

링크드 데이터 방식을 통한 서지 정보의 확장에 관한 연구*

Extending Bibliographic Information Using Linked Data

박지영(Ziyoung Park)**

초 록

본 연구에서는 서지 정보를 확장하기 위한 방안으로 링크드 데이터를 선정하였다. 링크드 데이터는 웹 공간을 통해 공유 가능한 식별기호와 데이터 구조 및 링크 정보를 제공하기 때문이다. 특히 링크드 데이터는 서지 온톨로지와 결합하여 서지데이터를 확장시키는데 유용하다. 이에 링크드 데이터와 서지 온톨로지를 분석하고, 연계 가능한 링크드 데이터를 검토하였다. 그리고 이를 바탕으로 링크드 데이터로 구축된 기존의 전거 데이터 및 서지 데이터를 연계하였다. 이러한 실험적 연계를 통해 향후 링크드 데이터를 효과적으로 활용하기 위한 과제를 도출할 수 있었다. 즉, 1) 다양한 링크드 데이터 중에서 각 기관에서 적합한 데이터를 선정할 수 있어야 하며, 2) 선정된 링크드 데이터를 연계하기 위한 기준을 정립해야 하고, 마지막으로 3) 자관의 고유한 데이터를 개발하여 이를 다시 공유해 나가야 할 것을 제안하였다.

ABSTRACT

In this study, Linked Data was used for extending bibliographic data, because Linked Data provides shareable identifiers, data structures, and link information. Linked Data is especially efficient in expanding bibliographic data integrated with bibliographic ontology. Therefore, Linked Data and bibliographic ontologies were analyzed and available Linked Data was suggested. By linking between meta-data schemes, bibliographic data, and authority data, issues for the effective Linked Data sharing were suggested: 1) selecting proper Linked Data for each bibliographic organization, 2) linking between different Linked Data, and 3) developing their own Linked Data for each bibliographic organization.

키워드: 링크드 데이터, 서지 온톨로지, 전거 데이터

Linked Data, Bibliographic Ontology(Bibo), FRBR-Aligned Bibliographic Ontology (FaBiO), DBpedia, open library, Faceted Application of Subject Terminology (FAST), Virtual International Authority File(VIAF), British National Bibliography(BNB)

* 본 연구는 한성대학교 교내연구장려금 지원과제임.

** 한성대학교 지식정보학부 전임강사(zgpark@hansung.ac.kr)

■ 논문접수일자: 2012년 2월 22일 ■ 최초심사일자: 2012년 2월 28일 ■ 게재확정일자: 2012년 3월 9일
■ 정보관리학회지, 29(1), 231-251, 2012. [http://dx.doi.org/10.3743/KOSIM.2012.29.1.231]

1. 서론

도서관 목록을 둘러싼 최근의 변화를 고려한다면, 도서관 목록의 발전을 논의하기 위해서는 무엇보다도 관점의 변화가 필요해 보인다. 이제 상업적으로나 비상업적으로나 서지정보를 제공하는 주체가 다양화되었으며, 이러한 추세는 도서관 목록의 품질을 평가하는 관점에도 변화를 요구하고 있다. 이용자는 목록규칙에 따라 기술요소를 정확히 입력한 도서관의 서지 레코드를 표제나 저자명을 이용해 검색하는 데에 만족하지 않기 때문이다. '서평'이나 '별집', '구매행태를 반영한 링크'와 같은 부가 정보, 검색결과를 페이스북이나 블로그에 바로 답아주는 웹서비스는 '도서의 크기'나 '자료종별', '발행지' 정보보다 매력적이다. 그러나 목록의 품질을 향상시키기 위해 서지기관에서 더 많은 데이터를 입력하기는 어렵다. 기관의 예산과 관련하여 생각해 보면, 목록 작성과 관리에 드는 비용은 늘릴 수 있는 비용이 아니라 가능하면 줄여야 하는 항목으로 인식되고 있다. 따라서 Osborn(1941)이 제기했던 '목록의 위기'는 현재에도 진행 중이거나 심화된 것으로 보인다.

하지만 이 모든 위기의식에는 도서관이 정보 서비스의 중심에 있어야 하고, 이용자의 일차적인 검색지점이 되어야 한다는 관점이 담겨 있다. 이용자의 관점에서 보면 웹 포털이나 인터넷 서점의 서지 정보의 확산은 위기가 아니다. 그리고 우리가 다른 분야의 서비스를 협력 대상으로 보는 것이 바람직하다. 도서관 서비스는 상업적인 분야에서 포함하지 못하는 부분을 채워줄 수 있으며, 저작의 집중과 같은 목록의 기능과 그 기능을 실현하기 위한 전거제어

와 같은 기법들은 앞으로 더욱 발전가능성이 큰 우리의 자산이기 때문이다. 링크드 데이터의 관점에 따르면, 개별 기관이나 개인이 각기 구축한 데이터를 가지고 경쟁하는 것은 중요하지 않다. 데이터의 활용 측면에서 고립된 정보(data silo)를 연계된 정보 공간으로 끌어내는 것에 중점을 두는 것이다. 따라서 이 연구의 목적은 링크드 데이터를 통해 개별 도서관의 서지정보를 다른 도서관이나 관련 분야에서 제공하는 서지정보와 연계하여 확장시키는데 있다.

그리고 본 연구에서는 다양한 분야 간의 협력 수단으로 링크드 데이터를 선정하였다. 링크드 데이터는 데이터를 개방하여 연계하고, 공유하고, 재사용하기 위한 구조를 제공하기 때문이다. 이를 위해서 우선 링크드 데이터의 원칙과 현황을 분석하여, 연계 가능한 링크드 데이터를 선정하였다. 그리고 링크드 데이터 형식으로 서지정보를 표현하기에 적합한 서지 온톨로지를 선정하였으며, 온톨로지 구축도구인 'Protégé'를 이용하여 이용 가능한 링크드 데이터들을 실험적으로 연계하였다. 그리고 그 과정에서 도출된 향후 과제를 제시하였다.

2. 이론적 배경

2.1 링크드 데이터

시맨틱 웹은 기존의 문자열 매칭 방식으로 정보를 탐색해야 했던 웹의 한계를 개선하기 위해 대두된 것으로, 기계가 이해할 수 있는 형식으로 데이터를 기술하는 방법을 제시하고 있다 (Antoniou & Harmelen, 2008). 그런데 시맨

틱 웹은 하나의 단일한 형식으로 이루어지는 것이 아니라 시맨틱 웹 표준을 준수하는 다양한 형식의 데이터가 연결됨으로써 이루어지며, 웹의 발전과 함께 시맨틱 웹도 함께 발전하고 있다. Berners-Lee를 중심으로 확산되는 링크드 데이터 역시 시맨틱 웹의 발전 과정에 속해 있는데, 링크드 데이터는 웹 공간에서 데이터를 공유하고, 연계하는 수단을 제공해 준다(Bizer, Heath, & Berners-Lee, 2009). 링크드 데이터의 특징은 '링크드 데이터의 구축원칙'을 통해 살펴볼 수 있으며(Berners-Lee, 2006; 조명대, 2010), 각 원칙을 <표 1>의 예시와 함께 설명하면 다음과 같다.

첫째, 링크드 데이터에서는 웹에서 개체를 고유하게 식별하기 위해 URI를 사용한다. 이 URI는 웹에서 개체를 식별하는 고유기호로서, URL로 식별되는 특정 웹문서에 종속되지 않는다. <표 1>의 기술대상은 웹 문서의 주소가 아니라, 서지 개체로서 URI를 식별기호로 갖는다. 따라서 이 개체는 다른 URL에 속하는 웹 문서에서도 사용될 수 있으며, 웹에서 동일한 URI는 동일한 개체로 인식된다.

둘째, 링크드 데이터에서는 정보 교환을 위

한 데이터 구조인 RDF(Resource Description Framework)를 사용한다. <표 1>에서 URI를 부여받은 기술대상은 주어가 되며, 각 기술요소는 술어에, 요소 값은 목적어에 해당하는 트리플 구조를 가지고 있다. 이 구조는 3개의 요소로서 모든 설명이 가능하다는 장점이 있다. 하나의 문장을 하나의 트리플로 표현함으로써 정보를 분해하거나 결합하기에 용이해지기 때문이다.

셋째, 링크드 데이터는 HTTP 프로토콜을 준수하는 URI를 사용한다. 이는 웹을 통해 별도의 제약 없이 데이터에 접근하기 위한 것이다.

마지막으로, 링크드 데이터는 상이한 URI를 이어주는 링크 정보 제공한다. 링크드 데이터의 원칙 중 하나인 '링크를 통한 데이터 연계'는 기존 시맨틱 웹이 링크드 데이터가 되기 위한 가장 중요한 요건이다. 이 때, 링크드 데이터에서 연계 방식에는 <표 2>와 같이 3가지가 있다.

- ① RDF의 관계 정보를 이용한 연계: RDF의 SPO 구문을 이용한 방법이다. <표 2>의 첫째 예시는 'Bob'이란 이름을 가진 누군가가 'Alice'를 포함하는 URI를 가진 누군가를 알고 있다는 나타내고 있다.

<표 1> Cambridge 대학 도서관의 링크드 데이터 예시

<http://data.lib.cam.ac.uk/id/entry/cambrdgedb_784321>

주어: <http://data.lib.cam.ac.uk/id/entry/cambrdgedb_784321>에 대한 정보

술어	목적어
http://purl.org/dc/terms/title	Annual report label: text
http://purl.org/dc/terms/type	http://data.lib.cam.ac.uk/id/type/1cb251ec0d568def6a929b520c4aed8d1 label: continuing
http://purl.org/dc/terms/type	http://data.lib.cam.ac.uk/id/type/791b0c20a65a95dc903f4928be58156c
http://purl.org/dc/terms/identifier	UkCU784321
http://purl.org/dc/terms/language	http://id.loc.gov/vocabulary/iso639-2/eng

〈표 2〉 데이터의 연계 방식(Heath, Tom, & Bizer, 2011)

<ul style="list-style-type: none"> • Relationship Links: RDF(SPO 구문) 이용 :me foaf:name "Bob". :me foaf:knows <http://alice.exa.org/i>. • Identity Links: URI aliases 이용 <rdf:Description rdf:about="#William_Jefferson_Clinton"> <owl:sameAs rdf:resource="#BillClinton"/></rdf:Description> • Vocabulary Links: 동등관계를 이용한 느슨한 연계 <footballTeam owl:equivalentClass us:soccerTeam/>
--

- ② URI를 직접 연계: 'owl:sameAs'를 이용한 방법이다. 〈표 2〉의 두 번째 예시를 보면, 'William Jefferson Clinton'에 대한 기술문 다음에 'sameAs'를 통해 'BillClinton'이 같은 인물이라는 것을 나타내고 있다.
- ③ 동등관계를 이용하는 방법: 'owl:equivalentClass'와 같은 관계어휘를 이용한다. 둘째 경우인 'sameAs'보다는 느슨한 연계 방법으로써, 〈표 2〉의 세 번째 예시를 보면, 'footballteam'이 'soccerTeam'과 유사하다는 것을 나타내고 있다.

이와 같이 링크드 데이터는 이용자 측면과 목록데이터 작성기관의 측면, 그리고 목록작성자 측면에서 이점을 갖는다. 우선 이용자는 도서관 간의 경계, 도서관과 관련 타 분야와의 경계를 넘어선 데이터의 향해가 가능한데, 이를 여행에 빗대어 'toURISM'이라고도 부른다. 그리고 목록데이터 작성기관 측면에서도 이점이 있다. 전통적인 방식으로 목록데이터의 품질을 향상시키기 위해서는 더 많은 비용이 필요한데, 이와 관련 비용을 추가적으로 지불하기는 어렵다. 따라서 다양한 방식으로 도서관 외부에서 생산된 링크드 데이터를 수집하여 조직한다면, 개별 목록데이터 작성비용을 절감하면서 목록

데이터의 범위 확장 및 품질 향상을 생각할 수 있다. 마지막으로 목록작성자 링크드 데이터를 통해 목록데이터 작성에 필요한 다양한 자원을 직접 제공 받을 수 있는 이점이 있다. 이는 타 분야에서 이미 구축한 데이터를 재생산하는 과정을 축소시킴으로써, 도서관 분야에 특화된 데이터 생산에 주력할 수 있다는 이점이 있다 (Baker et al., 2011).

그런데 도서관 링크드 데이터는 아직 초기 단계이다. 링크드 데이터에 대한 기본 원칙은 Berners-Lee(2006)가 제시하고 있으며, 링크드 데이터를 활용하기 위한 도구들은 Bizer, Heath, & Berners-Lee(2009)의 글에서 소개하고 있다. 그리고 우리 분야의 시각으로 링크드 데이터의 장점을 강조한 연구로는 Singer(2009)와 Byrne과 Goddard(2010)의 논문이 있다. 2011년 전반기에는 링크드 데이터를 다룬 상세한 개론서로 Stuart(2011)의 저서가 출간되었으며, 하반기에는 링크드 데이터에 대한 포괄적인 조사 결과인 W3C의 3중 보고서가 발표되었다(Baker et al., 2011; Isaac et al., 2011; Suero, 2011). 이 보고서는 '최종 보고서'와 '기술요소 및 어휘체계(Datasets, Value Vocabulary, and Metadata Element Sets)', '활용 사례(Use Cases)'로 구성되어 있으며, 이

보고서들을 통해 우리는 도서관 링크드 데이터에 관한 전반적인 개념과 사례를 파악할 수 있다.

국내에서 발표된 관련된 연구도 초기 단계인데, 조명대(2010)의 연구와 오삼균, 김성훈, 장원홍의 공동연구(2011)가 있다. 전자는 링크드 데이터의 원칙과 사례를 소개하고, 이를 우리 분야에 적용하기 위한 모델을 제시하였다. 그리고 후자에서는 W3C의 '도서관 링크드 데이터 보고서'(Library Linked Data Incubator Group Final Report)를 중심으로 링크드 데이터를 소개하고, 이를 국내에 확산시키기 위한 방안을 제시하였다. 도서관 현장에서도 아직 링크드 데이터에 대한 인식은 낮은 편이다. 노동조와 조철현(2011)은 도서관 3.0 기반 서비스에 대한 대학도서관 사서의 인식에 관한 결과를 발표했다. 이 연구에 따르면, 조사 항목 중 인지도가 높은 항목은 '모바일 도서관', '시맨틱 검색', '인공지능' 순이었으며, 링크드 데이터에 해당하는 '링크드 도서관' 항목은 가장 인지도가 가장 낮았다. 그러나 국립중앙도서관에서 링크드 데이터 구축 사업에 착수하였으므로 도서관계에서 링크드 데이터에 대한 인식을 높아질 것으로 보인다(박진호, 2011). 또한 한국과학기술정보연구원(KISTI)에서는 링크드 데이터를 구축하고, 이를 발전시키기 위해 노력하고 있다(김평, 2011). 이 외에도 멀티미디어 콘텐츠의 브라우징 시스템에 링크드 데이터를 적용하려는 연구도 발표되었다(이연호, 오경진, 신위살, 조근식, 2011). 앞으로 링크드 데이터는 협업을 통해 효과적으로 정보를 활용하고자 노력하는 여러 분야에서 점차 확산될 것이다.

2.2 서지 온톨로지

2.2.1 Bibo

'Bibliographic Ontology'(Bibo)는 서지 개체를 시맨틱 웹에 기술하기 위한 구조를 제공하므로, 서지 데이터를 링크드 데이터로 발전시킬 수 있는 기반 구조를 제공한다. Bibo의 기본 구조를 보면, 기술대상인 '문헌'이나 '그림' 등을 개체로 정의한다. 속성은 두 가지 유형이 있는데, '표제'나 '발행처' 등은 '객체 속성'(object properties)으로 정의하며, ISBN과 같은 식별 기호나 발행년과 같은 기술요소는 '데이터 속성'(data properties)으로 정의한다. 이 외에 주 기사항과 같은 비구조화된 기술요소는 '주석 속성'(annotation properties)으로 구분하고 있다. 또한 Bibo는 'owl'이나 'rdf/rdfs', 'skos'를 비롯하여 'dc', 'foaf'와 같은 다양한 입력체계를 이용한다.

Bibo의 데이터 일부를 제시하면 <표 3>과 같다. '사례1'을 보면, Bibo의 '문헌' 개체는 'owl'의 클래스에 해당되며, 'Document'라는 영문 레이블을 갖는다. 그리고 'owl:equivalentClass' 속성을 이용해 'foaf:Document'와의 연계 정보를 제공한다. '사례2'에서 객체속성인 'translationOf'는 'translation of'라는 영문 레이블을 갖고, 번역서와 원서를 연결하는 기능을 담당한다. 또한 'isVersionOf'의 하위 속성이며, 도메인과 범위는 모두 'bibo:Document'이므로, 이 속성을 이용하여 문헌과 문헌을 연결할 수 있다(D'Arcus & Giasson, 2009).

2.2.2 FaBiO

서지 온톨로지 중 FaBiO(FRBR-Aligned Bi-

〈표 3〉 Bibo의 'Document' 개체와 'Image' 개체 사례

〈http://code.google.com/p/bibotools/source/browse/bibo-ontology/tags/1.0/bibo.n3〉

[사례1]

```
bibo:Document a owl:Class :
  vs:term_status "stable" :
  rdfs:label "Document"@en : r
  dfs:comment ""A document (noun) is a bounded physical representation of
    body of information...(주석 일부 생략).""@en :
  owl:equivalentClass foaf:Document :
  rdfs:isDefinedBy 〈http://purl.org/ontology/bibo/〉.
```

[사례2]

```
bibo:translationOf a owl:ObjectProperty :
  vs:term_status "stable" :
  rdfs:label "translation of"@en :
  rdfs:comment "Relates a translated document to the original document."@en :
  rdfs:isDefinedBy 〈http://purl.org/ontology/bibo/〉 :
  rdfs:subPropertyOf dcterms:isVersionOf :
  rdfs:domain bibo:Document :
  rdfs:range bibo:Document .
```

bibliographic Ontology)는 Bibo와 달리 저작이 나 표현형, 구현형 등 FRBR에 속한 서지 개체들이 갖는 관계를 구조화하여 기술할 수 있다. 또한 FaBiO는 주로 인쇄형태의 출판물을 대상으로 하지만, 디지털 출판물을 비롯하여 도서관 목록과 같은 서지정보도 개체로 기술할 수 있다 (Shotton & Peroni, 2011). 예를 들어, 〈그림 1〉은 서지 개체인 'article'의 정보인데, URI와 간략

한 설명, 상·하위 클래스와의 관계를 확인할 수 있다. 'article' 하위 클래스에는 'journal article'와 'magazine article' 등이 있는데, 각각 고유한 URI가 부여되어 있다. 〈그림 2〉는 하위 클래스 중 'journal article'에 대한 정보이다. 〈그림 2〉에는 부분-전체 관계로서 'journal article'이 저널의 한 호를 의미하는 'journal issue'의 일부임을 나타내고 있다.



〈그림 1〉 FaBiO의 'article' 클래스의 구조

journal article ^C	
IRI: http://purl.org/spar/fabio/JournalArticle	back to ToC or Class ToC
An article, typically the realization of a research paper reporting original research findings, published in a journal issue.	
has super-classes	
<ul style="list-style-type: none"> • article^C • is part of^{OP} some journal issue^C 	

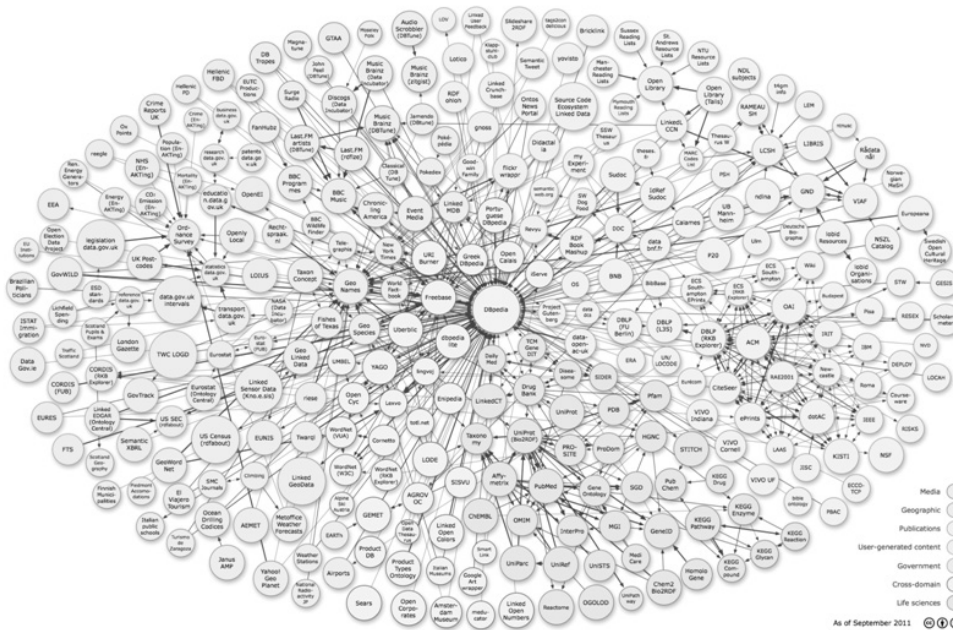
〈그림 2〉 FaBiO의 'article'의 하위 클래스인 'journal article'의 구조

3. 연계 대상 링크드 데이터

지금까지 구축된 주요 링크드 데이터는 'Linking Open Data'(LOD) 클라우드로 확인할 수 있다(〈그림 3〉 참조). 이 다이어그램에서 원은 하나의 링크드 데이터를 나타내는데, 원의 크기는 데이터에 포함된 트리플의 개수에 비례하

고, 원과 원을 연결하는 화살표의 굵기는 연계 강도에 비례한다. 다이어그램 내의 링크드 데이터라도 완전히 개방되지 않은 경우가 있는데, 각 데이터의 상세한 이용방법은 데이터마다 별도로 제시되어 있다. 참고로, 우리나라의 데이터로는 'KISTI'와 'bible ontology'가 있다.

LOD 클라우드의 가장 중앙에 위치한 원은



〈그림 3〉 Linking Open Data 클라우드(Cyganiak and Jentzsch 2011)

‘디비피디아’(DBpedia)이다. 가장 데이터의 양이 많고, 외부에서 들어오는 링크의 수도 가장 많기 때문이다. 약 360만개의 개체가 RDF 트리플 구조로 구축되어 있으며, 이 중 약 180만개는 인명, 지명, 음악앨범, 영화와 같은 분류정보도 포함되어 있다(DBpedia Team, 2010).

3.1 전거 데이터

3.1.1 FAST

링크드 전거 데이터는 링크드 데이터로 구조화시킨 서지 레코드의 기술요소나 입력체계를 의미하며, 현재 구축된 도서관 분야의 링크드 데이터의 대부분이 이에 해당된다(Baker et al., 2011). 이 중 FAST(Faceted Application of Subject Terminology)는 미국국회도서관의 주제명표(Library of Congress Subject Headings, LCSH)를 패싯화하여 재구조화한 것이다. OCLC에서는 LCSH에 이어 FAST도 링크드 데이터로 구축하였다(OCLC, “FAST Linked Data”).

〈표 4〉는 FAST 링크드 데이터의 예시로, ‘서울’이라는 개념의 URI를 이용하면, 주제명이나 지명으로 사용되는 ‘서울’에 대한 고유한 식별과 다양한 형식에서의 연계가 가능하다. 또한 ‘Has Exact Match’ 항목에서는 FAST의 주제명과 일치하는 LCSH의 주제명을 제공하고 있으며, ‘Is In Scheme’ 항목을 통해 이 주제명이 FAST의 지리 패싯에 속하는 것을 알 수 있다.

또한 FAST는 지명 데이터베이스인 ‘GeoNames’와 연계되어 있다. FAST의 ‘Focus’ 링크를 이용하면, 〈그림 4〉와 같이 관련 지리 정보를 함께 확인할 수 있다.

3.1.2 VIAF

가상의 국제전거파일(The Virtual International Authority File, VIAF)은 OCLC를 중심으로 각국의 국가도서관이 협력하여 구축한 전거레코드의 집합이다. VIAF는 웹에 공개되어 있으므로 개별 기관의 전거파일 구축비용을 절감할 수 있고, 다양한 전거파일을 링크하므로 다양한 표목의 연계가 가능하다(OCLC, “Virtual

〈표 4〉 FAST 링크드 데이터의 예시

〈<http://experimental.worldcat.org/fast/1410733/>〉

Korea (South) – Seoul (Information about the Concept)
Preferred Label: Korea (South)--Seoul
Linked Data Identifier: http://id.worldcat.org/fast/1410733
Identifier: 1410733
Alternative Label: Korea--Seoul, Korea--Sye-ul, Korea--Wang-ching (외 다수)
Has Exact Match: Seoul (Korea)
Focus: Seoul
Is In Scheme:
http://id.worldcat.org/fast/ontology/1.0/#fast
http://id.worldcat.org/fast/ontology/1.0/#facet-Geographic
Note:
GeoNames [algorithmically matched] - ppl:37°33'59"N 126°59'59"E (en) (이후 생략)



〈그림 4〉 FAST와 'GeoNames'의 연계
 〈<http://www.geonames.org/1835848/seoul.html>〉

International Authority File”). 특히, VIAF는 전거파일의 기본 유형인 개인명이나 단체명 외에도, 표제를 비롯하여 FRBR 제1집단의 상위 개체인 저작과 표현형에 대한 전거도 제공하고 있다. 예를 들어, 소설 ‘해리포터’ 시리즈를 VIAF에서 검색하면 〈표 5〉와 같은 유형

의 전거 데이터가 제공된다. 이 데이터를 보면, 각 개체에 대해 ID와 퍼머링크가 부여되어 있다. 이 정보를 통해 웹 공간에서 개체를 고유하게 식별할 수 있으며, 이 데이터에 포함된 모든 정보를 직접 끌어오거나 연계하여 활용할 수 있다.

〈표 5〉 VIAF 링크드 데이터의 저작과 표현형의 연계 예시
 〈<http://viaf.org/viaf/>〉

• 저작 ‘해리포터와 마법사의 돌’
Harry Potter and the sorcerer’s stone
VIAF ID: 198559295 (Work)
Permalink: http://viaf.org/viaf/198559295
- 표현형1 ‘해리포터와 마법사의 돌’의 독일어판
Rowling, J. K. Harry Potter and the philosopher’s stone. German
VIAF ID: 184879110 (Expression)
Permalink: http://viaf.org/viaf/184879110
- 표현형 2 ‘해리포터와 마법사의 돌’의 스페인어판
Rowling, J. K. Harry Potter and the philosopher’s stone. Spanish
VIAF ID: 180020947 (Expression)
Permalink: http://viaf.org/viaf/180020947 (.....)

3.2 서지 데이터

3.2.1 Open Library

‘오픈 라이브러리’(Open Library)는 인터넷 아카이브(Internet Archive)의 공개 프로젝트로서 대량의 서지데이터를 보유하고 있으며, 위키 방식을 도입하여 이용자가 데이터를 추가하거나 편집할 수 있다(Internet Archive “Open Library”). 따라서 저작이나 저자에 따라 데이터의 상세성에 차이가 있다. 단, 위키피디아와 달리 오픈라이브러리는 계정 없이도 편집이 가능하며, 이 경우에는 편집자의 IP주소가 공개된다.

오픈 라이브러리에서 제공하는 서지 데이터의

특징으로는 RDF 형식으로 데이터를 공개한다는 것 외에도, <표 6>과 같이 FRBR 모형의 제1집단 개체 구분을 적용하였다는 점이 있다. 따라서 오픈 라이브러리의 서지 데이터를 FaBiO와 같이 FRBR 구조를 지원하는 온톨로지에 반입할 경우, FRBR 구조를 유지할 수 있다.

3.2.2 BNB

영국 국가도서관에서는 방대한 양의 국가서지(British National Bibliography, BNB)를 링크드 데이터로 구축하였는데, 약 26만 개의 레코드를 대상으로 8천여 개의 트리플이 있다. BNB에서 제공하는 링크 정보에는 ‘VIAF’, ‘LCSH’

<표 6> 오픈 라이브러리의 링크드 데이터 일부

<http://openlibrary.org/works/OL15702773W>

```

<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdrel="http://RDVocab.info/RDARelationshipsWEMI/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcam="http://purl.org/dc/dcam/"
  xmlns:ov="http://open.vocab.org/terms/"
  xmlns:frbr="http://purl.org/vocab/frbr/core#">

  <frbr:Work rdf:about="http://openlibrary.org/works/OL15702773W/">
    <dcterms:title>Please Look After Mom</dcterms:title>
    <dcterms:creator
      <rdf:Description rdf:about="http://openlibrary.org/authors/OL7078492A">
        <rdf:value>Kyung-sook Shin</rdf:value>
      </rdf:Description>
    </dcterms:creator>
  </frbr:Work>

  <rdf:Description rdf:about="http://openlibrary.org/books/OL24627332M/">
    <rdrel:workManifested>http://openlibrary.org/works/OL15702773W/
    </rdrel:workManifested>
    <dcterms:title>Please Look After Mom</dcterms:title>
    <dcterms:date>2011</dcterms:date>
  </rdf:Description> (일부 생략)

```

'GeoNames'를 비롯하여 언어입력 체계인 'Lexvo', 'MARC의 국가코드 및 언어코드', 'DDC', 'RDF Book Mashup' 등 매우 다양하다. 현재는 단행본의 서지레코드를 중심으로 구축되어 있으나, 향후 연속간행물이나 가제식 출판물과 같은 계속자원 등도 포함할 계획이다("British Library Metadata Service" [Online]).

BNB 링크드 데이터에서는 두 가지 유형의

메타데이터 요소집합을 이용하여 서지데이터를 제공하고 있다. <표 7>을 보면, 첫째 유형인 'Linked Open BNB'는 'owl:sameAs' 속성을 이용하여 동일 개체를 표현하였고, 'rdfs:seeAlso' 속성을 이용하여 또 다른 링크드 데이터에 포함된 개체와의 연계 정보를 표현하고 있다. 반면에 두 번째 예시에서는 기본적인 RDF 구문만 표현하고 있음을 알 수 있다.

<표 7> BNB의 링크 정보 유형

<http://www.bl.uk/bibliographic/datasamples.html>

• **Linked Open BNB in RDF/XML**

- 서지 온톨로지 (Bibliographic Ontology)
- 전기정보 용어 (Bio: A Vocabulary for Biographical Information)
- 영국 국가도서관 용어 (British Library Terms)
- 더블린 코어
- 이벤트 온톨로지 (Event Ontology)
- 인맥정보 온톨로지 (FOAF: Friend of a Friend)
- 기관 온톨로지 (Org: An Organization Ontology)
- OWL
- SKOS
- RDF 스키마
- 지리 정보 (S84 Geo Positioning)
- 국제표준서지기술 (ISBD) 및 자원의 접근과 기술(RDA)

[예시]

```
<rdf:Description rdf:about="http://bnb.data.bl.uk/id/resource/007833980">
<blt:bnb>GBB179171</blt:bnb>
<owl:sameAs rdf:resource="http://bnb.data.bl.uk/id/resource/GBB179171"/>
<rdf:type rdf:resource="http://purl.org/dc/terms/BibliographicResource"/>
<rdf:type rdf:resource="http://purl.org/ontology/bibo/Book"/>
<bibo:isbn10>0954465407</bibo:isbn10>
<rdfs:seeAlso rdf:resource="http://www4.wiwiwiss.fu-berlin.de/bookmashup/books/0954465407"/>
<dc:title>The Battle of Britain</dc:title>
```

• **BNB in Basic RDF/XML**

- 더블린 코어
- SKOS
- OWLT, Time Ontology in OWL
- RDF 스키마
- 서지 온톨로지 (bliographic Ontology)
- 국제표준서지기술 (ISBD) 및 자원의 접근과 기술(RDA)

[예시]

```
<rdf:Description>
<dcterms:title>The Battle of Britain</dcterms:title>
```

4. 링크드 데이터의 실험적 연계

4.1 실험적 연계

4.1.1 데이터 입력체계 연계

Bibo와 FaBiO를 중심으로 링크드 데이터를 구축할 플랫폼으로는 온톨로지 구축도구인 Protégé를 선정하였다. 서지 기술요소를 반영한 화면은 <그림 5>와 같은데, FRBR 제1집단은 'endeavour' 클래스에 속하고, 제2집단은 'responsible entity' 클래스에, 제3집단은 'subject' 클래스에 속하는 것을 확인할 수 있다. 그리고 FaBiO는 각 클래스의 '대상속성'과 '데이터속성'도 함께 제공하며, 각 개체에 대한 정보를 주석을 통해서도 표현하고 있다.

FaBiO를 먼저 반영한 이후, Bibo와 같은 관련 온톨로지를 새로 반영함으로써 둘 이상의 온톨로지를 병합할 수도 있다. 이와 같은 Protégé의 온톨로지 반영 기능은 기존에 반영한 온톨로지에 새로 반영한 온톨로지를 병합할 수 있도록 지원한다. 특히 웹에 존재하는 온톨로지의 경우, URI를 통해 직접 온톨로지를 반영할

수 있다.

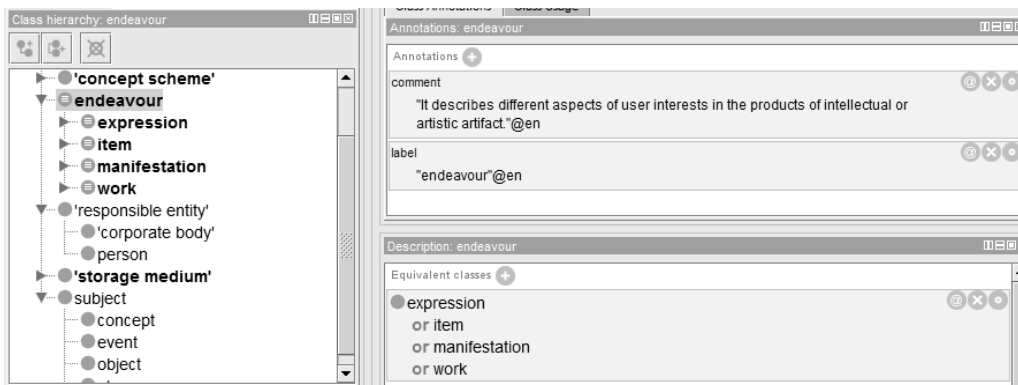
4.1.2 서지 데이터 연계

연계대상 데이터로는 톨킨(Tolkien, J. R. R.)의 대표 저작인 『호빗(Hobbit)』과 『반지의 제왕(The Lord of the Rings)』을 선정하였다.

1) 오픈 라이브러리의 링크드 데이터

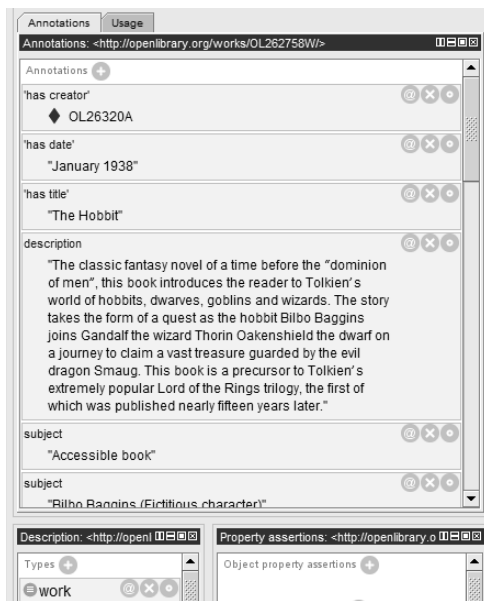
오픈 라이브러리는 아래 예시와 같이 저작 정보는 FOAF 체계를 적용하였고, 저작 정보는 FRBR에서 규정한 서지 개체를 반영할 수 있도록 구조화시키고 있다.

- 저작 '톨킨' - <foaf:Person rdf:about="http://openlibrary.org/authors/OL26320A"> <foaf:name>J. R. R. Tolkien </foaf:name>
- 저작 '호빗' - <frbr:Work rdf:about="http://openlibrary.org/works/OL262758W/"><dcterms:title>TheHobbit</dcterms:title>



<그림 5> Protégé에 FaBiO 온톨로지를 반영한 화면

이 데이터의 URI를 이용하여, FaBiO와 Bibo를 중심으로 구조화된 틀에 반입하면 <그림 6>과 같은데, 저작의 표제나 발행년과 같은 기본적인 서지사항을 비롯하여, 주제명과 저자 정보 등을 확인할 수 있다. 저자의 경우에는 '톨킨'에 대한 URI로 연계되어 있으므로, 저작과 저자간의 연계가 URI를 통해 이루어지고 있다. 저자 정보에는 기본적인 전기 정보와 관련 저작, 다양한 형식의 저자명을 제공하고 있다. 그리고 저작 정보는 FaBiO의 'endeavour'의 하위에 있는 'work' 클래스의 인스턴스가 되고, 저자 정보는 Bibo의 'person' 클래스의 인스턴스가 되었다. 이는 오픈 라이브러리가 제1집단만 반영했기 때문이다.



<그림 6> 오픈 라이브러리의 링크드 데이터를 반입한 화면

2) BNB의 링크드 데이터
오픈 라이브러리가 대표적인 위키 형식의 웹

서지 데이터라고 한다면, BNB는 국가서지에 해당하는 데이터를 링크드 데이터로 변환한 것이므로, 두 데이터를 연계하여 상호보완적으로 사용하는 것이 바람직할 것이다. BNB도 오픈 라이브러리와 같이 아래와 같은 데이터를 반입하였는데, URI와 기술요소명이 상이한 것을 알 수 있다.

- 저자 '톨킨'
<http://bnb.data.bl.uk/id/person/TolkienJRR%28JohnRonaldReuel%291892-1973>
- 저작 '호빗'의 영문판(London: HarperCollins, 2008)
<rdf:Description rdf:about="http://bnb.data.bl.uk/id/resource/014469086">
<rdfs:label>The hobbit / J. R. R. Tolkien</rdfs:label>
- 저작 '반지의 제왕'의 영문판(London: HarperCollins, 2007)
<rdf:Description rdf:about="http://bnb.data.bl.uk/id/resource/009702968">
<rdfs:label>The Lord of the Rings / J. R. R. Tolkien</rdfs:label>

BNB의 특징은 DDC, LCSH와 같은 전통적인 서지데이터 요소를 충실히 포함하고 있다는 점이다. <그림 7>은 BNB 서지 데이터를 반입한 화면의 일부인데, 'Agent' 클래스에서 원저자와 역자를, 'BibliographicResource' 클래스와 'Book' 클래스에서 저작정보를 인스턴스로 제공하고 있다. 그리고 각 인스턴스는 복수의 입력체계에 속하며, 전기 정보와 서지 기술 정보를 포함하고 있다. 또한 'TopicDDC'와



〈그림 7〉 BNB 링크드 데이터 반입 화면

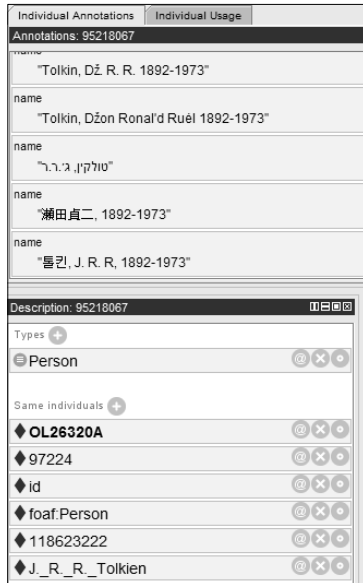
‘TopicLCSH’ 클래스를 통해 분류기호와 주제 명표목을 제공하는 것을 확인할 수 있다.

그런데 BNB는 구현형을 기반으로 작성된 서지 링크드 데이터이므로, 반입 이후 구현형 클래스에 속하는 것임을 반영해 주어야 한다. 이는 ‘Book’ 클래스를 FaBiO를 반영한 온톨로지의 ‘Manifestation’ 클래스와 연계함으로써 가능하다. 그리고 오픈 라이브러리에서 가져온 저작 정보와 연계하면, 오픈 라이브러리의 저작정보와 BNB의 구현형 정보를 계층관계로 연계할 수도 있다.

4.1.3 전거 데이터 연계

1) VIAF의 링크드 데이터

오픈 라이브러리나 BNB가 서지데이터를 중심으로 구축되었다면, VIAF는 저자 전거를 중심으로 저작을 비롯한 다양한 전거데이터를 제공하고 있다. 또한 VIAF의 경우, ‘Tolkien’과 같은 인명은 ‘FOAF’ 및 FRBR 모형의 제2집단인 ‘person’과 연결되어 있다. 그리고 저작 정보인 ‘호빗’이나 ‘반지의 제왕’은 독일 국가도서관의 레코드와 ‘owl:sameAs’ 관계로 연결되어 있다.



〈그림 8〉 VIAF의 링크드 데이터를 반입한 화면

예를 들면, VIAF 파일의 인명인 '톨킨' 데이터는 〈그림 8〉과 같이 'person' 클래스의 인스턴스로 반입된다. 그러나 저작권자인 '호빗'과 '반지의 제왕'은 'work' 클래스에 반입되지 않았으며, 이 저작과 관련된 외부 체계인 미국국회도서관(LC), 독일국가도서관(DNB), 호주국립도서관(NLA)의 전거데이터와의 링크 정보가 'concept' 클래스의 인스턴스로 반입되었다. 그리고 모두 'concept'로 처리되었다. 이는 VIAF에서 저작에 대한 식별기호를 제공하면서, 이를 FRBR와 관련된 외부 입력체계와 연계하지 않았기 때문이다. 오픈 라이브러리의 경우는 아래와 같이 'frbr:work'라고 명시했기 때문에 해당 개체가 '저작'임을 Protégé가 인식할 수 있었다. VIAF에서 저작과 표현형에 대한 정보를 구축했다면, 오픈 라이브러리의 경우와 같이 연계정보를 함께 제공해야 할 것이다. 따라

서 이에 Protégé의 'type' 메뉴를 이용하여, 클래스를 할당받지 못한 저작 인스턴스에 'work' 클래스를 지정할 수 있다.

2) FAST 링크드 데이터

FAST는 LCSH를 바탕으로 한 주제 전거 데이터인 만큼 풍부한 주제명을 포함하고 있으며, 각 주제명표목에 대해 패싯 정보를 제공하고 있다. 단, FAST의 주제명은 기본적으로 SKOS 체계를 기반으로 하므로, 패싯 정보가 온톨로지 클래스로 구체화되어 있지는 않는다. 반입한 데이터는 모두 'concept' 클래스에 속하게 되며, 필요한 경우 패싯을 클래스로 추가하여 주제명과 연계할 수 있다. 〈그림 9〉는 FAST에서 'Fantasy fiction, English'이라는 주제명 데이터를 반입한 화면인데, 주제명과 관련된 다양한 정보를 확인할 수 있다.



〈그림 9〉 FAST의 링크드 데이터를 반입한 화면

그런데, FAST와 같은 전거 데이터 자체에 서지 데이터와의 연계 정보는 포함되어 있지 않다. 따라서 오픈 라이브러리나 BNB 링크드 데이터와 같은 서지데이터와 연계해서 사용할 필요가 있다. 링크드 데이터 형식의 주제명을 통해 상이한 서지 데이터가 연계된다면, 도서관 중심의 서지 데이터 이외에 다양한 정보가 동일한 주제로 응집될 수 있을 것이다.

4.2 향후 과제 및 제언

4.2.1 적합한 데이터의 선정

링크드 데이터는 공유할 수 있는 형식으로 공개된 데이터로서 많은 장점이 기대된다. 그러나 '공유할 수 있다'는 가능성 자체가 각 기관에 적합한 서지데이터를 자동으로 제공한다는 의미는 아니다. 링크드 데이터는 누군가 가져가서 유용하게 쓸 수 있는 데이터이므로, 특정 분야나 특정 기관에 최적화시켜 제공하는 것이 아니기 때문이다. 링크드 데이터의 원칙도 해당 데이터를 공유하고 해석할 수 있는 프레임워크를 제공하는 것이므로, 데이터의 내용 측면까지 보장할 수는 없다. 따라서 우리 분야에 적합한 링크드 데이터를 검색하고 활용하기 위해서는 다음 두 가지가 요구된다.

첫째, 우리 분야의 시각으로 데이터를 해석하고 선정할 수 있어야 한다. 분야별로 정보 자료를 분석하는 관점에 차이가 존재하므로, 각 데이터에는 그 데이터를 생산한 기관이나 개인의 시각이 존재한다. 예를 들어, '오픈 라이브러리'는 서지 개체를 FRBR 모형에 따라 구분하고 있으므로, FRBR 모형에 대한 이해가 필요하다. 오픈 라이브러리와 연계성을 통해 구현

형 중심의 서지 데이터에 저작 정보를 연계할 수도 있을 것이다. 그리고 FAST는 LCSH와 달리 패시를 적용하여 후조합 형식을 취하고 있다. 따라서 LCSH의 세목을 패시로 변환한 차이점이나 색인 유형의 특징, 패시의 장점에 대한 이해가 필요하다. 자관의 서지데이터에 포함된 주제명이 후조합이라면 FAST와의 연계가 더 효과적일 것이다. BNB 링크드 데이터는 구현형을 기반으로 하는 MARC 데이터를 링크드 데이터화로 구축한 것이지만, 'VIAF'와 같은 전거 데이터나 'FOAF'와 같은 인맥정보, 'GeoPositioning'과 같은 지리정보 등 다양한 외부 정보원과의 링크를 포함하고 있다. 단순히 MARC 데이터를 공개된 온톨로지로 변환한 것보다 매우 풍부한 연계 정보를 제공하는 것이다. 따라서 영국국가도서관의 서지 데이터 자체는 필요하지 않은 기관이라도, 구현형 기반의 서지 데이터를 연계함으로써 다양한 외부 정보원과의 링크 정보를 부가로 얻을 수도 있을 것이다. 단순히 공개된 데이터를 연계하여 사용하는 것이 아니라, 데이터를 분석하여 자관에 필요한 데이터를 선별하는 것이 더 중요할 것이다.

둘째, 선정된 링크드 데이터의 유효성을 지속적으로 점검해야 한다. 링크드 데이터의 유효성은 외부정보원의 신뢰성과도 관련이 있다. 오픈 라이브러리는 이용자가 직접 서지데이터를 편집할 수 있기 때문에 풍부한 데이터가 저장될 수 있지만, 반면에 전문가가 작성한 데이터가 아니라는 한계도 있다. 또한 매우 풍부한 데이터가 저장된 도서가 있는 반면, 저자명도 제대로 입력되지 않은 도서가 있는 등 편차가 크다. 그리고 링크드 데이터의 대부분이 영어문화권

을 중심으로 구축되어 있기 때문에 문화적인 편견으로부터도 자유롭지 않을 것이다. 또한 웹에 존재하는 데이터의 공통된 사항으로 링크정보가 유효한지도 확인해야 한다.

이와 같이 링크드 데이터를 효과적으로 사용하기 위해서는 우리 분야의 관점에서 그 데이터가 어떠한 장점과 단점을 가지고 있는지를 파악하고 있어야 할 것이다.

4.2.2 데이터간의 연계 기준 정립

기관별 특성과 외부 링크드 데이터의 특성을 분석하여, 적합한 데이터를 선정한 다음에는 각 링크드 데이터를 연계하는 기준을 정립해야 한다. 링크드 데이터의 특성 상 둘 이상의 데이터를 연계하는 경우가 많으며, 다양한 분야의 데이터가 존재하기 때문이다. 또한 동일 분야의 링크드 데이터라도 각 데이터의 구조가 상이한 경우가 많다. 몇 가지 원칙과 데이터 구조를 준수하면 모두 링크드 데이터가 될 수 있기 때문이다. 비교적 상세한 입력체계인 MARC도 국가별, 기관별로 상이한 것을 감안하면, 링크드 데이터의 다양성의 범위는 더욱 클 것이다.

따라서 개별 분야나 기관이라는 구체적인 맥락을 가정한다면, 링크드 데이터는 다시 '링크'되어야 한다. 즉, 링크드 데이터는 '연계 가능한' 데이터이지, 자동으로 이미 연계되어 있는 것은 아니다. 만일 Protégé에 FaBiO 온톨로지와 Bibo 온톨로지를 모두 반입하였다면, FaBiO를 기준으로 작성된 데이터는 FaBiO의 클래스에, Bibo를 기준으로 작성된 데이터는 Bibo의 클래스에 해당하는 인스턴스가 될 것이다. 이 경우 Protégé와 같은 플랫폼에서 제공하는 기능을 활용하여 상이한 기술요소간을 연계할 필요가

있다. 예를 들어, FaBiO의 경우 '인명'과 '단체명'은 FRBR 제2집단에 속하므로, 'responsible entity'의 하위 클래스인 반면, Bibo는 인명과 단체명이 별도의 클래스로 구성되어 있다. 따라서 'equivalent classes' 관계를 맺어줄 필요가 있다. 이는 링크드 데이터의 연결 기법 중 하나인데, 모든 링크가 미리 저장된 것이 아니라 필요에 따라 추가될 수 있는 유동적인 것임을 의미한다.

기술요소 뿐 아니라 서지 데이터의 경우도 연계가 필요하다. 오픈라이브러리 링크드 데이터는 'dcterms:title' 요소를 이용하여 저작의 표제를 기술하는 반면, BNB 링크드 데이터는 'rdfs:label' 요소를 이용하여 표제를 기술한다. 따라서 표제 정보도 연결이 필요하다. 또한 저작에 해당하는 정보가 SKOS 체계에 따라 모두 'concept' 클래스의 인스턴스가 되는 경우가 많다. 그러나 이용자의 요구에 따라 저작이나 주제명을 세분할 필요가 있을 경우에는, 이를 조정할 수 있어야 한다. 이 외에도 여러 가지 경우가 가능한데, 점차 FRBR 모형이 확산되면서 FRBR 모형을 적용한 서지데이터와 그렇지 않은 서지데이터의 차이도 뚜렷해 질 것이다. FRBR 모형을 적용한 서지데이터도 구체적인 적용 방법의 차이로 인해, 다시 연결할 필요가 생길 수 있다. 따라서 둘 요소가 같다는 연계 정보를 맺어 줄 수 있어야 한다.

데이터 연계에는 같은 것을 이어주는 것 외에도, 중복 데이터를 체크하여 통합하거나 제거하는 과정도 포함될 것이다. 웹에 존재하는 데이터의 중복은 다양성과 링크를 위한 연결고리로 이용할 수 있지만, 이를 그대로 이용자에게 보여주는 것은 바람직하지 않기 때문이다.

4.2.3 서지데이터의 재구조화와

고유데이터의 개발

효과적인 서지 데이터를 이용자에게 제공하기 위해서는 적합하고 유효한 링크드 데이터를 선정하고, 각 링크드 데이터를 다시 '연계'한 것만으로는 충분하지 않다. 물론 우리는 이미 풍부한 자체 데이터를 보유하고 있다. 데이터뿐 아니라 데이터의 구축 방식, 공유 방식에도 우리 나름의 기준과 노하우가 있음은 물론이다. 그런데 우리의 서지데이터에는 한계가 있다. 우리의 데이터 구조가 외부 데이터와 연계하거나 이용자가 이용하기 어렵다는 점이다. 이미 'MARC' 형식의 특수성은 많이 거론되어 온 부분이다. 더불어 '국제표준서지기술'(International Standard Bibliographic Description, ISBD)의 특수성도 고려해야 한다. ISBD의 주 기술영역은 이용자가 자료에 접근하고 이용하는 주요 요소가 아닐 수 있기 때문이다. 이용자에게는 주 기술영역에 속하는 '형태사항'보다 구조화되지 않는 방식으로 다양한 정보를 담아 둔 '주기사항'에 들어가는 정보가 더 유용할 수 있다. 예를 들어 수상주기의 수상내역은 동일한 상을 수상했거나 후보에 오른 저작들과의 연계에 필요하다. 상을 받았다는 것 자체가 정보의 품질을 높여주기도 할 것이고, 어떤 상을 받았느냐에 따라 그 저작의 주제 정보를 예측할 수도 있다. 또한 주기사항에는 연주자나 배역진 주기도 포함되어 있는데, 이러한 정보는 멀티미디어 정보에 접근하고 이용하는 데 발행년보다 더 중요할 수 있다.

따라서 이미 우리가 가진 기술표준과 그 표준에 따라 생산된 콘텐츠만으로 링크드 데이터를 구축하는 것은 충분하지 않다. 기존에 생산된 데이터의 가능성을 충분히 활용하는 것도

물론 중요하지만, 새로운 관점에서 우리 분야에 특화된 데이터를 개발할 필요도 있는 것이다. 링크드 데이터는 우리의 고유 데이터에 다른 분야의 데이터를 연계하여 확장하기 위한 것이기 때문이다. 이를 위한 출발점은 기존 링크드 데이터의 특징을 이해하고, 자관의 기존 데이터를 재구조화하는 것이다. 그리고 우리 분야에 특화된 데이터에 대해 고민해야 할 것이다. 링크드 데이터에 대한 분석이 외부 세계와의 소통의 방식이 되어, 우리 분야 내에서 볼 수 없었던 한계를 볼 수 있는 새로운 시각을 제공하기를 제안한다.

5. 결론

본 논문에서는 도서관 목록을 둘러싼 정보환경의 변화를 활용하기 위해 링크드 데이터 방식을 분석하고 시범적으로 적용해 보았다. 링크드 데이터는 데이터의 구축 방식 뿐 아니라 데이터 자체를 가리키는데도 사용되지만, 본 연구는 링크드 데이터를 우리 분야의 데이터를 확장하고 개선하기 위한 수단으로 간주하였다. 이에 링크드 데이터의 개념과 원칙을 바탕으로 링크드 데이터 방식으로 서지 데이터를 기술하기 위한 서지 온톨로지를 검토하였다. 그리고 전거 데이터와 서지 데이터를 중심으로 우리가 활용할 수 있는 기존 링크드 데이터를 분석하였으며, 대표적인 링크드 데이터를 온톨로지 플랫폼인 Protégé를 이용하여 반입하고 연계해 보았다.

실험적 연계 결과, 앞으로 링크드 데이터를 우리 분야에 효과적으로 활용하기 위한 세 가

지 제안을 하였다. 첫째로는 많은 링크드 데이터 중 연계할 가치가 있고, 연계할 수 있도록 개방된 데이터를 선정하는 일이 필요하다. 링크드 데이터도 크기와 품질에 따라 선별할 필요가 있으며, 이 때 필요한 것은 링크드 데이터에 대한 기법보다 자료를 조직하는데 우리 분야에서 쌓아온 노하우와 전문 지식이 바탕이 되어야 하기 때문이다. 다음으로는 선정한 데이터를 연계하는 기준과 그 방식을 정하는 것이다. 링크드 데이터는 '연계 가능한' 데이터이며, 각 사용자에게 맞춤형 데이터가 아니기 때문이다. 인명이나 단체명, 표제 정보와 같은 고유명에서부터 주제명이나 언어 부호에 이르기까지 다양한 입력체계가 존재하므로, 이를 연계할 수 있는 정보를 부가하는 과정이 필요하다. 마지막으로 각 기관의 특성과 이용자의 요구를 반영한 고유 데이터를 개발해야 한다. 링크드 데이터를 통해 기존에 존재하는 다

양한 데이터를 반입하여 사용하는 것도 중요하지만, 필요한 데이터가 웹상에 존재하지 않는다면 새로 만들어야 할 것이다. 이 부분은 특정 분야나 기관의 고유한 기능을 위해 반드시 필요한 데이터를 생산하는 것이다. 특정 기관에게 할당된 인력이나 예산은 기존에 있는 정보를 중복하여 생산하는데 사용할 것이 아니라, 그 기관만이 생산할 수 있는 특화된 정보에 배분되어야 할 것이다. 그리고 이는 전체 정보 생태계 내에서 특정 분야의 기여도를 평가할 수 있는 부분이기도 하다.

시맨틱 웹은 앞으로 더욱 확대될 것이고, 이러한 경향의 일부분으로써 링크드 데이터 방식으로 데이터를 생산하고 공유하는데 참여하는 개인이나 단체도 점차 증가할 것이다. 이러한 흐름을 효과적으로 이용하면서, 우리 분야의 고유성을 지켜나가기 위한 시도가 앞으로 더 많이 이루어지기를 바란다.

참 고 문 헌

- 김평 (2011.12.7). Linked Data 기반 상호 연계를 위한 개체 식별. 제2회 시맨틱 웹 컨퍼런스 발표자료, 연세대학교 공학원, 서울.
- 노동조, 조철현 (2011). 도서관 3.0 기반 서비스에 대한 대학도서관사서의 인식에 관한 연구. 정보관리학회지, 28(4), 263-278. doi: 10.3743/KOSIM.2011.28.4.263
- 박진호 (2011.12.7). Library Linked Data: 사서가 소통하는 법. 제2회 시맨틱 웹 컨퍼런스 발표자료, 연세대학교 공학원, 서울.
- 오삼균, 김성훈, 장원홍 (2011). 도서관 링크데이터(Library Linked Data) 현황분석과 통합 LLD 설계 및 확산방향에 대한 고찰. 한국도서관·정보학회지, 42(4), 331-351.
- 이연호, 오경진, 신위살, 조근식 (2011). 링크드 데이터를 이용한 협업적 비디오 어노테이션 및 브라우징 시스템. 지능정보연구, 17(3), 203-219.

- 조명대 (2010). 도서관에서의 Linked Data 활용방안에 관한 연구. 한국문헌정보학회지, 44(1), 181-198.
- Baker, T., Bermès, E., Coyle, K., Dunsire, G., Isaac, A., Murray, P. ... Zeng, M. (2011). Library Linked Data Incubator Group Final Report. Retrieved from <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>
- Berners-Lee, T. (2006). Linked Data: Design issues. Retrieved from <http://www.w3.org/DesignIssues/LinkedData.html>
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data: The story so far. International Journal on Semantic Web and Information System, 5(3), 1-22.
- British Library Metadata Service. Retrieved from <http://www.bl.uk/bibliographic/datafree.html>
- Byrne, G., & Goddard, L. (2010). The strongest link: Libraries and Linked Data. D-Lib Magazine, 16(11/12). Retrieved from <http://www.dlib.org/dlib/november10/byrne/11byrne.html>
- Cyganiak, R., & Jentzsch, A. (2011). The linking open data cloud diagram. Retrieved from <http://richard.cyganiak.de/2007/10/lod/>
- D'Arcus, B., & Giasson, F. (2009). Bibliographic ontology specification. Retrieved from <http://bibliontology.com/specification>.
- DBpedia Team (2010). DBpedia. Retrieved from <http://thedatahub.org/dataset/dbpedia>
- Heath, T., & Bizer, C. (2011). Linked Data: Evolving the Web into a global data space. Morgan & Claypool.
- Internet Archive. Open Library. Retrieved from <http://openlibrary.org/>
- Isaac, A., Waites, W., Young, J., & Zeng, M. (2011). Library Linked Data incubator group: Datasets, value vocabularies, and metadata element sets. W3C Incubator Group. Retrieved from <http://www.w3.org/2005/Incubator/ld/XGR-ld-vocabdataset-20111025/>
- OCLC. FAST Linked Data. Retrieved from <http://experimental.worldcat.org/fast>
- OCLC. Virtual international authority file. Retrieved from <http://viaf.org/>
- Osborn, A. D. (1941). The crisis in cataloging. The Library Quarterly, 11(4), 393-411.
- Shotton, D., & Peroni, S. (2011). FaBiO, the FRBR-aligned bibliographic ontology. Retrieved from <http://www.essepuntato.it/lode/http://purl.org/spar/fabio#d3e2414>.
- Singer, R. (2009). Linked library data now! Journal of Electronic Resources Librarianship, 21(2), 114-126. Retrieved from <http://www.dlib.org/dlib/november10/byrne/11byrne.html>
- Stuart, D. (2011). Facilitating access to the Web of data: A guide for librarians. London: Facet Publishing.
- Suero, D. V. (2011). Library Linked Data incubator group: Use cases. W3C Incubator Group. Retrieved from <http://www.w3.org/2005/Incubator/ld/XGR-ld-usecase-20111025/>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Cho, Myungdae (2010). A Study on Applications for Linked Data in Libraries. *Journal of Korean Society For Library and Information Science*, 44(1), 181-198.
- Kim, Pyung (2011.12.7). Entity Identification for Interlink based on Linked Data. Paper presented at the 2nd Semantic Web Conference, Yonsei University, Seoul.
- Lee, Yeon-Ho, Oh, Kyeong-Jin, Sean, V., & Jo, Geun-Sik (2011). A collaborative video annotation and browsing system using Linked Data. *Journal of Intelligence and Information Systems*, 17(3), 203-219.
- Noh, Dong-Jo, & Cho, Chul-Hyun (2011). A study of the awareness focusing on the library 3.0 for the academic librarians. *Journal of the Korean Society for Information Management*, 28(4), 263-278. doi: 10.3743/KOSIM.2011.28.4.263
- Oh, Sam Gyun., Kim, Seonghun., & Jang, Wonhong (2011). Analyzing current state of library linked data, designing an integrated LLD, and thoughts on ways to expand LLD. *Journal of Korean Library and Information Science Society*, 42(4), 331-351.
- Park, Jinho (2011.12.7). Library Linked Data: The communication way between librarian. Paper presented at the 2nd Semantic Web Conference, Yonsei University, Seoul.

