

웹 학술정보자원의 효율적인 검색을 위한 자료유형 분류체계에 관한 연구

A Study on the Classification Schemes of the Type of Scientific Resources on the Web for the Effective Retrieval

김소형, 안혜연(성균관대학교 대학원 문헌정보학과)

So Hyeong Kim, Hye Yeon An

Dept. of Library and Information Science, Graduate School of
Sung Kyun Kwan University

본 연구의 목적은 웹 상의 다양한 학술정보자원들을 자료유형별로 효율적인 검색이 가능하도록 다양한 형태의 웹 학술정보자원에 대한 자료유형 분류체계 방안을 제시하는데 있다.

본 연구에서는 DCMI(DublinCore Metadata Initiative)를 비롯한 국내외 주요 기관이 연구·개발한 7개의 자료유형 요소 셋을 조사·분석하여, 이를 의미적으로 매핑시키고 그 결과를 참조하여 웹 학술정보자원의 특성에 맞게 10개의 범주 내에 37개의 자료유형 요소를 도출하였다. 본 연구에서 제시하는 자료유형 분류체계는 웹 학술정보자원에 공통적으로 포함되는 요소이므로 각 학문주제별로 적용시킬 경우에는 자료유형 요소의 추가 및 변경을 요구할 수도 있다.

1. 서론

본 연구의 목적은 웹 학술정보자원의 자료유형별 분류 및 검색이 가능할 수 있도록 다양한 형태의 웹 학술정보자원에 대한 자료유형 분류체계를 제시하는데 있다.

지금까지 웹 자원에 대한 자료유형 분류체계는 체계적이고 일관성 있는 기준에 의한 것이라기보다는 웹 학술정보자원의 대표적인 몇몇

자료유형에만 치중되거나 기존의 포털사이트에서 제공하는 자료유형 형태를 그대로 답습하는 경우가 대부분이었다.

이에 본 연구는 먼저 국내 및 해외의 대표적인 학술기관에서 이미 개발되어 사용 중인 웹 정보자원의 자료유형을 찾아 각 요소에 대한 정의와 설명을 정리하여 종합·분석하였다. 그리고 이를 바탕으로 웹 상의 학술정보자원이 라는 특수한 성격에 맞추어 1차 자료, 2차 자료, 3차 자료로 정보의 속성을 나누어 살펴본

후, 웹 학술정보자원에 대한 표준적인 자료유형 분류체계를 제시하고자 한다.

본 연구에서 제시될 웹 학술정보자원의 자료유형별 분류체계는 추후 웹 학술정보자원의 메타데이터를 구축함에 있어 자료유형 메타데이터 요소의 통제어를 지정하는데 유용한 참고자료로 활용될 수 있을 것이다.

2. 국내의 웹 자원 자료유형 요소 분석

국내의 대표적인 학술기관에서 제시하고 있는 웹 자원 자료유형 요소들을 추출·분석하는 과정은 다음과 같다.

① 국내의 대표적인 학술 기관에서 연구·개발한 웹 자원 자료유형을 조사하여 각각 요소의 의미를 정리한다.

② DCMI(DublinCore Metadata Initiative)에서 제시하는 9개의 기본 자료유형 요소를 기준으로 각기 다른 자료유형 요소들을 상호 매핑시켜 고유한 의미요소들을 추출하고 그것들의 의미가 비슷한 자료유형들끼리 범주화시켜 정리한다.

※ 이때 범주화하는 과정에서 DCMI에서 제시하는 9개의 기본 자료유형 요소 외에 추가 기준 항목을 필요로 하였기 때문에 임의로 기관/단체, 커뮤니티, 기타의 3개 항목을 범주 기준 요소로 추가하였다.

[표 1]은 DCMI에서 제시하는 9개의 자료유형 요소의 정리 및 각 요소에 대한 설명이다.

(1) 국외의 웹 자원 자료유형 조사·정리

국외의 교육 및 학술 분야의 대표적인 기관에서 자료들에 적용하기 위하여 만들어진 웹 자원 자료유형 요소들을 조사하고, 각 요소에 대한 의미와 특성을 파악하였다.

모든 요소에 대한 의미 분석과 DCMI의 자

료유형 요소를 기준으로 의미요소로의 매핑 과정을 거친 결과 모든 요소를 총망라하여 정리한 결과물은 [표 2]와 같다.

요소	설명
Collection	자원이 그룹화된것. 항목들의 집합(각기 검색되거나 기술될 수 있음)
Dataset	리스트, 표, 데이터베이스
Event	전시회, 공연, 컨퍼런스, 워크샵
Image	사진, 그림, 회화자료, 그래픽, 영화, 다이어그램, 지도
Interactive Resource	멀티미디어 학습물, 채팅서비스, 가상현실, 애플릿
Service	사진복사, 은행, 인증
Software	컴퓨터 프로그램
Sound	음악파일, 오디오디스크, 녹음된 연설이나 음악
Text	도서, 레터, 논문, 시, 신문, 학위논문
기관/단체(임의)	개인, 기관 정보나 홈페이지 등
커뮤니티(임의)	유즈넷, 커뮤니티 등
기타 (임의)	어떤 자료유형에도 속하지 않는 자원

[표 1] DCMI 자료유형 요소 설명

본 연구에서 조사·분석한 자료유형의 종류는 다음과 같다.

- ① DCMI(DublinCore Metadata Initiative) Type Vocabulary
- ② OBSOLETE: Dublin Core Standard Resource Types
- ③ Social Science Information Gateway (SOSIG)
- ④ Gateway to Educational Materials (GEM)
- ⑤ VW96 schema description

	DCMI	SOSIG	ROADS Project	GEM	VW96 schema
1	Collection	Articles / Papers / Reports (collections) Journals (full text)	Collection Proceedings Journal Magazine	Collection Serial	Journal Magazine
2	Dataset	Bibliographic Databases Bibliographies Data	Bibliography Dataset	Data set Research study Catalog record	Catalog Index Database Hypercatalog
3	Event	Research Projects		Event Activity Project	
4	Image		Advertisement Image Video	Image set	
5	Interactive Resource			Form Environment	SearchEngine Mall
6	Service		Service	Service Tool	
7	Software	Software		FAQ	Archive Linecard Keybank
8	Sound		Music		
9	Text	Articles / Papers / Reports (individual) Books / Book Equivalents Educational Materials Government Publications Journals (contents and abstracts) News Reference Materials Resource Guides	Article Book Booklet Course Material Honours Thesis In Book In Collection In Proceedings Manual Masters Thesis Message On Moderated Mailing List Message On Unmoderated Mailing List Newspaper PhD Thesis Poem Preprint Research Paper Tech Report Unpublished Unrefereed Article	Reference Story Study guide Educator's guide Lesson plan Unit of instruction Primary source Secondary source Curriculum support Literature	RFC Document Dictionary Manual Book HOWTO
10	기관/단체	Companies Organisations / Societies Research Centres Governmental Bodies	Organisation Info Personal Info		Homepage
11	커뮤니티	Mailing Lists / Discussion Groups	Posting To Moderated Newsgroup Posting To Unmoderated Newsgroup	Community	
12	기타		Misc	Course Curriculum Realia Artifact Best practice	World RealWorld

[표 2] 국외의 웹 자원 자료유형 조사·정리

(2) 국내의 웹 자원 자료유형 조사·정리

국내의 대표적인 학술기관들을 중심으로 웹 자원 자료유형 요소들을 조사하고, 각 요소에 대한 의미와 특성을 파악하였다.

모든 요소에 대한 의미 분석과 DCMI의 자료유형 요소를 기준으로 의미요소로의 매핑 과정을 거친 결과 모든 요소를 총망라하여 정리한 결과물은 [표 3]과 같다.

본 연구에서 조사·분석한 자료유형의 종류는 다음과 같다.

- ① METADB
- ② 한국교육학술정보원(riss4u.net)
- ③ 기타 포탈사이트

(3) 분석 결과 추출해 고유한 자료유형 요소 집합

국내의 주요 기관들에서 연구·개발한 모든 자료유형 요소에 대한 의미 분석과 DCMI의 자료유형 요소를 기준으로 의미요소로의 매핑 과정을 거친 결과, 중복되는 의미를 지닌 요소들이 하나의 의미요소로 묶이게 되었다. 그 결과 [표 4]와 같이 7개 세트의 총 164개의 요소들로부터 추출된 81개의 고유한 의미요소들의 집합인 1차 최종 자료유형 요소들이 정리되었다. 이것은 국내의 주요 기관들에서 연구·개발한 웹 자원에 대한 모든 자료유형 요소들을 총망라하여 정리한 결과물이다.

3. 웹 학술정보자원 자료유형 분류체계를 위한 Map 구성

(1) 1차 최종 자료유형 81개 요소에 대한 분석

대학원생 12명이 [표 4]에서 정리된 81개 각 자료유형 요소마다 실제 자료를 모으는 과정에서 내용 및 형태를 분석한 후, [표 4]의 자료유

형 요소표에 있는 요소들 중 필요하다고 판단되는 요소에 체크를 하고, [표 4]의 자료유형 요소표에는 없었으나 추가할 필요가 있다고 생각되는 요소를 유형별로 뽑아내었다.

	DCMI	METADB	RISS4U
1	Collection	잡지	학술지
2	Dataset	정보원목록 기사색인목록 주소록 인명록 상품목록 통계	주제별 해외DB
3	Event		
4	Image		비도서
5	Interactive Resource		통합검색 해외DB 통합검색(시험서비스)개최정보검색 국내의 검색엔진 웹강의실
6	Service		Call For Paper
7	Software		
8	Sound		비도서 동영상/발표자료
9	Text	백과사전 어학사전 용어사전 약어사전 시소러스 법령집 정책보고서 연감/백서 논문 보고서 교재 신문 eBook 특허/실용신안 표준/규격 회의자료/이벤트 자료	단행본 학술지수록논문 학위논문 강의계획서 강의록 해외박사학위논문 (DDOD) KERIS 연구보고서 용어사전
10	기관/ 단체	개인 기관/단체 기업	학회정보 교수, 연구실 홈페이지 대학검색 대학도서관
11	커뮤니티		
12	기타		

[표 3] 국내 웹 자원 자료유형 조사·정리
각 유형별로 1차 최종 요소표 중에서 체크된

요소들 중 일정 수준 이상 공통적으로 나타나는 요소들을 뽑아 기본요소로 삼고, 그 외에 유형마다 특이한 요소들은 논의를 거쳐 필요성이 있다고 판단되는 요소에 한해 각 유형별 특수요소로 선정하였다.

DCMI	의미요소명	DCMI	의미요소명			
Collection	논문/보고서 묶음자료 정기간행물(묶음자료) 정부간행물(묶음자료) 일반잡지(묶음자료) 프로시딩(묶음자료)	Text	학술논문(개별) 도서 소책자 강의자료 (강의계획서) 사전(백과, 어학, 용어, 시소러스 등) 메일링 리스트의 메시지 정부간행물(개별) 우수 논문 도서 일부 묶음자료 일부 프로시딩 일부 메뉴얼 석사논문 신문 학술잡지 초록 박사학위 논문 시 프리프린트 원정보 참고자료 연구논문 연구가이드 원정보에 대한 2차자료(비평, 리뷰) 기술 보고서 미간본 평가되지 않은 논문 ebook 회의자료/이벤트 자료			
Dataset	목록 DB 데이터셋(주소록, 인명록, 통계 등) 연구결과물(데이터셋만 포함된것) 서지.목록(정보원, 기사색인, 상품목록 등)		커뮤니티	메일링리스트/ 토론그룹 유스그룹		
Event	연구활동(이벤트) 프로젝트			기타	Misc(어떤 유형에도 속하지 않는 자료) 학기과정(코스) 커리큘럼 실물교재 모조품 모범사례 World RealWorld	
Image	광고 그림/사진(이미지) 비디오 비도서		Software		Archive Linecard Keybank 소프트웨어	
Interactive Resource	데이터 입력폼(온라인 설문 등) 채팅, 게임 검색엔진 쇼핑몰				Sound	음악자료 동영상발표자료
Service	서비스 연구도구(서비스의 일종) FAQ				기관/단체	정부기관 기업 조직/단체 개인 연구센터

[표 4] 1차 최종 자료유형 요소표

(2) 웹 학술정보자원에 맞는 자료유형 분류 체계 구성

[표 4]의 1차 최종 자료유형 요소가 일반적인 웹 자원에 모두 적용시킬 수 있는 자료유형 분류체계라면 본 연구에서는 웹 학술정보자원이라는 특수한 성격을 고려하여 이에 맞게 자료유형 분류체계를 재구성하기로 하였다.

먼저 (1)에서 81개 자료유형 요소에 대한 추가·삭제의 분석과정을 마친 임의의 자료유형 요소표를 가지고 자료가 가지는 속성에 따라 [표 5]에서 보는 바와 같이 자료 유형의 성격을 1차 자료, 2차 자료, 3차 자료의 세가지 범주로 나누어 MAP을 그렸다.

1차 자료의 범주에는 원래의 연구성과물을 기록한 개별적인 자료가 이에 속하며 2차 자료의 범주에는 개별 자료인 1차 자료의 묶음자료이면서 개별자료로의 접근을 위한 관문 역할을 하는 자료(자료실의 개념)와 1차 자료를 효과적으로 찾아볼 수 있는 자료가 이에 속하며, 3차 자료에는 참고적인 성격을 띠는 3차 자료로 나누었다.

[표 5]의 MAP을 기초로 하여 웹 학술정보자원의 자료형식적 성격이 위의 세가지 범주 내에서 뚜렷하게 나뉘어질 수 없다는 점과, 자료의 내용유형 및 형식유형을 보다 세분화할 필요성을 고려하여 [표 6]의 최종적인 자료유형 분류체계를 도출하였다.

4. 결론

전문 학술분야의 정보수요가 급증하는 시점에서 오늘날 인터넷이 제공하는 웹 학술정보자원은 양과 질적인 면에서 풍족해졌지만 주제 분류체계에 비해 자료유형에 대한 체계적인 분류체계가 부족한 실정이다.

본 연구에서 제시하는 학술정보자원에 대한 자료유형 분류체계는 이용자들이 자료를 찾을

1차자료	2차자료	3차자료
사전	데이터베이스	
강의계획서	서지.목록	
강의 및 강의관련 자료	개인홈페이지의 자료실	기관 홈페이지
학생보고서	(교수/전문가/학생)	(연구기관/도서관/학과/학회 등)
학위논문	원문제공기관	개인 홈페이지
학회지/학술지/컨퍼런스논문	학과 자료실	(교수/전문가/학생)
기관보고서/프로젝트 보고서	Q & A	학문분야별 포털사이트
실무자료(매뉴얼 등)	게시판 형태의 자료실	
정책보고서/법령집	포털사이트의 자료실	
일반학습자료	학회지/학술지	
동영상자료	제공문	
ebook		
음성자료		

[표 5] 자료의 속성별 자료유형 Map에 있어 자료유형별 검색이 가능하도록 하고, 추후 웹 학술정보자원의 메타데이터를구축함에 있어 자료유형 메타데이터 요소의 통제어를 지정하는데 유용한 참고자료로 활용이 가능하다.

본 연구에서 제시하는 자료유형 분류체계는 웹 학술정보자원에 대한 공통적인 항목들을 중심으로 구분되었기 때문에 추후 학문주체별로 적용을 시킬 경우에는 각 학문의 특성에 따라 부분적으로 요소의 추가/삭제/변경이 필요하다.

[참고자료]

- <http://www.roads.lut.ac.uk/Metadata/DC-ObjectTypes.html>
- <http://www.dublincore.or.kr>
- <http://www.dublincore.org/documents/1999/08/05/resource-typelist/>
- <http://www.loc.gov/marc/dc/typequalif.html>
- <http://www.roads.lut.ac.uk/Metadata/DC-ObjectTypes.html>
- <http://www.sosig.ac.uk/help/search.html#type>
- <http://vancouver-webpages.com/META/VW96-scheme.html>
- <http://www.metadb.net>

웹 학술정보자원 자료유형 분류		
강의자료	강의계획서 강의관련자료	'강의관련자료'는 반드시 '강의'만을 위해 만든 자료임(강의에 참고할 수 있는 모든 자료를 포함하는 것이 아님)
논문	학위논문 일반연구논문 회의/컨퍼런스 논문	
사전/디렉토리	용어사전 약어사전 시소러스 인명록/주소록	
일반자료	법령집 정책자료 통계 연감/백서 보고서/제안서 기사자료 회의정보 에세이/칼럼	'회의정보'는 '회의논문'을 제외한 회의결과 및 공지에 관련한 회의정보에 관한 웹문서를 뜻한다.
간행물	신문 학술지 기관지 프로시딩 웹진	'학술지'의 경우 학술지에 실린 개별논문들을 따로 구분하여 취급하지 않고 '묶음'자료로써 취급함.(개별논문은 '논문'유형에 포함시킴)
2차정보원	온라인목록 데이터베이스 자료실	
실무자료	매뉴얼/업무편람 표준/규격자료 사례집	학문별로 특성화하여 유형 구분에 있어 추가/삭제/변경이 가능하다.(예:문헌정보학의 경우 분류표/목록규칙, 저자기호표 등이 추가될 수 있다)
홈페이지	개인 기관/단체 기업 포탈사이트	학문별로 특성화하여 유형 구분에 있어 추가/삭제/변경이 가능하다.(예:사회복지학의 경우 교육기관, 노인복지시설, 장애인복지시설, 아동복지시설, 여성복지시설, 종합복지시설 등이 추가될 수 있다)
커뮤니티	Q & A 소모임 메일링리스트/토론그룹	
멀티미디어	영상자료 사진자료	

[표 6] 웹 학술정보자원에 대한 최종 자료유형 분류체계