

OAI, OpenURL 기반 개방형 학술정보 유통 아키텍처

설계 및 구현

이상기⁰ 최성원* 조성남* 진성진* 최진영** 정택영*

한국과학기술정보연구원 표준화기술지원실, 고려대학교 정보통신대학 컴퓨터학과
(sklee swchoi sncho sjibun tvchung}@kisti.re.kr choi@formal.korea.ac.kr

Design and Implementation of an Open Academic Information Dissemination

Architecture based on OAI and OpenURL

Sanggi Lee⁰, Seongwon Choi, Sunnam Cho, Sungjin Jhun, Taikyeong Chung

Technical Support for Standardization Dept., Korea Institute of Science and Technology Information

Jinyoung Choi

Dept. of Computer Science & Engineering, Korea University

요 약

상업출판사에 의한 학술정보의 무기화 및 독점화로 학술정보 유통이 일대 위기를 맞고 있는 가운데 누구나 자유롭게 학술정보를 이용하자는 운동이 활발하다. 이를 기술적 표준적으로 지원하기 위해 결성된 OAI에서는 OAI 아키텍처와 표준 프로토콜을 개발하였다. OAI는 프로토콜 표준화를 통해 메타데이터의 상호운용성은 크게 증진하였으나, 메타데이터와 전자원문 간 연계를 간과함으로써 정보자원 간 연계가 취약한 맹점이 있다. 본 논문에서는 OAI 아키텍처의 메타데이터와 전자원문 간 연계성을 개선하고 유사 자원으로서의 확장성을 제공하기 위해, 국제표준 프레임워크인 OpenURL을 접목한 개방형 학술정보 유통 아키텍처를 설계 및 구현하였다. 이를 통해 분산 환경에서의 정보자원 간 상호운용성 및 유사자원으로서의 확장성을 제고하였다.

1. 서론

상업출판사에 의한 학술정보의 무기화 및 독점화로 학술정보 유통이 일대 위기를 맞고 있는 가운데 누구나 자유롭게 학술정보를 이용하자는 운동(Open Access)이 활발하며, 이를 기술적 표준적으로 지원하기 위해 OAI(Open Archives Initiative)[1]가 결성되었다.

OAI에서는 디지털콘텐츠의 효율적인 유통 및 활용을 목적으로 상호운용성을 갖는 표준을 개발하고 보급하는 역할을 수행하고 있으며, 학술정보의 자유로운 유통 및 상호운용성을 제고하기 위해 OAI 아키텍처와 OAI 프로토콜(OAI-PMH[2])을 개발하였다.

분산된 디지털 콘텐츠의 상호운용성을 제고하기 위해서는 궁극적으로 정보자원을 어떻게 표현할 것인가와 표현된 자원을 어떻게 연계할 것인가가 중요하다. OAI는 전자의 경우 비교적 충실하나 후자를 간과한 나머지 정보자원 간 연계가 취약한 실정이다. 규정된 스펙이 없었기 때문에 메타데이터와 전자원문 간 연계가 제각각이며 표준화되지 못하고 있다.

본 연구에서는 메타데이터뿐만 아니라 전자원문까지 상호운용성을 제고하고 유사 자원으로서의 확장 기반을 확보하기 위해 OAI 아키텍처에 정보자원 간 연계 프레임워크인 OpenURL[3]을 접목한 개방형 학술정보 유통 아키텍처를 설계 및 구현하였다.

2. 관련 연구

2.1 OAI 아키텍처

OAI 아키텍처는 그림 1과 같이 3개 층으로 구성된다. 3개 층은 ①이용자가 정보를 검색하고 활용하는 영역인 사용자 영역(User Application), ②실제 콘텐츠를 생산·저장·관리하고 서비스제공자의 요청에 따라 메타데이터를 표준적으로 노출(제공)하는 데이터제공자(Data Provider, DP) 영역, ③데이터제공자로부터 메타데이터를 주기적으로 수집하여 부가적인 서비스를 하는 서비스제공자(Service Provider, SP) 영역이다.

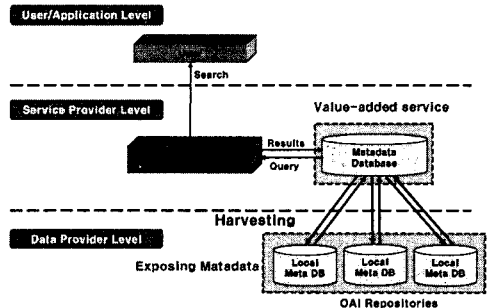


그림 1 OAI 아키텍처

정보선진국에서는 기관에서 생산된 지식정보를 상호공유하기 위하여 OAI 아키텍처를 기반으로 기관 저장소(Institutional Repository)를 개발·보급하고 있으며, 영국의 ePrints[4]와 미

국의 Dspace[5], 한국교육학술정보원의 Dcollection이 대표적이다. 또한, OAI 핵심 사항은 디지털라이브러리 등 관련 분야로 확산되어 오픈디지털라이브러리인 ODL(Open Digital Library)[6] 등에 응용되고 있다.

OAI 아키텍처의 핵심이랄 수 있는 OAI 프로토콜은 확장성이 뛰어나고 간결(simple)하며 국제표준인 XML, Dublin Core(DC), HTTP에 기반하고 있다. 그림 2는 서비스제공자가 OAI 프로토콜을 사용하여 데이터제공자로부터 메타데이터를 수집하는 과정을 도식화한 것이다. 서비스제공자는 OAI 프로토콜을 사용하여 데이터제공자로부터 주기적으로 메타데이터를 수집한 후 이를 저장·색인하며 부가적인(통합검색 등) 서비스를 제공한다[7].

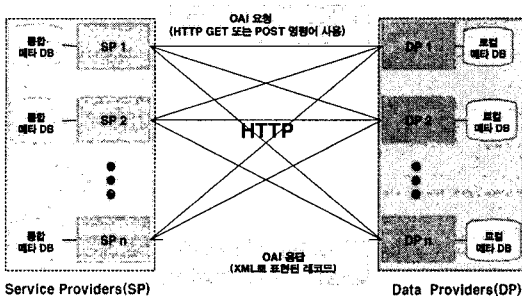


그림 2 OAI 프로토콜 모델

OAI 프로토콜은 메타데이터의 상호운영과 관련된 최소한(6가지)의 스펙만을 규정하고 있으며, 최근버전은 OAI-PMH v2.0이다(2002. 6). OAI-PMH v2.0과 OAI-PMH v1.0의 가장 큰 차이점은 v2.0부터 XML 스키마를 지원한다는 점이다. 표 1은 OAI 프로토콜의 6가지 기능 요약이다.

표 1 OAI 프로토콜(6가지)

| 요청 | 기능 설명 |
|----------------------|----------------------|
| GetRecord | 저장소로부터 하나의 레코드를 수집 |
| ListRecords | 한꺼번에 대량의 레코드를 수집 |
| ListIdentifiers | 저장소에서 수집 가능한 식별자를 검색 |
| ListMetadata Formats | 메타데이터 포맷을 검색 |
| Identify | 저장소의 기본정보를 검색 |
| ListSets | 저장소의 셋(set)구조를 검색 |

2.2 OpenURL

OpenURL은 메타데이터를 이용하여 전자정보 서비스 기관이 제공하는 정보자원으로 연계하는 일종의 해석기(resolver)이며, 이것은 기존 경직되고 단일적인 전자정보 연계 방식의 문제점을 해결하기 위해 벨기에 Ghent 대학의 Herbert Van De Sompel과 동료 연구자들에 의해 개발된 NISO 국제표준 정보 자원 연계 프레임워크이다[8].

그림 3은 일반적인 링크로 특정 자원에 대한 단일 링크만을 제공하며 유사정보 등 다른 자원으로의 확장이 어렵다.



그림 3 일반적인 링크 개념도

반면 그림 4는 OpenURL 프레임워크로 소스에서 메타데이터를 전달받아 이를 해석한 후 복수의 정보자원 링크를 산출 제공한다. 또한 전달하는 메타데이터에 다양한 식별자를 포함하여 다른 식별체계와 연동할 수 있다. OpenURL 구분구조는 BASE-URL과 Query라는 두 부분으로 구성된다.

<구분구조>

OpenURL ::= BASE-URL '?' Query

BASE-URL은 OpenURL Query를 입력받을 수 있는 OpenURL 서버의 위치정보 즉, URL을 의미하며, Query는 메타데이터를 OpenURL 구분구조로 기술하는 부분이다.

<예시>

<http://openurl.co.kr/link?isbn=123456789&genre=ebook>

<http://openurl.co.kr/link?id=doi:10.1000/6789-1>

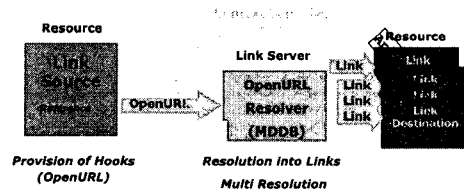


그림 4 OpenURL 프레임워크

OpenURL은 외국의 경우 SFX, UKLON, CrossRef에서 정보 자원 연계에 사용 중이며, 국내는 한국과학기술정보연구원과 (주)KINS에서 학술정보 연계에 활용하고 있다.

3. 시스템 설계 및 구현

3.1 시스템 구성

그림 5는 한국과학기술정보연구원에서 개발한 개방형 학술정보 유통 시스템 구성도이다. 데이터제공자는 메타데이터와 전자정보문을 생산·저장·관리하며, 서비스제공자의 요청이 있을 때 표준적인 방식으로 정보를 제공한다. 아울러 서비스제공자에게 자신의 존재를 알리기 위해 OAI 레지스터리[9]에 등록한다. 그러면 서비스제공자는 OAI 레지스터리에 등록된 기관(269개, 2005. 4 기준)을 대상으로 OAI 프로토콜을 사용하여 메타데이터를 주기적으로 수집·저장·색인한다. 참고로 OAI 레지스터리에 등록된 자료는 누구나 자유롭게 이용할 수 있는 무료 정보(Open Access)이다.

이용자는 서비스제공자의 사용자 영역(User Application)에서 메타데이터를 검색한 후 전자정보문을 요청하면, OpenURL 책버

는 관련된 정보자원(전자원문)으로의 링크를 제공한다.

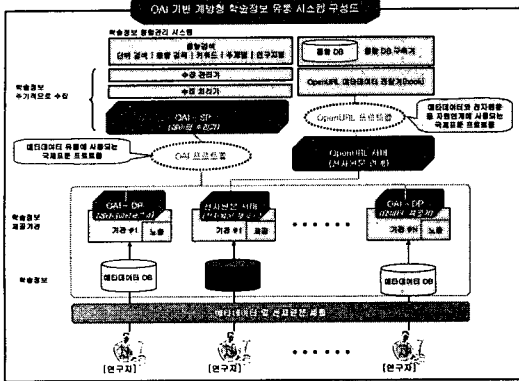


그림 5 시스템 구성도

하는 타겟정보원(데이터제공자)으로부터 전자원문을 이용하게 된다.

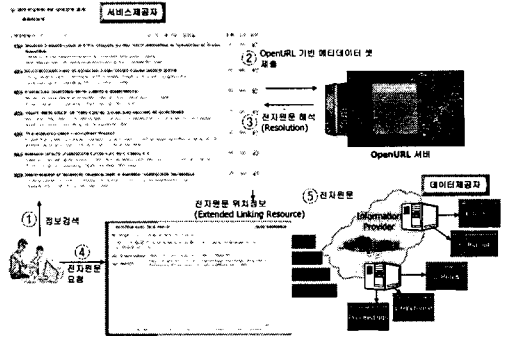


그림 7 OpenURL 기반 정보자원 연계

3.2 메타데이터 수집 프로세스

그림 6은 서비스제공자가 데이터제공자로부터 메타데이터를 수집하는 과정을 도식화한 것이다.

수집기(harvester)는 처음에는 데이터제공자의 모든 메타데이터를 일괄 수집하지만 이후부터는 추가되거나 변경된 것만을 선별 수집한다. 데이터제공자는 서비스제공자에게 XML과 Dublin Core(DC) 형태로 정보를 제공하며, 서비스제공자는 데이터제공자로부터 표준적인 방식으로 메타데이터를 수집하여 통합 검색, 전자원문 링크 등 부가적인 서비스를 제공함으로써 상호운용성이 확보되는 방식이다.

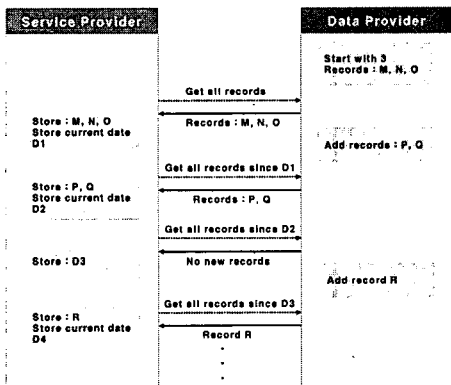


그림 6 메타데이터 수집 프로세스

4. 결론

개방형 학술정보 유통 시스템이 증가일로에 있으며 이들 간 상호운용성을 제고하기 위해 OAI에서는 표준 아키텍처와 프로토콜을 제정·보급하고 있으나, 정보자원 간 연계가 매우 취약하였다. 이를 개선하기 위해 본 연구에서는 OAI 아키텍처에 OpenURL을 접목함으로써 다음과 같은 성과를 얻었다.

첫째, 정보자원 간 연계를 표준화함으로써 분산된 학술정보가 메타데이터뿐만 아니라 전자원문까지 표준적으로 유통되도록 하였다. 둘째, DOI 등 다양한 식별체계와 통합운영(multi resolution)이 가능함으로써 레거시 시스템의 재사용성을 제고하였다. 셋째, OpenURL의 특징인 일대다(1:M) 연계 기능을 활용하여 유사자원(e-book, 원문복사신청, 인용색인 DB 등)으로 확장을 용이하게 하였다.

향후 OAI 표준화위원회의 공개포럼에 표준초안(working draft) 제출을 통한 국제표준으로 자리매김하기 위한 노력과 심층연구가 필요하다.

참고문헌

- [1] <http://www.openarchives.org/index.html>
- [2] <http://www.openarchives.org/OAI/openarchivesprotocol.html>
- [3] http://www.exlibrisgroup.com/sfx_openurl_syntax.htm
- [4] <http://www.eprints.org/>
- [5] <https://dspace.mit.edu/index.jsp>
- [6] 이상기, "OAI 기반 Open Digital Library 연구", 정보관리연구, 제35권, 제3호, 2004.
- [7] 이길신, "국가 학술연구 DB 생성 및 유통체계 1차 Workshop:OAI 프로토콜 소개", 한국교육학술정보원, 2003
- [8] 안계성 "정보자원 연계를 위한 OpenURL", URI 표준화포럼 및 워크숍, 2003.
- [9] <http://www.openarchives.org/Register/BrowseSites>

3.3 OpenURL 기반 학술정보 연계

그림 7은 OpenURL 프레임워크를 사용하여 학술정보를 연계한 실제 화면으로 ①이용자가 메타데이터를 검색한 후 ②OpenURL 메타데이터 제출기(hook)를 통해 메타데이터 셋을 OpenURL 구분구조에 맞게 작성하여 OpenURL 서버로 넘겨주면 ③OpenURL 서버는 제출받은 메타데이터를 분석하여 이용자 상황에 맞는 복수의 정보자원(타겟 정보원)을 산출 제공한다. 그러면 ④이용자는 산출된 정보자원 중 자신에게 가장 적합한 정보자원(전자원문 등)에 접근하여 ⑤ 전자원문을 서비스