

링크드 데이터에서 인물 정보의 식별 및 연계 범위 확장에 관한 연구: 국립중앙도서관 링크드 데이터를 중심으로*

Expanding the Scope of Identifying and Linking of Personal Information in Linked Data: Focusing on the Linked Data of National Library of Korea

이성숙 (Sungsook Lee)**

박지영 (Ziyoung Park)***

이혜원 (Hyewon Lee)****

초 록

본 연구에서는 국립중앙도서관 링크드 데이터를 대상으로 인물 정보가 표현되고 연계되는 방식을 분석하고, 이를 확장하기 위한 방안을 제안하였다. 분석 결과, 저자로서의 인물 정보는 링크드 데이터에서도 인물을 표현하는 어휘와 연계되어 기술되고 있는 반면에, 주제로 표현된 인물은 개념으로만 취급되고 있었다. 또한 링크드 데이터 구축과정에서 기존의 전거 정보를 변환한 것 외에는 별도의 부가 정보를 추가했는지를 확인할 수 없었다. 이에 본 연구에서는 저자로서의 인물 정보뿐 아니라 주제로서의 인물 정보도 서지 정보에 포함시키고, 저자로서의 인물 정보와 주제로서의 인물 정보를 연계할 때 링크드 데이터의 품질을 제고할 수 있다고 판단하였다. 그리고 이와 더불어 인물과 관련된 부가 링크 정보를 함께 구축하고 이를 활용하여 서지데이터 검색의 접근점을 확장하는 방안도 함께 제안하였다.

ABSTRACT

This study analyzed the methods for representing and linking personal information in the linked data of National Library of Korea and provided suggestions for expanding the scope of identifying and linking of the personal information. As a result, the personal information as a subject has been dealt with a concept, where the personal information as a contributor has been linked with a vocabulary of personal name. In addition, there have not been assured of including additional information except existing authority data in the process of building the linked data. Therefore, this study suggested that linking personal information as a subject and personal information as a contributor was essential for the quality of linked data. In addition, we proposed to provide additional information related to the person in linked data for expanding the scope of access points in information discovery.

키워드: 도서관 링크드데이터, 인물 정보, 전거 데이터, 국립중앙도서관, 가상의 국제전거 파일, 국제이름식별기호 library linked data, personal information, authority data, National Library of Korea, VIAF (Virtual International Authority Files), ISNI (International Standard Name Identifier)

* 본 연구는 2015년 충남대학교 교내연구비 지원에 의해 수행되었음.

본 연구는 한성대학교 교내학술연구비 지원과제임.

** 충남대학교 사회과학대학 문헌정보학과 교수(infolee@cnu.ac.kr) (제1저자)

*** 한성대학교 인문대학 응용인문학부 문헌정보전공 부교수(zgpark@hansung.ac.kr) (공동저자)

**** 서울여자대학교 사회과학대학 문헌정보학과 부교수(hwlee@swu.ac.kr) (교신저자)

■ 논문접수일자: 2017년 2월 20일 ■ 최초심사일자: 2017년 9월 6일 ■ 게재확정일자: 2017년 9월 20일

■ 정보관리학회지, 34(3), 7-21, 2017. [http://dx.doi.org/10.3743/KOSIM.2017.34.3.007]

1. 서론

시맨틱웹에서는, 사람은 사람끼리 장소는 장소끼리 식별정보와 연계정보를 통해 구조화된다. 이 점은 일반 웹이 정보자원을 구성하는 개별 개체를 인식하지 못하고, 문자열의 배열순서만으로 매칭시키는 것과 명확히 비교된다. 링크드 데이터는 이와 같은 시맨틱웹의 가능성을 최대한 활용할 수 있는 하나의 방안이다. 일반적으로는 유사한 데이터를 다루는 기관이라도 개별 기관의 정책과 핵심 이용자의 요구에 따라 기관별로 구축된 데이터는 다양한 모습을 갖는다. 따라서 복수의 데이터 저장소에 존재하는 데이터를 통합할 경우에는, 대표 기관에서 통합시스템을 구축하여 기관별 데이터를 하나의 틀에 넣는 방식을 사용했었다. 이와 같은 방식은 독자적인 데이터 관리 체계를 구축하기 어려운 개별 기관이나 중앙의 제어 시스템을 구축한다는 측면에서는 나름의 장점을 가질 수 있는 반면, 개별 기관의 개성을 약화시키거나 통합시스템과 기관별 시스템을 이중으로 유지해야 하는 업무 부담을 낳기도 한다. 그러나 링크드 데이터는 개별 기관의 데이터를 유지한 채 연계를 위한 큰 우산에 해당되는 데이터 구조를 모델링하고 링크 정보를 추가하는 방식으로 상이한 데이터를 통합하여 이용자에게 제공할 수 있는 방식이다. 즉 개별기관의 독립성과 자율성을 전제로 하여, 복수의 기관 간의 협업을 지원하는 것이다.

본 연구는 이와 같이 새로운 공유 플랫폼으로 활용될 수 있는 링크드 데이터가 도서관 및 유관 분야에서는 어떻게 구축되고 있는지에 대한 조사에서부터 시작하였다. 특히 국내에서는 다수의 공공기관이 유사한 시기에 대규모의 트리

플을 구축하며 링크드 데이터를 구축하였는데, 구축된 트리플 규모에 비해 인터링킹되는 트리플의 비중이 상당히 적다는 점에 주목하였다. 중점적인 분석 대상으로는 인물 정보를 선정하였는데, 이는 인물 정보가 인터링킹의 핵심적인 요소로 활용되기 때문이다. 인물은 그들의 생애를 통해 다양한 지적·예술적인 결과물을 남기는 한편 개인적인 생활인으로서의 면모도 가지고 있는데, 다양한 활동의 결과물이 복수의 기관에 분산되어 관리되고 제공되고 있다.

이에 본 연구에서는 국가대표도서관인 국립중앙도서관의 링크드 데이터를 대상으로 인물 정보가 표현되고 연계되는 방식을 분석하고자 하였다. 도서관에서 인물은 특정 저작의 저자나 주제로 취급되어 적합한 서지정보로 안내하는 역할을 주로 담당해 왔는데, 인물 정보가 링크드 데이터로 구축되거나 개방되면 저자나 주제가 아닌 인물 그 자체로서의 정보와 연계가 가능해지기 때문이다. 이를 위해 본 연구에서는 인물 정보가 식별되고 연계되는 방식에 대한 사례 분석을 수행하고, 현재 국립중앙도서관 링크드 데이터에서 인물 정보가 기술되고 연계되는 현황을 조사하였다. 그리고 그 결과를 바탕으로 효과적인 인물 정보의 구조화 방안에 대해 제안하고자 한다.

2. 인물 정보의 식별과 연계

2.1 인물 정보의 식별 데이터

2.1.1 가상의 국제 선거파일

가상의 국제 선거파일(Virtual International


Authority Files, VIAF)은 각국의 전거파일을 통합하여 클러스터화한 웹 기반 서비스이다. VIAF는 일종의 전거 데이터이므로 인물 정보는 인물 자체로 제시하고, 이 전거 데이터를 이용하는 도서관에서는 해당 인물 정보를 저작의 저자나 주제로 활용할 수 있도록 한다(VIAF, 2017). 예를 들어 ‘랑가나단(Ranganathan, Shiyali Ramamrita, 1892-1972)’의 경우에는 콜론분류법을 비롯한 많은 저서의 저자일뿐 아니라 후대 연구자들의 연구대상이 되고 있다. 랑가나단을 VIAF에서 검색하여 인명이 포함된 서지정보를 확인할 수 있다. 특히 ‘Title’ 필드를 보면 랑가나단의 저서뿐 아니라 랑가나단을 기념하는 저서와 관련 서지정보도 함께 검색된다(〈그림 1〉참고).

VIAF 데이터가 MARC 형식의 도서관 전

거레코드를 바탕으로 구축되었으므로 국가별 MARC 필드가 다양할 수 있지만, 대체로는 인물이 저자로 사용된 경우에는 100/700 필드에 랑가나단이 포함된 서지레코드의 표제를 가져오고, 주제로 사용된 경우에는 600 필드에 랑가나단이 포함된 서지레코드의 표제를 가져온 것으로 볼 수 있다. 이는 ‘랑가나단’이라는 인물 정보가 저자뿐 아니라 저작의 주제로도 사용될 수 있고, 전거 데이터에서는 이를 모두 연결해 주어야 하기 때문이다. 즉, VIAF가 제공하는 퍼머링크와 같은 식별기호와 ISNI 번호와 같은 연계정보는 랑가나단이 저자와 주제로서 다양한 서지정보와 결합할 수 있도록 지원할 수 있다.

2.1.2 국제표준이름식별기호

국제표준이름식별기호(International Standard



Ranganathan, Shiyali Ramamrita, 1892-1972

Ranganathan, S.R. (Shiyali Ramamrita), 1892-1972

Ranganathan, S.R., 1892-1972

Ranganathan, S.R.

Ranganathan, Shiyali R. 1892-1972

S. R. Ranganathan Indian mathematician and librarian.

Ranganathan, Shiyali Ramamrita

1972-1892 رانجاناثان، ش. ر.

Ranganathan, S. R. (Indian scholar, librarian, 1892-1972)


Ranganathan, Shiyali Ramamrita 1892-1972

VIAF ID:49268668 (개인)

링크://viaf.org/viaf/49268668

ISNI:0000_0001_2131_9756

VIAF의 랑가나단 전거 데이터

제목	출처
Developing horizons in library & information science : Ranganathan's tenth death anniversary commemoration volumes	'주제'로서의 랑가나단
Dictionary catalogue code by S. R. Ranganathan,....	'저자'로서의 랑가나단
Difficulties of government documents : James Childs	
Documentation : genesis and development	

〈그림 1〉 VIAF에서의 인물 정보의 표현(저자와 주제로서의 인물)

Name Identifier, ISNI)는 연구자나 작가, 예술가와 같은 콘텐츠 창작자 및 배포자를 식별하기 위한 ISO 표준 식별번호로 16자리 숫자로 구성된다(ISNI International Agency, 2017). ISNI 데이터의 상당 부분이 VIAF를 바탕으로 하고 있기 때문에, 특정 인물에 대한 ISNI 엔트리를 보면, 저자로서 인물과 주제로서 인물을 모두 확인할 수 있다. 예를 들어 ISNI에서 '랑가나단'의 검색결과를 보면, ISNI 번호 및 동일 인물의 다양한 표현과 함께 관련 저작의 표제를 함께 확인할 수 있다(〈그림 2〉 참고). 관련 표제 정보는 저자와 주제의 역할이 이용자가 확인할 수 있도록 구분되어 있지 않았으나, 동일 인물을 통해 저자와 주제가 모두 연계될 수 있는 것을 확인할 수 있다.

2.2 인물 정보의 연계 어휘집

2.2.1 FOAF

'친구의 친구'라는 의미인 FOAF는 친구를 통해 친구를 만들어가는 사회적 관계를 OWL/RDF(Resource Description Framework)로 표현한 온톨로지이다. 다양한 소셜 네트워크 서비스에서 사용자들을 통합하고 이들 간의 관계를 설정하기 위한 사용자 프로필을 기술하는데 활용된다. FOAF는 13개의 클래스와 62개의 속성으로 구성되는데, 가장 핵심 클래스인 Person 클래스는 Agent의 하위 클래스이며 knows와 함께 사용된다(Brickley & Miller, 2014). 이와 같이 FOAF는 도서관이나 기록관과 같은 특정 도메인을 제한하지 않고 전거 데이터로서

ISNI: 0000 0001 2131 9756

Name: Ranga-Nathan, Siyari Ramamrita
Ranganathan
Ranganathan (matematico e bibliotecario indiano)
Ranganathan, S. R.
Ranganathan, Schijali Ramamrita
Ranganathan, Shijali Ramamrita
Ranganathan, Shiyali R.
Ranganathan, Shiyali Ramamrita
Ranganathan, Shiyali Ramamrita (Rao Sahib)
Ranganathan, Shiyali Ramamrita
S. R. Ranganathan (Indiaas wiskundige (1892-1972))
S. R. Ranganathan
S. R. Ranganathan
Shiyali Ramamrita
Shiyali Ramamrita
Shiyali Ramamrita
Шижали Раманатан
ر. رانجاناثان، ش.
ر. رانجاناثان، ش.
ر. رانجاناثان، ش.

ISNI의 랑가나단 전거 데이터

Related names

Titles:

- 8. All India Library Conference 20-22. Jan. 1949, Nagpur ; Library needs of renascent India ; Presidential address
- Bhagavad gita, The
- Cataloguing practice
- Cinc lleis de la biblioteconomia
- Classified Catalogue Code
- Colon classification
- Conflict of authorship; corporate body vs corporate body
- Developing horizons in library & information science : Ranganathan's tenth death anniversary commemoration volumes
- Dictionary catalogue code, by S. R. Ranganathan,...
- Difficulties of government documents : James Childs
- Documentation and its facets, being a symposium of seventy papers by thirty-two authors
- Documentation genesis and development
- Dr. S. R. Ranganathan : a bibliography : to mark the inauguration ceremony of International Conference on Ranganathan's Philosophy, 11th Nov., 1985.
- Education for leisure.

'저자'로서의 랑가나단

'주제'로서의 랑가나단

〈그림 2〉 ISNI에서의 인물 정보의 표현(저자와 주제로서의 인물)

인물 정보 자체를 기술하므로, 서지 데이터와 연계할 경우에는 저자나 주제를 나타내는 인물 정보와 링크할 수 있다. 그런데 서지 데이터에서 FOAF 어휘는 주로 서지 데이터의 인물 정보를 표현하는 데에 활용된다. <그림 3>은 KERIS에서 제공하는 학술관계분석서비스인 SAM LOD의 검색화면으로 서지정보에서 저자 데이터를 확장하는 것을 보여주고 있다. 'rdf:type'는 자료유형을 보여주는 것으로, 해당 데이터인 '박지영(Zi-young Park)'이 FOAF의 Person 클래스의 인스턴스이며 Person의 속성을 통해 관련 정보가 제공될 수 있음을 확인할 수

있다.

<그림 4>는 FOAF 웹사이트에서 제공하는 예시로, 인물 정보를 상세하게 표현할 수 속성들을 보여주고 있다. 이러한 속성들은 전거레코드의 데이터로 활용될 수 있으며, 이는 외부 데이터와의 연계를 통해 확보할 수도 있다.

2.2.2 SKOS

SKOS(Simple Knowledge Organization System)의 목적은 시맨틱웹에서 시소러스나 분류표, 택사노미 등과 같은 지식조직체계(knowledge organization systems, KOS)의 사용을 지

<그림 3> KERIS LOD 인물 정보의 표현(저자로서의 인물)

```
<foaf:Person rdf:about="#danbri" xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <foaf:name>Dan Brickley</foaf:name>
  <foaf:homepage rdf:resource="http://danbri.org/" />
  <foaf:openid rdf:resource="http://danbri.org/" />
  <foaf:img rdf:resource="/images/me.jpg" />
</foaf:Person>
```

<그림 4> FOAF 예시

(출처. http://xmlns.com/foaf/spec/#term_Person)

원하는 명세를 개발하기 위한 것으로, RDF를 사용해서 KOS를 표현하는 표준 방식을 제공한다. 이 때 SKOS 개념 모델의 가장 기본 요소인 skos:Concept 클래스는 객체(object)나 의미(meaning), 아이디어 또는 이벤트를 표현하는데 사용된다(W3C, 2012).

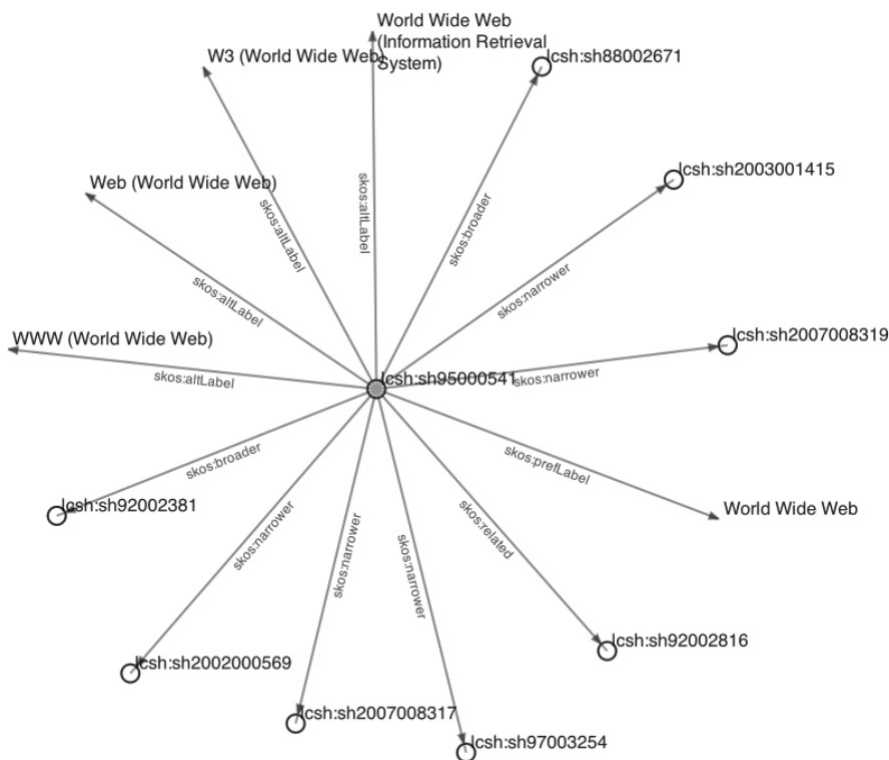
〈그림 5〉는 LCSH를 SKOS에 적용한 예로, 우선어와 비우선어 그리고 상위 및 하위관계의 용어들을 하나의 그림으로 표현한 것이다. 이러한 관계는 인물 정보를 표현하는데도 적용할 수 있다. 대표 접근점은 우선어('preLabel')로 지정하고 이외 접근점은 'altLabel' 속성으로

제시할 수 있다. 또한 단체명인 경우에는 계층 표현을 통해 상위기관이나 하위기관과의 연결을 시도할 수 있다.

2.3 서지 LOD에서 인물 정보의 기술

2.3.1 미국국회도서관 LOD

미국국회도서관(LC)에서는 2012년부터 Linked Data Service: Authorities and Vocabularies (LDS)를 통해 다양한 전거데이터를 제공하고 있다(Library of Congress, 2016). LC는 LCSH를 RDF로 표현하여 링크드 데이터를 발행하고



〈그림 5〉 LCSH를 SKOS에 적용한 예시
 (출처, Summers, Isaac, Redding, & Krech(2008), p. 28
<http://dcpapers.dublincore.org/pubs/article/viewFile/916/912>)

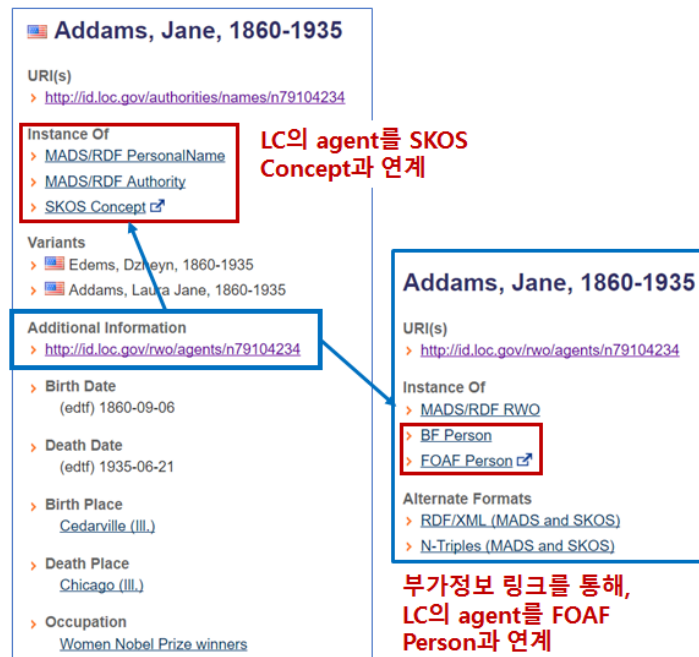
있으며 MARC 21 포맷으로 되어 있는 전거 데이터를 SKOS 어휘를 이용하여 RDF로 변환하고 이를 웹상에 발행하고 있다(Library of Congress, 2016). LC LOD에서는 하나의 인물에 대해 SKOS Concept와 FOAF Person으로 동시에 연계하고 있다. <그림 6>을 보면, 'Addams, Jane'이라는 인물이 기본적으로 'PersonalName'이면서 'SKOS Concept'임을 표시하였고, 부가정보 메뉴를 통해 'FOAF Person'과 'BF(Bibframe) Person'으로도 연계하고 있다.

2.3.2 영국국가도서관 LOD

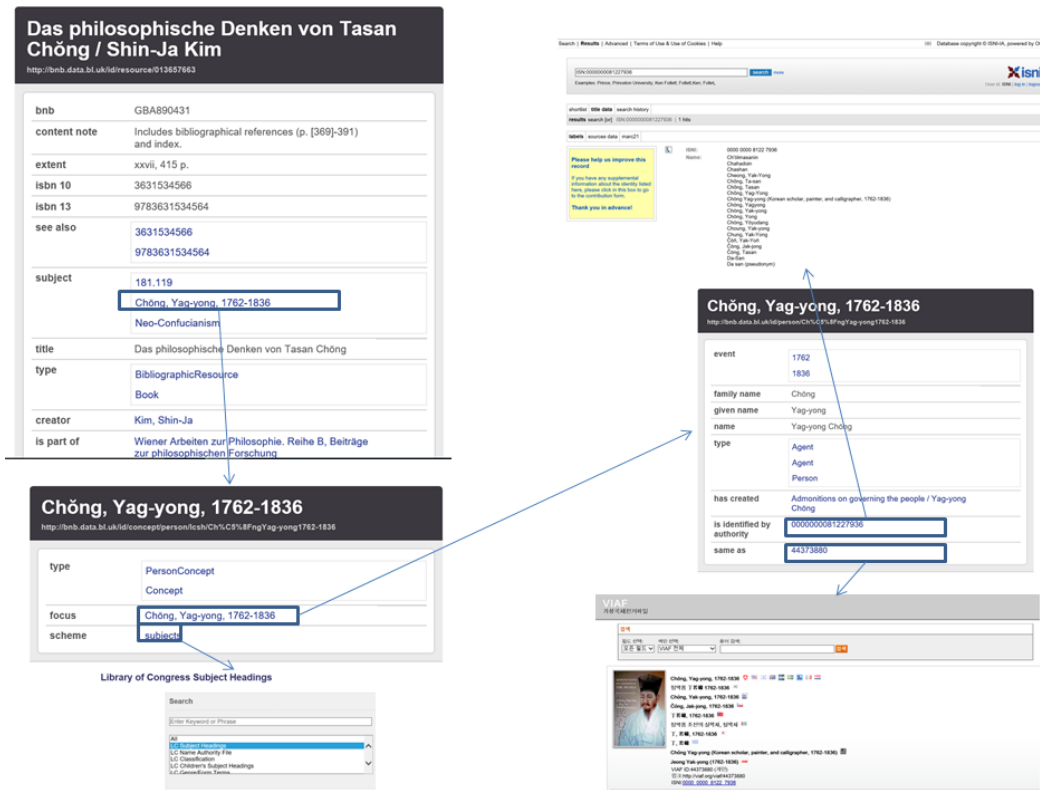
영국국가서지(British National Bibliography, BNB) LOD 프로젝트는 2011년 착수하여, 2013년 서비스를 시작하였다(British Library, 2016a).

데이터 기술을 위해서는 Bibliographic Ontology, British Library Terms, Dublin Core, Event Ontology, FOAF: Friend of a Friend, RDF Schema, SKOS 등 여러 어휘를 사용하였다(British Library, 2016b). BNB LOD에서 정약용의 검색결과를 보면, 검색어 'Ch'ong, Yag-yong'은 인물(person)과 주제(concept)로서 2건이 검색된다. 먼저 저자로서 검색결과를 보면, '유형' 값인 Agent와 Person은 FOAF와 연계된다. 또한 'same as'는 VIAF와 연계되며, 'is like'는 ISNI와 연계된다. 주제로서 검색결과를 보면, '유형' 값은 PersonConcept, skos:Concept와 연계된다. BNB LOD에서도 인물 정보가 저자로 표현될 때와 주제로 표현될 때를 연계하고 있다(<그림 7> 참고).

미국국회도서관 저자 LOD 일부



<그림 6> LC LOD 서비스의 저자 정보 표현 예시



〈그림 7〉 BNB LOD의 정약용 검색결과 - 상세화면
 <http://bnb.data.bl.uk/id/person/Ch%C5%8FngYag-yong1762-1836>
 <http://bnb.data.bl.uk/id/concept/person/lcsh/Ch%C5%8FngYag-yong1762-1836>

3. 인물 정보의 식별 및 연계 범위 확장

3.1 인물 정보의 기술 현황 분석

3.1.1 저자로서의 인물 정보 표현 및 연계
 국립중앙도서관 LOD(국립중앙도서관 n.d.-b)에서 '정약용'을 검색하여 저자로서의 인물 정보 표현 방식을 분석하였다. 저자로서 활용된 결과를 보면, 이름 전거 데이터에 해당하는 기본 정보를 확인할 수 있으며, 연결 어휘는 '같은

개체' 항목에서 확인할 수 있다(〈그림 8〉 참고). 이 때 '자료유형' nlon:Author는 foaf:Person과, '태어난 해와 사망한 해' nlon:birthYear, nlon:deathYear는 foaf:Person과 연계된다. '같은 저자가 저작한 도서'의 항목값인 'nlk:KCF200500121'을 선택하면 저작의 상세정보로 연결된다. '같은 개체'는 owl:sameAs를 사용하여 kdata, LC LOD, 한국사 LOD, nla 저자 정약용, VIAF로 연결되지만, 연계에 인물 자체만을 접근점으로 삼고 있어 같은 개체의 효과적인 연계방안이 필요하다.

정약용	
http://lod.nl.go.kr/resource/KAC200105569 (정약용)	
저자명 foaf:name	정약용
명칭 rdfs:label	Chông, Yag-yong Jeong, Yakyong Jung, Yakyong Tasan 귀농, g歸農 다산, g茶山 문도, g文度 문암일인, g門巖逸人 미용, g美鏞 [more...]
태어난_해 nlon:birthYear	1762 (xsd:int)
사망한_해 nlon:deathYear	1836 (xsd:int)
같은_저자가_저작한_도서 nlon:create	nIk:JMO201203423 nIk:KCF200100385 nIk:KCF200200480 nIk:KCF200302365 nIk:KCF200500121 nIk:KDM198923023 nIk:KDM199116832 nIk:KDM199118153 nIk:KDM199305560
LOD 발행일 nlon:datePublished	2016-12-08T10:22:31 (xsd:dateTime)
수정날짜 dcterms:modified	2016-11-08T16:21:06 (xsd:dateTime)
같은개체 owl:sameAs	kdata:Jeong_Yak-yong http://lod.koreanhistory.or.kr/resources/인명_정약용 nla:정약용 http://viaf.org/viaf/44373880
정보원(dc_source) dc:source	목민심서, (늘푸른소나무), 2012 (xsd:string) 위키백과(http://ko.wikipedia.org/wiki/정약용) (xsd:string) 한국문학 작가명 권거 데이터베이스 파일 (xsd:string) 流配地에서 보낸 편지, (詩人社), 1979 (xsd:string)
자료유형 rdf:type	nlon:Author

〈그림 8〉 국립중앙도서관 정약용 저자검색결과 - 상세화면
 〈<http://lod.nl.go.kr/page/KAC200105569>〉

3.1.2 주제로서의 인물 정보 표현 및 연계
 ‘정약용’에 대한 주제 검색 24건을 보면, 정약용의 자와 호에 대한 검색결과가 나타났다. 주제 검색의 결과 값은 저자명으로 검색할 경우보다 복잡했다. 검색된 2건 중 하나는 〈http://lod.nl.go.kr/subject/정약용_丁若鏞 (정약용[丁若鏞])〉으로, 다른 하나는 〈<http://lod.nl.go.kr/resource/KSH00273758> (정약용[丁若鏞])〉으로 URI는 다르지만, 클릭하면 둘 다 〈<http://lod.nl.go.kr/resource/KSH00273758> (정약용[丁若鏞])〉으로 이동했다. 주제명의 경우 SKOS 어휘를 활용하여 표현되고 있었지만, ‘같은개체 (owl:sameAs)’의 경우에는 링크 정보가 외부 인터링킹이 아닌 동일 URI인 〈<http://lod.nl.go.kr/resource/KSH00273758> (정약용[丁若鏞])〉으로 연계되었다. 또한 국립중앙도서관 LOD

에서는 저자로서의 정약용과 주제로서의 정약용의 연계 정보를 제공하지 않고 있다(〈그림 9〉 참고).

3.2 인물 정보의 식별 및 연계 범위 확장 방안

3.2.1 주제로서의 인물정보 표현 방식 개선
 인물정보의 연계성을 향상시키기 위해서는 ‘인물 정보’가 주제명으로 표현되었을 때, 이를 ‘개념’(concept)뿐 아니라 ‘인물’(person)을 나타내는 LOD 어휘로 식별할 수 있어야 한다. 즉, 저자로서의 인물정보를 보면, 〈그림 10〉과 같이 KORMARC 245와 700 필드에 기술된 인물정보(국립중앙도서관 n.d.-a)가 LOD에서도 ‘dc:creator’와 ‘dcterms:creator’로 표현되어 있다. 그런데

정약용[丁若鏞]	
http://lod.nl.go.kr/resource/KSH00273758 (정약용[丁若鏞])	
명칭 rdfs:label	정약용[丁若鏞]
SKOS_prefLabel skos:prefLabel	정약용[丁若鏞]
변경사항_changeNote skos:changeNote	nlk:bksh_00273758
정의 skos:definition	1762~1836. 조선 후기의 학자·문신. 본관 나주(羅州). 자 미용(美鏞)·송보(頌甫). 호자 귀농(歸農). 호 다산(茶山)·삼미(三眉)·여유당(與猶堂)·사암(侯菴)·자하도인(紫霞道人)·탁옹(鑿翁)·태수(苔叢)·문암일인(門巖逸人)·철마산초(鐵馬山樵). 가톨릭 세례명 요안.
SKOS_Scheme skos:inScheme	nlk:subjectTermScheme
해당_주제의_도서 nlon:isSubjectOf	nlk:CNTS-00030341059 nlk:CNTS-00030415126 nlk:CNTS-00030534125 nlk:CNTS-00054074793 nlk:CNTS-00056837640 nlk:CNTS-00056965138 nlk:CNTS-00057006993 nlk:CNTS-00059880758 nlk:CNTS-00059880759 [more...]
같은개체 owl:sameAs	nl:정약용_丁若鏞
자료유형 rdf:type	nlon:Concept

〈그림 9〉 국립중앙도서관 정약용 주제검색결과 - 상세화면
(<http://lod.nl.go.kr/page/KSH00273758>)

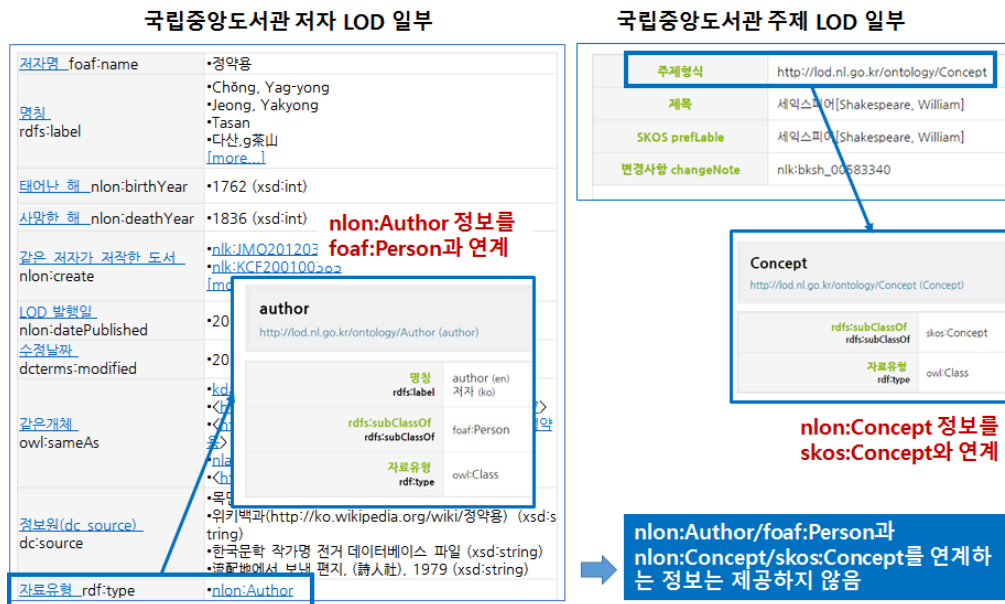
245	00	▼a목민심서/▼d정약용 지음;▼e이민수 옮김
250		▼a3판
260		▼a서울;▼b범우사;▼c2000
300		▼a364p.;▼b도판;▼c21cm
440	00	▼a사르비아총서;▼v201
653		▼a목민심서
700	1	▼a정약용;▼d1762-1836▼0KAC200105569

속성	값
제목 dcterms:title	목민심서
저자명 dc:creator	이민수 정약용
발행지 nlon:publicationPlace	서울
발행처 dc:publisher	범우사
발행년 dcterms:issued	2000
판표시 bibo:edition	3판
크기(수량) bibframe:extent	364p 21cm 도판
총서표제 nlon:titleOfSeries	사르비아총서
총서번호 nlon:volumeOfSeries	201
주제명 dcterms:subject	nlk:KSH00365456
키워드 dc:subject	목민심서
저자명 dcterms:creator	nlk:KAC200105569 nlk:KAC201313019

〈그림 10〉 MARC 기반 서지정보와 LOD 기반 서지정보에서 인물-저자정보의 변환 사례

인물은 특정 정보자원의 저자일뿐 아니라 다른 정보자원의 주제 역할도 수행할 수 있다. 더욱이 특정 개인의 지명도가 높을수록 해당 인물은

저자뿐 아니라 주제명으로도 빈번히 나타날 수 있다. 그러나 〈그림 11〉과 같이 기존의 MARC 데이터를 LOD로 변환할 때, 저자로서의 인물은



〈그림 11〉 국립중앙도서관 LOD에서 인물정보의 표현 예시

‘author’나 ‘person’으로 표현하고, 주제로서의 인물은 ‘concept’로 표현하고 있다. 인물은 복수의 LOD간 인터링킹의 주요 접점이 될 수 있지만 이와 같이 주제로서의 인물이 ‘개념’(concept)으로만 표현된다면, 개체의 식별과 연계가 어렵다.

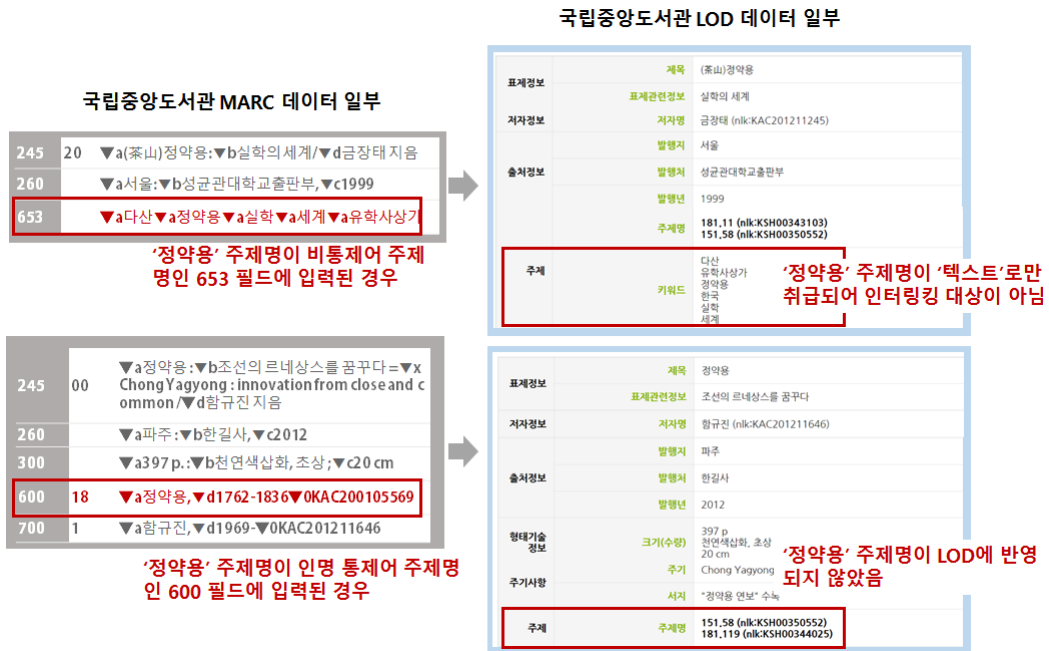
3.2.2 저자로서 인물정보와 주제로서

인물정보의 연계

도서관 분야에서 인물정보는 저자 이외에 주제명으로도 취급된다. 즉, MARC 형식의 데이터에서 주제로서의 인물정보는 〈그림 12〉와 같이 비통제어와 통제어 주제명 필드에 입력된다. 그러나 MARC 데이터를 LOD로 변환한 결과를 보면, 비통제어 색인의 경우에는 문자열로만 취급되므로 URI 기반 인터링킹이 불가능하다. 게다가 통제어 색인으로 600 필드에서 전거 번호를 가지고 있던 경우에도 LOD에 반영되

지 않고 있어, 역시 인터링킹 대상이 되지 못한다. 이와 같은 문제점은 복수의 LOD 간에 체계적인 인터링킹 정보 구축을 어렵게 할 것이다. 또한 〈그림 13〉을 보면, MARC 데이터의 오류가 LOD에 반영되기도 한다. 인물 정보를 주제 접근점으로도 표현하기 위해서는 우선 원 데이터에서 인물이 주제명으로 부여되어 있어야 한다.

LOD 생태계에서 인물을 표현하는 방법은 매우 다양하다. 각 분야의 관점을 반영하여, ‘author’나 ‘creator’로 표현할 수 있고, 간단히 ‘related Person’ 정도로 나타낼 수도 있다. 따라서 도서관 분야에서는 분야의 관점을 명확히 살린 어휘를 기본적으로 사용하되, 중립적인 인물 표현을 통해 허브 역할을 할 수 있는 LOD 어휘와의 링크정보를 제공하는 것이 바람직하다. 인물이 저자로서의 사람과 주제로서의 사람 역할을 주로 하고 있다면, LC의 사례와 같이 FOAF:



<그림 12> MARC 기반 서지정보와 LOD 기반 서지정보에서 인물-주제정보의 변환 사례



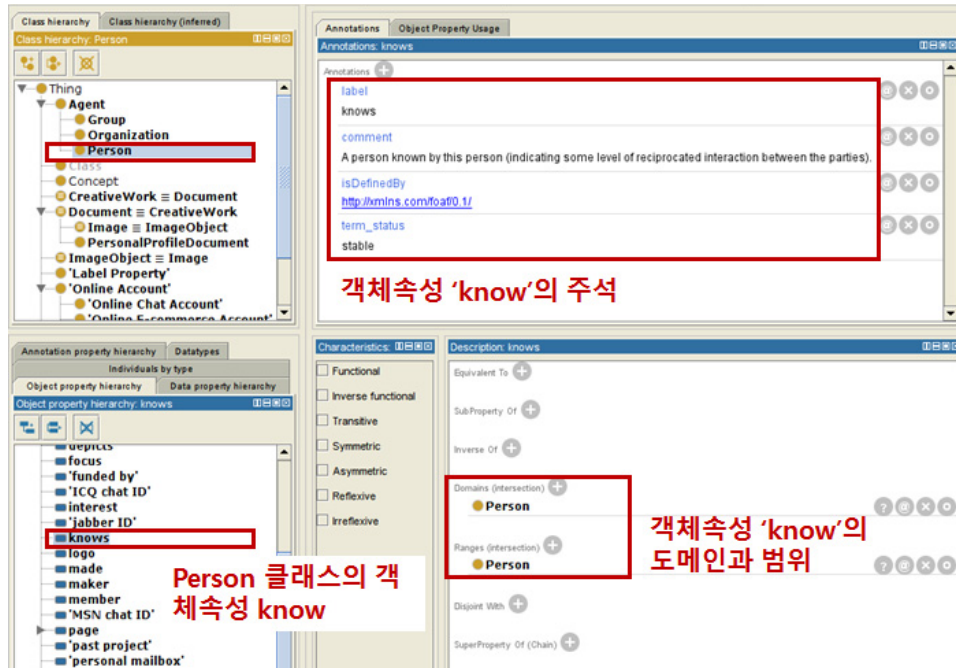
<그림 13> MARC 데이터의 주제명 오류가 LOD로 이어진 사례

Person과 SKOS:Concept을 연계하는 것이 바람직하다. FOAF는 인맥관리를 위한 메타데이터 어휘집합을 제공하며, LOD로 발행되었다(Brickley & Miller, 2014). 인물정보의 인터링킹에 기본적으로 FOAF 어휘집합을 사용할 수 있는 것이다. FOAF 기반 인물 기술정보의 어휘집합(클래스, 객체속성, 데이터속성)을

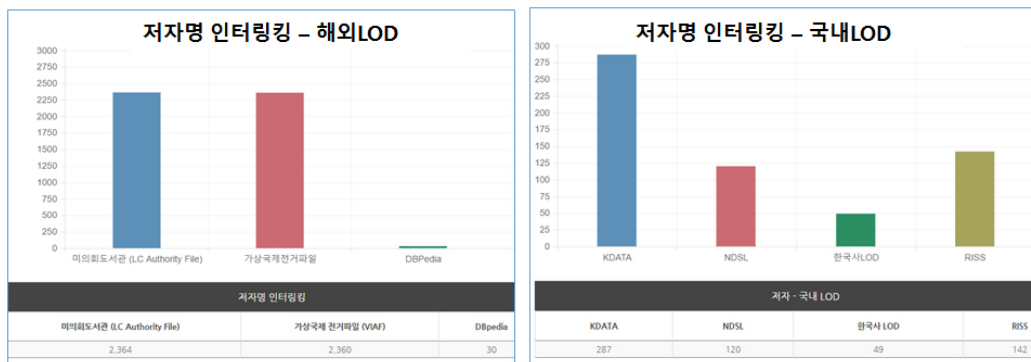
Protege 프로그램을 통해 반입하면, <그림 14>와 같다.

3.2.3 인물 관련 부가 링크 정보의 구축 및 연계

인물과 관련된 부가 정보는 인물정보의 연계 범위를 확장시킬 수 있는데, LOD에서의 저자명



〈그림 14〉 FOAF 어휘집합을 protege에 반입한 화면



〈그림 15〉 국립중앙도서관 저자명 인터링킹 통계
 〈<http://lod.nl.go.kr/home/about/stats.jsp>〉

인터링킹은 'owl:sameAs'를 통한 맵핑으로 이루어진다. 국립중앙도서관 LOD의 경우 저자명 간의 인터링킹 트리플수는 4,754개인데, 국외에서는 LC와 VIAF의 동일 저자명 인터링킹, 국내에서는 KDATA, NDSL, 한국사LOD, RISS 중심

의 동일 저자명 인터링킹이 있다(그림 15) 참고). 그런데 상이한 LOD 간의 연계에 인물 자체만을 접근점으로 삼는 것은 인터링킹의 확장을 제한할 수 있다. 예를 들어, 국립중앙도서관 LOD에서 제공하는 '정약용'의 인터링킹 정보는 아

래와 같다. 동일 인물 매핑을 바탕으로 하므로, 5개의 URI는 모두 인물 '정약용'을 나타내는 데이터이다. 그런데 분야마다 인물정보를 기술하는 관행은 다양하다. 아래의 데이터가 모두 정약용을 가리키고 있지만, 각 데이터가 제공하는 부가정보는 상이할 것이다.

- 국립중앙도서관: <<http://lod.nl.go.kr/resource/KAC200105569>>
- LC: <<http://id.loc.gov/authorities/names/n81046897>>
- 한국사 LOD: <http://lod.koreanhistory.or.kr/resources/인명_정약용>
- VIAF <<http://viaf.org/viaf/44373880/>>
- KDATA: <http://data.kdata.kr/page/Jeong_Yak-yong>

그리고 이러한 부가정보는 2단계나 3단계를 거쳐서 인물과 인물을 다시 연결해 줄 수 있다. 인물은 그가 남긴 다양한 작품과 관련 활동을 이용자와 연계해 주는 중요한 연결고리가 될 것이며, 인물명만으로는 연계 정보를 확보하기가 어렵다. LOD의 핵심은 상이한 주체들이 구축한 데이터의 개방적인 공유이다. 이와 같은 LOD의 속성이 인물의 부가정보 구축에 기여할 수 있을 것이다.

4. 결론

도서관 분야에서는 정보자원의 지적·예술적 책임을 진 주체를 서지정보와 분리하여 전거레

코드에 구조화하고, 이를 서지레코드와 연계해 왔다. 이와 같은 방식은 오랜 기간 동안 메타정보를 구축하는 과정에서 축적된 노하우를 바탕으로 한 데이터의 정규화로 간주되며, 매우 효과적이고 체계적인 데이터 관리 방식이다. 그런데 이 과정에서 전거 정보는 주로 서지 정보에 도달하기 위한 중간 과정으로 취급되었다. 즉, 전거 정보의 대부분을 차지한 인물 정보를 정보자원의 저차로만 관리한 것이다. 게다가 이와 같은 관점은 기존의 전거 데이터나 서지 데이터를 링크드 데이터로 변환한 이후에도 유지되는 경향이 있다. 그리고 이로 인해 이름전거 정보에 포함된 인물과 주제전거 정보에 포함된 인물을 서지정보에서 통합적으로 다루지 못하는 문제점이 발생할 수 있다.

본 연구에서는 위와 같은 문제점을 집중적으로 분석하기 위해 국립중앙도서관의 링크드 데이터를 대상으로 인물 정보가 표현되고 연계되는 방식을 확인하였다. 그 결과, 저자로서의 인물 정보는 링크드 데이터에서도 인물을 표현하는 어휘와 연계되어 기술되고 있는 반면에, 주제로 표현된 인물은 국립중앙도서관의 주제명 부여 방식의 특성으로 인해 개념으로만 취급되고 있었다. 이에 본 연구에서는 주제로서의 인물 정보를 표현할 수 있도록 하고, 저자로서의 인물 정보와 주제로서의 인물 정보가 연계될 때 링크드 데이터의 가능성을 더욱 확장시킬 수 있다고 판단하였다. 그리고 이와 더불어 인물과 관련된 부가 링크 정보를 풍부하게 제공함으로써 궁극적으로 국가서지의 중요한 부분을 차지하는 국립중앙도서관의 서지 데이터의 활용도를 제고하기 위한 제안을 하였다.

참 고 문 헌

Brickley, D., & Miller, L. (2014, January 14). FOAF vocabulary specification 0.99. Retrieved from <http://xmlns.com/foaf/spec/>

Summers, E., Isaac, A., Redding, C., & Krech, D. (2008, September). LCSH, SKOS and linked data. Paper presented at the Proc. Int'l Conf. on Dublin Core and Metadata Applications, Berlin. Retrieved from <http://dcpapers.dublincore.org/pubs/article/viewFile/916/912>

[참고 웹사이트]

국립중앙도서관 (n.d.-a). Retrieved from <http://www.nl.go.kr/>

국립중앙도서관 (n.d.-b). 링크드데이터. Retrieved from <http://lod.nl.go.kr>

국립중앙도서관 (2017). 국가전자과일. Retrieved from <http://www.nl.go.kr/authorities/naf/history>

국사편찬위원회 (n.d.). 한국사 LOD. Retrieved from <http://lod.koreanhistory.or.kr>

한국교육학술정보원 (n.d.). SAM Linked Open Data. Retrieved from http://data.riss.kr/sam_lod/main.do

British Library (2016a). Retrieved from <http://www.bl.uk/>

British Library (2016b). British national bibliography as linked open data. Retrieved from <http://bnb.data.bl.uk/>

FOAF (n.d.). Retrieved from <http://www.foaf-project.org>

ISNI International Agency (2017). Retrieved from <http://www.isni.org>

KDATA (n.d.). Linked Data for Korea. Retrieved from <http://kdata.kr>

Library of Congress (2016). Linked data service. Retrieved from <http://id.loc.gov/>

VIAF (2017). Retrieved from <http://viaf.org>

W3C (2012). SKOS Home. Retrieved from <http://www.w3.org/2004/02/skos/>

- | |
|---|
| <ul style="list-style-type: none">• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean) |
|---|

KERIS (n.d.). SAM Linked Open Data. Retrieved from http://data.riss.kr/sam_lod/main.do

National Institute of Korean History (n.d.). Korean History Linked Open Data. Retrieved from <http://lod.koreanhistory.or.kr>

NLK (n.d.-a). Retrieved from <http://www.nl.go.kr/>

NLK (n.d.-b). LOD. Retrieved from <http://lod.nl.go.kr>

NLK (2017). Authorities. Retrieved from <http://www.nl.go.kr/authorities/naf/history>

