

OAI-PMH 기반의 학습·학술 정보 공유시스템의 설계 연구

김영직*

*고려대학교 컴퓨터정보통신대학원
e-mail:capbang@korea.ac.kr

A Study on Design of a Learning and Scholarly Information Sharing System based on OAI-PMH

Young-Jik Kim*

*Graduate School of Computer and Information Technology, Korea University

요 약

IEEE LOM 기반의 ADL SCORM, KEM 등을 통해 생산된 학습 콘텐츠는 데이터의 표준을 준수하여 이종의 학습시스템에서도 연동 될 수 있도록 상호운용성이 크게 향상되었고, 이를 통해 학습 콘텐츠의 유통과 검색, 접근이 매우 수월하게 개선되었다. OAI-PMH는 XML 기반의 학술 정보에 대한 메타데이터 통합을 통하여 개방형 검색이 가능하도록 시스템을 구현하기 위한 프레임워크로써 학술 정보의 통합 유통 체제에 매우 활용도가 높다. 본 논문에서는 OAI-PMH의 SP/DP 개념의 메타데이터 수확(Harvesting) 기술을 활용하여 국내의 교육기관(학교), 연구 기관, 학술단체, 기업연수원 등에서 보유하고 있는 SCORM, KEM 기반의 학습 콘텐츠와 DC(Dublin Core) 기반의 학술 정보를 통합하여 사용자에게 정보 서비스를 제공하기 위한 개방형 학습·학술 정보 공유 시스템 설계를 제안하고자 한다.

1. 서론

최근 컴퓨터의 데이터 처리 속도의 향상과 대용량 데이터베이스 저장 기술, 인터넷과 네트워크를 이용한 데이터 전송 기술과 전송 속도의 향상은 이미 엄청난 량의 정보를 쏟아내고 있다. 누군가가 만들어 놓은 학습 콘텐츠 또는 학술 정보를 통합하여 쉽게 검색하고 활용할 수 있으면 얼마나 좋을까? 분산되어 있는 교육 자료를 통합하고자 할 때 먼저 고려되어야 할 사항이 바로 자료의 메타데이터(Metadata)이다[1]. 교육 자료에 대한 메타데이터 표준을 위한 연구들이 수행되었고, 주로 표준화를 통한 자원의 효율적 발견과 메타데이터간의 상호 호환성을 확보함을 목적으로 하였다. XML기반의 대표적인 메타데이터 표준으로는 DC(Dublin Core)와 IMS, IEEE LOM(Learning Object Metadata), ADL SCORM(Sharable Content Object Reference Model)이 대표적이다. 국내에는 2004년 KS 표준을 획득한 초·중등교육정보 메타데이터인 KEM(Korea Educational Metadata)이 있다. 각기 독립된 공간

에 구축된 학습 자료들의 통합을 통한 자료 검색은 사용자로 하여금 필요한 정보를 쉽게 찾고, 접근하고, 유통 시킬 수 있는 부가가치가 높은 서비스를 지향하게 되었다. 그같은 학습 콘텐츠 저장소(Repository)들간의 통합 활동으로 개방형 학술 정보의 통합 유통을 지향하는 OAI(Object Archives Initiative)가 있다. 원래 OAI는 디지털도서관의 학술논문, 연구보고서등의 학술정보의 통합 유통을 위한 개발을 목적으로 하였지만, 자료의 통합 프로세스와 아키텍처의 효율성으로 인해 최근에는 LOM 기반과 OAI-PMH(OAI Protocol for Metadata Harvesting)을 활용한 학습 콘텐츠들간의 통합 기술에도 응용되어지고 있다. 본 논문에서는 OAI-PMH의 SP/DP 개념의 메타데이터 수확(Harvesting) 기술을 활용하여 국내의 교육기관(학교), 연구 기관, 학술단체, 기업연수원 등에서 보유하고 있는 개방형 학습 콘텐츠와 학술 정보를 통합하여 사용자에게 유용하고 양질의 학습 정보를 제공하기 위한 개방형 학습·학술 정보 공유 시스템 설계를 제안하고자 한다.

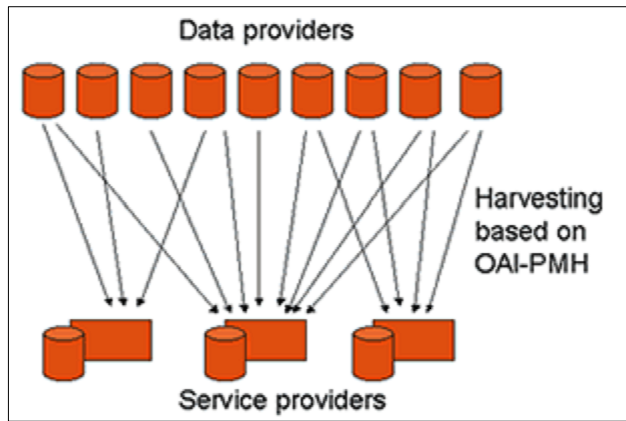
2. OAI-PMH 의 이론적 배경

2.1 등장배경

1999년 Santa Fe Convention 은 OAI-PMH(Open Archives Initiative - Protocol of Metadata Harvesting)의 초기 버전이다. 주요 검색 대상은 학술 문헌이었고, RePEc/SODA Service/Data provider 모델 구조를 갖고, Dienst 프로토콜이라고 지칭했다. OAI-PMH 1.0 은 메타데이터의 상호운용성을 위해 Dublin Core Element Set을 채택하였고, OAI-PMH 1.1 에서는 1.0 스펙에 XML Schema 스펙을 추가하였다. OAI-PMH 1.0 과 OAI-PMH 1.1 은 아직까지는 실험적 개발 수준에 머물러 있었다. 이후 1.x 구현자들의 심도 있는 토의와 alpha-tester 들의 테스트 결과가 반영되어 프로토콜에 매우 많은 수정이 이뤄졌고, 그 결과 안정적인 프로토콜인 OAI-PMH 2.0 이 탄생하게 되었다.

2.2 모델

SP(Service Provider)는 DP로부터 수집한 메타데이터를 통하여 검색과 원문 보기 등의 사용자 서비스를 제공한다. DP(Data Provider)는 각각의 저장소(Repository)를 의미하며 원문 자료와 메타데이터를 저장하고 있다.



<그림 1> SP/DP 메타데이터 수집 모델[2]

2.3 적용 사례

▪ RLLOMAP

영국 JISC는 Resource Discovery Network(RDN) 과 Learning and Teaching Support Network(LTSN) 의 공유 시스템 구축에 관한 연구 사업을 추진하고, 메타데이터의 통합을 위해 OAI의 OAI-PMH 를 이용하여 레코드들을 공유할 수 있게 하였다. IEEE LOM과 UK LOM, DC(Dublin Core) 메타데이터를

통합한 RLLOMAP(RDN/LTSN LOM Application Profile) 를 정의하였다[3]

▪ dCollection

2003년 국내의 대학기관과 교육유관기관이 참여한 시범 사업으로, 국가적 차원의 학술연구정보의 생성 및 유통 체계를 시범적으로 구축하는 것을 목표로 삼고 있다. 이후 2004년부터는 기능개선과 참여기관의 확대 보급을 위한 사업을 진행하고 있다[4]. '생성유통 체계'(일명 dCollection)로 명명된 사례 모델은 OAI에서 제시하는 전형적인 SP/DP 모델을 준수하고 있다.

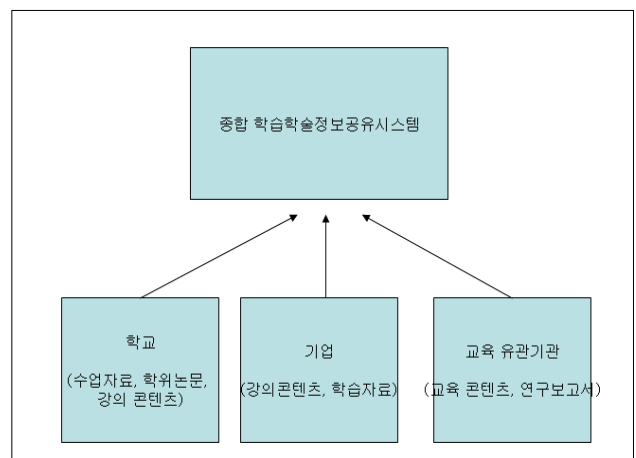
▪ stOAI

stOAI는 OAI 프로토콜을 이용한 과학기술유통시스템으로 수집(Harvester) 시스템과 저장(Repository) 시스템으로 구성된다. 이 시스템에서 OAI 프로토콜을 이용하여 수집, 저장된 콘텐츠는 과학기술정보 포털 서비스(<http://www.yeskisti.net>) 논문검색의 OA(Open Access) 메뉴에서 제공된다.[5]

3. OAI-PMH 기반의 학습·학술 정보 공유 시스템 설계

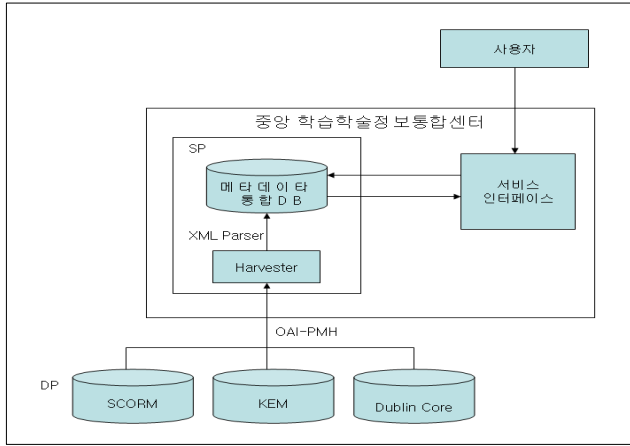
3.1 목표 시스템

OAI 프로토콜의 Data Provider 와 Service Provider 의 시스템 아키텍처를 활용하여 학교 및 유관 기관에서 보유하고 있는 학습 콘텐츠와 대학, 연구기관 등에서 보유하고 있는 학술 정보의 통합 메타데이터를 설계하고 이에 대한 통합 DB를 설계한다. 또, 주기적인 OAI-PMH 수집(harvesting)을 통한 통합 메타데이터를 수정, 갱신, 삭제하고 사용자에게 검색 서비스를 제공할 수 있는 학습·학술 공유 시스템 설계를 목표로 한다.



<그림 2> 학습·학술정보공유시스템 개념

3.2 시스템 아키텍처



<그림 3> SP/DP 시스템 구조도

OAI-PMH의 DP/SP 모델을 기반으로 다음과 같은 시스템 아키텍처를 갖는다. 우선 SCORM, KEM, Dublin Core를 메타DB로 갖는 가상의 DP가 있고 이들의 메타데이터를 수집(harveting)하는 SP가 있다. 각각의 DP와 SP는 HTTP 프로토콜을 통하여 통신하며, 데이터의 요청과 응답시 GET, POST 방식을 사용한다.

3.3 통합 메타데이터 기본요소 설계

각 DP로부터 전송받은 메타데이터를 SP에 일관성 있게 저장하고 검색하기 위한 통합 메타데이터를 설계한다. 통합 메타데이터는 SCORM, KEM, Dublin Core 메타데이터 엘리먼트(element)를 통합, 단순화하여 통합 검색이 가능하도록 함을 목적으로 한다. 통합 메타데이터의 기본 요소는 검색에 필요한 항목과 디스플레이 항목을 정의한 후 DP에서 제공하는 SCORM, KEM, Dublin Core 메타데이터로부터 공통된 엘리먼트를 추출하여 정의하였다.

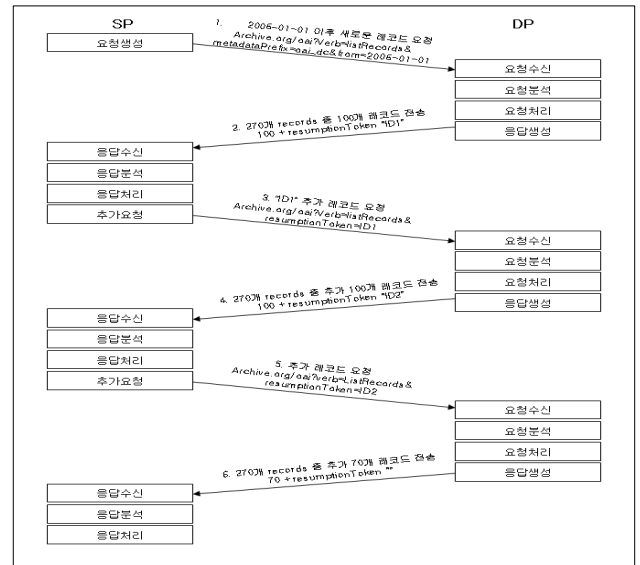
번호	요소	설명
1	identifier	자원의 유일한 식별 기호
2	title	자원의 제목
3	Description	자원에 대한 내용 설명
4	keyword	자원을 설명하는 키워드
5	author	자원의 제작자 또는 소유자
6	format	자원의 기술적인 데이터 형식
7	size	자원의 디지털 크기
8	location	자원을 접근하기 위한 위치
9	resource type	학습 자원의 특정 유형

<표 1> 통합 메타데이터 기본 요소

3.4 SP/DP 간 데이터 처리 프로세스

데이터 제공자로부터의 처음 자료 제공은 모든 데이터를 일괄 수집하는 형태이지만, 이후 부터는 신

규, 수정, 삭제된 정보만을 선택적으로 수집한다. SP의 자료 수집기는 DP의 전송 자료 형태(metadataPrefix)를 분석하여 통합 메타 DB에 저장된 엘리먼트 값을 추출한다.



<그림 4> SP / DP 간 요청 처리 프로세스

4. 결론

본 논문에서는 인터넷과 네트워크를 기반으로 각 분야의 학습 콘텐츠와 학술 정보의 통합을 위해 OAI-PMH 기반의 학술·학습 정보 공유 시스템의 설계에 관한 연구를 수행하였다. 이러한 공유시스템에서는 통합 메타데이터 요소 정의와 그에 대한 검증이 매우 중요하겠으며, 메타데이터의 수집 과정에서 XML 파싱을 통한 통합 메타데이터 요소 추출을 위한 XML 바인딩 기술이 매우 중요하게 생각된다. 향후 방대한 분야의 지식 정보의 공유를 위한 범 세계적인 지식 정보 공유 시스템의 설계 연구가 수행되기를 희망한다.

참고문헌

- [1] 최현중, 황성욱, 김태영, "XML 웹서비스와 JDBC를 이용한 분산 메타데이터 검색 시스템의 설계 및 구현", 2004
- [2] OAI, <http://www.openarchives.org/>
- [3] Andy Powell and Phil Barker, "RDN/LTSN Partnerships: Learning resource discovery based on the LOM and the OAI-PMH", 2004
- [4] KERIS, "생성유통체계", dCollection(국가지식정보생성 및 유통체계)시스템
- [5] 윤준원, 이용식, 이상기, 신기정, "과학기술정보 유통을 위한 OAI기반 지식정보 수집·저장시스템 구축", 2006