

차세대 서지 기술 형식으로서의 BIBFRAME 모형 연구

BIBFRAME Model for the Next Generation

Bibliographic Description Format

박지영, 한성대학교 지식정보학부, zgpark@hansung.ac.kr

Ziyoung Park, Division of Knowledge & Information Science at Hansung University

본 연구에서는 웹 환경에 적합한 서지 기술을 위해 LC에서 개발한 BIBFRAME을 분석하고, BIBFRAME이 개방형 서지 데이터의 생성에 적합한지를 검토하였다. 그리고 BIBFRAME의 등장 배경 및 특징을 FRBR 모형과 비교하였다. 그 결과 BIBFRAME은 1)서지 객체의 지적 내용과 물리 매체를 구분하고 있었는데, FRBR의 제1그룹보다 단순한 구조를 지니고 있었다. 2) 인명전거와 주제전거를 포함하는 넓은 개념으로 전거를 정의하고 있었으며, 서지데이터 작성자의 관점과 이용자의 관점을 함께 고려하였다. 3) 다양한 외부 데이터와의 연계 및 개별 도서관에 특화된 데이터를 추가할 수 있는 주식 개념을 도입하였다. BIBFRAME은 MARC 레코드에서 변환하거나 직접적으로 서지데이터를 작성하는 방안이 모두 가능하며, 웹 환경에 적합하다고 판단된다.

1. 서론

도서관 분야에서는 웹 정보환경에 적합한 서지 정보를 생산할 수 있는 서지데이터 모형을 구축하려는 노력을 꾸준히 해 오고 있다. 이에 LC에서는 2011년에 Bibliographic Framework Initiative를 시작하여, 2012년에는 그 결과물인 Bibliographic Framework (BIBFRAME)를 발표했다. 그 후 MARC 데이터의 BIBFRAME 변환이나 모델링 관련 이슈에 대한 활발한 토론이 이루어지고 있다 (BIBFRAME LISTSERV, 2013).

이에 본 연구에서는 이 새로운 모형이 지닌 특징과 개방형 서지 형식으로서의 가능성을 검토하고자 한다. 우리 분야는 서지 정보의 기술을 위해 다양한 매체와 기술방식을 개발해 왔으며, BIBFRAME도 여러 서지 기술 방식 중에 하나이기 때문이다. 현재 BIBFRAME은 초안 단계로서 국내에서는 국가서지 개방을 위한 세미나 발표자료에서 간략히 언급된 정도이다(박진호, 2013).

2. 이론적 배경

2.1 MARC의 확산과 지속적인 개정

1960년대에 개발된 MARC 형식은 그 시대에 적합한 기술을 이용하여 데이터를 교환하고 공유하기 위해 개발되었으며, 카드목록 중심의 서지 기술 방식을 크게 변화시켰다. 현재까지도 MARC은 대부분의 도서관에서 기본 서지 기술 형식으로 사용되고 있으며, 관련 규정은 ISBD를 수용한 각국의 목록규칙에 따라 국가별로 개발되어 있다. 국내의 경우 서양서에는 MARC21 형식이 주로 사용되고, 동양어나 국내서에는 KORMARC 형식이 사용되고 있다. 또한 MARC 형식은 정보매체의 발달과 정보환경의 변화를 수용하여 지속적으로 개정되어 왔다. 최근에도 개정작업이 계속되고 있는데, 전자책이나 웹자원과 같은 새로운 매체를 기술하기 위한 규정을 추가하거나 새로운 목록규칙인 '자원의 기술과 접근'(RDA)을 수용하기 위한 필드를 추가하는 것에 중점을 두고 있다.

2.2 서지 기술 형식의 다양화

MARC의 지속적인 개정 작업이 이루어지고 있지만, 한편에서는 서지 기술을 위한 새로운 포맷이 개발되고 있다. 레코드 단위로만 교환이 가능한 MARC 데이터의 경직성이나 표시 기호나 식별기호 등의 가독성 문제, 비구조화된 텍스트 형식의 데이터 입력 등 MARC이 구조적으로 안고 있는 문제들은 부분 개정 작업으로는 해결하기 어렵기 때문이다. 이에 국내에서도 학위논문이나 PDF와 같은 전자파일은 MODS로 기술하는 도서관이 늘고 있으며, 전자저널이나 전자책은 외부 벤더가 별도의 메타데이터를 구축하여 도서관에 제공하고 있으므로 MARC을 사용하지 않는 경우가 많다. 즉, 단행본이나 도서관에서 구입하여 소장하고 있는 자료 외에는 MARC의 적용범위가 감소하는 추세이다.

3. BIBFRAME의 특징 및 구조

3.1 BIBFRAME의 목적

BIBFRAME에서는 서지 정보를 웹 정보환경에 적합하도록 구축하기 위해서 1) 개념적 콘텐츠와 물리적 구현물의 구분, 2) 정보 개체를 식별하기 위한 중의성 해소, 3) 개체 간의 관계 정보 표현 방안을 제시하였다.

따라서 앞으로는 지적 내용인 저작과 물리적 특징인 구현형을 구분하고, 기본표목과 같

은 문자열을 식별기호로 사용하는 대신 기계가 해석할 수 있는 식별기호를 사용해야 한다. 또한 기존의 서지 데이터에서는 부족했던 연관 데이터가 풍부해져야 한다. 그러면 도서관의 서지 데이터는 웹 환경에서 다양한 외부 자원과 연계가 가능한데, 예를 들면, 도서관의 대출기록과 아마존의 구매기록이 연계도 시도할 수 있다.

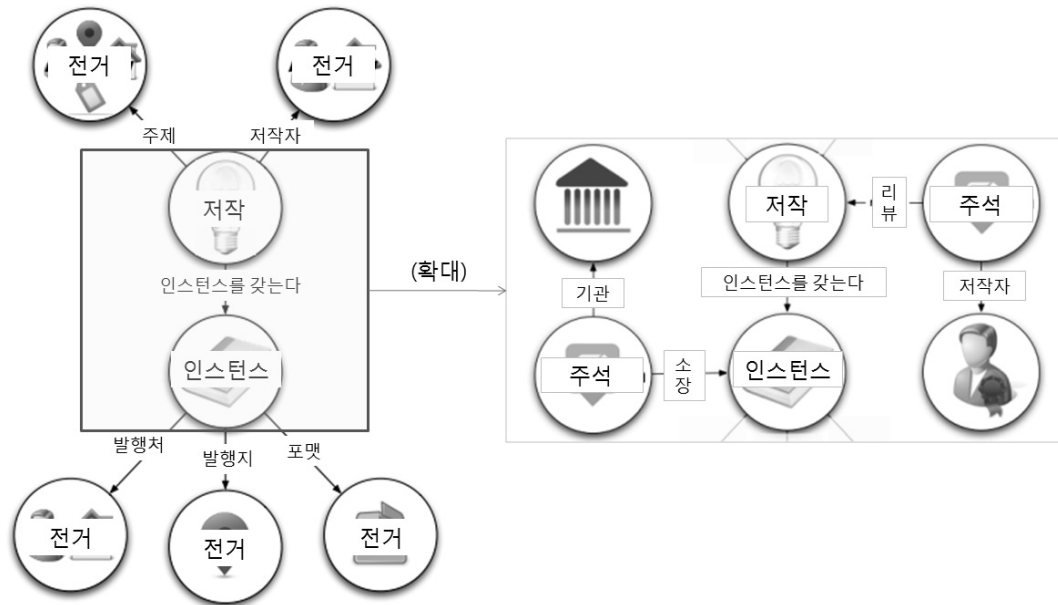
3.2 BIBFRAME의 구조

BIBFRAME은 FRBR 모형을 반영한 개체 관계 모형으로서, 서지 정보를 RDF로 표현하여 웹을 통해 식별하고 접근할 수 있게 한다. 기본구조는 4개의 기본 클래스를 통해 서지 정보에 필요한 자원을 구성하고, 각 클래스 간의 관계를 모형화하는 것이다(Library of Congress 2012)<표 1> 및 <그림 1> 참고). 우선 저작 클래스는 개념으로만 존재하는 것으로서, 복수의 인스턴스 클래스를 모을 수 있는 웹기반의 제어점(control point)로만 존재한다. 또한 저작은 주제나 저작의 생산에 관여한 개체를 속성으로 가지며, 다른 저작과 부분-전체 관계 등을 갖는다.

인스턴스 클래스는 특정 매체에 담긴 개별 저작이므로 물리적 대상이나 디지털 상태로 존재한다. 저작과 마찬가지로 웹기반 제어점으로 존재한다. 그리고 인스턴스의 속성은 발행이나 생산, 배포와 관련된 전거 클래스와 관계를 맺을 수 있다.

<표 1> BIBFRAME 모형의 기본 클래스(Library of Congress 2012, 8)

클래스명	특징
저작(Creative Work)	목록작성대상 자료의 개념적 특징을 반영하는 자원
인스턴스(Instance)	매체에 담긴 개별 저작을 반영하는 자원
전거(Authority)	저작과 인스턴스가 지닌 관계들을 정의한 핵심 전거 개념을 반영하는 자원. 사람, 장소, 주제, 기관 등을 포함함.
주석(Annotation)	부가 정보를 통해 다른 BIBFRAME의 자원을 수식하는 자원. 도서관 소장 정보, 표지 그림, 리뷰 등을 포함함.



<그림 1> BIBFRAME 클래스간의 관계 (Library of Congress 2012, 9-12)

전거 클래스는 저작과 인스턴스 내의 관계를 나타내는데, 발행처나 발행지 뿐 아니라 포맷까지 포함되는 넓은 개념으로 사용되고 있었다. 또한 다양한 관점을 고려했는데, 목록작성자의 관점에서 전거는 중의성 해소와 유사한 개념의 동조에 사용된다. 그런데 이용자의 관점에서 전거는 관련 저작과 인스턴스를 향해하고 맥락 정보를 얻기 위한 제어점이 된다는 점을 고려하였다.

주석 클래스는 리뷰나 논평 등을 포함하여 전문가가 생성한 메타데이터 뿐 아니라, 이용자가 생성한 메타데이터도 수용하기 위한 것이다. 다양한 부가 정보를 표현할 수 있는 클래스로서 도서관 분야가 아닌 외부에서 생산된 정보도 허용한다. 주석 클래스는 링크드 데이터 원칙에서 강조하는 다양한 링크 정보를 표현하는데 적합하며, 이 클래스를 활용하여 자관에 특화된 정보를 구축할 수도 있다.

3.3 BIBFRAME과 FRBR

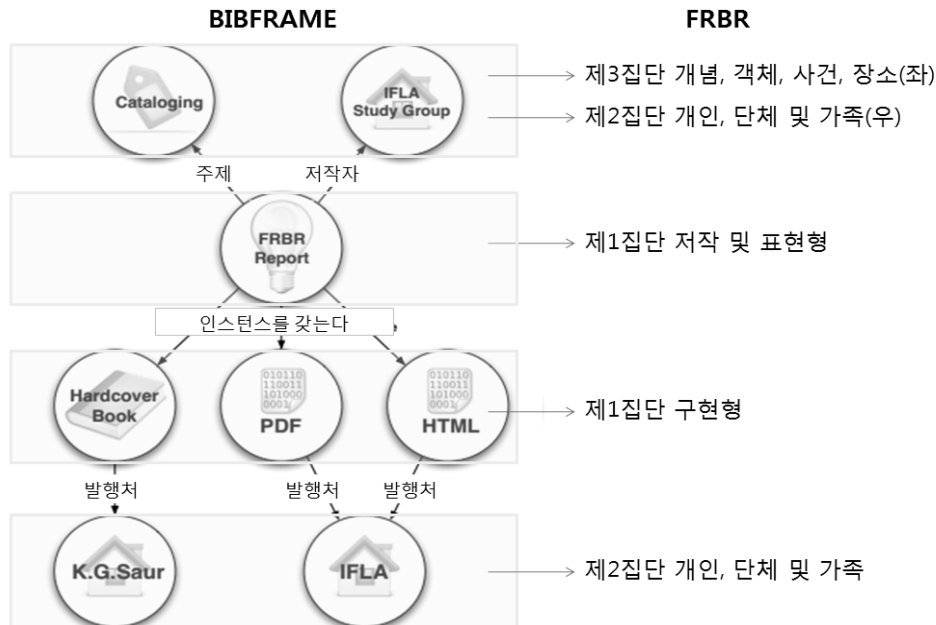
이 절에서는 BIBFRAME의 예시를 살펴보

고, 이와 함께 BIBFRAME 모형과 FRBR 모형을 비교하였다. <그림 2>는 *FRBR Report*를 BIBFRAME 모형으로 표현한 것인데, 하나의 저작이 3개의 상이한 인스턴스를 가지고 있다. 그리고 주제와 저작자 및 발행처를 나타내는 전거 클래스와 연계되어 있다.

또한 BIBFRAME은 이 예시를 RDF 구문으로도 제공하고 있는데, RDF 구문을 확인해보면 BIBFRAME의 저작은 FRBR의 저작과 표현형에 해당한다. 표제와 저작자, 언어 정보가 저작에 포함되어 있기 때문이다.

그리고 인스턴스는 FRBR의 구현형에 해당되며, BIBFRAME 클래스에서 제외된 FRBR 제1집단의 개별자료 개체는 BIBFRAME의 주석 클래스를 이용하여 소장정보를 추가하는 방식으로 표현할 수 있다.

또한 전거 클래스 중 저작자와 발행처는 FRBR의 제2그룹에, 주제와 장소는 FRBR 제3그룹에 해당된다. 즉, BIBFRAME에서는 인명전거를 개인이나 단체, 가족으로 구분하지 않으며, 주제전거를 별도의 클래스로 구분하지 않고 전거 클래스에 통합하였다.



<그림 2> BIBFRAME과 FRBR의 비교 (Library of Congress 2012, 19; IFLA 1998)

마지막으로 주석 클래스는 FRBR에서 개체의 속성으로 간주된다. BIBFRAME은 효과적인 데이터 연계를 위해 주석 정보를 개별 클래스로 독립시킨 것이다.

4. 결론

이 연구에서는 차세대 서지 기술 형식으로서의 BIBFRAME의 특징과 기본 구조를 살펴보고, FRBR 모형의 기본 개체와 비교하였다.

그 결과 BIBFRAME은 RDF 구문을 기반으로 하며, 저작과 인스턴스를 구분하여 저작의 개념적 요소와 물리적 요소를 분리하였다. 또한 전거 클래스를 통해 개체를 식별할 수 있도록 했는데, 이름이나 주제, 매체까지 포함하는 넓은 개념이었다. 또한 주석 클래스를 통해 기존의 서지 데이터 및 외부 데이터와 연계할 수 있거나 개별 기관에 특화된 데이터를 추가할 수 있도록 하였다.

아직은 BIBFRAME이 초기 단계이고, 앞으로 도 도서관의 기본 기술 형식인 MARC은 유지

될 것으로 보인다. 그러나 다양한 서지 기술 형식을 고민하고, 우리 분야의 경계를 넘어서 여러 분야와 데이터를 공유하기 위한 가능성을 모색하는 일은 더욱 확대될 것으로 판단된다.

참고문헌

BIBFRAME LISTSERV (Bibliographic Framework Transition Initiative Forum) (2013). <LISTSERV.LOC.GOV>.

IFLA (1998). Functional Requirements for Bibliographic Records. IFLA Study Group on the FRBR. (As amended and corrected through February 2009).

Library of Congress (2012). Bibliographic Framework as a Web of Data Washington, DC: Library of Congress.

박진호 (2013). 국가서지정보 Open Public Data Platform 구축. 공공정보 활용 사업화 전략 세미나 발표자료. 2013년 3월 29일, 대전시청.