

6장. 요소기술

부산대학교 문헌정보학과

이수상 교수

sslee@pusan.ac.kr



■ 요소기술

- 컴포넌트(component)
- 사전적 의미의 컴포넌트는 '구성요소'이다. 이것은 시스템이나 기기를 구성하는 요소 즉, 다른 요소와 구분이 되며, 재활용이 가능하며, 경우에 따라서는 다른 것으로 교환이 가능하다.

■ 디지털도서관의 요소기술

- 디지털도서관 시스템 구성요소
- 디지털도서관 시스템에서 특정 기능을 담당하는 모듈(프로그래밍 블록)이나 도구(어플리케이션)
 - ✓시스템 내부의 요소로 구현
 - ✓기 구현된 것을 도입하여 사용
 - ✓외부에 존재하는 것을 네트워크로 연계하여 사용
- 디지털도서관 시스템을 구성하는 요소기술은 다양한 영역에서 단위기술로 구현될 수 있다.

■ 디지털도서관의 식별자

- 디지털도서관 영역에서 식별자는 디지털 콘텐츠의 효율적인 식별, 접근, 관리를 위한 식별기호, 식별주소 또는 식별코드
- 인터넷 환경에서의 디지털 콘텐츠의 식별자: URI(Uniform Resource Identifier)라는 영역에서 다양한 식별자 형태가 만들어지고 사용
- URI
 - ✓ 통일된(uniform) 콘텐츠 자원(resource)에 대한 식별자(identifier)라는 의미
 - ✓ 인터넷으로 유통되는 콘텐츠 자원을 나타내는 유일한 주소
 - ✓ URL(Uniform Resource Locator)과 URN(Uniform Resource Name)으로 구분

■ URL

- URL은 인터넷 상에서 자원에 접근하는 실제적인 네트워크 경로 즉 주소 (address)
- 주소의 형식: <프로토콜/서버주소/디렉토리/파일>
 - ✓ 프로토콜은 대부분 'http'이며, 경우에 따라서는 ftp, gopher, telnet 등을 사용하기도 한다.
 - ✓ 서버주소는 'pulip.pusan.ac.kr'과 같은 DNS(Domain Name Server) 이거나 '164.125.100.101'과 같은 IP주소로 표현된다.
 - ✓ 디렉토리/파일 부분은 'mylibrary/index.html'와 같이 특정한 파일이 위치한 디렉토리나 파일의 이름을 나타낸다.

■ URN

- URN은 URL보다 영구적인(persistent) 주소체계
- URN은 자원 그 자체에 부여된 독립적인 이름 → URN은 자원의 위치와 상관없이 자원에 영구적인 고유한 식별기호로 접근을 가능하게 해준다.
- URN의 형식: <urn:NID:NSS>
 - ✓ NID(Namespace Identifier)는 이름공간 식별자
 - ✓ NSS(Namespace Specific String)는 NID 내의 고유한 값
- 디지털도서관 영역에서 가장 대표적인 URN은 DOI, ISBN 등
 - ✓ 이를 처리하기 위한 요소기술이 요구된다.
 - ✓ 사례: '978-89-92885-01-0'의 ISBN 번호는 <urn:isbn:978-89-92885-01-0> 형식의 URN

■ DOI(Digital Object Identifier)

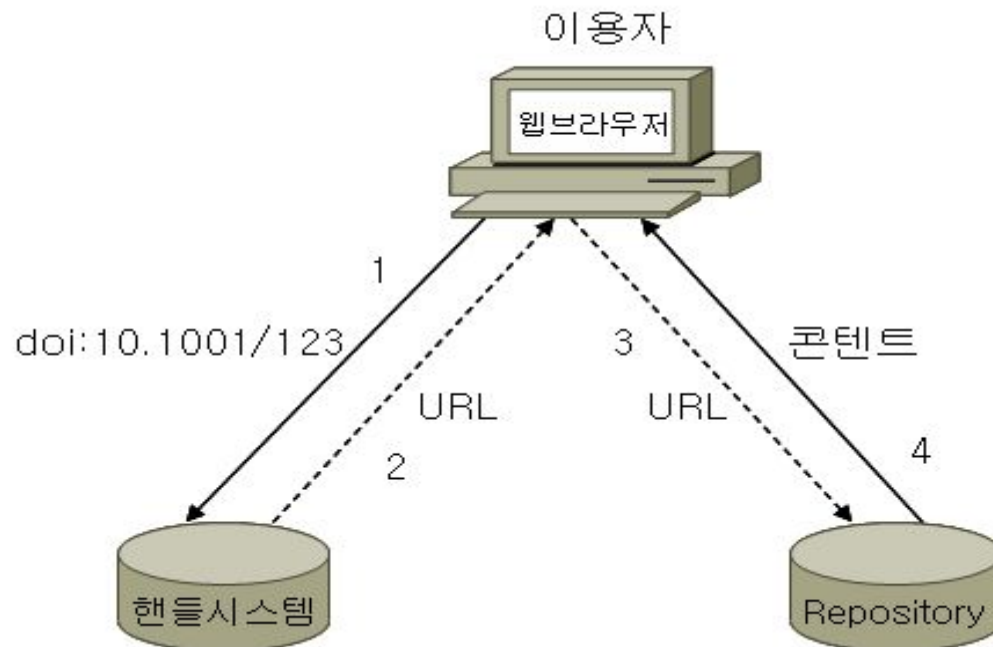
- 콘텐츠에 지속적으로 존재하는 식별자를 지칭하며, 서비스를 유통하기 위한 인터넷 상의 식별자를 처리하는 시스템이다.
- 콘텐츠를 직접 표시하는 것이 가능하므로 위치와 독립적으로 디지털 콘텐츠를 관리하는 구조가 된다.

- DOI의 전형적인 형태

prefix / suffix
10.1001 / 123456
디렉토리.등록기관번호/[코드]항목ID

- DOI 접두어(prefix)는 다시 두 부분으로 나뉜다. DOI를 관리하고 접두어의 할당을 담당하는 디렉토리 관리자 코드와 디렉토리 관리자가 디지털 객체를 보유하고 있는 기관에 부여한 등록기관 코드이다.
- DOI 접미어(suffix)는 접두어를 부여받은 기관(등록자)이 개별 자원에 대해 부여하는 식별기호이다.

■ DOI 변환기



**** 참조: KISTI의 DOI센터:** <http://doicenter.ndsl.kr/>

- KoreaScience <http://koreascience.kr>

■ URL의 사용의 문제점과 한계성

- URL은 **위치 의존적(location-dependent)**이어서 네트워크상의 자원의 위치가 변경되면 해당 자원을 지시하는 모든 URL링크는 자원의 새로운 위치를 반영하기 위해 갱신되어야만 한다. 그렇지 않으면 네트워크상의 '**절단링크(broken links)**'가 발생하게 된다.
- **자원의 식별과 위치연결을 동일시한다.** URL은 정보자원 자체를 식별하지 못하고 자원의 위치정보만을 식별할 뿐이다. URL은 분산 웹 환경에서 동일한 자원에 대한 여러 위치에 존재할 수 있는 복수 표현물 (incarnation)들 중 하나의 위치만을 지시할 수 있을 뿐이다. 즉, 동일한 자원에 대한 표현물이 네트워크 상에서 여러 위치에 존재한다면 각각의 표현물을 지시하기 위해서는 각기 개별적인 URL 링크가 필요하다.
- URL 링크에서는 **자원에 접근하고자 하는 이용자의 상황(User's Context)**이 전혀 고려되지 않는다. 즉, URL 링크가 자원의 표현물에 대한 위치정보를 정확하게 지시하고 있더라도 이용자가 해당 자원에 대한 접근 권한을 가지고 있지 않다면 이러한 URL 링크는 이용자에게는 '**데드링크(dead-links)**'가 된다.

■ 적합자원 (appropriate copy) 문제

- 특정 기관에 소속됨으로 해서 이용자에게 부여되는 접근권한 내에 있는 정보자원의 식별, 자원에 대한 기관의 소유상태, 정보서비스(문헌전달 서비스 등)에 대한 기관의 선호도 등 이용자 정보 환경을 고려하여 이용자에게 적합한 링크만을 선별 제공할 수 있는 것을 말한다.
- 이용자 정보환경을 고려하는 "상황인식연계(context-sensitive linking)"
- '상황인식연계'를 구현하기 위하여 사용되는 기술
 - ✓ URL의 문제점 및 한계성들을 해결하고자 하는 방향으로 발전하고 있다.
 - ✓ OpenURL 프레임워크는 참조대상이 되는 연계자원에 대한 메타데이터와 식별자를 기반으로 하여 '상황인식(context-sensitive)연계'를 지원하는 아키텍처이다.
 - ✓ OpenURL 프레임워크는 링크 생성자가 주기적으로 링크의 연결여부를 관리해야만 하는 부담을 줄임과 더불어 안정적인 자원 연결 서비스를 지원한다.

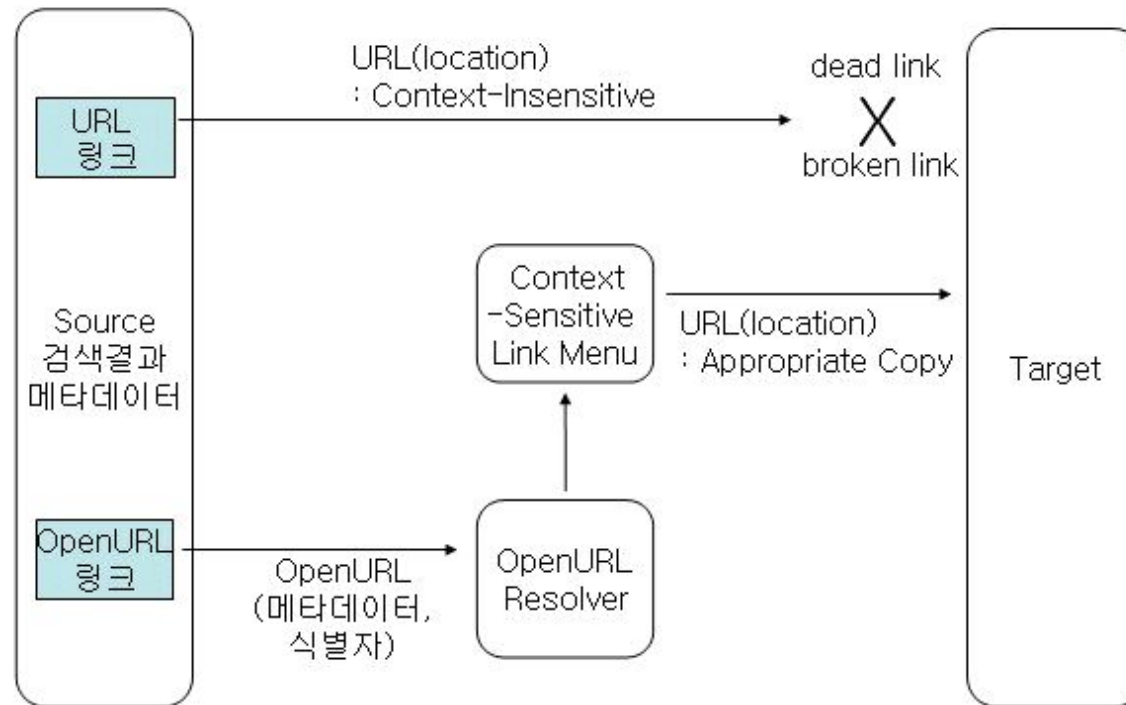
■ OpenURL의 초기 개념

- 겐트(Ghent)대학의 줌펠(Herbert Van de Sompel)과 호츠스텐바(Patrick Hochstenbach), 그리고 ExLibris 사의 베이트-아리(Oren Beit-Arie)에 의해서 구상되었다.
- 이들은 '상황인식연계'를 수행하는 SFX라고 명명된 OpenURL 변환 시스템을 개발하였다. 그 후 몇몇 실험적 연구 프로젝트와 시스템 개발사들이 연구와 상업적 제품개발에 OpenURL 명세서(OpenURL Specification)를 활용하였다.
- 2001년에 NISO(National Information Standards Organization)는 OpenURL 구문에 대한 미국 표준을 개발하기 위한 표준화 위원회(Standard Committee AX)를 발족하였으며, 이 위원회에서는 SFX를 위해서 개발된 OpenURL 명세서를 표준화 작업의 토대로 하여 그 발전적 방향을 모색하고 있다.

■ OpenURL 프레임워크

- OpenURL 프레임워크의 기본 개념은 "연결대상이 되는 정보자원은 특정 메타데이터 집합(예: issn, title, year, volume, page)이나 아니면 하나 이상의 고유한 식별자(예: DOI)로 식별가능하다"는 것이다.
- 그러므로 OpenURL은 위치정보가 아닌 정보자원 자체의 식별을 위해서 필요한 메타데이터 집합이나 고유한 식별자를 통해서 정보자원을 연결한다. 기존의 URL과는 달리, OpenURL은 연결 대상 자원의 고정된 위치 정보를 지시하지 않는다.
- OpenURL은 다음과 같은 내용을 포함한다.
 - ① OpenURL을 입력값으로 받는 OpenURL 변환기의 주소
 - ② 이용자 소속을 식별하는 메타데이터
 - ③ 연결대상 자원을 식별하기 위한 정보

■ OpenURL을 통한 자원접근 개념도



**** 참조: ExLibris의 SFX:** <http://www.exlibrisgroup.com/category/SFXOverview>

KISTI의 링크센터(CLICK) <http://koreascience.kr>

■ OpenURL 구성요소: 소스 정보원, OpenURL 서버, 타겟 정보원

- **소스 정보원**은 이용자가 정보를 획득한 최초의 전자정보 서비스를 총칭한다. 예를 들어 전자도서관 시스템에서 이용자가 도서관 소장 자료를 검색하고자 했고 검색결과에서 외부의 전자정보 서비스 제공자가 보유한 학술지 정보로 접근한다면 도서관 정보 시스템이 OpenURL프레임워크에서 소스 정보원이 된다.
- **OpenURL 서버**는 소스 정보원으로부터 전달받은 OpenURL을 해석하고, 입력된 OpenURL에 적합한 전자정보 목록을 산출한다. OpenURL의 해석은 일정한 구문으로 입력되는 메타데이터를 분석하여 적합한 전자정보를 선별하고 최종적으로 타겟 정보원(사이트)로 링크하기 위한 OpenURL을 생성한다.
- **타겟 정보원**은 OpenURL에 의해 최종적으로 이용자가 접근할 수 있는 전자 정보서비스 제공자의 웹 사이트를 의미한다. 이러한 타겟 정보원의 서비스 유형에는 전자저널 서비스, 웹학술 데이터베이스, 인용색인 정보, 인터넷 서점, 도서관 서지정보, 원문복사 서비스 기관 등이 있다.

■ 포괄적인 의미에서 검색엔진(search engine)

- 인터넷(네트워크), 독립된 DB, 혹은 개인 컴퓨터 등에 있는 어떠한 정보 자원들 중에서 원하는 사람이 찾고자 하는 데이터(정보레코드)를 정보자원의 집합에서 추출하여 보여주는 것, 혹은 그 핵심 알고리즘, 또는 프로그램을 말한다.

■ 검색엔진이라는 용어의 사용 영역 2가지

- 첫째, 인터넷 상에 산재해 있는 제반 정보를 미리 수집하고 이를 체계적으로 저장한 후, 이용자가 원하는 정보를 수시로 찾을 수 있도록 해주는 일종의 인터넷 데이터베이스 시스템이다. 인터넷에 흩어져 있는 수많은 정보자원들은 조직적으로 관리되고 검색될 수 없는 형태이다. 검색엔진은 이들을 수집하고 체계적으로 관리하여 데이터베이스화하고, 이용자가 검색하도록 해주는 정보환경을 말한다. 구글, 네이버, 야후, 콰에르 등이 가장 대표적인 유형이다.
- 둘째, 데이터베이스에 소장된 정보에 대한 검색요구를 효율적으로 처리해주는 소프트웨어 또는 도구를 말한다. 기술적인 의미의 검색엔진이다. 이러한 소프트웨어 또는 도구는 개발자가 직접 제작할 수도 있지만, 상용 또는 오픈소스 형태의 제품을 사용할 수 있다.

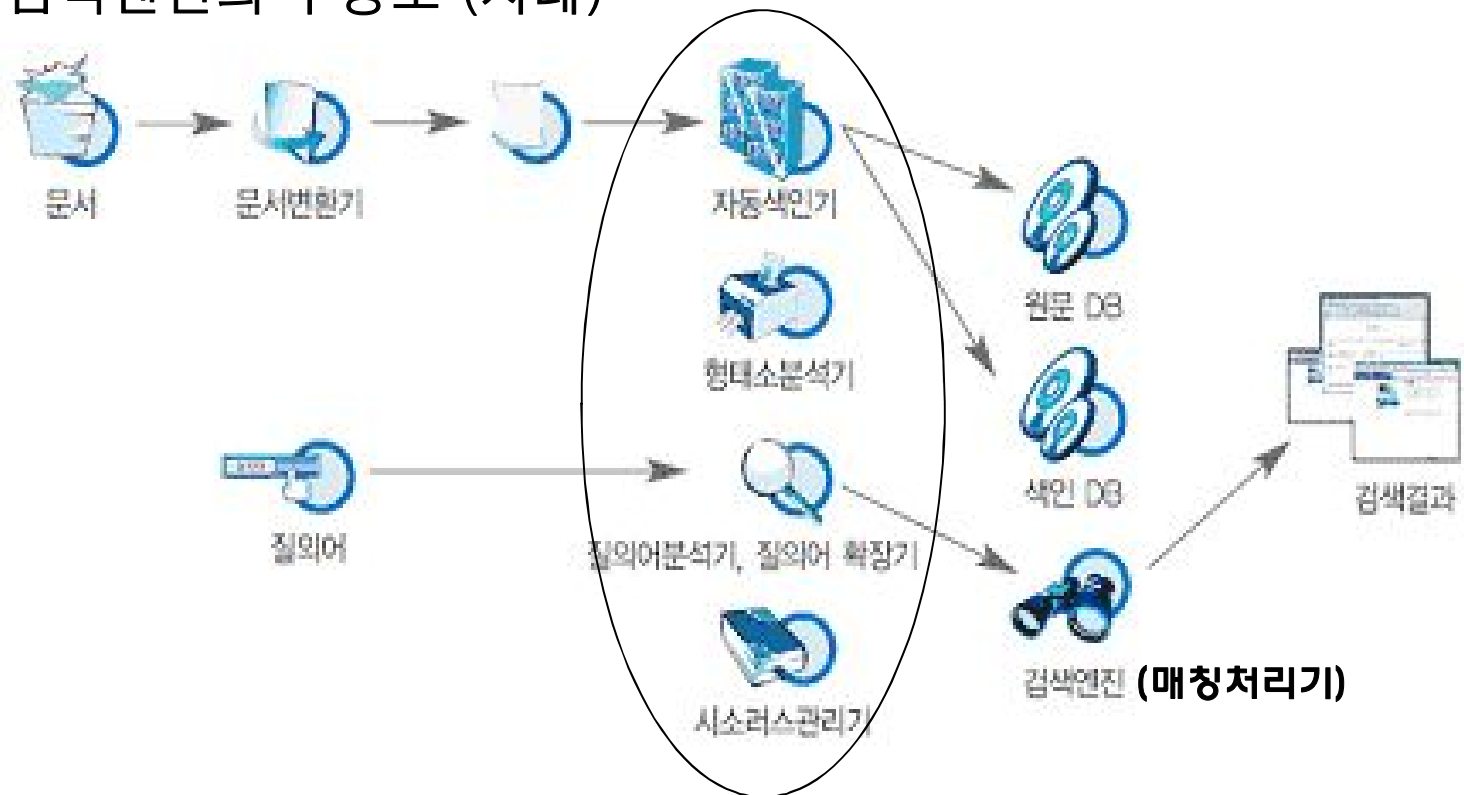
■ 정보검색과 검색엔진

- 정보검색에서 검색엔진은 핵심적인 컴포넌트이다. 이용자가 원하는 정보를 찾아주는 역할을 수행하기 때문이다. 이러한 검색엔진은 크게 키워드형 검색엔진과 디렉토리형 검색엔진으로 구분된다.
 - ✓ 키워드형 검색엔진: 키워드로 입력되는 검색요구를 처리하는 S/W
 - ✓ 디렉토리형 검색엔진: 분류체계를 이용하여 정보자원을 분류해서 이용자가 선택하도록 하는 S/W
 - ✓ 대부분의 정보검색시스템에서는 키워드형 검색엔진과 디렉토리형 검색엔진의 기능을 모두 제공하고 있다.
- 최근 동향: 시맨틱 검색엔진(IN2, OntoFrame 등)

■ 요소기술로서의 검색엔진의 사례

- 외산: K2(www.3soft.com), 잉크토미(Inktomi Enterprise Search)
- 국산: 레이다(한국정보공학), KRISTAL-II(한국과학기술정보연구원)
- 오픈소스: 루신(Lucene)

■ 검색엔진의 구성도 (사례)



■ 검색엔진의 주요 기능: 색인기, 검색기, 수집기

- **색인기**는 수집된 정보자원을 검색하는데 필요한 키워드를 색인하여 색인DB를 구축하는 기능을 처리한다. 또한 색인기는 검색질의에 표현된 각종 검색어휘들을 색인하여 검색기에 전달하여 정보가 검색되도록 한다. 색인기가 키워드를 추출하기 위해서는 형태소분석 방식 또는 n-그램 방식의 알고리즘을 사용하며, 스템밍(stemming)과 같은 키워드 용어의 각종 어미변환과 활용형들에 대한 처리를 한다.
- **검색기**는 색인기가 이용자가 입력한 검색질의에서 추출한 색인들을 토대로 검색 알고리즘(랭킹 알고리즘)을 이용하여 가장 적합한(일치하거나 유사한) 정보자원을 추출해 주는 기능이다. 따라서 질의분석을 위한 파싱 처리, 대량의 정보를 빠른 시간안에 찾아내는 검색 알고리즘 처리 능력이 요구된다.
- **수집기**는 인터넷 검색포털에서 검색엔진은 색인기와 검색기 이외에도 인터넷 상에 존재하는 각종 웹페이지들을 수집하는 기능이다. 흔히들 웹로봇(web robot)이라 하며, 크롤러(crawler), 스파이더(spider)라고 한다.

■ 다국어 정보검색

- 다국어 정보검색은 질의어의 언어와 상관없이 여러 언어로 만들어진 정보를 검색하는 시스템을 말한다. 텍스트를 기본으로 하여 음성, 이미지도 검색 대상으로 다루지만 주로 텍스트를 위주로 한다.
- 간단하게 말하자면 우리가 한글로 '잠수함'을 검색한다면 영어나 독일어, 프랑스어 등 여러 언어에서 의미가 '잠수함'인 단어들로 검색을 한다는 의미이다. 우리가 흔히 쓰는 한영사전이나 영한사전이 온라인으로 넘어와서 DB의 검색에 이용된다고 보면 될 것이다.
- 다언어 정보검색
 - 다국어 정보검색 방법
 - 언어 변환의 방식에 따라서 두 가지 방식으로 구분
 - ✓ 문서 번역(Document translation)방식
 - ✓ 질어의 번역(Query translation)방식

■ 문서 번역 방식

- 문서 번역 방식은 검색 대상 문서를 질의어의 언어로 변환하는 방식이다. 이를 위해서 해당하는 **기계번역 시스템**을 이용한다.

■ 질의어 번역 방식

- 질의어 번역 방식에서는 질의어가 임의의 통제 어휘(Controlled vocabulary)로만 구성이 되는지 아니면 자유 어휘(Free text)들을 모두 사용하는지에 따라 분류가 된다.
- 통제 어휘를 이용한 다국어 문서검색이란 미리 정한 어휘들만을 이용하여 검색 문서들을 모두 수작업으로 색인한 후 이용자의 질의어도 같은 어휘들로 표현하여 검색하는 것을 말한다.
- 자유 어휘 방식은 질의어를 검색 대상 문서의 언어로 변환하기 위해 활용하는 언어 자원의 종류에 따라서 세 가지 방법을 사용
 - ✓사전 기반 방법
 - ✓시소러스 기반 방법
 - ✓코퍼스 기반 방법

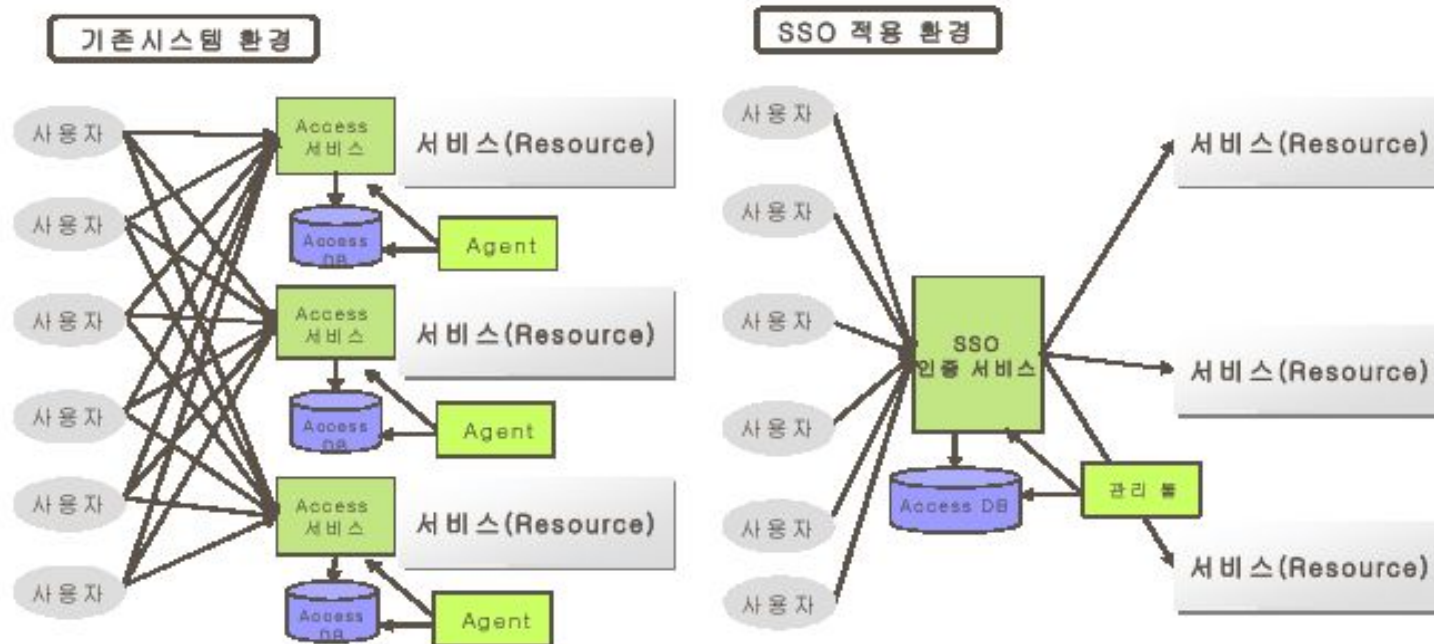
■ 이용자 인증

- 이용자 인증은 하나의 이용자(참가자, Party 또는 Claimant)에게 검증자(Verifier)가 다른 참가자와의 식별을 보증하는 절차이다.
- 일반적으로 검증자는 인증 요구자들이나 참가자들에게 알려져 있으며 그들의 신뢰를 받고 있다고 가정한다. 하나 또는 다수의 참가자들은 그들의 식별을 인증하고자 하는데, 단일(unilateral) 또는 상호(Mutual)인증을 제공할 것을 요청한다.

■ 디지털도서관의 이용자 인증

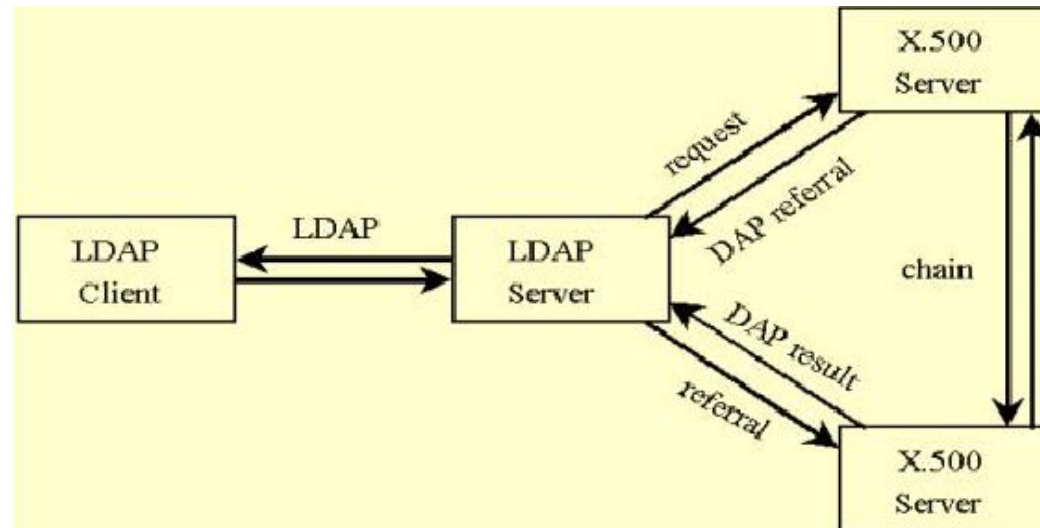
- SSO(Single Sign On)라는 개념을 구현하는 것을 원칙으로 한다.
- SSO는 한번의 로그인을 통해서 모든 서버에 접속할 수 있는 권한을 가지도록 하는 개념이다. 즉, 집안에 있는 모든 문을 열수 있는 만능키라고 보면 될 것이다. 이는 지금 현재 우리가 흔히 쓰고 있는 포털사이트에서 사용되고 있는 개념으로 우리가 네이버나 다음에 접속해서 한번의 로그인으로 모든 서비스를 이용할 수 있는 것으로 이해하면 될 것이다.
- SSO 구현을 위한 프로토콜 : LDAP 프로토콜, Shibboleth 프로토콜

■ SSO 인증의 개념도



■ LDAP

- LDAP(Lightweight Directory Access Protocol)은 각 서버별로 존재하는 회원정보를 통합하여 통합DB를 구축하고 빠른 인증서비스를 제공하기 위하여 사용되는 디렉토리 서비스의 표준 규약이다.
- 이는 모든 형태의 디렉토리형 자료를 표준화된 방식으로 저장하고 검색하기 위한 통신규약으로 미국 미시간 대학에서 개발된 통신규약이다.
- LDAP은 ITU-T의 X.500을 근거로 개발되었다.
- LDAP 계층도



■ 십볼릿

- 십볼릿(Shibboleth)은 오픈소스이며, 표준 기반의 솔루션으로 조직이나 기관이 자신의 이용자 정보를 안전한 방식의 교환을 가능하게 한다. 즉, 통합환경의 이용자 인증 솔루션 시스템이다.
- 각종 기관이나 벤더들이 십볼릿 통합시스템에 가입하면, 도서관 게이트웨이는 도서관 이용자와 개별 자원들간에 인증을 처리할 필요가 없게 된다. 십볼릿에서 제공하는 속성들을 이용하여 고객에게 맞추어진 각종 서비스를 제공한다.

■ 이용자 인증 방법: 로그인 방법

- 개개인에게 할당된 ID와 PASSWORD를 입력하는 방식
- 은행이나 카드결제시 사용되는 인증서방식
- 텔레뱅킹에 사용되는 방식처럼 매번 다른 비밀번호를 요구하는 일회성 PASSWORD와 공개키를 혼합한 방식
- TAG카드를 이용한 인증 방법
- 홍채나 지문을 이용한 방법 등

■ 이용자 인증 방법 구분(비밀번호와 암호화 방식에 따라)

- 질의-응답 방식
- 질의-응답 암호화 방식(공개키 방식 등)
- 스마트카드 방식
- 생체정보인증 방식

■ Z39.50

- Z39.50은 1988년에 국가정보표준기구(National Information Standards Organization, NISO)에서 승인되었고, 현재는 미국국가표준 (American National Standard) 이며, 공식적인 명칭은 "ANSI/NISO Z39.50-1995-Information Retrieval(Z39.50) : Application Service Definition and Protocol Specification" 이다.
- 데이터베이스 탐색과 정보검색을 위하여 통신하는 두 개의 시스템이 지켜야 할 규칙과 프로시저 등이 명시되어 있는 표준
 - ✓Z39.50을 이용할 경우 정보검색절차와 방법이 표준화되기 때문에 대규모 정보데이터베이스를 이용하는데 훨씬 쉬워진다.
 - ✓특히 Z39.50은 분산형 클라이언트/서버 환경에서의 정보검색을 지원한다. 서버용 소프트웨어는 한 개 이상의 데이터베이스를 탐색하여 탐색 명령어를 만족시키는 검색결과를 클라이언트에 돌려준다.
- Z39.50을 이용할 경우 OS, 하드웨어, 탐색엔진, DBMS 등이 서로 다른 시스템들끼리 어려움 없이 상호작용할 수 있다.

■ ZING

- 웹 기반 기술을 사용하는 개발자 및 정보제공자에게는 Z39.50의 유용성이 크게 인정받지 못했다. 이러한 배경에서 ZING이 출현했는데, ZING(Z39.50 International Next Generation)은 그 이름처럼 차세대 Z39.50 개발을 위하여 수행되는 여러 가지 실험적인 연구 및 프로젝트들을 통칭하는 포괄적인 용어이다.
- 현재 ZING으로 수행 중인 다섯 가지 실험적 연구 주제들은 다음과 같다.
 - ✓SRW/SRU : Search / Retrieve for the Web Service / URL Service
 - ✓ (웹 서비스와 URL 서비스를 위한 검색 프로토콜)
 - ✓CQL : Common Query Language
 - ✓ZOOM : the Z39.50 Object-Oriented Model
 - ✓ZeeRexL Z39.50 Explain, Explained and Re-Engineered in XML
 - ✓ex3950L Simple Implementation of Z39.50 over SOAP using XER

■ OAI 프로토콜

- 새로운 학술커뮤니케이션 모델인 개방접근(Open Access) 운동을 체계적으로 지원하기 위한 기술로 개발되었다.
- OAI 프로토콜은 가볍고 기술장벽이 낮으며, OAI 저장소간 정보를 쉽게 공유할 수 있는 매우 유용한 프로토콜이다. 기존 Z39.50 프로토콜이 정교한 메타데이터인 MARC을 채택함으로써 비용 부담이 컸던 반면, OAI 프로토콜은 더블링크어 메타데이터를 사용하여 상대적으로 저비용이다.
- OAI 프로토콜은 통합검색을 위해 수확검색(Harvesting Search) 방식을 사용하고 있으며 서비스 제공자가 데이터 제공자로부터 메타데이터를 주기적으로 수확한 후 통합검색을 하게 된다.
 - ✓수확검색은 서비스 제공자가 메타데이터를 수확하여 이를 로컬 저장소에 저장한 후 검색에 이용하는 방식이다.
 - ✓수확검색은 각 저장소가 개별적으로 디지털도서관의 요청에 따라 메타데이터를 전송할 수 있어 부담이 적다.

■ OAI 프로토콜의 특징

- 첫 번째로 HTTP, XML, XML 스키마와 같은 웹 표준을 기반으로 동작하기 때문에 구현하기가 쉬우며,
- 둘째로 프로토콜에 의해 지원되는 동작체계가 단순하다.

■ OAI 프로토콜 버전

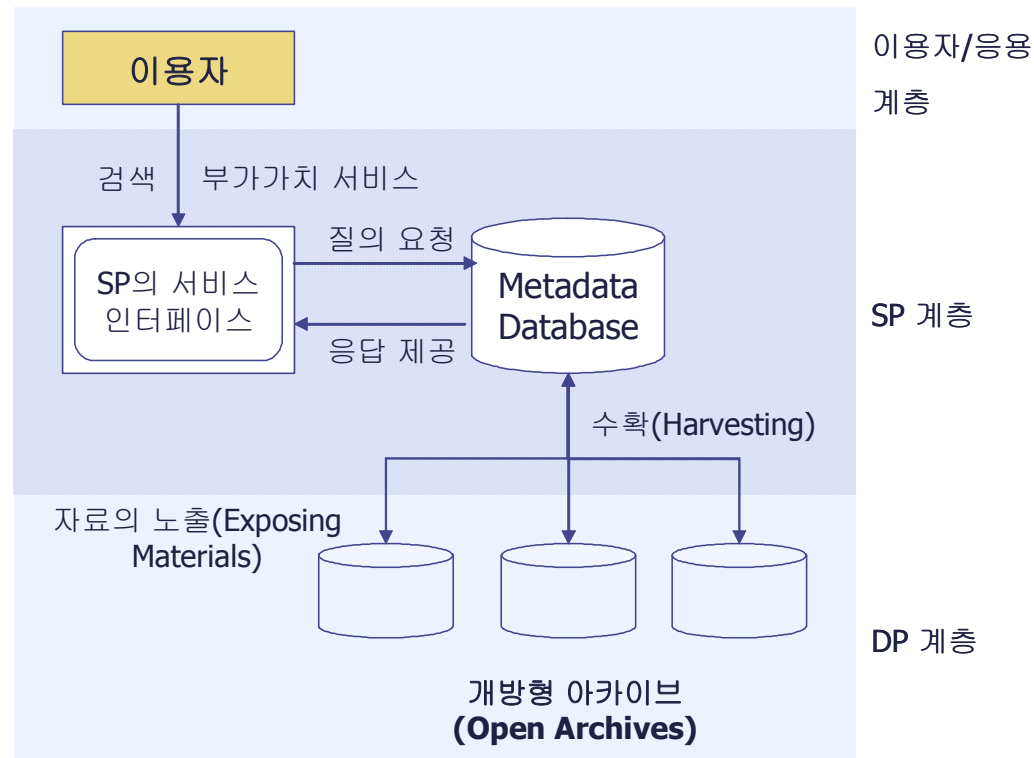
- OAI-PMH는 버전 1.0을 2001년 1월 21일에 발표하였으며 버전 1.1을 2001년 7월 2일에, 버전 2.0을 2002년 6월 14일에 발표하였다.
- OAI-PMH 버전 1.0과 2.0의 주요한 차이점은 OAI-PMH 버전 2.0에서는 XML 스키마를 사용하였다는 것이다.

■ OAI 프로토콜 운용모델

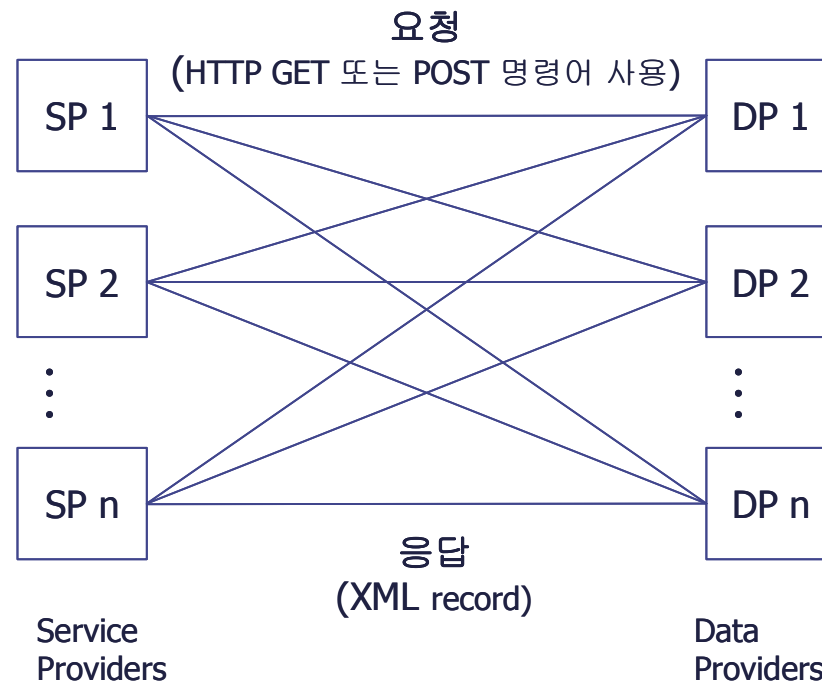
- OAI 프로토콜의 운용은 SP/DP라는 아주 단순한 모델로서 설명된다.
- SP(Service Provider)는 DP들로부터 수확한 메타데이터를 기반으로 검색, 브라우징, 연계 및 원문제공 등과 같은 부가가치 서비스를 제공하는 기능을 수행한다.
- DP(Data Provider)는 자체적으로 각종 디지털 컬렉션을 수집, 보유하고 있는 시스템이며, SP의 수확요청에 대응하여 적합한 메타데이터를 제공하는 역할을 수행한다.
- 이러한 SP/DP 모델은 각각의 역할 또는 기능에 따라 구분되는 개념이지, 그것이 바로 시스템인 것은 아니다.

■ SP/DP 모형

- SP/DP 모형은 각각의 역할 또는 기능에 따라 구분되는 개념이지, 그것이 바로 시스템인 것은 아니다.



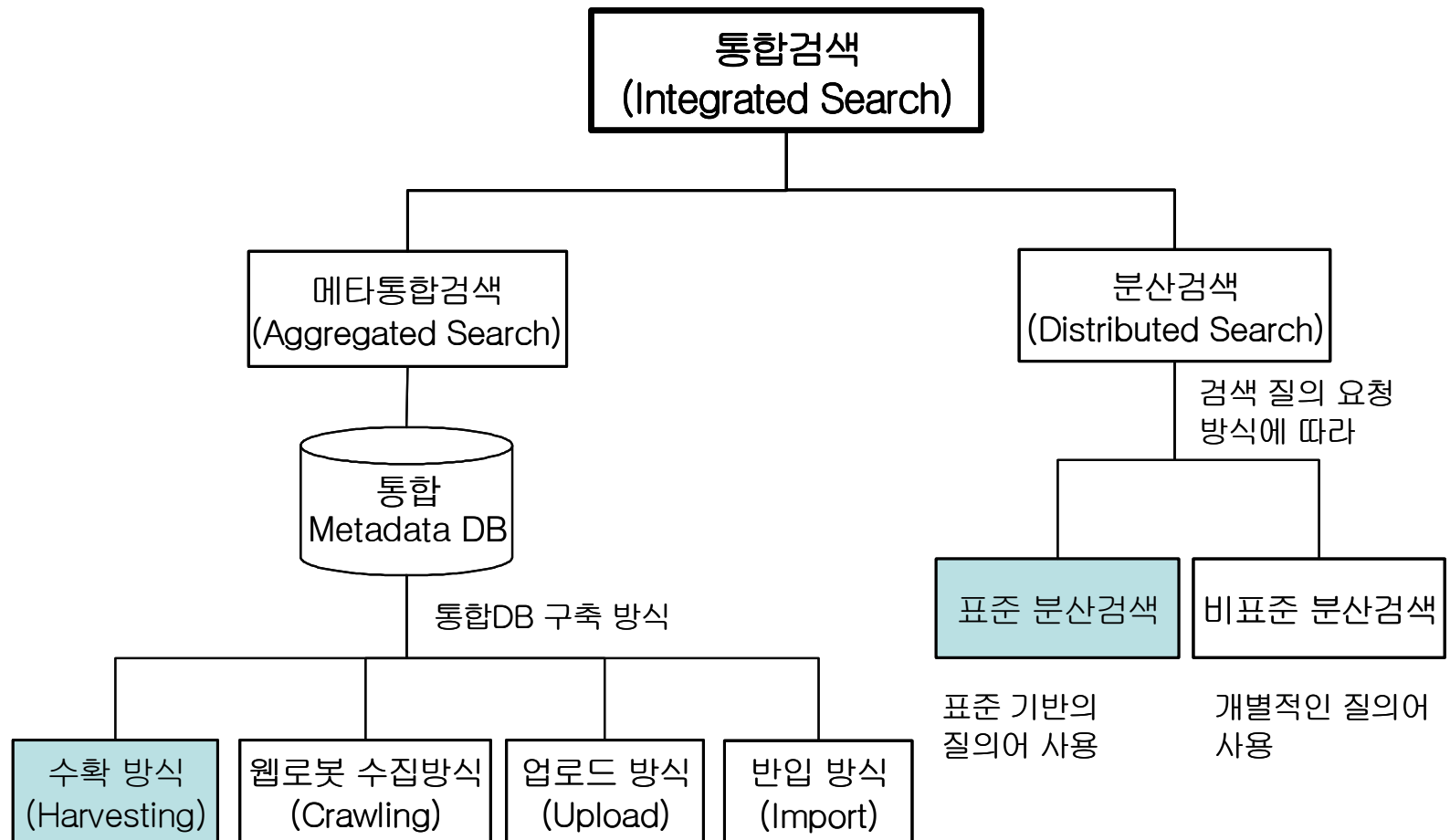
■ SP/DP 커뮤니케이션 방식



■ 6가지 요청 명령어와 기능

명령어	기능
Identify	아카이브에 관한 일반정보 검색
GetRecord	아카이브에서 하나의 레코드 검색
ListIdentifiers	아카이브에서 수확 가능한 레코드의 식별자 검색
ListRecords	아카이브에서 복수의 레코드를 수확
ListSets	아카이브의 세트 구조정보를 검색
ListMetadata Formats	특정 레코드에서 사용하고 있는 메타데이터 포맷 종류를 검색

■ 통합검색 기법의 유형



■ DLF의 ERMI(1)

- ERMI(Electronic Resource Management Initiative)는 DLF(Digital Library Federation) 프로젝트 중의 하나이며, XML 기반의 메타데이터 스키마 /DTD와 라이선스 관련 데이터를 관리하는 각종 도구들을 개발하는 것을 목적으로 한다.
- 워크플로우, 모범사례 조사, 표준화 촉진, 도서관시스템 벤더들에게 ERM 시스템(ERMS) 개발에 참여하도록 하는 등의 과제들을 수행한다.
- 제웰이 코넬 대학의 정보기술 담당 사서인 샐러(Adam Chandler)와 공동으로 2001년에 DLF의 지원으로 ERMI 프로젝트를 시작하였다.

■ DLF의 ERM(2)

- 2년 동안 ERM 프로젝트는 ALA 컨퍼런스, 워크숍 등을 개최하여 도서관 사서, 시스템 개발자, 데이터베이스 벤더 등과 같은 이해관계자들의 의견을 구했으며, NISO와 함께 전자자원 관리기능과 표준화가 요구되는 데이터 요소를 파악하는 조사작업도 수행하였다. 이 협력조사 결과는 "NISO White Paper - The Exchange of Serials Subscription Information"라는 백서로 발간하게 된다.
- 이러한 활동과 문서들을 정리하여 2004년에 "Electronic Resource Management: Report of the DLF ERM Initiative"라는 초안 버전의 보고서를 온라인으로 발표하였고, 각종 피드백을 수용하여 2005년에는 온라인과 인쇄형 보고서를 발간하였다.

■ 전자자원의 이용통계 분석

- 전자자원 이용통계 데이터의 추출과 분석 문제는 최근 도서관계의 주요 관심 영역이 되고 있다.
- ERMI 프로젝트 동안 실무추진단 구성원들은 도서관 사서들에게서 다양한 정보원들로부터 통계데이터를 저장하고 표현하는 공통의 프레임워크가 필요하다는 요구를 받았다.
- 그래서 ARL의 E-metrics 프로젝트나 COUNTER 프로젝트와 같은 활동들에 관심을 기울이고, SUSHI(Standardized Usage Statistics Harvesting Initiative) 활동에는 직접적으로 참가하고 있다.
- 이 협력 작업은 도서관 벤더의 시스템과 ERM 시스템 간에 COUNTER 보고서를 직접 전송하는 웹 프로토콜을 개발하는 것으로, ERMI 2 프로젝트의 주요 기능의 하나이다.

■ ERM I 2 프로젝트

- ERM I 2 프로젝트는 데이터 표준화, 라이선스 용어에 대한 전문가 훈련, 전자자원 이용통계, ERM I 작업과 다른 관련 프로젝트들과의 조장작업 등에 대한 문제들을 해결하는 것을 목표로 하고 실무추진단을 구성하였으며, 전체 8명이 참여하고 있다.

■ ERMS

- 기관이 전자저널, 전자책, DB 등의 디지털 자원 자체에 대한 수집, 유지, 서비스, 그에 따른 라이선스 등의 라이프사이클을 관리하는 시스템

■ COUNTER

- 온라인 정보자원의 이용은 급속도로 성장해 왔으며 정보의 생산자와 구매자는 온라인 정보자원에 대한 이용이 지속적인 방법으로 측정되어야 한다는 사실에 의견 일치를 보였다. 이에 벤더, 사서, 정보중개자들은 온라인 정보자원의 이용성을 측정할 수 있는 국제표준을 개발하기 위해 국제적인 노력이 필요함을 자각했다.
- COUNTER(Counting Online Usage of Networked Electronic Resources)는 JISC, ALPSP, 출판협회(Publishers Association)에 의해 형성된 PALS(Publisher and Librarian Solutions) 그룹과 함께 2002년 3월에 영국에서 시작되었으며, 온라인 이용 통계의 기록과 교환을 용이하게 함으로써 사서와 출판업자와 정보중개자를 지원할 목적으로 추진된 국제적인 프로젝트이다(<http://www.projectcounter.org>).
- COUNTER 시행령(code of practice)의 목적은 지속적이고 신뢰할 수 있는 벤더가 온라인 이용 통계를 준비하는 데 있어 국제 공개 표준과 프로토콜을 수립함으로써 온라인 이용 데이터의 기록, 교환 및 해석을 용이하게 하는 것이다.

■ 장서등록기

- 장서등록기(Collection Registry)는 IMLS DCC 프로젝트에서 처음으로 사용하는 이름이며, 장서정보의 데이터베이스이다. 장서정보의 검색과 서비스에 관련된 기능을 제공한다.

■ IMLS DCC 장서등록기

- IMLS(Institute of Museum and Library Services)는 NLG(National Leadership Grant) 프로그램을 통해 수백 건의 디지털 장서를 개발하고 있다. 이처럼 IMLS가 투자한 디지털 장서를 통합 관리할 수 있는 레지스트리(registry)를 의미
- IMLS 장서 레지스트리 환경에서의 장서란,
 - ① 주제영역이든 자료유형이든 결합할 수 있어야 하고,
 - ② 별개의 장서로 검색 가능해야 하며,
 - ③ 요소의 유일한 접근점을 통해 이용할 수 있는 것이다.

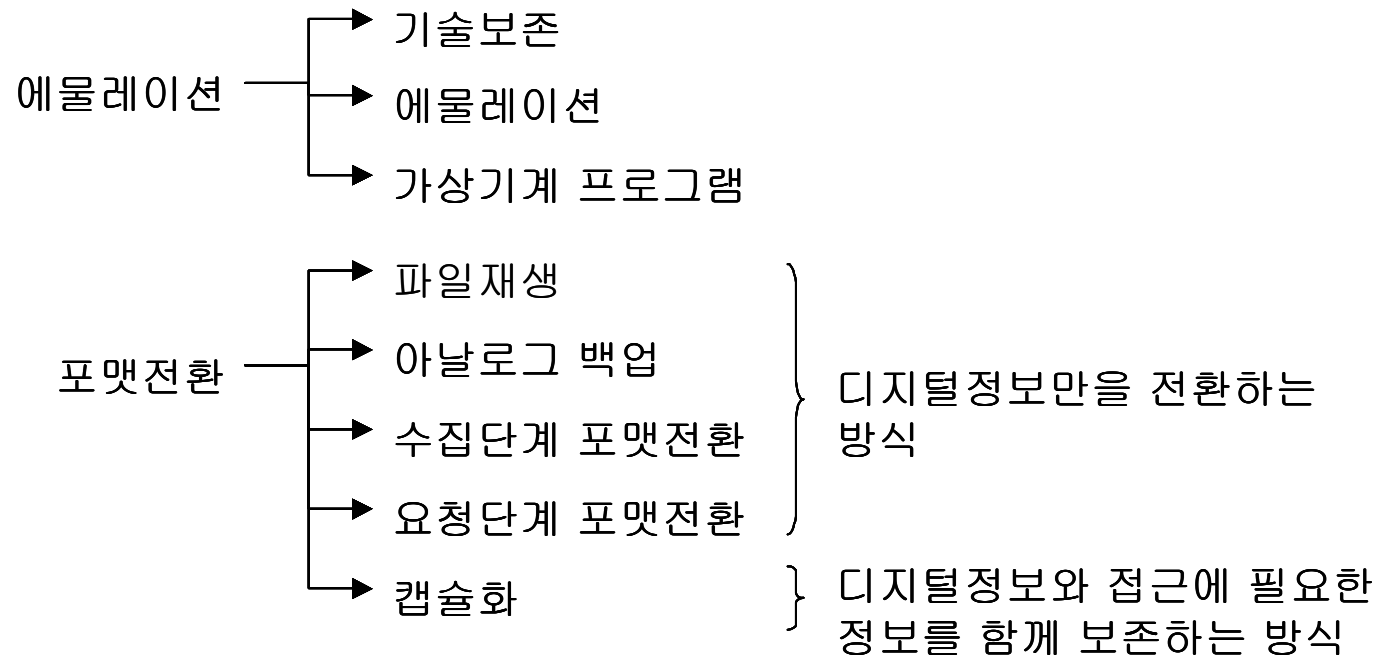
■ 장서등록기의 기술요소

- RSLP(Research Support Libraries Programme)의 장서기술 스키마(Collection Description Schema; 이하 RSLP CD 스키마)와
- Dublin Core의 장서기술 응용 프로파일(Collection Description Application Profile; 이하 DC CD AP)을 기반으로
- IMLS DCC(Digital Collections and Content) 장서기술 메타데이터(IMLS DCC CD 스키마)를 개발

■ IMLS DCC 장서등록기의 기능

- 주제 영역별 브라우징
- 장서들 간의 관계 표현
- 기간, 지역, 이용자, 자료의 유형에 의한 제한 검색 등
- 장서 레지스트리 레코드의 수정 등

■ 보존처리 기술 방식



■ 3가지 보존처리 기술 방식의 장단점

구분	장점	단점	대상 영역
에뮬레이션	-보고 느끼는 형태 그대로 유지	-에뮬레이션 기술사양의 구축이 복잡함 -많은 양의 정보를 보존하여야 함 -오래된 소프트웨어 사용	-응용 소프트웨어 -실행파일을 포함하는 복잡한 디지털 자원 -충분한 정보가 부족한 자원 -미래의 가치와 활용이 불확실한 자원
포맷전환	-원본의 응용 환경을 보유할 필요가 없음 -적극적인 접근과 관리활동을 지원	-많은 비용 소요 -정보내용의 퇴화 -보존 메타데이터의 부족 -관리자에 의한 지속적인 관심 요구	-과학데이터와 DB처럼 적극적인 접근과 관리가 필요한 자원 -충분히 알려진 포맷의 자원
캡슐화	-보존정보를 유지	-포맷에 관한 지식을 보존하여야 함 -디지털정보를 추출하는 기능 필요	-적극적인 접근과 관리가 요구되지 않는 자원 -잘 알려지지 않은 포맷의 자원