

# 국회도서관 학술연구자정보

## 공유사업

송미경  
국회도서관 전자정보정책과

### I. 서론

현대사회는 첨단기술의 발달로 정보화사회에서 지능정보사회로 급격히 전환되어 가고 있고, 이에 따른 정보환경도 급격히 변하고 있다. 특히, 지식의 공유 및 커뮤니케이션, 상호연결성이 강조됨에 따라 다양한 분야 간의 협력 및 융합이 필수적이다.

국회도서관에서도 연결과 공유의 시대에 맞아 다양한 변화를 시도하고 있다. 그 일례가 여기서 소개하고자 하는 학술연구자정보 공유사업이다. 이 사업을 통해 국회도서관은 연구자에게 국제표준에 맞는 고유식별번호를 발급하여 개인 식별성을 높이고 각 개인이 저술한 학술정보와 연계하여 클라우드 상에서 공유·활용할 수 있도록 할 예정이다.

또한 국회도서관이 미래창조과학부, 한국정보화진흥원과 함께 진행 중인 「국가학술정보 클라우드사업」(이하 클라우드사업)과 연계하여 연구자를 중심으로 한 학술논문 검색, 빅데이터 분석기법을 이용한 학술데이터 분석 및 시각화 등 다양한 변화를 시도할 계획이다.

국회도서관은 학술연구자정보 공유사업과 클라우드사업을 통해 국가지식정보를 널리 확산하고 국내외 연구자들의 활발한 학술연구활동을 지원하는데 기여할 수 있을 것이라고 생각된다.

현재 학술연구자정보 공유사업은 초기 단계로 연

구용역사업을 진행 중이다. 향후 시스템 설계와 개발을 마치고 사업이 본격적으로 시작되면 다시 한번 더 소개할 수 있는 기회를 가질 수 있기를 기대해 본다.

전문도서관에서도 국회도서관이 추진하고 있는 다양한 공유·협력사업에 관심을 가져주시기를 바라며, 향후 이들 정보를 자관에 활용하여 어떠한 서비스가 가능할 지를 같이 고민하고 공유하는 시간을 가질 수 있기를 바란다.

### II. 사업 소개

#### 1. 개요

학술연구자정보 공유사업은 연구자에게 국제표준에 맞는 고유식별번호를 발급하여 개인의 식별성을 높이고 각 개인이 저술한 학술정보와 연계하여 클라우드 상에서 공유하기 위한 사업이다.

동 사업은 국회도서관이 전문도서관, 대학도서관 등 유관기관과 협력하여 고유식별기호를 매개로 연구자정보를 공유하고 융합하는 기반을 조성하는 것을 목표로 한다. 이를 위해 국회도서관은 자관이 보유한 저자명DB를 국제표준에 맞게 재정비하고 있다.

## 2. 필요성

전자데이터는 검색의 정확성을 향상시켜 이용자의 편의성을 높여주는 유용한 도구이며, 국회도서관은 저자명DB를 구축하고 있다.

국회도서관 외에 국립중앙도서관, 서울대학교, 이화여자대학교, 한국과학기술정보연구원(KISTI) 등에서도 전자전거를 구축하고 있으나 기관마다 목적에 따라 형식과 구축방법이 다르고, KISTI를 제외한 대부분이 단행본 저자를 중심으로 구축되어 있다. 반면, 수시로 발간되는 대량의 학술논문을 대상으로 전자전거를 구축하는 것은 많은 인력과 노력이 필요하며, 어느 한 기관이 모든 데이터를 구축하고 관리하기는 더욱 어렵다.

국회도서관은 국내 최대의 저자명데이터 약 62만 건을 보유하고 있으며, 데이터는 국회도서관 서지데이터와 연계되어 있다. 이들 데이터를 LOD로 발행하여 클라우드사업을 통해 클라우드 상에 공개하여 누구나 학술정보 데이터를 융합하고 활용할 수 있도록 제공함으로써 중복 구축 방지 및 공동 활용을 도모할 수 있다.

또한, ISNI, UCI 등 국제표준식별기호를 연결매개체로 사용함으로써 각 기관이 공유할 수 없는 비공개정보는 포함하지 않고서도 동일저자와 저작물을 식별하여 효과적으로 인터링크할 수 있다.

## 3. 사업내용

### 3.1. 저자명DB 구축

국회도서관은 필명(筆名), 호(號), 한자명(漢子名), 외국어명, 이전 명칭, 약칭 등 다양한 이형(異形)을 가진 저자를 식별하기 위하여 2007년부터 저자명DB(구 검색어참조DB)를 구축하기 시작하였다.

2016년 6월말 기준으로 인명 524,855건과 단체명

95,877건, 총 62만 건의 저자명 데이터가 구축되었고, 정보의 현행화를 위하여 지속적으로 정비하고 있다. 저자명과 연계된 서지데이터(단행본, 학위논문, 학술논문, 세미나 자료)의 양은 약 250만 건이다.

〈표 1〉 국회도서관 전자 구축 현황 (2016. 6. 30. 기준)

구분	인명	단체명	합계
구축 건수	524,855	95,877	620,732

저자명DB는 인명·단체명에 대한 필수정보(4개 항목)와 참고정보(5~9개 항목)에 저작물 서지정보가 연결된 구조를 가진다.

〈표 2〉 저자명DB 데이터 요소

구분	인명	단체명
필수 정보	자관 식별번호, 대표명(성 이름), 생년, 몰년	자관 식별번호, 대표명(full name), 시작년, 폐지년
참고 정보	이형(필명, 호, 한자, 외국어명), 국가, 성별, 직업, 직위, 전공, 소속, 활동분야, 표준식별번호(ISNI, UCI 등)	이형(약칭, 외국어명), 이전기관명, 단체명유형, 연혁, 표준식별번호(ISNI, UCI 등)
저작물	단행본(도서, 비도서), 세미나, 학위논문, 학술논문	

\* 단체명 유형: 입법기관, 사법기관, 중앙행정기관, 연구단체, 대학교 등 총 20개로 구분

### 3.2. 국제표준식별기호 도입

도서관에서는 기관의 필요에 따라 서로 다른 전거 데이터 요소와 구조를 가진 전거데이터를 가지고 있다. 통일된 전거형식이 없는 상황에서 국제표준식별기호의 도입은 전거데이터 간의 매칭처리 및 링크드 데이터 구축을 쉽게 할 수 있는 기반이 된다.

국회도서관은 국제표준이름식별자(ISNI, International Standard Name Identifier) 등록기관이자 디지털콘텐츠식별체계(UCI, Universal Content Identifier) 등록기관이다.

#### 3.2.1. ISNI(International Standard Name Identifier)

ISNI는 연구자, 발명가, 작가, 예술가, 영상제작자, 연주자, 제작자, 출판사 등 창작물과 창작행위에 기여한 수많은 사람을 식별하기 위하여 국제표준화기구(ISO)가 인증한 국제표준번호로, 16자리(15자리 숫자+체크숫자)의 숫자로 구성된다.<sup>1)</sup>

ISNI는 개방형 표준으로 링크드 데이터 및 시맨틱 웹 응용 프로그램의 중요한 구성 요소이자 연계식별자(bridge identifier)이며, 여러 기관이 공개정보만으로 저자전거데이터를 교환 할 수 있다.

국회도서관은 클라우드사업을 통해 학술정보와 저자전거 등을 LOD로 발행할 때 연계식별자의 하나로 ISNI를 도입하였고, 2016년 4월 ISNI 등록기관(ISNI Registration Agency)으로 지정되었다.<sup>2)</sup>

#### 3.2.2. UCI(Universal Content Identifier)

UCI란 URN 기반의 국가표준디지털 콘텐츠 식별체계를 말한다. 인터넷 상에서의 디지털 콘텐츠의 효율적인 유통과 활용을 위해 디지털콘텐츠에 유일하고 영구한 코드를 부여하고 이를 관리해 주는 체계 또는 상이한 식별 체계의 연계표준을 말한다.<sup>3)</sup>

UCI는 개별 콘텐츠를 식별하기 위한 ID를 발급하고 메타데이터(항목의 속성정보)를 관리하며, 디지털 네트워크 상에서 ID의 변환 과정을 통해 다양한 정보를 제공하는 서비스 체계이다.

2012년부터 한국저작권위원회가 UCI 총괄기관으로, UCI에 대한 전반적인 권한과 보급 확산을 책임을 지고 있다.

국회도서관은 2012년부터 UCI 콘텐츠의 등록관리 업무를 수행하는 UCI 등록기관으로 지정되었고, 식별자 연계를 위한 최소한의 메타데이터를 총괄기관과 공유한다. 국회도서관은 국내 최초로 학술연구자에 대한 UCI를 발급하고, ISNI와 같이 연구자와 저작물을 연결하는 요소로 활용할 계획이다.

### 3.3. 국가학술정보 클라우드사업 연계

국가학술정보 클라우드사업은 국회도서관의 학술데이터와 관련 통제어휘(사람, 장소, 기관, 주제 등), 그리고 유관기관의 학술데이터를 클라우드 기반에서 LOD로 발행하고 공유하기 위한 사업이다.

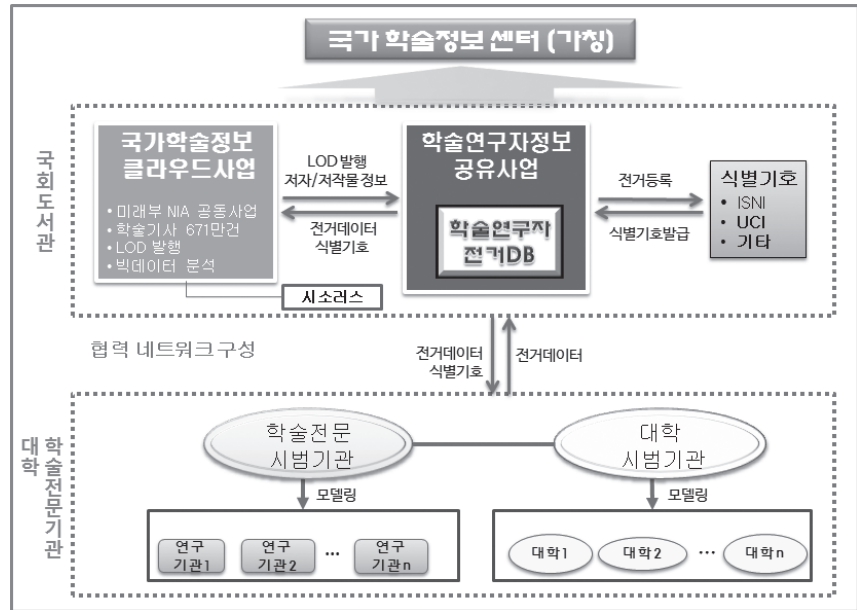
이용자는 클라우드사업을 통해 산재한 학술정보를 클라우드 상에서 공동 활용할 수 있고, 자관에 맞게 데이터를 재사용할 수 있으며, 빅데이터 분석을 통해 연구동향 예측, 의미기반 맞춤형 서비스 등 과학적인 데이터 분석서비스를 제공받을 수 있다.

국회도서관과 미래창조과학부와 한국정보화진흥

1) ISNI <http://www.isni.org/>

2) ISN > ISNI Registration Agencies <http://www.isni.org/content/isni-registration-agencies>

3) UCI <http://www.uci.or.kr/kor/file/main/main.jsp>



〈그림 1〉 학술연구자정보 공유사업 체계

원이 주관하고, 한국연구재단, 한국과학기술정보연구원, 성균관대학교(이하 협력기관)가 시범기관으로 참여하고 있다.

동 사업은 2016년과 2017년, 2단계에 걸쳐 추진한다. 사업 대상 데이터는 학술논문 약 671만건과 국회도서관의 저자명DB 약 62만건, 성균관대가 보유한 학술연구자정보이다.

전 분야에 걸친 학술정보를 공유·활용하기 위해서는 클라우드 환경에서 학술정보를 통합할 필요가 있으며, 데이터를 연결하고 공유하기 위해서는 전거가 필수적이다. 국회도서관에서는 저자명DB와 용어관계사전(시소러스)을 클라우드사업에 활용하여 이용자 친화적 서비스 환경을 구현할 예정이다.

〈표 3〉 국가학술정보 클라우드사업 대상 데이터

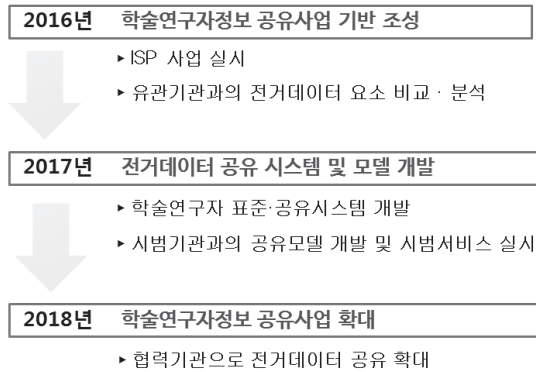
기관명	데이터(건)
국회도서관	학술기사 390만건, 저자명 62만건, 시소러스 18만건
한국연구재단	등재학술지 113만건
한국과학기술정보원	과학학술지 161만건
성균관대학교	연구자 및 학술정보 4만건

### III. 학술연구자정보 관리·공유시스템 구축방안

국회도서관은 2016년 7월부터 「학술연구자정보 관리·공유시스템」 구축을 위한 연구용역을 실시하고 있다. 11월 말까지 연구용역을 통해 국내외의 표준 운영사례를 조사·분석하고, 국제표준에 적합한 학술연구자정보 관리·공유시스템 구축 설계안을 마련할 예정이다. 학술연구자정보의 고유식별번호 발급과 협력기관과의 연계·공유 방안을 구체화할 것이다.

#### IV. 추진 일정

2016년 연구용역 결과를 기반으로 2017년부터는 시스템 개발에 착수하여 본격적으로 ISNI와 UCI 등 국제표준식별번호를 학술논문 저자 중심으로 발급하고 클라우드 기반으로 LOD를 발행하여 서비스 할 예정이다.



〈그림 2〉 학술연구자정보 공유사업 추진 일정

#### V. 결론

국내 학술연구자정보는 국회도서관을 비롯하여 각 기관들이 모두 상이한 식별체계를 가지고 있어 식별자에 대한 호환성이 보장되지 않고, 데이터 융합에도 걸림돌이 되고 있다.

데이터의 연결과 융합이 보편화됨에 따라 데이터를 식별하고 검색을 효율적으로 하기위한 다양한 노력이 시도되고 있으며, 국회도서관에서도 62만 여건의 저자명DB를 ISNI, UCI 등 국제표준식별기호를 통해 동일저자의 고유성을 확보하여 검색의 효율성을 확보하고 이용자들에게 검색의 편의성을 제공할 예정이다.

학술연구자정보 공유사업을 성공적으로 수행하기 위하여 국회도서관은 2015년부터 학술연구자 중심의 저자명DB를 구축하고, 학술논문을 연계하는데 주력

하고 있다.

향후, 동사업의 협력기관을 확대하여 국가학술정보 센터로서 역할을 충실히 수행하고, 국가연구활동 기반을 조성하기 위해 전문도서관과 협력관계도 지속적으로 강화해 나갈 것이다.