인문학 연구업적 데이터베이스 구축 방안

A Study on the Implementation of Database for the Research Results in Humanities

2002. 10. 24

인 문 사 회 연 구 회

인문학 연구업적 데이터베이스 구축 방안

A Study on the Implementation of Database for the Research Results in Humanities

2002. 10. 24

연구책임자 : 권기원 (성균관대학교 인문학부)
공동연구원 : 고영만 (성균관대학교 인문학부)
연구보조원 : 민혜령 (성균관대학교 정보관리연구소)
신성기 (성균관대학교 정보관리연구소)

요 약 문

1. 연구 과제명

인문학 연구업적 DB 구축 방안

2. 연구의 중요성 및 목적

인문학 분야 연구업적의 사회적 활용을 활성화시키기 위해서는 여러 기관에서 생산되는 연구업적의 체계적이고 효율적인 관리가 필요하며, 이는 데이터베이스의 분산 구축과 통합 서비스를 전제로 한다. 또한 이러한 체계를 갖추기 위해서는 연구업적 정보의 데이터베이스 입력 형식에 관한 기준과 데이터베이스 구축의 규격화가 시급한 과제이다. 그러나 현재 인문학 분야의 연구업적은 해당 기관의 독자적인 방법과 체계에 따라 관리되고 있으며 관련기관 간의 연계 또한 제대로 이루어지지 않고 있다. 본 연구는 이러한 문제점을 극복하는 방안을 찾는 것으로서, 인문학 분야 연구업적에 해당하는 여러 종류의 정보자원을 체계적으로 선별하여 그 속성들을 기술할 수 있는 메타데이터 표준 모형을 개발하고, 연구업적 원문의 데이터베이스 구축방안을 제시하는 것을 그 목적으로 한다.

3. 연구의 내용, 방법 및 범위

본 연구의 내용은 크게 세 가지로 나뉜다. 첫째는 인문학 연구업적 관리 및 데이터베이스 구축과 관련된 문제점을 파악하기 위한 기초자료 조사와 현황 분석을 하는 것이며, 둘째는 인문학 연구업적 데이터베이스 구축의 공유체계 확립을 위한 분류체계의 표준 모델을 제시하고 메타데이터 기술 요소와 원문데이터의 디지털 표준 포맷형식을 개발하는 것이고, 셋째는 인문학 연구업적 데이터베이스의 기술적 구축방안을 제시하는 것이다.

기초자료 조사 및 현황 분석에는 인문학 분야 국내 기관의 연구업적 데이터베이스 구축 현황, 디지털 포맷 형식, 운영 방안을 조사하여 비교 분석하는 방

법이 적용된다. 메타데이터 기술요소의 선정 방법은 국내외 기관의 연구업적 데이터베이스 목록 요소들을 상호 비교하여 1차 요소를 추출하고, 더블린 코어와 같은 국제적인 메타데이터 표준 요소와 상호 비교해서 2차 요소를 추출한 후 1차 요소와 2차 요소를 다시 상호 대조하여 3차 요소를 추출하는 방식으로 진행된다. 데이터베이스 구축 방안에서는 기술적 측면에서 입력, 변환, 검색 시스템 설계 모형을 제시하고 정책적 측면에서의 고려 사항을 추출하였다.

본 연구의 대상이 되는 연구업적의 범위는 인문사회연구회 소속 연구기관, 한국역사정보시스템 참여기관, 학회 등에서 발간되는 학위논문, 학술논문 및 연구보고서로 한정된다. 연구의 방향은 이러한 연구업적의 메타데이터 및 원문 데이터베이스 구축을 위한 기술적 방안을 제시하는 것이다.

4. 주요 연구 결과

4.1 인문학 연구업적 관리 현황

(1) 현황 요약

- 연구업적 정보의 생성 : 연구업적 정보의 이용자는 일반적으로 학위논문, 학술논문 및 연구보고서를 선호하고 있으며, 이 자료들은 대학과 학회 및 연구기관에서 데이터베이스를 구축하고 있으나 많은 부분이 디지털화되어 있지 않은 채로 남아 있다.
- 연구업적 정보의 축적 : 연구업적은 여러 기관에서 수집되고 축적되고 있으나 기관별 서비스로 인해서 많은 이용자들이 불편을 겪고 있으며, 인문학 분야만을 총괄해서 축적할 수 있는 조직이나 시스템이 없다.
- 연구업적 정보의 유통 및 활용 : 기관들의 시스템 관리 체계가 기초 수준이고 기관간의 정보공유가 이루어지지 않고 있다.

(2) 문제점

- 연구업적 정보의 생성 : 인문학 연구업적 정보는 기관별로 생산, 수집되고 있으며, 관련기관의 자체 예산과 기술력으로는 연구업적 정보의 완전한 디지털화가 어렵다.
- 연구업적 정보의 수집 : 연구업적 정보의 수집과 디지털화가 개별적으로 이루어지고 있으며 주관기관이 불명확하다.

- 연구업적 정보의 축적 : 생산되고 축적되는 연구업적 정보에 대한 통합관리 및 통합검색 체계가 필요하다.
- 연구업적 정보의 유통 : 연구업적 정보 생성기관과 축적기관 간의 정보 공유를 위한 연계가 미흡하다.

(3) 대안

- 연구업적 정보의 생성 : 각 기관에서 생성되는 연구업적 정보의 디지털화를 위한 예산 지원 및 생성 단계에서 디지털화의 의무화를 추진하는 것이 바람직하다.
- 연구업적 정보의 수집 : 개별적으로 수집된 연구업적 정보의 통합 및 디지털화의 주관은 해당 정보의 생성기관에서 담당하는 것이 바람직하다.
- 연구업적 정보의 축적 : 주관센터의 추진 또는 주관기관의 지정을 통한 공유, 유통체제의 확립과 통합검색을 위한 통합 메타데이터 체계를 구축해야 한다.
- 연구업적 정보의 유통 : 원활한 유통을 위한 저작권 관리체제 도입 및 기관 간 연계를 위한 지원 체제를 정립하는 것이 필요하다.

4.2 인문학 연구업적 데이터베이스 표준화 방안

(1) 인문학 정보의 분류체계 표준화

인문학 분야 정보의 분류체계 표준화를 위해 기초로 한 참고자원은 학술진흥재단의 연구분야분류표와 한국교육학술정보원의 학술분야정보분류체계표준(안), 그리고 교육인적자원부 학술분야정보분류체계(안) 등이다. 이 중 학문과 지식의 정보화를 위해 개발된 분류체계는 한국교육학술정보원에서 제안하는 인문학술분야정보분류표준안을 기본모형으로 삼아 수정, 보완하는 것이 바람직하다.

(2) 인문학 연구업적 정보의 메타데이터 표준(안) 개발

① 인문학 연구업적 DB의 기본요소 메타데이터 세트 : 각기 다른 메타데이터 세트의 요소들을 상호 매핑시켜 고유한 의미요소들을 추출하고 이를 DC 기반의 확장 스키마 형태로 정리하여 1차 요소표를 작성한 후, 여기서 선택되고 추가된 요소들 가운데 인문학 연구업적 DB의 기본요소 메타데이터 세트

로 선정된 13개의 공통 요소를 기본요소 메타데이터 세트로 삼는 표준안을 작성하였다 (표1 참조).

② 인문학 연구업적 DB의 자료유형별 특수요소 메타데이터 세트 : 학위논문, 학술지논문, 보고서의 자료유형별 특수요소로 선정된 13개의 요소를 추가한 특수요소 메타데이터 세트를 작성하였다 (표2 참조).

표 1 : 인문학 연구업적 정보의 메타데이터 일반요소 표준 모형(안)

번호	요소명	설명	기타
1	**자료 유형**	자료의 내용적 성격, 장르(자료 유형 기술)	학위논문, 일반논문, 프로시딩, 사전, 법령집, 보고서, 학술지 등
2	**자료명**	자원에 부여된 제목	버전정보 포함함
2.1	대체 자료명	자원의 공식적인 이름을 대신하는 대체 자료명(다른 언어로 번역된 제목, 간략제목 등)	
3	**작성자명**	자료를 현 상태로 작성한 사람 또는 기관의 이름	
3.1	작성자 역할	작성자의 역할	저자,번역자, 자료가공자
3.2	작성자 유형	작성자의 유형	교수, 학생, 강사/연구원/기타, 국/공립기관, 교육/연구기관, 기업/민간기관
3.3	작성자 이메일	작성자의 이메일주소	
3.4	작성자 홈페이지	작성자의 홈페이지 주소	
4	**주제명**	자원의 내용이 지닌 주제(topic)에 관한 사항 (주제 리스트에서 값을 취함)	
5	**기술사항**	자원의 내용과 관련된 설명, 요약(목록자가 작성함)	
5.1	목차	자원의 내용을 나타내는 하부 단위들의 리스트	
6	**발행자명**	자원의 발행자 및 온라인상의 배포자의 이름	
6.1	발행자 유형	발행자 유형	교수, 학생, 강사/연구원/기타, 국/공립기관, 교육/연구기관, 기업/민간기관
6.2	발행자 이메일	발행자의 이메일	
6.3	발행자 홈페이지	발행자의 홈페이지	

7	**기여자명**	작성자 요소에 명시된 개체 이외에, 자원의 내용에 대해 2차적인 책임이 있는 기타 개체의 이름	
7.1	기여자 역할	기여자의 역할	번역자, 논문수여기관, 기타
7.2	기여자 유형	기여자의 유형	교수, 학생, 강사/연구원/기타,국/ 공립기관, 교육/연구기관, 기업/민간기관
7.3	기여자 이메일	기여자의 이메일	
7.4	기여자 홈페이지	기여자의 홈페이지 주소	
8	**날짜**	날짜 종류에 대한 실제 값	--년 --월 --일 등 표현 형식 정하기
8.1	날짜의 유형	자원의 존재 기간 동안 어떠한 사건이 발생한 날짜의 유형	작성일, 발행일, 최종갱신일
9	**자료형식**	자원의 표현형식. 기술상(technical)의 데이터 유형	text : plain, html, sgml, xml 등 application : excel, postscript, pdf, powerpoint 등 image: jpeg, pjpeg, gif, tiff, png, bmp 등 audio : au, wav, x-aiff, x-midi, realaudio 등 video : mpeg, quicktime, vhs 등
10	**원정보 제목**	현재 자원의 출처가 되는 원정보자원 또는 현재의 디지털 자원을 생성하는 데에 사용되었던 출처 자원의 제목	
10.1	원정보저자	원 자원의 내용에 주된 책임을 가진 개체	
10.2	원정보 작성일	출처 자원의 생성일	
11	**언어**	자원의 지적인 내용을 기술하고 있는 언어의 이름	한국어, 영어, 일본어, 중국어, 기타
12	**관련자료 제목**	기술되고 있는 자원의 관련자료 제목	
12.1	관계 종류	기술되고 있는 자원에 대한 관련자원의 관계유형	
13	**이용조건**	자원을 이용하기 위해 요구되는 조건	

표 2 : 인문학 연구업적 정보의 메타데이터 특수요소 표준 모형(안)

자 료 유 형	요 소 명		설 명
학위논문	1	수 여 기 관	학위논문을 수여한 대학 및 기관
	2	학 위 유 형	학, 석, 박사 논문의 구분
	3	초 록	자원의 내용에 대한 요약 (영문초록, 국문초록 등)
	4	키 워 드	자원의 내용을 설명하는 주요 단어
학술지논문	1	간 기	자원의 출처가 되는 학술지의 발행 주기
	2	자료수록 범위	자원이 수록된 권, 호 정보(예: vol, no.)
	3	초 록	자원의 내용에 대한 요약 (영문초록, 국문초록 등)
	4	평 가 정 보	자원의 출처가 되는 학술지의 평가정보(예: SCI 등재 여부)
	5	식 별 자	자원의 출처가 되는 학술지 식별자(예: ISSN, DOI 등)
보고서	1	간 기	자원의 출처가 되는 보고서의 발행 주기
	2	공 개 유 무	자원발행의 공개여부(예: 회색문헌 등)
	3	초 록	자원의 내용에 대한 요약 (영문초록, 국문초록 등)
	4	식 별 자	보고서의 내부 식별자 및 외부 식별자 정보

(3) 인문학 연구업적 원문데이터의 표준화 방안

① PDF를 기반으로 하는 원문 DB 구축 방향 : 현재 구축되어 있는 원문 자료의 대부분이 PDF와 TIFF로 구축되어 있는 실정이다. 또한 향후 구축하고 희망하는 원문 자료의 포맷도 PDF와 TIFF가 다수를 차지하고 있다. 원문 포맷 구축자의 입장에서 원문 구축에 대한 경제성, 기술적 안정성, 향후 표준화 가능성 등을 고려해 본다면 PDF, TIFF 등이 효과적인 원문 포맷이 될 것이다. 이용자의 입장에서도 지명도가 높고 뷰어가 프리웨어인 PDF가 가장 편리하게 이용할 수 있는 포맷이라고 할 수 있다. 따라서 현재 구축된 많은 원문 정보의 포맷이며 경제성, 효율성, 이용자 편이성 등의 측면을 고려할 때 원문 포맷의 표준화로는 PDF 포맷 중심으로 지향하는 것이 타당한 것으로 평가된다.

② 향후 연구 방향 : XML과 PDF의 장점을 최대한 유지하며 상호 보완적인 형태로 진행되는 것이 바람직하며, 이를 위해서는 XML문서가 PDF와 호환될 수 있는 전용 XSL 및 PDF XSL-FO Processor의 개발이 필요하나 아직은 구상 단계에 머물고 있는 실정이다.

4.3 인문학 연구업적 데이터베이스 구축 방안

(1) 구축 원칙

인문학 분야의 연구업적을 대상으로 하는 디지털 데이터베이스의 구축은 기여도, 역량, 합의의 극대화를 원칙으로 삼아야 한다. 이러한 구축원칙에 대한 구체적 내용은 다음과 같다 :

① 자원생산기관의 전반적인 디지털 역량 강화
- ◦ 디지털화 비율의 향상
- ◦ 디지털화 예산의 지원
- ◦ 연구업적 정보의 온라인 서비스 확충
- ◦ 정보획득 실패 시 대체수단과의 연계성 강화

② 자원 생산기관의 연구업적 이용자 및 자원 공유체제 구축에 대한 기여도
- ◦ 연구업적 정보의 연간 및 총 생산량
- ◦ 연구업적 이용자의 선호도

◦ 연구업적 정보 연계 및 공유에 대한 의지

② 디지털화 및 그 혜택을 최대한 보장하는 네트워크 내의 합의

◦ 원문 및 관련정보의 디지털 포맷 표준화

◦ 연구업적 정보의 저작권

(2) 정책적 지원

① 관리체제의 체계화 : 인문학 분야의 DB 구축을 성공적으로 추진하기 위해서는 첫째, 집중 투자에 의한 연구업적 DB 자원을 대폭 확충하는데 있다. 인문학 분야의 디지털화 정도를 획기적으로 제고할 수 있고, 투자효과를 극대화할 수 있도록 집중적으로 투자해야 한다. 둘째, 연구업적 데이터베이스의 유통을 활성화시켜야 한다. 이는 기관별로 분산되어 있는 학술 자원과 데이터베이스를 마치 하나의 데이터베이스를 이용하는 것과 같은 시스템 환경으로 조성하여 시스템 유지관리 측면에서의 비용절감 뿐만 아니라 최종 이용자인 인문학 연구자들의 편의를 도모함으로써 자발적이고 적극적인 관심과 참여를 유도해야 한다. 셋째, 인문학 관련 지식정보자원 관리체제의 정비가 있어야 한다. 실례로 한국전산원의 국가지식정보통합검색시스템과 국가정부출연연구기관인 지식정보관리시스템인 IKIS의 관리체제를 보다 효율적으로 정비할 필요가 있다.

② 관련 기관간의 연계 : 인문학 연구업적 DB를 구축하기 위해서는 먼저 인문학 관련 지식정보자원 관리기관과의 연계를 강화해야 한다. 교육학술분야 종합정보센터로는 한국교육학술정보원과의 연계를, 역사분야 종합정보센터로는 국사편찬위원회와 연계를, 그리고 문화예술분야 종합정보센터로는 문화관광부와의 연계를 강화하는 것이 인문학 연구업적 DB를 성공적으로 구축하기 위한 필수 방안이라고 볼 수 있다.

③ 법, 제도의 정비 및 표준화 : 구축된 데이터베이스가 원활하게 관리되고 유통되기 위해서는 우선적으로 지식기반사회로의 전환을 위한 법과 제도를 정비해야 한다. 저작권이 보호되면서 디지털 자원의 공유와 유통이 활성화 될 수 있도록 관련법과 제도를 현실에 맞게 정비해야 한다. 이와 더불어 지식정보자원 관리의 표준화 및 기술개발이 시급하다고 볼 수 있다. 새로운 정보통신 환경에 적합한 분야별 메타데이터 표준화가 시급히 이루어져야 하고, 저작권보호 및 이용의 편리성, DB 구축의 경제성 등의 문제를 모두 극복할 수 있는 원문 디지털 포맷의 표준화가 빨리 정착되어야 한다.

초록

본 연구에서는 인문학 분야 연구업적의 사회적 활용을 활성화시키기 위한 인문학 분야의 기관 간 정보자원 공유의 체계화 방안과 표준화된 연구업적 데이터베이스의 구축방안이 논의되었다. 인문학 연구업적 정보의 통합 데이터베이스를 구축하기 위해서는 무엇보다도 정보분류체계, 메타데이터 형식, 원문의 디지털화 형식의 표준화가 매우 시급하며, 본 연구에서 제안하는 내용은 다음과 같다 : (1) 인문학 분야 연구업적 정보의 분류표준으로서 한국교육학술정보원에서 제안한 표준(안)을 토대로 인문학 분야의 표준안을 개발, 발전시키는 것이 바람직하다. (2) 인문학 연구업적 정보의 메타데이터 형식의 표준으로는 더블린 코아(DC)를 기반으로 연구업적 정보의 특성과 새로운 매체에의 상호 운영성을 고려한 표준안이 바람직하며, 시범적 표준(안)을 제시하였다. (3) 원문의 디지털화을 위한 표준 포맷으로는 PDF 형식이 경제적, 기술적, 인력의 운용 용이성, 저작권 보호 측면에서 타당한 것으로 평가된다. 이와 더불어 정책적으로는 인문학 관련 지식정보자원 관리체제의 체계화, 관련 기관간의 연계, 법·제도 및 표준화를 지원하는 것이 매우 필요한 것으로 분석되었다.

Abstract

In this study, we tried to find methods for the systematic implementation of database that holds the research results with various formats and from various institutions in the field of humanities. By analyzing the current situation and the tendency of the technology development, this study presents some directions for the implementation of database that holds in common research results between various institutions. Major findings are (1) the classification system, metadata format and digital full-text format of the various research results in humanities should be standardized, (2) for the standardization of classification, it is desirable to develop the stardard based on the classification system proposed by KERIS, (3) the metadata format based on DC is evaluated the best standard model in the various network environment, (4) PDF format is recommended for dizitizing the full-text relating to economic, technical, manpower, and plagiarism problems.

차례

요약문

초록

1. 서론 ... 1
  1.1 연구의 필요성 ... 1
  1.2 연구의 목적 ... 2
  1.3 연구 방법 및 범위 ... 2

2. 인문학 연구업적 관리 현황 및 문제점 ... 3
  2.1 인문학 연구업적 생산 현황 ... 3
  2.2 인문학 연구업적 수요 현황 ... 5
  2.3 인문학 데이터베이스 유통 현황 ... 6
  2.4 인문학 연구업적 디지털 포맷 현황 ... 11
  2.5 인문학 연구업적 관리의 문제점 ... 13

3. 인문학 연구업적 데이터베이스 표준화 방안 ... 15
  3.1 인문학 정보의 분류체계 표준화 ... 15
  3.2 인문학 연구업적 정보의 메타데이터 표준화 방안 ... 17
  3.3 인문학 연구업적 원문데이터의 표준화 방안 ... 28

4. 인문학 연구업적 데이터베이스 구축 방안 ... 35
  4.1 기술적 측면 ... 36
    4.1.1 입력 시스템 ... 36
    4.1.2 변환 시스템 ... 38
    4.1.3 검색 시스템 ... 39
    4.1.4 저작권관리 시스템 ... 44
  4.2 정책적 측면 ... 44
    4.2.1 관리체제의 체계화 방안 ... 44
    4.2.2 타기관과의 연계 방안 ... 45

5. 결론 ... 48

참고문헌 ... 50

부록 : 메타데이터 형식 참고 자료

1. 서론

1.1 연구의 필요성

일반적으로 지식은 목적 그 자체로 보다는 기능으로 역할을 한다. 오늘날에는 경제적, 기술적 목표에 초점을 두는 경향이 강하게 나타난다. 그러나 지식은 실용적이며 도구적인 기능 이외의 다른 기능들도 가지고 있는데, 대표적인 경우가 인문사회과학적 지식이다. 이중에서도 인문학의 탐구 대상은 인간의 사회, 문화 생활에 있어서의 가치의 세계라고 할 수 있으며, 그 기능은 인적, 문화적 자원에 대한 지식과 기초적 연구역량의 증대를 담당한다고 할 수 있다.

총체적 의미에서 문화란 어떤 집단이 지니고 있는 정신적 가치이자 전형적 생활 형식이다. 그리고 어떤 사회적 집단에서 살고 있는 모든 사람들의 창의성 촉진을 통해 문화는 그 사회의 복지 수준과 번영의 토대가 되기도 한다. 문화는 또한 개인들에게 스스로의 자주성 확보를 위한 자유영역을 형성시키려는 목표를 가지게 하여 지속적인 자기 발전을 이루게 하는 동력이기도 하다. 이와 관련하여 정보시대의 인문학에 부여되는 중요한 과제중 하나는 인문학 분야에서 생산되는 연구업적의 사회적 활용을 활성화시키는 일이며, 이는 결국 문화적 자원에 대한 기초적 연구 역량의 증대와 그 사회 문화수준의 향상을 꾀하는 것이라 할 수 있다.

인문학 분야 연구업적의 사회적 활용을 활성화시키기 위해서는 여러 기관에서 생산되는 연구업적의 체계적이고 효율적인 관리가 필요하며, 이는 데이터베이스의 분산 구축과 통합 서비스를 전제로 한다. 또한 이러한 체계를 갖추기 위해서는 연구업적 정보의 데이터베이스 입력 형식에 관한 기준과 데이터베이스 구축의 규격화가 필요하다. 그러나 현재 인문학 분야의 연구업적은 해당 기관의 독자적인 방법과 체계에 따라 수집, 관리되고 있으며 관련기관 간의 연계 또한 제대로 이루어지지 않고 있다. 인문학 분야의 연구업적 관리를 위한 데이터베이스 구축은 특히 타학문 분야에 비해서도 매우 더딘 것으로 나타나고 있다. 따라서 인문학 분야 연구업적의 정보를 기술하는데 필요한 메타데이터 요소의 개발 및 데이터베이스의 구축의 규격화 방안을 제시하는 일은 인문학의 활성화를 위해서 매우 시급하게 이루어져야 할 기초작업이라 할 수 있다.

1.2 연구의 목적 및 내용

본 연구의 목적은 인문학 분야 연구업적에 해당하는 여러 종류의 정보자원을 체계적으로 선별하여 그 속성들을 기술할 수 있는 메타데이터 표준 모형을 개발하고, 연구업적 원문의 데이터베이스 구축방안을 제시하는 것이다.

본 연구의 내용은 크게 세 가지로 나뉜다. 첫째는 인문학 연구업적 관리 및 데이터베이스 구축과 관련된 문제점을 파악하기 위한 기초자료 조사와 현황 분석을 하는 것이며, 둘째는 인문학 연구업적 데이터베이스 구축의 공유체계 확립을 위한 분류체계의 표준 모델을 제시하고 메타데이터 기술 요소와 원문데이터의 디지털 표준 포맷형식을 개발하는 것이고, 셋째는 인문학 연구업적 데이터베이스의 기술적 구축방안을 제시하는 것이다.

본 연구에서 제시되는 모형은 인문학 분야 정보자원의 관리를 담당하는 기관의 측면에서는 인문학 분야 정보자원 공유의 체계화를 앞당기는 효과를 가져올 수 있을 것이며, 이용자 측면에서는 인문학 분야 연구업적 데이터베이스의 신속하고 정확한 서지사항 검색 및 원문 이용의 활성화를 지원할 수 있을 것이다.

1.3 연구 방법 및 범위

기초자료 조사 및 현황 분석에는 인문학 분야 국내 기관의 연구업적 데이터베이스 구축 현황, 디지털 포맷 형식, 운영 방안을 조사하여 비교 분석하는 방법이 적용된다. 두 번째의 연구 내용인 메타데이터 기술요소의 선정 방법은 국내외 기관의 연구업적 데이터베이스 목록 요소들을 상호 비교하여 1차 요소를 추출하고, 더블린 코어와 같은 국제적인 메타데이터 표준 요소와 상호 비교해서 2차 요소를 추출한 후 1차 요소와 2차 요소를 다시 상호 대조하여 3차 요소를 추출하는 방식으로 진행된다. 세 번째 연구 내용인 데이터베이스 구축 방안에서는 기술적 측면에서 입력, 변환, 검색 시스템 설계 모형을 제시하고 정책적 측면에서의 고려 사항을 추출하였다.

본 연구의 대상이 되는 연구업적의 범위는 인문사회연구회 소속 연구기관, 한국역사정보시스템 참여기관, 학회 등에서 발간되는 학위논문, 학술논문 및 연구보고서로 한정되며, 연구의 방향은 이러한 연구업적의 메타데이터 및 원문 데이터베이스 구축을 위한 기술적 방안을 제시하는 것이다.

2. 인문학 연구업적 관리 현황 및 문제점

2.1 연구업적 생산 현황

대학교, 정부출현 연구기관, 학회 등에서 생산되는 인문학 연구업적 정보의 생산 현황을 연구업적의 정보유형별 연간 생산 현황, 인문사회연구회 소속 연구기관의 연구업적 정보 총 자료량, 국내 학회발간물 및 발표논문 수로 나누어 살펴보면 다음과 같다.

가. 연구업적의 정보유형별 연간 생산 현황

2002년 한국교육학술정보원의 정보화전략계획수립 완료보고서에 따르면 연구업적 정보유형별 총 정보 생산량에서는 <그림 1>에서 보는 바와 같이 대학학위논문, 학회지, 연구보고서 순으로 많이 생산되는 것으로 나타나고 있다.

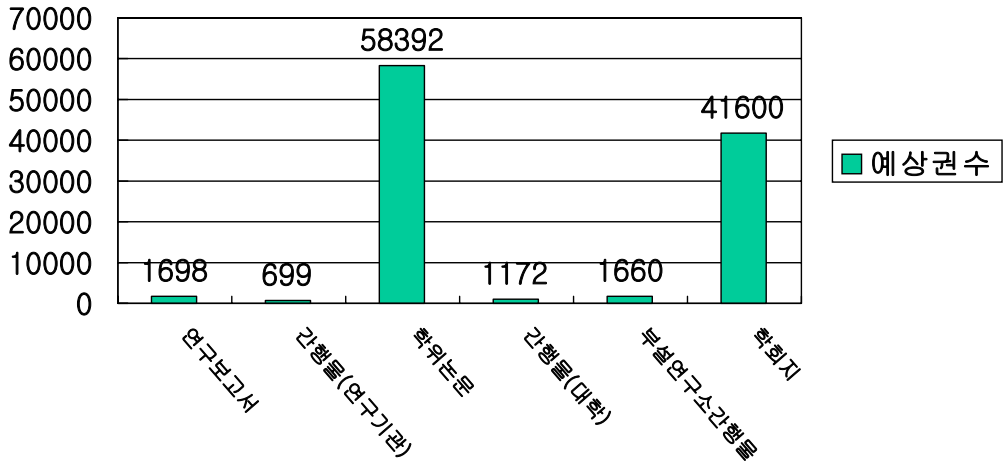


그림 1 : 정보유형별 인문학 연구업적정보 연간 생산현황 (추정치)

나. 인문사회연구회 소속 연구기관의 연구업적 정보 총 자료량

인문사회연구회 소속 9개 연구기관에 대한 연구업적 정보의 보유 자료형태 및 자료량을 조사한 결과는 <표 1>과 같다. 9개 연구기관 모두 연구업적 정보

의 주요 내용은 연구결과물이며, 보유 자료형태는 문서자료이고, 가장 많은 자료를 보유하고 있는 기관은 한국교육개발원과 한국여성개발원 순으로 조사되었다. 그리고 9개 연구기관이 보유하고 있는 총 자료량의 합계는 708,168 건으로 나타났다.

표 1 : 인문사회연구회 소속 연구기관의 연구업적정보 총 자료량

기관명(9개기관)	주요내용	자료형태	총자료량
통일연구원	연구결과물	문서자료	72,799
한국교육개발원	연구결과물	문서자료	163,327
한국교육과정평가원	연구결과물	문서자료	79,522
한국법제연구원	연구결과물	문서자료	54,124
한국여성개발원	연구결과물	문서자료	115,589
한국직업능력개발원	연구결과물	문서자료	35,401
한국청소년개발원	연구결과물	문서자료	58,952
한국행정연구원	연구결과물	문서자료	38,527
한국형사정책연구원	연구결과물	문서자료	89,927
합계			**708,168**

다. 국내 학회발간물 및 발표논문 수

국내 학문 분야별 학회발간물 및 발표논문수를 조사한 결과 인문분야에서는 <표 2>에서 보는 바와 같이 406개의 학회에서 237개의 정기간행물이 발간되며, 국내학술회의 발표논문수는 213개, 국제학술회의 발표논문수는 159개로 나타났다.

표 2 : 국내학회 발간물 및 발표논문 수 (한국학술진흥재단 학회 통계, 2002년 7월 현재)

분 야	학회수	정기간행물 발간물수	국내학술회의 발표논문수	국제학술회의 발표논문수
인 문	406	237	213	159
사회과학	448	208	213	141
자연과학	103	65	65	49
공 학	155	94	99	65
의약 학	158	78	83	45
농수해양	54	25	26	17
예술체육	107	33	33	27
복합 학	29	14	11	10
합 계	**1,459**	**754**	**743**	**513**

2.2 연구업적 수요 현황

2002년 한국교육학술정보원의 정보화전략계획수립 완료보고서에 나타난 연구업적 정보의 이용자 수요 현황을 분석하면 다음과 같다.

가. 선호하는 연구업적 유형

연구업적 정보 이용자 중 전체의 22%가 국내학위논문, 20%가 국내학술논문, 14%씩 외국학술논문, 단행본 자료의 순으로 선호하는 것으로 나타났다. 이들 4개 유형이 전체 선호하는 정보의 과반수를 넘게 차지하고 있다.

나. 선호하는 국내 학술 정보 사이트

이용자 중 전체의 22%가 RISS, 19.9%가 소속 기관 웹사이트, 19.1%가 국회도서관, 14.3%가 국립중앙도서관을 국내 학술정보 사이트로 선호하는 것으로 나타났다.

다. 소속 기관에서 입수하지 못한 자료 획득방법

이용자 중 전체의 37.1%가 일반 웹 검색 사이트에서, 23.9%가 일반 학술정보 서비스 사이트에서, 15.7%가 국회도서관에서 자료를 획득하는 것으로 나타났다.

라. 시급히 디지털화가 진행되어야 할 분야

시급히 디지털화가 진행되어야 할 분야로 23.8%가 국내학위논문을, 22.1%가 국내학술논문을 지적하고 있는 것으로 나타났다.

인문학 연구업적 정보 이용자에 대한 수요현황 조사결과를 종합하면 가장 선호하는 연구업적 유형으로는 국내학위논문과 국내학술논문으로 나타나고 있으며, 시급히 디지털화가 진행되어야 할 분야 또한 국내학위논문과 국내학술논문이라고 나타나 이러한 유형의 데이터베이스에 대한 이용자의 요구가 큰 것을 알 수 있다. 또한 소속 기관 웹 사이트를 통해 자료를 입수하는 경우가 비교적 높은 것으로 나타나 소속 기관의 연구업적 데이터베이스의 구축이 시급함을 알 수 있다.

2.3 인문학 데이터베이스 유통 현황

인문학 데이터베이스 유통 현황을 인문학 관련 지식정보 유통기관과 국내 학술논문의 원문 서비스 현황으로 나누어 분석한 결과는 다음과 같다.

가. 인문학 관련 지식정보 유통기관

인문학 관련 지식정보 유통기관에는 (1) 인문사회연구회, (2) 한국역사분야정보화추진위원회, (3) 한국역사정보시스템 참여기관, (4) 교육학술정보원, (5) 문화관광부, (6) 국가정보통합검색시스템 (한국전산원) 등이 있으며 각 기관별 특징 및 주요 DB 사업내용을 조사한 결과는 다음과 같다.

(1) 인문사회연구회 소속 연구기관

총리실 산하 인문사회연구회에는 9개의 정부출연 연구기관이 소속되어 있으며, 이들 기관에서의 인문학 관련 대표적인 지식정보 유통기관인 인문사회연구회 소속 9개 연구기관의 메타데이터 개발은 <표 3>에서와 같이 모두 MARC 포맷이 표준 역할을 하는 것으로 나타나고 있다. 원문 데이터베이스 개발은 PDF, HTML, TXT, HWP 등의 포맷이 혼재되어 표준적 역할을 하는 포맷이 없는 것으로 조사되었으며, 서비스 형태는 대부분의 연구기관이 소장도서목록과 서지의 상세정보를 제공하고 있는 것으로 나타났다.

표 3 : 인문사회연구회 메타DB 및 원문 DB 디지털 포맷 현황

기관명(9개기관)	메타DB개발(포맷)	원문 DB(포맷)	서비스형태
통일연구원	MARC	구축(XML)	소장도서목록 정간물기사색인
한국교육개발원	MARC	구축(hwp, txt, zip, jpg)	소장자료목록
한국교육과정평가원	MARC	구축(PDF)	연구결과물초록집
한국법제연구원	MARC	구축(hwp)	간행물상세정보
한국여성개발원	MARC	일부구축 (html, hwp, pdf)	상세정보
한국직업능력개발원	MARC	구축(PDF, hwp)	서지상세정보
한국청소년개발원	MARC	구축(html)	소장자료목록 서지상세정보
한국행정연구원	MARC	일부 구축(PDF, hwp)	서지상세정보
한국형사정책연구원	MARC	일부 구축(txt, html)	서지상세정보

(2) 한국역사분야정보화추진위원회

한국역사분야정보화추진위원회는 한국 역사 분야의 공통 관심과 업무의 협의 조정을 통한 디지털화의 중복방지 및 표준화를 추진하고 한국역사분야의 정보화추진을 위한 중장기 계획을 수립하기 위해 구성되었으며, 여기에는 한국역사정보시스템 참여기관(국사편찬위원회, 민족문화추진위원회, 서울대 규장각, 정신문화연구원, 한국국학진흥원), 고려대 아세아문제연구소, 국립중앙도서관, 전쟁기념사업회, 여성개발원, 독립기념관 등이 참여하고 있다. 그러나 인센티브나 강제규정이 없어서 기관간의 원활한 협조를 위한 제도적인 장치가 필요한 실정이며, 현재 본격적으로 가동되지 않고 있다.

(3) 한국역사정보시스템 참여기관

한국역사정보시스템은 “지식정보자원관리사업”의 일환으로 구축되는 시스템으로서, 여기에는 국사편찬위원회, 민족문화추진위원회, 서울대 규장각, 한국정신문화연구원, 한국국학진흥원 등의 5개 기관이 참여하고 있다. 참여기관 별 데이터베이스 구축현황은 다음과 같다.

① 국사편찬위원회 : 사료적 가치가 높은 원본소장자료와 위원회가 엄선편찬한 기간행물, 한국사연표광장, 그리고 보존상태가 열악한 근현대 자료의 원문 DB화를 추진 중에 있다.
② 민족문화추진위원회 : 고전국역총서와 한국문집총간의 주요논자 1268명의 문집에 대한 원문의 DB화를 추진 중에 있다.
③ 서울대 규장각 : 국내 유일본 중 사료가치가 높고 국내외 한국학자에게 학술적 이용가치가 높은 자료를 우선적으로 원문 DB화 추진 중에 있다.
④ 한국정신문화연구원 : 이용빈도가 높은 유일본과 학술적 이용가치가 높은 자료, 그리고 타기관에서 구축되지 않은 자료, 상업성은 없으나 활용가치가 높은 자료의 원문 DB화를 추진 중에 있다.
⑤ 한국국학진흥원 : 안동문화권을 중심으로 한 유교관련 고서적의 발굴과 수집을 통한 원문 DB화를 추진 중에 있다.

한국역사정보시스템의 문제점으로는 우선 참여기관의 역할에 대한 명시적 규정이 없는 점을 들 수 있다. 기술적 측면의 문제점으로는 참여기관의 통합검색은 가능하지만 시스템 초기화면의 경우 참여기관의 DB 명칭을 그대로 나열하고 있는 수준에 그치고 있으며, 기관에 따라 상이한 소장자료와 상이한 분류체계를 가지고 있으므로 통합검색의 효용성이 떨어지는 점이다. 또한 대부분의 시스템에서 목차, 서지사항, 해제 정도의 서비스를 제공하고 있으며 원문 서비스의 제공은 매우 저조하며 이미지와 텍스트 형식이 주를 이루고 있다. 전산원에서 실시한 이용자 만족도 조사에 의하면 한국역사정보시스템에 축적되는 정보의 유용성에 대한 만족도는 높게 나타나고 있으나, 정보의 최신성 측면에서는 이용자 만족도가 가장 낮게 나타나고 있다.

(4) 한국교육학술정보원(KERIS) : 국가학술연구데이터베이스

한국교육학술정보원 역시 지식정보자원관리사업의 일환으로 “국가학술연구데

이터베이스"를 구축하였다. 이 데이터베이스에는 인문학 분야의 업구업적물이 포함되어 있으며, 자원의 내용과 사업 내용은 <표 4>와 같다.

표 4 : 교육학술정보원 : 국가학술연구 DB

자원 명칭	보유기관	지정자원	자원의 내용
국가학술 연구DB	한국교육 학술정보원	학술지<목록유, 권호소장유>/학술논문<목록유, 원문PDF>/교육학술동영상자료유/논문참고문헌DB	국내외에서 생산되는 학술연구정보의 목록 및 원문, 학술지, 학술논문(대학부설연구소 • 대학 연구논총 등), 교육학술동영상자료, 논문참고문헌, 해외DB통합색인

과 제 명	주관기관	소요예산	사업내용
국가학술연구 DB 구축 2차사업	한국교육 학술정보원	24.4 (억원)	국내외 학술연구정보를 망라하는 통합검색 및 원문제공 서비스 구축 및 국내대학 석박사 학위논문 및 학술세미나 정보 통합 서비스 기반마련
정부출연연구기관 지식정보 DB구축 3차사업	산업연구원	21.6 (억원)	국무조정실 산하 경제사회연구회 14개 및 인문사회연구회 9개 연구기관의 연구결과물의 지식정보 DB를 구축하여 전국민에게 전문웹서비스를 제공

(5) 문화관광부 : 문화예술종합정보시스템

문화관광부에서 지식정보자원관리(3차)사업으로 추진하고 있는 문화예술종합정보시스템 구축 사업의 경우 인문학 분야, 특히 문화예술 분야의 연구업적물에 대한 디지털화 작업을 하고 있으며, 사업 내용은 <표 5>와 같다.

표 5 : 문화관광부 : 문화예술종합정보시스템

자원 명칭	보유기관	지 정 자 원	자원의 내용
문화예술 종합정보	문화 관광부	공연,전시정보/문화예술교육자료/실물공연용품/무대용품디자인/사진촬영미술정보/문화상품기획자료/문화예술정책자료/문화통계/북한문화예술정보/한국영화자료/기타문화자료	국립극장, 국립현대미술관, 국립국악원,문예진흥원, 문화정책개발원, 예술의 전당, 한국미술협회, 한국영상자료원 등이 소장하고 있는 자료

과 제 명	주관기관	소요예산	사 업 내 용
국가문화유산종합정보시스템구축(3차)사업	문화 관광부	35.7 (억원)	전국박물관에 소장되어 있는 유물 및 국보, 보물, 주요 지정문화재 등 소중한 국가자산인 문화유산정보를 지식DB로 구축하여 범국가적인 통합관리 및 종합서비스 체계마련
문화예술 종합정보 시스템구축 (3차)사업	문화 관광부	15 (억원)	문화예술분야의 핵심테마를 중심으로 하는 가치정보와 대표 문화예술기관의 추가적인 기반정보 및 고품질 문화예술정보 포털을 구축

(6) 한국전산원 : 국가정보통합검색시스템

한국전산원에서는 지식정보자원관리사업으로 구축되는 데이터베이스의 통합검색시스템을 운영하고 있으며, 여기에는 과학기술, 문화, 역사, 정보통신 분야에서 구축된 시스템들이 연계되어 있다.

표 6 : 국가정보통합검색시스템 및 연계 시스템

구분	구축 내용	인터넷 주소	주관 기관
통합검색	국가통합검색시스템	www.knowledge.go.kr	한국전산원
과학기술	과학기술종합정보시스템	www.kisti.re.kr	한국과학기술정보연구원
교육학술	학술연구정보서비스	www.riss4u.net	한국교육학술정보원
문화	문화정보통합검색시스템	www.culture.go.kr	문화정책개발원
역사	한국역사정보통합시스템	www.koreanhistory.or.kr	국사편찬위원회
정보통신	정보통신통합정보시스템	www.itfind.or.kr	한국전자통신연구원

나. 국내 학술논문의 원문 서비스 현황 (2002년 8월 현재, 추정량)

2002년 한국교육학술정보원의 조사에 따르면 국내 학술논문의 원문서비스 현황은 <표 7>과 같이 나타나고 있다.

표 7 : 국내 학술논문의 원문서비스

서비스 연도	구분	학회	대학	계	비고
2002년	원문구축	119기관 138종	435기관 771종	554기관 909종 258,940편	무료
	원문연계	488기관 570종	63기관 67종	551기관 637종 300,000편	유료 (민간기관과 연계제공)
	계	607기관 708종	498기관 838종	1,105기관 1,546종 **458,940편**	국내 학회지40%원문제공

2.4 연구업적 디지털 포맷 현황

한국교육학술정보원이 2002년 국내연구기관 및 대학을 대상으로 조사한 연구업적의 디지털화에 관해 조사한 내용을 연구업적 디지털 포맷 현황과 국내 학술논문의 원문 디지털화 현황으로 나누어 분석하면 다음과 같다.

가. 인문학 연구업적 분야 디지털 포맷 현황

정보생산기관을 대상으로 연구업적의 원문 및 메타데이터의 디지털 포맷 현황을 조사한 결과는 <표 8>과 같다. 원문 포맷의 경우 정부출연 연구기관은 PDF, HTML, TXT의 3가지 포맷이 분할 구도를 형성하고 있지만, 대학도서관에서는 PDF가 약 85%를 차지하면서 나름대로 표준의 역할을 수행하고 있다. 메타데이터(목차, 초록 등 포함)의 경우 정부출연 연구기관과 대학도서관 모두

메타데이터로는 MARC를 가장 많이 사용하는 것으로 나타났다.

표 8 : 연구업적 디지털 포맷 현황

- **정부출연 연구기관**
  - 메타데이터 DB
    - MARC가 표준역할
  - 원문 DB
    - 원문 PDF, HTML, TXT, HWP 포맷이 혼재되어
    - 표준적 역할을 하는 포맷이 없는 실정
- **대학**
  - 메타데이터 DB
    - MARC 86%, 자체개발 12%, DC 2%
  - 원문 DB
    - PDF 84%, TIFF 13%
    - 원문 포맷의 경우 대학은 PDF가 표준 역할을 하고 있음

나. 국내 학술논문의 원문 디지털화 현황 (2002년 8월 현재 추정량)

국내학술논문의 원문을 디지털화하는 곳은 한국교육학술정보원(KERIS), 한국과학기술정보원(KISTI), 국회도서관, 한국학술정보, 교보문고/누리미디어, 페이퍼리스 등 6개 기관이다. 이중 한국학술정보, 교보문고/누리미디어, 페이퍼리스는 상업적 기관으로서 학술논문의 디지털화를 위한 판권을 취득하여 상업적 목적으로 디지털화한 원문을 상업적으로 판매하고 있다. 디지털 포맷으로는 PDF와 TIFF가 주를 이루며, XML 포맷은 한국교육학술정보원에서 시범적 사업을 수행하고 있다.

표 9 : 국내 학술논문 원문 디지털화 현황

기관명	학술지	학술논문	원문포맷	비고
KERIS	814종	130,000 편	PDF, XML	인터넷서비스
KISTI	250종	200,000편	TIFF	인터넷서비스
국회도서관	164종	190,000편	TIFF	관내서비스
한국학술정보	900종	500,000편	PDF	판매용
교보문고/ 누리미디어	150종	45,000편	TIFF	판매용
페이퍼리스	90종	27,000편	PDF	판매용

2.5 인문학 연구업적 관리의 문제점

가. 일반적

(1) 현황 요약

① 연구업적 정보의 생성 : 연구업적 정보의 이용자는 일반적으로 학위논문, 학술논문 및 연구보고서를 선호하고 있으며, 이 자료들은 대학과 학회 및 연구기관에서 데이터베이스를 구축하고 있으나 많은 부분이 디지털화되어 있지 않은 채로 남아 있다.
② 연구업적 정보의 축적 : 연구업적은 여러 기관에서 수집, 축적되고 있으나 기관별 서비스로 인해서 많은 이용자들이 불편을 겪고 있으며, 인문학 분야만을 총괄해서 축적할 수 있는 조직이나 시스템이 없다.
③ 연구업적 정보의 유통 및 활용 : 기관들의 시스템 관리 체계가 기초 수준이고 기관간의 정보공유가 이루어지지 않고 있다.

(2) 문제점

① 연구업적 정보의 생성 : 인문학 연구업적 정보는 기관별로 생산되고 수집되고 있으며, 관련기관의 자체 예산과 기술력으로는 연구업적 정보의 완전한 디지털화가 어렵다.
② 연구업적 정보의 수집 : 연구업적 정보의 수집과 디지털화가 개별적으로 이루어지고 있으며 주관기관이 불명확하다.
③ 연구업적 정보의 축적 : 정보 수요자에 대하여 생산되고 축적되는 연구업적 정보에 대한 통합관리 및 통합검색 체계가 필요하다.
④ 연구업적 정보의 유통 : 연구업적 정보 생성기관과 축적기관 간의 정보 공유를 위한 연계가 미흡하다.

(3) 대안

① 연구업적 정보의 생성 : 각 기관에서 생성되는 연구업적 정보의 디지털화를 위한 예산 지원 및 생성 단계에서 디지털화의 의무화를 추진하는 것이 바람직하다.
② 연구업적 정보의 수집 : 개별적으로 수집된 연구업적 정보의 통합 및 디지털화의 주관은 해당 정보의 생성기관이 담당하는 것이 바람직하다.
③ 연구업적 정보의 축적 : 주관센터의 추진 또는 주관기관의 지정을 통한 공유, 유통체제의 확립과 통합검색을 위한 통합 메타데이터 체계를 구축해야 한다.
④ 연구업적 정보의 유통 : 원활한 유통을 위한 저작권 관리체제 도입 및 기관 간 연계를 위한 지원체제를 정립하는 것이 필요하다.

3. 인문학 연구업적 데이터베이스 표준화 방안

인문학 분야 연구업적 데이터베이스를 구축하기에 앞서 반드시 선행되어야 할 작업은 인문학 정보의 분류체계 표준화와 인문학 연구업적 정보의 메타데이터 표준화 방안 및 인문학 연구업적 원문데이터의 표준화 방안이라는 세 분야에 대한 표준화 작업이다. 이는 현재 분산 구축되어 있는 연구업적 정보의 통합 구축을 위한 필수 선행 단계라고 할 수 있다. 따라서 이들 세 분야의 현황 분석을 하면 다음과 같다.

3.1 인문학 정보의 분류체계 표준화

가. 개요

정보의 다양화와 정보 통신 환경의 급격한 변화에 대응하기 위해서는 정보분류체계의 표준화가 선결과제이다. 이는 향후 데이터베이스 구축 과정과 이후에 이루어질 통합검색시스템 등의 구축 시에 매우 중요한 선결 작업이 될 것이며, 추후 인문학 DB의 주제별 검색, 디렉토리 서비스 제공 등 이용자 중심의 정보서비스를 제공함에 있어서도 매우 중요한 역할을 하게 될 것이다. 정보의 분류체계 표준화를 위해서는 다음과 같은 지침을 참고하여 적용해야 할 것이다.

- ○ 이용자의 편이성을 우선시하는 서비스에 적합한 분류체계 마련 : 체계적인 분류 기준을 마련하여 이용자가 손쉽게 정보에 접근하고 이용하도록 유도
- ○ 정보를 효과적이고 경제적으로 관리할 수 있는 분류체계
- ○ 다양한 인문학 연구업적 정보를 포함하기 위해 기존의 학문분류체계를 응용한 분류체계
- ○ 향후 새로이 출현하는 인문학 분야의 지식을 수용할 수 있도록 최신성 프로세스가 부가된 분류체계

나. 정보의 분류체계 표준화 현황 분석

현재 인쇄매체뿐만 아니라 디지털 자원까지 포괄할 수 있는 정보분류체계의 현황은 주제 분류의 체계화가 이루어지지 않고 있는 것으로 나타났다. 기존의 분류체계는 대부분 인쇄매체의 분류를 위한 체계라 할 수 있으며, 새로운 매체

를 포괄하는 분류 역시 각 기관마다 다른 체계를 사용하거나 아니면 자체적으로 기존 분류체계를 가감하여 사용하고 있다. 도서관의 경우 양서는 DDC, 동서는 KDC를 사용하는 예가 많으며, 일부에서는 LCC 등을 사용하기도 하고, 자체적으로 분류체계를 개발하여 사용하고 있는 사례도 있다. 따라서 인쇄매체만이 아니라 디지털 자원 등 다양한 정보 매체를 수용할 수 있으며 인문학 분야의 표준으로 정립될 수 있는 정보분류체계의 개발이 필요하다고 할 수 있다.

다. 정보의 분류체계 표준화 방안

인문학 분야 정보의 분류체계 표준화를 위해 기초로 한 참고자원은 학술진흥재단의 연구분야분류표와 한국교육학술정보원의 학술분야정보분류체계표준(안), 그리고 교육인적자원부 학술분야정보분류체계(안) 등이다. 이 중 학문과 지식의 정보화를 위해 개발된 분류체계는 한국교육학술정보원에서 제안하는 인문학술분야정보분류표준안이다. 따라서 인문학 정보분류체계의 표준은 또다른 표준화 작업이 가져오는 표준안의 중복을 방지하기 위해 한국교육학술정보원의 안을 기본모형으로 삼아 이를 정련하는 것이 바람직하다.

표 10 : KERIS 인문학술분야 정보분류표준안

철학/종교/윤리	
	동양철학
	서양철학
	종교학
예술	
	체육학
	무용
	국악
	음악
	미술
	연극영화
	기타(예술체육)
문화	
	인류학
	민속학

언어/문학	
	언어학
	국어학
	국문학
	중어중문학
	일어일문학
	불문학
	독어학
	독문학
	영어학
	영문학
	인도어문학
	아랍어문학
	마인어문학
	태국어문학
	베트남어문학
	기타동양어문학
	서반어문
	노어노문
	이탈리어문
	기타제어문

역사	
	한국사
	동양사
	서양사
기타	

3.2 인문학 연구업적 정보의 메타데이터 표준화 방안

가. 개요

현재 각 기관마다 사용하고 있는 메타데이터 형식은 통일되어 있지 않다. 대부분의 메타데이터가 MARC/KORMARC를 기반으로 작성되고 있으나 최근 DC 기반의 메타데이터 또는 자체 개발 메타데이터를 사용하고 있는 기관이 증가하고 있어서 메타데이터의 혼재에 따른 중복과 유사의 문제점이 대두되고 있다. 불필요한 노력과 투자를 방지하기 위해서는 상호 호환이 가능한 메타데이터의 표준화가 시급한 실정이다.

전통적으로 메타데이터는 데이터에 대한 데이터(Data of Data)라는 의미로 정의되지만, 오늘날에는 주로 네트워크 상의 디지털 형태의 정보 자원에 대한 소재나 내용을 기술(Description)하고 또 네트워크 환경에서 이를 탐색하고 이용하기 위한 요소들의 집합이라는 의미로 통용되고 있다. 도서관 환경에서의 정보자원을 관리하기 위한 방법으로서 AACR, KCR, ISBD 등의 목록규칙이 개발되었으며, 이들 규칙을 토대로 정보자원을 자동적으로 관리하고자 하는 노력이 진행되었다. 그 결과 MARC가 개발되었으며, 현재 대다수의 도서관에서는 정보자원을 관리하기 위한 시스템의 기반기술로서 MARC를 활용하고 있다.

MARC와 같은 메타데이터가 지니는 주요한 기능의 하나는 탐색하고자 하는 정보자원을 적절하게 검색할 수 있도록 해주는 것이며, 다른 하나는 검색된 자원 중에서 서로 연관성이 있는 정보를 그룹화하여 보여줄 수 있도록 하는 것이다. 그러나 MARC는 책과 같은 전통적인 정보자원의 기술에는 적합할 수 있으나 필름과 같은 특수한 형태의 정보자원이나 네트워크 상의 정보자원이 지니는 다양한 특성을 기술하기 어려우며, 또한 정보자원 간의 연결관계를 제대로 기술할 수 없는 한계성을 지니고 있다. 정보자원의 전달 매체와 그 형태가 다양해지면서 이에 대한 관리 방법도 다양해져 가고 있기 때문이다. MARC 외에도 현존하는 목록 규칙들은 특히 정보자원의 연계와 그룹화를 제대로 구현하는데 있어서 많은 문제점을 보이고 있다.

메타데이터 모델링 분야의 연구는 지속적으로 확산되고 있으며, 특히 인터넷의 확산과 더불어 상호운영성 문제가 새로운 연구분야로 주목받고 있다. 인문학 연구업적 메타데이터 표준화의 방향은 새로운 매체와 관련된 상호운영성의 측면에서 볼 때 많은 장점을 가지고 있는 DC에 기반한 메타데이터 구조를 도입하는 것이 바람직하다. DC 기반의 메타데이터가 지니는 장점은 다음과 같다:

첫째는 단순성(Simplicity)으로 더블린 코어는 자원 기술에 있어 비전문가도 사용할 수 있도록 간단하게 설계되었다. 대부분의 요소는 도서관 목록카드의 개략적인 이해를 바탕으로 하고 있다. 둘째는 의미론적 상호운용성으로 인터넷 자원들이 모두 서로 다른 기술형식에 의해 표현되면 일관된 탐색이 어렵고 검색효율이 떨어진다. 더블린 코어는 일반적으로 의미론적 상호운용성을 증가시키도록 서로 다른 데이터의 내용 표준을 단일화하는 것을 지원한다. 셋째는 국제적 합의로서 더블린 코어는 남미, 유럽, 오스트리아 및 아시아 20여개 국가가 참여하고 있는 국제적인 합의에 의한 표준이다. 넷째는 확장성으로 더블린코어는 도서관계의 MARC 목록과 같은 기술 모델을 더욱 정교하게 만드는 경제적인 대안을 제공하는 등 확장의 가능성이 크다. 다섯째는 웹 메타데이터 모듈을 제공한다는 점을 들 수 있다.

나. 인문학 연구업적 DB의 메타데이터 요소 분석과정

(1) 국내 기관의 연구업적 DB의 목록 요소 비교 분석

국내 인문사회연구회 산하 9개 기관에서 제공하고 있는 DB의 목록 요소를 나열・비교・분석하여 불필요한 요소를 배제하고 공통 요소와 필수 요소를 선정하고 추출하였다.

(2) 해외 기관의 연구업적 DB의 목록 요소 비교 분석

국제적으로 해외에서 널리 통용되고 있는 관련 해외기관 인문학 분야 DB를 선정하여 목록 요소를 나열・비교・분석한 다음, 불필요한 요소는 배제시키며 필수 요소를 선정하고 추출하였다.

(3) 더블린코어(Dublin Core)를 비롯한 각종 메타데이터 세트의 요소 분석

일반적인 인문학 분야 학술자원의 기술을 위해 사용되는 요소들은 무엇인지를 조사한 후 공통적인 요소들이 무엇인지, 여러 메타데이터 세트에서 공통적으로 고유한 의미로 사용되는 요소들은 무엇인지를 조사하는 것은 매우 중요하다. 이를 위해서는 타 메타데이터 세트에 대한 요소들을 나열하고, 요소들을 조사 및 정리하는 것이 필수적이라고 할 수 있다. 인문학 분야 학술자원에 대

한 DB 구축에서 이러한 과정은 향후 다른 학술정보자원에 대한 DB들과의 연계성 및 상호 운용성을 높여주는 역할을 할 것으로 기대된다. 본 연구에서 메타데이터 요소분석에 참고한 메타데이터 리스트는 표 11과 같다.

표 11 : 메타데이터 형식 참고 리스트

스키마 이름	국가	분야
Dublin Core	세계적	일반 웹 자원
AHDS : Arts & Humanities Data Service	세계적	고고학, 역사학, 원문연구, 행위예술, 시각예술
IEEE Learning Object Metadata (3.6)	세계적	교육자료
GEM Element Set and Profile(s)	미국	교육자료
FIU DL Metadata	미국	일반 웹 자원
KERIS 학술논문 메타데이터	미국	학술논문 자료
KERIS 학술지 메타데이터	미국	학술지

① Dublin Core : 웹 자원을 기술하기 위한 대표적인 메타데이터 세트로서 인문사회연구회 산하기관의 DB 목록 요소 및 타 메타데이터 세트와의 매핑을 위한 기준 도구로서 사용하였다.

② AHDS(Arts & Humanities Data Service) : AHDS는 Dublin Core를 기반으로 한 프로젝트 수행 사례로서 중앙 행정 부서와 5개의 서비스 부서로 구성된 연합기관으로 고고학, 역사학, 원문연구 그리고 행위 예술과 시각 예술 분야의 전자 자료를 서비스하고 있다. AHDS는 Dublin Core기반의 메타데이터 세트를 통하여 각 서비스 부서가 제공하는 전자 자료들을 모두 이용할 수 있도록 하는 통합시스템의 구축을 지원하는 것이다.

③ IEEE Learning Object Metadata (LOM) : 급속히 증가하는 교육 분야의 디지털 자원의 효율적인 관리와 검색의 용이성 증대를 위해 IEEE에 의해 개발된 메타데이터 세트이다. LOM은 다양한 형태의 교육 자료들에 필요한 최소한의 속성만을 정의함으로써 폭 넓은 적용성과 확장성을 갖도록 하는 데

에 초점을 두었다. 특히 교육 분야의 디지털 자원이 급속도로 증가하고 있는 상황에서 그에 대한 정보나 메타데이터 없이는 자원의 관리와 검색에 많은 어려움이 있기 때문에 IEEE는 이러한 문제점을 해결하고자 LOM을 만들게 되었다.

④ GEM Element Set and Profile(s) : 미국 교육부의 지원을 받아 ERIC Clearinghouse와 Syracuse 대학 문헌정보학과 등이 공동으로 시작한 교육 정보의 활용을 위해 시작된 프로젝트로서 인터넷을 통해 이용할 수 있는 강의 계획서, 커리큘럼 유닛, 그리고 그 밖의 다른 교육 자료들에 대한 이용자의 접근성을 강화하기 위하여 시작되었으며 인터넷 상의 모든 교육 자료에 대한 one-stop, any stop access를 실현하기 위한 목표를 가지고 있다.

⑤ FIU(Florida International University) Digital Library : FIU의 전자도서관 프로젝트에서는 Dublin Core를 기반으로 하되 몇 개의 자체적인 요소와 한정어를 추가하는 방식으로 새로운 메타데이터 세트를 만들었다.

⑥ KERIS 메타데이터 : 한국교육학술정보원 KERIS에서 교육자료에 대한 효율적인 메타데이터 구축을 위해 개발한 메타데이터 세트로서 학술지 메타데이터, 학술논문 메타데이터, 멀티미디어 교육 자료 메타데이터의 세가지 세트를 구분하여 개발하였으며, 본 연구에서는 학술지 메타데이터와 학술논문 메타데이터 세트를 참조하였다.

다. 인문학 연구업적 DB의 표준 메타데이터 요소 추출과정

① 국내 인문사회 연구회 산하기관의 DB 목록 요소를 나열하여 DC의 제 요소와 비교하였다. 인문사회연구회 산하 9개 기관에서 기 구축되어 운영중인 DB의 메타데이터 요소를 조사해 본 결과 체계적인 메타데이터 요소의 적용이 이루어지지 않음은 물론 각 기관간의 비표준의 문제가 심각하여 통합 DB운용에 상당한 제한 사항으로 지적되고 있다. 각 기관의 메타데이터 요소의 적용 정도를 Dublin Core 요소를 기준으로 정리하면 <표 12>와 같다.

표 12 : 국내 기관별 메타데이터 요소 적용 (Dublin Core 기준)

Dublin Core 요소	요소별 적용기관	적용 기관수
Title	전 기관	9
Creator	전 기관	9
Subject	한국형사정책연구원, 통일연구원, 한국행정연구원	3
Discription	한국행정연구원, 한국법제연구원	2
Publisher	한국형사정책연구원, 한국행정연구원, 한국직업능력개발원, 통일연구원, 한국법제연구원	5
Contributor	한국직업능력개발원	1
Type	전 기관	9
Date	한국교육과정평가원, 한국형사정책연구원, 한국행정연구원, 한국직업능력개발원, 통일연구원, 한국법제연구원, 한국교육개발원	7
Format	한국교육과정평가원, 한국직업능력개발원, 한국교육개발원, 한국법제연구원	4
Identifier	한국형사정책연구원, 한국직업능력개발원, 통일연구원, 한국법제연구원	4
Source	없음	0
Language	없음	0
Relation	없음	0
Coverage	없음	0
Rights	없음	0

② 해외에서 통용되고 있는 관련 메타데이터 스키마들을 조사하여 각각의 요소를 정리하였다. AHDS, IEEE, GEM, FIU DL 등의 해외 메타데이터 스키마 각각의 메타데이터 요소들을 Dublin Core 제 요소를 기준으로 매핑하여 정

리한 결과는 <표 13>과 같다.

표 13 : 해외 메타데이터 요소 매핑

Dublin Core	AHDS	LOM	GEM	FIU DL	KERIS	
					학술논문	학술지
Title	Title	Title	Title	Title	Title	Title
Creator	Creator	Contribute	Creator	Creator	Creator	Creator
Subject	Subject	Classification	Subject	Subject	Subject	Subject
Discription	Discription	Discription	Discription	Discription	Discription	Discription
Publisher	Publisher	Contribute	Publisher	Publisher	Publisher	Publisher
Contirbutor	Contirbutor		Contirbutor	Contirbutor	Contirbutor	Contirbutor
Type	Type		Type	Resource Type	Type	Type
Date	Date		Date		Date	Date
Format	Format	Format	Format	Format	Format	Format
Identifier	Identifier	Identifier	Identifier		Identifier	Identifier
Source	Source		Source	Source	Source	Source
Language	Language	Language	Language	Language	Language	Language
Relation	Relation	Relation	Relation	Relation		Relation
Coverage	Coverage	Coverage	Coverage	Coverage place Coverage time		Coverage
Rights	Rights	Rights	Rights	Access Ristrictions		

③ 각기 다른 메타데이터 세트의 요소들을 상호 매핑시켜 고유한 의미요소를 추출하였으며, 이를 DC 기반의 확장 스키마 형태로 정리하여 1차 요소표를

작성하였다.

- 앞의 (1) 과 (2)에서 조사·분석하여 추출된 요소들을 상호 매핑하여 각 메타데이터 세트에서 동일한 의미를 가진 요소들을 추출하고, 이를 그룹화하여 DC 기반의 확장 스키마 형태로 정리하여 1차 요소표를 작성하였다.

- 공통요소 이외의 특수한 요소들은 그 개념을 최대한 살려 요소들을 그룹화하였으며, 이렇게 추출된 요소들은 Dublin Core의 틀에 맞추어 상위요소와 하위요소로 구분하여 나열하였고, Dublin Core에 없는 특수한 요소들은 상위요소 또는 하위요소로 추가하는 방식으로 요소표를 작성하였다.

- 메타데이터 요소표 및 각각의 요소의 출처는 <표 14>와 같다.

표 14 : 1차 요소표 정리 결과 및 출처

번호	요소명	국내 기관	해외 메타데이터 셋	비고
1	자료유형	전기관 적용	DC, GEM, FIU, KERIS, AHDS	
2	자료명	전기관 적용	DC, LOM, GEM, FIU, KERIS, AHDS	
2.1	대체 자료명	없음	FIU, KERIS, AHDS	
3	작성자명	전기관 적용	DC, LOM, GEM, FIU, KERIS, AHDS	
3.1	작성자 역할	없음	LOM, GEM, FIU, AHDS	
3.2	작성자 유형	없음	LOM, FIU, KERIS, AHDS	
3.3	작성자 이메일	없음	GEM, KERIS, AHDS	
3.4	작성자의 홈페이지	없음	GEM, KERIS, AHDS	
4	주제명	한국형사정책연구원, 통일연구원, 한국행정연구원	DC, LOM, GEM, FIU, KERIS, AHDS	
5	기술사항	한국행정연구원, 한국법제연구원	DC, LOM, GEM, FIU, KERIS, AHDS	
5.1	목차	없음	KERIS	
6	발행자명	한국형사정책연구원, 한국행정연구원, 한국직업능력개발원, 통일연구원, 한국법제연구원	DC, LOM, GEM, FIU, KERIS, AHDS	
6.1	발행자 유형	없음	LOM	
6.2	발행자 이메일	없음	GEM, KERIS	
6.3	발행자 홈페이지	없음	GEM, KERIS	
7	기여자명	한국직업능력개발원	DC, GEM, FIU, KERIS, AHDS	
7.1	기여자 역할	없음	GEM, FIU, KERIS	
7.2	기여자 유형	없음	GEM, FIU, KERIS	
7.3	기여자 이메일	없음	GEM, KERIS	
7.4	기여자 홈페이지	없음	GEM, KERIS	
8	날짜	한국교육과정평가원, 한국형사정책연구원, 한국행정연구원, 한국직업능력개발원, 통일연구원, 한국법제연구원, 한국교육개발원	DC, GEM, KERIS, AHDS	

번호	요소명	국내기관	해외 메타데이터 셋	비고
8.1	날짜 유형	없음		자체 생성
9	자료형식	한국교육과정평가원, 한국직업능력개발원, 한국교육개발원, 한국법제연구원	DC, LOM, GEM, FIU, KERIS, AHDS	
10	원정보 제목	없음	DC, GEM, FIU, KERIS, AHDS	
10.1	원정보 저자	없음		자체 생성
10.2	원정보 작성일	없음	FIU	
11	언어	없음	DC, LOM, GEM, FIU, KERIS, AHDS	
12	관련자료 제목	없음	DC, LOM, GEM, FIU, KERIS, AHDS	
12.1	관계 종류	없음	LOM, GEM	
13	이용조건	없음	DC, LOM, GEM, FIU, AHDS	

④ 유형마다 각각의 실제 자료를 모아 내용 및 형태를 분석한 후, 1차 요소표에서 필요하다고 판단되는 요소를 선택하고 추가할 필요가 있는 요소를 유형별로 추출하였다.

⑤ 1차 요소표에서 선택된 요소들 중 기본요소와 유형별 특수요소를 선정하였다.

<표 15>는 1차 요소표에서 선택되고 추가된 요소들 가운데 인문학 연구업적 DB의 기본요소 메타데이터 세트로 선정된 13개의 공통 요소를 보여주며, <표 16>은 학위논문, 학술지논문, 보고서의 자료유형별 특수요소로 선정된 13개의 요소를 보여준다.

표 15 : 기본 요소 메타데이터 세트

번호	요소명	설명	기타
1	**자료 유형**	자료의 내용적 성격, 장르(자료 유형 기술)	학위논문, 일반논문, 프로시딩, 사전, 법령집, 보고서, 학술지 등
2	**자료명**	자원에 부여된 제목	버전정보 포함
2.1	대체 자료명	자원의 공식적인 이름을 대신하는 대체 자료명(다른 언어로 번역된 제목, 간략제목 등)	
3	**작성자명**	자료를 현 상태로 작성한 사람 또는 기관의 이름	
3.1	작성자 역할	작성자의 역할	저자,번역자, 자료가공자
3.2	작성자 유형	작성자의 유형	교수, 학생, 강사/연구원/기타, 국/공립기관, 교육/연구기관, 기업/민간기관
3.3	작성자 이메일	작성자의 이메일주소	
3.4	작성자 홈페이지	작성자의 홈페이지 주소	
4	**주제명**	자원의 내용이 지닌 주제(topic)에 관한 사항 (주제 리스트에서 값을 취함)	
5	**기술사항**	자원의 내용과 관련된 설명, 요약(목록자가 작성함)	
5.1	목차	자원의 내용을 나타내는 하부 단위들의 리스트	
6	**발행자명**	자원의 발행자 및 온라인상의 배포자의 이름	
6.1	발행자 유형	발행자 유형	교수, 학생, 강사/연구원/기타, 국/공립기관, 교육/연구기관, 기업/민간기관
6.2	발행자 이메일	발행자의 이메일	
6.3	발행자 홈페이지	발행자의 홈페이지	

7	**기여자명**	작성자 요소에 명시된 개체 이외에, 자원의 내용에 대해 2차적인 책임이 있는 기타 개체의 이름	
7.1	기여자 역할	기여자의 역할	번역자, 논문수여기관, 기타
7.2	기여자 유형	기여자의 유형	교수, 학생, 강사/연구원/기타, 국/공립기관, 교육/연구기관, 기업/민간기관
7.3	기여자 이메일	기여자의 이메일	
7.4	기여자 홈페이지	기여자의 홈페이지 주소	
8	**날짜**	날짜 종류에 대한 실제 값	--년 --월 --일 등 표현 형식 정하기
8.1	날짜의 유형	자원의 존재 기간 동안 어떠한 사건이 발생한 날짜의 유형	작성일, 발행일, 최종갱신일
9	**자료형식**	자원의 표현형식. 기술상(technical)의 데이터 유형	text : plain, html, sgml, xml, sgml 등 application : excel, postscript, pdf, powerpoint 등 image: jpeg, pjpeg, gif, tiff, png, bmp 등 audio : au, wav, x-aiff, x-midi, realaudio 등 video : mpeg, quicktime, vhs 등
10	**원정보 제목**	현재 자원의 출처가 되는 원정보자원 또는 현재의 디지털 자원을 생성하는 데에 사용되었던 출처 자원의 제목	
10.1	원정보저자	원 자원의 내용에 주된 책임을 가진 개체	
10.2	원정보 작성일	출처 자원의 생성일	
11	**언어**	자원의 지적인 내용을 기술하고 있는 언어의 이름	한국어, 영어, 일본어, 중국어, 기타
12	**관련자료 제목**	기술되고 있는 자원의 관련자료 제목	
12.1	관계 종류	기술되고 있는 자원에 대한 관련자원의 관계유형	
13	**이용조건**	자원을 이용하기 위해 요구되는 조건	

표 16 : 연구업적 유형별 특수 요소 메타데이터 세트

자료유형	요소명		설 명
학위논문	1	수여기관	학위논문을 수여한 대학 및 기관
	2	학위유형	학, 석, 박사 논문의 구분
	3	초록	자원의 내용에 대한 요약 (영문초록, 국문초록 등)
	4	키워드	자원의 내용을 설명하는 주요 단어
학술지논문	1	간기	자원의 출처가 되는 학술지의 발행 주기
	2	자료수록 범위	자원이 수록된 권, 호 정보(예: vol, no.)
	3	초록	자원의 내용에 대한 요약 (영문초록, 국문초록 등)
	4	평가정보	자원의 출처가 되는 학술지의 평가정보(예: SCI 등재 여부)
	5	식별자	자원의 출처가 되는 학술지 식별자(예: ISSN, DOI 등)
보고서	1	간기	자원의 출처가 되는 보고서의 발행 주기
	2	공개유무	자원발행의 공개여부(예: 회색문헌 등)
	3	초록	자원의 내용에 대한 요약 (영문초록, 국문초록 등)
	4	식별자	보고서의 내부 식별자 및 외부 식별자 정보

3.3 인문학 연구업적 원문데이터의 표준화 방안

가. 개요

현재 국내 학술정보의 원문 포맷은 PDF, TIFF, TEXT, HWP 등 매우 다양한 것으로 조사되었다. 이미 구축된 원문의 89% 이상은 PDF, TIFF 포맷이며, 향후 구축을 희망하는 원문 포맷의 경우도 학위논문은 PDF 포맷 95%, 보고서 및 간행물은 PDF 또는 TIFF 포맷이 90% 이상을 차지하는 것으로 나타났다 (2.4, 표 8 참조). 이는 해당 기관의 형편과 정보 환경에 따라 각각 다른 원문 포맷을 적용한 결과이다. 향후 디지털화된 자원에 대한 원문 포맷 표준화가 확산될 전망이며 분산되어 있는 원문 정보의 교환을 위해서는 원문 포맷에 대한 표준화가 시급한 실정이다.

나. 원문데이터 파일포맷 종류 비교

원문데이터를 디지털화하는 파일포맷 종류는 크게 이미지기반 방법과 아스키 기반 방법 그리고 마크업 기반방법의 셋으로 나뉘어진다. 각각의 파일포맷 방식의 특징 및 장단점은 다음과 같다.

(1) 이미지기반 방법

① 특징 및 장·단점

장점	▪ 해당 페이지를 그림, 표, 사진 등을 포함하여 레이아웃 정보 스타일로 저장 ▪ 키워드 검색을 사용하여 검색속도 빠름 ▪ 문헌의 이미지를 담은 저장매체를 따로 보관하여 관리용이
단점	▪ 기계가독형태의 자료가 아니기 때문에 특정 페이지의 검색을 위해 별도의 색인어 제시 ▪ 모든 데이터가 이미지로 입력되기 때문에 대규모의 컴퓨터 기억공간이 요구, 전송 시 시간이 많이 걸림 ▪ 파일의 표준화가 이루어지지 않아 시스템간의 호환성이 적음 ▪ 검색된 데이터를 이용자가 필요 시 가공해서 처리할 수 없음 ▪ 이용자 터미널 화면에서의 해상도가 낮음
고려사항	▪ 이미지변환기술과 파일압축기술 필요 ▪ 이미지 파일 포맷 종류 (GIF, JPEG, TIFF) ▪ TULIP(The University Licensing Program)프로젝트의 경우 TIFF 형식의 비트맵포맷으로 만들어 아스키포맷과 혼합하여 아스키 데이터를 이용한 전문탐색가능 & 비트맵데이터를 이용하여 책자형에서와 같은 화상형태로 학술지 기사 이용 가능하게 함

② 파일포맷 종류 비교

	GIF (Graphics Interchange Format)	JPEG(Joint Photographic Experts Group)	TIFF(Tagged-Image File Format)
장점	▪ 이미지용 표준 포맷으로 컴퓨터 장치에 무관하게 데이터 저장 ▪ JPEG 가 등장하기 전까지 최고의 압축률 ▪ 애니메이션 기능이 추가되어 인터넷에서 영상을 다루는데 널리 사용	▪ 압축률이 가장 뛰어남 ▪ 인터넷에서 거의 이미지 표준포맷으로 자리매김 ▪ 비슷한 화질과 색상수로 저장할 때 다른 포맷의 1/10 이하로 크기가 작음 ▪ 24비트 트루컬러 지원, 멀티미디어 환경에서도 대중적임 ▪ 실압축방식으로 화질의 저하없이 파일 압축 ▪ 컬러데이터 처리할 때 많이 쓰임	▪ 모든 스캐너에서 TIFF 이미지 생성가능 ▪ 다양한 압축방법제공- 다양한 용도로 사용가능 ▪ JEPG처럼 화질이 떨어지지도 않고, 모든 색상 표현가능, 압축률도 어느정도 뛰어나 컴퓨터출판 이미지는 대부분 이 포맷 사용
단점	▪ 비트맵 방식의 압축 형식으로 8비트 256 컬러밖에 지원 못함 ▪ JPEG에 비해 파일크기가 크고 전송속도 느림		▪ TIFF 파일을 압축하면 파일 크기는 줄지만 파일을 열고 저장할때 요구되는 시간 길어짐

(2) 아스키기반 방법

장점	▪ 원문헌에서 사용된 단어를 기반으로 색인이 작성되므로 전문검색과 주제검색이 가능 ▪ 검색시간과 전송속도가 빠르고 저장공간이 절약 ▪ 이용자들이 원하는 정보를 검색해서 자기가 필요로 하는 형태로 가공, 변환 가능
단점	▪ OCR과정에서 문서에 담긴 문자들의 인쇄상태나 글자간격, 글자체 등에 따라 인식의 정도가 달라짐 ▪ 서명이나 장(章)의 제목, 표 등의 정보 누락 가능 ▪ 문자의 오류를 수정하는 데 시간과 노력이 많이 듦
고려 사항	▪ PDF는 텍스트 파일 포맷으로 디지털화 된 문헌의 외형적인 질과 검색기능 및 출력기능 등의 측면에서 가장 우수하다고 인정되고 있음

(3) 마크업기반 방법

① 특징 및 장·단점

장점	▪ 문헌이 보유한 내용과 구조를 인지하여 네트워크상에서 데이터전송을 가능하게 함 ▪ 기술적(descriptive) 마크업은 문헌이 전달하고자 하는 다양한 정보를 구조화하여 정보검색 시 이용자에게 제공함으로써 전자문헌의 구축과 검색에 보다 유리
단점	▪ 사용이 어렵고 복잡하며 전문적인 각 마크업 언어에 대한 지식이 요구 ▪ 마크업 언어에 따라 특정 판독기(viewer, browser 등)가 필요함 ▪ 구축시 비용이 많이 듦 ▪ 기 구축된 MARC와의 상호호환성 문제
고려사항	▪ 기술적 마크업 언어의 대표적 예로 SGML과 HTML이 있는데 HTML은 SGML의 축약판이라고 할 수 있음 ▪ HTML의 경우 특수문자나 기호들을 표현하지 못하므로 이러한 것들은 비트맵 등을 이용해서 표현해야 함

② 각 마크업 언어 비교

	HTML	SGML	XML
출력형식언어	CSS	DSSSL	XSL
링크	HTML(only simple link)	Hytime	XLL
스펙	단순	복잡	사용자 정의 가능, SGML보다는 제한적
태그사용	사용자 정의 불가능,제한적	사용자 정의 가능, 무제한	웹브라우저
브라우저	웹브라우저	전용 브라우저	웹브라우저
데이터 공유	불가능(단방향)	가능	가능(양방향)
문서작성	간단 용이함.논리구조 작성 어려움	복잡	작성용이
문서검색	효과적 검색 어려움	정확한 검색 가능, 문서구조 검색 가능	정확한 검색 가능, 문서구조 검색 가능
문서재사용	불가능	가능	가능
응용분야	단순구조 문서	방대한 내용, 구조를 요하는 기술적인 문서 및 전자	방대한 내용, 구조의 기술 문서, 웹상의 교환문서

(4) 아스키기반 방법과 마크업기반 방법의 비교

※ PDF 와 HTML의 비교

	PDF	HTML
사용언어	PostScript	Hyper Text Markup Language
Data	Text, Image, Vector Graphic이 파일 내 존재	각자의 File로 존재하며 Link하여 Image 주소를 갖는다.
Browser	Acrobat Reader, Exchange, Netscape/Explorer	Netscape/Explorer
통신	On-line(Netscape/Explorer)	On-line(Netscape/Explorer)
	Off-line(Acrobat Reader, Exchange)	
System	Windows, MAC, OS/2, Unix	Windows, MAC
File Save	PDF만 저장(PDF내 모두 존재)	HTML, Image 각각 저장

	PDF	HTML
Graphic	Vector-based line art	Bitmap
Image 표현	EPS, JPEG, GIP, TIF	JPEG, GIP
압축기능	있음	없음
보안기능	있음	자체 보안기능은 없다.
확대 / 축소	가능	불가능
WYSIWYG	가능	불가능
Off-line 작업	가능	불가능
디지털 서명 지원	가능	불가능
안전한 JavaScript	가능	불가능

다. XML 기반 원문 DB 구축 문제점

① XML 원문 DB 구축의 경제적 문제점 : 디지털 텍스트가 없는 경우 PDF로 변환하는 것에 비해 약 10~15배, 디지털 텍스트가 있을 경우에는 PDF로 변환하는 것에(1면 당 200원) 비해 약 20~25배 정도의 비용이 소용된다 (1면 당 약 3,500~5,000원, 1인 당 1일 기준 평균 25면 정도 작업 가능).

② XML 원문 DB 구축의 기술적 문제점 : DB의 무결성 유지가 어려워 자료수집, 분류, 제작, 검수, 서비스제공까지 최소 3~4개월이 소요된다. 그리고 XML DTD의 불완전성으로 인해 기술개발 방향이 XSL 및 RDF 방향으로 바뀌고 있는 등 실제 적용이 불안정하며, 디지털 텍스트에 수식, 한자, 그림 등이 포함된 경우 서비스 제공 시 폰트와 플러그인 설치가 필요하다.

③ XML 기반 원문 DB 구축과 이용의 현실적 문제점 : 학술적 가치가 매우 높거나 최신성이 필요한 자료의 경우 구축 의미가 있으나 이러한 자료는 국학이나 첨단과학 분야에 국한되며, 첨단분야도 제작과정 4개월이 경과하면 그 의미가 없다고 볼 수 있다. 또한 인력 확보가 어려운 문제점이 있다. XML 원문DB 구축에 투입될 인력은 상당한 지식수준과 경험이 필요하나, 업무 성격은 단순 업무이므로 인력을 지속적으로 유지하는데 현실적 어려움이 발생할 수 있다. 디지털 문서형태의 표준을 위한 기술 축적 측면에서 시범 사업

의 의미는 가질 수 있으나, 현실 적용을 서두를 경우 표준화를 위한 표준화가 될 수 있음을 고려해야한다. 그리고 XML 기반 DB의 구조검색시스템에 대한 이용자 만족도가 키워드 검색시스템보다 낮으며, 구조검색시스템을 이용할 수 있는 이용자는 매우 높은 지적 수준과 시스템 이용 경험을 보유해야 함을 명심해야 한다.

라. PDF를 기반으로 하는 원문 DB 구축 방향

현재 구축되어 있는 원문 자료의 대부분이 PDF와 TIFF로 구축되어 있는 실정이다. 또한 향후 구축하고 희망하는 원문 자료의 포맷도 PDF와 TIFF가 다수를 차지하고 있다. 원문 포맷 구축자의 입장에서 원문 구축에 대한 경제성, 기술적 안정성, 향후 표준화 가능성 등을 고려해 본다면 PDF, TIFF 등이 효과적인 원문 포맷이 될 것이다. 이용자의 입장에서도 지명도가 높고 뷰어가 프리웨어인 PDF가 가장 편리하게 이용할 수 있는 포맷이라고 할 수 있다. 따라서 현재 구축된 많은 원문 정보의 포맷이며 경제성, 효율성, 이용자 편이성 등의 측면을 고려할 때 원문 포맷의 표준화로 PDF 포맷 중심으로 지향하는 것이 타당할 것이다.

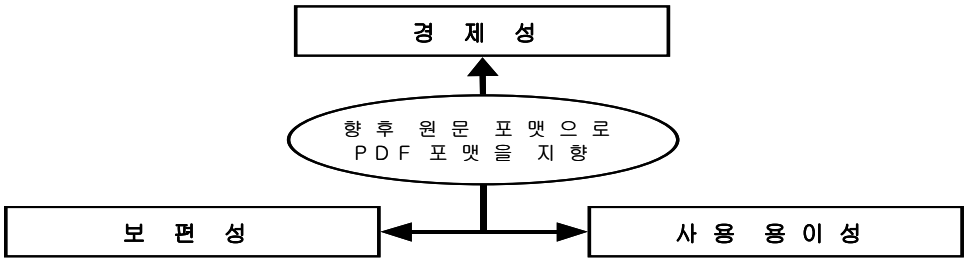
물론 XML과 PDF가 상호 장점을 최대한 유지하며 보완적인 형태로 진행되는 것이 바람직하며, 이를 위해서는 XML문서가 PDF와 호환될 수 있는 전용 XSL 및 PDF XSL-FO Processor의 개발이 필요하나 아직은 구상 단계에 머물고 있는 실정이다.

※ 원문의 디지털화시 비용 측면 고려

▶ 원문의 형태가 문서인 경우에도 적은 비용으로 구축 가능

▶ 원문 구축을 위한 기반 장비에 대한 투자는 필요하지만 미약한 수준임

▶ 원문 뷰어가 프리웨어로 제공되어 비용이 필요 없음



※ 해당기관의 선호도 측면 고려

▶ 다수의 기관이 선호하여 향후 지식생성기관 등에서 자체적으로 생산된 원문에 대한 교환 및 공유가 용이함

※ 이용자입장에서 편리한 정보접근측면고려

▶ 다수의 정보이용자가 공통적으로 사용하는 원문뷰어 사용

▶ 디지털화된 원문 조회시 뷰어자체에서 다양하고 편리한 기능들을 제공

4. 인문학 연구업적 데이터베이스 구축 방안

인문학 분야의 연구업적을 대상으로 디지털 데이터베이스의 구축은 기여도, 역량, 합의의 극대화를 원칙으로 삼아야 한다.

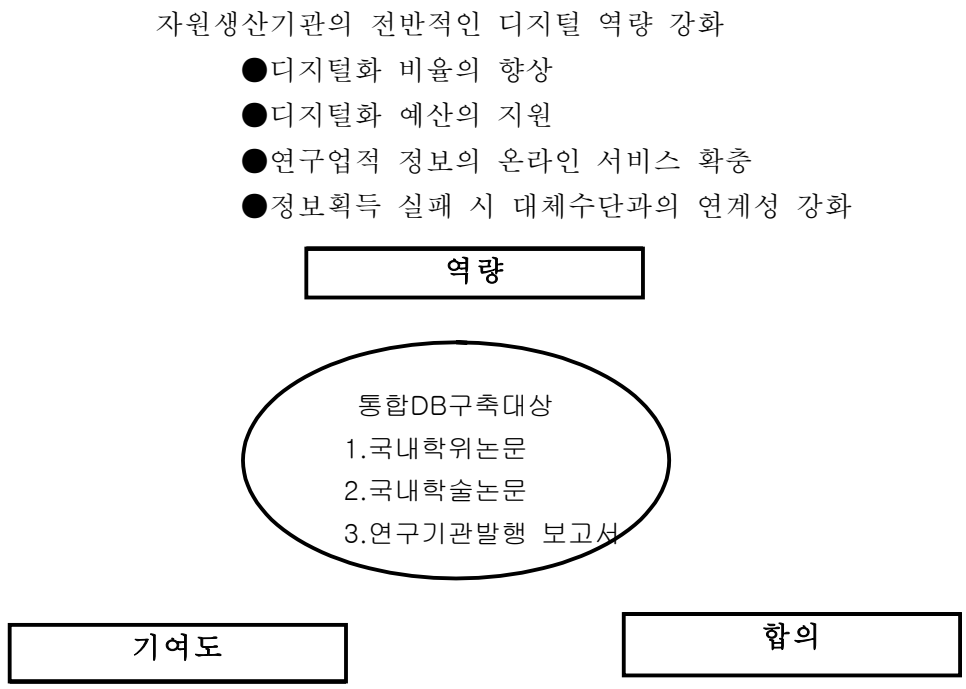


그림 2 : DB 구축 원칙

<그림 2>가 보여주는 원칙을 바탕으로 현재 각 기관마다 분산되어 처리되고 있는 DB를 연계시킬 수 있는 인문학 연구업적 네트워크를 구축하여 통합 관리와 검색이 가능한 서비스 시스템을 구축하는 것이 필요하다 (<그림 3> 참조).

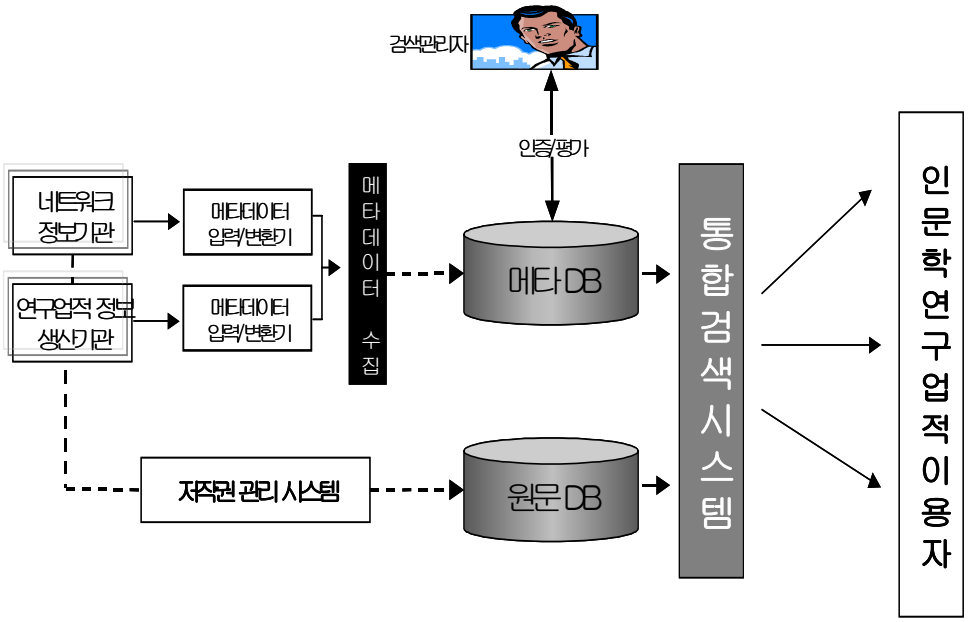


그림 3 : 인문학 연구업적 DB 구축 개요도

4.1 기술적 측면

가. 입력 시스템

메타데이터 입력기란 각 인문학 분야 네트워크 소속기관 혹은 관리기관에서 생성하는 디지털 자원에 대한 메타데이터를 메타 DB에 입력하는 도구를 말한다. 이는 표준화된 메타데이터 생성을 유도하여 통합 메타 DB의 일관성을 보장하게 된다. 메타데이터의 기능으로는 메타데이터 입력, 저장, 삭제/편집, 조회, 요소 편집, 메타데이터/스키마 반입/반출, 사용자 관리 등의 기능이 있게 된다.

메타데이터 입력기 개발시 다음과 같은 사항을 고려해야 한다. 템플리트 메타데이터 스키마 정의로 메타데이터의 설계 용이성을 확보해야 하며, 손쉬운 메타데이터 스키마 생성 인터페이스를 제공해야 한다. 또한 다양한 메타데이터 유형에 대하여 동일한 입력 화면을 제공해야 하며, 내부적인 데이터 교환 포맷으로 XNL을 지향하고 DC 기반의 XML 구조의 메타데이터 저장 구조를 지원해야 한다. <그림 4>는 메타데이터 입력기의 기본 구조도이며, <그림 5>는 메

타데이터 입력기의 실제 예시 화면이다.

그림 4 : 메타데이터 입력기 기본 구조도

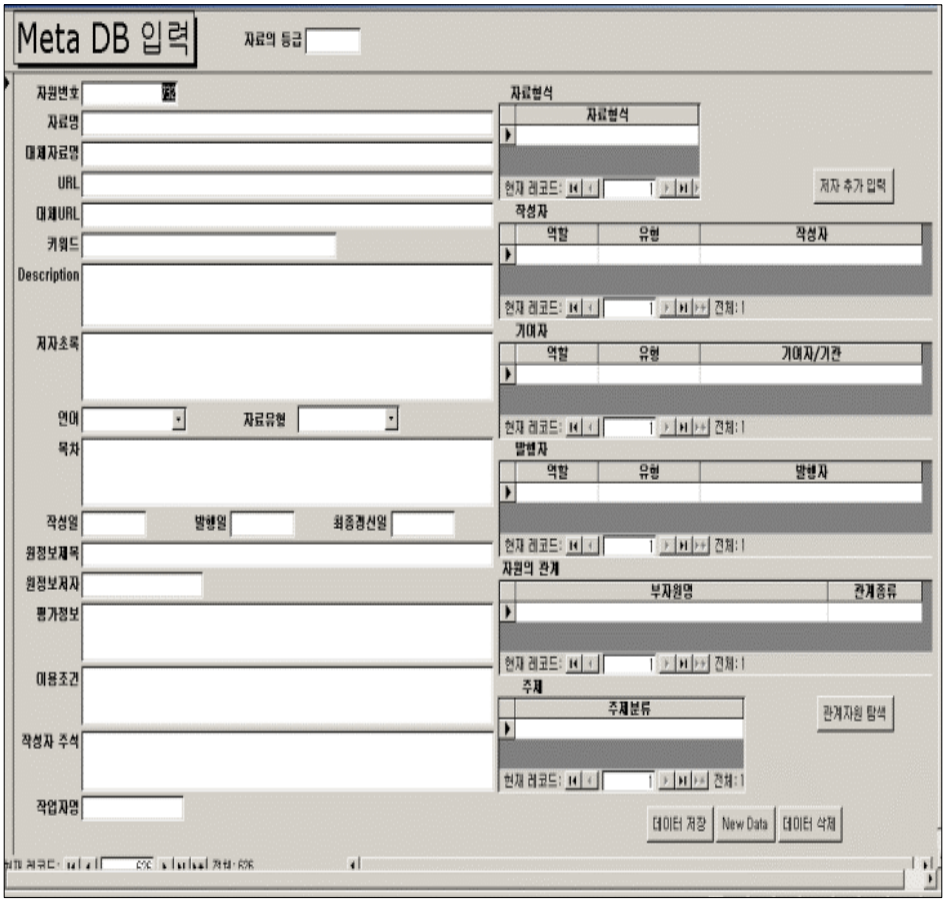


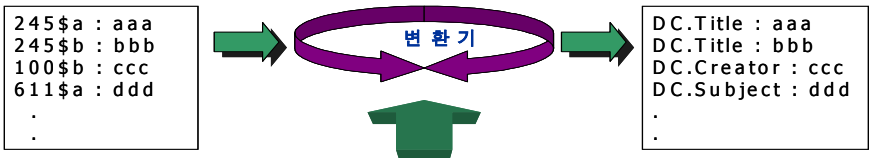
그림 5 : 메타데이터 입력기 실제 예시 화면

나. 변환 시스템

메타데이터 변환기란 각 인문학 분야 네트워크 소속기관의 해당 기관별로 관리되는 메타데이터에 대한 통합 DB를 구축하기 위한 도구이다. 이는 다양한 형태의 표준에 의해 만들어진 메타데이터에 대하여 통합된 메타 DB의 일관성을 보장하게 한다

메타데이터 변환기의 기능으로는 메타데이터 반입, 조회, 변환 맵 정의/생성, 변환, 적재, 반출 등이 있다. 그 구조는 크게 메타데이터 변환 맵과 일괄 배치 프로그램으로 나누어 볼 수 있다. 메타데이터 변환 맵은 현재 다수의 기관들이 사용하고 있는 MARC 포맷에 대하여 새로이 개발, 정의한 메타데이터 구조로 매핑시킬 수 있는 변환 맵이 있어야 한다. 일괄 배치 프로그램은 미리 정의한 데이터 요소에 대한 반입/반출을 위한 교환 포맷을 정의하여 각 기관에서 반출된 메타데이터를 표준 메타데이터 형태로 자동 변환시키는 프로그램이다.

변환기 프로그램을 통하여 기존 메타데이터 DB에 저장된 메타데이터는 변환기를 통하여 XML 형태의 새로운 메타데이터 포맷으로 전환되어 신규 메타데이터 DB에 저장되게 된다. <그림 6>은 메타데이터 변환기 구조를 나타낸다.



MARC 포맷	DC 포맷
245$a	DC.Title
245$b	DC.Creator
100$a	DC.Subject
…	…

그림 6 : 메타데이터 변환기 구조

다. 검색 시스템

산재되어 있는 인문학 연구업적 정보를 효율적으로 관리하고 검색할 수 있는 분산 통합 검색 시스템을 개발할 필요가 있다. 인문학 분야의 연구업적 정보 중에서 검색 대상을 지정하고 검색 대상에 따른 검색 방식의 다양성을 수용하도록 한다. 이는 언제, 어디서나 원하는 정보를 편리하게 얻을 수 있는 One-stop 서비스가 될 것이다. <그림 7>은 통합 검색 시스템의 개요도를 나타낸 것이다.

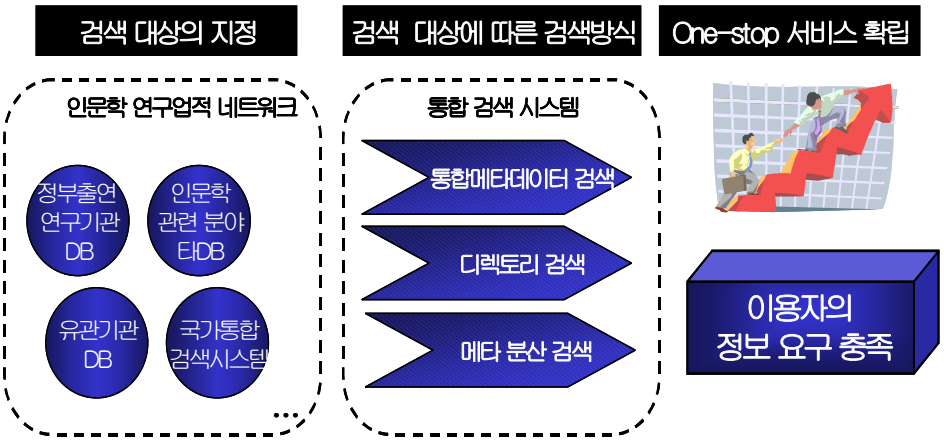


그림 7 : 통합 검색시스템의 개요도

(1) 통합메타데이터 검색

통합메타데이터 검색은 메타데이터 데이터베이스에 모든 자료를 저장하고 특정 단어(키워드, 주제어 또는 검색어)를 입력함으로써 원하는 정보를 찾는 방법이다. 근래에는 로봇 에이전트 혹은 웹 에이전트를 이용하여 인터넷 상의 정보와 연계하여 모든 자료(파일, 웹 관련자료, RDB)를 수집하여 저장구조에 저장하는 형태로 구성된다. 통합메타데이터 검색시스템의 기능은 다음과 같다.

※ 단일 서버에서 메타데이터가 관리될 수 있으므로 검색 전략과 검색시스템 운용이 용이
※ 동의어/유사어 등 확장 검색 이용시 적용이 용이

※ 서로 다른 형식의 문서들의 메타데이터를 통합 관리하므로 검색 용이
※ 방대하고 복잡한 정보환경에서 수집된 문서들을 단일화된 구조의 메타데이터로 작성하므로 일관성이 있음
※ 상대적으로 다른 검색 방식보다 검색 성능이 우수

통합메타데이터 검색시스템의 장단점 및 고려사항은 다음과 같다.

※ 장점
▶ 단일 서버에서 색인이 관리되므로 검색 전략과 검색시스템 운용이 용이
▶ 동의어/유사어 등 확장 검색 이용시 적용이 용이
▶ 서로 다른 형식의 문서들에 대한 통합된 메타데이터의 적용 및 검색 용이
▶ 방대하고 복잡한 정보환경에서 수집된 문서들을 단일화된 구조로 메타데이터화해 일관성이 있음
▶ 상대적으로 다른 검색 방식보다 검색 성능이 우수

※ 단점
▶ 메타데이터가 정확하지 않은 경우, 원치 않는 결과를 초래
▶ 메타데이터 결과가 방대하여 컴퓨팅 자원이 다른 메타데이터에 비해 많이 소요
▶ 하드웨어 문제 발생시 전체 검색 및 정보자원과 관련된 모든 서비스가 중단됨
▶ 정보제공기관에서의 주기적인 자료 제공 필요(이용자에게 최신자료를 제공하기 위하여)

※ 고려사항
▶ 문서량의 증가에 따라 하드웨어를 증설해야 하며, 초기 Sizing이 매우 중요함
▶ 장애발생을 대비 별도의 장애관리 솔루션을 도입해야 함
▶ 문서량의 증가에 따라 대용량의 시스템이 필요
▶ 타 기관으로부터 데이터를 원활히 제공 받을 수 있는지 등 데이터 수집 방안이 수립되어야 함
▶ 동의어/유사어(시소러스 기능)를 활용하기 위해서 인문학 분야의 시소러스 구축이 선행되어야 함

▶ 동일한 하드웨어 환경에서 자원(CPU, Memory, Disk 등)을 효과적으로 이용할 수 있는 솔루션을 채택

(2) 디렉토리 검색

디렉토리 검색은 수집된 정보를 보고 이들 정보에 대한 분류를 통해 데이터베이스화 및 색인하여 다단계 메뉴형태로 정보를 제공하는 것이다. 다른 검색방식에 비해 이용자가 찾고자 하는 정보에 대해 어느 정도 지식을 가지고 있다면 쉽게 정보를 찾을 수 있는 검색 방식이다. 디렉토리 검색시스템의 기능은 다음과 같다.

※ 온라인 상에 산재되어 있는 정보를 일정한 방식으로 군집하여 서비스를 함
※ 이용자는 분류된 디렉토리 구조에 따라 관련된 지식정보를 찾아가면서 쉽게 검색
※ 검색어에 대한 자료를 분류체계에 따른 자료가 저장된 데이터베이스에서 조회하여 결과를 보여줌으로서 키워드 검색의 약점을 보완
※ 문서의 상호참조 및 다중분류를 쉽게 구현
※ 분류체계 변경시 쉽게 재분류가 가능하며, 분류를 균형 있게 유지

디렉토리 검색시스템의 장단점 및 고려사항은 다음과 같다.

※ 장점
▶ 온라인 상에 산재되어 있는 정보를 일정한 방식으로 군집하여 서비스를 함
▶ 이용자는 분류된 디렉토리 구조에 따라 관련된 지식정보를 찾아가면서 쉽게 검색
▶ 검색어에 대한 자료를 분류체계에 따른 자료가 저장된 데이터베이스에서 조회하여 결과를 보여줌으로써 키워드검색의 약점 보완
▶ 문서의 상호참조 및 다중분류를 쉽게 구현함
▶ 분류체계 변경시 쉽게 재분류가 가능하며, 분류를 균형있게 유지함

※ 단점
▶ 이용자가 잘못된 분류체계로 들어갈 경우 잘못된 결과를 얻음

▶ 지식정보를 수집하고 분류하는 작업은 일정 수준의 전문가가 분류해야 하며, 이에 대한 인적자원 소요비용이 큼
▶ 상대적으로 유지보수 인원이 많이 필요함

※ 고려사항
▶ 인문학 분야 지식정보자원에 대한 분류체계 확립이 선행
▶ 운영 인력/조직이 확보되어 있는가에 대한 확인이 필요함
▶ 자동분류기능 도입의 가능성을 점검
▶ 분류를 자동화할 경우 비용절감 효과가 매우 큼
▶ 대량의 문서들에 대해서 빠른 시간 내에 처리하는 관점에서는 오히려 수작업보다 편리하지만 이론적으로는 일관성 및 정확성이 떨어짐

(3) 메타 분산 검색

메타 분산 검색은 통합메타데이터 검색(키워드 검색)처럼 데이터베이스에 자료를 저장하는 방식이 아니라 데이터베이스에 자료를 가지고 있지 않으며, 리모트의 검색 프로그램들을 이용하고 각 검색엔진으로부터 회신된 결과들을 종합하여 좀더 통일된 형태의 결과물을 제공하는 것이다. 메타 분산 검색시스템의 기능은 다음과 같다.

※ 검색 기능을 제공하는 단위 시스템들을 통합 검색하기 위한 가장 빠른 구축 방법
※ 신규로 추가적인 시스템 구축 시 통합 검색을 위해서 고려할 사항이 거의 없는 상태
※ 특정 시스템의 일시정지 및 오류발생시에도 다른 시스템에 대한 검색은 정상적으로 수행
※ 막연하고 모호하며 불분명한 주제를 검색할 때 용이
※ 다른 검색 방식으로 어떤 결과도 찾지 못했을 때 용이
※ 하나의 검색식으로 많은 결과를 얻을 수 있음
※ 한번의 클릭으로 여러 개의 검색 엔진이나 디렉토리를 한꺼번에 이용하는 효과

메타 분산 검색시스템의 장단점은 다음과 같다.

※ 장점

▶ 검색기능을 제공하는 단위 시스템들을 통합 검색하기 위한 가장 빠른 구축 방법
▶ 신규로 추가적인 시스템 구축 시 통합 검색을 위해서 고려할 사항이 거의 없는 상태임
▶ 특정 시스템의 일시정지 및 오류발생 시에도 다른 시스템에 대한 검색은 정상적으로 수행됨
▶ 단위시스템 관리자와의 업무협의 부분이 거의 없으므로 사업 추진이 용이
▶ 막연하고 모호하며 불분명한 주제를 검색할 때 용이
▶ 검색해서 어떤 결과도 못 찾을 때 용이
▶ 하나의 검색식으로 가능한 많은 자료를 얻을 수 있음
▶ 한번의 클릭으로 여러 개의 검색엔진이나 디렉토리를 한꺼번에 이용하는 효과

※ 단점

▶ 동일한 키워드에 같은 결과가 중복 발생할 수 있으며 동일한 검색결과를 보장할 수 없음
▶ 특별한 형태의 검색식을 쓸 때는 검색 수행에 제한을 받음
▶ 각각의 검색엔진이 구현할 수 있는 연산자나 검색 기능을 효과적으로 살릴 수 없어 복잡한 검색식 구현이 어려움
▶ 제한된 시간 내에 검색되는 자료를 가져오기 때문에 각 하위의 시스템 및 네트워크 상황에 따라 검색속도가 달라질 수 있어 성능 유지가 어려움
▶ 검색속도가 통합대상 시스템 중 가장 느린 시스템에 의해서 결정되며, 대부분 웹을 통한 인터페이스이므로 느린 편에 속함
▶ 일반적인 검색엔진에서 제공하는 정확율에 따른 순위 매김이 어려움
▶ 통합분류체계의 제공이 제한적임(정보제공기관별 분류만 가능함)
▶ 서로 다른 검색엔진의 사용으로 정확율 향상을 위한 검색전략 수립이 어려움

라. 저작권관리 시스템

디지털 정보 서비스는 이용자로 하여금 편리하고 신속하게 각종 정보를 획득할 수 있도록 지원함으로써 사회, 경제적 효과를 극대화하고 있다. 그러나 디지털이라는 매체의 특수성으로 인해 복제나 불법 활용이 매우 유용하며 일단 복제되면 그 전달이 대량으로 확산될 수 있다는 특징을 가지고 있다. 만약 체계적인 저작권 관리 및 보호 제도가 없다면 저작자 및 대행사에 막대한 손해를 끼칠 수 있다.

DRM(디지털 지적재산권 관리 시스템)의 도입으로 체계적인 디지털 저작물의 저작권 정보를 관리하며 각 저작물별 저작권 관리 및 조건들을 이용자 환경에 적용하여 저작물 자체에 대한 보호 환경을 지원할 수 있다.

따라서 이는 디지털 저작물 구축 및 이용에 필수적인 시스템으로 제시된다. 디지털 저작권 관리 시스템은 저작물에 대한 설명, 식별, 권한, 조건, 거래 등의 다양한 저작권 정보를 통하여 저작물의 생성에서 활용까지의 시스템 흐름이 필요하며 지속적인 관리를 운영함으로써 저작물의 활용 기간 동안 자신을 보호하고 관리될 수 있도록 구축되어야 한다.

또한 디지털 저작물이 저작권자로부터 이용자에게 전달되는 과정에서 안전, 보안, 보호 및 다양한 관리가 이루어 질 수 있는 운영 시스템의 구현이 필요하다. 디지털 저작권 관리 시스템의 기능으로는 디지털 저작권 정보 관리 기능, 저작물 이용자 정보 관리 기능, 보고 및 통계 정보 관리 기능 등이 있다.

4.2 정책적 측면

가. 관리체제의 체계화 방안

인문학 분야의 DB 구축을 성공적으로 추진하기 위해서는 첫째, 집중 투자에 의한 연구업적 DB 자원을 대폭 확충하는데 있다. 인문학 분야의 디지털화 정도를 획기적으로 제고할 수 있고, 투자효과를 극대화할 수 있도록 집중적으로 투자해야 한다. 둘째, 연구업적 데이터베이스의 유통을 활성화시켜야 한다. 이는 기관별로 분산되어 있는 학술 자원과 데이터베이스를 마치 하나의 데이터베이스를 이용하는 것과 같은 시스템 환경으로 조성하여 시스템 유지관리 측면에서의 비용절감 뿐만 아니라 최종 이용자인 인문학 연구자들의 편의를 도모함으

로써 자발적이고 적극적인 관심과 참여를 유도해야 한다. 셋째, 인문학 관련 지식정보자원 관리체제의 정비가 필요하다. 실례로 한국전산원의 국가지식정보통합검색시스템과 국가정부출연연구기관인 지식정보관리시스템인 IKIS의 관리체제를 보다 효율적으로 정비할 필요가 있다.

표 17 : 관리체제의 체계화를 위한 전문조직 및 인력확보 추진과제

구분	세부추진과제
연구기능 확보	• 자료 보존 및 관리에 대한 연구인력 확보 • 연구사업 시행
전문인력 네트웍 구성	• 전문인력 Database 구축 • 전문인력네트웍 구축 및 운영 • 자문, 공동연구, 위탁 등 연계 강좌
전문인력 양성	• 전문인력 양성에 관한 연구 실시 • 인력양성 프로그램 개설 - 인문사회연구회와 연계
전담부서 신설	• 인문학 연구업적 통합관리 전담부서 신설 • 전담 CIO 임명

나. 관련 기관의 연계 방안

인문학 연구업적 DB를 구축하기 위해서는 먼저 인문학 관련 지식정보자원 관리기관과의 연계를 강화해야 한다. 교육학술분야 종합정보센터로는 한국교육학술정보원과의 연계를, 역사분야 종합정보센터로는 국사편찬위원회와 연계를, 그리고 문화예술분야 종합정보센터로는 문화관광부와의 연계를 강화하는 것이 인문학 연구업적 DB를 성공적으로 구축하기 위한 필수 방안이라고 볼 수 있다.

표 18 : 관련기관 연계 세부 추진 과제

구분	세부추진과제
협의체 구성	• 국내 협의체 구성 운영 • 국제 협력을 위한 개별 접촉 및 협력 유도
정보통합관리시스템 구축 및 지원	• 종합관리시스템 구축 • 관련기관 정보화 추진을 위한 시스템 및 운영 지원
영문 데이터베이스 구축	• 용어와 고유명사에 대한 통일 작업
해외 연계시스템 개발 및 서비스	• DC, IFLA 등의 표준화 작업에 동참 • 메타데이터 추출 및 디지털화 과정에서 국제적인 공유 • 한국의 인문학을 필요로 하는 외국 기관과의 연계

다. 법, 제도 정비 및 표준화

구축된 데이터베이스가 원활하게 관리되고 유통되기 위해서는 우선적으로 지식기반사회로의 전환을 위한 법과 제도를 정비해야 한다. 저작권이 보호되면서 디지털 자원의 공유와 유통이 활성화 될 수 있도록 관련법과 제도를 현실에 맞게 정비해야 한다. 이와 더불어 지식정보자원 관리의 표준화 및 기술개발이 시급하다고 볼 수 있다. 새로운 정보통신 환경에 적합한 분야별 메타데이터 표준화가 시급히 이루어져야 하고, 저작권보호 및 이용의 편리성, DB 구축의 경제성 등의 문제를 모두 극복할 수 있는 원문 디지털 포맷의 표준화가 빨리 정착되어야 한다.

라. 단계별 추진 계획

인문학 연구업적 데이터베이스 구축을 성공적으로 수행하기 위하여 2003년부터 2005년까지 기반조성, 안정화, 활성화 등의 목표를 부여한 단계별 구체화 방안은 표 19와 같다.

표 19 : 인문학 연구업적 단계별 구체화 방안

구분	2003년 기반조성	2004년 안정화	2005년 활성화
메타데이터 관련	◦ 메타데이터 표준화 ◦ 인문학 학술정보 분류체계 표준화	◦ 표준 메타데이터 적용 시범시스템 개발 ◦ 메타데이터 표준 적용 확산 ◦ 메타데이터 DB 구축 ◦ 메타데이터 입력/변환기 개발	◦ 메타데이터 표준 적용 확산 (계속) ◦ 기존 데이터 변환
원문데이터 디지털화	◦ 원문 디지털화 포맷 표준화 ◦ 매체별 관리체계 확립	◦ 원문 디지털화 포맷 적용 시범시스템 개발 ◦ 표준 적용 확산 ◦ 원문데이터 DB 구축	◦ 원문데이터 DB 구축 확산 (계속) ◦ 기존 데이터 변환
서비스	◦ 통합색인 설계 및 구축 ◦ 서비스 뷰 설계 ◦ 통합검색 시스템 구축 ◦ 회원관리 및 웹서비스 시스템 구축	◦ 인문학 정보 포털시스템 구축	◦ 국내외 유관기관 정보공유시스템 구축 ◦ 웹 LOD 및 CRM 구축 ◦ 개인화 맞춤형 서비스 구축
연계방안	◦ 협의체 구성 운영	◦ 정보통합관리시스템 ◦ 협의체 구성 운영(계속)	◦ 관련기관 정보화시스템 정비 ◦ 영문 DB 구축 ◦ 해외 연계시스템 구축
전문조직 및 인력확보	◦ 전담부서 및 CIO 확보 ◦ 전문인력 네트워크 구축 ◦ 전문인력 DB 구축	◦ 전문인력 네트워크 구축(계속) ◦ 연구부서 신설 ◦ 인력양성 프로그램	◦ 연구사업 시행
법, 제도 개선	◦ 법, 제도 개선 추진위 구성 ◦ 법, 제도 개선 세미나 개최 ◦ 저작권 관련 법규 개정	◦ 인문학 정보 전문기관화	◦ 전문기관 위상 확립

5. 결론

본 연구에서는 인문학 연구업적 관리 현황 및 문제점을 살펴보고 이러한 제 문제를 해결할 수 있는 데이터베이스 표준화 방안을 제시한 후, 이를 기초로 하여 인문학 연구업적 데이터베이스 구축방안을 기술적 측면과 정책적 측면으로 나누어 제시하였다.

본 연구의 궁극적인 목적을 요약하자면 인문학 분야의 기관 간 정보자원 공유의 체계화를 이룩하고, 표준화된 연구업적 데이터베이스를 구축하여 결국 정보를 가진 기관의 입장에서는 연구업적 정보의 효율적인 관리를 가능하게 하며, 최종이용자의 입장에서는 편리하게 연구업적 정보의 소재를 파악할 수 있고 또 원문을 바로 이용할 수 있는 방안을 마련하는 것이라 할 수 있다. 이러한 목적을 구현하기 위해서는 먼저 데이터베이스의 분산 구축과 통합 서비스 체제가 기본적으로 전제되어야 한다. 그러나 현황조사의 분석 결과는 인문학 분야의 연구업적 정보를 관리하고 있는 기관들의 독자적 기준에 따라 연구업적 정보를 수집・구축・서비스를 하고 있는 것으로 나타나고 있으며, 연계 시스템에 대한 계획도 마련되어 있지 않은 것으로 나타나고 있다.

따라서 인문학 연구업적 정보의 통합 데이터베이스를 구축하기 위해서는 무엇보다도 정보분류체계, 메타데이터 형식, 원문의 디지털화 형식의 표준화가 매우 시급한 것으로 보이며 이를 위해서 본 연구에서는 다음과 같은 제안을 하였다 : (1) 인문학 분야 연구업적 정보의 분류표준으로서 한국교육학술정보원에서 제안한 표준(안)을 토대로 인문학 분야의 표준안을 개발, 발전시키는 것이 바람직하다. (2) 인문학 연구업적 정보의 메타데이터 형식의 표준으로는 더블린 코어(DC)를 기반으로 연구업적 정보의 특성과 새로운 매체에의 상호운영성을 고려한 표준안이 바람직하며, 본 연구의 결과로서 시범 모형을 제시하였다. (3) 원문의 디지털화을 위한 표준 포맷으로는 PDF 형식이 경제적, 기술적, 인력의 운용 용이성 측면에서 가장 타당한 것으로 평가된다.

기관간의 비 표준화 및 유사, 중복 등의 문제점을 최소화하기 위해서는 관련 기관간의 철저한 협력과 업무분담이 전제되어야 한다. 기관간의 업무분담 방법에는 연구업적 정보의 수집 및 디지털화의 주관은 해당 정보의 생성기관이 담당하고, 연구업적 정보의 축적 및 유통은 주관센터나 주관기관을 지정하여 이를 통한 공유와 유통체제를 확립한 다음, 통합검색을 위한 통합 메타데이터 체계를 구축하는 것이 바람직하다. 또한 연구업적 정보의 수집방법에 있어서 연구자들의 자발적인 참여를 유도하여 연구자 본인의 정보 및 연구업적자료를 등

록/수정 할 수 있도록 함으로써 각 기관간의 역할 분담을 보다 효율적으로 완성할 수 있을 것이다.

정책적 측면에서는 무엇보다도 관리체제의 체계화가 시급하다. 이를 위해서는 집중 투자에 의한 연구업적 DB 자원을 대폭 확충, 연구업적 데이터베이스의 유통 활성화, 인문학 관련 지식정보자원 관리체제의 정비가 필요하다. 관리체제의 체계화와 함께 고려되어야 할 다른 하나의 측면은 관련 기관들의 연계를 강화하는 것이다. 이를 위해서는 인문학 관련 지식정보자원 관리기관과의 연계 강화, 지식기반사회로의 전환을 위한 법과 제도 정비, 지식정보자원 관리의 표준화 및 기술개발이 시급하게 이루어져야 한다.

[참고문헌]

1. 김남호, 이봉형, 김일남, 황의천, 왕수현, “인터넷을 통한 연구센터 데이터베이스 시스템 구축”, 연구보고서, Vol.2 No.1, 1997.
2. 김미현, “학술정보시스템의 온라인 인터페이스에 관한 연구 : 적응형 인터페이스를 위한 정보조직 및 활용”, 한국문헌정보학회지 32(4) pp. 259-276, 1998.
3. 김형래, 임기영, “관계형 데이터베이스 관리시스템을 사용한 지식베이스시스템 설계에 관한 연구”, 논문집, 1991.
4. 서은경, “정보검색시스템에서의 이용자 인터페이스 기능에 관한 분석적 고찰”, 정보관리학회지 16(4) pp. 125-150, 1999.
5. 오삼균. 1999. "인터넷 대학강의안의 검색을 위한 Metadata DB 구축."정보관리학회지16(2) : 149-164.
6. 주경희, 김홍식, 서재현, “웹과 DB상의 효율적인 정보검색을 위한 데이터 통합 시스템 설계”, 자연과학 Vol.4 No.1, 2000.
7. 한국 데이터베이스 진흥센터. 1998. "제3부: 메타데이터 명세 및 메타모델의 표준화". 표준화 연구 보고서 Web 데이터베이스의 효율적인 검색을 위한 분류체계 및 초록 작성 지침.
   <http://www.dpc.or.kr/02_biz/report/webdb.html>
8. 한국 데이터베이스 진흥센터. 1999. "제1부: 분야별 정보의 공동활용을 위한 데이터 요소 표준 연구". 표준화 연구 보고서 메타데이터 레지스트리, 데이터 요소 및 디지털 도서관 표준화 연구.
9. 한국과학재단 소식지 Kosefnews. 2002(1).
   <http://kosef.nextdata.co.kr/menu2/c0001/2002_01_F.htm>
10. 한국과학재단. 2001(11). "국내과학기술인력 종합 DB 시스템 개발에 관한 연구."
11. Day, Michael. 1999. "Metadata for digital preservation: an update." Ariadne, 22(December)
12. FIU DL Metadata Creation & Maintenance Manual
    <http://www.fiu.edu/~diglib/metadata/index.html>.
13. Fox et al., “Users, user interface, and objects : envision, a digital library”, Journal of the American Society for Information Science 44(8) pp. 480-491, 1993.

14. Gateway to Educational Materials. 2001 <http://www.geminfo.org>.
15. Hendry, D.G. and Harper, D.J., “An information-seeking environment”, Journal of American Society for Information Science 48(11) pp. 1036-1048, 1997.
16. IEEE P1484.12 Learning Object Metadata Working Group. 2001 <http:ltsc.ieee.org.wg12/index.html>.
17. Peter Rob, Carlos Coronel, “Database Systems - Design, Implementation, and Management, Thomson Learning”, 2000.
18. Weibel, S. 1998. "DC5: The Helsinki Metadata Workshop." D-Lib. Magazine.
19. Weibel, S. 1999. "The State of the Dublin Core Metadata Initiative." D-Lib Magazine.
20. Weibel, Stuart L., Koch, Traugott. 2000. "The Dublin Core Metadata Initiative: Mission, Current Activities, and Future Directions." D-lib Magazine. December.

[인터넷 참고사이트]

국가과학기술인력 종합DB시스템(http://i2s.kordic.re.kr/KHuman/index.jsp)
코리아인터넷 닷컴 <http://korea.internet.com>
한국과학재단지정 의약품연구정보센터 연구인력정보 DB
(http://dric.sookmyung.ac.kr/)
한국법제연구원(http://www.klri.re.kr/)
한국여성개발원(http://www.kwdi.re.kr/)
한국청소년개발원(http://www.youthnet.re.kr/)
한국학술진흥재단(http://www.krf.or.kr/)
http://vancouver-webpages.com/META/VW96-schema.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.metadb.net
http://www.roads.lut.ac.uk/Metadata/DC-ObjectTypes.html
http://www.sosig.ac.uk/help/search.html#type

부 록

메타데이터 형식 참고 자료

1. IEEE Learning Object Metadata (3.6)
2. GEM Element Set and Profile(s)
3. FIU DL Metadata
4. KERIS 학술지 메타데이터
5. KERIS 학술논문 메타데이터

부록 1. IEEE Learning Object Metadata (3.6)		
번호	요 소 명	정 의
1	General	상황의 구애를 받지 않는 자원의 전체적인 일반적 특성
1.1	Identifier	자원의 고유한 식별자(label)
1.2	Title	자원에 부여된 이름
1.3	CatalogEntry	자원에 주어진 기호(designation)
1.3.1	Catalogue	값으로 갖는 문자열의 근원. (ex) ISBN, ARIADNE
1.3.2	Entry	실제 문자열 값
1.4	Language	자원의 언어
1.5	Description	자원의 내용에 대한 설명
1.6	Keywords	자원을 설명하는 키워드
1.7	Coverage	자원의 내용상의 지리적 혹은 시간적 특성
1.8	Structure	자원의 조직상의 기본적인 구조
1.9	Aggregation Level	자원의 기능상의 규모
2	LifeCycle	자원의 life cycle과 관련된 특성
2.1	Version	자원의 버전(판)
2.2	Status	자원의 지위, 상태
2.3	Contribute	자원에 기여한 바가 있는 사람 또는 기관 (저술, 편집, 발행 등)
2.3.1	Role	기여(역할)의 종류
2.3.2	Entity	관련된 개체(사람, 기관)들로서, 관련이 큰 순서대로 기재
2.3.3	Date	기여한 날짜
3	MetaMetaData	자원보다는 오히려 메타데이터 그 자체에 관한 특성
3.1	Identifier	메타데이터의 고유한 식별자(label)
3.2	Catalog Entry	메타데이터 인스턴스에 주어진 명칭(designation)
3.2.1	Catalogue	따라나오는 문자열 값의 근원을 지시
3.2.2	Entry	실제 문자열 값

3.3	Contribute	메타데이터에 기여한 바가 있는 사람 또는 기관
3.3.1	Role	기여의 종류
3.3.2	Entity	연관된 개체(들)로서, 관련이 큰 순서대로 기재
3.3.3	Date	기여한 날짜
3.4	Metadata Scheme	메타데이터 구조의 이름 (version 포함) (ex) LOM-1.0
3.5	Language	메타데이터 인스턴스의 언어
4	**Technical**	자원의 기술적인 특성, 기술적인 요구사항
4.1	Format	자원에 있어서 기술상(technical)의 데이터 유형 (ex) video, html
4.2	Size	디지털 자원의 크기(단위는 byte)
4.3	Location	자원의 위치 또는 그 위치로 변환시키는 방법 (위치가 우선 순위)
4.4	Requirements	자원에 접근하기 위해 필요한 것 (복수일 때는 AND가 논리적 연산자)
4.4.1	Type	requirement의 유형 (ex) browser, OS
4.4.2	Name	요구되는 아이템의 이름
4.4.3	Minimum Version	요구되는 아이템의 가장 낮은 버전
4.4.4	Maximum Version	요구되는 아이템의 가장 높은 버전
4.5	Installation Remarks	자원을 설치하는 방법에 대한 설명
4.6	Other Platform Requirements	다른 소프트웨어 및 하드웨어 요구사항에 대한 정보
4.7	Duration	원래의 속도로 자원을 연속 재생시킬때 걸리는 시간
5	**Educational**	자원의 교육적, 교육학적, 또는 교수법 상의 특징
5.1	Interactivity Type	자원이 지원하는 상호작용의 유형 (ex) simulation, exercise, essays, video clips 등
5.2	Learning Resource Type	자원의 특정 유형. 지배적인 유형을 먼저 기재함 (ex) 표, 그래프, 실험 등
5.3	Interactivity Level	최종 이용자와 자원 간의 상호작용 수준

5.4	Semantic Density	자원의 크기나 시간과 비교하여 자원이 갖는 유용성을 주관적으로 측정
5.5	Intended end user role	자원의 일반적인 이용자. 가장 지배적인 것을 먼저 기재 (ex) author,teacher, learner, manager
5.6	Learning Context	학습자의 일반적인 유형 (ex) Primary Education, Secondary Education 등
5.7	Typical Age Range	일반적인 기대 이용자의 연령
5.8	Difficulty	일반적인 타겟 이용자들이 자원을 이용할 때의 난이도
5.9	Typical Learning Time	자원을 이용할 때 걸리는 대강의 시간
5.10	Description	자원을 이용하는 방법에 대한 주석
5.11	Language	자원의 기대 이용자들이 사용하는 언어
6	**Rights**	자원의 이용에 대한 조건, 지적 재산권
6.1	Cost	자원을 이용할때 비용을 지불해야 하는지의 여부
6.2	Copyright and Other Restrictions	저작권 및 기타 제한점들이 자원의 이용에 적용되는지의 여부
6.3	Description	자원의 이용과 관련된 조건에 대한 설명
7	**Relation**	다른 자원들과의 관계에 있어서의 특성
7.1	Kind	자원들 간의 관계 유형(성질)
7.2	Resource	관계가 지시하고 있는 자원
7.2.1	Identifier	다른 자원의 고유 식별자
7.2.2	Description	다른 자원에 대한 설명
8	**Annotation**	자원의 교육상의 용도에 대한 주석, 논평
8.1	Person	논평을 한 사람
8.2	Date	논평이 만들어진 날짜
8.3	Description	논평의 내용

9	Classificaion	자원의 성격을 분류체계 내의 항목(entry)들로써 기술한 것. 즉, 특정한 분류체계 안에서 그 자원이 위치하는 곳을 설명
9.1	Purpose	분류 항목에 의해 기술되는 자원의 성격 (ex) dicipline, idea, prerequisite, educational objective, skill level 등
9.2	TaxonPath	특정 분류체계 내에서의 taxonomic path
9.2.1	Source	특정한 분류체계 (ex) ACM, MESH, ARIADNE
9.2.2	Taxon	분류체계 내의 항목(entry). 순서 있는 taxon list 에 의해 taxonomic path가 만들어짐 (ex) medicine/diagnostics/instruments
9.2.2.1	Id	해당 분류체계 내에서의 그 taxon의 식별자 (ex)300
9.2.2.2	Entry	taxon에 대한 이름(문자) (ex) Social Science
9.3	Description	기술하고 있는 특성에 대한 설명
9.4	Keywords	특성을 설명하는 키워드. 가장 관련이 큰 것부터 기재

부록 2. GEM Element Set and Profile(s)		
번호	요 소 명	정 의
1	Title	그 자원에 부여된 제목
2	Creator	자원의 내용에 주된 책임을 가진 개체
	namePersonal	creator가 개인일때, 그 이름
	nameCorporate	creator가 회사 또는 단체일 때, 그 이름
	email	E-Mail 주소
	postal	우편 주소
	phone	전화번호
	fax	팩스 번호
	affiliation	creator가 소속된 학회, 단체
	homepage	홈페이지 주소
	role	creator 의 역할에 대한 설명 (free text)
3	Subject	자원의 내용이 지닌 주제(topic). 통제어휘, 공식적인 분류 스킴을 사용할 것이 권장됨
	subjectPath	시소러스나 통제어휘집으로 부터의 용어나 어구
	keywords	자원의 text로 부터 선정한 용어나 어구 (통제되지 않음)
4	Description	자원의 내용에 대한 설명
5	Publisher	자원을 현재의 형태로 이용가능하게 만든 실체
	nameCorporate	자원의 발행자 또는 온라인 상의 배포자 명
	role	역할을 지시하는 말 ("onlineProvider" 또는 "publisher")
	email	E-Mail 주소
	postal	우편 주소
	phone	전화번호
	contact	publisher(기관)의 담당자 이름
	homepage	publisher의 홈페이지

6	**Contributor**	제작자 요소에 명시된 개체 이외에, 자원의 내용에 기여한 책임이 있는 기타 개체
	namePersonal	이름
	nameCorporate	이름
	email	E-Mail 주소
	postal	우편 주소
	phone	전화번호
	fax	팩스 번호
	affiliation	contributor의 배경이 되는 학회, 단체
	homePage	홈페이지 주소
	role	contributor 의 역할에 대한 설명 (free text)
7	**Date**	자원의 존재 기간 동안 어떠한 사건이 발생한 날짜.
	recordCreated	메타데이터 레코드가 작성된 날짜
	created	자원이 처음 생성된 날짜
	modified	자원이 가장 최근에 수정된 날짜
	placedOnline	자원이 처음 온라인 상에 올려진 날짜
	validFrom	자원이 유효하게 된 시작일
	validTo	자원이 유효한 마지막 날짜
	created	자원이 처음 생성된 날짜(간단한 textual 지시)
	modified	자원이 가장 최근에 수정된 날짜(간단한 textual 지시)
	placedOnline	자원이 처음 온라인 상에 올려진 날짜(간단한 textual 지시)
8	**Type**	자원 내용의 성격 또는 장르 (ex) homepage, novel, working paper, technical report 등
9	**Format**	자원의 물리적 표현형식 및 디지털 표현형식
	contentType	자원의 파일 포맷
	platform	플랫폼
	fileSize	파일 크기(byte로 표시)
10	**Identifier**	자원을 가리키는 지시자

	locSpec	자원을 찾도록 도와주는 모든 식별자, 또는 자원의 위치를 알 수 있도록 변환시켜주는 메커니즘
	public_id	일반적으로 널리 알려진 고유 식별자
	SID	GEM의 자원에 대하여 GEM에서 부여한 식별자
	SDN	GEM 회원의 특정 자원을 위하여 부여한 회원의 고유 식별자
11	Source	현재 자원의 출처가 되는 원정보자원으로의 참조
	description	자원에 대한 간단한 설명
	version	버전 넘버
	locSpec	출처 자원을 찾도록 도와주는 모든 식별자, 또는 출처 자원의 위치를 알 수 있도록 변환시켜주는 메커니즘
	public_id	일반적으로 널리 알려진 고유 식별자
12	Language	자원의 지적인 내용을 기술하고 있는 언어
	code	사용 언어를 나타내는 코드 값
	text	사용 언어의 원래 이름
13	Relation	관련 자원들로의 참조
	kind	기술되고 있는 자원과 관련 자원간의 관계 유형을 나타내는 용어
	description	관계에 대한 간단한 설명
	locSpec	지시되고 있는 자원의 위치 또는 그 위치로 변환시켜주는 방법(URL, DDC number)
	public_id	일반적으로 널리 알려진 고유한 지시자
14	Coverage	자원의 내용이 다루고 있는 범위 (시간적, 공간적)
	spatial	내용상의 공간적 특성 (free text)
	temporal	내용상의 시간적 특성 (free text)
15	Rights	그 자원에 대해서 갖고 있는 권리에 관한 정보
	locSpec	자원의 지적재산권 또는 자원의 이용에 있어서의 제한점을 언급하고 있는 진술의 위치, 또는 그 위치로 변화시켜주는 값
	agent	권리를 소유하고 있는 에이전트 명
	use	자원의 지적재산권 또는 자원의 이용에 있어서의 제한점 기술
	priceCode	자원의 비용을 지시하는 코드

1	GEM.Audience	자원에 대한 특정 이용자 (통제 어휘 사용)
	toolFor	자원을 사용하는 집단 (ex: "teacher")
	beneficiary	자원의 궁극적인 수혜자 (ex: "disabled student")
2	GEM.Cataloging	GEM 목록 레코드를 작성한 기관에 관한 기본 정보
	nameCorporate	목록 기관명
	namePersonal	목록작성자의 이름
	email	목록 기관의 이메일 주소
	postal	목록 기관의 우편 주소
	phone	목록 기관의 전화번호 (표준 형식 준수)
	fax	목록 기관의 팩스번호 (표준 형식 준수)
	homePage	목록 기관의 홈페이지 주소
	application	자원의 목록기술에 사용된 어플리케이션 (ex: "GEMCat")
	version	어플리케이션의 버전 번호
3	GEM.Duration	자원에서 설명하고 있는 대로 강의나 활동(activity)을 따라할 때, 소요되는 시간.
4	GEM.Essential Resources	선생님이 그 자원을 효과적으로 사용하기 위해서 반드시 필요한 자료 (free text 리스트)
5	GEM.Grade	이용자의 등급(학년), 학년 범위, 교육상의 레벨, 나이
	grade	자원이 특정한 학년 또는 학년 범위를 대상으로 하는 경우에 그에 관한 정보를 기술
	level	학년과 대조하여, 자원이 어떤 레벨에 초점을 맞추고 있는 경우
	age	특정 나이 또는 나이 범위를 대상으로 하는 경우
6	GEM.Pedagogy	교수 방법, 평가 방법, 지도를 위해 학생들을 그룹화 하기(조편성), 선수과목 등에 대한 정보
	grouping	학습 지도를 위해 학생들을 그룹화하는 것에 대한 지시
	teachingMethods	교수 방법에 대한 지시
	assessment	학생들의 평가 방법에 대한 지시
	prerequisite	학습에 필요한 선수과목(사전에 알고 있어야 할 지식)에 대한 지시
7	GEM.Quality	교육 자료의 질을 평가하기 위한 수단이 되는 항목임
	criteria	자원을 평가하고 등급을 매기는 데 기초가 된 다양한 기준(criteria)

	category	등급 (ex : "poor", "fail", "average", "good", "excellent" 등)
	authority	교육자료의 질을 평가한 기관 또는 사람의 이름
	scale	평가에 사용된 척도에 대한 간략한 설명 (ex : "1-5")
	homePage	평가 기관이 그러한 평가 기준을 사용한 것에 대한 설명이 온라인 상에 있을때 그 주소
8	GEM.Standards	기술되고 있는 자원에 매핑될 수 있는 국가적인 학문 표준이 있을때
	authority	자원을 그러한 표준과 연결시킨 기관 또는 사람의 이름
	correlation	자원이 그 표준과 어느 정도로 상관관계가 있는지에 대한 진술 (ex : "slightly correlated")
	identifier	표준을 제정한 기관에서 부여한 그 표준의 지시자
	discipline	표준의 주제 분야 (ex : "Science", "Mathmatics" 등)
	level	표준에서 밝히고 있는 등급, 수준 (자원의 등급, 학년 범위와는 상응하지 않을 수도 있음)
	standardPath	계층안에서의 연속적 기호

부록 3. FIU DL Metadata		
번호	요 소 명	정 의
1	Record Identifier	데이터베이스 내의 각 레코드에 할당된 고유한 숫자 (9자리 수)
2	Record URL	Record URL은 외부의 서버에서 FIU 디지털 도서관의 특정 자원으로 접근하고자 할 때 필요한 접근점임. 그러므로 Record Identifier가 도서관 내부에서 자원의 탐색에 사용되는 것과 구별됨
3	Access Restrictions	이 레코드에 대한 접근이 FIU 커뮤니티로만 제한된다면, "FIU community only" 라는 표현을 기재해야 함
4	Title	자원의 저자가 부여한 자원의 이름 또는 제목
5	Title, Alternative	이용자가 이 자원을 찾기 위해서 사용할 수도 있다고 여겨지는 모든 방식의 제목(주요 Title과는 다름)
6	Creator	자원의 지적 내용의 저작에 주요한 책임을 가지고 있는 사람 또는 기관 (책의 저자, 사진가, 연주가 등등)
	creator qualifier	생몰년, creator의 완전한 이름, 컨퍼런스정보 등, creator를 구분할 수 있는 정보
	creator role	자원과 creator의 관계
7	Contributor	creator 요소에 명시된 개체 이외에, 자원의 내용에 대해 2차적인 책임이 있는 기타 개체
	contributor qualifier	생몰년, contributor의 완전한 이름, 컨퍼런스정보 등을 구분할 수 있는 정보
	Contributor role	자원과 contributor의 관계
8	Publisher	자원을 현재의 형태로 이용 가능하게 만든 실체
9	Language	자원의 지적 내용에 사용된 언어. 비디오나 오디오 자료의 구어도 이에 해당. language code 와 language 명으로 구성됨
10	Collection Name	기술되고 있는 자원이 보다 큰 합집(전집)에 속해있을 때, 그 합집(전집)의 고유한 title
11	Subject	자원의 내용이 지닌 주제(topic)
	Subject thesaurus	subject를 위해 사용된 시소러스
12	Description	제목만으로는 알 수 없는, 자원의 내용에 관한 설명

13	Relation	기술되고 있는 자원이 FIU나 그 밖의 다른 곳에 있는 자원과 어떤 관계가 있는 지를 기술. (공식적 관계 : 버전관계, 부분/전체관계 등, 비공식적 관계 :이용자에게 흥미있을 법한 다른 관련 자원)
	relation number	relation 요소에서 기술되고 있는 자원의 record identifier
14	Public Note	다른 메타데이터 요소에는 적합하지 않은 부가적인 정보를 제공하거나, 통제어휘를 사용하는 어떤 요소를 더 상세히 기술하고자 하는 경우에 이 요소를 사용
15	Cataloger's Note	메타데이터의 생성 및 유지 관리, 또는 자원의 관리를 맡고 있는 사람들을 위해서 목록자가 제공하는, 내부적 관리를 위한 주석문. 이용자에게는 보이지 않음
16	Source Standard Number	자원과 관계된 표준의 번호. ISBN, ISSN, Sudoc number 등
17	Source	현재의 디지털 자원을 생성하는 데에 사용되었던 출처 자원에 대한 기술 (제목, 발행정보, 표준번호, 물리적 특성에 관한 설명 등)
	date of source	출처 자원의 생성일
	end of source date range	기간인 경우 마지막 날짜
	date type	날짜 기술의 유형
	date of digitization	어떤 개체가 디지털 형태로 변환되었거나 디지털 형태로 만들어진 날짜
	part number	하나의 자원에 여러 개의 이미지, 사운드 파일 등이 속할 때, 각 부분에 대한 번호 또는 파일명
	instance	각 부분에 대한 하나 이상의 인스턴스에 관한 정보 (ex : archive copy, internet copy, thumbnail)
	format	데이터/파일의 형태. 필요한 소프트웨어, 하드웨어를 알려주기 위한 정보임
	file size	자원의 인터넷 인스턴스에 있어서의 파일 크기
18	Resource Type	자원의 지적 내용의 주요 유형 및 범주. (2가지 구성 단계를 가짐). (ex : video-interview)

19	Subject Coverage	FIU를 위한 고유 범주로서, 주제 브라우징의 계층을 위해 만들어졌음. subject coverge는 Florida 주립대학에서 사용되고 있는 Florida Statewide Course Numbering System을 기초로 함
20	Coverage, Place	자원의 내용이 다루고 있는 지역, 위치(발행지가 아님). 네가지 계층 중 첫번째로서, 대륙에 해당
	Coverage, Place-2nd level	국가
	Coverage, Place-3rd level	주
	Coverage, Place-4th level	도시
21	Coverage, Time	자원의 내용이 다루고 있는 시대(발행일이 아님)
	coverage,time-beginning date	기간의 첫 날
	coverage,time-end date	기간의 마지막 날

부록 4. KERIS 학술지 메타데이터		
1. 표제사항		
1.1	Title	본표제
1.2	Title.parallel	대등표제
1.3	Title.subtitle	기타표제
2. 저자사항	**저자사항**	
2.1	Creator.personalName	개인저자명
2.2	Creator.personalName.alternative	개인저자의 다른 이름
2.3	Creator.personalName.email	개인저자의 이메일
2.4	Creator.personalName.affiliation	개인저자의 소속
2.5	Creator.personalName.homepage	개인저자의 홈페이지
2.6	Creator.personalName.major	개인저자의 전공 및 학과명
2.7	Creator.personalName.fax	개인저자의 팩스
2.8	Creator.personalName.phone	개인저자의 전화번호
2.9	Creator.corporateName	단체저자명
2.10	Creator.corporateName.alternative	단체저자의 다른이름
2.11	Creator.corporateName.email	단체저자의 이메일
2.12	Creator.corporateName.affiliation	단체저자의 소속

2.13	Creator.corporateName.homepage	단체저자의 홈페이지
2.14	Creator.corporateName.fax	단체저자의 팩스
2.15	Creator.corporateName.phone	단체저자의 전화번호
3. 주제사항	**저자사항**	
3.1	Subject	주제어
	FreeText	프리텍스트
	Theme	theme
	LCSH	LCSH
	Degree	Degree
	DDC	DDC
	LCC	LCC
	UDC	UDC
	KDC	KDC
4. 기술사항	**저자사항**	
	Description	주기사항
	FreeText	프리텍스트
	URL	URL
	URN	URN
	Description.abstract	초록
	FreeText	프리텍스트
	URL	URL
	URN	URN
	Description.tableOfContents	목차
	FreeText	프리텍스트
	URL	URL
	URN	URN

5. 발행사항	**발행사항**	
	Publisher	발행자명
	Publisher.alternative	발행자의 다른 이름
	Publisher.email	발행자의 이메일
	Publisher.homepage	발행자의 홈페이지
	Publisher.address	발행자의 주소
	Publisher.fax	발행자의 팩스
	Publisher.phone	발행자의 전화번호
	Publisher.country	발행 국가
6. 기타저자사항	**기타저자사항**	
	Contributor.personalName	개인 기여자명
	Contributor.personalName.alternative	개인 기여자의 다른 이름
	Contirbutor.personalName.email	개인 기여자의 이메일
	Contributor.personalName.affiliation	개인 기여자의 소속 기관
	Contributor.personalName.role	개인 기여자의 역할
	Contributor.personalName.homepage	개인 기여자의 홈페이지
	Contributor.personalName.fax	개인 기여자의 팩스
	Contributor.personalName.phone	개인 기여자의 전화번호
	Contributor.corporateName	단체 기여자명
	Contributor.corporateName.alternative	단체 기여자의 다른이름
	Contributor.corporateName.email	단체 기여자의 이메일

	Contributor.corporateName.affiliation	단체 기여자의 소속 기관
	Contributor.corporateName.role	단체 기여자의 역할
	Contributor.corporateName.homepage	단체 기여자의 홈페이지
	Contributor.corporateName.fax	단체 기여자의 팩스
	Contirbutor.corporateName.phone	단체 기여자의 전화번호
7. 자원관련일자	**자원관련일자**	
	Date.Created	발행년
	ISO 8601	
	Date.discontinued	폐간년
	ISO 8601	
	Date.republication	복간년
	ISO 8601	
	Date.metadataCreated	메타데이터 최초작성일
	ISO 8601	
	Date.metadataModified	메타데이터 최종수정일
	ISO 8601	
8. 자원유형	**자원유형**	
	Type	자원의 범주나 장르
	Type.interval	자원의 현재 간기
9. 자원형태	**자원형태**	
	Format	자원의 형태
	IMT	
	Format.page	자원의 전체 페이지 수

10. 식별자	**식별자**	
	Identifier	식별 항목
	ISSN	
	LCCN	
	URL	
	URN	
11. 원정보자원		
	Source	원정보자원
	FreeText	
	ISSN	
	Source.type	원정보자원의 유형
	Source.volume	원정보자원의 권호정보
	Source.interval	원정보자원의 간기
12. 언어사항	**Language**	
	Language	기술언어
13. 관계		
	Relation	관련자원
14. 지역사항	**Coverage**	
	Coverage.country	자료에서 취급된 지역 및 시대
15. 정보이용사항	**Rights**	
	Rights	원문서비스의 이용현황
	Rights.start	원문서비스 개시일
	Rights.expire	원문서비스 만료일

부록 5. KERIS 학술논문 메타데이터		
번호	**요 소 명**	**정 의**
1	**서명사항**	
1.1	Title	논문 원제목
1.2	Title.parallel	논문 대등 제목
1.3	Title.subtitle	기타 표제
2	**저자사항**	
2.1	Creator.personalName	개인저자명(저자, 편저자)
2.2	Creator.personalName.alternative	개인 저자의 다른 이름
2.3	Creator.personalName.email	개인 저자의 이메일
2.4	Creator.personalName.affiliation	개인 저자의 소속
2.5	Creator.personalName.homepage	개인 저자의 홈페이지
2.6	Creator.corporateName	단체저자명
2.7	Creator.corporateName.alternative	단체저자의 다른 이름
2.8	Creator.corporateName.email	단체저자의 이메일
2.9	Creator.corporateName.affiliation	단체저자의 소속
2.10	Creator.corporateName.homepage	단체저자의 홈페이지
3	**주제사항**	
3.1	Subject	
	FreeText	자유주제어
	KDC	KDC 분류기호

	LCC	LCC 분류기호
	UDC	UDC 분류기호
	THEME	주제분류
	LCSH	LCSH
4	**기술사항**	
4.1	Description	기술사항
	FreeText	자유기술어
	URL	자유기술어 URL
	URN	자유기술어 URN
4.2	Description.abstract	초록
	FreeText	초록
	URL	초록에 대한 URL 정보
	URN	초록에 대한 URN 정보
4.3	Description.content	목차
	FreeText	목차
	URL	목차에 대한 URL 정보
	URN	목차에 대한 URN 정보
5	**발행사항**	
5.1	Publisher	발행자 이름
5.2	Publisher.alternative	발행자의 다른 이름
5.3	Publisher.email	발행자의 이메일
5.4	Publisher.homepage	발행자의 홈페이지
5.5	Publisher.place	발행국명

6	**기타저자사항**	
6.1	Contributor.personalName	개인 기여자명
6.2	Contributor.personalName.alternative	개인 기여자의 다른 이름
6.3	Contributor.personalName.email	개인 기여자의 이메일
6.4	Contributor.personalName.affiliation	개인 기여자의 소속기관
6.5	Contributor.personalName.role	개인 기여자의 역할
6.6	Contributor.corporateName	단체 기여자명
6.7	Contributor.corporateName.alternative	단체 기여자의 다른 이름
6.8	Contributor.corporateName.email	단체 기여자의 이메일
6.9	Contributor.corporateName.affiliation	단체 기여자의 소속기관
6.10	Contributor.corporateName.role	단체 기여자의 역할
7	**날짜 사항**	
7.1	Date.created	발행년
7.2	Date.metadata.created	메타데이터 최초 작성일
7.3	Date.metadata.modified	메타데이터 최종 수정일

8	**자원 유형**	
	Type	자원의 범주나 장르(연구논문, 대학석박사학위 논문 등)
9	**자원 형태**	
	Format	자원의 형태
	Format.page	시작-끝 페이지
	Format	자원의 형태
	Format.page	전체페이지 수
10	**식별자**	
	Identifier	식별자
11	**원정보자원**	
	Source	
	FreeText	원정보자원의 서명
	ISSN	원정보자원의 ISSN
	Source.Type	원정보자원의 유형
	Source.volume	원정보자원의 권호정보
	Source.interval	원정보자원의 간기
12	**언어사항**	기술언어