되는 문헌번호(들)을 기술하게 하였다. 이러한 자료를 바탕으로 검색시간, 재현율 및 정확률을 계산하였다. 설문지는 두 가지 유형으로 작성하였다. 즉, FRBR 기반 시스템과 OPAC 기반 시스템 중 어떤 시스템을 먼저 사용하느냐에 따라서 그 결과값에 영향을 받는다고 판단하여 한 유형은 FRBR 기반 시스템을 사용하도록 하였고, 다른 유형은 OPAC 기반 시스템을 사용하도록 하였다. 표현형과 구현형에 관한 질문을 소개하면, "표현형: Beethoven(베토벤)의 piano concerto no.5(피아노 협주곡 제5번) 중에 Brendel(Alfred Brendel)과 함께 연주한 연주단체(오케스트라)들을 모두 찾습니다."와 "구현형: Furtwangler(Wilhelm Furtwangler)가 지휘하고 Vienna Philharmonic Orchestra가 연주한 Beethoven(베토벤)의 표제

13) 부록 2 참조.

명이 "Symphony: No.1~No.9"인 EMI에서 1988년 발행한 LP판을 찾습니다"이다. 이외에 FRBR 기반 시스템과 OPAC 기반 시스템의 좋은 점(들)과 불편한 점(들) 또는 개선되어야 할 점(들)을 자유롭게 기술하도록 하였다.

나. 피조사자

설문지의 내용상 피조사자의 자격으로 정보검색과 FRBR 모형에 대한 기초 지식이 필요하다고 판단되었다. 따라서 피조사자는 임의로 선정한 6개의 대학 도서관에서 각 1명의 사서, 총 6명의 사서와 FRBR 관련 과목을 수강한 M대학의 문헌정보학과의 22명의 학부 및 대학원생들을 포함하여 전체 28명을 대상으로 하였다. 22명의 학부 및 대학원생들에게는 이 두 시스템에 대한 소개를 30분간 하고 설문지를 배포한 후 이메일로 회수하였다. 사서인 경우 조사자가 개별적으로 이 두 시스템에 대해서 30분간 설명을 한 다음 설문지를 배포하고 직접 또는 이메일을 통해서 설문지를 회수하였다. 조사는 8월 1일부터 16일 까지 진행되었다.

2. 데이터 분석과 가설 검증

가. 중다변량분석

중다변량분석(MANOVA)은 변량분석(ANOVA)의 일반화된 형태로 독립변인의 주효과와 상호작용 효과를 검증하는 일차적인 목적을 가지고 있으며, 여러 종속변인의 선형조합에 대한 독립변인의 효과의 검증이 가능하다. 본 연구에서는 독립 변인으로는 시스템 유형(FRBR 및 OPAC 기반 시스템), 질문 유형(표현형 및 구현형 질문) 그리고 이 독립변인들의 상호작용으로 구성되며, 종속변인은 재현율, 정확률 및 검색 시간이 된다. 분석 결과, 재현율에서 시스템 유형, 질문 유형은 물론 두 독립변인의 상호작용도 종속 변인에 모두 영향을 미치는 것으로 나타났고, 정확률과 검색 시간에서는 시스템 유형과 질문 유형의 상호 작용만이 종속 변인에 미치는 효과가 통계적으로 유의미한 것으로 나타난다. 분석 결과는 가설 검증에서 상술된다.

〈표 2〉 중다변량분석 결과

종속변인 / 독립변인	재현율 f(p)	정확률 f(p)	시간(초) f(p)
시스템유형	1144.38(0.00)**	1.26(0.27)	3.30(0.08)
질문유형	150.19(0.00)**	0.29(0.59)	3.82(0.06)
시스템유형 * 질문유형	227.54(0.00)**	16.49(0.00)**	4.13(0.05)*

**는 0.01 수준에서 유의미하고 *는 0.05 수준에서 유의미함.

나. 가설 검증

독립변인과 종속변인간의 관계를 보여주는 중다변량분석 자료를 활용하여 연구 가설을 검증하기 위해서 t-검증을 실시하였다. 재현율, 정확률 및 검색 시간 카테고리에서 평균의 최대값과 최소값은 굵은 글씨체와 밑줄로 표시하여 구분하였다. 〈표 3〉의 각 평균값(예, FRBR의 표현형 질문의 재현율 평균값)은 12개의 질문을 두 개의 질문 유형 즉, 표현형(6개)과 구현형(6개)으로 각각 구분하고 28개 사례의 평균을 계산하여 측정하였다. 두 평균값의 차이를 나타내는 t값과 p값은 오른쪽 항목에 각각 기술하였다. 예를 들어서, FRBR과 OPAC의 표현형 질문의 재현율 평균값은 "0.92", "0.51"로 이 두 평균에 대한 t-검증 결과, t(p)값은 "28.14(0.00)"으로 나와 유의확률수준 0.05에서 통계적으로 유의미한 차이를 보였다.

〈표 3〉 t-검증 결과

	재현율 평균(표준편차)			정확률 평균(표준편차)			시간 평균(표준편차)		
	FRBR	OPAC	t(p)	FRBR	OPAC	t(p)	FRBR	OPAC	t(p)
표현형	0.92 (0.05)	0.51 (0.06)	28.14** (0.00)	0.99 (0.04)	0.97 (0.05)	1.66 (0.11)	83.55 (38.74)	82.96 (40.28)	0.090 (0.93)
구현형	0.92 (0.03)	0.75 (0.05)	20.68** (0.00)	0.96 (0.03)	0.99 (0.04)	3.55** (0.00)	105.59 (36.94)	83.76 (53.77)	2.32* (0.03)

**는 0.01 수준에서 유의미하고 *는 0.05 수준에서 유의미함.

(1) 연구 가설 1과 2: 재현율

"표현형과 구현형의 경우 모두 FRBR 기반 실험시스템이 OPAC 기반 시스템 보다 평균 재현율이 더 높게 나타날 것이다."의 연구 가설 1과 2를 검증하기 위해서 t-검증을 한 결과, 표현형과 구현형 질문에서 FRBR 기반 시스템이 OPAC 기반 시스템 보다 평균 재현율이 더 높게 나타났으며 이 차이를 분석한 결과 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 이 두 가설은 검증되었다(표 3 참조). 중다변량분석 결과 재현율인 경우 시스템 유형, 질문 유형은 물론 두 독립변인의 상호작용 요인도 종속변인에 모두 영향을 미치는 것으로 나타난다. 이 결과를 t-검증 결과와 함께 해석하면 시스템에 따라 재현율이 차이가 나며(예, FRBR 기반 시스템 경우 재현율이 높음), 질문 유형에 따라서도 차이가 난다(예, 구현형 질문의 경우 재현율이 높음). 더 나아가 이 독립 변인들의 상호 작용 효과도 나타나는데 FRBR 기반 시스템이 표현형 질문(또는 구현형 질문)과 결합될 때 재현율은 높게 나타난다(표 2와 표 3 참조).

(2) 연구 가설 3과 4: 정확률

"표현형 질문의 경우 FRBR 기반 실험시스템이 OPAC 기반 시스템 보다 평균 정확률이 더 높

게 나타날 것이다"는 가설 3은 FRBR의 평균 정확률(0.99)이 OPAC의 평균 정확률(0.97) 보다 높게 나타났으나 통계적으로 유의미한 결과를 보이지 못해서 검증되지 못했다. "구현형 질문의 경우 OPAC 기반 시스템이 FRBR 기반 실험시스템 보다 평균 정확률이 더 높게 나타날 것이다"는 가설 4는 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 검증되었다(표 2 참조). 중다변량분석 결과 시스템과 질문 유형의 상호 작용만이 정확률에 미치는 효과가 통계적으로 유의미한 결과를 보였다. 즉, FRBR 기반 시스템이 구현형 질문과 상호작용할 때 정확률은 낮아지는 것으로 나타난다(표 2와 표 3 참조).

(3) 연구 가설 5와 6: 검색 시간

"표현형 질문의 경우 FRBR 기반 실험시스템이 OPAC 기반 시스템 보다 평균 검색 시간(초)이 더 짧을 것이다"는 가설 5는 예측과는 달리 FRBR의 평균 검색 시간(83.55초)이 OPAC의 평균 검색 시간(82.96초) 보다 더 길게 나타나 검증되지 못했다(표 2 참조). 이러한 결과에 대한 여러 해석을 할 수 있는데 FRBR 시스템의 단점에서 지적된 것처럼 이용자가 익숙하지 않은 그룹화 기능에 적응하는데 어느 정도 시간이 걸린 것으로 사료된다. "구현형 질문의 경우 OPAC 기반 시스템이 FRBR 기반 실험시스템 보다 평균 검색 시간(초)이 더 짧을 것이다"는 가설 6은 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 검증되었다(표 2 참조). 중다변량분석 결과 시스템 유형과 질문 유형의 상호 작용만이 시간에 미치는 효과가 통계적으로 유의미한 것으로 나타난다. 즉, FRBR 기반 시스템이 구현형 질문과 상호작용할 때 탐색 시간이 길어지는 것으로 나타난다(표 2와 표 3 참조).

3. 분석 결과에 따른 시스템간 상호 비교

FRBR 기반 시스템의 장점으로는 많은 사람들이 한 연주가나 단체가 연주한 자료를 한 눈에 볼 수 있도록 하는 관련 자료 그룹핑 기능, 전거통제로 인한 높은 재현율, 다양한 접근점 제시, 데이터의 정확성 그리고 간략목록과 상세목록을 동시에 보여 주는 유용한 인터페이스를 선정하였다. 이에 반해서, OPAC 기반 시스템의 장점으로는 단순하고 익숙한 인터페이스로 학습이 용이한 점, FRBR 시스템에 비해 검색 건수가 적어 원하는 자료를 쉽게 찾을 수 있는 점 그리고 명확한 답을 도출하는데 유리하다는 점을 들었다.

FRBR 기반 시스템의 단점으로는 세부사항 확인 불편, 구현형 확인 어려움, 익숙하지 않은 그룹화 기능을 들었다. 즉, 비슷한 결과를 그룹핑 한다는 점에서는 장점이 있으나, 역시 세부사항에 있어서는 상세목록까지 확인을 거쳐야만 결과의 정확도를 확인할 수 있다는 점을 들었다. 또한 같은 자료라고 해도 출판사 등에 따라 다른 자료로 취급이 되는데 이를 한꺼번에 브라우징하지 못하고

일일이 봐야하는 것이 불편했다고 지적하였고 자료에 대한 지식이 부족한 사람의 경우 브라우징을 통해서 선택하는 경우가 많기 때문에 이를 고려한 시스템의 보완이 필요하다고 제안하였다. 익숙하지 않은 그룹화 기능을 단점으로 제시했으나 시스템에 일단 익숙해지면 좋은 기능으로 인식될 것으로 보았다. 한편 OPAC 기반 시스템에 대해서는 적합문헌 확인의 어려움, 검색결과의 연관성이 부족한 점, 검색결과를 한 눈에 알아보기 어려운 점, 낮은 재현율, 전거통제 미흡 등을 열거하였다. 즉, 검색결과가 단순히 나열되어 있어 정확한 자료를 찾기가 어렵고, 더욱이 검색 및 결과의 범위가 넓어졌을 경우 자료를 파악하기가 매우 힘들어진다고 하였다. 또한 똑같은 검색 키워드를 입력해도 FRBR 시스템에서 검색된 자료가 OPAC 시스템에서 검색되지 않는 점을 들었고, 대부분의 경우 OPAC 시스템이 FRBR 시스템 보다 검색결과가 만족스럽지 못하다고 지적하기도 하였다.

V. 결 론

본 연구에서는 FRBR 모형 기반 서지검색시스템의 검색 효율성을 평가하기 위해서 두 개의 실험 시스템을 구축하였다. 즉, 387건의 KORMARC로 코딩된 음악 자료를 데이터베이스로 활용하여 FRBR화 알고리즘을 적용한 FRBR 기반 시스템과 일반 OPAC 기반 시스템을 구성하였다. 이 두 시스템의 특성을 이론적으로 파악하여 6개의 연구 가설을 세우고 28명의 피조사자와 6개의 표현형 질문과 6개의 구현형 질문으로 구성된 총 12개의 검색 질문을 이용하여 이 두 시스템을 재현율, 정확률 및 검색 시간을 측정 기준으로 하여 비교, 분석하였다.

실험 결과를 가설 검증과 연계 시켜 이해하기 쉽게 기술하기 위해서 중다변량분석과 t-검증을 수행하였다. 먼저 검증된 4개의 가설들을 중심으로 설명하면, 재현율에서 FRBR 시스템이 질문의 유형을 가리지 않고 OPAC 시스템 보다 평균 재현율이 단연 앞서는 것으로 나타났다. 이는 FRBR 기반 시스템이 관련 자료를 유형별로 그룹핑하여 데이터베이스를 구성하고 저자와 서명을 전거통제했기 때문으로 생각된다. 정확률과 검색 시간에서 구현형 질문의 경우 OPAC 기반 시스템이 FRBR 기반 실험시스템 보다 평균 정확률이 조금 높았고 평균 검색 시간이 더 짧게 나타났다.

검증되지 못한 연구 가설 3은 예측한 대로 FRBR의 평균 정확률(0.99)이 OPAC의 평균 정확률(0.97) 보다 높게 나타났으나 유의확률수준 0.05에서 통계적으로 유의미한 결과를 보이지 못해서 검증되지 못했다. 연구 가설 5는 예측과 달리 표현형 질문에서 FRBR의 평균 검색 시간(83.55초)이 OPAC의 평균 검색 시간(82.96초) 보다 더 길게 나타나 검증되지 못했다. 이는 FRBR 시스템의 단점에서 지적된 것처럼 이용자가 익숙하지 않은 그룹화 기능에 적응하는데 어느 정도 시간이 걸린 것으로 생각할 수 있으며, FRBR 기반 시스템이 보편화되는 검색 환경에서는 이와는 다른 결과를 예상해 볼 수 있다.

결론적으로 FRBR 기반 시스템은 재현율에서 질문의 유형을 가리지 않고 OPAC 시스템 보다 높은 성능을 보이고 있고, OPAC 시스템은 구현형 질문인 경우 정확률과 시간에서 FRBR 시스템 보다 다소 효율적인 것으로 나타났다. 이러한 실험 결과를 현 검색 환경에 적용시키기 위해서는 FRBR 기반 시스템이 서지적 연관성이 높은 분야에 활용될 수 있으며 이러한 분야에 적용될 때 OPAC 시스템을 대체하기 보다는 이를 보완하면서 FRBR 시스템의 특성을 살려 질의 유형이 표현형인 경우 또는 높은 재현율을 요하는 경우 특히 효율적으로 적용될 수 있을 것이다.

〈참고문헌은 각주로 대신함〉

〈부록 1〉 FRBR화 알고리즘 적용

수집한 387건의 저작 그룹을 본 연구에서 제안한 알고리즘을 이용하여 표현형으로 구분한 다음 각 표현형별로 구현형들을 그룹핑하였다. 피아노 협주곡의 경우 연주단체, 연주자 및 지휘자 중 연주단체와 연주자를 결합하여 하나의 표현형으로 확인하였고, 교향곡의 경우 연주단체, 연주자 및 지휘자 중 연주단체와 지휘자를 결합하여 하나의 표현형으로 확인하였고, 피아노 소나타인 경우는 연주자 정보를 이용하여 표현을 기술하였다. 알고리즘을 실제 387건의 음악자료 레코드에 적용해 본 결과 악보는 평균 성공률이 66.5%, 공연/연주는 평균 성공률이 79%로 나와 전체 평균 성공률은 72.6%이다.

가. 표현형 확인

(1) 악보

1) Leader/06을 이용하여 악보인지 확인하였다. Leader/06 정보가 없는 경우나 단행본 KORMARC 포맷을 사용하는 경우에는 300ab를 분석하여 "music", "parts", "악보" 등의 값을 갖는 레코드를 선택하였다.
2) 악보임을 확인한 후에 이를 다시 008/35-37을 사용하여 언어별로 구분하였다. 008/35-37 정보가 없는 경우에는 041$a를 이용하여 언어를 확인하여 악보-독일어판, 악보-한국어판 등으로 구분하였다.
3) 확인 작업
   (가) 모차르트의 피아노 협주곡 20번: 53건의 자료 중 악보로 확인된 자료는 총 17건이었고 이 중 두 개의 자료는 녹음자료가 잘못 코딩되어 있어서 제외시키고 15개 자료가 악보 자료로 확인되었다. Leader/06를 통해서 확인된 자료는 7건이었고, 300$a을 이용하여 검색한 결과 총 7건 중 일차 검색에서 누락된 3건이 새롭게 검색되었고 마지막으로 300$b를 이용한 결과 7건이 검색되었고 이 중 5건의 새로운 자료가 확인되었다. 악보로 확인된 15건 중 언어가 확인이 안 된 3개의 자료를 제외하고 총 12건의 자료를 최종적으로 확인하여 80%의 성공률을 보였다. 구체적으로 Leader/35-37을 통해서 확인된 자료는 11건 이었고 나머지 한 건은 041태그를 이용하여 언어 코드를 확인하였다. 이들 자료는 "악보-독일어", "악보-일본어", "악보-영어", "악보-영어/독일어", "악보-한국어"의 다섯 가지 표현형으로 구분되었다.
   (나) 모차르트의 피아노 소나타 제11번: 25건의 자료 중 악보로 확인된 자료는 총 5건이었다. 300$a를 통해 확인한 자료는 5건이었고, Leader/06과 300$b를 통해서는 자료를 확인할 수 없었다. 악보로 확인된 5건 중 언어가 확인된 자료는 4건으로 80%의 성공률을 보였다. 모두 Leader/35-37을 통해서 확인하였다. 이 자료는 모두 "악보-한국어"의 하나의 표현형

으로 구분되었다.

(다) 모차르트의 교향곡 제41번: Leader/06을 통해 6건의 자료를 확인하였다. 그리고 300$a를 통해 확인한 자료 9건 중 1차 검색에서 누락된 6건이 새롭게 확인되었고, 300$b를 통해 확인한 17건 중 16건을 새로 확인하여 총 28건을 확인하였다. 악보로 확인된 28건 중 언어가 확인된 자료는 22건으로 약 79%의 성공률을 보였다. Leader/35-37을을 사용하여 언어를 확인하여 악보-영어, 악보-한국어, 악보-일본어로 구분되었다.

(라) 베토벤의 교향곡 6번: 54건의 자료 중 악보로 확인된 자료는 총 3건이었다. Leader/06를 통해서 확인된 자료는 1건이었고, 300$a을 이용하여 검색한 결과 1건이 새롭게 검색되었고 300$b를 이용한 결과 1건이 새로 검색되었다. 악보로 확인된 3건 중 언어가 확인이 안 된 1개의 자료를 제외하고 총 2건의 자료를 최종적으로 확인하여 67%의 성공률을 보였다. 2건 모두다 Leader/35-37을 통해서 언어 코드를 확인하였다. 이들 자료는 "악보-한국어", "악보-영어"의 두 가지 표현형으로 구분되었다.

(마) 베토벤의 피아노 협주곡 제5번: Leader /06을 통해 9건의 자료를 확인하였다. 300$a를 통해 확인한 자료 8건 중 1차 검색에서 누락된 3건이 새롭게 검색되었고, 마지막으로 300$b를 통해 확인한 자료 13건 중 10건의 새로운 자료를 확인하여 총 22건을 확인하였다. 악보로 확인된 22건 중 언어가 확인된 자료는 15건으로 약 68%의 성공률을 보였다. Leader/35-37을을 사용하여 언어를 확인. 악보-독일어, 악보-영어, 악보-한국어로 구분되었다.

(사) 베토벤의 피아노 소나타 제23번: Leader/06을 통해 4건의 악보자료를 확인하였다. 300$a를 통해 1건의 자료를 검색하였지만 중복자료로 확인되었다. 악보로 확인된 4건 중 언어가 확인된 자료는 1건으로 25%의 낮은 성공률을 보였다. Leader /35-37을 사용하여 언어를 확인, 악보-영어의 표현형으로 구분되었다.

(2) 공연/연주

1) 'Leader/06'을 활용하여 그 값이 평면영사자료(g), 녹음자료(음악)(j)이면 공연/연주 자료로 확인하였다. Leader 정보가 없는 경우나 단행본 KORMARC 포맷을 사용하는 경우에는 245$h 또는 300$a$b를 분석하여 "녹음자료", "sound recording", "Audio CD", "비디오", "sound disc", "음반", "stereo", "digital" 등의 값을 갖는 레코드를 선택하였다.

2) 공연/연주로 확인되면 700adem 및 710abem 또는 511$a 태그를 이용하여 연주자, 지휘자, 연주 단체를 확인한다. 연주자와 지휘자를 확인하기 위해서 700adem를 확인해 보았는데 공연/연주에 해당되는 총 89건(모차르트: 38건, 베토벤: 51건) 중에서 24건만이 700$e를 사용하여 지휘자의 역할어(예, conductor, cond., 및 지휘)를 사용하고 있었다. 연주자와 배역진 주기를 기술하는 511 태그에서도 정보를 명확하게 확인하기가 어려운데, 이는 511

태그에서는 연주자와 배역진을 세부적으로 구분하는 역할어를 별도로 사용하고 있지 않고 있으며 '$a' 식별기호를 사용하여 전체적으로 표현하고 있기 때문이다(예, $a비엔나 필하모닉 관현악단, 칼뵘 지휘). 따라서 역할어가 생략된 경우에는 차선책으로 표현형을 확인하기 위해서 KORMARC 데이터베이스를 분석하여 "Concentus Hungaricus(연주단체)-Jenoe Jando(연주자)-Andras Ligeti(지휘자)"와 같은 표현형 목록을 구성한 다음 700$a, 710$a 및 511$a의 데이터를 차례로 검색하여 해당 연주단체와 지휘자가 검색되면 동일한 표현형으로 구분하였다. 한 레코드에서 연주자, 지휘자 및 연주 단체를 모두 기술하는 레코드가 많지 않아서 본 연구에서는 피아노 협주곡인 경우는 연주 단체(오케스트라 명칭)와 연주자(피아노)만 일치하여도 동일한 표현형으로 구분하고, 교향곡인 경우에는 연주 단체(오케스트라 명칭)와 지휘자만 일치하여도 동일한 표현형으로 분류하였다.

3) 확인 작업

(가) 모차르트의 피아노 협주곡 20번: 53건의 자료 중 악보로 확인된 15건 자료를 제외한 38건의 자료가 공연/연주 자료로 확인되었다. 구체적으로 38건의 자료 중 일차적으로 Leader/06을 이용하여 33건을 확인하였고 245$h로 검색한 19건 중 4건이 새로운 자료였다. 300$a을 이용하여 검색한 24건 중 새로운 자료는 하나도 없었고 마지막으로 300$b를 이용하여 검색한 25건 중 한 건이 새로운 자료였다. 공연/연주로 확인된 자료 38건의 자료 중 24건만 연주단체 및 연주자 정보를 확인할 수 있어서 약 65%의 성공률을 나타냈다. 확인된 24건은 "Concentus Hungaricus(연주단체)-Jenoe Jando(연주자)", "Philharmonia Orchestra(연주단체)-Vladimir Ashkenazy(연주자)" 등의 19개의 각기 다른 표현형으로 기술되었다.

(나) 모차르트의 피아노 소나타 제11번: 25건의 자료 중 악보로 확인된 5건을 제외한 20건의 자료가 연주/공연으로 확인되었다. Leader/06을 통해 20건을 모두 확인하였다. 245$h를 통해 13건, 300$a를 통해 13건, 300$b를 통해 17건을 확인하였으나 새로운 자료는 없었다. 이와 같이 공연/연주로 확인된 20건의 자료 중 12건에서만 연주자 정보를 확인할 수 있어 60%의 성공률을 나타냈다. 확인된 12건은 "Backhaus, Wilhelm(연주자)" 등의 10개의 표현형으로 기술되었다.

(다) 모차르트의 교향곡 제41번: Leader/06을 통해 58건의 자료를 확인하였다. 245$h를 통해 42건을 확인하였고, 이 가운데 첫 번째 검색에서 누락된 1건의 자료를 확인하였다. 300$a를 통해 34건을 확인하였고, 이 가운데 2건을 새로 확인하였다. 300$b를 통해 41건을 확인하였으나 새로 검색된 자료는 없었다. 위의 검색결과를 통해 총 61건의 연주/공연자료를 확인하였다. 확인된 61건의 자료는 연주/공연: Berlin Philharmonic Orchestra (연주단체)-Herbert Von Karajan(지휘자), 연주/공연: London Philharmonic Orchestra (연주단체)-Alfred Scholtz (지휘자) 등의 30개의 표현형으로 기술되었다. 성공률은 100%를 나타냈다.

(라) 베토벤 교향곡 제6번: 54건의 자료 중 악보로 확인된 3건을 제외한 51건의 자료가 공연/연주 자료로 확인되었다. 구체적으로 51건의 자료 중 일차적으로 Leader/06을 이용하여 48건을 확인하였고 245$h로 검색한 39건 중 2건이 새로운 자료였다. 300$a를 통해 검색한 35건의 자료 중 1건이 새로운 자료로 확인되었고, 300$b를 통해 검색한 34건의 자료 중에서도 새로운 자료는 없었다. 이와 같이 공연/연주로 확인된 51건의 자료 중 42건만 연주단체와 지휘자 정보를 확인할 수 있어서 약 82%의 성공률을 나타냈다. 확인된 42건은 "Detroit Symphony Orchestra(연주단체)-Paul Oaray(지휘자)", "Gewandhausorchester Leipzig (연주단체)-Kurt Masur(지휘자)" 등의 22개의 각기 다른 표현형으로 기술되었다.

(마) 베토벤의 피아노 협주곡 제5번: 총 98건의 자료 중 악보자료 22건을 제외한 나머지 66건의 자료는 Leader/06을 통해 모두 연주공연자료임을 확인하였다. 공연/연주로 확인된 66건의 자료 중 54건에서 연주단체 및 연주자 정보를 확인할 수 있어 약 82%의 성공률을 나타냈다. 확인된 54건의 자료는 "Vienna Philharmonic Orchestra(연주단체)-Maurizio Pollini(연주자)", "Chicago Symphony Orchestra(연주단체)-Alfred Brendel(연주자)" 등의 38개의 표현형으로 기술되었다.

(사) 베토벤의 피아노 소나타 제23번: Leader/06을 통해 55건의 연주/공연 자료를 확인하였다. 245$h를 통해 확인한 자료 49건 중 4건의 새로운 자료를 확인하였다. 그리고 300$a를 통해 확인된 자료 50건 중 새로 1건의 자료를 확인하여 전체 64건의 자료 가운데 악보 4건을 제외한 60건의 자료를 모두 연주/공연자료로 확인하였다. 60건의 연주/공연 자료 중 51건에서 연주자 정보를 확인할 수 있어 85%의 성공률이 확인된 51건의 자료는 26개의 표현형으로 구분되었다.

나. 구현형 확인

각 표현형에 속한 구현형들을 확인하여 배열한다.

1) 악보인 경우에는 260bc$g 또는 300$abcef$g을 확인하여 구현형간의 차이를 확인하고 그룹핑하였다.
2) 공연/연주인 경우에는 260bc$g 또는 300$abcef$g을 확인하여 구현형간의 차이를 확인하고 그룹핑하였다.
3) 본 실험에서는 표본 선정 절차에서 중복데이터를 제거하였기 때문에 본 과정에서 중복데이터 처리작업은 생략하였다.

〈부록 2〉 설문지 질문(12개)

Ⅰ. 표현형 질문(6개)

1. Beethoven(베토벤)의 piano concerto no.5(피아노 협주곡 제5번) 중에 Brendel(Alfred Brendel)과 함께 연주한 연주단체(오케스트라)들을 모두 찾습니다.

2. Beethoven(베토벤)의 Symphony no.6(교향곡 제6번)을 Berlin Philharmonic 오케스트라(연주단체)와 공연/연주한 지휘자들을 모두 찾습니다.

3. Mozart(모짜르트)의 Symphony no.41(교향곡 제41번)의 공연/연주 중에서 bohm이 지휘한 연주단체(오케스트라)들을 모두 찾습니다.

4. Beethoven(베토벤)과 Mozart(모짜르트)가 작곡한 Piano concerto(피아노 협주곡) 중 한국어로 된 모든 악보들을 찾습니다(베토벤과 모짜르트 자료 모두 포함).

5. Beethoven(베토벤)과 Mozart(모짜르트)의 Piano Concerto(피아노 협주곡)을 Ashkenazy(Vladimir Ashkenazy)가 연주한 전체 공연/연주 자료를 찾습니다(베토벤과 모짜르트 자료 모두 포함).

6. Karajan(Herbert von Karajan)이 지휘하고, Vienna(또는 Wiener) Philharmonic 오케스트라(연주단체)가 연주한 Mozart(모짜르트)의 Symphony no.41 Jupiter [교향곡 제41번 쥬피터(또는 주피터)]의 전체 공연/연주 자료를 찾습니다.

Ⅱ. 구현형 질문(6개)

7. EMI사에서 발간한 Beethoven(베토벤)의 Piano Sonata(피아노 소나타)의 모든 sound disc(들)을 찾습니다.

8. "Edwin F. Kalmus"사에서 발간한 Mozart(모짜르트)의 교향곡 제41번(symphony no.41)에 대한 모든 악보(들)을 찾습니다.

9. Arrau(Claudio Arrau)가 연주한 Beethoven(베토벤)의 Piano Concerto no.5(피아노 협주곡 제5번)을 Philips사에서 발행한 모든 disc(video 또는 sound 또는 sound compact) 자료를 찾습니다.

10. 동경 음악지우사(音樂之友社)에서 1951년 발행한 Mozart(모짜르트)의 표제명이 "Symphony no.41 K.V. 551 C-major = モ-シァルト交響曲 第査十一番 ジュピタ- 八長調: jupiter ; 牧定忠 解說"인 악보를 찾습니다(1건).

11. Kempff(Wilhelm Kempff)가 연주한 Beethoven의 표제명이 "Piano sonata No.8, No.14, No.23"인 Windy Sound Record에서 발간한 sound disc 자료를 찾습니다(1건).

12. Furtwangler(Wilhelm Furtwangler)가 지휘하고 Vienna Philharmonic Orchestra가 연주한 Beethoven(베토벤)의 표제명이 "Symphony: No.1~No.9"인 EMI에서 1988년 발행한 LP판을 찾습니다(1건).