

관측분야 과학데이터 관련 메타데이터 상호운용성 확보를 위한 OpenURL 메타태그 연구

A Study on the OpenURL META-TAG of Observation Research Data for Metadata Interoperability

김 선 태* · 이 태 영**

Sun-Tae Kim · Tae-Young Lee

차 례

- | | |
|--------------------|----------------|
| 1. 서 론 | 5. ORD 메타태그 개발 |
| 2. 과학데이터 양상 | 6. 결 론 |
| 3. 과학데이터의 메타데이터 분석 | · 참고문헌 |
| 4. OpenURL 메타태그 분석 | |

초 록

여러 소스에 출현하는 관측분야의 과학데이터가 OpenURL 서비스 아키텍처에서 원활하게 유통되도록 하기 위해, OpenURL 소스와 서비스 구성요소간 상호운용성 확보를 할 수 있도록 Key/Encoded-Value(KEV) 형식으로 OpenURL 핵심 메타태그를 제안하였다. 현재의 OpenURL은 문헌위주의 학술 정보 유통에 그 초점이 맞추어져 있으며, 과학데이터 유통을 위한 메타태그는 제공하지 못하고 있다. 과학데이터 기술을 위한 DataCite 컨소시엄의 메타데이터 요소를 기준으로 하고, 더블린 코어의 핵심 메타데이터 요소, 데이터세트와 테이블을 인용하기 위해 OECD에서 만든 메타데이터 요소, 지구관측 분야의 메타데이터 표준인 DIF 메타데이터 요소를 비교, 분석하여 관측분야의 과학데이터를 위한 OpenURL 핵심 메타태그를 도출하였다.

키 워 드

연구데이터, 과학데이터, 메타태그, OpenURL, 관측연구

* 한국과학기술정보연구원 정책연구실 선임연구원
(Senior Researcher, Dept. of Policy Research, Korea Institute of Science and Technology Information, stkim@kisti.re.kr)

** 전북대학교 문헌정보학과 교수
(Professor, Library and Information Science, Chonbuk National University, taehyun@jbnu.ac.kr)

• 논문접수일자: 2011년 3월 2일
• 최종심사(수정)일자: 2011년 3월 20일
• 게재확정일자: 2011년 4월 2일

ABSTRACT

This paper presents a core meta-tag of OpenURL written in Key/Encoded-Value format in the field of observation research, to distribute the scientific data, produced in many experiments and observations, on the OpenURL service architecture. So far, the OpenURL hasn't supplied a meta-tag represented scientific data because it has focused on circulation of scholarly and technological information extracted from thesis, proceedings, journals, literatures, etc. The DataCite consortium metadata were analyzed and compared with the Dublin Core metadata, OECD metadata, and Directory Interchange Format metadata to develop a core meta-tag in observation research.

KEYWORDS

OpenURL, META-TAG, Scientific Data, Research Data, Z39.88, Observation Research Data

1. 서론

과학 연구자가 수행하는 연구 패러다임이 지속적으로 변하여 최근에는 데이터가 연구의 중심이 되고 있다. Green(2009)은 데이터 수집이 연구의 대부분이라고 주장하였으며 학위 논문이나 저널 논문, 단행본도 데이터 없이는 쓰일 수 없음을 강조하였다. 인터넷 이전에는 데이터를 손쉽게 얻는 것이 불가능 하였지만, 1995년 이후 인터넷이 활성화되고 최근 들어 하드웨어의 눈부신 발전과 고성능 네트워크, 센서장비의 확산이 여러 분야에서 다양한 포맷의 데이터를 발생시키고 있다. 따라서 도서관이나 정보센터는 연구자들의 새로운 연구

패러다임을 지원하기 위하여 기존의 학술논문이나 특허, 연구보고서 중심의 서비스에 데이터¹⁾ 서비스를 추가 확대할 필요가 있다.

도서관이나 정보센터에서 데이터 서비스를 추가하는 것은 문헌위주의 과학기술정보 서비스 체계를 구성하는 각각의 구성요소가 관찰/실험 데이터도 처리할 수 있게 확장되어야 함을 의미한다. 과학데이터 관련 서비스를 확대하는 데 있어 OCLC(2005)의 보고서가 지적인 사항을 고려해야 한다. 그 보고서에 따르면 학술연구자들의 1%만이 기관에서 제공한 OPAC을 활용한 서비스를 이용하고 있으며, 연구자들의 90%는 검색포털 엔진(예: 구글, 야후 등)이나 이메일 등을 검색의 접점으로 이용하는 것

1) 연구자가 연구과정을 통해 생성한 모든 데이터는 연구데이터(Research Data)와 과학데이터(Scientific Data)로 불린다. 본 논문에서는 이를 혼용하여 사용한다.

을 알 수 있다. 이 결과는 연구자들이 선호하는 정보 접근 방식을 연구자 소속 기관이 충분히 검토 수용해야 함을 시사하고 있다. 다시 말하면, 연구자들이 점점으로 사용하고 있는 서비스와 연구자의 상황을 분석하여 최적의 서비스를 제공할 수 있도록 서비스 구성요소 간 상호운용성 확보가 중요하다. OpenURL²⁾은 이런 환경에서 연구자의 상황을 인지하여 서비스 할 수 있는 적절한 해결책 중 하나가 될 수 있다. 그러나 현재의 OpenURL은 학술정보 유통에 그 초점이 맞추어져 있으며, 과학데이터 유통을 위한 메타태그는 제공하지 못하고 있다.

본 연구에서는 메타태그나 데이터 기반을 알기 위해 NISO OpenURL Version 0.1과 2004년 NISTO AX위원회에서 표준으로 제정한 Z39.88-2004를 검토하였다. 과학데이터 기술을 위한 DataCite³⁾ 컨소시엄의 메타데이터 요소를 기준으로 하고, Dublin Core(이하 DC)의 핵심 메타데이터 요소, 데이터세트와 테이블을 인용하기 위해 개발된 OECD 메타데이터 요소, 지구관측 분야의 메타데이터 표준인 Directory Interchange Format(이하 DIF) 메타데이터 요소를 비교 분석하여 관측분야의 과학데이터를 위한 OpenURL 핵심 메타태그

를 Key/Encoded-Value(KEV) 형식으로 제안하였다.

2. 과학데이터 양상

2.1 과학데이터의 중요성

데이터 중심의 제 4세대 R&D 패러다임이 도래하면서 과학데이터에 대한 관심이 급증하고 있으며, 과학데이터에 대한 보존, 공유, 재사용에 대한 연구가 활발하게 진행되고 있다. 특히, 메타데이터 통합관리를 위한 표준화, 데이터에 영속적 접근제공 방법, 데이터의 인용에 관한 연구가 활발하게 진행되고 있다. 이를 하나씩 살펴보면 아래와 같다.

첫째, 분야별로 과학데이터에 대한 표준화를 주도하고 데이터를 공유하는 움직임이 활발하게 전개되고 있다. 특히 학제 간 연계·융합 연구를 위해서 이질적인 메타데이터를 의미적으로 통합하여 연구자들이 관련된 데이터를 동시에 검색하고 발견할 수 있도록 하는 서비스의 개발이 필요하게 되었다. 둘째, 발생된 데이터에 대한 영구적 접근을 제공하기 위해 데이터 식별자 부여에 대한 관심이 급증하고 있다. 국

2) Herbert Van de Sompel(<http://public.lanl.gov/herbertv/>) 박사에 의해 주창된 것으로 학술정보에 대한 서지 상황을 표준 된 포맷으로 전송하기 위해 파라미터(예를 들어 서명, 저자명, 시작페이지 등)를 정의하였다. 초기 버전은 0.1 버전으로 현재까지도 가장 많이 사용되고 있으며, NISO에 의해 OpenURL 1.0이 2004년에 표준(Z39.88-2004)으로 제정되었다.

3) DataCite는 국제적인 컨소시엄으로서 2010년 8월 말 현재, 9개국에서 12개 기관이 참여 중이며, DOI(Digital Object Identifier)가 부여된 80만 건 이상의 레코드를 확보하고 있다. DataCite는 과학데이터에 DOI를 부여하는 에이전시 역할을 수행하고 있다(김선태 외 2010, 87).

제적인 데이터 공유와 접근을 위해 독일의 국가과학기술도서관(TIB)에서 운영대행을 하고 있는 DataCite 컨소시엄이 구성된 것도 바로 이러한 이유에서이다. 셋째, 데이터의 인용을 위한 표준 포맷에 관한 연구도 활발하게 진행 중에 있다. DataCite는 과학데이터를 기술하기 위한 스키마를 최근에 발표하였으며, 데이터 인용을 위한 핵심 요소와 그 형식을 제안하였다. OECD의 경우도 데이터세트와 테이블이 인용되기 위한 핵심 요소와 그 형식을 제안하였으며, DIF의 경우도 <Data_Set_Citation>⁴⁾ 요소를 활용하여 데이터 인용과 관련된 정보를 기술할 수 있도록 하고 있다.

과학데이터와 관련하여 출판업계의 움직임도 활발하다. ALPSP⁵⁾(2008)의 3번째 보고서에 따르면 전 세계의 과학·기술·의료 분야의 400개 출판사를 대상으로 조사한 결과 160개(45%) 출판사가 그들이 발행하는 저널의 논문과 관련되어 있는 데이터로의 접근을 제공하고 있다. 한 가지 사례를 보면 '임상조사 저널(Journal of Clinical Investigation)'의 경우 1999년부터 2009년까지 데이터가 딸린 논문의 양이 0%에서 87%로 늘었으며, 2010년엔 거의 100%에 이른다(Joan Starr, Angela Gastl 2011, 1). 이러한 조사결과는 학술정보 유통에서 연구데이터의 중요성이 심화되어 가는 것을 의미한다. 또한 관련 커뮤니티에서 연

구데이터 이용의 활성화를 준비해야 함을 시사한다.

2.2 과학데이터 현황

연구패러다임의 변화 속에서 범지구적 문제의 해결을 위한 이해와 요구가 데이터를 중심으로 한 연구를 만들어 내고 있다. 특히 기후변화, 황사, 해양 오염 등 지구적 문제해결을 위해 많은 연구자들이 과학데이터를 공유하고 있으며, 대형 장비의 공동 활용을 시도하고 있다. 예를 들어 휴먼 게놈 연구 결과의 공유 등이 그러하다. 이러한 글로벌 동향의 중심에 과학데이터가 있다. 과학데이터란 연구자의 연구 활동 과정 중 생성되는 다양한 유형의 사실적 기록을 의미한다. 즉, 연구 활동을 통하여 생산된 연구 활동의 기록물로서 관측, 감시, 조사, 실험, 분석, 계산 등의 과정을 통하여 생산된 문자, 이미지, 오디오, 동영상 등의 아날로그 및 디지털 형식을 포괄하는 데이터라고 종합하여 정의할 수 있다. 따라서 그 형식이 복잡하고 널리 배포되며 범위가 넓은 것이 큰 특징이다. 또한 자료의 크기 또한 다양하다. 과학 분야의 대표 학술지인 Science는 최근 자매지 Science Signaling, Science Translational Medicine과 함께 데이터의 홍수 속에서 부각되고 있는 여러 가지 이슈사항을 다루기 위해 연구데이터에

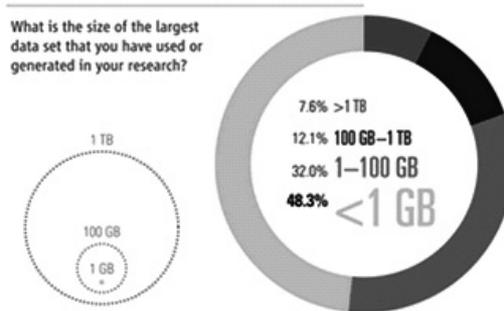
4) 본 논문에서는 가독성을 위해 모든 메타데이터 요소는 '<'과 '>' 기호로 쌍따옴표로 표현한다.

5) ALPSP(The Association of Learned and Professional Society Publishers)는 1972년에 설립된 국제적 비영리 출판사 협회이다. 2011년 1월 현재, 36개국에서 360개 이상의 멤버가 가입되어 활동하고 있다. 멤버들은 1만종 이상의 저널을 발행하고 있으며, 단행본과 보고서, 데이터베이스를 발행하고 있다(web-00001).

대한 특집호를 기획하여 발행하였다. 해당 호에는 과학데이터에 대한 설문조사 결과가 포함되어 있다. 내용은 <<http://bit.ly/pvDIAh>>에서 확인 가능하다. 설문조사는 2010년 Science 학술지 동료 연구검토자(peer reviewers)들을 표본으로 선정하여 진행되었으며, 1,700명의 연구자가 응답을 한 결과이다. <그림 1>에서와 같이 연구자들이 사용하고 있는 과학데이터는

그 크기가 다양하다. 1TB(TeraByte) 이상의 대용량 데이터를 사용하는 연구자도 7.6%이며, 1GB(GigaByte) 미만의 데이터를 사용하는 연구자도 48.3%에 이른다.

관측분야의 과학데이터가 생산되는 자료형식은 관측분야와 관측요소에 따라 그 형태가 매우 다양하다. <표 1>은 국내 지구 관측 분야별 자료형식 일부를 나타낸 것이다.



<그림 1> 과학데이터의 크기
(Science staff 2011)

<표 1> 국내 지구 관측 분야별 자료형식(이혜영, 곽승진 2007, 263 일부발췌)

관측분야	관측요소 (데이터그룹)	관측주기	자료형식
보건의료	1	실시간	DBF
에너지자원	3	-	dxf, shp
기후	3	월, 년	텍스트
수자원	17	월~연간	텍스트
기상	27	30분~12시	GRIB, BUFR
생태계/생물다양성	13	10년	GIS-DB
농업	5	5~40년	GIS파일, Excel..
산림	2	10, 100년	GIS파일
해양	32	1초~1년	DB, Seg-Y
국토공간정보	11	5~10년	백터, 라스터 ...
재해	2	-	Excel

3. 과학데이터의 메타데이터 분석

범용적인 OpenURL 핵심 메타태그를 개발하기 위하여 과학데이터 기술을 위한 DataCite 컨소시엄의 메타데이터 요소를 기준으로, DC의 핵심 메타데이터 요소와 데이터세트와 테이블을 인용하기 위해 개발된 OECD 메타데이터 요소, 지구관측 분야의 메타데이터 표준인 DIF 메타데이터 요소를 중심으로 조사하였다.

3.1 DataCite 데이터세트 분석

DataCite는 국제적인 컨소시엄으로서 2011년 1월 말 현재, 9개국에서 12개 기관이 멤버로 활동 중이며, 2개 기관이 협력멤버로 참여 중이다. DOI(Digital Object Identifier)가 부여된 80만 건 이상의 레코드를 확보하고 있으며, 과학데이터에 DOI를 부여하는 역할을 수행하고 있다. 2010년 8월 DataCite는 메타데이터 커널(버전 1.1)을 발표하였다. 이것의 주요한 기능은 크게 4가지로 요약할 수 있다. 첫째, DOI 등록을 위해 필요한 최소한의 요소를 이용해 데이터세트를 인용하는 표준을 제시하는 기능, 둘째, 데이터 관리를 위한 다양한 스키마들과 상호 운용성 확보 기능, 셋째, 다른 자원과의 관계를 기술하는 것과 같이 데이터에 대한 융통성 있는 기술이 가능하도록 필수적인 요소를 제공함으로써 데이터세트의 발견이 더욱 용이하도록 하는 기능, 넷째, 용어의

통제어 리스트 사용이나 표준 어휘집 사용을 통해 미래 서비스를 준비하는 기능으로 구분된다(Joan Starr, Angela Gastl 2011).

DataCite 메타데이터 스킴의 핵심 속성은 데이터세트의 인용을 위해 준비된 5개의 요소로서 <Identifier>, <Creator>, <Title>, <Publisher>, <PublicationYear>이다. 특이한 점은 데이터세트 인용을 위한 <Identifier> 이외에도 DOI를 대체할 수 있는 <AlternatedIdentifier> 속성도 제공한다. 또한 관련된 자원과의 연계를 위해 <RelatedIdentifier>를 제공하고 있다. Joan Starr, Angela Gastl(2011)은 데이터세트에 DOI를 부여하여 DataCite에 메타데이터를 기탁하기 전 독자적인 시스템에서 관리하던 식별자를 <AlternatedIdentifier>에 기술함으로써 데이터의 수명주기를 지원할 수 있다고 주장하였다.

3.2 DC 데이터세트 분석

더블린 코어(Dublin Core, 이하 DC)는 ISO 15836으로 표준화된 메타 데이터요소 집합이다. DC의 가장 큰 장점인 단순성, 확장성, 구문 독립성과 같은 특징 때문에 동영상, 소리, 이미지, 텍스트, 웹 페이지 등의 디지털 매체들을 기술하는 데 널리 사용되고 있다. 많은 분야에서 메타데이터의 상호 운용성 확보를 위해 연계테이블(Crosswalks)을 작성할 때 그 기준으로 DC를 사용하는 것이 그 이유이다. DC의 목표는 데이터의 형식과 구조를 단순화하

〈표 2〉 DataCite 스킴에 누락된 DC의 주요 요소

DC	설명	DIF
coverage	대체로 공간적 위치(지명이나 지리적 좌표), 시간적 시기(시대명, 날짜, 혹은 기간) 혹은 관할지역(행정구역의 이름) 등이 포함	Temporal_Coverage / Spatial_Coverage / Location
audience	자원이 의도되고 있거나 또는 이 자원이 그들에게 유용하리라고 생각되는 등급 또는 수준	×

여 원문의 저자나 발행자가 메타데이터를 직접 작성하고, 이를 바탕으로 특정분야에서 요구되는 수준으로 확장하여 사용할 수 있도록 하는 것이다. 단순하고 확장 가능한 것이 큰 장점으로 인정되는 반면 디지털도서관, 교육매체 관리 등과 같은 응용 영역에는 데이터 요소가 부족하다는 비판이 있지만 단순한 구조의 시스템 개발, 교환 및 통합용 메타데이터, 응용 프로파일 개발의 기본 메타데이터로 활용되고 있어 유용성이 매우 많다. DC 표준에는 두 가지 단계가 들어 있다. 하나는 단순 더블린 코어로 15개의 요소로 이루어져 있고, 다른 하나는 한정적 더블린 코어로 기본 더블린 코어에 3개의 요소를 추가하고 요소가 적용되는 범위를 제한하기 위한 한정자(qualifier)를 추가한 것이다.

15개 기본요소 중 <coverage>를 제외한 14개 요소가 DataCite 요소와 매핑이 되었다. 이 분석 과정에서 주목할 만한 사항들이 아래와 같이 발견되었다.

- DataCite 요소 <PublicationYear>의 의미는 데이터가 대외적으로 공표되어 활용

가능해진 날짜(년도)를 의미하기 때문에 DC의 대응 요소로 <date> 요소보다는 <available> 요소가 더욱 적합함.

- DataCite 요소와 기본요소 15개 중 14개와 추가요소 1개로 매핑이 가능함.
- coverage 요소는 관측분야의 데이터 요소에서 권고되는 항목으로 DataCite 요소로 추가가 필요함.

〈표 2〉는 DataCite 스킴에 누락된 DC의 주요 요소와 그에 대응되는 후술할 DIF 요소이다. DC의 <coverage>는 공간적 위치와 시간적 시기, 관할지역 등을 포함할 수 있다. <audience>는 이용자에게 적합한 관련 콘텐츠를 추천하는 서비스 설계 시 필요한 항목으로 DataCite 메타데이터 요소에 포함되면 그 활용도가 매우 높을 것이다.

3.3 OECD 데이터세트 분석

Toby(2009)는 데이터세트나 테이블을 출판, 표현, 인용하는 데 있어 표준적인 방법이 없음을 지적하였다. 심지어 OECD에서 만들

어떤 같은 소스를 전혀 다른 방식으로 인용하고 있는 사례들을 들며 이를 위한 표준화 작업이 필요함을 강조하였다. 이러한 현상에 대한 이유로 Toby(2009)는 저자와 출판사가 데이터소스를 어떻게 인용해야 하는지 모른다는 점을 들었다. 이러한 이유로 OECD 메타데이터 요소가 개발되었다. 인용을 위한 주요 메타데이터 항목을 기술한 것으로서 DataCite나 DC 데이터세트 보다 더욱 단순하게 구성되어 있다.

데이터세트와 테이블을 출판하기 위한 OECD의 메타데이터 요소는 Altman과 King(2007)의 제안 요소 중 <UNF(Universal numeric fingerprints)>만 제외한 모든 필드를 포함하고 있다. Toby(2009)는 <UNF> 또한 <Supersedes>와 <is continued by> 요소로 대체 기술이 가능하다고 하였다. OECD 메타데이터 요소는 <Identifier> 필드와 같은 11개의 필수 필드와 <Subtitle>과 같은 15개의 선택필드로 구성되어 있다. Toby(2009)는 세부적으로 dataset, collection of datasets, key table collection, key tables, Key table edition으로 나누어 메타데이터 요소를 제안하였다. 각각의 메타데이터 요소들은 '출판용 메타데이터(Publishing metadata)' 항목과 타 정보원과의 '연결(Links)' 항목으로 구분하여 정의하였다. 본 연구에서는 DataCite 메타데이터 요소를 기준으로 각각의 메타데이터 요소 묶음에서 추출한 17개의 필드를 대상으로 연계를 하였다.

3.4 DIF 요소 분석

관측분야에서 표준으로 사용하고 있는 DIF(Directory Interchange Format) 포맷은 1987년 2월에 개최된 ESADS(Earth Science and Applications Data Systems Workshop)의 결과물로서 1988년 CI(Catalog Interoperability) 과학자문위원회에서 최종 승인을 받아 채택되었다. 1989년 지구관측위성 위원회 데이터 워킹그룹(CEOS, Committee on Earth Observation Satellites Data Working Group)에서 CI 워크숍에 적극적으로 참여하여 DIF구조에 많은 의견을 주었다. 이후 2004년 6월에 ISO 19115/TC211로써 표준으로 공표되었다. DIF 포맷은 Entry_ID 등 8개의 필수 요소와 18개의 적극 추천요소, 10개의 추천 요소로 구성되었다. 본 연구는 관측분야의 과학데이터를 위한 메타태그를 제안하는 것이 목적이기 때문에 관측분야의 메타데이터 기술 표준인 DIF에서 사용하는 요소가 반드시 고려되어야 한다. 아래 <표 3>은 DIF에서 사용하는 요소 중 DataCite, DC, OECD에 포함되어 있지 않으면서, 과학데이터용 OpenURL 메타태그에 추가시킬 요소들을 정리한 것이다. 분석에 활용된 메타데이터 요소들의 샘플 데이터는 <http://bit.ly/iWiUkV>에서 확인 가능하다.

〈표 3〉 DataCite 스키마에 누락된 DIF의 주요 요소

DataCite	DC	OECD	DIF	설명
×	×	×	Metadata_Name	메타데이터를 기술한 메타데이터 스키마의 이름을 기술함.
×	×	×	Metadata_Version	메타데이터를 기술한 메타데이터 스키마의 버전을 기술함.
×	×	×	Data_Resolution	데이터의 해상도 정보를 기술하고 특정해상도 이상의 데이터 검색 제한에 활용할 수 있음.
×	×	×	Project	데이터 생성과 관련된 프로젝트명을 기술함. 데이터 검색에 중요한 요소로 활용가능 하며, 프로젝트 관련 데이터 일괄 검색이 가능함.
×	×	×	Sensor_Name	데이터 획득에 사용된 장비정보를 기술함(크게 3가지로 구분). 장비관련 데이터 일괄 검색이 가능함.
×	×	×	Source_Name	데이터를 획득하는데 사용된 플랫폼정보를 기술함(크게 11개로 구분됨). 플랫폼 데이터 일괄 검색이 가능함.
×	×	×	Quality	데이터 품질 정보 기술. 현재는 자유기술 형식으로 되어 있지만, 데이터 품질이 과학데이터 유통에 새롭게 부각되고 있어 추가할 필요가 있음.

4. OpenURL 메타태그 분석

4.1 Z39.88

4.1.1 KEV 메타데이터 형식

연구데이터를 위한 메타데이터 형식 개발이 필요하다. 2011년 2월 현재 OpenURL Framework Registry에 등록된 메타데이터 형식은 〈표 4〉와 같이 단행본, 학위논문, 저널, 특허, 서비스 타입이 전부이다. 연구데이터를 기술할 수 있는 범용의 메타데이터 포맷을 개발하여 레지스트리에 등록할 필요가 있다. 문헌위주의 학술정보를 기술하기 위한 4가지 포맷이 존재하지만 다양한 유형의 과학데이터를 기술하기 위해서 레지스트리에 등록된 메타데이터 포맷은 없다. 따라서 다양한 유형의 과학데이터를

ContextObject에 기술하기 위해 두 가지 방법을 사용하고 있다. 그 중 한 가지 방법은 단행본용 메타데이터 포맷을 사용한다고 지정한 후 장르(genre) 키 값을 'unknown'으로 지정하거나, 전송하지 않고 제목과 저자 정보만으로 ContextObject를 생성하는 것이다. 또 다른 방법으로 그 단순성과 명료성을 인정받은 DC 메타데이터 포맷을 사용하는 것이다. 이 방법은 DC의 〈title〉, 〈creator〉, 〈type〉, 〈format〉을 사용할 수 있기 때문에 첫 번째 방법보다 더욱 상세한 기술이 가능하다. 일정한 표준이 없이 이러한 기술방법을 제시하는 것은 연구데이터를 위한 메타데이터 형식이 개발되어 레지스트리에서 관리되고 있지 않기 때문이다.

〈표 4〉 Key/Encoded-Value 메타데이터 포맷

(NISO AX Committee 2007, 10)

메타데이터 포맷	레지스트리 식별자
단행본	info:ofi/fmt:kev:mtx:book
학위논문	info:ofi/fmt:kev:mtx:dissertation
저널	info:ofi/fmt:kev:mtx:journal
특허	info:ofi/fmt:kev:mtx:patent
학술적 서비스 유형	info:ofi/fmt:kev:mtx:sch_svc

〈표 5〉 학술적 서비스 유형을 위한 key 구성

(NISO AX Committee 2007, 15)

Key	내용
초록(abstract)	Referent의 초록정보
인용(citation)	Referent의 서지 인용정보
원문(full-text)	Referent의 원문정보
소장정보(holdings)	Referent 관련 소장정보
도서관상호대차(ill)	Referent에 대한 도서관 상호대차
기타(any)	다른 유형의 서비스

4.1.2 학술서비스 유형 형식 분석

다음으로 학술서비스 유형 형식의 확장이 필요하다. 현재의 학술적 서비스 유형은 〈표 5〉와 같이 총 6가지의 키로 구분되어 있다. 〈abstract〉, 〈citation〉, 〈fulltext〉, 〈holdings〉, 〈ill〉, 〈any〉이다. 과학데이터는 생성된 과학데이터 자체가 연구의 결과물인 경우도 있지만 대부분 연구의 내용을 정리한 보고서, 논문 등에 더욱 자세히 기술되어 있다. 따라서 논문과 관련된 데이터에 대한 이용자의 수요가 반드시 존재하며, 이러한 이용자의 요구가 서비스 서버에 전송될 수 있어야 한다. 하지만 현재의 여섯 가지 키로는 이를 만족시킬 수 없다. ‘any’ 키를 활용할 수도 있지만 ContextObject에 정 확한 요구를 담기에는 한계가 있다.

4.2 OpenURL 0.1 분석

4.2.1 글로벌-식별자-영역

(GLOBAL-IDENTIFIER-ZONE)

OpenURL에서 글로벌 식별자로 사용되는 메타태그는 id이다. 이에 대한 메타 값은 글로벌 네임스페이스와 글로벌 식별자의 조합이다. 현재 네임스페이스로 사용할 수 있는 것은 DOI (digital object identifier), pmid(PubMed identifier), bibcode(identifier used in Astrophysics Data System), oai(identifier used in the Open Archives initiative)이다(Herbert Van de Sompel 외 2000). 개별 과학데이터를 손쉽게 식별하기 위해서 ID는 필수적으로 사용되어야 한다. 과학데이터의 글로벌 유통을

위해서 DataCite는 과학데이터에 DOI를 부여하고 있기 때문에 과학데이터용 메타태그로서 id는 사용될 수 있으며, DOI 네임스페이스를 그대로 활용할 수 있다.

4.2.2 객체-메타데이터-영역

(OBJECT-METADATA-ZONE)

OpenURL 메타태그 중 genre 태그의 값을 보면 개별 아이템의 묶음으로써 journal, book, conference를 제시하며, 개별 아이템으로서 article, preprint, proceeding, bookitem을 제시하고 있다. 이것들은 문헌정보만을 위한 메타태그의 값들로서 본 논문에서 다루고 있는 과학데이터 용으로는 부적합하다.

메타태그 중 저자 관련 태그는 <aulast>, <aufirst>, <auinit>, <auinit1>, <auinitm> 이렇게 5가지를 제공하고 있다. 이것은 제 1저자의 성과 이름을 메타 값으로 전달하기 위해 사용한다. 이것은 2가지 측면에서 문제점이 존재한다. 첫째, 서비스 컴포넌트 측면에서의 문제이다. 문헌정보와 과학데이터 정보를 웹을 통해 제공하는 서비스 중 검색제한으로 제 1저자를 지정하는 사례가 극히 드물다. 즉 검색을 위한 색인 작업을 할 때 콘텐츠 저작자에 대한 제 1저자만 별도로 색인을 잡지 않는 것을 의미한다. 둘째, 서비스 컴포넌트에게 콘텐츠 관련 정보를 전달하게 되는 정보시스템 측면에서의 문

제이다. 제 1저자의 정보를 전달하기 위해서는 별도로 관련 정보를 관리해야 하는 부가적인 프로세스가 발생하게 된다. 이는 해당 메타태그의 활용을 저해하는 요인이 되고 있다.

메타태그 중 식별자 정보를 제공하기 위한 태그로는 <issn>, <eissn>, <coden>, <isbn>, <sici>, <bici>가 있다. 단행본의 일부 장에 부여되는 <bici>를 제외하곤 모두 아이템의 묶음인 문헌정보를 식별하기 위한 식별자들이다. 모두 특정한 목적을 갖고 식별자의 표준으로 사용되고 있기 때문에 과학데이터를 식별하기 위한 식별자로서 사용할 수 없다.

메타 값의 조합으로 개별 아이템을 식별할 수 있는 태그로서 <volume>, <part>, <issue>, <spage>, <epage>, <pages>, <artnum>, <date>, <ssn>, <quarter>를 제공하고 있다. 아이템의 묶음인 문헌정보를 식별하기 위한 식별자와 마찬가지로 대부분은 메타태그가 문헌정보를 식별하기 위해 쿼리 조합에 필요한 것이다.

5. ORD(Observation Research Data) 메타태그⁶⁾ 개발

5.1 메타태그 설계원칙

관련 분야의 표준을 최대한 반영하고자 노

6) 관측분야 과학데이터의 유통을 위한 OpenURL 핵심 메타태그를 본 논문에서는 ORD(Observation Research Data) 메타태그라 한다.

력하였다. 특히 메타데이터 요소 비교의 기준으로 사용된 DataCite 메타데이터 요소를 거의 모두 수용하고자 하였다. 해당 분야의 전문가들이 연구현장의 의견을 반영하여 지속적인 갱신을 해오고 있는 DataCite, DC, OECD, DIF 메타데이터 요소 하나하나에 대해서 그 용도와 가치를 판단하는 것은 불필요하다고 전제하였다. 다만 본 연구에서는 OpenURL 소스와 서비스 구성 요소간 상호운용성 확보를 할 수 있도록 OpenURL 핵심 메타태그를 KEV 형식으로 제안하는 것이기 때문에 OpenURL 메타태그로 적합한 메타데이터 요소인지만 판단하였다. 다음은 ORD KEV 메타데이터 정의를 위해 준수한 설계방향이다.

- 데이터 인용을 위해 제시된 메타데이터 요소는 해당 분야 전문가들에 의해서 각 고의 연구 노력으로 도출된 요소라고 전제(DataCite, OECD, DIF)
- 과학데이터 기술을 위한 최근 연구 동향을 반영하여 메타데이터 요소를 설계(DataCite)
- 관측분야 표준메타데이터의 필수요소와 적극 추천하는 요소는 해당분야의 전문가들에 의해서 오랜 시간 동안 검토되고 수정되어 제시된 메타데이터 요소라고 전제(DIF)
- 웹상의 전자자원을 기술하는 데 가장 많이 활용되고 있는 DC의 핵심메타데이터는 다른 정보자원 서비스와 연계를 위해

서라도 핵심요소 그대로 인정(DC)

- Z39.88 표준을 준수하여 KEV 메타데이터 제한언어는 “The Z39.88-2004 Matrix Serialization(식별자: info:ofi/fmt:kev:mtx)” 형식을 따름(Z39.88)
- Z39.88 표준을 준수하여 메타태그명을 수정함(Z39.88)

상기의 기준으로 DataCite 메타데이터 요소를 기준으로 DC, OECD, DIF 각각의 메타데이터 요소를 분석하여 연결테이블(Crosswalks)을 작성하였다. 다음으로 DataCite 메타데이터 요소를 대상으로 OpenURL 메타태그로 적합한 메타데이터 요소인지 판단하여 1차 핵심 메타태그를 도출하였다. DataCite 메타데이터 요소 중 <Right>와 <Description>은 OpenURL 소스와 서비스 구성 요소간 상호운용성 확보를 위해 불필요한 요소로 판단하여 선정하지 않았다. DataCite 메타데이터 요소를 기준으로 핵심 메타태그를 선정 한 후 마지막으로 검토 대상이 되었던 DC와 OECD, DIF에서 관측분야의 핵심 메타태그로 추가될 수 있는 요소를 도출하였다. 다음 <표 6>은 DataCite 메타데이터 요소와 DC, OECD, DIF 각각의 메타데이터 요소를 비교한 연결테이블(Crosswalks)이다.

〈표 6〉 DataCite, DC, OECD, DIF 스키마 연결테이블(Crosswalks)

ID	DataCite	DC	OECD	DIF
1	Identifier 초기에는 DOI ⁷⁾	identifier	Identifier	Entry_ID는 로컬식별로 사용가능함
2	Creator	creator	Author(s)	Data_Set_Citation(Dataset_Creator) / Personnel(First_Name,Middle_Name,Last_Name,)
3	Title	title	Main Title Subtitle	Entry_Title
4	Publisher	Publisher	×	Data_Set_Citation(Dataset_Publisher)
5	PublicationYear 초기에는 PublicationDate	available	Publication date	Data_Set_Citation(Dataset_Release_Date)
6	Subject 초기에는 Discipline	subject	JEL Classification Theme(s)	Parameters / ISO_Topic_Category / Keyword
7	Contributor	contributor	Is Edited By	Data_Center
8	Date	date	Period:Start Year Period:End Year	Temporal_Coverage(Start_Date,Stop_Date)
9	Language	language	Languages	Data_Set_Language
10	ResourceType	type	×	Distribution(Distribution_Media)
11	AlternateIdentifier 초기 검토 시기에는 없었음	identifier	Identifier	×
12	RelatedIdentifier	relation / source	Belongs to Has Main Parent Supersedes Is continued By	Related_URL(URL_Content_Type(Type),URL,Description)
13	Size	extent	Size	Distribution(Distribution_Size)
14	Format	format	×	Data_Set_Citation(Data_Presentation_Form) / Distribution(Distribution_Format)
15	Version	relation	×	Data_Set_Citation(Version) / Data_Set_Progress
16	Rights	rights	×	Access_Constraints / Use_Constraints / Private
17	Description	description	Keyword(s)	Summary(Abstract, Purpose)

7) 2010년 DataCite 메타데이터 스키마 공식적으로 발표 되었을 때 핵심 DOI가 핵심 속성이었으나 DataCite 메타 데이터 워킹 그룹에 참여하고 있는 Joan Starr, Angela Gast(2011)의 연구에서는 Identifier로 변경되었음을 알 수 있다.

5.2 핵심 메타태그

핵심 메타태그를 도출하기 위하여 DataCite 메타데이터 커널을 기준으로 DC 메타데이터 요소, OECD 메타데이터 요소와 DIF 메타데이터 요소들을 비교하여 공통요소를 도출하였다. 이후 비교 대상으로 검토되었던 스키마에서 과학데이터 관련 핵심 메타태그를 도출하여 추가하였다.

현재의 OpenURL 레지스트리에 등록된 메타데이터 형식 중 학술 커뮤니티에서 자주 사용될 수 있는 서비스 유형에 대한 형식도 존재한다. 하지만 본 연구는 OpenURL소스와 서비스 구성요소의 상호운용성 확보를 위한 핵심 메타태그를 개발하는 것이 주된 목적이다. 따라서 데이터 관련 서비스 유형 확장에 대한 연구는 별도로 진행하고자 한다.

핵심 메타태그를 도출할 때 검토 스키마에 모두 존재하는 것은 해당 요소가 데이터를 기술하는 데 기본적인 기술항목임을 의미하는 것이며, 또한 학술 커뮤니케이션에서 기본 항

목임을 의미하기 때문에 핵심메타태그로 선정하였다. 이런 기준으로 추가된 요소는 <Identifier>, <Creator>, <Title>, <Publication-Year>, <Subject>, <Contributor>, <Date>, <Language>, <RelatedIdentifier>, <Size>이다. 다만 DataCite의 <Description> 요소만 제외를 하였다. 다음 <표 7>은 검토 스키마 중 하나의 스키마에서도 누락된 태그들이 선정된 이유를 설명하고 있다.

정의된 ORD KEV 메타데이터 형식은 <Summary> 필드를 제외하고 DIF 필수요소를 모두 포함하도록 설계되었다. <Entry_ID>는 Identifier, <Entry_Title>은 <Title>, <Parameters>와 <ISO_Topic_Category>는 <Subject>, <Data_Center>는 <Contributor>, <Metadata_Name>과 <Metadata_Version>은 <MetadataName>과 <MetadataVersion>으로 매핑 되었다. 다만 <Summary>는 OpenURL 소스와 서비스 구성 요소간 상호운용성 확보를 위해 불필요한 요소로서 선정하지 않았다.

<표 7> ORD 메타데이터 포맷의 Key 선정 기준

출현 스키마	요소	선정 이유
DataCite, DC, DIF	Publisher	OECD 스키마에서는 지원하지 않으나, 학술 커뮤니티에서 주요한 요소임으로 추가하였음.
DataCite, DC, DIF	ResourceType	OECD 스키마에서는 지원하지 않으나, 학술 커뮤니티에서 주요한 요소임으로 추가하였음. DataCite 스키마는 통제어 리스트를 제공
DataCite, DC, OECD	AlternateIdentifier	DIF 스키마에서는 지원하지 않으나, 학술 커뮤니티에서 주요한 요소임으로 추가하였음. 연구 분야별로 존재할 수 있는 식별체계를 활용할 수 있음.
DataCite, DC, DIF	Format	동일한 데이터에 대해서 표현된 형태가 다양할 수 있으며, 연구자는 특정 형태의 데이터를 원할 수 있음.

출현 스키마	요소	선정 이유
DataCite, DC, DIF	Version	다양한 버전의 데이터가 존재할 수 있으며, 이용자는 특정 버전의 데이터 및 그것과 관련된 정보를 원할 수 있음.
DC, DIF	coverage	공간적 위치(지명이나 지리적 좌표), 시간적 시기(시대명, 날짜, 혹은 기간) 혹은 관찰지역(행정구역의 이름) 등이 포함. 관측용 데이터 이외에도 활용 가치가 높음. 예를 들어 사회과학 분야의 설문조사 범위(연령, 지역 등)에 사용될 수 있음.
DC	audience	현재는 DC 스키마만 기술되어 있지만, DLESE ⁸⁾ 와 같이 교육 커뮤니티에서 데이터가 활용될 경우를 대비해서 선정함.
DIF	Metadata_Name	관측분야의 메타데이터 포맷이기 때문에 DIF 필수요소 선정. OpenURL을 전송 받을 서비스에서 검색 제한자로 사용 가능.
DIF	Metadata_Version	관측분야의 메타데이터 포맷이기 때문에 DIF 필수요소 선정. OpenURL을 전송 받을 서비스에서 검색 제한자로 사용 가능.
DIF	Data_Resolution	데이터 검색에 중요한 요소로 활용가능. 특히, 프로젝트 관련 데이터 일괄 검색이 가능 및 검색 제한자로 활용 가능.
DIF	Project	데이터 검색에 중요한 요소로 활용가능. 특히, 프로젝트 관련 데이터 일괄 검색이 가능 및 검색 제한자로 활용 가능.
DIF	Sensor_Name	데이터 검색에 중요한 요소로 활용가능. 특히, 장비관련 데이터 일괄 검색 가능 및 검색 제한자로 활용 가능.
DIF	Source_Name	데이터 검색에 중요한 요소로 활용가능. 특히, 플랫폼 데이터 일괄 검색 가능 및 검색 제한자로 활용 가능.
DIF	Quality	DIF 스키마에 자유기술 형식으로 되어 있지만, 데이터 품질이 과학데이터 유통에 새롭게 부각되고 있어 예비로 선정함.

5.3 KEV 형식 정의

ORD KEV 메타데이터 정의에 사용한 제한 언어(Constraint Language)로는 OpenURL 레지스트리에서 관리되고 있는 “The Z39.88-2004 Matrix Serialization(식별자: info:ofi/fmt:kev:mtx)” 형식을 사용하였다. Key 이름은 표준에 맞게 변경하였다. 모두 소문자로 변경하였으며, ‘_’ 문자를 삭제하였다. 또한 모

든 공백을 삭제하였다. ‘_’ 문자가 삭제된 기존 요소는 <Metadata_Name>, <Metadata_Version>, <Data_Resolution>, <Sensor_Name>, <Source_Name>이다. <표 8>은 “Z39.88-2004 Matrix Serialization” 언어에 맞게 작성된 ORD KEV 메타데이터 정의이다. 실제 정의되어 있는 모습은 <http://bit.ly/euht9q>에서 확인 가능하다. Min, Max 열의 값은 DataCite에서 설계한 메타데이터 기술 시 발생빈도를

8) DLESE는 미국 국립과학재단의 재정 지원을 받아 개발되었으며, 현재는 국가대기연구소에서 운영 중에 있다. 연구자가 연구수행 중 생성한 모든 데이터를 연구데이터로 정의할 때 DLESE에서 관리 및 유통되는 자료 중 과학데이터가 상당하다고 볼 수 있다. 또한 DLESE의 지구과학 교육 지원 내용 중 교육에 효율적으로 활용 될 수 있도록 여러 도구와 인터페이스를 포함한 데이터세트, 이미지 제공은 과학데이터 서비스의 한 사례로 전혀 손색이 없는 예라 할 수 있다.

〈표 8〉 ORD를 표현하기 위한 KEV 형식의 매트릭스 정의

Delim	Key	Equals	Value	Min	Max	Description
#						이하 DataCite 스킴에서 추출된 Key
&	identifier	=	<data>	0	1	데이터의 유일한 문자열 식별자
&	creator	=	<data>	0	n	데이터생산에 참여한 주요 연구자
&	title	=	<data>	0	n	데이터가 알려진 이름
&	publisher	=	<data>	0	1	데이터 소유기관 및 배포기관
&	publicationyear	=	<data>	0	1	데이터의 사용가능 시작년도
&	subject	=	<data>	0	n	데이터 설명(주제, 키워드, 분류코드 등)
&	contributor	=	<data>	0	n	데이터 수집이나 개발에 책임이 있던 기관이나 사람
&	date	=	<data>	0	n	데이터와 관련된 날짜
&	language	=	<data>	0	1	데이터의 언어
&	resourcetype	=	<data>	0	1	자원(데이터)의 유형
&	alternateidentifier	=	<data>	0	n	DOI 이외의 데이터 식별자
&	relatedidentifier	=	<data>	0	n	관련 자원의 식별자
&	size	=	<data>	0	1	데이터의 크기
&	format	=	<data>	0	1	데이터의 기술적인 형태(from MIME Media Types list)
&	version	=	<data>	0	1	데이터의 버전
#						이하 DC 스킴에서 추출된 Key
&	coverage	=	<data>	0	1	공간, 시간 혹은 관찰지역
&	audience	=	<data>	0	1	데이터 이용 대상
#						이하 DIF 스킴에서 추출된 Key
&	metadataname	=	<data>	0	1	메타데이터 이름
&	metadataaversion	=	<data>	0	1	메타데이터 버전
&	dataresolution	=	<data>	0	1	데이터 해상도
&	project	=	<data>	0	1	데이터 생성과 관련된 프로젝트 이름
&	sensorname	=	<data>	0	1	데이터 획득에 사용된 장비정보
&	sourcename	=	<data>	0	1	데이터를 획득을 위한 플랫폼정보
&	quality	=	<data>	0	1	데이터 품질 정보

최대한 수용하였으나, <size>와 <format> 요소는 Max 값을 '1'로 한정하였다. Referent나 ReferringEntity로 지정되는 특정 데이터는 하나의 크기(size)와 형식(format)을 갖기 때문이다.

정의된 ORD KEV 메타데이터 형식은 다음의 다섯 가지 사항을 고려한 것이다. 첫째, ANSI/NISO Z39.88-2004 표준을 준수하여

관련 커뮤니티 연구자들과 종사자들이 쉽게 이해할 수 있도록 하였다. 둘째, 메타태그의 최신성이다. 연구데이터의 출판에 앞장서고 있는 DataCite의 최근 표준을 기준 스키마로 사용하기 때문에 데이터를 기술하되 필요한 메타태그에 대한 최근 연구결과를 반영하였다. 셋째, OECD와 DataCite, DIF의 데이터 인용을 위한 필수 요소 항목들을 함께 검토하였다. 따라

서 OpenURL 소스에서 데이터 인용 관련 요소가 모두 활용될 수 있도록 설계하였다. 넷째, 관련분야 표준을 최대한 반영하였다. 관측분야에서 표준으로 사용되고 있는 DIF 형식의 요소 중 주요 요소가 추가되어 관측분야에서 활용하는데 있어 기존 데이터베이스에 구축된 정보가 활용될 수 있도록 하였다. 다섯째, 미래를 준비하는 예비성을 들 수 있다. <audience>와 <quality> 메타태그는 통제가 기준이 마련되어 있지 않아 지금 당장 효용성이 없더라도 향후에 연구데이터의 활용이 활성화되었을 때 필요한 태그라 판단하여 예비 메타태그로 선정되었다.

5.4 개발 메타태그 평가 방안

본 개발 결과의 검증과 구현은 앞으로 DataCite 기술 관련 워킹 그룹, 현재 OpenURL 레지스트리 운영을 맡고 있는 OCLC, 세계적 Linking Resolver 솔루션 개발사인 엑스리브리스, 한국과학기술정보연구원의 전자자원 호스팅 사업 커뮤니티, 그리고 관련 종사자들과의 협력 연구를 거쳐 이루어질 것으로 사려 된다. 구체적인 과정을 서술하면 다음과 같다.

첫째, DataCite 기술 관련 워킹 그룹에 본 연구의 결과를 제출하고 관련 연구자와 후속 연구 진행이 필요하다. 둘째, 지구 관측 분야에서 2011년 2월 현재 OpenURL 레지스트리 운영에 책임을 맡고 있는 OCLC(Online Computer Library Center, Inc.)에 개발된 내용을 제출하여 표준화 작업을 진행하는 것이다. 셋째,

Linking Resolver 솔루션을 개발하여 세계적으로 보급하고 있는 엑스리브리스사의 개발팀과 긴밀한 협력을 통해 ORD 개발 내용이 국제 사회 커뮤니티에서 활발하게 논의될 수 있도록 해야 한다. 넷째, 한국과학기술정보연구원에서 운영하는 전자자원 호스팅 사업의 커뮤니티를 활용하는 것이다. 2007년에 사업을 시작하여 2011년 2월 현재 국내 57개 기관이 호스팅 사업에 참여하고 있다. 참여기관의 관리자들을 위한 교육이 매년 정기적으로 진행되고 있어, 이러한 채널을 이용하여 ORD의 실험적 적용을 통해 문제점을 지속적으로 개선시킬 수 있을 것이다. 마지막으로, ORD KEV 메타데이터 형식에 대한 관련분야 종사자의 검토가 무엇보다도 필요하다.

6. 결론

문헌위주의 학술정보 유통에 초점이 맞추어져 개발된 OpenURL에 과학데이터 유통을 위한 메타태그를 접목하는 시도를 하였다. ORD 메타데이터의 객관성과 OpenURL에서의 상호운용성을 확보하기 위하여 DataCite 컨소시엄의 메타데이터 요소, DC의 핵심 메타데이터 요소, OECD 메타데이터 요소, 그리고 DIF 메타데이터 요소를 비교, 분석하였다.

DataCite 메타데이터 요소를 기준으로 상각 기관의 요소들과의 연결테이블(Crosswalks)을 작성하였다. 각각의 요소들을 비교한 후 공

통요소를 찾아내고 이후 비교 대상으로 검토되었던 스키마에서 과학데이터 관련 핵심 메타태그를 도출 정립하였다. 핵심 메타태그를 결정할 때의 가장 큰 원칙은 검토 스키마에 모두 존재하는 것을 우선적으로 선택하는 것이었다. 다음으로 학술 커뮤니케이션에서의 주요한 요소, 관측분야의 메타데이터 포맷 등이 있었다.

ORD KEV 메타데이터(<http://bit.ly/euht9q>)의 정의는 “Z39.88-2004 Matrix Serialization(식별자: info:ofi/fmt:kev:mtx)” 형식을 제한 언어로 사용하였다. Key 이름은 표준에 맞게 변경하였으며 메타데이터 형식은 첫째, ANSI/NISO Z39.88-2004 표준 준수, 둘째, 메타태그의 최신성 보전, 셋째, OECD와 DataCite, DIF의 필수 요소 항목 참고, 넷째, 관련분야 표준의 적극적 반영, 다섯째, 미래 대비와 같은 다섯 가지 사항을 고려하였다.

본 개발 연구의 평가는 후속되는 연구에서 유관 기관과 관련 종사자들을 연계하여 평가할 것이다. 후속 연구로는 본 연구 결과의 활용과 확산에 있어 다양한 형식의 제한 언어를 통한 연구결과 공유를 위해 info:ofi/fmt:kev:xml 형식의 제한 언어 확장이 필요하다. 또한 현재의 학술적 서비스 유형이 총 6가지의 키로 구분되어 있는데, 모두 문헌을 위주로 한 서비스 유형들만 포함되어 있기 때문에 이에 대한 연구가 필요하다. 이를 위해서는 Scholarly Service Type Keys의 확장 연구가 필요하다. 왜냐하면 과학데이터에 대한 이용자의 다양한 서비

스 수요가 존재하며, 이러한 이용자의 요구가 서비스 서버에 전송될 수 있도록 서비스 타입 확장이 필요하기 때문이다.

참고문헌

- 곽승진. 2004. 청소년 과학정보 메타데이터 요소 및 데이터베이스 구축 연구. 『한국문헌정보학회지』, 38(1): 263-279.
- 김선태, 한선화, 이태영, 김용. 2010. 과학데이터 보존 및 활용모델에 관한 연구. 『한국비블리아학회지』, 21(4): 81-93.
- 심경. 2003. 메타데이터 통합 방안. 『한국도서관·정보학회지』, 34(3): 169-192.
- 이혜영, 곽승진. 2007. 지구 관측자료 공유를 위한 메타데이터 연구. 『한국문헌정보학회지』, 41(2): 257-276.
- 황종엽. 2003. 해양과학기술 R&D 결과정보 데이터베이스 구축 연구. 『한국도서관·정보학회지』, 34(2): 209-231.
- Herbert Van de Sompel, Patrick Hochstenbach, Oren Beit-Arie. 2000. OpenURL Syntax Description, version OpenURL/1.0f.
- Joan Starr, Angela Gastl. 2011. “isCitedBy: A Metadata Scheme for DataCite.” *D-Lib Magazine*, 17(1/2).
- Joan Starr et al. 2010. *DataCite Metadata Kernel for the Publication and Citation of Research Data*. DataCite Metadata

- Working Group.
- Lois Mai Chan, Marcia Lei Zeng. 2006. "Metadata Interoperability and Standardization-A Study of Methodology Part I Achieving Interoperability at the Schema Level." *D-Lib Magazine*, 12(6).
- Micah Altman, Gary King. 2007. "A Proposed Standard for the Scholarly Citation of Quantitative Data." *D-Lib Magazine*, 13(3/4).
- Marcia Lei Zeng, Lois Mai Chan. 2004. "Trends and Issues in Establishing Interoperability Among Knowledge Organization Systems." *Journal of the American Society for Information Science and Technology*, 55(5): 377-395.
- NISO AX Committee. 2004. Z39.88-2004: The OpenURL Framework for Context-Sensitive Services-The Key/Encoded-Value(KEV) Format - Implementation Guidelines. v1.7.
- OCLC. 2005. *Perceptions of Libraries and Information Resources*. Dublin, Ohio: OCLC.
- OCLC. 2006. *Process for the Evaluation and Approval of New OpenURL Community Profile Submittals by the Z39.88 Open-URL Maintenance Agency*. Dublin, Ohio: OCLC.
- Priscilla Caplan. 2003. *Metadata Fundamentals for All Librarians*. Chicago: American Library Association.
- Science Staff. 2011. "Challenges and Opportunities." *Science*, 331(6018): 692-693.
- Toby Green. 2009. "We Need Publishing Standards for Datasets and Data Tables." *OCCD Publishing White Paper*.
- Tom Kuipers, Jeffrey van der Hoeven. 2009. PARSE.insight: Insight into digital preservation of research output in Europe.
- UKRDS. 2010. Proposal and Business Plan for the Initial Pathfinder Development Phase.
- about ALPSP. [cited 2011. 1. 20]. <<http://bit.ly/eB7yGE>>.
- What is DataCite. [cited 2011. 1. 20]. <<http://bit.ly/hqfzMJ>>.
- What is a DIF? [cited 2011. 1. 20]. <<http://bit.ly/eb7AeS>>.
- 더블린 코어(Dublin Core, DC). [cited 2011. 1. 20]. <<http://bit.ly/fn0dCH>>.
- What is the maximum length of a URL? [cited 2011. 1. 20]. <<http://bit.ly/gTEAFZ>>.