정보관리연구, vol.42, no.4 2011, pp.39-58
http://dx.doi.org/10.1633/JIM.2011.42.4.039

연구데이터와 관련된 OpenURL의 학술서비스 유형 메타태그의 확장에 대한 연구

A Study on the Expansion of Meta-Tag for Research Data in Scholarly Service Type of OpenURL

김 선 태* · 이 태 영**
Sun-Tae Kim · Tae-Young Lee

차 례

1. 서 론
2. 연구데이터 출판
3. 학술 서비스 유형 도출
4. KEV 메타데이터 형식
5. KEVFormat: Sch-Svc 확장
6. 결 론
· 참고문헌

초 록

최신의 학술서비스 유형과 연구데이터 출판 및 서비스를 위한 DataCite 메타데이터 요소를 분석하여 OpenURL 학술서비스 유형을 기술하기 위한 메타태그를 Key/Encoded-Value(KEV) 형식으로 확장 제안하였다. 현재의 OpenURL Z39.88 표준의 KEVFormat: Sch-Svc은 문헌중심의 학술서비스 유형 6가지만을 지원하고 있어 연구데이터의 원활한 유통을 위해서는 확장이 필요하다. 본 연구에서는 학술서비스 유형 분석을 위해 Scopus와 Web of Science, NDSL 서비스를 비교 검토하여 8개의 학술서비스 유형을 도출하였다. 또한 연구데이터 기술을 위한 DataCite 컨소시엄의 〈RelatedIdentifier〉 메타데이터 요소를 집중적으로 분석하여 9개의 대표속성을 도출하였다.

키 워 드

연구데이터, 과학데이터, 메타태그, OpenURL, 학술서비스유형

* 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 Univ., taehyun@jbnu.ac.kr)
• 논문접수일자: 2011년 4월 25일
• 최종심사(수정)일자: 2011년 8월 20일
• 게재확정일자: 2011년 8월 28일

ABSTRACT

This paper presents a meta-tag expanded from scholarly service types of OpenURL written in Key/Encoded-Value format, after analyzing new scholarly service types and DataCite meta-data elements which are for research data publishing and services. So far, OpenURL Z39.88 standard, KEVFormat: Sch-Svc, supporting six scholarly service type only, the expansion of this standard is needed for a research data circulation. New eight scholarly service types were extracted, after analyzing and comparing with the Scopus, Web of Science, and NDSL services. And nine representative attributes were extracted, after analyzing intensively the DataCite's 〈RelatedIdentifier〉 elements.

KEYWORDS

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

1. 서 론

니콜라스 네그로폰테(1974) MIT대학 교수는 유비쿼터스적인 분산된 형태의 컴퓨터를 보게 될 것이며, 아마 컴퓨터라는 것이 장난감, 아이스박스, 자전거 등 가정 내 모든 물건과 공간에 존재하게 될 것이라 예견하였다. 2011년 3월 현재 마크 와이저(1988)의 세 가지 철학 중 '보이지 않는 컴퓨팅'이 실현되고 있으며, 세상의 모든 것이 도처에서 데이터를 생산하여 유통시키고 있다. 고성능 컴퓨터와 네트워크, 자동화 센서, 고성능 관측 및 실험 장치는 대량의 연구데이터를 쏟아내고 있기 때문에 연구의 패러다임도 변해야 할 시점에 와 있다.

전통적으로 과학자들은 연구를 함에 있어, 먼저 가설을 수립하고 수립된 가설을 증명하기 위해 실행 방법론을 설정하고 그 조건에 맞는 데이터를 수집하여 분석 처리한 후 결과와 결론을 도출하였다. 즉 연구 양상이 ① 가설, ② 수행 방법론, ③ 데이터 수집, ④ 분석 가공 처리, ⑤ 결과, ⑥ 결론의 순서로 진행되었다. 이 과정 중에서 Green(2009)은 데이터를 수집하는 시간이 연구의 대부분을 차지한다고 주장하였으며 학위논문이나 저널 논문, 단행본도 데이터 없이는 쓰여 질 수 없음을 강조하였다. 그러므로 현대와 같은 초스피드 시대에서는 데이터 수집 및 작성에 투여되는 시간을 여하히 줄이느냐가 관건으로 등장하며, 수많이 생산되어지는 데이터를 사장시키지 않고 효과적으로 이용하는 방법이 모색되어야 한다.

그 돌파구의 하나로 같거나 비슷한 연구 주제에서는 데이터를 공유하는 개념을 도입하면 데이터 수집 기간을 없애거나 줄일 수 있어 그만큼의 시간이 실제 연구에 투자됨으로 보다

우수한 결과물을 기대할 수 있다. 말 그대로 데이터 중심의 연구(data-intensive research)에서는 공유로 인해 풍부해진 데이터에 새롭고 다양한 가설을 실험해 볼 수 있다.

데이터 공유 개념은 학술 커뮤니케이션에 있어서도 많은 변화를 요구한다. 문헌중심의 유통에서 연구데이터 등 다양한 자원으로의 연계를 통해 연구자가 논문을 검색한 이후 자연스럽게 관련 데이터를 발견할 수 있는 서비스의 개발을 요구한다. 이것은 도서관이나 정보센터가 연구자들의 새로운 연구 패러다임을 지원하기 위하여 기존의 학술논문이나 특허, 연구보고서 중심의 서비스에 데이터 서비스를 추가 확대할 필요가 있음을 의미하기도 한다. 즉, 문헌위주의 과학기술정보 서비스 체계를 구성하는 각각의 구성요소가 연구데이터도 처리할 수 있게 확장되어야 함을 의미한다. 그러기 위해서는 여러 가지 커뮤니케이션 상황을 분석하여 서비스의 핵심역할을 수행하는 링킹서버의 역할이 매우 중요하다. 또한 링킹서버와 목표서비스를 제공하는 시스템 간 상호운용성 확보를 위해 시스템 간 교환하는 메타태그의 확장도 수반되어야 한다.

본 연구에서는 2004년 NISO AX위원회에서 표준으로 제정한 Z39.88-2004의 학술 커뮤니케이션 유형(scholarly service type)을 검토하였다. OpenURL 레지스트리에 등록된 KEVFormat:Sch-Svc 표준을 기준으로 하고, 연구데이터 기술을 위한 DataCite 컨소시엄의 메타데이터 요소 중 자원연계를 위한 〈RelatedIdentifier〉의 'type' 속성과 Scopus, Web of Science(이하 WoS), National Discovery for Science Leaders(이하 NDSL) 데이터베이스의 서비스 유형을 비교・분석하여 OpenURL KEVFormat:Sch-Svc 표준의 확장을 Key/Encoded-Value(KEV) 형식으로 제안하였다. 제안되는 연구결과의 활용은 DataCite 기술 관련 워킹 그룹과 현재 OpenURL 레지스트리 운영을 맡고 있는 OCLC, 세계적 링킹 리졸버(Linking Resolver) 솔루션 개발사인 엑스리브리스, 한국과학기술정보연구원의 전자자원 호스팅 사업 커뮤니티, 그리고 관련 전문가들과의 협력을 통한 후속 연구를 필요로 한다.

2. 연구데이터 출판

2.1 학술정보 출판유형의 변화

2010 DataCite Summer Meeting에서 Smit(2010)는 과학기술의료분야의 해외 유수 출판사들이 인쇄기반 출판에서 디지털 기반 출판으로 성공적인 전환을 해오고 있다고 발표 하였다. 최근 국제 STM 출판사협회의 조사(2009)에 따르면 이러한 성공적인 전환은 기술이 주도적인 역할을 하는 것으로 분석된다. 학술 커뮤니케이션에 있어서 연구수행과 커뮤니케이션의 심오한 변화와 저널출판에 기술이 큰 영향을 끼치고 있는 것이다. 이러한 기술의

힘을 입어 전 세계 출판시장을 선도하는 해외 학술출판사들이 새로운 플랫폼을 발표하고 있다. 새로운 서비스 플랫폼을 발표한 곳은 스프링거(Springer), 엘스비어(Elsevier), 와일리(Wiley), 프로퀘스트(Proquest) 등이다.

스프링거는 2010년 8월 24일 보도 자료에서 스프링거링크(SpringerLink) 리뉴얼을 발표했다. 리뉴얼에서 나타나는 가장 큰 특징은 시맨틱 링크(Semantic link)다. 시맨틱 링크를 통해 전자책과 학술논문을 넘나들며 자료를 연결시킨다는 것이다. 논문 상세 페이지에서는 탭을 이용해 서로 다른 정보 사이를 이동할 수 있다. 엘스비어는 2010년 8월에 콘텐트를 통합하여 제공하고 콘텐트의 연계를 강화시키는 새로운 플랫폼 사이버스(SciVerse)를 출시했다. 자사가 생산하여 제공하는 사이언스다이렉트(ScienceDirect), 스코푸스(Scopus), 사이토픽(SciTopics) 외에 엘스비어의 과학기술 웹 검색엔진인 사이러스(Scirus) 기능을 통합하고 제3자가 제공하는 콘텐트를 통합하여 제공한다. 연구자가 검색하고 관련정보를 발견할 수 있도록 하기 위해 모든 과학기술 콘텐트를 통합적으로 제공하고자 하는 것을 목표로 하고 있다. 와일리는 간결함과 쉬운 이용 방법에 대한 이용자의 요구(기대)와 새로운 웹 기술 적용에 대한 압력 등으로 인해 새로운 서비스 플랫폼을 출시하게 되었다고 한다. 콘텐트에 대한 웹 주소는 OpenURL 방식을 사용하도록 개선하였다. 프로퀘스트는 자사의 모든 콘텐트를 통합하는 것을 특징으로 2010년 6월에 새로운 플랫폼 출시 계획을 발표하였다.

이상의 출판유형 변화의 시사점은 크게 3가지로 정리될 수 있다. 첫째, 검색과 발견의 공존이다. 연구자가 원하는 정보를 획득하기 위해 가장먼저 수행하는 것은 검색이다. 여전히 검색은 연구자들이 연구를 수행하는데 있어 필수적으로 거쳐야 하는 프로세스 중 하나이다. 검색이후에는 연구자들이 필요로 하는 정보를 발견 할 수 있도록 하는 서비스가 제공되고 있다. 둘째, 자원의 연계를 이용한 학술정보 서비스가 자리매김 했음을 알 수 있다. 셋째, 콘텐트 확장이 다양해지고 있다. 연구자가 검색한 콘텐트와 연계된 데이터, 웹 자원, 단행본, 특허, 관련 논문 등, 제공하는 콘텐트가 다양해지고 있다.

이상의 시사점으로 살펴본 검색과 발견의 공존, 자원연계 서비스, 콘텐트 확장은 링킹서버의 중요성을 더욱 부각시킨다. 연구자의 선호도와 이용권한을 분석하고, 이를 기반으로 소속 기관이 제공 가능한 서비스를 조합하여, 연구자가 관심을 둔 콘텐트 이외에 다양한 서비스 제공이 가능하도록 하는 툴이 링킹서버이다. 이 링킹서버의 기능은 연구데이터 유통을 위해서 더욱 확장되어야 할 것이다. 예를 들어, 연구데이터와 관련된 OpenURL 메타태그가 확장이 되면 이를 입력 값으로 받아들여 해석할 수 있어야 하며, 목표서비스에 종속적인 URL 패턴추가 등 지식베이스(knowledge base)의 확장이 필요하다.

2.2 연구데이터 유통

연구데이터의 유통에 있어 활용측면의 핵심 요소로는 정보의 출판, 식별, 인용이라 할 수 있다. 데이터의 출판은 연구결과를 공개하는 필수 절차로서 연구의 마지막 단계라 할 수 있다. 데이터 식별은 출판된 데이터의 영속적 접근점을 제공하는 수단으로서, 그 중요성이 날로 더해지고 있다. 데이터 인용은 데이터 활용의 확산을 이끌어 낸다. 본 논문에서는 연구데이터 유통 현황을 활용의 측면에서 분석하기 위해 데이터의 출판과 식별, 인용 분야에서 대표적인 활동을 조사하였다.

출판 부문의 경우 대표적인 사례인 PANGAEA는 2010년 8월 말 현재 170개의 프로젝트 데이터를 서비스하고 있다. 세계 해양환경과학데이터센터(World Data Center for Marine Environmental Sciences, WDC-MARE)는 PANGAEA를 데이터 아카이빙과 배포시스템으로 사용하고 있으며, Earth System Science Data(ESSD)는 해당 저널의 아카이브로 PANGAEA를 지정하여 서비스하고 있다. PANGAEA서비스는 ScienceDirect와 연동되어 데이터와 논문, 논문과 데이터의 연계서비스가 가능하다. 이렇듯 문헌과 데이터가 연계되어 새로운 서비스 패러다임을 만들어 내고 있다.

데이터 식별의 경우 대표적인 사례는 DataCite를 들 수 있다. DataCite는 국제적인 컨소시엄으로서 2010년 8월 말 현재, 9개국에서 15개 기관이 참여 중에 있다. D-Lib 저널을 통해 Starr와 Gastl(2011)이 메타데이터 스킴을 발표하였으며, 2011년 7월 현재 메타데이터 커널 버전 2.2를 발표하고 개선작업을 진행 중에 있다. DataCite는 연구데이터에 DOI(Digital Object Identifier)를 부여하는 에이전시 역할을 수행하고 있으며, 식별자로서 DOI가 부여된 80만 건 이상의 레코드를 확보하고 있다. 이외에도 연구데이터 세트를 적재, 식별, 인용할 수 있도록 지원하고 있으며, 문헌 이외의 모든 데이터에 대한 기술내용 공개 및 서비스를 지원하고 있다. 데이터 공개를 위해 필요한 프로세스와 표준을 지원하며, 논문의 근거자료가 되는 연구데이터와 논문간의 연계를 제공하고 있다. DataCite 컨소시엄 및 하버드대학교의 Dataverse, PANGAEA는 연구데이터의 영속적 접근을 제공하기 위하여 DOI 등 글로벌 식별자를 등록하여 서비스에 활용하고 있다. DataCite의 경우 연구데이터에 글로벌 식별자 DOI를 부여하여 데이터 유통에 새로운 패러다임을 만들고 있다. 이는 CrossRef이 논문에 대해 DOI를 유통시킨 점과 유사하며, 연구데이터에 특화되어 있다는 것이 차이점이다. DOI를 통해 데이터의 물리적인 위치의 변화에 관계없이 데이터에 대한 영속적인 접근을 연구자에게 제공할 수 있게 되었으며, 기존 학술정보 유통 채널과의 연계를 통해 연구자에게 영구적인 접근을 허용함으로써 보다 나은 안정적 연구 환경을 제공할 수 있게 되었다.

데이터의 인용을 위한 표준 포맷에 관한 연구도 활발하게 진행 중에 있다. 주로 DataCite와 OECD, DIF가 주도적인 표준화 작업을 진행하고 있다. DataCite는 연구데이터를 기술하기 위한 스키마를 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〉 요소를 활용하여 데이터 인용과 관련된 정보를 기술할 수 있도록 하고 있다(김선태, 이태영 2011).

이상에서 살펴본 연구데이터의 출판, 식별, 인용은 학술정보 출판유형의 변화와 무관하지 않다. 데이터를 학술정보 출판에 포함시키려는 움직임은 이미 학술정보에 새로운 서비스 유형으로 자리를 잡았다. 일례로 화학분야의 국제적인 학술출판사로서 2011년 3월 현재 32종의 학술지를 발행하고 있는 영국왕립화학회(2010)는 논문의 주요 화합물 정보를 제공하기 위하여 저널 플랫폼을 개선하였으며, 검색결과를 재조정하고 직관적인 필터기능을 추가하여 약 100만 건의 논문에 대한 검색서비스를 개선시켰다. 주요 개선항목으로 화학 구조나 특성, 관련 특허 등에 접근할 수 있게 하였다. 엘스비어는 '미래의 논문' 프로젝트를 2009년 7월에 발표했다. 과학 논문을 온라인에서 어떻게 표현할지를 재정의 하기 위해 엘스비어사와 과학 커뮤니티 간 협력하여 진행 중인 프로젝트이다. 주요 특징은 문자와 그림을 계층적으로 표현한 것이다. 이 조직구조는 전통적인 인쇄형 논문이 가지고 있는 선형조직을 벗어나는 중요한 출발점이다. 이 새로운 조직구조는 핵심이 되는 텍스트와 부가적 자료들을 단일 구조 내에서 통합시킨다. 이 구조는 온라인 환경의 장점을 최대한 이용하고 있다. 이것은 논문과 관련된 자원들의 통합발견이 가능하도록 하는데 큰 의미가 있다.

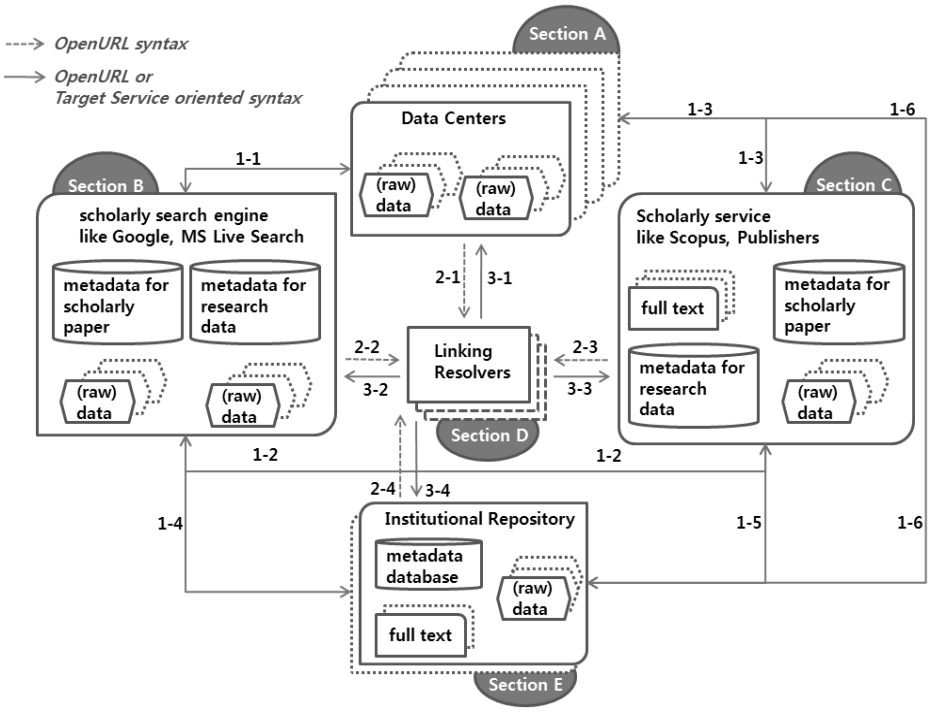
학술정보 출판 유형이 연구데이터를 포함하고, 연구자가 관심을 둔 콘텐트 이외에 새로운 콘텐트를 발견할 수 있도록 변하고 있다. 따라서 학술정보 활용에 있어 링킹서버 및 OpenURL 표준인 Z39.88의 확장이 필요하다. 이하에서는 Z39.88의 학술서비스 유형을 확장하기 위해 최근 서비스되고 있는 학술서비스의 유형을 구체적으로 살펴보고, 새로운 관련자원의 발견을 향상시킬 수 있는 DataCite 스킴의 〈RelatedIdentifier〉 요소에 대해 분석한다.

2.3 연구데이터 유통과 링킹서버 역할

문헌콘텐트 중심의 학술정보 유통에서 링킹서버는 세 가지 측면에서 역할을 수행하고 있다. 첫째, 가장 기본적인 역할로서 정보소스와 목표서비스를 연결해주는 매개자 역할을 수행한다. 둘째, 연구자가 활용할 수 있는 관심 콘텐트의 복사본 문제를 해결하는데 중심적인 역할을 수행하고 있다. 문헌콘텐트의 확산으로 이해관계를 가진 문헌콘텐트 유통 주체들이 연구자가 관심을 둔 콘텐트를 중복으로 서비

스할 경우에 발생할 수 있는 문제를 해결하는 것이다. 셋째, 연구자가 소속된 기관의 라이선스 정보를 종합하여 연구자에게 발견의 기회를 제공하는 역할을 수행하고 있다. 연구데이터 유통에서 링킹서버의 역할은 변함없이 요구되고 있다. Ware와 Mabe(2009)는 STM 보고서에서 데이터가 독립적인 출판물이 될 것이라 예측하였다. 〈그림 1〉은 연구데이터 유통 주체들과 링킹서버의 관계를 도식화해 표현한 것이다. 〈그림 1〉에서 OpenURL 구문이 사용되는 것은 점선 화살표로 표기되어 있으며, OpenURL 및 목표서비스에 기반 한 구문 전달은 실선 화살표로 표기되어 있다. 섹션별 상세설명은 다음과 같다.

섹션A는 연구데이터를 가지고 있는 데이터 센터들을 의미한다. 연구데이터 센터들의 허브일 수 있으며, 독자적으로 데이터 센터 역할을 수행하는 기관의 서비스일 수 있다. 섹션B는 구글과 같은 검색엔진 서비스를 의미한다. 검색엔진의 특성상 주로 학술논문이나 특허, 연구보고서 등 학술 출판물과 연구데이터에 대한 메타데이터 검색관련 서비스를 제공하는 부문을 의미한다. 〈그림 1〉에서 〈1-1〉, 〈1-2〉, 〈1-3〉, 〈1-4〉, 〈1-5〉, 〈1-6〉이 의미하는 것은 각 섹션이 링킹서버와 통신하지 않고 직접적인 통신을 하는 것을 의미한다. 〈2-1〉, 〈2-2〉, 〈2-3〉, 〈2-4〉는



〈그림 1〉 연구데이터 유통 주체들과 링킹서버의 관계

각 섹션이 링킹서버에게 OpenURL 메타태그를 전달하는 것을 의미한다. 〈3-1〉, 〈3-2〉, 〈3-3〉, 〈3-4〉, 〈3-5〉는 링킹서버가 연구자에게 목표서비스(target service) 목록을 제공하고 연구자가 선택한 목표서비스에게 메타태그를 전달하는 것을 의미한다. 〈1-x〉와 〈3-x〉의 경우는 OpenURL 메타태그가 사용될 수도 있으나, 주로 목표서비스에서 별도로 지정한 서비스 API 구문을 준수한다. 섹션B는 구글이나 마이크로소프트의 라이브서치와 같은 학술정보 검색을 제공하는 검색포털의 영역이다. 2008년 1월에 구글은 공개적으로 접근이 가능한 연구데이터에 대한 호스팅 서비스를 계획하고 있음을 선언하였고, 또한 마이크로소프트사는 연구자들의 연구성과물을 관리하는 저장소를 개발 중에 있기 때문에(CUL Data Woking Group 2008) 섹션B의 경우 (원시)데이터를 포함시켰다. 섹션C는 Scopus나 WoS와 같은 인용색인 데이터베이스나 출판사 서비스를 의미한다. 다양한 검색 방법을 제공하며, 학술논문의 원문뿐만 아니라 (원시)데이터 및 가공데이터를 가지고 있을 수 있다. ALPSP(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. 이러한 조사결과는 학술정보 유통에서 연구데이터의 중요성이 심화되는 것을 의미한다. 또한 관련 커뮤니티에서 연구데이터 이용의 활성화를 준비해야 함을 시사한다. 섹션E는 연구자가 소속된 기관의 연구데이터 저장소를 의미한다. 2010 DataCite 여름 회의에서 세계 과학・기술・의료 출판연합회 이사 Smit(2010)의 발표 자료에 의하면 세계 과학기술의료 분야의 70%의 출판사 중 90% 저널이 데이터를 받고 있으며, 그중 95% 출판사는 논문을 데이터에 연결시켜 서비스하고 있다. 또한 5% 미만의 출판사가 데이터세트를 관리하기 위한 인프라를 보유하고 있다. 60% 출판사는 데이터세트 유지 및 관리 책임이 저자나 연구기관에 있다고 생각한다. 이러한 발표 내용은 출판사는 연구데이터 유통에 대한 의지는 있으나 관리에 대한 의지가 없음을 의미한다. 이것을 통해 연구자가 소속된 기관이나 분야별 데이터센터, 데이터 호스팅 기관이 연구데이터 유통에서 중요한 역할을 수행하게 될 것을 알 수 있다. OpenDOAR(Directory of Open Access Repositories)에 의하면 2010년 7월 기준으로 1,345개 기관에서 1,696개의 학술관련 개방 접근 리포지터리가 등록되어 있다. 대륙별로는 유럽이 49%로 가장 많은 비중을 차지하고 있다. 단일 국가로는 미국이 371개 리포지터리를 운영하고 있어 가장 활발한 연구데이터 개방을 추진하고 있는 것을 확인할 수 있다. 학술관련 리포지터리와 링킹서버의 연계도 매우 중요하다. 섹션D는 학술정보 유통에 있어 연구데이터의 물리적인 위치와 복사본의 문제, 연구자 소속기관의 연구데이터 라이선스 문제를 지능적으로 해결하여 연구자 상황에 맞는 서비스를 연구자가 선택할 수 있게 하는 링킹서버 부

문이다. 학술논문의 경우 직접적인 출판사 서비스, 출판사의 원문정보를 모아 서비스하는 애그리게이터(aggregator) 서비스가 있듯이 연구데이터의 경우도 데이터 소유권과 관련하여 새로운 유통 서비스 체제가 만들어 질 것이다. 따라서 문헌중심의 학술서비스 유형의 OpenURL 메타태그가 연구데이터의 서비스 유형을 포함하도록 확장되어야 한다.

3. 학술서비스 유형 도출

3.1 학술서비스 유형 분석

범용적으로 사용될 수 있는 OpenURL 학술 커뮤니티 서비스 유형의 핵심 메타태그를 개발하기 위하여 대표적인 학술정보서비스인 해외 Elsevier 출판사의 Scopus와 ISI의 Web of Science(이하 WoS), 국내 과학기술정보 대표 포털서비스인 NDSL(National Discovery for Science Leaders, 이하 NDSL)을 중심으로 조사하였다. Scopus와 WoS는 초록・인용 데이터베이스 중에 세계 최고의 서비스로 인정받고 있으며, NDSL은 국내에서 제공되는 대표적인 과학기술정보서비스이기 때문에 분석대상으로 선정하였다. 각각의 서비스는 다양한 유형의 서비스를 제공하고 있지만, 본 연구에서는 OpenURL 메타태그로 적합한 항목만 도출하였다.

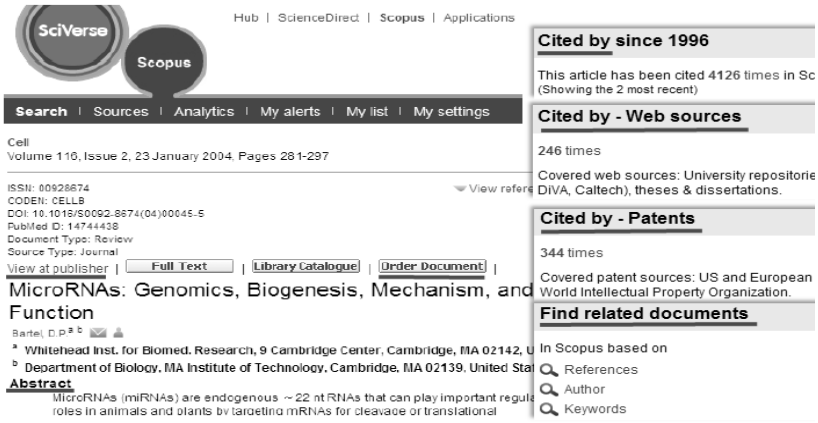
3.1.1 Scopus

Elsevier가 제공하는 Scopus는 우수한 연구논문과 양질의 웹 기반 정보들로 이루어진 최대의 초록・인용 데이터베이스이다.

과학자들이 원하는 정보를 효율적으로 찾을 수 있도록 고안된 Scopus는 연구과정에서의 자료 수집을 빠르고 쉽게 도와준다. 매일 업데이트되는 Scopus를 통해 연구자들은 항상 최신의 정보를 검색하며, 원문 출판사 사이트로의 링크를 따라 Full-text를 받아볼 수 있다. 논문뿐 아니라 저자별로도 피인용 분석을 할 수 있는 서비스도 제공한다. 5,000여 출판사에서 발간하는 약 18,000여 종의 학술정보와 350권의 단행본 시리즈, 360만 건의 컨퍼런스 논문을 서비스 하고 있다. 1996년 이후 자료는 2천만 건의 참고문헌까지 제공하고 있으며, 1826년부터 1996년까지의 논문초록을 구축하여 서비스하고 있다.

〈그림 2〉는 Scopus에서 저자ID가 '7006581498'로 식별되는 Bartel의 논문을 검색하여 학술정보서비스 유형의 항목을 표기한 것이다. 실제 서비스 화면의 관련 부분을 캡처하여 재구성한 화면으로써 실제 모습과 다르다.

Scopus에서 서비스하고 있는 대표적인 학술정보서비스 유형은 그림에서 밑줄로 표기되어 있으며, 이외에도 다양한 서비스가 존재하지만 OpenURL 메타태그로 적합한 것만을 표기하였다.

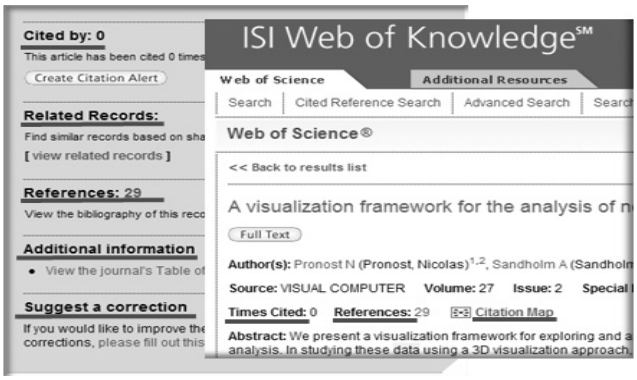


〈그림 2〉 Scopus 논문 상세화면 기능구성(논문: MicroRNAs: Genomics, Biogenesis, Mechanism, and Function, 검색일: 2011.3.14)

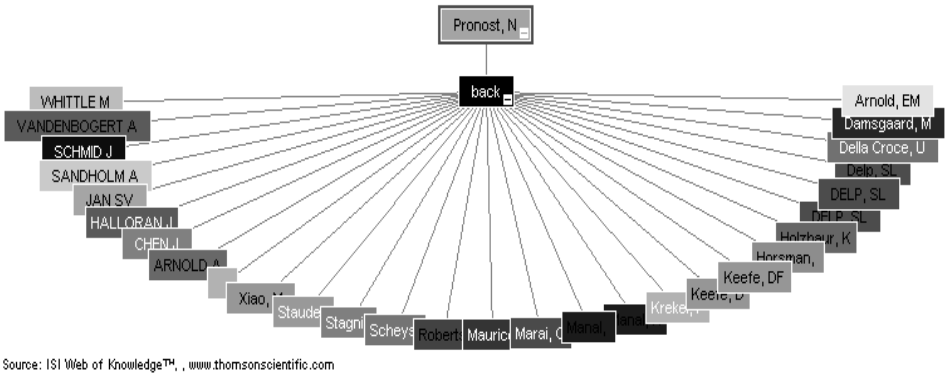
3.1.2 Web of Science

WoS는 세계적으로 권위를 인정받은 선도적 인용 데이터베이스로써 빠르고 용이한 접속 방식과 검색으로 256개 분야의 포괄적인 콘텐츠를 제공하고 있다. ISI가 제공하는 인용색인 데이터베이스인 SCIE(Science Citation Index Expanded), SSCI(Social Sciences Citation Index), A & HCI(Art & Humanities Citation Index)를 웹에서 동시에 검색할 수 있는 웹 데이터베이스이다. 오픈액세스 저널 및 회의록 정보와 전 세계 영향력 높은 1만 종 이상의 과학, 사회과학, 인문예술 분야 저널을 수록하고 있으며, 1900년도까지의 과거 정보를 제공하고 있다.

〈그림 3〉은 WoS에서 글로벌 식별자인 논문의 DOI가 '10.1007/s00371-010-0534-y'로



〈그림 3〉 WoS 논문 상세화면 기능구성(논문: A visualization framework for the analysis of neuromuscular simulations, 검색일: 2011.3.12)



〈그림 4〉 Citation Map(논문: A visualization framework for the analysis of neuromuscular simulations, 옵션: Backward Only, 검색일: 2011.3.12, Cited by 정보는 없으므로 Forward Only 정보는 없음)

식별되는 논문을 검색하여 학술정보서비스 유형의 항목을 표기한 것이다. 실제 서비스 화면의 관련 부분을 캡처하여 재구성한 화면으로써 실제 모습과 다르다. WoS에서 서비스하고 있는 대표적인 학술정보서비스 유형은 그림에서 밑줄로 표기되어 있으며, 이외에도 다양한 서비스가 존재하지만 OpenURL 메타태그로 적합한 것만을 표기하였다. 〈그림 4〉는 WoS에서 차별화 서비스로 제공하고 있는 Citation Map 서비스의 모습을 보여 준다.

3.1.3 National Discovery Science Leaders

NDSL은 국가 과학기술 R&D 활동을 지원하기 위해 KISTI가 개발·운영 중인 대표 서비스이다. NDSL에서는 국내논문, 해외논문, 국내연구보고서, 동향, 특허, 산업표준, 사실정보와 같은 다양한 유형의 정보를 통합적으로 검색할 수 있으며, 각 정보 유형별 검색 또한 가능하다. NDSL은 논문과 특허, 특허와 사실정보와 같이 서로 다른 종류의 콘텐트를 연결하고, 정보를 제공할 때 연결된 정보를 함께 제공함으로써 높은 정보접근성(information accessibility)을 제공할 수 있을 뿐만 아니라 차별화된 개인별 맞춤 서비스 역시 가능하게 한다.

NDSL 서비스는 Open API, OpenURL과 같은 개방형 정보기술을 활용하여 국내 다른 기관에서 쉽게 활용할 수 있다. NOS(NDSL Open Service)로 불리는 개방형 서비스를 통하여 국내 기관들은 각 기관의 홈페이지에서 KISTI 콘텐트를 마치 자신들이 보유한 정보처럼 편리하게 이용할 수 있으며, OpenURL을 통해 자관의 서비스에서 NDSL 서비스로 직접적인 연계가 가능하다.

〈그림 5〉는 논문 검색화면에 대한 예시로서 연구자가 이용할 수 있는 학술정보서비스 유형을 보여준다. 실제 서비스 화면의 관련 부분을 캡처하여 재구성한 화면으로서 실제 모습과 다르다.

〈그림 5〉 NDSL 논문 상세화면 기능구성
(논문: DOI 수집 API 개발에 관한 연구, 검색일: 2011.3.18)

3.2 DataCite 스킴 분석

DataCite 메타데이터 커널은 5개의 필수요소 〈Identifier〉, 〈Creator〉, 〈Title〉, 〈Publisher〉, 〈PublicationYear〉와 12개의 선택요소 〈Subject〉, 〈Contributor〉, 〈Date〉, 〈Language〉, 〈ResourceType〉, 〈AlternateIdentifier〉, 〈RelatedIdentifier〉, 〈Size〉, 〈Format〉, 〈Version〉, 〈Rights〉, 〈Description〉으로 구성되어 있다. 이중 〈RelatedIdentifer〉 요소는 관계유형(relation type) 속성을 갖는다. 관계유형의 값으로 사용될 수 있는 값은 모두 18개로 구성되어 있다. 이러한 요소와 속성은 연구논문과 구성요소로서의 그림과 표, 데이터셋 그리고 다른 보조자원을 모두 엮은 복합문서(composite document)를 기술하는 데 유용하게 사용될 것이다.

〈RelatedIdentifier〉요소와 관계유형 속성을 함께 사용함으로 학술논문과 그 연구의 기초자료가 되었던 데이터의 관계를 기술할 수 있다. 또한 여러 버전의 자원들과의 관계와 다양한 유형의 자원과의 관계 또한 기술할 수 있다. 하나의 자원을 구성하는 요소자원을 기술할 수 있으며, 기술된 자원의 보충 자원도 기술 가능하다. 이 요소는 반복적으로 기술이 가능하며, 폭넓은 자원을 대상으로 하고, 미묘한 관계의 유형을 자세하게 기술할 수 있게 한다(Starr and Gastl 2011).

4. KEV 메타데이터 형식

2011년 4월 현재 OpenURL 프레임워크 레지스트리에 등록된 학술서비스 유형 메타데이터 형식은 〈표 1〉과 같이 총 6가지의 메타태그 〈abstract〉, 〈citation〉, 〈fulltext〉, 〈holdings〉, 〈ill〉, 〈any〉로 구분되어 있다. 주로 문헌 중심의 연구 결과물과 관련된 정보만을 학술적 서비스 유형으로 보았기 때문에 6개의 메타태그로 구성되었지만 최근의 학술정보서비스의 변화와 연구자가 관심을 갖는 레퍼런트(referent) 유형의 다변화는 학술서비스 유형을 위한 메타태그 확장을 요구하고 있다. 따라서 연구데이터와 관련된 학술서비스 유형을 요청할 수 있는 메타데이터 형식 개발이 필요하다.

연구데이터는 생성된 데이터 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. 이에 따라 현재 운용하고 있는 여섯 가지의 메타태그만으로는 이용자요구를 수용하는데 한계를 갖고 있다. 왜냐하면 'any' 키를 활용하여 그 이상의 의미를 수용할 수 있지만, 'any'키 값의 특성상 이용자들이 요구하는 깊이 있는 의미를 하나의 키로 모든 것을 표현하는데 제약을 갖고 있기 때문이다.

5. KEVFormat: Sch-Svc 확장

5.1 메타태그 설계원칙

메타태그를 설계하는 데 있어서 관련 분야의 표준을 최대한 반영하고자 노력하였다. 특히 정보서비스 사이에 요청될 수 있는 서비스 유형을 도출하기 위해서 대표적인 국내・외 학술서비스를 조사하였으며, 연구데이터 기술을 위한 최신의 메타데이터 포맷을 반영하였다. 이를 위해, 세계적인 초록・인용 데이터베이스로 인정받고 있는 해외의 Scopus와 WoS, 국내의 NDSL 서비스를 조사 대상으로 선정・분석하였다. 또한 연구데이터 기술을 위해 2011년 2월에 발표된 DataCite 메타데이터 요소 중 〈RelatedIdentifier〉의 속성값을 모두 수용하였다. DataCite 스킴은 연구

〈표 1〉 학술적 서비스 유형을 위한 key 구성(NISO AX Committee 2007, 15)

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

데이터를 기술하기 위한 다양한 요소들을 가지고 있으며 특히, 자원들의 관계를 기술하기 위한 〈RelatedIdentifier〉요소의 값은 최근에 학술 서비스 유형의 핵심이 되고 있는 자원 연결을 통한 추천서비스의 필수요소로 활용될 수 있다. 이것은 Starr와 Gastl(2011)이 DataCite 스킴의 세 번째 기능으로 제시한 것으로서, 데이터에 대한 융통성 있는 기술을 가능하도록 한다. 즉, 다른 자원과의 관계를 기술함으로써 데이터세트의 발견이 더욱 용이하도록 하는 기능을 강조한 것이다. 따라서 본 연구에서는 관련 값들을 그룹화하여 대표 속성 도출 후 메타태그로 선정하였다.

본 연구의 목적은 OpenURL 소스와 서비스 구성요소가 학술서비스 유형 전달에 있어서 상호운용성을 확보를 할 수 있도록 OpenURL 핵심 메타태그를 KEV 형식으로 확장하여 제안하는 것이다. 다음은 KEVFormat: Sch-Svc 확장을 위해 준수한 설계방향이다.

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

나. Z39.88 표준을 준수하여 메타태그명을 수정함(Z39.88)

다. 최근의 국내・외 학술서비스 유형을 파악하고 정보서비스 소스와 목표서비스 사이 요청될 수 있는 서비스 유형을 도출(Scopus, WoS, NDSL)

라. 연구데이터 기술을 위한 최근 메타데이터 포맷을 반영(DataCite)

마. 기존 Z39.88을 구현한 링킹서버와 연동되는 레거시 시스템의 호환성을 위해 기존 표준 메타태그를 모두 포함(KEVFormat: Sch-Svc)

바. 특정 서비스 데이터베이스에 의존적이며 독자적인 학술서비스 유형은 제외

상기의 기준으로 'KevFormat: Sch-Svc' 메타태그를 기준으로 Scopus, WoS, NDSL 서비스의 학술서비스 유형을 분석하여 연결테이블(Crosswalks)를 작성하였다. OpenURL 메타태그로 적합한 항목만 도출하였다. 다음으로 DataCite 〈RelatedIdetifier〉요소의 관계유형 속성을 분석하여 대표속성을 도출하였다. 다음 〈표 2〉는 Z39.88의 학술서비스 유형 메타태그를 기준으로 세계적인 초록・인용데이터베이스인 Scopus, WoS와 국내의 과학기술정보 서비스인 NDSL의 학술서비스 유형을 비교한 연결테이블(Crosswalks)이다.

본 연구에서 메타태그 선정 시 연구자들의 의견을 수렴하지 않은 이유는 크게 두 가지 이다. 첫째, 학술서비스 유형 메타태그의 특성상, 연구자에게 제공할 수 있는 서비스의 유형은 최신의 서비스 유형을 검토하여 확장요소를 도출할 수 있기 때문이다. 둘째, DataCite에서 발표한 메타데이터 형식이 자원들의 관계를 기술하기 위해 부족함이 없으며, 최신의 추천 서비스 동향을 반영하고 있기 때문이다. 하지만

〈표 2〉 학술정보 제공서비스 비교(기준 2011.3.9)

KEVFormat:Sch-Svc http://bit.ly/hE48SC	Scopus http://www.scopus.com	WoS http://isiknowledge.com/	NDSL http://scholar.ndsl.kr
abstract	show abstract	abstract	초록
fulltext	view at publisher / Full Text	Full Text	원문
citation	Cited by in Scopus	Citedby(from Web of Science)	-
holdings	-	-	소장처 리스트
ill	-	-	-
any	Cited by - Web sources	-	-
	Cited by - Patents	-	연관자료(특허)
	-	-	연관자료(보고서)
	-	-	연관자료(동향)
	-	-	연관자료(표준)
	Find related documents	Related Records	추천정보보기
	Citation counts in references	References	(국내논문 일부제공)
	order document	-	원문복사
	-	-	NTIS 과제정보
	-	-	연관검색어 리스트
	-	Citation Map	-
	view author details	-	-
	저자의 다른 논문리스트	저자의 다른 논문리스트	-
	References	References(from Web of Science)	참고문헌

연구데이터 식별을 위한 메타태그는 연구 분야, 연구데이터의 특성 등에 따라 다를 수 있기 때문에, 메타태그 도출 과정에서 관련분야 전문가의 검토가 꼭 필요하다.

5.2 핵심 메타태그

세계적인 초록・인용데이터베이스인 Scopus와 WoS, 국내에서 대표적인 과학기술정보 서비스로 인정받는 NDSL의 학술서비스 유형을 비교하여 핵심 메타태그를 도출하였다. 〈표 3〉은 학술정보서비스 비교를 통해 대표속성을 선정한 이유를 설명한다. SCOPUS 서비스의 경우 비교대상 서비스와 크게 차별화할 수 있는 것은 보다 구체적인 인용정보 제공을 꼽을 수 있다. Scopus 데이터베이스에서 인용된 정보와 웹 자원으로부터 인용된 정보, 특허로부터 인용된 정보로 나누어 인용정보를 제공하고 있다. 이러한 서비스 유형은 정보서비스가 검색에서 이용자의 발견과 이용자의 상황을 인지하여 추천하는 서비스로 진화하기 있기 때문에 그 의미가 매우 크다. WoS 데이터베이스의 경우 타 서비스에서 제공하지 않는 인용정보 지도(Citaton Map) 서비스를 제공하고 있다. 이것은 포워드 인용(Citing)과 백 워드 인용(Cited) 정보를 지도로 살펴볼 수 있는 서비스로서 인용의 깊이를 2단계까지 지정하여 논문의 인용정보를 제공하는 것이다. 연구데이터의 경우도

〈표 3〉 학술정보서비스 비교를 통한 대표속성 선정

Scopus	WoS	NDSL	대표속성	선정 이유
○	○	○	abstract	KEV Format: sch_svc과의 호환
○	○	○	fulltext	KEV Format: sch_svc과의 호환
○	○	-	citation	KEV Format: sch_svc과의 호환
-	-	○	holdings	KEV Format: sch_svc과의 호환
○	-	-	webcitation	Discovery 서비스 핵심요소로서 인용정보 제공의 세분화 필요
○	-	○	patentcitation	Discovery 서비스 핵심요소로서 인용정보 제공의 세분화 필요
-	-	○	reportcitation	Discovery 서비스 핵심요소로서 인용정보 제공의 세분화 필요
○	○	○	recommendation	Discovery 서비스 핵심요소로서 조사대상 서비스 모두 제공
○	-	○	rdds	문헌중심에서 연구데이터 확장 필요
-	○	-	citationmap	연구자 그룹 및 연구동향을 파악하는데 핵심요소
○	-	-	authordetail	연구데이터에 고유 식별자를 부여 하듯이 연구자 개인에게 식별자를 부여하는 것이 새로운 추세 임. 연구자 식별을 위한 핵심요소
○	○	-	authorwork	Discovery 서비스 핵심요소로서 저자의 연구데이터 제공
○	○	○	reference	DataCite 대표속성과 동일함. Discovery 서비스 핵심요소

데이터의 인용을 위한 표준이 마련되고 있으므로 향후 학술커뮤니케이션 분야에서 데이터에 대한 인용이 활발해지고 인용된 데이터를 생산한 연구자에 대한 기여도가 제대로 평가받는다는 가정에서는 매우 중요한 서비스 요소로 활용될 수 있다. NDSL은 네 가지의 특화서비스를 제공하고 있다. 소장처 리스트 제공과 원문복사, NTIS 과제정보, 연관검색어 리스트이다.

소장처 리스트 제공 서비스는 NDSL이 KESLI 컨소시엄을 운영하면서 컨소시엄에 참여하는 대학, 연구소, 병원 등 국내기관으로부터 수집한 인쇄학술지 소장정보 데이터베이스를 기반으로 서비스 하는 항목이다. 물론 원문복사 서비스의 경우도 같은 데이터베이스를 기반으로 한다. NDSL서비스의 '연관자료(동향)', '연관자료(표준)', 'NTIS 과제정보', '연관검색어 리스트'는 메타태그 설계방향 '바'항목에 해당함으로 선정하지 않았다. 3개 서비스에서 공통적으로 제공한 참고문헌 건수(Citation counts in references)는 학술서비스 유형에 부합되지 않아 선정하지 않았다.

〈표 4〉는 DataCite 스킴 중 〈RelatedIdentifier〉 요소의 type속성 값과 그에 대한 설명 그리고 사용가능 값을 그룹화 과정을 통해 대표속성을 도출한 결과를 보여준다. 자원 A와 B가 특정 관계를 맺고 있다고 가정할 경우 'A 속성 값 B' 구문의 의미를 설명하고 있다. 대표속성 중 'citation'은 기존 표준 메타태그 KEVFormat: Sch-Svc에 포함되어 있으며, 'Reference'는 학술서비스 유형 메타태그에서도 도출되어 그 중요성이 매우 크다. 총 18개의 사용가능 값은 두 개의 쌍으로 구성되어 있는데 각각을 대표하는 하나의 대표속성을 도출하였다. 사용가능

〈표 4〉 DataCite 스킴 RelatedIdentifier 요소의 type 속성 값과 대표속성 도출

사용가능 값	설명	대표속성
A IsCitedBy B	A를 인용한 저작(work)	citation
A Cites B	A가 인용한 저작	
A IsSupplementTo B	A가 B의 보충 자원	supplement
A IsSupplementedBy B	B가 A의 보충 자원	
A IsContinuedBy B	A 다음으로 B가 연속적으로 생성됨	continuation
A Continues B	A가 B에 이어 연속적으로 생성됨	
A IsNewVersionOf B	B의 수정이나 갱신 후 A가 생성됨	version
A IsPreviousVersionOf B	A가 B의 이전 버전임	
A IsPartOf B	A가 B의 일부임	part
A HasPart B	A가 B를 포함	
A IsReferencedBy B	A가 B의 정보소스로 사용되었음	reference
A References B	A가 B를 정보소스로 사용	
A IsDocumentedBy B	A에 대한 설명을 B가 하고 있음	document
A Documents B	A가 B를 설명함	
A IsCompiledBy B	B의 편집 및 컴파일작업 후 A가 생성됨	compilation
A Compiles B	A는 B를 만들어내기 위해 편집 및 컴파일 됨	
A IsVariantFormOf B	A는 B의 다른 형태의 자원	form
A IsOriginalFormOf B	A는 B의 최초 형태의 자원	

값의 대표 속성을 메타태그 설계원칙 '가' 항목에 맞도록 명명하였다.

5.3 KEV 형식 정의

OpenURL 학술서비스 유형 메타태그 정의에 사용한 제한 언어(Constraint Language)로는 OpenURL 레지스트리에서 관리되고 있는 "The Z39.88-2004 Matrix Serialization (식별자: info:ofi/fmt:kev:mtx)" 형식을 따랐다. 키 이름은 표준에 맞게 변경하였다. 모두 소문자로 변경하였으며, '_' 문자와 공백을 사용하지 않았다. 〈표 5〉는 "Z39.88-2004 Matrix Serialization" 언어에 맞게 작성된 확장된 OpenURL 학술서비스 유형 메타태그 정의 이다. 실제 정의되어 있는 모습은 http://bit.ly/giyqa4에서 확인 가능하다. 'Min', 'Max' 열의 값은 학술서비스 유형 메타태그의 특성 상 서비스 요청 여부만을 나타내기 때문에 모두 '0'과 '1'로 설정하였다.

확장 정의된 OpenURL 학술서비스 유형 메타태그의 우수성은 크게 네 가지이다. 첫째, ANSI/NISO Z39.88-2004 표준을 준수하기 때문에 관련 커뮤니티 연구자들뿐 아니라 관련분야 개발자들이 쉽게 이해할 수 있다. 둘째, 메타태그의 최신성이다. 연구데이터의 출판에 앞장서고 있는 DataCite의 최근 스키마에서 자원의 관계를 기술하는 〈RelatedIdentifier〉

〈표 5〉 Scholarly Service Type를 표현하기 위한 KEV 형식의 메트릭스 정의

Delim	Key	Equals	Value	Min	Max	Description
#						이하 기존에 활용되는 Key
&	abstract	=	“yes”, “no”	0	1	연구데이터 초록정보 요청 여부
&	fulltext	=	“yes”, “no”	0	1	연구데이터 원문정보 요청 여부
&	citation	=	“yes”, “no”	0	1	연구데이터 인용정보 요청 여부
&	holdings	=	“yes”, “no”	0	1	연구데이터 소장처정보 요청 여부
&	ill	=	“yes”, “no”	0	1	연구데이터 상호대차 요청 여부
&	any	=	“yes”, “no”	0	1	연구데이터 관련 요청 여부
#						이하 DataCite 스키마에서 추출된 Key
&	supplement	=	“yes”, “no”	0	1	보충자원 정보제공 요청 여부
&	continuation	=	“yes”, “no”	0	1	연속성 정보제공 요청 여부
&	version	=	“yes”, “no”	0	1	버전 정보제공 요청 여부
&	part	=	“yes”, “no”	0	1	파트(부분) 정보제공 요청 여부
&	reference	=	“yes”, “no”	0	1	레퍼런스 정보제공 요청 여부
&	document	=	“yes”, “no”	0	1	연구데이터 설명자료 요청 여부
&	compilation	=	“yes”, “no”	0	1	연구데이터 편집정보 요청 여부
&	form	=	“yes”, “no”	0	1	연구데이터 형태정보 요청여부
#						이하 학술정보서비스에서 추출된 Key
&	rdds	=	“yes”, “no”	0	1	연구데이터 전송요청 여부
&	webcitation	=	“yes”, “no”	0	1	연구데이터 관련 웹정보 요청 여부
#	patentcitation	=	“yes”, “no”	0	1	연구데이터 관련 특허정보 요청 여부
&	reportcitation	=	“yes”, “no”	0	1	연구데이터 관련 보고서정보 요청 여부
&	recommendation	=	“yes”, “no”	0	1	연구데이터 관련 추천정보 요청 여부
&	authordetail	=	“yes”, “no”	0	1	연구데이터 관련 저자정보 요청 여부
&	citationmap	=	“yes”, “no”	0	1	연구데이터 관련 인용맵 요청 여부
&	authorwork	=	“yes”, “no”	0	1	저자의 다른 연구데이터 요청 여부

요소의 속성 값과 관련 된 최근 연구결과를 반영하였다. 또한 학술 커뮤니케이션에서 널리 활용되고 있는 초록・인용데이터베이스의 최신 학술서비스 유형을 반영하였다. 셋째, 메타태그의 상호운용성이다. 메타태그 설계원칙 ‘마’ 항목을 준수하여 기존 Z39.88을 구현한 링킹서버와 연동되는 레거시 시스템의 호환성을 위해 기존 표준 메타태그를 모두 포함시켰다. 넷째, 미래를 준비하는 예비성을 들 수 있다. 〈reference〉와 〈recommendation〉 메타태그 이외에는 검토된 최신 학술서비스 유형들이 모두 매핑 되는 요소가 없다. 이것은 선정된 대표속성이 향후 서비스들의 개선을 예측하여 선정되었음을 의미한다. 〈rdds〉의 경우 Scopus (order document)와 NDSL(원문복사서비스)에만 출현하는 요소이지만 향후 연구데이터의 요청을 포함하는 메타태그로 판단하여 예비 메타태그로 선정되었다.

6. 결 론

OpenURL 메타데이터 레지스트리에 등록된 문헌위주의 학술서비스 유형 메타태그 KEVFormat: Sch-Svc를 기반으로 연구데이터 유통을 위한 메타태그를 확장하였다. 최신의 학술서비스 유형을 도출하기 위해 해외의 대표적인 초록·인용 데이터베이스인 Elsevier 출판사의 Scopus와 ISI의 WoS, 국내 과학기술정보 대표 포털서비스인 NDSL을 비교 검토하여 연결테이블(Crosswalks)을 작성하였다. 이후 메타태그의 상호운용성을 확보하기 위해 KEVFormat: Sch-Svc에서 제시된 메타태그를 포함하고, OpenURL 메타태그로 적합한 학술서비스 유형을 도출하였다.

연구데이터 출판을 위해 DataCite에서 발표한 최신의 메타데이터 스킴 중 자원간의 관계를 구체적으로 기술하는 〈RelatedIdentifier〉 요소의 속성 값을 분석하여 대표속성을 도출하여 추가하였다. "Z39.88-2004 Matrix Serialization(식별자: info:ofi/fmt:kev:mtx)" 형식을 제한 언어(Constraint Language)로 사용하여 확장된 학술서비스 유형 메타태그(http://bit.ly/eddqUB)를 정의하였다.

후속 연구로는 김선태, 이태영(2011) 연구와 본 연구의 결과를 토대로 하여, 관측분야의 과학데이터와 관련된 메타태그를 전송할 수 있는 Z39.88 개발 연구가 필요하다. 특히 시스템의 상호운용성 확보를 위해서는 메타데이터 스킴 분석과 학술정보서비스 요소를 분석하는 것은 물론, 웹서비스 표준 기술에 대한 분석도 필요하다. 또한 과학데이터 사이의 관계를 기술하는 용어들에 대한 분석이 필요하다. 이러한 '하이브리드적인 메타데이터 요소 도출 방법(이하 하이브리드 메타태그 도출방법이라 함)을 사용하여 메타태그를 선정하고, 과학데이터에 관한 메타태그를 전송하기 위한 Z39.88을 개발해야 한다. 물론 하이브리드 메타태그 도출방법에 대한 객관적인 검증과 도출된 메타태그를 수용하기 위한 지식베이스 설계 및 실험 연구도 필요하다.

참고문헌

김선태, 이태영. 2011. 관측분야 과학데이터 관련 메타데이터 상호운용성 확보를 위한 OpenURL 메타태그 연구. 『정보관리연구』, 42(3): 147-165.

김선태, 한선화, 이태영, 김용. 2010. 과학데이터 보존 및 활용모델에 관한 연구. 『한국비블리아학회지』, 21(4): 81-93.

Starr, Joan and Angela Gastl. 2011. "isCitedBy: A Metadata Scheme for DataCite." *D-Lib Magazine*, 17(1/2).

Smit, Eefke. 2010. "The STM View on Data Publication." *DataCite Summer Meeting*.

Starr, Joan, Jan Ashton, Brase Jan, Paul Bracke, Angela Gastl, Jacqueline Heller, lfred Heller, Birthe Krog, Lynne

McAvoy, Karen Morgenroth, Elizabeth Newbold, Madeleine de Smaele, Anja Wilde, Scott Yeadon, Wolfgang Zenk-Möltgen. 2010. "DataCite Metadata Kernel for the Publication and Citation of Research Data." *DataCite Metadata Working Group*.

Ware, Mark and Michael Mabe. 2009. "An Overview of Scientific and Scholarly Journal Publishing." *The STM Report*.

Green, Toby. 2009. "We Need Publishing Standards for Datasets and Data Tables." *OECD Publishing*.

NISO AX Committee. 2004. "Z39.88-2004: The OpenURL Framework for Context-Sensitive Services - The Key/Encoded-Value(KEV) Format - Implementation Guidelines. v.1.7."

OCLC. 2006. "Process for the Evaluation and Approval of New OpenURL Community Profile Submittals by the Z39.88."

[참조사이트]

유비쿼터스. [cited 2011. 2. 8]. 〈http://bit.ly/oHlj9w〉.

해외 학술출판사들, 잇따른 신규 플랫폼 발표. [cited 2011. 3. 29]. 〈http://bit.ly/fglPqA〉

About ALPSP. [cited 2011. 2. 3]. 〈http://bit.ly/eB7yGE〉.

Announcing launch of Release 6.1.2 of RSC Publishing Platform. [cited 2011. 3. 20]. 〈http://bit.ly/gERVuN〉

Data, data everywhere. [cited 2011. 3. 28]. 〈http://econ.st/eNjRcn〉.

Elsevier Announces the "Article of the Future". [cited 2011. 3. 28]. 〈http://bit.ly/dF7jSB〉.

KESLI 설명. [cited 2011. 3. 28]. 〈http://bit.ly/gxfTT0〉.

Scopus 설명. [cited 2011. 3. 28]. 〈http://bit.ly/et7mqy〉.

The data deluge. [cited 2011. 3. 28]. 〈http://econ.st/fIQR7G〉.

The Directory of Open Access Repositories - OpenDOAR. [cited 2011. 3. 21]. 〈http://opendoar.org/〉

What is DataCite. [cited 2011. 3. 28]. 〈http://bit.ly/hqfzMJ〉.

WoS 설명. [cited 2011. 3. 28]. 〈http://bit.ly/fnyxbQ〉.