

BIBFRAME 구축 사례 분석을 통한 국내 적용방안에 관한 연구*

Analyzing BIBFRAME Cases for the Development of BIBFRAME Application Plans in Korea

이 미 화(Mihwa Lee)**

<목 차>

I. 서론	IV. 국내에서 BIBFRAME 적용방안
II. 도서관 링크드데이터	1. 국내 어휘집의 링크드데이터로 변환 구축
1. 발전 과정	2. BIBFRAME 온톨로지 적용 및 시스템 개발
2. 어휘집과 온톨로지	3. 목록사서 교육
III. BIBFRAME 구축 사례 분석	V. 결론
1. 미의회도서관 BIBFRAME 구축	
2. Linked Data for Production (LD4P)	

초 록

본고는 링크드데이터를 위한 도서관 분야의 온톨로지 BIBFRAME 개발에 따라 국내에서 BIBFRAME 적용을 위한 구체적인 방안을 마련하는 것이다. 이를 위해 문헌연구, 사례조사, 설문조사를 실시하였고, 사례조사로 BIBFRAME을 개발한 미의회도서관과 BIBFRAME 프로젝트를 진행하는 LD4P를 분석하고, 설문조사에서는 국내 목록사서를 대상으로 링크드데이터 관련 용어의 이해도, 링크드데이터 구축을 위해 필요한 조건 등에 대해 조사하였다. 이를 바탕으로 국내에서 BIBFRAME을 도입하기 위한 적용방안을 다음과 같이 제시하였다. 첫째, 이집트, 주제명 등의 기존 전거데이터를 링크드 데이터로 출판하는 것뿐만 아니라 도서관에서 통제어휘로 사용하거나 데이터 값으로 사용하는 용어집도 링크드데이터로 생성해야 한다. 둘째, 국내에서 BIBFRAME의 적합성을 판단하고, 확장 모델링을 개발할 필요가 있으며, KORMARC과의 매핑 테이블 작성과 변환기 및 입력기의 시스템 개발도 요구된다. 셋째, BIBFRAME과 같은 링크드데이터 구축이 사서의 고유 업무가 될 수 있도록 사서를 대상으로 한 체계적인 교육프로그램이 개발되어야 한다. 따라서 본 연구는 국내에서 BIBFRAME 구축을 위한 실질적인 적용방안을 모색하였다는 점에서 그 의미가 있다.

키워드: 링크드데이터, 비브프레임, 온톨로지, 어휘집

ABSTRACT

This study is to suggest the concrete application plan of BIBFRAME under the development of BIBFRAME as library specific ontology for linked open data. The several research methods are used as the literature reviews, the case study of LC and LD4P, and the survey of cataloging librarians which is to grasp understanding level of the linked data related terms and requirements for constructing LOD. The application plan is suggested as follows. First, publishing name authority data and subject heading in LOD are prominent as the startup with creating terms list or vocabulary in LOD that has been used in library for controlled vocabulary and data value. Second, it is needed to develop BIBFRAME application and extension modeling in Korea, to map KORMARC and the properties and classes of BIBFRAME, and to develop the editor and MARC to BIBFRAME Transformation Tools. Third, the systematical training for cataloging librarians is designed to regard BIBFRAME related works as the librarian's main field. Therefore, this study would contribute to seek the practical application plan for BIBFRAME in Korea.

Keywords: BIBFRAME, SKOS, FOAF, Ontology, Vocabulary, LD4P, Bibliotek-o, BIBO, LDS, RDF, MADSRDF, Authority control

* 이 논문은 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A5A8018778)

** 공주대학교 문헌정보교육과 부교수(leemh@kongju.ac.kr)

•논문접수: 2018년 5월 8일 •최초심사: 2018년 5월 28일 •게재확정: 2018년 6월 18일

•한국도서관정보학회지 49(2), 59-78, 2018. [<http://dx.doi.org/10.16981/kliss.49.201806.59>]

I. 서론

링크드데이터는 상호 연결된 웹을 지향하는 모형으로, URI와 RDF의 트리플 구문을 이용하여 데이터를 구조화하여 기술함으로 정보의 배포와 접근을 향상시킬 수 있다. 링크드데이터 구축을 통해 한 분야에서 구축된 정보는 다른 분야에서 재사용이 가능하므로 특히 데이터 구축이 중복적인 도서관 및 박물관과 같은 기관에서 링크드데이터의 활용가능성이 매우 크다. 도서관에서 구축해온 서지데이터와 전거데이터를 링크드데이터에 맞게 변환하여 제공한다면 여러 분야에서 활용이 가능하고, 타 분야의 링크드데이터를 도서관에서도 활용할 수 있을 것이다.

미의회도서관에서는 서지 및 전거데이터를 링크드데이터로 변환하기 위해 전거데이터를 MADSRDF를 이용해 링크드데이터로 출판하고, BIBFRAME 온톨로지를 개발해 서지레코드를 링크드데이터로 변환하였다. BIBFRAME은 메타데이터 인코딩과 교환 표준으로 MARC21을 대체하고, 웹과 네트워크 환경에서 발생하는 서지 기술의 미래를 위한 기반이 되고 있다. BIBFRAME 온톨로지를 이용하여 LC를 비롯해 OCLC, Casalini Libri, LD4 프로젝트, Zepheira의 Library.Link 등에서 링크드데이터를 개발 중이며(Lovins and Hillmann 2017), 미국 대학도서관에서도 BIBFRAME을 이용한 프로젝트를 진행하고 있어 BIBFRAME이 도서관 특화된 링크드데이터 온톨로지로 자리매김하고 있는 상황이다.

국내 도서관의 링크드데이터 구축은 2013년 국립중앙도서관, KERIS, KISTI을 중심으로 공공부문에서 시작하였다. 이들은 표준 온톨로지를 개발하는 대신 BIBO, DC와 같은 기존 온톨로지를 재사용하여 서지데이터를 링크드데이터로 변환하였고, 이러한 재사용 온톨로지에는 도서관 특성이 반영된 요소들이 포함되지 않아 자체적으로 별도의 온톨로지를 추가 개발하였다. 그러므로, 국내에서 도서관 링크드데이터의 체계적인 모델링과 완전한 상호운용성을 위해 도서관 특화된 표준 온톨로지로서 BIBFRAME의 적용방안을 모색해야 한다.

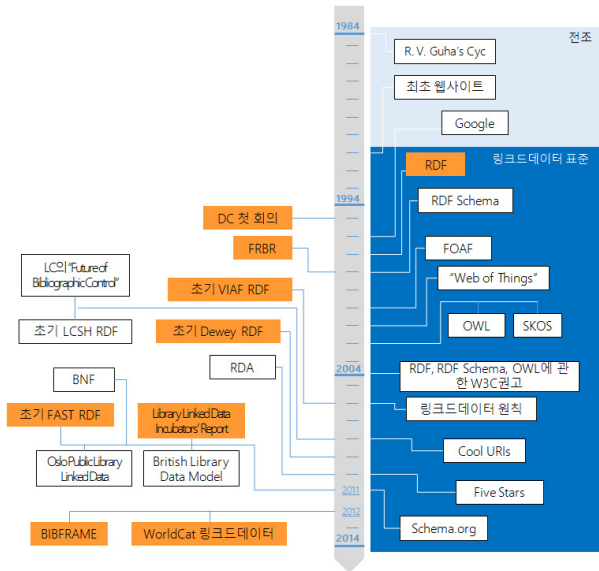
이에 본고에서는 BIBFRAME의 국내 적용을 위한 방안을 마련하고자 문헌연구, 사례조사, 설문조사를 실시하였다. 문헌연구와 사례조사를 통해 국외 도서관에서 링크드데이터 구축 및 실행과정을 살펴보았다. 사례조사로 BIBFRAME을 개발한 미의회도서관과 BIBFRAME 프로젝트를 진행하는 LD4P를 분석하였다. 설문조사에서는 국내 목록사서를 대상으로 링크드데이터 관련 용어의 이해도, 링크드데이터 구축을 위해 필요한 조건 등에 대해 조사하였다. 본 연구는 BFRAME의 구축 사례를 분석하여 국내에서 BIBFRAME 구축을 위한 실질적인 적용방안을 구체적으로 모색하였다는 점에서 그 의의가 있다.

II. 도서관 링크드데이터

1. 발전 과정

의미적으로 풍부한 구조화된 데이터의 이용은 지식의 발견을 향상시킬 수 있다는 진제하에 시맨틱웹이 탄생하였고, 시맨틱웹을 위한 웹친화적이고 구조화된 데이터 기술을 위해 RDF, OWL, SKOS, FOAF 등의 공식적 표준이 개발되었다. RDF(Resource Description Framework, 이하 RDF)는 링크드데이터를 위해 선택된 언어로 RDF의 기본 구축 단위는 주어(S) + 술어(P) + 목적어(O)의 트리플로 구성된 문(statement)이다. 문은 다이어그램과 Turtle, N-Triples, RDF/XML, JSON-LD의 직렬화 형태로 표현할 수 있는데 RDF/XML을 주로 사용한다(Coyle 2010, 14). OWL(Web Ontology Language, 이하 OWL)은 RDF를 확장하는 표준이며 특정 시맨틱웹 온톨로지를 정의하는데 사용된다(Coyle 2012a, 11). 즉 OWL은 온톨로지라는 공식적 시맨틱 통제어휘를 생성하는 지식표현 언어로 시맨틱웹의 제한 사항과 속성을 추가하는 매커니즘을 제공한다(Huerga and Lauruhn 2016, 64). 이후 RDF를 위한 메타언어로 SKOS(Simple Knowledge Organization Scheme, 이하 SKOS)와 FOAF(Friend of a Friend, 이하 FOAF)가 개발되었다. SKOS는 시소러스와 다른 구조적인 어휘집을 RDF로 표현하기 위한 표준이고, FOAF는 소셜네트워크를 기술하기 위한 것으로 2000년에 소개되었다(Godby, Wang, and Mixer 2015, 10). SKOS는 시소러스, 분류표, 주제명표목표와 같은 조직화된 데이터를 나타내는 표준으로, 계층적 관계를 나타낼 수 있고, 색인어, 엔트리 어휘, 정의를 제공할 수 있다. SKOS의 기반은 RDF이며, SKOS는 유형, 속성, 값의 RDF 개념을 이용한다. SKOS는 웹상에서 LCSH의 실행을 위해 사용되고, RDA의 일부인 어휘집의 인코딩을 위해서도 사용된다(Coyle 2010, 17).

도서관 표준도 <그림 1>과 같이 많은 변화를 이어오고 있는데 어휘집, 전거파일, 메타데이터 표준간의 매핑과 같은 도서관에서 관리해온 자원은 도서관과 웹을 통합하는 환경 하에서 웹친화적 포맷으로 재구성되어야 했다. 이에 미의회도서관은 2008년 보고서 *On the Record* 이후 서지프레임워크 변환 프로젝트를 구성하여 FRBR, RDA와 같은 개체-관계 모델의 이용을 권고하고, MARC를 대체할 BIBFRAME을 2012년 개발하였다. 도서관 표준 전문가들은 RDF 인코딩을 위한 모델을 개발하고, OMR(Open Metadata Registry, 이하 OMR)과 같이 웹기반 온톨로지와 어휘집 저장소를 생성하고, 도서관과 다른 문화유산기관에서 운영하는 비전통적인 자원의 웹친화적 모델을 구축하였다. 이에 LCSH, VIAF, FAST와 같은 도서관의 전거데이터는 새로운 아키텍처에 맞게 재구조화되었다(Codby, Wang, and Mixer 2015, 12-13). 링크드데이터 환경에서는 이미 정의된 온톨로지나 데이터를 재사용하는 것이 선호되므로 BIBFRAME과 같은 온톨로지와 링크드데이터로 구축된 LCSH와 같은



〈그림 1〉 도서관 링크드데이터 개발의 개요

※출처: Codby, Wang and Mixter(2015)

집 모두 어휘라는 용어를 사용할 수 있으나 그 기능이 다르기 때문에 구분을 위해 메타데이터 스키마 및 요소는 온톨로지로, 통제어휘집은 어휘집으로 정의한다. 링크드데이터 구축을 위해 도서관 특화된 온톨로지를 선택하고, 해당 요소의 값을 통일성 있게 입력할 수 있는 어휘집을 구축해야 한다.

가. 어휘집

서지레코드를 링크드데이터로 변환하기 위해서는 먼저 어휘집의 링크드데이터 구축이 필요하다. LC에서는 BIBFRAME을 개발하기 전 통제어휘집인 LDS(Linked Data Services, 이하 LDS)를 구축하였다. LDS는 미의회도서관에서 반포한 표준과 어휘집으로 주제명, 이름 전거와 같은 통제어휘와 장르, 자원유형 등 다양한 데이터 값이 링크드데이터 형태로 구축되어 있다. OCLC의 VIAF, 국제표준이름식별자 ISNI도 이러한 어휘집의 일종이라 할 수 있다.

나. 온톨로지

서지레코드를 링크드데이터로 변환하기 위해 도서관 도메인에 특화된 온톨로지를 개발해야 한다. 미의회도서관에서 도서관 서지레코드를 위한 온톨로지로 BIBFRAME을 개발하였고, 선거통제를 위한 온톨로지로 MADSRDF를 사용하였다. OCLC에서는 온톨로지로 Schema.org, OWL, SKOS 등의 일반적 스키마를 사용하였다. 특히, LOV(Linked Open Vocabularies), OMR(Open Metadata Registry)에는 여러 다양한 메타데이터 온톨로지를 제공하고 있어 검색을 통해 원하는 온톨로지를 찾을 수 있다.

어휘집은 링크드데이터를 생성하려는 기관에서 이용할 수 있는 보편적 자원이 된다.

2. 어휘집과 온톨로지

메타데이터 스키마(metadata schema)는 데이터요소가 정의된 요소의 집합이다. 시맨틱웹에서는 데이터요소를 속성(properties)과 클래스(classes)로 정의하고, 메타데이터 스키마를 어휘(vocabulary) 혹은 온톨로지(ontology)라고 명명한다. 그러나, 어휘라는 용어는 데이터로 사용되는 통제어휘집에도 사용되고 있다(Coyle 2012b, 15). 메타데이터 스키마와 통제어휘

Ⅲ. BIBFRAME 구축 사례 분석

1. 미의회도서관 BIBFRAME 구축

가. 어휘집 및 온톨로지

미의회도서관은 링크드데이터 구축을 위해서 링크할 자원의 필요성을 인식하고, BIBFRAME 개발 이전에 기존 도서관이 구축한 이름전거, 주제명, 장소 등의 어휘집을 링크드데이터로 변환하기로 결정하였다. 2012년 URI를 포함하는 전거 및 통제어휘 리스트를 MADSRDF 표준을 사용하여 RDF로 변환하고, 이를 LDS라고 명명하였다. LDS에는 LC/NACO 전거파일과 LCSH를 비롯하여 RDA의 여러 가지 소규모 어휘리스트가 포함되어 있다.

어휘집 구축 이후 도서관 특화된 온톨로지로서 BIBFRAME을 2012년 개발하였는데, 이는 저작, 인스턴스 등의 서지 관련 개체를 표현하기 위해 개발된 클래스와 속성을 이용하여 RDF에 기반을 둔 그래프기반 데이터모델이다. 서지기술을 명확한 트리플 문장으로 해체하고, 텍스트나 문자 값 대신 LDS의 URI 사용하며, 내용규칙으로 RDA를 이용한다(McCallum 2017). BIBFRAME 온톨로지는 2016년 BIBFRAME 2.0으로 개정되었고, 이에 따라 BIBFRAME 클래스, 속성의 주요 개체도 저작, 인스턴스, 개별자료, 주제, 에이전트, 이벤트로 변경되었다.

나. 시스템 개발

LC의 Network Development and MARC Standards Office(NDMSO)에서 BIBFRAME 구현을 위한 프로그램으로 서지레코드를 BIBFRAME으로 변환하는 MARCXML to BIBFRAME Transformation(이하, BIBFRAME 변환기), BIBFRAME 형식에 맞게 입력하는 BIBFRAME Editor(이하, BIBFRAME 입력기)를 개발하였다. BIBFRAME 변환기는 MARCXML 서지레코드를 BIBFRAME으로 변환하고, BIBFRAME 입력기는 단행본, 악보, 연속간행물, 지도, 녹음자료, 동영상(블루레이 DVD), 동영상(35mm 필름), 화상자료의 8개 자료유형별 프로파일에 따라 입력한다.

1) BIBFRAME 변환기

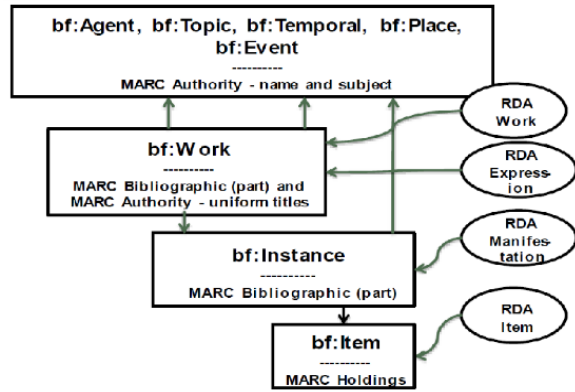
서지레코드의 일부와 전거레코드 중 통일표제는 BIBFRAME의 저작으로, 서지레코드의 일부는 BIBFRAME 인스턴스로, 소장 정보는 BIBFRAME 아이টে으로 매핑된다(<그림 2> 참조). 현재, MARC를 BIBFRAME 데이터로 변환하였지만 MARC 데이터와 BIBFRAME 링크드데이터를 각각 관리하고 있다(McCallum 2017, 80).

2) BIBFRAME 입력기

BIBFRAME 입력기는 MARC 태그 대신 RDA 요소명과 요소별로 연결된 RDA 규칙을 제

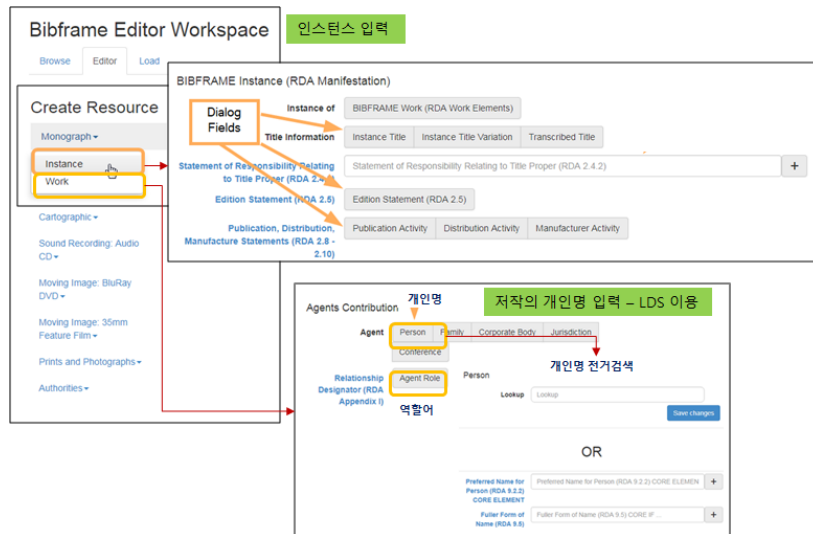
6 한국도서관·정보학회지(제49권 제2호)

공하고 있어 입력이 용이하다. <그림 3>과 같이 단행본, 악보, 연속간행물 등의 BIBFRAME 프로파일 유형을 우선 선택하고, 이후 저작, 인스턴스를 선택하여 입력한다. 만일 저작이 검색되면 기존 저작을 재사용하고 인스턴스 부터 입력을 시작하지만, 저작이 검색되지 않으면 저작의 채택표제, 저자, 저작의 형식, 저작 일자 등의 저작을 입력하고 인스턴스를 입력한다. 특히, BIBFRAME 입력기는 이름전기, 주제명 등을 검색할 수 있



<그림 2> MARC, RDA와 BIBFRAME 매핑
 ※출처: McCallum(2017)

는 LDS를 이용하여 링크드데이터를 생성하도록 모듈화되어 있어, 저작, 저자, 단체명, 역할, 지명 등을 입력할 때 LDS를 검색하여 해당 자원이 있으면 이를 선택하고, 없으면 직접 입력한다 (Carlton, Hawkins, and Frank 2017).



<그림 3> BIBFRAME 입력기
 ※출처: Carlton, Hawkins, and Frank(2017)

다. 테스트 및 교육

1) 1차 파일럿 테스트

BIBFRAME 1.0 환경에서 Network Development and MARC Standards Office와 Acquisitions and Bibliographic Access Directorate Divisions(이하, ABA)의 Cooperative

and Instructional Programs Division(이하, COIN)은 2015년 9월 8일부터 2016년 3월 31일까지 테스트를 진행하였다. 목표는 BIBFRAME 입력기로 단행본, 연속간행물, 지도자료, 음악악보, 녹음자료, 동영상, 화상자료 등의 다양한 포맷과 언어 자료를 BIBFRAME에 맞게 기술하는 것이었다. ABA내 Collections and Services Special Format Divisions 부서에서 약 40명 목록사서 및 테크니션이 참여해 여러 다양한 언어/문자, 포맷의 자료를 목록하였다. 사서들은 Voyager ILS 시스템을 통해 MARC 포맷과 BIBFRAME 입력기를 통해 BIBFRAME의 2가지로 데이터를 입력하면서 MARC와 BIBFRAME 데이터를 상호 비교하고, 데이터 간의 상호 연계성을 파악하였다. 사서들이 MARC 레코드를 작성하면서 동시에 BIBFRAME 데이터를 작성하였기 때문에 업무과정에서의 변경은 없었다(Library of Congress. Acquisitions and Bibliographic Access Directorate 2016).

2) 2차 파일럿 테스트

미의회도서관은 1차 파일럿 테스트의 경험, BIBFRAME을 테스트한 여러 기관의 의견, 메타데이터 전문 컨설턴트의 검토를 바탕으로 BIBFRAME의 클래스와 속성을 변경하여 BIBFRAME 2.0을 2016년 발표하였다. BIBFRAME 2.0의 발표와 함께 시작한 파일럿 테스트의 목표는 MARC를 참조하지 않고 BIBFRAME으로 데이터를 생성하는 환경을 구축하는 것으로 우선 전체 MARC데이터를 BIBFRAME으로 변환하고, 이후 BIBFRAME으로 데이터를 생성하는 것이다. MARC를 BIBFRAME 온톨로지로 변환하기 위해 변환 프로그램을 다시 작성하되 변환에 따른 데이터 손실이 없도록 상세한 변환 명세서가 마련되었고, 변환시 데이터 변환뿐만 아니라 데이터에 해당하는 URI도 제공되었다. 약 19,000,000 MARC 레코드의 변환이 완료되었지만 변환 명세서와 프로그램은 지속적으로 갱신되었다. 변환시 문제점은 RDA, BIBFRAME 데이터모델에 맞도록 에이전트, 토픽, 장소 등과 연계된 저작, 인스턴스, 아이টে므로 데이터를 만드는 것이다. 통일표제 전거파일을 유지하는 LC에서 통일표제 레코드는 BIBFRAME 저작으로 변환되고 추가적인 저작 기술은 서지레코드의 일부 데이터를 사용하며, 인스턴스도 서지레코드를 바탕으로 생성된다. 서지레코드에서 인스턴스 기술이 완성되며 인스턴스는 해당하는 저작과 연계되고, 인스턴스의 주제 정보는 인스턴스에서 저작으로 전달된다(McCallum 2017, 82-83).

3) 교육

2015년에 시작된 1차 파일럿에서는 16시간의 교육을 이수하도록 하였으며, COIN 부서의 사서 4명이 교육을 담당하였다. 교육내용은 시맨틱웹, 링크드데이터, BIBFRAME 입력기로 구성되었다. 주 단위 회의를 통해 참가 사서, 교육자, 개발자가 상호 의견을 교류하고 최선의 방안을 모색하였다. 사서는 경험을 바탕으로 BIBFRAME에 관련된 다양한 제안을 하였으며, 이러한 제안은 즉시 프로그램 개발에 적용되었다(Library of Congress. Acquisitions and

Bibliographic Access Directorate 2016). 2017년 6월부터 시작된 2차 파일럿에서는 링크드데이터, 시맨틱웹, BIBFRAME 2.0을 비롯하여 BIBFRAME 입력기와 데이터베이스에 대한 교육이 중심을 이루었다.

2. Linked Data for Production (LD4P)

LD4P는 Linked Data for Libraries라는 이전 멜론재단 기금을 받는 프로젝트가 확장된 것으로 도서관 자원의 링크드데이터 생성을 목표로 한다. 미의회도서관과 컬럼비아, 코넬, 하버드, 프린스턴, 스탠포드의 5개 대학교 도서관이 참여하였다. 2016-2018년 프로젝트의 주요 내용은 첫째, 링크드오픈데이터(LOD)로 메타데이터를 공동으로 생성하기 위한 표준, 가이드라인, 인프라를 개발하고, 둘째, 정보조직부서의 데이터 생성 환경 속에서 링크드데이터 생성을 위한 업무흐름도 개발하며, 셋째, 특정 도메인과 포맷의 도서관 자원을 기술하기 위한 BIBFRAME 온톨로지의 확장 방안을 마련하고, 넷째, 도서관 커뮤니티의 지속가능하고 확장 가능한 환경을 마련하는 것이다. 참여 도서관은 도서관 링크드데이터를 위한 온톨로지인 BIBFRAME을 사용하고, 이를 확장하는 방안을 각 특화된 도메인별로 연구하며, 메타데이터 생성한다. 또한, BIBFRAME 확장 모델링, 온톨로지 확장 모델링과 링크드데이터 생성을 위한 최선의 구축 사례, 링크드데이터 도구 평가, 링크드데이터 협력 생성을 위한 인프라 구조 프로토타입 개발을 공동으로 진행하고 있다(LD4P 2018).

6개 참여 도서관별로 프로젝트의 내용을 달성하기 위해 온톨로지 모델링, 메타데이터 생성, 특화된 영역별 온톨로지 모델링과 메타데이터 생성이라는 3가지 영역으로 나누어 연구를 진행한다(<표 1> 참조). 온톨로지 모델은 LC가 주축이 되어 BIBFRAME 온톨로지를 연구하며, BIBFRAME 온톨로지의 확장 모델에 관해서는 코넬대학교에서 bliotek-o 프로젝트를 진행하고, 메타데이터 생성에 관련하여 LC의 메타데이터 생성 파일럿과 스탠포드대학교의 트레이서블릿(Tracer Bullets)이 진행된다. 특화된 온톨로지 모델링과 메타데이터 생성의 경우 4개 기관에서 예술, 지도, 동영상, 연주음악, 희귀자료별로 모델링을 만들어 메타데이터를 생성한다. 컬럼비아 대학에서는 예술 분야의 도메인으로 예술 객체(화상자료, 3차원 실물)를 기술하기 위해 BIBFRAME 2.0에 예술 관련 속성을 확장한 ArtFrame을 개발한다. 하버드대학교는 지도자원 기술을 위한 BIBFRAME 확장 개발과 지도 메타데이터와 BIBFRAME 2.0의 매핑을 연구하며, 여성 감독의 영화자료 링크드데이터 기술도 함께 진행한다. 코넬대학교는 힙합컬렉션 중 LP로 생산된 비상업적인 자료 수백점을 RDF로 기술하고, 희귀본을 위한 BIBFRAME 2.0의 확장에 관해 연구한다. 프린스턴대학교는 철학자 데리다 개인 도서관의 16,000권의 도서나 자료 중에서 많이 인용된 자료를 대상으로 링크드데이터를 구축한다(LD4P 2018).

〈표 1〉 LD4P 프로젝트별 특성

구분	명칭	세부사항	
모델	BIBFRAME 온톨로지	미의회도서관	
	bibliotek-o	Cornell University Library	
메타데이터 생성	Stanford Tracer Bullets	Stanford University Library	
	BIBFRAME 메타데이터 생성과일럿	미의회도서관	
영역별 특화 온톨로지 모델링 및 메타데이터 생성	영역	모델	전문 장서별 메타데이터
	예술	ArtFrame	Columbia Art Properties
	지도	Cartographic Extension	Harvard Cartographic Materials
	동영상	Moving Image Extension	Harvard Film Archive
	연주음악	Performed Music Ontology	Cornell Hip Hop Archive
	희귀자료	Rare Materials Ontology Extension	Cornell Hip Hop Archive Princeton Derrida Archive

※출처: LD4P(2018)

가. Cornell의 bibliotek-o

1) 온톨로지 특징 및 모델링 패턴

① 온톨로지 특징

bibliotek-o는 BIBFRAME 온톨로지에 기반한 서지 메타데이터 모델링을 위한 프레임워크로 BIBFRAME의 대안 모델로 제안되었다. 2014에 진행된 LD4L 프로젝트에서 BIBFRAME 1.0이 도서관 자원을 기술하는데 충분하지 않다고 분석하고, 임시 온톨로지를 개발하였다. 2016년 LD4P가 시작되면서 LD4 온톨로지 그룹에서는 BIBFRAME 2.0을 사용하기 위해 분석한 결과 BIBFRAME 2.0이 BIBFRAME 1.0 보다는 향상되었으나 BIBFRAME 모델링의 수정이 필요하다고 판단하여 bibliotek-o를 개발하였다.

bibliotek-o 프레임워크의 핵심은 BIBFRAME이고, BIBFRAME 기반 위에 구축되며, BIBFRAME 패턴 없이 실행될 수 없지만, bibliotek-o 온톨로지는 모델링 원칙과 패턴에서 BIBFRAME과 다음의 차이가 있다. 첫째, 데이터 교환과 상호운용성을 위해 기존 외부 어휘를 재사용한다. 예를 들어, bf:suject를 사용하기 보다는 이미 정의된 dcterms:subject를 사용한다. 둘째, 외부 데이터 소스에서 재사용할 수 있도록 용어를 폭넓게 정의한다. 셋째, 도메인과 범위와 같은 OWL, RDF의 규칙을 이용하여 온톨로지와 데이터를 과도하게 제한하지 않고 표현성을 제공한다. 넷째, 객체 속성, 구조화된 데이터, 비구조화된 문자값 보다 통제어휘를 선호한다. 다섯째, 데이터를 표현할 수 있는 단순하고 통일된 모델을 선호하며, 질의 경로(query path)를 최소화하기 위해 관계나 속성을 표현하는데 하나의 방법을 채택한다. 여섯째, 클래스 ex:CartographicMovingImage 대신 bf:MovingImage와 bf:Cartography와 같이 복합 데이터 표현보다 가장 작은 단위로 데이터를 표현하는 것을 선호한다(Kovari, Folsom, and Younes 2017, 55).

BIBFRAME과 비교할 때 가장 큰 차이점은 기존 온톨로지의 재사용이라고 할 수 있다.

BIBFRAME에서는 기존 온톨로지를 재사용하지 않고 모두 새롭게 개발하였지만 bibliotek-o는 BIBFRAME 핵심 온톨로지를 기반으로 bibliotek-o(bibliotek-o 온톨로지는 bib로 표현), CIDOC-CRM, DCMI Terms, FOAF, Lingvo, Prov-o, RDAU, Schema.org, SKOS, VIVO, Web Annotation(LD4L Labs and LD4P Ontology Group 2017) 등의 온톨로지를 재사용하였다.

② 모델링 패턴

BIBFRAME은 공급활동(배포, 제작, 출판, 생산)과 창작활동간 차이를 두어 공급활동은 인스턴스, 창작활동은 저작에 적용되도록 하였다. LD4 그룹에서는 저작권 관련 활동과 인스턴스 관련 활동 간의 구분이 완전하게 지속되지 않는다는 것을 인식하고, 이러한 구분이 불필요하며 과도하게 복잡한 결과를 낼 수 있다고 판단하였다. 이에 활동패턴을 정의하고 이는 에이전트와 저작, 인스턴스, 아이템을 링크하는 bib:Activity 클래스의 하위클래스 bib:AuthorActivity, bib:PublisherActivity 등으로 명확히 역할에 대해 규정하고 공급활동과 창작활동의 구분을 삭제하였다. 이를 통해 활동 관계는 이벤트와 같이 자원의 다른 유형으로 확장이 가능하게 되었다(Kovari, Folsom, and Younes 2017, 55). 또한, BIBFRAME에서 내용유형, 수록매체유형, 매체유형을 bf:content/bf:Content, bf:carrier/bf:Carrier, bf:media/bf:Media로 모델링하였으나 bibliotek-o에서는 자원이 특별한 종류임을 선언하도록 rdf:type을 이용한다(Cornell University Library 2017).

2) 변환기 및 에디터

MARC와 bibliotek-o 매핑 명세서를 작성하고, bibliotek-o에서 사용되지 않는 BIBFRAME의 속성 리스트도 제공한다. 현재 bibliotek-o 프레임워크 기반 RDF 메타데이터 생성과 기존 메타데이터 변환 도구를 개발하고 있다. 변환기는 오픈소스로 전통적인 서지 메타데이터를 RDF와 링크드데이터로 변환하는 도구이며, 입력기는 VitroLib라는 온톨로지 기반 에디터로 Vitro라는 의미적 웹 온톨로지와 인스턴스에 기반한다(Kovari, Folsom, and Younes 2017, 59).

나. Stanford의 Tracer Bullets

1) 링크드데이터 생성 및 검색환경 구현

스탠포드 도서관의 링크드데이터 생성 프로젝트는 기술부서의 업무흐름에 중점을 둔 것으로 업무를 링크드데이터 생성 환경으로 변화시키기 위해 링크드데이터 생성 각 단계별로 업무과정을 검토한다. 따라서 핵심은 링크드데이터 생성과정을 순서대로 제시하면서 링크드데이터 생성에 따른 업무과정상의 변화를 파악하기 위해 모든 업무 내용을 추적한다. 특히, 원목, 카피목록, 자원의 디포짓, 컬렉션 자원의 디포짓 4가지 측면에서 업무과정의 변화를 제시한

다. 이를 통해 링크드데이터 생성을 위한 생성 워크플로우를 제시하고, 링크드데이터의 변환과 원목을 위한 도구를 제안하며, 링크드데이터 변환 전후 개선을 위한 최선의 방안을 검토하고, 자원을 링크드데이터로 기술한다.

특히, 벤더가 제공하는 카피목록과 원목 업무과정을 분석하고, 벤더제공 레코드의 링크드데이터로 변환시 업무과정을 제시하였다. RDF 이용을 위해 업무에 기반한 검색 환경의 요건을 마련하고, 원목을 위한 BIBFRAME 프로파일을 평가하였다. BIBFRAME 변환을 향상시키기 위해 벤더 제공 MARC 데이터를 개선하는 방안에 대해 벤더와 논의 중이다. Tracer Bullet 1에서는 LC의 변환프로그램을 이용해 MARC 레코드 38,000개를 BIBFRAME으로 변환하고, Blazegraph 그래프로 트리플로 표현하였으며, 자동 스크립트를 통해 Blacklight solr¹⁾ 환경에서 색인을 작성하였다. Tracer Bullet 2에서는 BIBFRAME 2.0 Editor로 50개 자원의 원목을 작성하고, Tracer Bullet 3에서는 CEDAR RDF Editor로 디지털 자원 30개 이상의 원목을 작성할 것이다. 뿐만 아니라 MARC 레코드를 BIBFRAME으로 변환하고, 색인 처리를 위한 자동화 방안을 모색할 것이다.

링크드데이터와 MARC 데이터가 혼용된 상태에서 인스턴스 기반의 검색 환경을 개발하고, 도서 자료를 대상으로 BIBFRAME 2.0에서 solr로 매핑을 작성하였다. 현재, 디지털자원을 RDF에서 solr로 변환하는 매핑 프로그램을 개발하였고, BIBFRAME 2.0에서 solr로 매핑 문서를 작성하고 있다(Stanford University Library 2018a).

2) BIBFRAME 변환 및 입력 도구 조사

링크드데이터 변환 및 생성프로그램의 요건을 조사하고, <표 2>와 같이 관련 프로그램의 테스트를 실시하였다. BIBFRAME 관련 도구로 LC BIBFRAME 입력기 및 변환기, 검색도구 SHARE Virtual Discovery Environment, 기타 도구로 ALIADA, CEDAR, KARMA, RML Editor, VitroLib, Tableau의 다양한 변환기, 편집기, 검색기 등이 있다. 뿐만 아니라 MARC-to-RDF 변환기를 위한 요건을 개발하였으며, LC의 BIBFRAME 2.0 입력기를 이용해 인스턴스를 생성하였다.

<표 2> 테스트 중인 도구(Tools)

BIBFRAME 관련 도구		
도구명	유형	기능
BIBFRAME 1.0 Converter	변환기	MARC에서 BIBFRAME 1.0으로 변환
BIBFRAME 2.0 Converter	변환기	MARCXML에서BIBFRAME 변환
BIBFRAME Editor	입력기	BIBFRAME으로 RDF 구조화함

1)open source enterprise search platform으로 자바로 작성

BIBFRAME Profile editor	킨피그레이션에디터	LC의 BIBFRAME 에디터를 이용하기 위한 프로파일을 구조화함
marc2bibframe-wraper	marc2bibframe 변환 유틸리티	대규모의 배치 레코드를 marc2bibframe 유틸리를 이용해 한번에 변환함. 한번에 처리되며, 개별 레코드의 에러는 별도의 XML 로그 파일에 저장됨
SHARE Virtual Discovery Environment	검색도구	2,029,218서지레코드와 3,601,327 전거레코드를 처리함. Casalini사와 @Cult사에서 참가도서관에서 데이터를 받아서 검색환경에 맞게 생성
기타 RDF 도구		
도구명	유형	기능
ALIADA	변환기, 출판사, 링커	MARC에서 커스터머된 RDF로 변환, 저장하고 다른 데이터셋과 링크
CEDAR	에디터	온톨로지에 기반하여 RDF 생성/편집
KARMA	멀티포맷 변환기	텍스트파일, XML, JSON, KML, Web APIs의 제한없이 데이터베이스, 스프레드시트를 포함하여 여러 데이터 소스로 부터 데이터를 빠르고 쉽게 통합하도록 함
RML Editor	변환기, 데이터 모델러	여러 종류의 데이터를 모델링화하는 GUI 매핑 언어로 RML을 이용하고, 다수의 이기종의 데이터 소스에서 매핑을 편집하도록 통일된 GUI를 제공
VitroLib	에디터	도서관목록을 위한 Vitro에디터
Tableau	시각화	데이터 시각화

*출처: Stanford University Library(2018b)

지금까지 살펴본 사례를 종합하면, 링크드데이터 구축을 위해 어휘집의 링크드데이터 변환, 온톨로지 개발 및 적절한 온톨로지 선택, 시스템 개발, 실재 데이터 입력의 과정이 있었다.

LC에서 링크드데이터 구축의 첫 단계로 LDS라는 통제어휘집을 링크드데이터로 제작하고, 서지레코드를 링크드데이터로 변환하기 위한 온톨로지로 BIBFRAME을 개발하였다. 이를 바탕으로 서지레코드를 링크데이터로 변환할 수 있는 BIBFRAME 변환기와 입력기를 개발하였으며, 목록사서 교육과 함께 파일럿 테스트를 진행하였다. 코넬에서는 BIBFRAME 확장의 필요성에 따라 bibliotek-o라는 온톨로지를 개발하고, 별도의 입력기와 변환기를 개발하였다. 스탠포드에서는 BIBFRAME 온톨로지를 그대로 적용하면서 Tracer Bullet이라는 구축 테스트를 통해 사서의 업무과정을 분석하였으며, 입력기, 변환기, 검색기 등의 도구를 테스트하였다(<표 3> 참조).

<표 3> 사례 분석 종합

분석항목	LC	LD4P	
		Cornell	Stanford
어휘집	LDS 구축		
온톨로지	BIBFRAME	bibliotek-o	BIBFRAME
시스템개발	BIBFRAME 변환기, 입력기	VitroLib 입력기, 변환기	다양한 검색도구 테스트
데이터 구축	파일럿 1, 2 진행		Tracer Bullet 1, 2

이러한 사례를 바탕으로 국내에서 BIBFRAME 적용방안을 모색한다면 먼저, 링크드데이터 구축의 전제조건인 어휘집 구축, BIBFRAME 온톨로지 적용 및 시스템 개발, 마지막으로 이를 실질적으로 구축하기 위해 목록사서의 교육 측면으로 나누어 살펴볼 수 있을 것이다.

IV. 국내에서 BIBFRAME 적용방안

1. 국내 어휘집의 링크드데이터로 변환 구축

링크드데이터 구축을 위해서는 링크할 자원이 필요하기 때문에 어휘집이 먼저 구축되어야 한다. 이에 국내에서도 링크드데이터 구축을 위해 이름전거, 주제명 등을 링크드데이터로 구축한 어휘집을 개발해야 하는데 우선적으로 필요한 어휘집의 종류를 규정하고 구축 방안을 모색해야 한다.

구축할 어휘집의 종류를 파악하기 위해 링크드데이터를 구축한 기관들이 사용한 어휘집을 조사한 결과 LD4L에서는 OCLC Fast, OCLC Works, Worldcat, ORCID, VIAF, LoC Subject Authority, LoC Name Authority, AGROVOC를 사용하였다. Smith-yoshimura (2016a)에 따르면 현재 링크드데이터 출판(publishing)시 가장 많이 사용되는 어휘집은 Europeana, Getty vocabularies, LC의 LDS, 일본국회도서관 NDL 순이며, 이외에도 OCLC WorldCat, VIAF, ISNI 데이터가 주로 이용되었다. 링크드데이터 이용(consuming)시 가장 많이 이용하는 어휘집은 VIAF > DBpedia > GeoNames > id.loc.gov > 자체에서 링크드데이터로 변환한 자원 > Getty's Art and Architecture Thesaurus > FAST Linked Data > WorldCat.org > data.bnf.fr > DNB Linked Data Services 순이다. DBpedia (위키에서 작성된 정보를 구조화한 지식베이스, <https://wiki.dbpedia.org>), GeoNames(무료 지명데이터베이스, <http://www.geonames.org>)를 제외하고는 도서관 분야의 자원을 이용하고 있다.

〈표 4〉 주요 기관의 링크드데이터 구축시 사용하는 어휘집

어휘집	VIAF	LC	FAST	WorldCat	BNF	DNB
AGROVAC		○			○	
BNF		○			○	
DBpedia			○	○	○	
DNB		○			○	
FAST				○		
GeoNames	○		○		○	
ISNI	○				○	
LDS	○	○	○		○	
VIAF		○	○	○	○	
WIKIDATA	○	○				
WorldCat.org	○			○		
WorldCat Work	○	○				
기타		○			○	

※출처: Smith-Yoshimura(2016b)의 내용을 기반으로 작성

링크드데이터 구축하는 6개 주요 기관 VIAF, LC, FAST, WorldCat, BNF, DNB에서 사용하는 어휘집을 분석한 결과 <표 4>와 같이 VIAF, LDS를 가장 많이 사용하는 것으로 나타났다. 대부분이 이름전거 및 주제명을 링크드데이터로 변환한 어휘집이었으며, WorldCat의 저작정보도 사용되었다. 뿐만 아니라 DBpedia, GeoNames, WIKIDATA와 같이 도서관 이외의 분야에서 구축한 자원을 링크드데이터 구축시 사용하고 있었다. 특히 대부분의 국가도서관 프로젝트나 서비스에서는 링크드데이터로 구축한 서지와 전거 데이터에 접근을 제공한다. 영국국가도서관은 링크드데이터로 국가서지를 작성한 최초의 도서관이며, ISNI, VIAF 식별자와 링크를 제공한다. 독일국가도서관은 링크드데이터를 구축하는 4개 프로젝트를 수행하는데 독일국가서지, 독일전거파일(GND), BIBFRAME 프로토타입, 엔티티 팩트(Entity Facts)이다(Smith-yoshimura 2016a).

이름전거, 주제명의 핵심 어휘집 이외에 필요한 통제어휘집을 파악하기 위해 LDS에서 제공하는 어휘집을 조사한 결과 분류표, 수록매체, 내용유형, 매체유형, 자원유형, 출판간기 등 다양한 소규모 용어집도 구축되어 있다(<표 5> 참조). 이러한 어휘집은 저작명, 저자명, 저지역할어, 지역명, 주제명, 분류번호 등의 데이터 입력시에 활용할 수 있다.

<표 5> LDS 어휘집

어휘집	어휘집	스키마 및 소규모 용어집
LC Subject Headings	MARC Relators	Identifiers
LC Name Authority File	MARC Countries	Carriers
LC Classification	MARC Geographic Areas	Content Types
LC Children's Subject Headings	MARC Languages	Media Types
LC Genre/Form Terms	MARC Genre Terms	Resource Types
LC Medium of Performance	ISO639-1 Languages	Description Convention
Thesaurus for Music	ISO639-2 Languages	Publication Frequencies
LC Demographic Group Terms	ISO639-5 Languages	Resource Components
Thesaurus for Graphic Materials	Extended Date/Time Format	BIBFRAME Ontology
AFS Ethnographic Thesaurus		MARC Genre/Form Schemes
Cultural Heritage Organizations		MARC Subject Schemes
		Classification Schemes

그러므로 국내에서 BIBFRAME을 구축하기 위한 첫 단계는 이름전거, 주제명 등 기존 전거데이터를 링크드데이터로 생성하는 것이다. 현재, 국립중앙도서관에서 구축한 링크드데이터를 활용할 수 있으며, 앞으로 국내 대학에서 구축해온 전거데이터에 대한 공개를 바탕으로 링크드데이터로 구축한다면 이를 공유할 수 있을 것이다. 국내에서 전거데이터의 통합구축시에는 VIAF와 같은 모델이 가장 적합한 형태이며, 이를 통해 개별 도서관에서는 링크드데이터 형태로 전거를 구축할 수 있을 것이다. 또한, 이름전거와 주제명 이외에도 도서관에서 통제어휘로 사용하거나 데이터 값으로 사용할 용어리스트를 링크드데이터로 구축해 공유해야 한다. 예를 들어, KDC를 비롯해 KORMARC에서 사용되는 발행국부호표, 한국대학부호표, 언어구분부호표, 한국정부기관부호표, 한국지역구분부호표, 외국지역구분부호표, 한국도서관

부호표, 역할어 부호를 링크드데이터로 구축해야 한다.

2. BIBFRAME 온톨로지 적용 및 시스템 개발

링크드데이터를 구축하기 위해서는 온톨로지를 사용해야 하는데 OCLC에서 사용중인 schema.org, SKOS, FOAF 등은 도서관 이외의 타 분야에서 주로 개발된 온톨로지로서 도서관 특화된 데이터를 기술하는데 한계가 있다. 반면, BIBFRAME은 도서관 특화된 온톨로지로서 도서관에서 필요한 데이터를 기술하도록 MARC를 모두 수용하고 있어 앞으로 도서관 분야에서 널리 사용될 것이다.

도서관 링크드데이터를 구축한 곳에서 사용되는 온톨로지를 살펴보면 LD4L에서는 BIBFRAME을 중심으로 Bibliographic Ontology(이하, BIBO), Open Annotations, VIVO Ontology, IIF, ORE를 사용하였고, bibliotek-은 BIBFRAME, dcterms, oa, OWL, prov, RDF, RDFs, schema-org, SKOS, vann, voaf, xmlschema를 사용하였다. Smith-Yoshimura (2016a)의 설문조사에서 도서관 링크드데이터를 구축하는 기관에서 가장 많이 사용하는 온톨로지는 SKOS, FOAF, DCMES, Schema.org, BIBO, RDA, EDM, ISBD, BIBFRAME, FRBR, Event ontology, MADS, CIDOC CRM 순이었다. 특히, 링크드데이터 구축하는 6개 주요 기관에서 사용되는 메타데이터 스키마를 분석한 결과 <표 6>과 같이 FOAF, SKOS, DC, schema.org가 주로 사용되었으며, BIBFRAME은 미국과 독일에서 사용되고 있었다.

국내 대표적인 링크드데이터 구축 도서관은 국립중앙도서관, KERIS, KISTI이며(한국정보화진흥원 2014), 이들 기관에서는 서지 관련 온톨로지로서 DC, BIBO를 사용하지만, 이러한 온톨로지는 도서관 특화된 요소를 충분히 반영하지 못하고 있어 자체적으로 개발한 온톨로지를 함께 사용한다. 궁극적으로 링크드데이터는 소스데이터와 링크드데이터의 매핑문제 및 어휘의 매핑이 이루어져야 하므로 표준화된 방법으로 데이터를 구축하고, 표준 온톨로지를 사용해야 하는데 자체 개발 온톨로지는 이러한 측면에서 적합하지 않다. BIBFRAME이 미국을 비롯하여 독일에서 사용되고 있고, 도서관에서 구축해온 서지데이터에 부합하는 온톨로지이므로 도서관에서는 BIBFRAME을 이용한 링크드데이터를 구축하는 것이 바람직하다.

그러므로 국내에서 BIBFRAME을 적용하기 위해 BIBFRAME의 적합성을 판단하고, 국내에 적합하게 확장 모델링을 개발할 필요가 있다. 뿐만 아니라 BIBFRAME과 KORMARC과의 매핑 명세서를 작성하고, 변환기를 개발해야 하며, 입력을 위한 입력기 개발도 요구된다. 이를 위해서는 현재 GitHub에 있는 LC 개발 변환기 및 입력기를 테스트해야 한다.

〈표 6〉 주요 기관의 링크드데이터 구축시 사용하는 온톨로지

온톨로지	VIAF	LC	FAST	WorldCat	BNF	DNB
BIBFRAME		○				○
BIBO	○				○	○
Biographical Ontology					○	
DC & DC terms	○		○	○	○	○
FOAF	○	○	○	○	○	○
FRBR					○	
ISBD						○
ISNI					○	
MADS/RDF		○				
Music Ontology					○	
OAI ORE Terms					○	
OWL	○				○	○
RDA					○	○
RDF schema	○	○			○	○
schema.org	○		○	○		
SKOS	○	○	○	○	○	○
WSGS84 Geo Positioning			○		○	

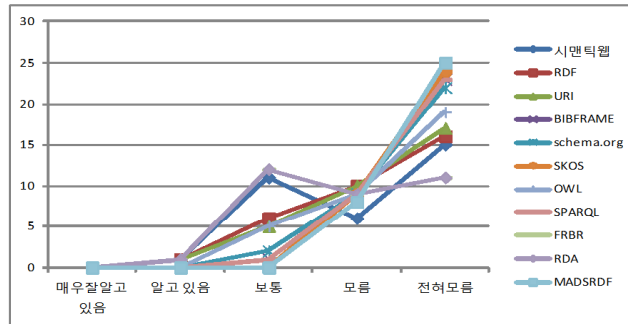
※출처: Smith-Yoshimura(2016b)의 내용을 기반으로 작성

3. 목록사서 교육

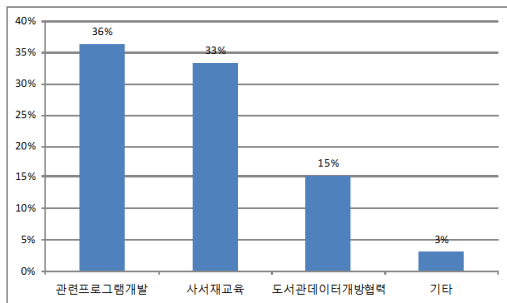
링크드데이터에 대한 국내 사서들의 링크드데이터 인식정도를 파악하기 위해 국립중앙도서관 목록교육에 참여한 33명의 사서를 대상으로 관련 용어 설문조사를 실시하였다.

링크드데이터 관련 용어 3개(시맨틱웹, RDF, URI), 온톨로지 5개(BIBFRAME, SCHEMA.ORG, SKOS, OWL, SPARQL), 목록관련 최신용어 3개(FRBR, RDA, MADSRDF)에 대해 이해정도를 5점 라이커트 척도로 표시하도록 한 결과 전반적으로 링크드데이터 관련 용어에 익숙하지 않은 것으로 나타났다(<그림 4> 참조). FRBR, RDA와 같은 목록 용어, 시맨틱웹, RDF, URI와 같은 링크드데이터 관련 일반 용어는 인식하고 있으나, BIBFRAME, schema.org와 같은 온톨로지에 대해서는 거의 알지 못하였다.

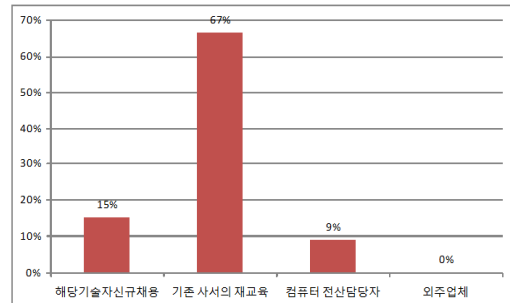
링크드데이터를 위한 데이터의 개방에 관한 질문에서는 개방이 필요하다고 인식하고 있었으며, 링크드데이터 구축을 위한 필수 요건에 대해서는 관련 프로그램 개발(36%) > 사서재교육(33%) > 도서관 데이터 개방 협력(15%) 순으로 나타났다(<그림 5> 참조). 링크드데이터의 새로운 기술을 담당할 직원의 수급에 대해서는 기존 사서의 재교육(67%) > 해당기술자 신규 채용(15%) > 컴퓨터전산담당자(9%) 순으로 답하였다(<그림 6> 참조). 이는 국외와 마찬가지로 링크드데이터 구축이 사서의 고유업무로 인식하고 있는 것이다.



<그림 4> 관련 용어에 대한 이해도



<그림 5> 링크데이터 구축을 위한 필수요건



<그림 6> 링크데이터 구축 담당자

앞으로 링크데이터가 사서의 고유업무가 되기 위해서는 사서를 대상으로 한 링크데이터 교육이 체계적으로 개발되어야 한다. 이는 학계를 포함하여 재교육 기관에서 링크데이터 교육이 상호 연계되어 체계적으로 이루어져야 한다. 학계에서는 이론을 바탕으로 교육하여 사서들이 SKOS, FOAF, RDF, XML 등의 개념을 올바르게 이해하도록 하며, 재교육 기관에서는 현장 적용이 가능하도록 실습위주의 교육을 마련해야 한다.

V. 결론

미의회도서관은 도서관 자료의 링크데이터 변환을 위해 도서관을 위한 특화된 온톨로지 BIBFRAME을 개발하고, LC의 전체 서지레코드를 링크데이터로 변환하는 테스트를 완료하였고, 독일국가도서관도 BIBFRAME을 이용해 링크데이터를 구축하고 있어 링크데이터를 위한 도서관 온톨로지로 BIBFRAME의 영향력이 매우 클 것으로 판단된다. 이에 본고에서는 BIBFRAME 구축을 위한 국내 도서관의 구체적인 적용방안을 마련하고자 문헌연구, 사례조사, 설문조사를 실시하였다. 문헌연구를 통해 국외 도서관에서 링크데이터 구축 및 실행과정을 살펴보고, 사례조사에서는 BIBFRAME을 구축한 미의회도서관과 LD4P 프로젝

트를 분석하였다. 이를 바탕으로 국내에서 BIBFRAME을 도입하기 위한 적용방안을 다음과 같이 3가지로 제시하였다.

첫째, 이름전거, 주제명 등의 기존 전거데이터를 링크드데이터로 출판하는 것뿐만 아니라 도서관에서 통제어휘로 사용하거나 데이터 값으로 사용하는 용어집도 링크드데이터로 생성해야 한다.

둘째, 국내에서 BIBFRAME의 적합성을 판단하고, 확장 모델링을 개발할 필요가 있으며, 이에 맞는 데이터 구축을 위해 KORMARC과의 매핑 테이블 작성과 변환기 및 입력기 등의 시스템 개발도 요구된다.

셋째, BIBFRAME과 같은 링크드데이터 구축이 사서의 고유한 업무가 될 수 있도록 사서를 대상으로 한 체계적인 링크드데이터 교육 프로그램이 개발되어야 한다.

링크드데이터 생성시 장애요인으로 데이터를 표현하는데 적절한 온톨로지의 선택, 시스템 구축에 관한 문헌이나 권고의 부재, 도구의 부족, 소프트웨어의 문제 등(Smith-yoshimura 2016a)이 지적되었는데 이는 링크드데이터 구축을 위해 해결되어야 할 주요 문제들이다. 특히, 링크드데이터 구축을 위해서는 실질적인 시스템 개발 이전에 적절한 온톨로지를 선택해야 하며, 이를 위해서는 국내에서 BIBFRAME의 적용가능성을 타진하는 다양한 실험연구들이 진행될 필요가 있다. 또한, 전거, 용어집과 같은 자원이 필요한데 이를 위해서는 도서관의 기본에 충실한 데이터 구축이 필요하다(이미화 2017, 124). 이러한 온톨로지 실험, 데이터 구축 및 변환을 통해 국내에서도 실질적인 링크드데이터 구축이 가능할 것이다.

참고문헌

- 이미화. 2017. BIBFRAME 2.0 특징 분석 및 BIBFRAME 구축시 고려사항에 관한 연구. 『한국도서관·정보학회』, 48(4): 107-127.
- 한국정보화진흥원. 2014. 『링크드오픈데이터 국내 구축사례집』. 서울: 한국정보화진흥원.
- Carlton, T., L. Hawkins and P. Frank. 2017. *Module2: The BIBFRAME Editor and the BIBFRAME Linked Data Store*. <<http://www.loc.gov/catworkshop/bibframe>> [cited 2018. 1. 15].
- Cornell University Library. 2017. *bibliotek-o : An Overview*. <<https://wiki.duraspace.org/display/LD4P/bibliotek-o+%3A+an+overview>> [cited 2018. 1. 15].
- Coyle, K. 2010. "Metadata Models of the World Wide Web." *Library Technology Reports*, 48(4): 12-19.
- Coyle, K. 2012a. "Semantic Web and Linked Data." *Library Technology Reports*, 48(4):

10-14.

- Coyle, K. 2012b. "Metadata Elements." *Library Technology Reports*, 48(4): 15-26.
- Godby, C. J., S. Wang and J. K. Mixer. 2015. *Library Linked Data in the Cloud*. California: Morgan and Claypool.
- Huerga, I. and M. P. Lauruhn. 2016. "Linked Data Implications for Authority Control and Vocabularies." *In Linked Data for Cultural Heritage*. Chicago: ALA.
- Kovari, J., S. Folsom and R. Younes. 2017. *Towards a BIBFRAME Implementation: the bibliotek-o*. <<http://dcpapers.dublincore.org/pubs/article/view/3854/2039>> [cited 2018. 1. 15].
- LD4L Labs and LD4P Ontology Group. 2017. *Bibliotek-o: a BIBFRAME Ontology Extension*. <<https://bibliotek-o.org/1.1/ontology.html>> [cited 2018. 1. 15].
- LD4P. 2018. *Linked Data for Production*. <<https://wiki.duraspace.org/pages/viewpage.action?pageId=74515029>> [cited 2018. 1. 15].
- Library of Congress. Acquisitions and Bibliographic Access Directorate. 2016. *BIBFRAME Pilot Phase One: Report and Assessment*. <<https://www.loc.gov/bibframe/docs/pdf/bibframe-pilot-phase1-analysis.pdf>> [cited 2018. 1. 15].
- Lovins, D. and D. Hillmann. 2017. "Broken-world Vocabularies." *D-Lib Magazine*, 23(3/4). <<https://doi.org/10.1045/march2017-lovins>> [cited 2018. 1. 15].
- McCallum, S. 2017. "BIBFRAME Development." *JLIS.it*, 8(3): 71-85. <<https://www.jlis.it/article/view/12415>> [cited 2017. 7. 15].
- Smith-Yoshimura, K. 2016a. "Analysis of International Linked Data Survey for Implementers." *D-Lib Magazine*, 22(7/8). <<http://www.dlib.org/dlib/july16/smith-yoshimura/07smith-yoshimura.html>> [cited 2018. 1. 15].
- Smith-Yoshimura, K. 2016b. *Linked Data Implementations: Who, What and Why?* <<http://campuspress.yale.edu/knowledgepermaculture/2016/12/14/oclc-linked-data-survey/>> [cited 2018. 1. 15].
- Stanford University Library. 2018a. *Stanford Tracer Bullets*. <<https://wiki.duraspace.org/display/LD4P/Stanford+Tracer+Bullets>> [cited 2018. 4. 15].
- Stanford University Library. 2018b. *Registry of Tools*. <<https://wiki.duraspace.org/display/LD4P/Registry+of+Tools>> [cited 2018. 4. 15].

국한문 참고문헌의 영문 표기

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

Lee, Mihwa. 2017. "A Study on the Considerations in Constructing BIBFRAME by Analyzing BIBFRAME 2.0." *Journal of Korean Library and Information Science Society*, 48(4): 107-127.

NIA. 2014. *Linked Open Data Construction Casebook in Korea*. Seoul: NIA.