

KORMARC TO BIBFRAME 변환을 위한 매핑시 고려사항에 관한 연구*

A Study on the Considerations in Mapping between
KORMARC to BIBFRAME

이 미 화(Mihwa Lee)**, 박 진 호(Jin Ho Park)***
이 종 육(Jongwook Lee)****, 송 민 선(Min Sun Song)*****

〈 목 차 〉

I. 서 론	IV. KORMARC TO BIBFRAME 변환시 고려사항
II. MARC TO BIBFRAME 변환	V. 결 론
III. MARC21 TO BIBFRAME 변환 매핑표 분석	

요 약 : 본 연구에서는 MARC21 TO BIBFRAME의 변환을 바탕으로 KORMARC TO BIBFRAME의 변환시 고려할 사항을 제안하고자 문헌연구 및 MARC21 TO BIBFRAME 매핑표의 표시기호, 식별기호 등으로 다층적 분석을 수행하였다. MARC21 TO BIBFRAME 변환의 특징은 중복데이터를 생성하지 않는 것, 미변환이 상당히 포함되어 있다는 것, 어휘 인코딩스킴의 사용을 바탕으로 한다는 점이 있으며 이를 바탕으로 국내에서 KORMARC TO BIBFRAME 변환을 위해 매핑표 개발시 고려사항으로 3가지를 제안하였다. 첫째, KORMARC용 어휘 인코딩스킴 개발, 둘째, KORMARC에만 있는 표시기호 및 식별기호 변환, 셋째, MARC21에서 미변환(nac, ignore) 요소 중에서 KORMARC에서 변환할 사항이다. 이를 통해 본 연구는 KORMARC 및 BIBFRAME 상호간의 매핑표 개발 및 변환기 개발에 기여할 수 있을 것이다.

주제어 : 서지프레임, 매핑표, 변환기, 통합서지용 KORMARC

ABSTRACT : This study aims to propose considerations for the conversion from KORMARC to BIBFRAME based on the conversion from MARC21 to BIBFRAME. The literature review and analysis of the MARC21 to BIBFRAME mapping table from multiple layers by field and subfield of MARC were conducted. The characteristics of the MARC21 to BIBFRAME conversion included not generating redundant data, the significant presence of unconverted elements, and the use of a vocabulary encoding scheme. Based on these analyses, three considerations were proposed for the development of a mapping table for the KORMARC to BIBFRAME conversion in Korea. First, the development of a vocabulary encoding scheme for KORMARC, second, the conversion of field and subfield unique to KORMARC, and third, those that need modification for KORMARC among unconverted elements from MARC21. Through this, the study might contribute to the development of a mapping table and converter for KORMARC.

KEYWORDS : BIBFRAME, Mapping Table, Converter, KORMARC, MARC21

* 본 연구는 국립중앙도서관의 『국가서지에 대한 차세대 서지 구조 적용방안 연구(2024)』의 일부를 수정 보완하여 학술논문으로 발전시킨 것임.

** 국립공주대학교 문헌정보교육과 교수(leemh@kongju.ac.kr / ISNI 0000 0004 6431 3495) (제1저자)

*** 한성대학교 지식정보문화트랙 교수(jhp@hansung.ac.kr / ISNI 0000 0004 7641 0372) (공동저자)

**** 경북대학교 문헌정보학과 부교수(jongwook@knu.ac.kr / ISNI 0000 0004 6830 6145) (공동저자)

***** 대진대학교 문헌정보학과 조교수(songsery@daejin.ac.kr / ISNI 0000 0004 9246 0812) (교신저자)

• 논문접수: 2025년 2월 19일 • 최초심사: 2025년 3월 4일 • 게재확정: 2025년 3월 11일

• 한국도서관·정보학회지, 56(1), 307-326, 2025. <http://dx.doi.org/10.16981/kliss.56.1.202503.307>

* Copyright © 2025 Korean Library and Information Science Society

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits use, distribution and reproduction in any medium, provided that the article is properly cited. the use is non-commercial and no modifications or adaptations are made.

I. 서 론

서지데이터를 링크드데이터로 변환하기 위한 표준으로 BIBFRAME이 거론되고 있고, 국내외 여러 국립도서관을 중심으로 MARC로 구축된 서지데이터를 BIBFRAME으로 변환하는 프로젝트가 진행되고 있다. 이러한 프로젝트의 핵심은 MARC로 구축한 레거시 데이터를 BIBFRAME 어휘를 이용한 RDF 형태의 서술문으로 변환하는 것이다. 이는 기존 MARC 레코드 기술 방식과는 완전히 다른 RDF 문장으로 변환되어야 하므로 가능한 데이터의 손실 없이 변환이 이루어져야 한다.

이러한 데이터 변환은 상호운용성 측면에서 다양한 방법이 가능하겠지만 주로 매핑표를 개발하고 이를 이용하여 변환기를 개발하는 형태이므로 MARC에서 BIBFRAME으로 매핑표가 우선 개발되어야 한다. BIBFRAME 프로젝트를 시작한 미의회도서관에서는 프로젝트 초기에 MARC21에서 BIBFRAME으로의 매핑표를 작성하였으며, BIBFRAME 2.0이 개발되면서 2017년 매핑표를 개정하여 매핑표의 현재 버전은 2.8이다. 매핑표는 한번의 개발로 끝나는 것이 아니라 데이터의 무결성있는 변환을 위해 지속적인 수정이 요구된다. 매핑표의 이러한 지속적인 변경은 MARC21 데이터의 누락없는 변환이 상당히 까다롭고 복잡한 작업이라는 것을 보여준다.

그러나, 미의회도서관에서 개발한 매핑표에서 MARC21의 일부 표시기호 및 식별요소의 매핑이 누락되었고, 미의회도서관이 작성한 BIBFRAME 매핑표가 전세계적인 BIBFRAME 매핑의 기준이 될 수 있는가에 대한 의문이 제기되었다. 이러한 상황에서 국내에서도 BIBFRAME의 변환이 필요한 만큼 기준의 MARC21 TO BIBFRAME 매핑을 분석하여 KORMARC를 BIBFRAME으로 변환하기 위한 매핑표의 개발이 요구되는 바이다.

이에 본 연구에서는 MARC21 TO BIBFRAME의 변환을 바탕으로 KORMARC TO BIBFRAME 변환시 고려할 사항을 제안하고자 문헌연구를 비롯해 MARC21 TO BIBFRAME 매핑표의 표시기호별, 식별기호별 등으로 다층적 분석을 수행하였다.

MARC에서 BIBFRAME 매핑표 및 변환과 관련된 선행연구로는 알버타 대학교도서관의 변환기와 Casalini Share 변환기를 분석한 연구(Bigelow et al., 2018), Hong Lou Meng 관련 문헌을 대상으로 MARC21 서지레코드에서 BIBFRAME으로 변환을 테스트 연구(Fu & Ke, 2021), LC, Share-VDE, LD4P2에서 MARC를 BIBFRAME으로 변환한 작업에 관한 연구(Samples & Bigelow, 2020), 일리노이 대학교 도서관에서 300,000개의 전자책 MARC를 BIBFRAME으로 변환한 후 이를 색인화한 연구(Jin et al., 2016) 등이 있다. 국내에서도 이성숙, 이지원(2018)은 MARC21과 BIBFRAME과의 매핑을 바탕으로 저작, 인스턴스 별로 매핑 사례를 제시하였고, 국립중앙도서관(2022; 2024)에서도 BIBFRAME 연구를 진행하면서 매핑연구를 함께 진행하였다. 이러한 선행연구와 비교할 때 본 연구는 MARC21 TO BIBFRAME 변환의 특징을 분석하고 이를 바탕으로 국내에서 변환시 고려할 사항을 거시적인 측면에서 제시하였으므로 국내 BIBFRAME

구축을 위한 실질적인 매핑표 개발에 기여할 수 있을 것이다.

II. MARC TO BIBFRAME 변환

1. 변환의 필요성

최근 정보환경이 데이터 개방과 공유 중심으로 변화하여 도서관 외부 데이터가 시맨틱웹을 위해 링크드데이터로 변환되고 있으므로 도서관의 서지데이터도 MARC에서 RDF 기반의 링크드데이터로 구축되어야 한다. 도서관에서 링크드데이터 구축을 위한 표준으로 BIBFRAME 사용이 확장되고 있다. BIBFRAME을 통해 레코드 기반의 MARC 구조에서 RDF 기반의 서술문으로 변환이 이루어지게 되면 시맨틱웹의 기반이 마련될 수 있다.

BIBFRAME은 개념모델과 MARC를 대체하는 인코딩포맷의 온톨로지라는 2가지 특징을 갖는다. 먼저 개념모델로 BIBFRAME은 서지 자원의 구조와 개체 상호 간의 관계를 기술하는 개체로 저작, 인스턴스, 개별자료, 에이전트, 주제, 이벤트 등을 갖는다. 또한, 자원 기술을 위한 온톨로지 어휘로 BIBFRAME 어휘는 RDF 모델링 방식을 활용하여 클래스 속성(클래스 간의 관계 포함)으로 서지자원을 기술하는 서술문을 작성하도록 한다(Joudrey & Taylor, 2018, 281-285). BIBFRAME 어휘는 기본 어휘와 미의회도서관이 BIBFRAME을 확장한 어휘를 같이 사용할 수 있다.

도서관을 둘러싼 제반 환경이 여전히 MARC를 유지하는 상태에서 MARC에서 BIBFRAME의 전면적인 변환은 상당한 시간이 지난 후 가능할 것이고, 변환을 위한 매핑표 및 프로그램은 프로젝트를 통해 지속적으로 개정되면서 완전하게 작성될 것이다.

2. 매핑표와 변환기의 관계

MARC에서 BIBFRAME으로 변환 처리를 위해서는 변환을 위한 매핑표가 필요하다. 미의회 도서관(<http://www.loc.gov/bibframe/mtbf/>)에서 매핑표와 변환기를 제공하고 있다. 미의회 도서관에서는 매핑표를 바탕으로 변환기를 개발하였고, 변환기는 변환의 내용을 확인할 수 있도록 뷰어(viewer)도 함께 제공하고 있다.

미의회도서관의 MARC21 TO BIBFRAME 변환 명세와 변환기는 St. Pierre & LaPlant(1998)의 권고를 준용한다. 크로스워크는 소스 메타데이터 표준(source metadata standard)의 요소 내용에 적용된 일련의 변환으로, 그 결과 대상 메타데이터 표준의 유사한 요소에 적절하게 수정된 내용으로 저장된다. 크로스워크는 의미론적 매핑(semantic mapping)과 메타데이터 변환 명세

(metadata conversion specification)으로 구성된다. 메타데이터 변환 명세는 소스 메타데이터 표준에 부합하는 메타데이터를 타겟 메타데이터 표준(target metadata standard)에 부합하는 데이터로 변환하는 데 필요한 명세를 포함한다(Xu et al., 2017, 228, 재인용).

매핑표와 변환기는 매우 밀접한 관계를 갖고 있어 미의회도서관에서는 매핑표를 기반으로 변환기를 개발하고 샘플을 변환한 후 검토의견을 바탕으로 매핑표를 지속적으로 수정하고 있다. BIBFRAME 2.0 개발 직후에서 나온 의견으로 041 언어 변환 매핑 문제, MARC 86X에서 변환 수정, xx9, x9x, 9xx의 로컬 필드의 매핑 제시 요구, 미변환 필드로 031, 034, 1xx, 7xx \$u 매핑 요구, 총서필드(490 및 8xx)의 매핑 요구(Xu et al., 2017, 239-244) 등이 있었으며, 이러한 의견은 다시 BIBFRAME 매핑 및 변환기 개발에 순기능으로 작용하게 되었다. 따라서 매핑표와 변환기는 계속적으로 갱신 및 수정이 이루어지는 관계를 갖는다.

3. 매핑표를 바탕으로 한 변환기

MARC를 BIBFRAME으로 변환하기 위해 매핑표를 기반으로 변환기가 개발되었고 이 변환기를 통해 현재 MARC에서 BIBFRAME으로 변환이 가능하다. 미의회도서관에서는 MARC - MARC XML - BIBFRAME으로 변환한다. 변환 프로그램은 GitHub를 통해 이용할 수 있으며 뷰어도 함께 제공하고 있다.

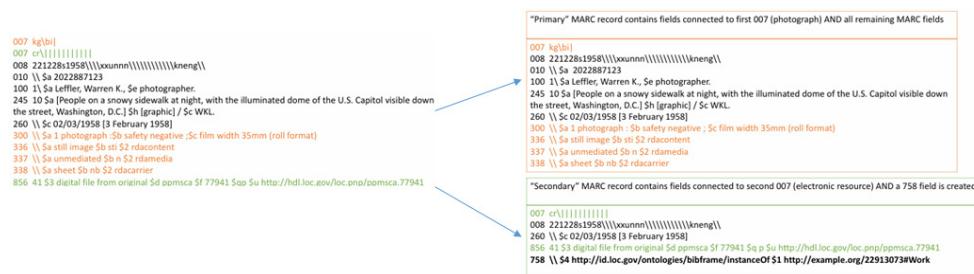
변환프로그램은 BIBFRAME 2.0 및 MADS/RDF 온톨로지를 사용하여 MARCXML 레코드를 RDF/XML로 변환하는 XSLT 1.0 애플리케이션이 포함된다. 입력은 MARCXML 레코드 또는 컬렉션이며, 출력은 RDF/XML 구문에서 RDF 트리플 집합으로 데이터를 표현하는 XML 문서이다.

MARC에서 BIBFRAME 어휘로 매핑을 위한 과정은 MARC에 URI 추가하는 배치처리와 RDF 서술문으로 자원 기술을 변환하는 것이다. 먼저, MARC에 URI를 추가하는 배치 처리하는 과정은 MARC의 데이터를 id.loc.gov와 같은 링크드데이터 서비스로 매칭하여 BIBFRAME으로 변환하기 전에 URI를 삽입하게 된다. 이후 BIBFRAME 변환기는 이러한 URI를 BIBFRAME RDF의 올바른 속성으로 처리할 수 있다. 이 과정은 기존 MARC 데이터를 BIBFRAME으로 변환하는 데 필요하다. 만약 LC ID 서비스에서 URI를 일치시킬 수 없는 개체나 필드 요소가 있다면, 변환기는 기본적으로 설정된 baseURI 매개변수의 접두사, 001 필드의 MARC 레코드 ID, 해시 URI를 결합하여 새로운 URI를 생성한다. 해시 URI는 요소 클래스, 필드, MARC 레코드에서 필드 위치(예: <http://example.org/ocn782927281#Topic650-42>)로 구성된다(Xu et al., 2017, 229).

다음으로 RDF 서술문으로 변환의 기능은 MARC 레코드의 필드를 처리하여 생성되는 bf:Work와 bf:Instance를 구축하는 프로그램이다. 특히, 옵션 전처리 단계는 변환기의 핵심 내용으로 볼 수 있다. 개별 MARC 레코드를 토대로 원본 또는 소스 MARC 레코드에서 동일한 저작의 여러

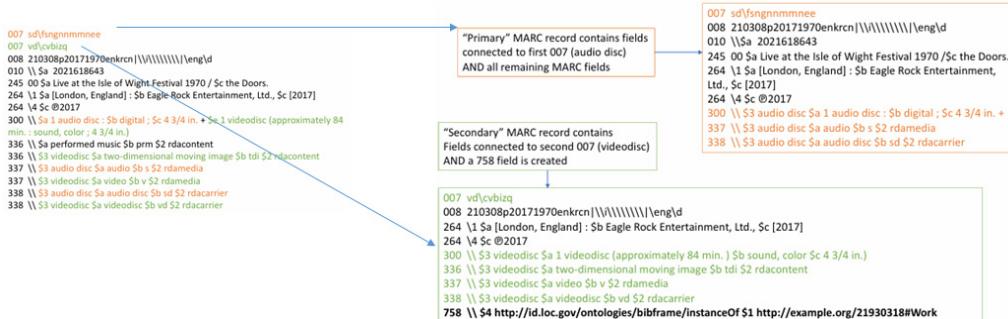
KORMARC TO BIBFRAME 변환을 위한 매핑시 고려사항에 관한 연구

인스턴스를 나타내는 추가 MARC 레코드를 갖는 다수의 MARC 레코드로 나누는 작업이다. 하나의 MARC/XML 레코드를 입력받아 하나 이상의 MARC 레코드로 구성된 marc:collection을 출력한다. 즉 007, 3XX, 856 필드에 기반하여 다수의 MARC 레코드로 분리한다. 이는 하나의 MARC 레코드에 목록된 다수의 자원 문제를 해결하기 위한 것으로 다수의 인스턴스를 생성하도록 한다. 〈그림 1〉과 같이 856 필드와 007 필드를 바탕으로 레코드가 분리되며, 〈그림 2〉와 같이 007 필드와 300 필드를 바탕으로 딸림자료가 분리된다(Ford & Williamschen, 2023).



〈그림 1〉 856 필드 및 007 필드를 이용해 다수의 인스턴스 생성

※ 출처: Ford & Williamschen(2023)



〈그림 2〉 007 필드와 300 필드(딸림자료)를 이용해 다수의 인스턴스 생성

※출처: Ford & Williamschen(2023)

III. MARC21 TO BIBFRAME 변환 매핑표 분석

1. 매팅표 개요

MARC21 TO BIBFRAME 2.0 변환 명세는 2017년 3월에 발행되었으며, MARC 관점에서 작성되었고, BIBFRAME으로 변환되지 않더라도 MARC의 각 요소가 고려되었다. 매핑표는

MARC를 중심으로 동일한 기능의 태그를 그룹핑하여 제시되어 있다(〈표 1〉 참조).

〈표 1〉 MARC21 TO BIBFRAME 매팅표 개요

매팅표(버전)	설명	참고
Leader (09/14/2022)	리더	
Fields 001-007 - Control, physical description (11/22/2023)	형태기술	
Fields 006, 008 - Fixed fields (11/20/2024)	고정장	
Fields 010-048 - Identifiers, etc. (11/20/2024)	식별자	
Field 048 - Instrument codes (10/21/2021)	기호	
Fields 050-088 - Class/call numbers, etc. (11/20/2024)	번호	
Fields 1XX, 7XX, 8XX - Names (11/20/2024)	이름	Process 0 - 이름/표제 Process 1 이름
Fields 200-24X, except 240 - Titles (11/20/2024)	표제	Process 0, Process 1
Fields 240, X30, etc. - Uniform titles (11/20/2024)	통일 표제	
Fields 250-270 - Edition, imprint, etc. (08/22/2024)	판	
Fields 3XX - Physical description, etc. (11/20/2024)	형태사항	
Fields 490, 510 - Links (11/20/2024)	링크	
Fields 5XX - Notes (11/20/2024)	주기	
Fields 600-662 - Subjects (11/20/2024)	주제	
Fields 720, 740-758 - Added entries (11/20/2024)	부출	
Fields 760-788 - Linking entries (11/20/2024)	연관저록	
Fields 841-887 - Holdings, location, alternate graphics, etc. (10/21/2021)	소장 등	

*기타 매팅: rdfType (08/09/2022), Numeric Subfield Conversions (11/20/2024), Pre-process notes (11/22/2023), Process notes (11/20/2024), Subject Thesaurus indicator (10/21/2021)

※ 출처: Library of Congress(2024)

리더의 경우 〈그림 3〉과 같이 변환 1, 변환 2, 주기가 있으며 내용은 W - adminMetadata - AdminMetadata - status - Status와 같은 BIBFRAME 어휘를 이용하고 <http://id.loc.gov/vocabulary/mstatus/c>와 같이 어휘 인코딩스킴을 사용한다.

A	B	C	D
1 LEADER - v2.1, 09/14/2022	Conversion 1	Conversion 2	Notes / LC stats
3 Character Positions			
4 00-04 - Logical record length	Ignore		
5 05 - Record status	W - adminMetadata - AdminMetadata - status - Status		
6 a - Increase in encoding level	## - http://id.loc.gov/vocabulary/mstatus/c ; add rdfs:label "changed"		
7 c - Corrected or revised	## - http://id.loc.gov/vocabulary/mstatus/c ; add rdfs:label "changed"		
8 d - Deleted	## - http://id.loc.gov/vocabulary/mstatus/d ; add rdfs:label "deleted or deprecated"		
9 n - New	## - http://id.loc.gov/vocabulary/mstatus/n ; add rdfs:label "new"		
10 p - Increase in encoding level from prepublication	## - http://id.loc.gov/vocabulary/mstatus/p ; add rdfs:label "prepublication"		
12 06 - Type of record	W - rdf:type (see ConvSpec-rdftype)		
13 a - Language material	see ConvSpec-rdftype for mapping with L/07		If MARC 336 \$a does not exist, W - content - Content - http://id.loc.gov/vocabulary/contentType/s/tx ; add rdfs:label "text"
14 b - Archival and manuscripts control	Ignore		
15 c - Notated music	see ConvSpec-rdftype for mapping with L/07		If MARC 336 \$a does not exist, W - content - Content - http://id.loc.gov/vocabulary/contentType/s/ntm ; add rdfs:label "notated music"

〈그림 3〉 리더 매팅표 일부

※ 출처: Library of Congress(2024)

‘001-007 - Control, physical description’, ‘006, 008 - Fixed fields’, ‘010-048 - Identifiers, etc.’, ‘048 - Instrument codes’, ‘050-088 - Class/call numbers, etc.’는 코드화된 정보가 많아 상당히 복잡하고, 특히, 007 및 008의 위치별로 해당하는 BIBFRAME과의 매핑이 제공된다.

‘1XX, 7XX, 8XX’는 이름(names) 변환을 위한 매핑표로 추가적으로 Processes 1과 0을 사용 한다. Processes 1과 0에는 bflc:nameXXMatchKey를 사용하여 이름 매칭키를 생성하거나, bflc:nameXXMarcKey를 사용하여 이름 MARC 키를 생성하거나, X00, X10, X11의 각 하위 필드에서 가져온 리터럴(literal)을 사용하여 이름 rdfs:label을 생성하거나, \$e(X00, X10) 또는 \$4를 기반으로 이름의 역할을 생성하며, bflc:PrimaryContribution 클래스를 사용하여 이름 주요 기여자를 생성하는 등의 작업이 포함된다. MARC에서 BIBFRAME 환경으로 가져오는 일부 요소는 유지될 필요가 없는 경우 병합 및 매칭이 완료되면 RDF에서 제거된다.

‘240, X30 등’은 통일 표제를 위한 것으로 bf:Work로 매핑되고, ‘200-24X (240 제외)’는 표제를 위한 것으로 bf:Instance로 매핑된다. 통일 표제인 130, 240은 Work - expressionOf - Hub의 bf:title로 매핑된다. 245는 bf:Instance - bf:title - bf:Title, bf:Work - bf: title - bf:Title로 인스턴스와 저작으로 매핑되는데 저작은 130, 240 필드가 없는 경우에만 매핑된다.

관계의 경우 ‘1XX’는 이름/표제 관계가 된다. ‘600-662’는 Subjects로 600/610/611에서 \$t가 있으면 W - bf:subject - bf:Hub로, 600에 \$t, \$v, \$x, \$y, \$z가 없으면 W - bf:subject - bf:Agent으로, 600 \$v, \$x, \$y, \$z가 있으면 W - bf:subject - bf:Topic으로 매핑된다. 700-730은 제2지시기호가 2인 경우 bf:hasPart를 갖고 그 외에는 relatedTo, 740은 bf:relation, 8XX는 bf:hasSeries, ‘Fields 760-788’ 연관저록은 필드 및 제2지시기호에 따라 관계가 결정된다. 510은 인용으로 W- references - W bnode로 해당 자원을 링크한다.

‘250-270’은 판사항, 발행사항을 위한 것으로 250 판사항은 I - editionStatement, 255 지도자원의 자료특성사항은 W - scale - Scale - rdfs:label, 260 발행사항은 I - provisionActivity - Publication에 매핑된다.

‘3XX’ 형태기술필드에서 식별기호 \$a는 I - extent - Extent, \$b는 I - note - Note - rdf:type rdf:resource = "http://id.loc.gov/vocabulary/mnotetype/physical" 및 rdfs:label을 사용하며, \$c는 I - dimensions - literal, \$e는 I - note - Note - rdf:type rdf:resource = "http://id.loc.gov/vocabulary/mnotetype/acmat" 및 rdfs:label로 변환한다.

‘5XX’ 주기로 주로 인스턴스와 연계되며 I - note - Note를 사용하거나 특정 속성으로 연결된다. 다만, 다음 주기는 인스턴스가 아닌 저작으로 연계된다.

502 W - dissertation - Dissertation

505 W - tableOfContents - TableOfContents

507 W - scale - Scale
518 W - capture - Capture
520 W - summary - Summary
521 W - intendedAudience - IntendedAudience
522 W - geographicCoverage - GeographicCoverage
525 W - supplementaryContent - SupplementaryContent
546, 580, 586, 587 W - note - Note

856, 880을 제외한 '841-887' 소장, 위치, 다른 그래픽 등은 'nac'로 변환되지 않는다. 856은 Process 3, 880의 식별기호 \$6은 명세를 참조하여 변환한다.

매핑표의 항목을 세부적으로 살펴보면, 첫 번째 컬럼은 MARC 표시기호, 식별기호, 지시기호이고, 다음 컬럼에는 변환 별로 BIBFRAME의 어휘가 제시된다. 표시기호 및 식별기호별로 변환 1(Conversion 1)은 BIBFRAME 모델을 대상으로 하는 매핑과 id.loc.gov의 어휘집을 나타내고, 변환 2(Conversion 2)는 조건에 따른 추가 매핑을 나타낸다. 속성은 전체 속성명으로 표기되지만, 저작, 인스턴스, 개별자료는 약어를 사용하여 Work는 W, Instance는 I, Item은 Item으로 표기되어 있다.

2. 표시기호별 분석

MARC21 TO BIBFRAME 매핑에 대한 대략적인 내용을 정리해 보면 〈표 2〉와 같다. 이는 표시기호를 중심으로 제시한 것이며, 'nac', 'ignore'에 해당하는 것은 제외하였다. 미의회도서관 레코드에서 요소 사용이 없거나 거의 없으면 'nac'(no attempt to convert)로, MARC 환경에서만 사용되는 데이터 요소는 'ignore'로 변환되지 않는다. 표시기호는 매핑이 이루어지지만 식별기호에서는 식별기호별로 매핑이 있는 것도 있고 없는 것도 있으므로 세부사항은 매핑표를 직접 참조해야 한다. 다만, 표시기호가 'nac', 'ignore'이면 식별기호의 매핑은 제공되지 않는다.

표시기호별로 약어를 제시하여 저작 속성인지 인스턴스 속성인지를 파악할 수 있다. 즉, 표시기호별로 저작에 해당하는 것은 W로, 인스턴스에 해당하는 것은 I로 클래스가 시작한다. 다만 051 태그는 로컬도서관에 해당하여 개별자료(Item)와 매핑된다.

RDF에서 주어 - 술어 - 목적어는 클래스 - 속성 - 클래스 혹은 리터럴이므로 클래스의 첫 글자는 대문자, 속성은 첫글자가 소문자이다. 목적어에 해당하는 것이 자원으로 클래스인 경우는 첫글자가 대문자로 시작하며, 문자열인 경우는 대쉬(-)만 표기한다. 클래스 다음에 나머지는 각 식별기호나 지시기호에 따라 RDF 문장이 생성된다. 예를 들어 020 ISBN인 경우 I - identifiedBy - Isbn이고, \$a는 # # - rdf:value 이고, \$c는 # # - acquisitionTerms - literal, \$q는 # # - qualifier - literal로 추가적으로 서술문이 완성된다.

KORMARC TO BIBFRAME 변환을 위한 매핑시 고려사항에 관한 연구

〈표 2〉 MARC21에서 BIBFRAME 서술문 매핑

(* nac, ignore 제외함)

MARC	BIBFRAME	MARC	BIBFRAME
리더	05-07, 17-18만 자리별로 변환이 다름		
001+	W - adminMetadata - AdminMetadata -	005	W - adminMetadata - AdminMetadata - changeDate - literal ^xs:dateTime
003	identifiedBy - Local - rdf:value		
007	SMD별로 변환이 다름	008	GMD별로 변환이 다름
010	I - identifiedBy - Lccn - rdf:value	034	W - scale - Scale
015	I - identifiedBy - Nbn	035	I - identifiedBy - Local
016	I - identifiedBy - CopyrightNumber	036	I - identifiedBy - StudyNumber
020	I - identifiedBy - Isbn	037	I - acquisitionSource - AcquisitionSource -
022	W - identifiedBy -	038	W - adminMetadata - AdminMetadata - bfcl:metadataLicenser - bfcl:MetadataLicenser
023	W - identifiedBy -	040	W - adminMetadata - AdminMetadata
024	I - identifiedBy - 지시기호에 따라 변경	041	W - language - Language - URI W - accompaniedBy - Work
025	I - identifiedBy - LcOverseasAcq	042	W - adminMetadata - AdminMetadata - descriptionAuthentication - DescriptionAuthentication
026	I - identifiedBy - Fingerprint	043	W - geographicCoverage - GeographicCoverage
027	I - identifiedBy - Strn	045	W - temporalCoverage
028	I - identifiedBy - 지시기호에 따라 변경	046	W - originDate - ^edtf:edtf W - validDate - ^edtf:edtf
030	I - identifiedBy - Coden	047	W - genreForm - GenreForm
032	I - identifiedBy - PostalRegistration	048	See ConvSpec-048
033	W - capture - Capture		
050	W - classification - ClassificationLcc	074	I - identifiedBy - Identifier - assigner - Agent
051	Item - classification - ClassificationLcc	080	W - classification - ClassificationUdc
052	W - geographicCoverage - GeographicCoverage	082	W - classification - ClassificationDdc
055	W - classification - ClassificationLcc	084	W - classification - Classification
060	W - classification - ClassificationNlm	086	W - classification - Classification
070	W - classification - ClassificationNal	088	I - identifiedBy - ReportNumber
072	W - subject - Topic		
100/	[Process0-이름/표제 및 Process1-이름 명세 참조] 이름: \$t 이전의 모든 식별기호(단, \$e, \$4, \$h, \$i, \$j(X11) 제외) 역할: \$e, \$4, \$j(X11) 표제: \$t 이후의 모든 식별기호(단, \$h \$v, \$x, \$y, \$z, \$w, \$0-8 제외) [제2지시기호 분출표목인 경우 Process 0.2 - 관계 구성요소 참조] 1XX 이름/표제 기술 대상		
700/	700-730 \$i 있으면 bfcl:relation 속성과 \$i 내용 기술함. \$i 없으면 제2지시기호가 2인 경우 bf:hasPart 매핑, 그외에는 bf:relatedTo 매핑 8XX bf:hasSeries 매핑		
800			
110/	710/ [제2지시기호 분출표목인 경우 Process 0.2 - 관계 구성요소 참조] 1XX 이름/표제 기술 대상 700-730 \$i 있으면 bfcl:relation 속성과 \$i 내용 기술함. \$i 없으면 제2지시기호가 2인 경우 bf:hasPart 매핑, 그외에는 bf:relatedTo 매핑 8XX bf:hasSeries 매핑		
710/			
811			
210	I - title - AbbreviatedTitle	245	(변환1) I - title - Title (변환2) W - title - Title
222	W - title - KeyTitle	246	(변환1) I - title - VariantTitle (변환2) W - title - VariantTitle
242	I - title - VariantTitle		
243	W - title - CollectiveTitle	247	W - title - VariantTitle
130	Work - expressionOf - Hub	700, 710, 711	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/relatedwork - associatedResource - Hub

MARC	BIBFRAME	MARC	BIBFRAME
240	Work - expressionOf - Hub	730	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/relatedwork - associatedResource - Hub
830	프로세스 6		
250	I - editionStatement	256	I - note - Note - rdf:type
254	I - editionStatement	257	I - originPlace - Place
	(식별기호별) W - scale - Scale - rdfs:label	260	I - provisionActivity - Publication
	W - cartographicAttributes - Cartographic - projection - Projection - rdfs:label	261	I - provisionActivity - Production
	W - cartographicAttributes - Cartographic - coordinates	262	I - provisionActivity - Publication
255	W - cartographicAttributes - Cartographic - ascensionAndDeclination	263	I - projectedProvisionDate - literal
	W - cartographicAttributes - Cartographic - equinox		(식별기호별)
	W - cartographicAttributes - Cartographic - outerGRing	264	I - provisionActivity - Production I - provisionActivity - Publication I - provisionActivity - Distribution I - provisionActivity - Manufacture 008/11-14에서 생성되지 않은 경우 I - copyrightDate - literal
	W - cartographicAttributes - Cartographic - exclusionGRing	265	I - acquisitionSource - AcquisitionSource
	(식별기호별)		
300	I - extent - Extent	341	W - contentAccessibility - ContentAccessibility
	I - note - Note - rdf:type		
	I - dimensions - literal	344	I - soundCharacteristic -
	I - note - Note - rdf:type	345	I - projectionCharacteristic -
306	I - duration - literal	346	I - videoCharacteristic -
310	I - frequency - Frequency	347	I - digitalCharacteristic -
321	I - frequency - Frequency		(식별기호별)
334	I - issuance - Issuance	348	W - musicFormat - MusicFormat W - notation - MusicNotation - rdfs:label
336	W - content - Content	351	W - collectionArrangement - CollectionArrangement
337	I - media - Media	352	I - digitalCharacteristic
338	I - carrier - Carrier	353	W - supplementaryContent - SupplementaryContent
	(식별기호별 차이)	362	(지시기호) I(1) - firstIssue - literal I(2) - lastIssue - literal I(3) - note - Note - rdf:type
	I - baseMaterial - BaseMaterial - rdfs:label	370	W - originPlace - Place - rdfs:label
	I - dimensions - literal	377	W - language - Language
	I - appliedMaterial - AppliedMaterial - rdfs:label	380	W - bf:genreForm - bf:GenreForm
	I - productionMethod - ProductionMethod - rdfs:label	382	W - musicMedium - MusicMedium
	I - mount - Mount - rdfs:label		(식별기호별)
	I - reductionRatio - ReductionRatio - rdf:value	383	W - musicSerialNumber - literal W - musicOpusNumber - literal W - musicThematicNumber - literal
	I - colorContent - ColorContent - rdfs:label	384	W - musicKey - literal
	I - systemRequirement - SystemRequirement - rdfs:label	385	W - intendedAudience - IntendedAudience
340	I - generation - Generation - rdfs:label	386	W - bfcl:creatorCharacteristic - bfcl:CreatorCharacteristic
	I - layout - Layout - rdfs:label		
	I - binding - Binding - rdfs:label		
	I - bookFormat - BookFormat - rdfs:label		
	I - fontSize - FontSize - rdfs:label		
	I - polarity - Polarity - rdfs:label		
	W - illustrativeContent - Illustration - rdfs:label		
	I - reductionRatio - ReductionRatio - rdfs:label		
490	Process 6		
500	I - note - Note	533	I - note - Note
501	I - note - Note	534	I - note - Note
502	W - dissertation - Dissertation	536	I - note - Note

KORMARC TO BIBFRAME 변환을 위한 매핑시 고려사항에 관한 연구

MARC	BIBFRAME	MARC	BIBFRAME
504	I - note - Note	538	I - systemRequirement - SystemRequirement
505	W - tableOfContents - TableOfContents	540	I - usageAndAccessPolicy - UsePolicy
506	I - usageAndAccessPolicy - AccessPolicy	541	Item - immediateAcquisition - ImmediateAcquisition
507	W - scale - Scale	544	I - note - Note
508	I - credits	545	I - note - Note
510	(자시기호에 따른 차이) W- bfcl: indexedIn - W bnode W- references - W bnode	546	(식별기호별) W - note - Note W - notation - Notation - rdfs:label
511	I - credits	547	I - note - Note
513	I - note - Note	550	I - note - Note
515	I - note - Note	555	I - note - Note
516	I - note - Note	556	I - note - Note
518	W - capture - Capture	561	I - custodialHistory
520	W - summary - Summary	563	I - note - Note
521	W - intendedAudience - IntendedAudience	580	W - note - Note
522	W - geographicCoverage - GeographicCoverage	581	I - note - Note
524	I - preferredCitation	583	I - note - Note
525	W - supplementaryContent - SupplementaryContent	585	I - note - Note
530	I - note - Note	586	W - note - Note
532	I - contentAccessibility - ContentAccessibility	587	W - note - Note
		588	I - note - Note
600	(변환) \$t 있지만 \$v, \$x, \$y, \$z 없으면 W - bf:subject - bf:Hub	651	(변환1) \$v, \$x, \$y, \$z 없으면 W - bf:subject - bf:Place (변환2) \$v, \$x, \$y, \$z 있으면 W - bf:subject - bf:Topic
610	(변환2) \$t, \$v, \$x, \$y, \$z 없으면 W - bf:subject - bf:Agent		
611	(변환3) \$v, \$x, \$y, \$z 있으면 W - bf:subject - bf:Topic		
630	(변환1) \$v, \$x, \$y, \$z 없으면 W - bf:subject - bf:Hub (변환2) \$v, \$x, \$y, \$z 있으면 W - bf:subject - bf:Topic	653	W - bf:subject : add rdf:type bfcl:Uncontrolled
647	(변환1) \$v, \$x, \$y, \$z 없으면 W - bf:subject - bf:Event (변환2) \$v, \$x, \$y, \$z 있으면 W - bf:subject - bf:Topic	655	(변환1) \$v, \$x, \$y, \$z 없으면 W - bf:genreForm - bf:GenreForm (변환2) \$v, \$x, \$y, \$z 있으면 W - bf:subject - bf:Topic
648	(변환1) \$v, \$x, \$y, \$z 없으면 W - bf:subject - bf:Temporal (변환2) \$v, \$x, \$y, \$z 있으면 W - bf:subject - bf:Topic	656	W - bf:subject - bf:Topic - madsrdf:ComplexSubject
650	W - bf:subject - bf:Topic - madsrdf:Topic	662	W - bf:subject - bf:Place - madsrdf:HierarchicalGeographic
720	W - contribution - Contribution - agent - Agent	753	I - systemRequirement
740	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/relatedwork - associatedResource - Work	758	기존 저작에 두번째 인스턴스로 생성
752	W - place - Place - madsrdf:HierarchicalGeographic		
760	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/series	775	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/otheredition
762	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/subseries	776	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/otherphysicalformat
765	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/translationof	777	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/issuedwith

MARC	BIBFRAME	MARC	BIBFRAME
767	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/translatedas	780 785	relation - Relation - relationship - Relationship - 관계는 제2지시기호에 따라 변경
770	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/supplement	786	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/datasource
772	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/supplementto	787	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/relatedwork
773	relation - Relation - relationship - Relationship - - http://id.loc.gov/vocabulary/relationship/part_of	788	필드는 nac이고 식별기호는 아래와 같음 변환1 # # - associatedResource - Work 변환2 # # - hasInstance
774	relation - Relation - relationship - Relationship - http://id.loc.gov/vocabulary/relationship/part		
856	Process 3	880	식별기호 \$6 명세

※출처: Library of Congress(2024)

MARC21에서 코드화된 정보를 포함하는 리더, 007, 008는 코드로 기술하며 자리 위치에 따라 해당 코드 정보가 다르므로 위치별로 BIBFRAME과 매핑된다. 특히 007, 008 모두 자료유형에 따라 내용이 다르기 때문에 자료유형에 따라 명확하게 매핑이 작성되어야 한다. <표 3>은 007(자료범주, 특정자료종별 제외)을 분석하여 정리한 것이며, <표 4>는 008의 내용 중 도서에 해당하는 매핑이다.

<표 3> MARC21 007 매핑

MARC21 007	BIBFRAME 클래스 및 속성	MARC21 007	BIBFRAME 클래스 및 속성
색채	I - colorContent - ColorContent	축소비율의 범위	I - reductionRatio - ReductionRatio
감광상태	I - polarity - Polarity	축소비율	I - reductionRatio - ReductionRatio - rdf:value
속도	I - soundCharacteristic - PlayingSpeed	필름의 기본재료	I - baseMaterial - BaseMaterial
물리적매체	I - baseMaterial - BaseMaterial	음향	I - soundContent - SoundContent
감광유제	I - baseMaterial - BaseMaterial I - emulsion - Emulsion	음향의 재생상태	I - soundCharacteristic - PlaybackChannels
규격	300 크기 내용이 없으면 생성 I - dimensions - "리터럴" 생성	제작요소	W - genreForm - GenreForm
크기	300 크기 내용이 없으면 생성 I - dimensions - "리터럴" 생성	복제단계	I - generation - Generation
테이프 폭	300 크기 내용이 없으면 생성 I - dimensions - "리터럴" 생성	복제유형	I - generation - Generation
필름의 폭	300 크기 내용이 없으면 생성 I - dimensions - "리터럴" 생성	제작/복제내용	I - productionMethod - ProductionMethod
기본재료	I - baseMaterial - BaseMaterial	음구의 폭/높이	I - soundCharacteristic - GrooveCharacteristic
스크린 형태	I - projectionCharacteristic - PresentationFormat W - aspectRatio - AspectRatio -	테이프 트랙의 수	I - soundCharacteristic - TapeConfig
녹음재생형태	I - soundCharacteristic - PlaybackChannels	음반/실린더/ 테이프의 종류	I - generation - Generation
비디오 녹화형식	I - videoCharacteristic - VideoFormat	재질	I - baseMaterial - BaseMaterial
음향유무	I - soundContent - SoundContent	음구의 깊임종류	I - soundCharacteristic - GrooveCharacteristic
음향수록매체	I - soundCharacteristic - RecordingMedium	특수재생장치특성	I - soundCharacteristic - PlaybackCharacteristic
틀의 재질	I - mount - Mount	녹음/저장 기법	I - soundCharacteristic - CaptureStorage

〈표 4〉 MARC21 008 매핑: 도서(Book) 예시

MARC21 008-Book		BIBFRAME
공통	00-05 Date entered on file	W - AdminMetadata - creationDate - literal ^^xs:date
	06 Type of date/Publication status	I - provisionActivity - Publication - date
	07-10 Date 1	I - provisionActivity - Publication - date
	11-14 Date 2	I - provisionActivity - Publication - date
	15-17 Place of publication, production, or execution	I - provisionActivity - Publication - place - Place
	35-37 Language	W - language - Language
	38 Modified record	ignore
도서	18-21 Illustrations	W - illustrativeContent - Illustration
	22 Target audience	W - intendedAudience - IntendedAudience
	23 Form of item	338 있으면 무시하고 없으면 I - carrier - Carrier
	24-27 Nature of contents	W - genreForm - GenreForm (with exceptions noted below)
	28 Government publication	W - bfcl:governmentPubType - bfcl:GovernmentPubType
	29 Conference publication	W - genreForm - GenreForm
	30 Festschrift	W - genreForm - GenreForm
	31 Index	W - supplementaryContent - SupplementaryContent - http://id.loc.gov/vocabulary/msupplcont/index
	32 Undefined	ignore
	33 Literary form	W - genreForm - GenreForm
	34 Biography	W - genreForm - GenreForm

3. MARC21 TO BIBFRAME 변환 매핑표 특징

가. 중복데이터를 생성하지 않음

MARC21에서 BIBFRAME으로 변환시 MARC 데이터에서 중복된 데이터가 BIBFRAME에서는 중복 생성되지 않는다. 즉, 중복되는 데이터 중에서 가장 우선하는 데이터를 BIBFRAME으로 변환하고, 우선하는 데이터가 없는 경우 다른 중복된 데이터로 변환하게 된다.

예를 들어 007/00, 01과 337(매체유형), 338(수록매체 유형) 필드의 내용이 중복성을 가질 수 있는데 이 중 우선하는 데이터를 336(내용유형), 337(매체유형), 338(수록매체 유형) 필드로 두었다. 이에 337 \$a나 336 \$a가 존재하지 않는 경우에 한해 007/00(자료범주)로 I - media - Media 혹은 W - content - Content로 변환한다. 또한, 007의 규격, 크기, 테이프폭, 월름폭은 300에 형태사항의 \$c(크기) 내용이 있으면 데이터를 변환하지 않고, 300 \$c가 없는 경우 007를 이용해 I - dimensions - "리터럴"로 크기 정보를 변환한다.

008/23 Form of item도 'If MARC tag 338 \$a exists, ignore'와 같이 338이 있으면 338 데이터로 BIBFRAME 변환을 생성하고, 338이 없는 경우에 008/23으로 변환한다. 또한, 008/06 t-발행연도/판권연도이고, 008/11-14에 판권연도가 있는 경우 이를 판권연도로 'I - copyrightDate - literal'와 같이 생성하고, 없는 경우 264에서 4 저작권일자를 이용하여 변환한다.

그러므로, 국내에서 KORMARC에서 BIBFRAME으로 변환하는 경우 이러한 사항을 고려하여

중복되는 데이터를 가능한 제외하고 생성하도록 개발해야 하며, MARC 데이터의 변환시 우선하는 데이터에 대한 기준도 마련해야 한다.

나. 미변환 항목

MARC21에서 BIBFRAME 변환 매핑표에 nac, ignore로 제시된 경우가 있는데 이는 BIBFRAME으로 변환되지 않는 MARC 데이터이다. 'nac'는 미의회도서관 레코드에서 요소 사용이 거의 없거나 전혀 없기 때문에 미변환하는 것이고, 'ignore'는 자체적으로 미정의되거나 MARC 이외에서는 의미 없는 데이터이기 때문에 미변환한다. 이러한 미변환은 표시기호, 식별기호, 코드 기술하는 자리별, 자리의 코드값별로 제시된다.

표시기호 자체가 nac, ignore인 경우 해당하는 식별기호의 매핑은 제시될 필요가 없다. 표시기호 뿐만 아니라 식별기호 수준에서도 'nac', 'ignore'가 제시된다. 즉, 표시기호는 BIBFRAME으로 매핑되지만 해당하는 식별기호 일부가 nac, ignore인 경우가 있는데 이는 식별기호의 일부가 매핑되지 않는 경우이다.

또한, 리더, 007, 008 등과 같이 부호로 입력하는 경우 부호에 해당하는 자리별로 nac, ignore가 제시되기도 하고, 자리에 해당하는 값이 nac, ignore인 경우도 있다(〈표 5〉 참조). 예를 들어 리더 00-04는 레코드의 길이로 이는 MARC에만 적용되므로 ignore로 미변환된다. 08 제어유형, 09 문자부호화체계, 10 지시기호 자리수, 11 식별기호 자리수, 12-16 데이터 기본번지, 20-23 엔트리 맵은 MARC를 위한 사항이므로 'ignore' 미변환이다. 19-멀티파트 자원 레코드 수준은 미의회도서관에서 사용이 없으므로 'nac' 미변환된다.

또한 자리에 해당하는 코드값이 없는 경우 매핑이 제공되지 않는다. 예를 들어, 008/24-27 Nature of contents에서 코드값 b - Bibliographies는 'W - supplementaryContent - SupplementaryContent - <http://id.loc.gov/vocabulary/msupplcont/bibliography> : add rdfs:label "bibliography"'이지만, 코드값 a - Abstracts/summaries는 'nac'로 매핑이 제공되지 않는다.

〈표 5〉 nac, ignore에 해당하는 표시기호 및 자리

구분	필드 및 자리	비고
리더	ignore: 00-04, 08-16, 20-23 / nac: 19	자리별
007	<ul style="list-style-type: none"> 전자자료: 이미지비트 깊이, 파일형식의 종류, 품질보증지표, 원시자료의 정보원, 압축수준, 재포맷의 품질 영화: 색채의 범주, 다채색 인화 유형, 변질 상태, 필름의 완전성, 필름 검사일자 원격탐사도: 탐사기고도, 탐사기각도, 구름비율, 탑재기유형, 탑재기이용범주, 탐사기유형, 데이터유형 축감자료: 점자표기의 유형, 축약수준, 점자악보형식, 형태적 특성 	자리별
008	<ul style="list-style-type: none"> 공통: 38(수정레코드), 39(목록전거) 자료별: 지도자료유형, 음악의 파트, 음악 편곡 	자리별
가변길이 데이터 필드	013, 018, 031, 044, 061, 066, 071, 085, 258, 270, 307, 335, 342, 343, 350, 355, 357, 363, 365, 366, 381, 387, 388, 514, 526, 535, 542, 552, 562, 565, 567, 584, 654, 657, 658, 688, 751, 754, 841, 842, 843, 844, 845, 850, 851, 852, 853, 854, 855, 863, 864, 865, 866, 867, 868, 876, 877, 878, 881, 882, 883, 884, 885, 886, 887	표시기호

다. 인코딩스킴 사용

리더, 007, 008의 코드값은 URI를 포함하는 어휘 인코딩스킴 및 구문 인코딩스킴을 주로 사용한다. 또한, 주제명, 에이전트명 등 도서관에서 생성하는 고유 전거파일이나 주제명의 URI를 이용하여 매핑한다. 특히, 에이전트 처리에서 일반적으로 RWO(Real World Object)가 \$1에 존재하거나 \$0에서 생성될 수 있는 경우, 해당 URI를 사용한다(〈표 6〉 참조).

〈표 6〉 인코딩스킴을 사용하는 항목(일부)

구분	인코딩스킴을 사용하는 요소
리더	05(레코드상태), 06(레코드유형), 07(서지수준), 17(입력수준), 18(목록기술형식)
007	00(자료범주), 01(특정자료종별), 색채, 감광상태, 속도, 물리적 매체, 감광유제, 기본재료, 녹음재생형태, 비디오녹화형식, 복제유형, 스크린 형태, 음향, 음향유무, 음구, 축소비율, 틀의재질, 음향수록매체, 지도 제작 및 복제 내용, 음향 재생상태, 테이프트랙 수, 녹음자료 음반 실린더 테이프의 종류, 영화 제작요소, 마이크로 감광유제, 녹음자료 재질, 복제단계, 음구 깎임 종류, 필름 기본재료, 녹음자료 특수재생장치, 녹음저장기법
008	(공통) 발행국명, 한국대학부호, 언어 (도서) 삽화, 이용대상자수준, 개별자료형태, 내용형식, 회의간행물, 색인, 문학형식, 전기
1xx	RWO(Real World Object)로 \$1, \$0의 URI

※자료범주에 따라 자리가 변경되어 007/00, 01 이외에는 자리수를 제시하지 않음. 008은 도서만 제시함

어휘 인코딩스킴으로 genreForms의 사용이 널리 사용되고 있는데 예를 들어, 지도자료에서 지도집, 도표, 지도, 측면도와 같은 경우는 genreForms의 용어가 사용된다. 회의간행물, 기념논문집, 색인, 문학형식, 전기과 관련된 사항도 모두 genreForms의 용어집을 근간으로 한다.

어휘 인코딩스킴 이외에도 연도 입력에서는 구문 인코딩 스키미 사용된다. 발행년 유형 부호별 변환에서 EDTF(Extended Date/Time Format) 명세를 이용하여 작성한다. 예를 들어, 발행년의 경우 I - provisionActivity - Publication - date : 날짜형식: <07-10>-<11-12>-<13-14>^^edtf:edtf를 제시한다.

IV. KORMARC TO BIBFRAME 변환시 고려사항

3장에서 분석된 MARC21 TO BIBFRAME의 특징을 바탕으로 국내에서 KORMARC TO BIBFRAME 변환을 위해 매핑표 개발시 고려사항으로 (1) KORMARC용 어휘 인코딩스킴 개발, (2) KORMARC에만 있는 표시기호 및 식별기호 변환, (3) MARC21에서 미변환(nac, ignore) 요소 중에서 국내에 변경할 사항으로 나누어 제안하고자 한다.

1. 어휘 인코딩스킴 개발

링크드데이터에서는 목적어가 다시 서술문을 갖기 위해서는 리터럴이 아닌 자원으로 기술되어야 한다. 즉 공백노드를 포함하여 노드가 되어야 하는데 이를 위해서는 모든 값공간의 값 기술시 리터럴이 아닌 어휘 인코딩스킴의 사용이 요구된다. 이를 위해 전 세계적으로 기 발행된 어휘 인코딩스킴을 활용하는 것과 자체적으로 새로 개발할 어휘 인코딩스킴을 개발하는 것을 모두 고려할 수 있다.

첫째, 기 발행된 어휘 인코딩스킴을 활용하는 경우 이미 미의회도서관에서 사용하는 ID라는 어휘인코딩스킴을 활용할 수 있으며, 다만 해당 용어를 국내에 맞는 용어로 변환하여 사용해야 한다. 즉, ID를 사용하되 필요시 표출하는 용어는 lang = "kor"를 이용해 한국어로 변경한다. 예를 들어, 008의 경우 자료유형에서 해당하는 어휘 인코딩스킴을 선택하고 <http://id.loc.gov/vocabulary/contentTypes/txt>: lang = "kor" rdfs:label "문자자료"와 같이 사용한다.

둘째, 국내에 맞는 어휘 인코딩스킴을 자체 개발하는 경우 웹으로 발행하여 전세계 도서관에서 사용할 수 있도록 발행한다. KORMARC을 위해 새로 개발할 사항으로는 <표 7>과 같은 자료범주 및 특정자료종별 등의 어휘 인코딩스킴이다. MARC21 TO BIBFRAME에서 가장 많이 사용되는 것이 장르 용어인데 이를 위해 국내에 맞는 장르용 어휘 인코딩스킴의 생성을 고려해야 한다. 특히, 국내에서 반드시 작성해야 하는 인코딩스킴으로 한국대학부호, 한국정부기관부호, 고서와 관련된 사항이다.

<표 7> 개발해야 할 KORMARC 어휘 인코딩스킴: 007, 008을 중심으로

구분	내용		비고
리더	레코드상태, 레코드유형, 서지수준, 입력수준, 목록기술형식		ID 활용
	MARC21의 어휘 인코딩스킴을 사용하되 한글화해야 함 고서의 경우 다음과 같이 추가 개발이 요구됨		
007	007/01 a - 권축장 b - 절첩장 c - 호첩장 d - 포배장 e - 선장 f - 족자 g - 접포 h - 모장 i - 낱장	어휘 인코딩스킴 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "권축장" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "절첩장" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "호첩장" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "포배장" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "선장" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "족자" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "접포" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "모장" 추가 # # - http://.....gov/authories/genreForms/ ... : rdfs:label "낱장" 추가	ID 활용 및 신규 개발 (고서 관련)
008	• 공통: 발행국명, 한국대학부호 , 한국정부기관부호 , 언어부호 • 도서: 삽화표시, 이용대상자수준, 개별자료형태, 내용형식, 회의간행물, 기념논문집, 색인, 문학형식, 전기 • 전자자료: 유형 • 지도자료: 모형도, 도법, 유형, 형태상의특징 • 음악녹음자료: 작곡형식, 악보형식, 악보파트보, 말립자료, 녹음자료의내용, 조높김 및 편곡 • 계속자료: 간행빈도, 정규성, 유형, 원자료 발간 형식, 내용형식, 표제의 언어, 레코드생성방식 • 시청각자료: 유형, 기술수준 • 고서: 삽화표시, 광과, 사경 및 사본의 계선, 판구, 어미, 내용형식, 판종, 문학형식		ID 활용 및 신규 개발 (밀줄친 부분)

2. KORMARC에만 있는 내용의 변환

KORMARC이 MARC21을 기본으로 개발되었지만 KORMARC에만 있는 특별한 내용(표시기호 및 식별기호 등)이 있으므로 이를 BIBFRAME과 매핑시키기 위한 방안이 필요하다. KORMARC에만 있는 표시기호의 BIBFRAME으로 변환 제안은 〈표 8〉과 같다. 012 국립중앙도서관 제어번호는 I - identifiedBy - Nlkcn - rdf:value로, 052 국립중앙도서관 청구기호는 ClassificationNlk라는 클래스를 개발하여 W - classification - ClassificationNlk로, 056 한국십진분류기호는 ClassificationKdc 클래스를 개발해 W - classification - ClassificationKdc로 매핑이 가능하다. 090 자관청구기호는 ShelfMarkKdc 클래스를 개발하여 Item - shelfMark - ShelfMarkKdc 형태로 매핑할 수 있다.

〈표 8〉 KORMARC에만 있는 표시기호의 BIBFRAME 변환

표시기호	BIBFRAME 제안
012 국립중앙도서관 제어번호	I - identifiedBy - Nlkcn - rdf:value
052 국립중앙도서관 청구기호	W - classification - ClassificationNlk
056 한국십진분류기호	W - classification - ClassificationKdc
049 소장사항	Item - heldBy - Agent Item - identifiedBy - rdfs:value Item - shelfMark - ShelfMark - note - Note - rdfs:label Item - shelfMark - ShelfMark - note - Note - rdfs:label Item - physicalLocation - rdfs:label
090 - 자관청구기호	Item - shelfMark - ShelfMarkKdc
590 - 소장본 주기	무시함

KORMARC에만 있는 식별기호의 경우도 매핑이 고려되어야 하는데 예를 들어, 020 ▼g - 부가기호의 경우 Note 클래스와 note 속성을 사용하여 매핑할 수 있다.

부호화정보필드 중 KORMARC에만 있는 자리수로 008/26-27 한국대학부호의 경우 인스턴스를 중심으로 I - provisionActivity - Publication - agent - Agent와 같이 개발할 수 있다. 또한, 008 고서의 경우 고서의 특성을 고려하여 고서의 삽화표시, 광과, 사경 및 사본의 계선, 판구, 어미, 내용형식, 판종, 문학형식을 개별적으로 개발하기 위해서는 새로운 속성과 클래스 개발이 필요하다. 다른 매핑 방법은 고서를 도서와 같은 유형으로 취급하는 것으로, 이 경우에는 고서의 특성을 반영하여 변환하기 어렵다.

다만, KORMARC에만 있는 내용은 추후 데이터를 상호 운용할 때 손실될 수 있는 사항이며, 손실되지 않도록 가능한 BIBFRAME의 클래스와 속성을 사용하여 개발할 필요가 있다. 속성과

클래스를 자체 개발하는 경우 이러한 데이터는 다른 국가와 상호운용할 때 적용되기 어렵다는 점을 고려해야 한다.

3. 미변환 데이터에서 고려사항

MARC21 TO BIBFRAME 매핑 문서를 검토한 결과 BIBFRAME 2.0에는 아직도 많은 MARC 필드가 BIBFRAME으로 변환되지 않고 있다. 이는 'ignore' 또는 'nac'으로 표시된 필드이다. 'ignore'는 MARC만을 위한 정보이므로 변환할 필요가 없지만 'nac'에 해당하는 것은 미의회도서관이 아닌 다른 기관에서는 필요한 데이터일 수 있으므로 국내에서는 고려해야 할 대상이 될 것이다.

MARC21에서는 미변환이지만 국내 상황에 따라 변환을 고려할 수 있는 표시기호로는 307 이용 시간, 335 확장계획, 342 지리공간참조정보, 343 평면 좌표 정보, 355 보안분류 통제, 357 원작자의 배포 제어, 363 권·연차, 연월차 부호, 365 판매 가격, 366 판매 정보, 381 저작 또는 표현형의 기타 구별 특성, 387 대표표현형 특성, 388 창작기간, 514 지리공간데이터 품질주기, 526 학습프로그램 주기, 535 원본/복제본 소재 주기, 542 저작권 관련 정보, 552 개체와 속성 정보 주기, 562 사본/판본 식별 주기, 565 사례조사파일 특성 주기, 567 방법 주기, 584 이용 누적 및 빈도 주기 등이 있다. 이 중에서 국내 도서관을 위해 매핑이 필요한 사항에 대해서는 매핑을 제공해야 한다.

사실 미변환 데이터에 대해서는 다양한 의견이 제안되고 있다. 첫째, 미의회도서관에 미래에 사용되지 않을 데이터를 BIBFRAME 변환에 추가하라는 제안을 하기 보다는 이러한 누락 부분에 대해 특정 커뮤니티에서 변환 사양을 보강한 매핑표를 개발하고 공유하여 안정화된 후에 업그레이드된 매핑표를 활용하도록 하는 것이다. 둘째, 커뮤니티가 기여하여 완벽한 매핑표를 작성하는 “슈퍼 매핑”을 하여 모든 사용 가능한 데이터를 변환 가능하도록 하는 것이다. 다만 문제는 로컬을 중심으로 한 매핑은 데이터가 공유되고 사용될 때 혼란과 표준화 부족을 초래할 수 있다는 한계는 있다.셋째, 다양한 커뮤니티의 변형을 허용하는 방법이다. 즉, 커뮤니티가 자체 변형 변환 사양을 설명하고 제시할 수 있는 플랫폼을 제공하거나 링크를 제공하는 것이다. 이는 더 많은 논의가 필요하지만, 기본 아이디어는 “모든 도서관에 동일한 방법이 반드시 맞지는 않는다”는 것에서 출발한다. 다만, 기본 출력은 필수적인 BIBFRAME 구조를 준용하도록 한다. 예를 들어, 음악 서지 담당자들이 음악 서곡의 매핑을 원할 수 있는 반면, 다른 그룹은 그렇지 않을 수 있다. RDF 그래프는 다른 어휘에서 추가 클래스와 속성을 포함함으로써 쉽게 확장하고 추가할 수 있기 때문에, 다양한 커뮤니티들의 자체 변환 사양을 유연하게 반영하도록 하고 이들 중 일부는 “BIBFRAME 확장”으로 지정하는 것이다(Xu et al., 2017, 241, 245-246).

그리므로, MARC21의 미변환 필드의 경우 KORMARC에서도 동일하게 적용할 것인지 아니면 매핑을 새롭게 개발할 것인지를 논의하여 국내에 맞는 매핑표를 개발하는 것이 필요할 것이다.

V. 결 론

본 연구에서는 MARC21 TO BIBFRAME의 변환을 바탕으로 KORMARC TO BIBFRAME의 변환시 고려할 사항을 제안하고자 문헌연구와 함께 MARC21 TO BIBFRAME 매핑표를 표시기호별, 식별기호별 등 다층적으로 분석하였다.

MARC21을 필드의 그룹별로, 표시기호별로 분석을 통해 MARC21 TO BIBFRAME 변환의 주요 특징을 다음과 같이 파악하였다. 첫째, 중복데이터를 생성하지 않는다는 점이다. MARC21에서는 중복데이터가 많았으나 BIBFRAME에서는 이를 중복 생성하지 않고, 우선순위에 따라 데이터를 생성하고 중복되는 다른 데이터는 미변환한다. 둘째, MARC 데이터가 BIBFRAME으로 변환되지 않는 미변환 항목(nac, ignore)이 상당히 포함되어 있다. MARC 특성에 따른 데이터와 미의회도서관에서 사용되지 않는 데이터가 미변환이다. 그러나 이러한 미변환 중에서 변환이 필요한 사항이 있을 수 있다는 점에서 검토가 필요하다. 셋째, BIBFRAME 변환은 어휘나 구문의 인코딩스킴의 사용을 바탕으로 한다. 발행년에는 구문 인코딩스킴이, 다양한 어휘는 통일을 위해 어휘 인코딩스킴이 사용되고 있다.

분석된 내용을 바탕으로 국내에서 KORMARC TO BIBFRAME 변환을 위해 매핑표를 개발하기 위한 고려사항으로 3가지를 제안하였다. 첫째, KORMARC용 어휘 인코딩스킴 개발이 필요하다. 인코딩스킴은 기존에 발행된 것을 사용하는 것과 새롭게 개발하는 방법이 있다. 둘째, KORMARC에만 있는 표시기호 및 식별기호 변환에 대한 고려가 필요하다. 국내에만 있는 데이터요소이므로 BIBFRAME 요소를 사용하여 매핑표를 개발해야 한다. 셋째, MARC21에서 미변환(nac, ignore) 요소 중에서 국내에 변경할 사항 재검토하여 적절한 매핑이 필요하다.

BIBFRAME으로 데이터를 변환하기 위해서는 변환기가 개발되어야 하고, 변환기는 매핑표를 바탕으로 하므로 국내에서도 매핑표 개발을 바탕으로 변환 테스트를 시도해 보아야 할 것이며 이를 통해 변환 매핑표의 지속적인 오류 검증 및 업그레이드가 필요할 것이다.

참 고 문 헌

국립중앙도서관 (2022). 차세대 서지 언어 BIBFRAME 명세서 번역 및 해설서 제작. 서울: 국립중앙도서관.

국립중앙도서관 (2024). 국가서지에 대한 차세대 서지 구조 적용방안 연구. 서울: 국립중앙도서관.

이성숙, 이지원 (2018). MARC 21과 BIBFRAME 2.0의 변환에 관한 기초 연구. 충남대학교 사회과학연구, 29(3), 73-93.

Bigelow, I., Davoodi, D., Farnel, S., & Sparling, A. (2018). Who will be our bf: comparing techniques for conversion from MARC to BIBFRAME. In WLIC, Kuala Lumpur. Available: <https://library.ifla.org/id/eprint/2194/1/141-bigelow-en.pdf>

Ford, K. & Williamschen, J. (2023, July 12). Breaking news: splitting MARC records to create better. In 2023 LD4 Conference on Linked Data. Available: <https://www.youtube.com/watch?v=9i3y23XiNQE>

Fu, Y. J. & Ke, H. R. (2021). A study of converting MARC 21 to BIBFRAME 2.0: a case of printed text material “Hong Lou Meng.” *Journal of Educational Media & Library Sciences*, 58(2), 157-191. [https://doi.org/10.6120/JoEMLS.202107_58\(2\).0056.RS.AM](https://doi.org/10.6120/JoEMLS.202107_58(2).0056.RS.AM)

Jin, Q., Hahn, J., & Croll, G. (2016). BIBFRAME transformation for enhanced discovery. *Library Resources & Technical Services*, 60(4), 223-235. <https://doi.org/10.5860/lrts.60n4.223>

Joudery, D. N. & Taylor, A. G. (2018). *The Organization of Information*. California: Libraries Unlimited.

Library of Congress (2024). MARC 21 to BIBFRAME Conversion Specifications 2.8. Available: <https://www.loc.gov/bibframe/mtbf/>

Samples, J., & Bigelow, I. (2020). MARC to BIBFRAME: converting the PCC to linked data. *Cataloging & Classification Quarterly*, 58(3/4), 403-417. <https://doi.org/10.1080/01639374.2020.1751764>

St. Pierre, M. & LaPlant, W. (1998). Issues in Crosswalking Content Metadata Standards. NISO White Paper.

Xu, A., Hess, K., & Akerman, L. (2017). From MARC to BIBFRAME 2.0: crosswalks. *Cataloging & Classification Quarterly*, 56(2/3), 224-250. <https://doi.org/10.1080/01639374.2017.1388326>.

• 국한문 참고문헌의 영문 표기

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

Lee, Sung Sook & Lee, Ji Won (2018). Study on conversion between MARC 21 and BIBFRAME 2.0. *Chungnam National University Social Science Research*, 29(3), 73-93.

National Library of Korea (2022). *Next-generation Bibliographic Language BIBFRAME Specification Translation and Commentary*. Seoul: National Library of Korea.

National Library of Korea (2024). *A Study on the Application of BIBFRAME to National Bibliography*. Seoul: National Library of Korea.