

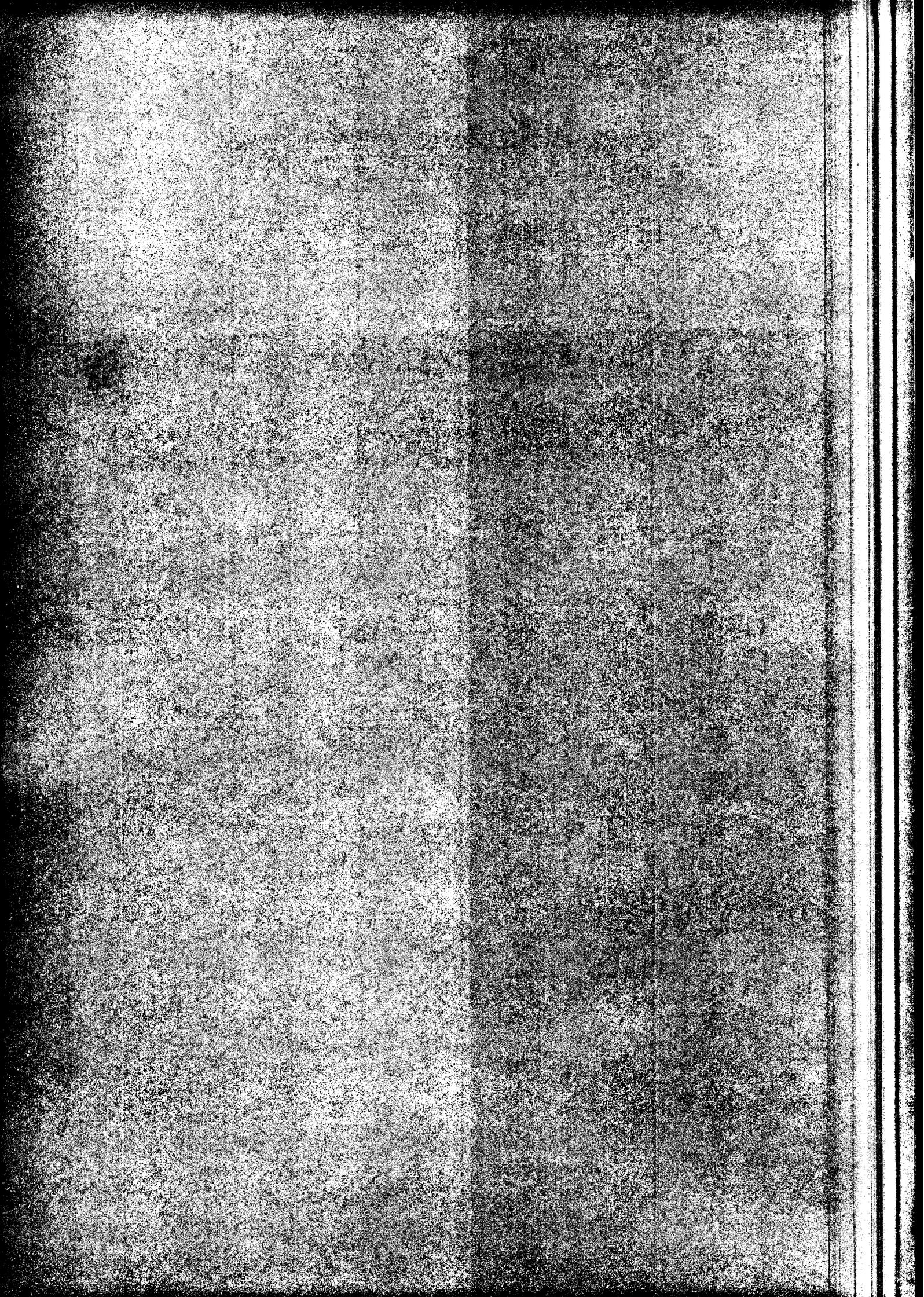
【서명표 1】



2003. 5

김성중

capkimo@anu.ac.kr



< 사례발표 1 >

서울대학교전자도서관의 기본 방향

2003. 5.

김성중
(capkimsj@snu.ac.kr)

차 례

1. 전자도서관의 개념	1
2. 시스템 구성(소프트웨어).....	1
3. 시스템 구성(장비).....	2
4. 콘텐츠 구축	2
5. 범용뷰어 사용	3
6. 서울대 구축 콘텐츠 서비스 방법.....	3
7. 학위논문 원문정보 상호이용에 관한 공동협약.....	3

1. 전자도서관의 개념

전자도서관은 광의 또는 협의의 개념에 따라 digital library, electronic library, virtual library, library without the wall, paperless library, multimedia library 등 다양한 견해가 있다. 서울대학교에서는 전자도서관을 콘텐츠 중심의 협의의 개념으로 사용하고 이를 점차 확대하여 나감.

1.1. 콘텐츠를 온라인을 통하여 서비스

- 도서관서비스 중 원문 콘텐츠를 서비스
- 도서관 서비스의 일부분
- 캠퍼스내 각 기관의 학술자료를 대상으로 함
도서관의 전자도서관 시스템이 아니라 서울대학교전자도서관시스템
<http://sdl.snu.ac.kr/>

1.2. 자체구축 콘텐츠, 구독 콘텐츠 포함

- Web DB, 전자저널, CD-ROM 등

2. 시스템 구성(소프트웨어)

2.1. 도서정보시스템을 기본으로하고 전자도서관은 도서정보시스템의 서브시스템

- 도서정보시스템의 UP-grade : SOLARS II

2.2. 도서정보시스템, 장비, 전자도서관시스템을 독립적으로 분리

2.3. 소프트웨어는 분야별 전문업체의 솔루션을 독립적으로 구성

- 메타데이터 및 콘텐츠관리
- 저작권 보호 : 소프트웨어, Watermark, DRM(Digital Rights Management)
- AOD/VOD 시스템
- 전자저널 통합검색
- 캠퍼스 웹문서 검색

2.4. 서지데이터 MARC를 기반으로 함

- 메타데이터는 MARC를 기반으로하고 MARC로 구축하기 적합지 않은 콘텐츠는 메타

데이터로 구축한다.

- 필요에 따라 서지(MARC)와 메타데이터 이원화
- 메타데이터는 XML스키마로 한다(국내 전자도서관시스템에서 최초 적용)

3. 시스템 구성(장비)

3.1. 시스템 장비는 도서정보시스템과 전자도서관시스템을 통합하여 구성한다.

- UNIX 서버 : 5대(메인서버, 이미지서버, 웹서버, 보조서버)
- NT 서버 : 5대
- 스토리지 10 TB(EMC 스토리지 2대)

3.2. 장비와 소프트웨어는 분리한다.

4. 콘텐츠 구축

4.1. 콘텐츠 구축 : 보존용 포맷과 서비스용 포맷으로 이원화

- 보존용 포맷 : 고�해상도
- 서비스용 포맷 : 저해상도
서비스포맷의 기술발전에 따라 보존용 포맷으로 서비스포맷 변경가능

4.2. 캠퍼스내 전체 학술자료를 대상으로 함.

도서관, 단과대학, 학과, 연구소, 박물관, 대학기록관 사료 등

4.3. 콘텐츠 구축량

- 서울대학교수여 학위논문
- 서울대학교 간행물
- 고문헌(1945년 이전자료)
- 학내 각 기관 소장자료

<기관소장 학술자료>

연 도 별	구 축 량	비 고
2003년 5월 현재	6,944,833면(원문) 76,211분(VOD/AOD)	
2003년 구축 예정	870,000(면) : 발주중 430,000(면) : 하반기 예정	발주중: 12월 까지 구축예정
2004-2005년	800,000면	
계	약 1천만 면 예상	

5. 범용뷰어 사용

- Adobe Acrobat Reader
- Windows에서 일반적으로 사용하는 뷰어
- 기타 Web에서 일반적으로 사용하는 뷰어
- 데자뷰(컬러이미지) : Windows에 플러그인으로 자동설치

6. 서울대 구축 컨텐츠 서비스 방법

- 6.1. 서울대 홈페이지에 접속하여 이용(저작권상 제한공개 컨텐츠 제외)
- 6.2. 각 대학에 원문정보링크데이터 제공

7. 학위논문 원문정보 상호이용에 관한 공동협약

- 7.1. 제공컨텐츠 : 학위논문, 서울대학교간행물, 학술자료, 기타
- 7.2. 공동 협약 대상기관
 - 전국국사립 대학도서관
 - 국가기관 및 전문도서관

7.3. 공동협약 절차

- 서울대와 공동협약 체결

서울대와 공동이용협약 체결 : CD로 제공

- 다자간 공동이용협약 체결

-공동이용협약 체결 대학 확정후 다자간 공동이용협약 체결

-다자간 협약체결 후 : 협약대학간 원문링크 정보데이터 제공

7.4. 사용방법

- 각대학도서관 시스템에 올려 자관 DB와 통합검색

- 캠퍼스내에서 사용

7.5. 공동이용에 관한 협의체 구성

- 공동이용협약 대상대학 확정후 구성

- 공동이용 방법 및 운영에 관한 사항

7.6. 공동이용에관한 연계방안

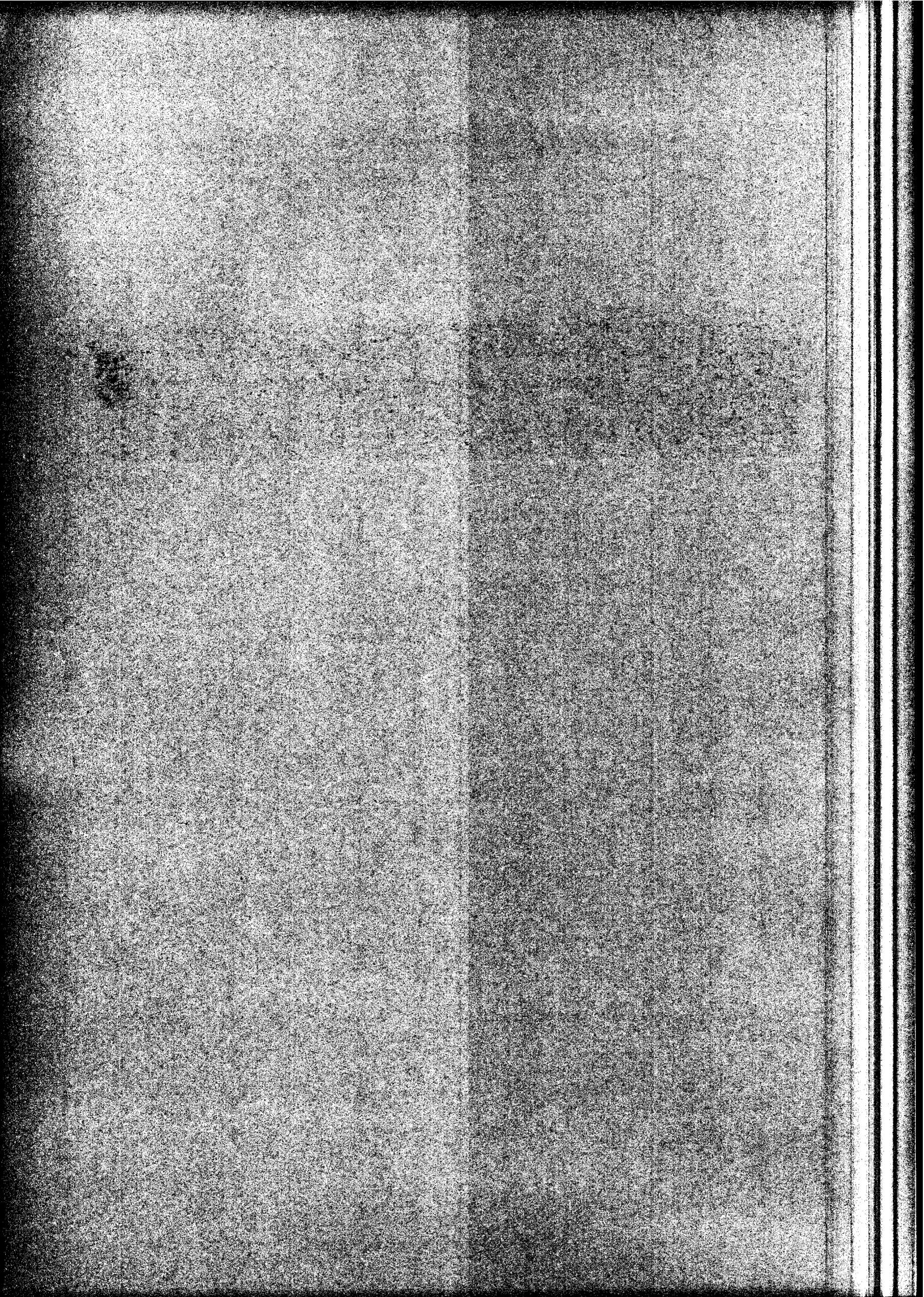
- 한국교육학술정보원 : 연계자금지원

- S/W개발회사, 각 대학

【 사례 발표 2 】

2003. 5

배종학
anslis@snu.ac.kr



< 사례발표 2 >

서울대학교전자도서관
학술정보 System 구성 및 도입 사례

2003. 5.

배종학
(anstls@snu.ac.kr)

차 례

1. 시스템 도입 배경.....	1
2. 시스템 도입 방향.....	1
3. 시스템과 S/W 분리 계약의 장단점	2
4. 시스템 도입과정	2
5. 시스템 구성도	5
6. 시스템 구성 특징.....	6

1. 시스템 도입 배경

- 1.1. 서지 중심의 학술정보와 전자도서관 운용을 위해서 고성능의 시스템 도입 필요
- 1.2. 분관, 학과 및 연구소 DB 통합과 다양하고 광범위한 멀티미디어 콘텐츠 구축에 따른 대용량 데이터 처리와 통합 검색시 처리속도 향상이 필요
- 1.3. 종합학술정보 및 전자도서관 서비스 시행시 예상되는 사용자수 증가를 수용할 수 있는 시스템 필요

1.3.1. 예상 사용자수

일반 관리자	중앙도서관, 분관, 학과 및 연구소	
	예상 사용자수	500 User
	동시 사용자수	단위업무당 최고 50 User 예상 (목록DB연관)
캠퍼스 내 사용자	학생 32,384명, 교수 2,012명, 직원 999명 (2003.4.1 기준)	
	예상 사용자수	35,395 User
	동시 사용자수	최고 283 User 예상 (총 인원의 0.8%)
캠퍼스 외 사용자	예상 사용자수	약 10만 User
	동시 사용자수	최고 5000 User 예상 (총 인원의 0.5%)

- 1.4. 복잡한 시스템의 효율적인 운용과 안정적인 데이터 처리가 가능해야 함

2. 시스템 도입 방향

- 2.1. 종합학술정보시스템과 전자도서관시스템의 효율적으로 수용 가능한 구조를 위해 상호 유기적 연관 관계를 수용할 수 있는 아키텍처로서 3-tier 방식의 설계를 지원
- 2.2. 분관 DB 및 전자도서관의 대용량 데이터 구축 및 활용을 위해 대용량 데이터 저장 및 고속 처리 장비 도입

- 2.3. 인터넷을 통한 다수의 이용자를 고려해 웹 중심의 서비스를 위한 구성 지원
- 2.4. 도입 후 최소 5년 간의 데이터 및 성능에 대한 보장성 확보
- 2.5. 장비 도입에 있어 중복된 부분의 최소화를 통한 구축예산의 효율적 활용과 유지보수 비용 최소화

3. 시스템과 S/W 분리 계약의 장단점

3.1. 장점

- 3.1.1. 시스템만 구입함으로써 시스템 전문업체끼리의 경쟁을 통해 성능이 우수한 시스템을 싸게 구입
- 3.1.2. 시스템을 먼저 구입한 후 S/W를 개발하였기 때문에 특정시스템에 종속되지 않는 S/W 개발이 가능하며 S/W 교체시에도 시스템은 교체하지 않고 사용 가능

3.2. 단점

- 시스템을 먼저 도입 후 전자도서관 S/W 개발시 추가적인 H/W가 필요할 수 있음

4. 시스템 도입과정

- 시스템도입 자문위원들과 전산소 시스템실의 도움과 협의를 통해 시스템 구성에 있어서 최적의 성능과 가격 결정이 가능하였음

- 4.1. 중앙도서관 전산실 자체적으로 시스템 구성 및 사양 검토 : 2001. 11월 ~ 12월
- 서버업체와 스토리지 업체들과 협의를 하였고, 타대학 시스템 구성 검토

4.2. 시스템 도입 규격검토위원 구성

- 4.2.1. 규격검토위원회 위촉(2001년 12월) - 자문위원 교수님 9명

- 4.2.2. 전자도서관 시스템 규격 검토 회의 2차례 실시(2001년 12월, 2002년 1월)

4.2.3. 참석자 : 도서관장, 학술정보원장, 자문위원 교수님(9명), 전산실 직원

4.2.4. 규격 검토 회의시 고려 사항

- 공통스펙으로 작성으로 특정업체 규격 배제하여 경쟁입찰 고려
- 유지보수기간을 3년으로 하여 유지보수비용 문제 해결
- 조달청 입찰 방식을 1차 기술평가 후 2차 최저가 입찰방식으로 하여 도서관 시스템에 적합한 H/W가 도입될 수 있도록 고려
- 도서정보 시스템의 중요성에 따른 고성능 시스템 구성 검토
- 향후 H/W 업그레이드시 계약당시의 D.C율을 적용하도록 함
- Unix 서버는 Sun, IBM, HP업체 고려
- NT서버는 Compaq, Dell, HP 등 국내판매 순위 10위까지 고려
- 스토리지는 EMC, 히타치, Sun, IBM, HP 등 스토리지 전문업체 고려

4.3. 시스템 도입과 관련하여 전산소 시스템실과 업무협의

4.3.1. 2001년 12월부터 2002년 1월까지 9차에 걸쳐서 업무협의

4.3.2. 협의내용

- UNIX서버를 도메인 기능이 가능한 한 대의 통합시스템 구성
- 스토리지를 EnterPrise급으로 구성
- 백업방안 검토(StorageTek)하여 전산소 시스템실 백업장비 사용
- 각 업체에게 도서관에 맞는 시스템 구성을 제안하게 한 뒤 협의

4.4. 시스템 도입 제안요청서 완성(2002. 1. 28)

- 제안서 60%, 가격 40% 적용 : 성과와 기술면에서 적합한 업체를 대상으로 가격 개찰하는 “규격,가격분리입찰”로 진행

4.5. 학술정보시스템 및 전자도서관 시스템 H/W 도입(조달구매) 추진(2002. 1. 31)

4.5.1. 제안서 평가를 통한 협상에 의한 계약체결 방법 요청

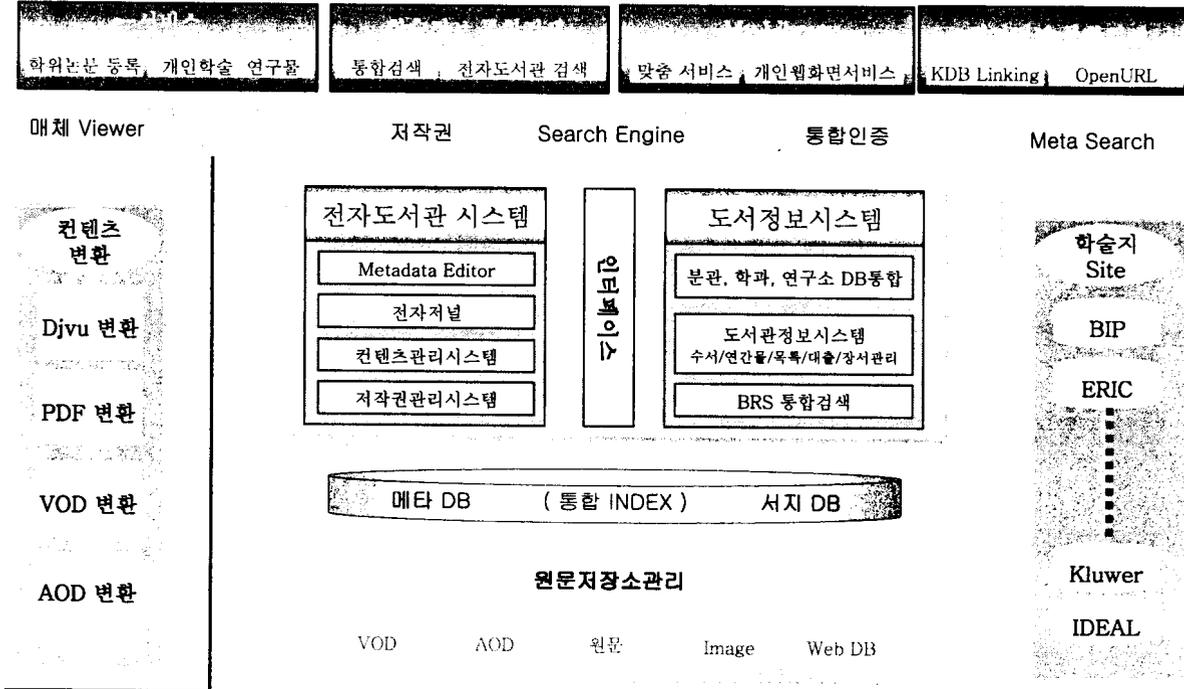
4.5.2. 개발된 학술정보시스템 환경(Oracle, BRS, Web Agent 등)을 수용하여야 하며, 도서관 데이터관작업 및 향후 운영시 업체의 기술력 평가가 필요

4.6. 전자도서관 시스템 가격 협상 회의 : 2002. 2. 19

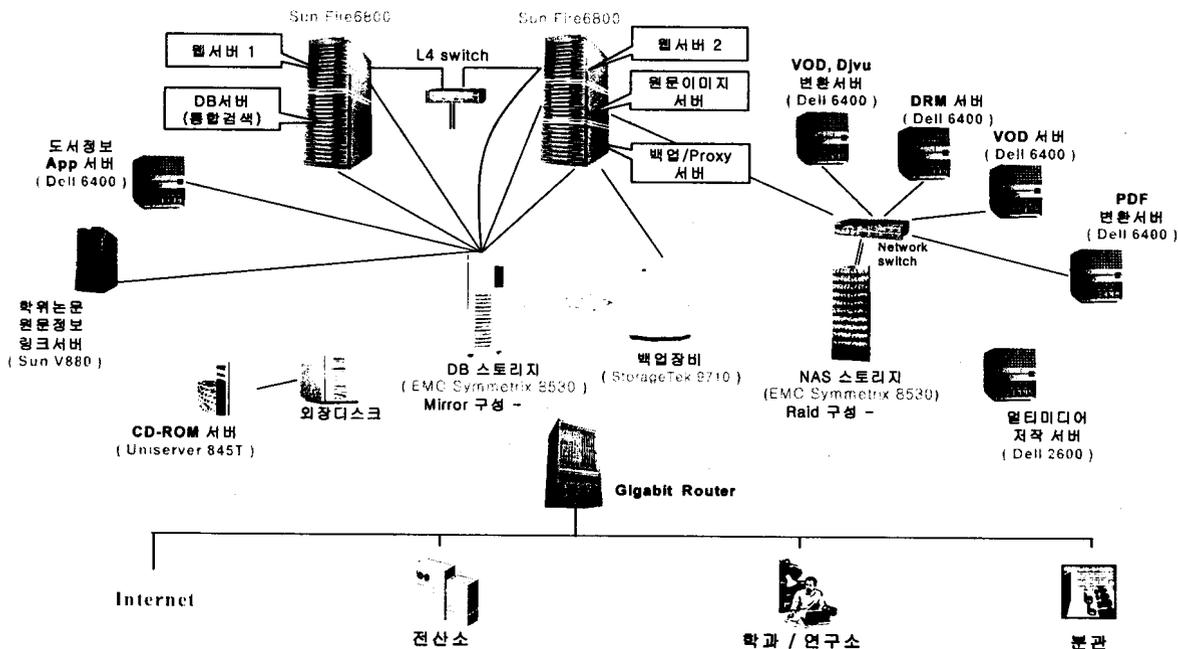
- 4.7. 전자도서관 시스템 제안서 평가위원(11명) 위촉 및 평가회의(2002.2.26)
- 4.8. 시스템 계약 체결(2002. 5. 13)
- 4.9. 시스템 설치 전 사전작업
 - 4.9.1. 전자도서관 시스템 설치일정 계획 작성 (2002.5.21)
 - 4.9.2. 시스템 도입 전에 분전반 및 전기배선 공사 (2002.5.20 ~ 5.23)
 - 4.9.3. 시스템 설치 및 UPS 연결 작업 (2002.5.24)
 - 4.9.4. 시스템에 OS 설치 및 환경 설정 (2002.5.27 ~ 5.31)
 - 4.9.5. S/W 설치 및 데이터 이관 작업 (2002.6.1 ~ 6.5)
 - 4.9.6. 시스템 통합 테스트 및 안정화 작업 (2002.6.7 ~ 6.10)
 - 4.9.7. 향온향습기 1대 추가 설치하여 2대로 시스템 운영 (2002.6.18)
- 4.10. 학술정보/전자도서관 시스템 설치 시작 및 검수계획 작성 (2002.5.27 ~)
- 4.11. 학술정보/전자도서관 시스템 설치에 따른 검수 완료 (2002. 6. 27)

5. 시스템 구성도

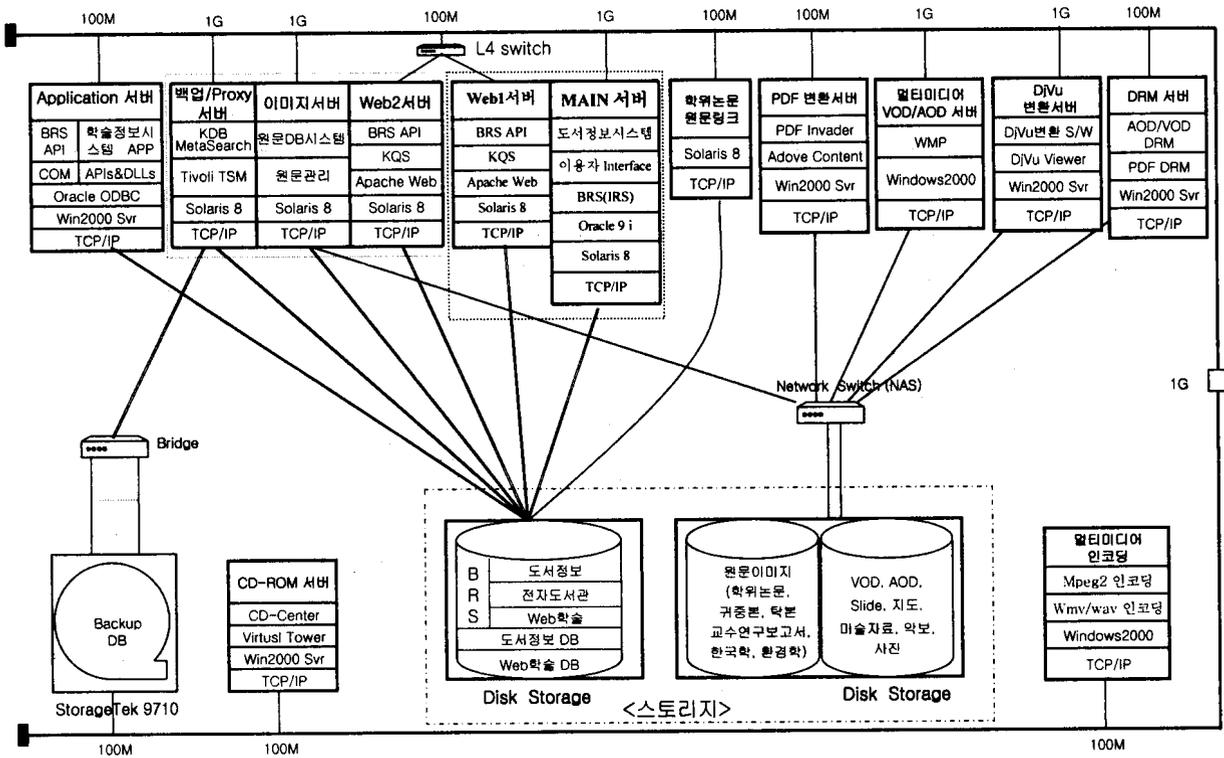
5.1. 시스템구성도



5.2. H/W 구성도



5.3. S/W 구성도



6. 시스템 구성 특징

6.1. 도메인 기능의 Unix 통합서버를 도입하여 여러대의 서버로 분리하지 않고 통합 구성

6.1.1. 통합서버는 고사양 아키텍처로 분리된 여러 서버에 비해 성능이 우수함

6.1.2. 관리비용 절감과 공간 절약

6.1.3. 자원 재할당 가능 : CPU, 메모리를 다른 서버로 재할당이 가능하여 서버에 부하가 많을 경우 시스템 자원에 여유있는 서버의 메모리와 CPU 할당이 가능

6.1.4. 손쉬운 확장 : 새로운 서버를 도입하고자 할때 통합서버에 도메인을 추가하여 추가 서버 구성이 가능

6.2. 웹서버 2대 구성

6.2.1. 웹 서버 2대를 L4스위치를 통해 분산하여 서비스 속도 향상

6.2.2. 웹서버 2대 중 1대에 장애가 발생하더라도 LA 스위치를 통해 웹서버 1대로 서비스가 계속되어 웹서비스에 중단이 없도록 구성

6.3. 스토리지 도입 : 대용량 디지털 콘텐츠 수용 및 안정성 고려

6.3.1. 서버와 스토리지를 분리하여 서버 장애시에도 데이터는 안전하게 스토리지 저장

6.3.2. 스토리지 자체 백업 S/W를 사용하여 스토리지 여분의 공간에 데이터 백업 활용

6.3.3. 대용량 데이터를 스토리지에서 관리하여 데이터 IO에 대한 서버 부하를 줄여 서버의 서비스 속도 증가

6.3.4. 서버에 독립적이기 때문에 서버(사용기간 : 4~5년) 교체시 스토리지(사용기간 : 6~7년)는 계속 사용 가능

6.3.5. 모뎀을 통해 온라인 유지보수가 가능하여 신속하게 장애 발견 및 대처가 가능

6.4. DB 저장용 스토리지와 콘텐츠 저장용 스토리지를 분리 도입

6.4.1. DB 저장용 스토리지

6.4.1.1. DB 저장용 스토리지는 데이터 안전성을 고려하여 Mirror구성의 스토리지를 도입하여 한 쪽 디스크에 장애시에도 데이터 및 서비스에는 문제가 없음

6.4.1.2. NT서버, Unix서버 등 기종에 관계없이 서버에 접속이 가능

6.4.2. 콘텐츠 저장용 스토리지

6.4.2.1. 많은 확장이 요구되는 콘텐츠 저장용 스토리지는 NAS전용 Mid-range급 스토리지로 구성하여 DB 저장용 스토리지에 비해 추가 확장 비용이 저렴

6.4.2.2. 콘텐츠 저장용 스토리지는 대용량이기 때문에 Mirror구성이 아닌 Raid 구성의 NAS스토리지 도입하여 여러 서버에서 공유할 수 있도록 하여 스토리지 용량을 효율적으로 관리

6.4.2.3. 데이터 및 서버가 증가하더라도 스토리지 용량을 증설하게 되면, 서버마다 디스크 용량을 확장해야 되는 불편과 비용이 감소

6.4.2.4. NT서버, Unix서버 등 기종에 관계없이 공유가 가능

6.5. Unix서버와 NT서버 구입

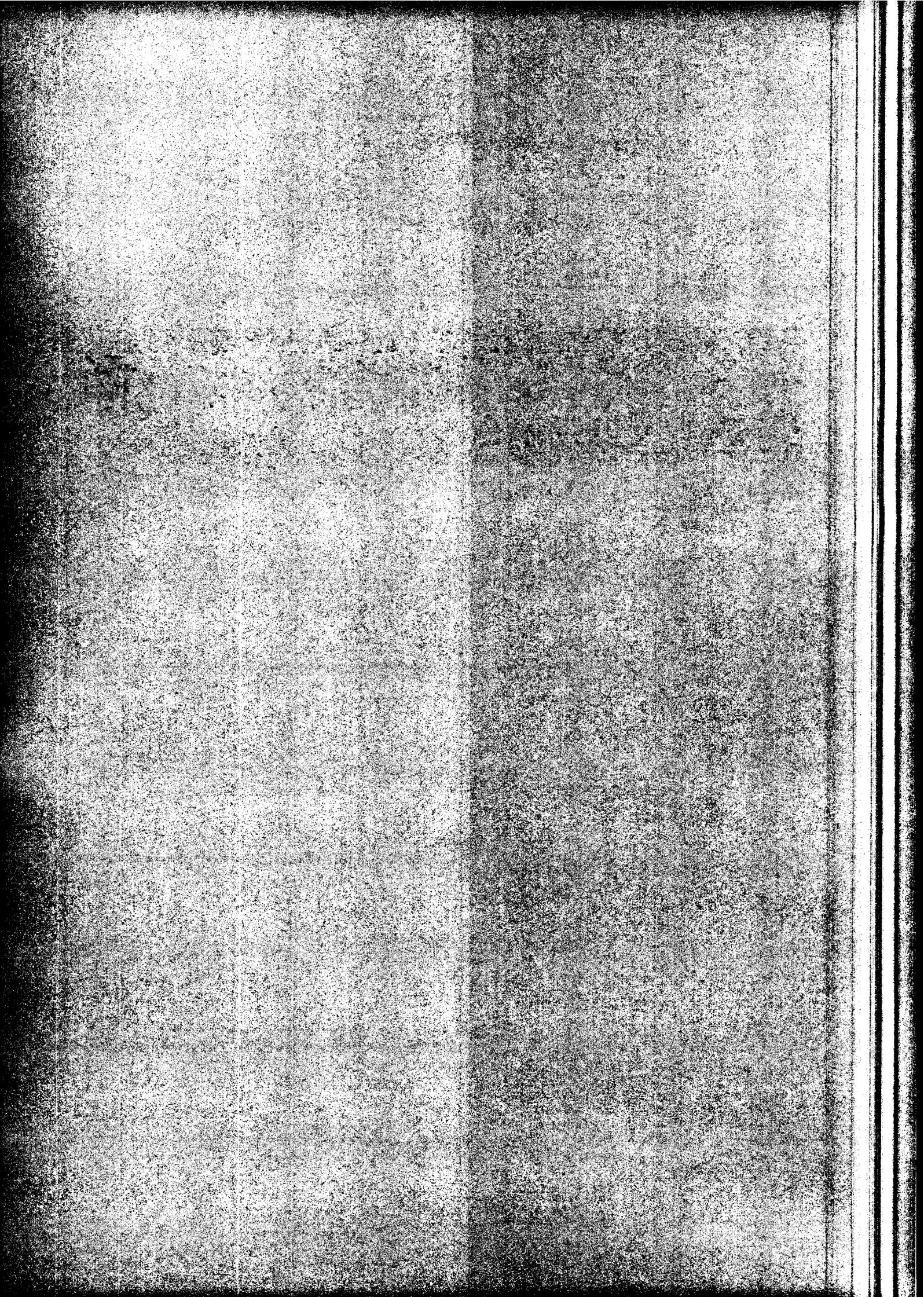
6.5.1. 안전성과 보안성이 필요한 DB서버와 이용자가 많은 웹서버, 이미지서버는 Unix서버로 도입

6.5.2. 전자도서관 S/W가 주로 NT서버 기반으로 개발되기 때문에 전자도서관 S/W 설치를 위해 NT서버를 도입

【서명부 3】

2003. 5

김희전
uri004@snu.ac.kr



< 사례발표 3 >

서울대학교전자도서관
XML Schema 채택 배경과 구현 사례

2003. 5.

김희진
(uri0804@snu.ac.kr)

차 례

1. 서울대학교 전자도서관 시스템의 XML Schema 채택 배경.....	1
1.1. XML Schema란?	1
1.2. XML Schema 채택 배경	1
2. 서울대학교 전자도서관 시스템의 개발 및 구조.....	3
2.1. XML Schema 개발 요구	3
2.2. XML Schema 구현	4
2.3. XML S초دم 운영	15
3. 문제점 및 개선방향	21

1. 서울대학교 전자도서관 시스템의 XML Schema 채택 배경

1.1. XML Schema란?

XML은 인터넷상에서 문서나 데이터를 교환 또는 배포할 때 표준이 될 수 있는 마크업 언어이다. 표준이 될 수 있다는 것은 어떤 플랫폼에서나 읽을 수 있는 포맷을 지원한다는 것으로 특정회사의 제품과 관련된 특정 환경에 얽매이지 않아도 된다는 것이다. 즉, XML로 짜여진 문서의 가장 큰 특징은 소수의 어플리케이션과 사용자들에게 국한되지 않는 확장성과, 보다 효율적이고 정확한 정보 전달을 위한 구조라는 것이다. 그러나, 이렇게 효율적이고 정확한 정보전달을 위해서는 기본적인 전제조건이 엘리먼트나 콘텐츠의 통일성이다. XML에서는 이러한 통일성을 위해 validation(검증)의 절차를 거친다. 이 검증의 절차를 거칠때 필요한 것이 바로 XML Schema인 것이다. 즉, XML Schema는 XML 문서구조를 정의하는 것으로서 데이터 입력과 XML 문서가 그 정의에 부합되게 만들어졌는지를 검증하는 도구가 된다. 이러한 도구에는 XML DTD와 XML Schema가 있으나 서울대학교에서는 XML Schema를 채택하여 개발하였다.

1.2. XML Schema 채택 배경

그렇다면, 왜 XML Schema를 기반으로 한 전자도서관 구축이 필요했던 것인가? 도서관에서는 정보자원관리를 위한 방법으로 MARC이 일반화되어 있다. MARC은 그 뿌리가 전통적인 서적형태의 정보자원에 대한 상세한 기술이므로 전자도서관에서 제공하고자 하는 디지털 자원들 - AOD, VOD, 각종 IMAGE 등 - 에 대한 새로운 형태의 정보유형을 기술하는 데에는 한계가 드러나고 있다. 또한, MARC의 데이터 모델은 자원기술의 대상을 단일 객체로 파악하는 단층 데이터 모델이므로 여러 객체간의 연결관계를 설정할 수 있는 다층 데이터 모델을 이용한 기술이 필요한 경우에는 적절하지 못하므로 이러한 한계점을 극복하고자 메타데이터(데이터에 대한 구조화된 데이터)를 연구하게 되었다.

그리고, 이렇게 연구된 메타데이터 셋은 서울대학교에 국한된 것이 아닌 국내 더 나아가 국제적인 메타데이터 셋으로의 발전을 바라며, 데이터 전송의 표준으로 인정받은 XML을 활용하여, 메타데이터 교환시 데이터의 무결성 보장을 위해 XML Schema를 채택하게 된 것이다. XML DTD가 아닌 Schema를 채택한 이유는 문서중심(Document Centric)인 XML DTD와는 달리, XML Schema는 데이터 중심(Data Centric)이며, XML DTD에서는 지원되기 힘든 Namespace 지원과 다양한 데이터 타입 지원 및 데이터의 상속이 가능하기 때문에 확장성 및 호환성이 뛰어나기 때문이다. 또한, XML Schema는 XML과 같은 문법 형식을 취하므로 응용프로그램은 파서를 통해 스키마를 다른 XML문

서와 유사한 방법으로 처리할 수 있으며, 반복 횟수 및 범위에 대한 세밀한 제약이 가능하고 별도의 프로그래밍적 코드가 필요 없어 응용면에서 뛰어난 기능을 가지고 있다. 아래의 <표 1>은 DTD와 XML Schema를 비교한 것이다.

<표 1> DTD와 Schema의 차이점

항 목	DTD	Schema
문법	XML과는 별도의 문법(EBNF)을 알아야 하므로 작성이 어렵다	XML 문법을 기반으로 하므로 XML 톨을 그대로 사용 가능
DOM (Document Object Model) ¹⁾ 지원	불가능	Schema도 XML문서이므로 DOM을 통해 조작 가능
컨텐츠 모델	간단한 순차 리스트 및 선택리스트만 제공	엘리먼트가 나타나는 횟수까지 정확하게 지정 가능
데이터 타입 정의	문자열, nametoken, ID와 일부 데이터 타입만 지원 : 응용이 어려움	문자열, 숫자, 날짜/시간 및 구조를 포함하는 널리 쓰이는 데이터 타입 지원 / 사용자 정의 가능
Namespace	Global만 사용	Global 및 local name 모두 사용
상속	불가능	가능
확장성	제한적 : 새로운 요소 추가시 새로 작성되어야 한다.	제한 없음
다중어휘	XML 문서에 하나의 DTD	XML 문서에 참조 Schema는 Namespace 기반으로 무한히 가능
동적연결	DTD는 읽기 전용(변경불가)	Runtime시에 동적으로 선택과 변경가능

1) "Document Object Model"이란 이름에서 알 수 있듯이 객체 모델로서 DOM은 valid HTML과 well-formed XML 문서를 객체를 사용하여 트리 구조로 모델화하고 만들어진 모델을 사용해서 실제 문서를 다루게 된다. DOM은 언어에 독립적인 객체 모델로 문서를 모델링하며, 모델링된 도커먼트 개체를 다루는 공통된 인터페이스를 제공함으로써 프로그래머들을 실제 문서의 문법으로부터 자유롭게 한다. (참고 : Opening XML, 이호경, 송순원 공저, 구민사, 2002)

2. 서울대학교 전자도서관 시스템의 개발 및 구조

2.1. XML Schema 개발 요구

서울대학교 전자도서관 시스템은 단위 시스템²⁾의 메타데이터를 XML Schema 기반으로 구축하고 서비스하는 시스템으로, 먼저 각 단위시스템별로 필요한 메타데이터를 연구하여 그 구성요소를 추출하고 이를 XML Schema로 정의한다. 정의된 XML Schema를 시스템에 적용시켜 XML 메타데이터를 입력하고 유효성(validation) 검사를 하며, 이렇게 입력된 데이터는 검색자 매핑정보에 의해 시스템 내부적으로 색인작업이 되어 사용자 서비스가 가능하게 된다. 이러한 전자도서관 시스템을 개발하는데 있어서 서울대학교에서 요구한 기능들은 아래의 <표 2>와 같다.

<표 2 > XML Schema 개발 요구안

기능명	요구사항
XML Schema 편집/ 저장	상용 프로그램을 사용하여 XML Schema 파일을 수정하거나 신규 생성함
XML Schema 등록	외부 스키마 파일이나 새로 생성된 스키마 파일을 등록하여 메타데이터를 구축할 수 있게 함
Encoding 스킴 등록	인코딩 스킴을 새로 등록하거나, 수정이 가능해야함
검색자 추출정보 매핑	XML Schema 요소의 간략리스트 항목 결정, 키워드 검색자/전방일치/제한 검색자 매핑 작업
Namespace 관리	SNU/SNUTERMS Namespace의 XML 파일 작성
	환경 관리 시스템에서 네임스페이스 관리를 위한 관리 모듈 필요
반입	XML Schema로 작성된 외부데이터를 DC 15개 항목에 맞춰 반입
반출	서울대학교 메타데이터를 DC 15개 항목에 맞춰 타기관으로 반출

위의 개발요구 내용을 보면 XML Schema 관리, encoding 스킴, 검색자 매핑, 네임스페이스 관리, 데이터의 반입·반출의 5가지로 나누어서 생각해볼 수 있다.

첫째, XML Schema 관리는 스키마를 새로 생성하거나 수정·편집하여 시스템상에 등록하기 위한 것으로 XML SPY라는 s/w를 이용한다. 처음 개발요구에서 제안했던 XML Schema 편집기의 개발은 구현의 어려움과 기존 Software 사용으로 충분히 대체 가능하다는 판단이 되어 개발계획에서 제외되었다.

2) 단행본, 연속간행물 등과 같은 디지털 콘텐츠 유형에 따른 분류로서 서울대학교 전자도서관의 단위시스템에는 음악자료, 미술작품, 학술행사, 대학사료, 기사, 학술지, 의학자료가 있다.

둘째, 인코딩 스킴은 요소의 내용 또는 값을 제어하는 것으로서 통제어휘집, 공식적인 표시법 또는 문장 분석 규칙 등을 포함한다. 그러므로, 인코딩 스킴을 사용하여 표현된 값은 통제 어휘집으로부터 선정된 하나의 신호(예를들어, 분류체계나 주제명표목표에서 가지고 온 용어), 또는 공식적인 표시법에 의거하여 형식화된 문자열(예를들어, 날짜를 표시하는 표준 규칙을 적용한 '2000-01-01')등이 될 것이다. 서울대학교 전자도서관 환경관리 시스템에서는 이러한 인코딩 스킴을 신규등록 및 수정이 가능하도록 구현하였다.

셋째, 검색자 매핑은 검색 간략리스트에 디스플레이 될 요소와 키워드/전방일치/제한 검색 항목으로 사용될 요소를 선정하고 관리할 수 있는 기능을 구현한 것으로 역시 환경관리 시스템에서 조정할 수 있다.

넷째, 네임스페이스 관리는 네임스페이스 자체가 XML 파일이므로 XML Schema와 마찬가지로 신규생성 및 수정은 상용 s/w를 통하여 편집하여 관리한다. 네임스페이스란, 데이터를 생성하고 또 그것을 다른 사람들과 공유하기 위하여, XML 파일 안에서 사용되는 특정한 데이터 요소 또는 속성들의 이름에 대해 정의 및 설명을 해주고 있는 특정한 웹 사이트를 가리킨다.

마지막인 데이터의 반입·반출은 XML Schema로 작성된 데이터의 상호호환을 위해 메타데이터의 기본이 되는 DC 15개 항목에 맞춰 반입·반출이 가능하도록 구현할 것을 요구한 것이다. 그러나, 이것의 전제는 DC 15개 항목의 규칙을 적용시킨 스키마를 기반으로 한 메타데이터만을 대상으로 하는 것이기 때문에 그 활용성은 아직 미지수이다.

2.2. XML Schema 구현

2.2.1. XML Schema 구현 원칙

서울대학교 전자도서관의 XML Schema는 UKOLN에서 제안한 다음과 같은 자료를 참고하여 작성하였으며, 이는 Dublin Core 기반의 메타데이터 집합을 구현한 것으로 서울대학교 XML Schema의 근간이라고 할 수 있다.

- Guidelines for implementing Dublin Core in XML³⁾
- Example Dublin Core XML Schema⁴⁾

위의 자료들을 참고하여 만들어진 서울대학교 전자도서관 XML Schema는 단순히 요소를 확정하고 정의하는 것이 아니라 몇가지 원칙에 의거하여 만들어졌다. 다음의 3가지 원칙이 서울대학교 전자도서관 XML Schema의 요소 집합(Element Set)과 한정어

3) <http://www.ukoln.ac.uk/metadata/dcml/dc-XML-guidelines/>

4) <http://www.ukoln.ac.uk/metadata/dcml/dcXML/examples.html>

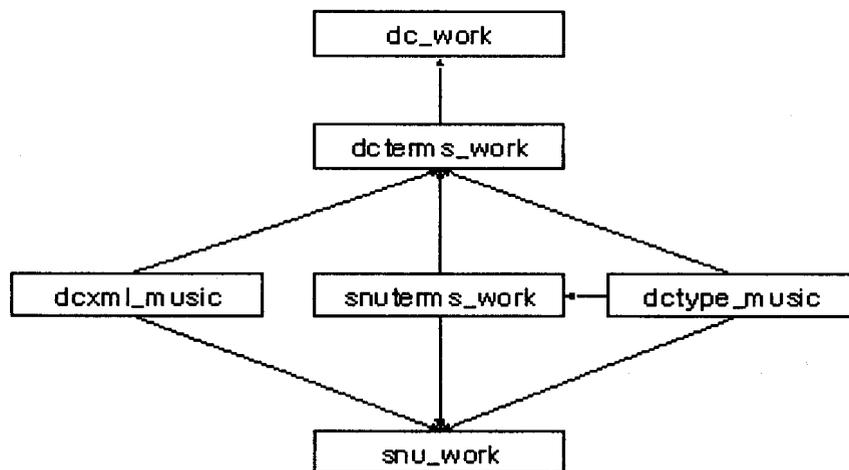
(Qualifiers)의 사용 원칙이다.

첫째, 하위 요소가 없는 요소는 실질적인 값을 취할 수 있지만 하위 요소가 있는 요소는 실질적인 값을 취하지 않는다. 만약, 검색시에 하위요소를 가지는 요소가 검색대상이 될 경우에는 덤다운(Dumb Down) 원칙에 의하여 하위요소에 있는 실질적인 값이 상위 요소의 값으로 대체가 되어 상위요소로의 검색도 가능하게 한다.

둘째, DC 요소 집합의 하위 요소로는 DC 한정어와 SNU 한정어가 올 수 있다.

셋째, SNU 요소 집합의 하위 요소로는 SNU 한정어가 온다.

즉, SNU 한정어는 DC와 SNU의 하위 요소로는 올 수 있으나 DC 한정어는 DC의 하위 요소로는 갈 수 있어도 SNU의 하위 요소로는 갈 수가 없다. 이는 DC 한정어가 덤다운 원칙에 적용되기 위해서는 상위요소로 반드시 DC가 와야하기 때문이다. 이러한 관계를 간단히 그림으로 나타내보면 다음과 같다.



이 그림은 음악자료 단위시스템의 work 메타데이터에 대한 XML 파일간의 연관관계를 나타낸 것으로 화살표는 import 관계를 나타내며, dcxml은 work 메타데이터의 입력을 위한 scheme 정의 파일이며 dctype은 서울대학교 XML Schema의 공통타입을 정의한 파일이다.

위의 원칙을 기본으로 하여 만들어진 서울대학교 XML Schema의 내용을 형태별로 살펴보면 다음과 같다.

- 자식 요소가 있는 경우

```

<xs:complexType name="titleType">
  <xs:sequence>
    <xs:element ref="snuterm:smainTitle"/>
    <xs:element ref="dcterms:alternative" minOccurs="0"
maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>

```

- 자식 요소가 없이 스킴만을 가지는 경우

```

<xs:complexType name="subjectType">
  <xs:simpleContent>
    <xs:extension base="dctype:elementType">
      <xs:attributeGroup ref="dcXML:subjectScheme"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>

```

- 자식 요소가 없이 스킴과 Dumb Down에 의한 Refinement Qualifier을 갖는 경우

```

<xs:complexType name="workRelationRefinementType">
  <xs:simpleContent>
    <xs:extension base="dctype:elementType">
      <xs:attribute name="relationType" type="relationTypeType"/>
      <xs:attributeGroup ref="dcXML:identifierScheme"/>
      <xs:attributeGroup ref="dcXML:relationRefines"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>

```

- 요소의 값으로 나열 값을 갖는 경우

```

<xs:simpleType name="relationTypeType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="derived"/>
    <xs:enumeration value="isMemberOf"/>
  </xs:restriction>
</xs:simpleType>

```

- 두 요소 중 하나만 자식 요소로 오는 경우

```
<xs:complexType name="relationType">
  <xs:choice>
    <xs:element ref="snuterm:isArticleOfConference"/>
    <xs:element ref="snuterm:isArticleOfSession"/>
  </xs:choice>
</xs:complexType>
```

- 스킴 정의 (여러 가지 스킴이 있을 경우)

```
<xs:simpleType name="subjectSchemeType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="LCSH"/>
    <xs:enumeration value="MESH"/>
    <xs:enumeration value="DDC"/>
    <xs:enumeration value="LCC"/>
    <xs:enumeration value="UDC"/>
  </xs:restriction>
</xs:simpleType>
<xs:attributeGroup name="subjectScheme">
  <xs:attribute name="scheme" type="subjectSchemeType"/>
</xs:attributeGroup>
```

- 스킴 정의 (한가지 스킴만 있을 경우)

```
<xs:attributeGroup name="publisherScheme">
  <xs:attribute name="scheme" type="xs:string" fixed="SNUCR"/>
</xs:attributeGroup>
```

- Dumb Down에 의한 Refinement Qualifier

```
<xs:attributeGroup name="titleRefines">
  <xs:attribute name="refines" type="xs:string"
    fixed="http://purl.org/dc/elements/1.1/title"/>
</xs:attributeGroup>
```

2.2.2. 단위시스템 및 스키마 모델

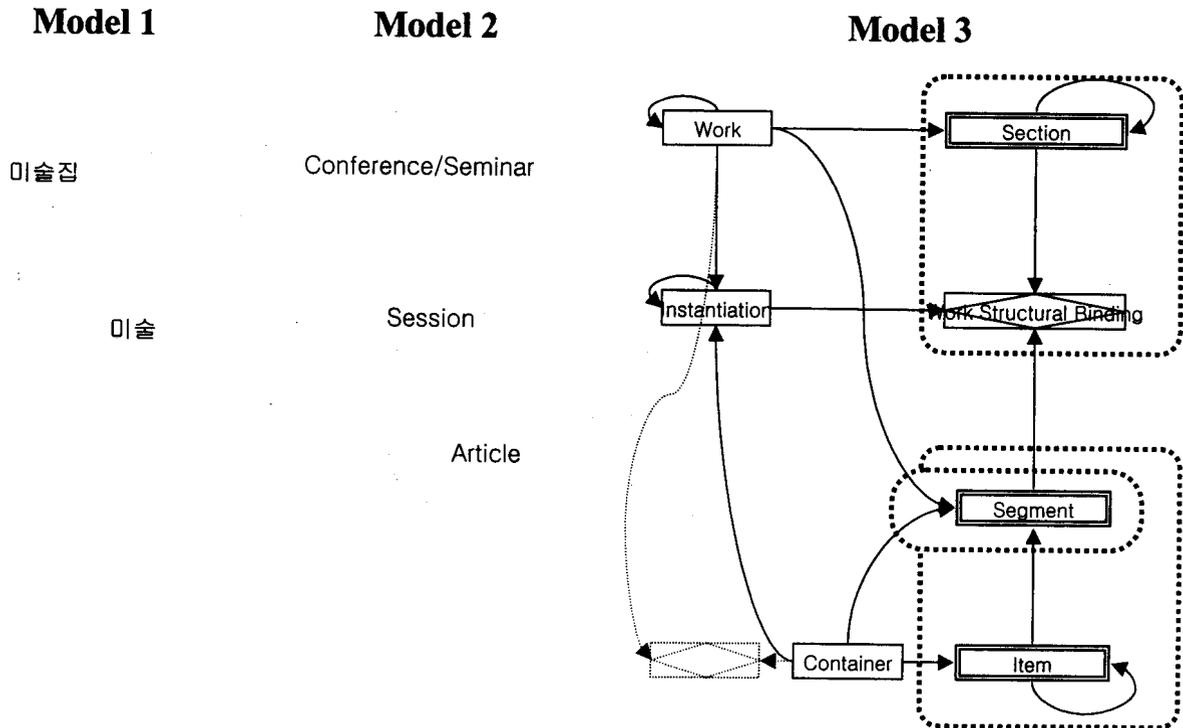
앞에서도 언급하였지만, XML Schema가 만들어지기 전에 선행되어야 하는 것은 단위시스템의 선정과 메타데이터 요소 확정이다. 서울대학교 전자도서관의 단위시스템은 그 데이터의 성격상 MARC만으로는 표현할 수 없는 다면적인 특성을 지닌 지식자원 - Audio, Video, slide, Image 등의 형태로 그 검색접근점이 다양하며 유관자료끼리의 링크가 복잡하게 얽혀있는 자원들 - 을 중심으로 선정되었다. 이렇게 선정된 단위시스템들은 그 특성에 따라 3가지 모델로 나누어 구분지을 수 있는데, 이 3가지의 구분은 메타데이터의 구분과 동일한 것으로, 각 단위시스템에 해당되는 자원의 성격에 따라 메타데이터가 결정되고 XML Schema가 그 메타데이터를 정의하기 때문이다.

<표 3> 서울대학교 전자도서관 단위 시스템

구 분	단 위 시 스템
Schema model 1	기사
	학술지
	대학사료
	미술작품
	의학자료
Schema model 2	학술행사
Schema model 3	음악자료

<그림1>과 같이 Schema model 1에 해당되는 단위시스템은 단순 데이터 모델 유형에 속하는 것으로 미술작품 단위시스템과 같이 미술작품집이라는 기술대상과 미술작품집 속에 포함되어 있는 개개의 미술작품에 대한 서지사항을 연계하여 구축되는 유형이며, Schema model 2에 해당하는 학술행사 단위시스템은 다층 데이터 모델 유형으로 학술/세미나(conference)를 기준으로 세션(session), 발표자료(article)에 대한 기술을 통해 상하간 연계되도록 구축하는 유형이다. Schema model 3에 해당하는 음악자료는 복합적 데이터 모델 유형으로 상하간의 관계뿐만이 아니라 수평적 메타데이터 셋끼리의 유기적인 구조를 갖기 때문에 model1과 model2보다 더 복잡하다. 각 단위시스템의 DB 테이블의 기본유형도 이러한 스키마 모델 유형에 따라 상하관계만 있는 model1과 model2와 달리 model3의 DB 테이블을 나타내는 <그림4>는 그 복잡도가 더하다.

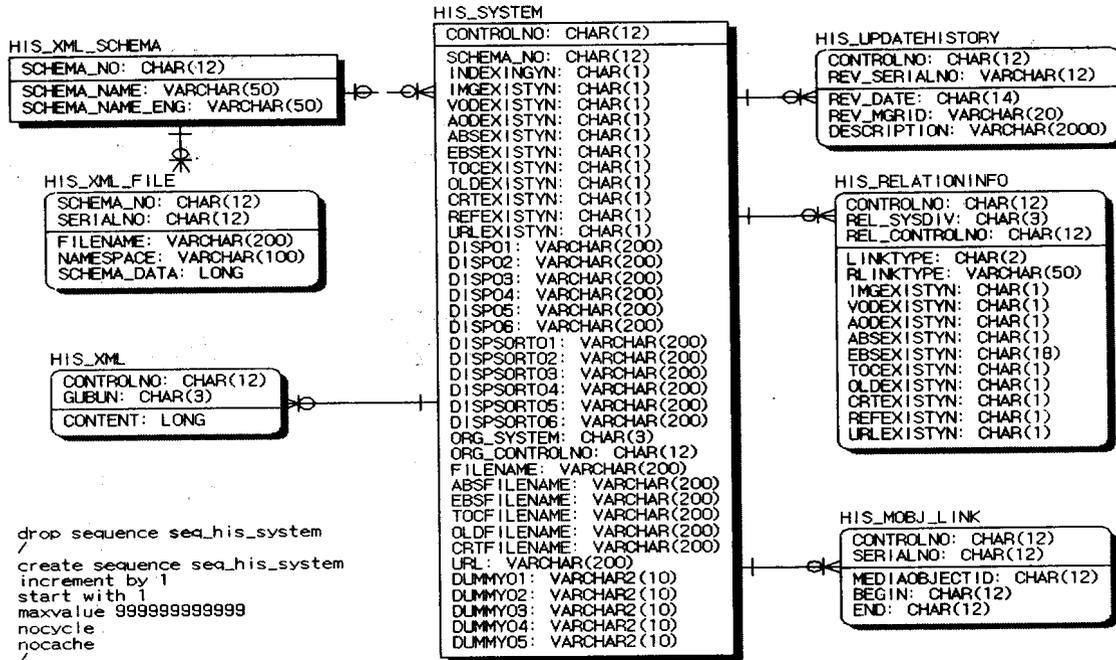
< 그림 1 > XML Schema 모델 >



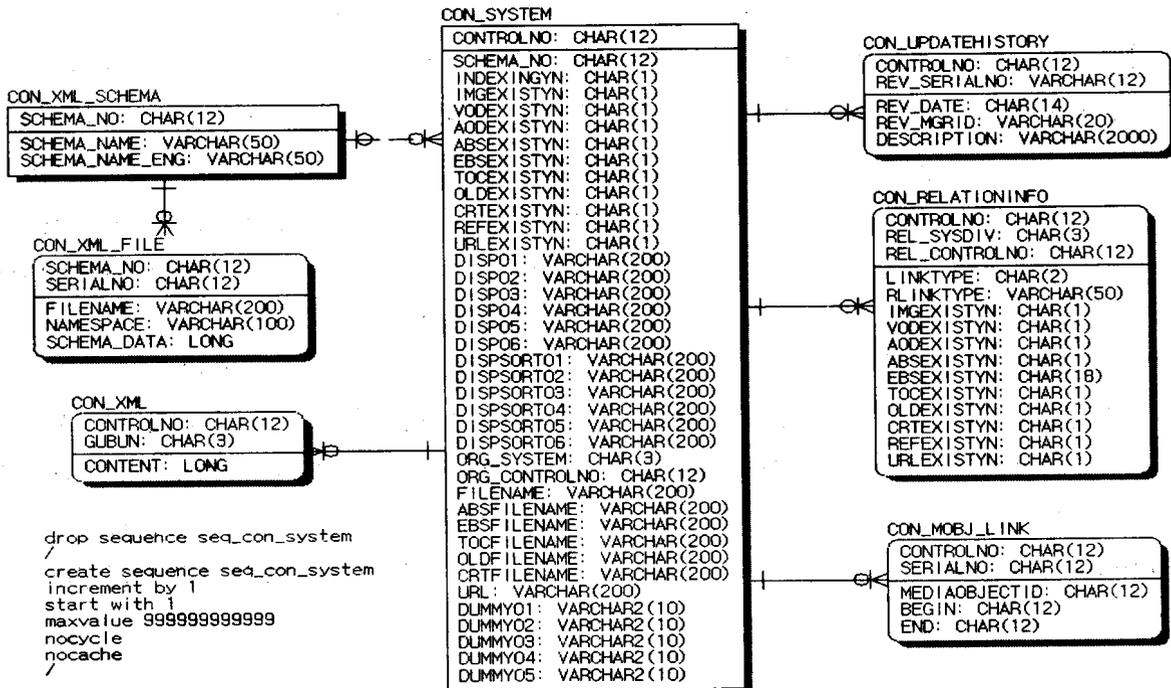
<그림2>~<그림4>에서 공통적으로 보이는 ~_SYSTEM 테이블은 그 단위시스템에 적용되는 스키마번호, 서비스등록여부, 각 매체의 존재유무, 간략리스트에서 디스플레이 될 간략항목, 소트용 간략항목, 연동시스템, 대표화면파일 등의 값들을 가지게 된다. 예를 들어, 클라이언트 프로그램 중 관리자용으로 개발된 환경관리 시스템에서 간략화면 디스플레이 항목을 선택 저장할 시 DISP01 ~ DISP06의 필드에 그 값이 들어가게 된다. 일반적으로 각각의 테이블은 제어번호(CONTROLNO)와 스키마번호 (SCHEMANO)로 연결될 수 있는데, ~_UPDATEHISTORY는 업데이트 정보를 담는 테이블이고, ~_RELATIONINFO는 관련시스템에 대한 정보를 담는 테이블이며, MOBJ_LINK 테이블은 해당 미디어오브젝트의 정보를 저장해서 검색시 속도를 보다 향상시켜주는 역할을 한다.

음악자료의 메타데이터를 담아야 하는 model3의 구조는 음악자료의 구조적인 특성으로 보다 많은 테이블로 구성된다. WORKSTRUCT_MAIN은 작품을, WORKSTRUCT_SECTION은 작품의 하부구조 정보를, WORKSTRUCT_BIND는 작품의 하부구조인 Section에 대한 연주정보를 담는 테이블이다. 그리고, CONTAINER_ITEMLIST는 음반/매체에 대한 정보를, CONTAINER_DIVLIST는 음반/매체에 수록되어 있는 연주구조정보를 담는 테이블이다.

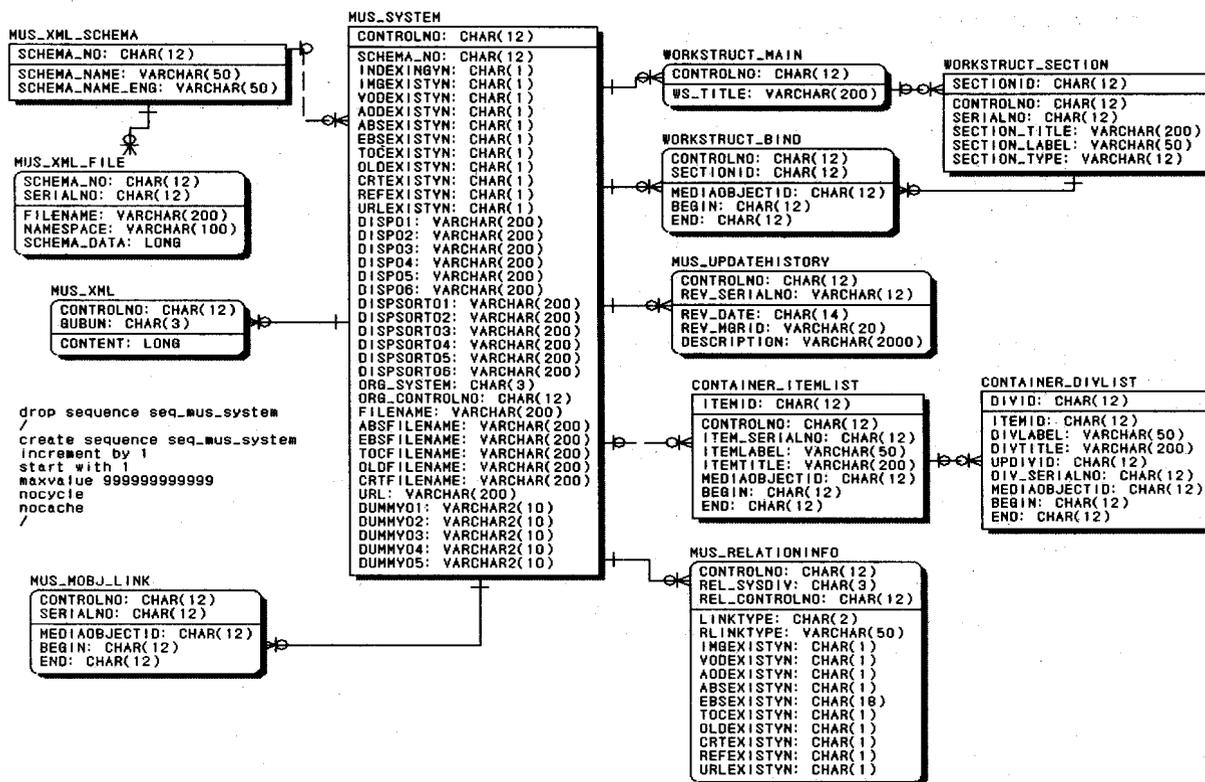
<그림2> 모델 1 테이블 스키마 - 대학사료(HIS)



<그림 3> 모델 2 테이블 스키마 - 학술행사(CON)



<그림 4> 모델 3 테이블 스키마 - 음악(MUS)



2.2.3. Encoding Scheme

Encoding Scheme이란, 의미를 명확히 하거나 정보자원 검색을 개선하기 위해서 요소나 요소 상세화의 내용 또는 그 값을 제어하는 것으로, 인코딩 스킴을 이용하여 표시된 값은 공식적인 표기(예를 들면, 날짜의 표준 표현으로서 '2003-01-01')에 따라서 정형화된 문자열이나 통제된 어휘로부터 선별된 것이다. 이러한 인코딩 스킴은 기계나 인간에 의해서 명확하게 해석되도록 설계되어야 하고, 메타데이터를 작성하는 사람들 뿐만 아니라 정보를 찾고자 하는 사람들에게도 유용해야 한다. 서울대학교 전자도서관 시스템에서 이용하는 인코딩 스킴은 <그림 5>에서 보는 바와 같이 그 종류가 다양하다. 여기에 나열된 스킴은 각각의 메타데이터의 입력시에 입력지침이 되는 것이며, 이 스킴에 해당되지 않는 데이터가 입력될 시에는 올바른 XML 파일이 만들어지지 않게 되어 유효성(Validation) 검사에서 걸러지게 된다. <그림 6>은 <그림 5>에 정리되어 있는 스킴들이 음악 Work메타데이터의 어떤 요소에 해당되는 것인지 보여주고 있다. 이 인코딩 스킴은 전자도서관 환경설정 시스템에서 수정 및 추가가 가능하도록 되어 있다.

<표 4> 서울대학교 전자도서관 Scheme의 종류

1. 서울대에서 사용하고자 하는 scheme의 종류

AACR2			
AAT	http://shiva.pub.getty.edu/aat/browse/titles.html	The Art & Architecture Thesaurus Browser	The Agnts Facet section of the Gettys Art and Architecture
ISO 639-2	http://lcweb.loc.gov/standards/iso639-2/langhome.html		언어를 표시하는 스킴
ISO 3166	http://www.din.de/gremien/nas/nabd/iso3166ma/codlstp1/index.html		
MARC Code List			
LCSH	http://www.mailbase.ac.uk/lists/dc-subdesc/files/wd-subdesc-qual.htm	Library of Congress Subject Heading	
MESH	http://www.mailbase.ac.uk/lists/dc-subdesc/files/wd-subdesc-qual.htm	Medical Subject Headings	
DDC	http://www.mailbase.ac.uk/lists/dc-subdesc/files/wd-subdesc-qual.htm	Dewey Decimal Classification	
LCC	http://www.mailbase.ac.uk/lists/dc-subdesc/files/wd-subdesc-qual.htm	Library of Congress Classification	
UCD	http://www.mailbase.ac.uk/lists/dc-subdesc/files/wd-subdesc-qual.htm	Universal Decimal Classification	
DCMI-Period	http://dublincore.org/documents/2000/07/28/dcmi-period/	시간에 대한 스킴	
W3CDTF	http://www.w3.org/TR/NOTE-datetime		
RFC 1766	http://www.ietf.org/rfc/rfc1766.txt		http://asg.web.cmu.edu/rfc/rfc1766.html
URI			
TGN	http://www.mailbase.ac.uk/lists/dc-coverage/files/wd-coverage-qual.htm		

2. 서울대 자체 구축 Scheme 중 음악자료 Scheme

SNUCR		서울대 목록규칙	자체 정리
SMF		음악 형식	서울대학교 음디와 협의하여 자체 구축
SMI		악기명	서울대학교 음디와 협의하여 자체 구축
SMK		주요 키값	서울대학교 음디와 협의하여 자체 구축
SMST		음악 석션 종류(type)	서울대학교 음디와 협의하여 자체 구축

<표 5> 음악자료 Work 메타데이터에서 사용하는 scheme의 예

요소	제1 하위요소	Scheme	Value	Definition
dc:title				작품의 제목
	snuterm:s:mainTitle(D)	AACR2, SNUCR		작품의 통일서명(Uniform Title), AACR2 목록규칙 표준을 충족시켜야만 한다
	dcterms:alternative(D)	AACR2, SNUCR		작품의 대등서명
dc:creator				editor, conductor 등 Work를 생성한 사람
	vCard:FN(D)	AACR2, SNUCR		editor, conductor 등 Work를 생성한 사람 이름
	vCard:N			
	vCard:BDAY			
	vCard:ROLE	MARC Code List, AAT		작품의 작곡/작사가의 작품에서의 역할
	vCard:EMAIL			
	vCard:ORG			
dc:subject			LC SH, MESH, DDC, LCC, UCD	작품의 주제
dc:description				작품에 대한 내용주기
dc:date				자원의 존재기간 동안 어떠한 사건이 발생한 날짜
	snuterm:s:created(D)	DCMI-Period, W3CDTF		작품의 창작일이나 창작일로 예측되는 일자
	snuterm:s:dateOfFirstPerformance(D)	DCMI-Period, W3CDTF		작품이 처음으로 연주된 일시
	snuterm:s:dateOfFirstPublication(D)	DCMI-Period, W3CDTF		작품이 처음으로 출판된 일시
dc:type			SMF	작품의 형식
dc:identifier				작품에 부여된 고유한 번호로 시스템상의 고유한 Record ID
dc:language			ISO 639-2, RFC 1766	작품의 창작에 사용된 언어
dc:relation				Work 와 관련된 자원들과 관계를 기술한다.
	dcterms:isPartOf	ID		작품유형이 Single 일때, Collection으로 기술된 작품의 메타데이터 ID기입
	dcterms:hasPart	ID		작품유형이 Collection일때, 이에 속한 Single 작품의 메타데이터 ID 기입
dc:coverage	snuterm:s:instantiation			작품의 내용이 다루고 있는 범위 (시간적, 공간적)

2.2.4. XML Namespace

XML Namespace란, 웹 상에서 데이터를 생성하고 또 그것을 다른 사람들과 공유하기 위하여, XML 파일 안에서 사용되는 특정한 요소 또는 속성들의 이름에 대해 정의 및 설명을 해주는 특정한 웹사이트를 가리키는 것이다. 이러한 네임스페이스는 여러 문서상의 용어 충돌을 막아 준다. 서울대학교 전자도서관의 기본이 되는 Dublin Core 요소집합과 한정어의 Namespace는 다음의 주소로써 식별된다.

- dc : <http://dublincore.org/documents/1999/07/02/dces/>
- dcq : <http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/>

DC로 기술하지 못하는 정보에 대해서는 서울대학교 자체적으로 연구하여 추가한 SNU, SNUTERMS를 이용하여 정보를 기술하도록 하였고, DC와 SNU에 속하는 요소의 의미를 세분화하기 위해서 다음과 같은 vCard⁵⁾, VRA⁶⁾를 이용하였다.

<표 6> vCard, VRA의 요소

구분	Element명
vCard	<input type="checkbox"/> FN - Formatted Name <input type="checkbox"/> BDAY - Birthday <input type="checkbox"/> TITLE - Position Title <input type="checkbox"/> ROLE - Position Role <input type="checkbox"/> TEL - Telephone <input type="checkbox"/> EMAIL - Email Address <input type="checkbox"/> N - Name <input type="checkbox"/> ORG - Organization
VRA	<input type="checkbox"/> Current Site <input type="checkbox"/> Former Site <input type="checkbox"/> Creation Site <input type="checkbox"/> Current Repository <input type="checkbox"/> Former Repository

- 5) vCard는 전자명함의 개념을 통해 대두된 기술요소로서, 네트워크를 통해 사업용 또는 개인용 명함을 교환하기 위한 규격이다. 서울대학교 전자도서관에서는 저자사항 한정요소를 위해 vCard를 차용하였다.
- 6) VRA(Visual Resources Association)는 미술작품이나 건축물 자원의 장소정보를 기술하기 위해 사용되는 요소로 서울대학교 전자도서관에서는 음악, 학술행사, 대학사료, 미술, 학술지 등의 메타데이터셋에서 장소와 관련된 한정 요소를 위해 차용하였다.

이처럼, 서울대학교 전자도서관에서 사용하는 메타데이터 셋인 DC, SNU, vCard, VRA 등을 설명해주는 Namespace의 URL은 다음 <표 7>과 같다.

<표 7> 서울대학교 Namespace의 URL

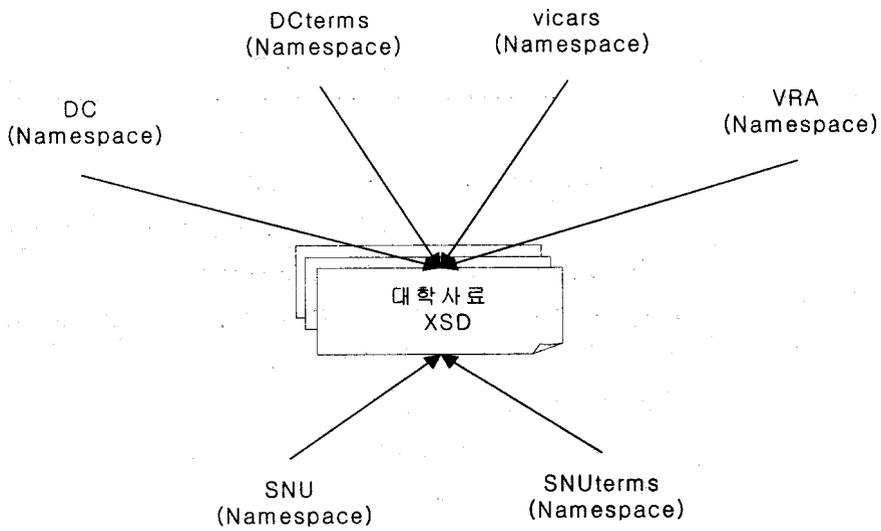
구분	설명
Dublin Core	XML Namespace Prefix : dc XML Namespace URI : http://purl.org/dc/elements/1.1
Dublin Core Qualifier	XML Namespace Prefix : dcterms XML Namespace URI: http://purl.org/dc/terms/
vCard	XML Namespace Prefix : vCard XML Namespace URI : http://www.w3.org/2001/vcard-rdf/3.0#
VRA	XML Namespace Prefix : vra XML Namespace URI: http://www.vraweb.org/
SNU	XML Namespace Prefix : snu XML Namespace URI: http://www.snu.ac.kr/metadata/
SNU Qualifier	XML Namespace Prefix : snuterms XML Namespace URI: http://www.snu.ac.kr/metadata/terms/
DcXML (인코딩 스킴)	XML Namespace Prefix : dcXML XML Namespace URI: http://purl.org/dc/XML/
Dctype (공통 타입)	XML Namespace Prefix : dctype XML Namespace URI: http://purl.org/dc/type/

2.3. XML Schema 운영

앞에서도 언급한 바와 같이 전자도서관 시스템의 개발 순서는 첫째로 메타데이터를 연구·확정하는 것이며, 그에 따라 XML Schema 파일이 작성되는 것이다. 확정된 메타데이터는 해당 네임스페이스를 참조하여 다음 그림과 같이 XML Schema로 정의(XSD : XML Schema Definition) 한다.

<그림 5>는 메타데이터 요소가 결정된 후 그에 대한 네임스페이스를 참조하여 xml schema로 정의하는 것을 나타낸 것으로서 대학사료를 예로 든 것이고, <그림 6>은 .xsd 파일을 만들어낼 수 있는 프로그램인 XML SPY에서 학술행사의 메타데이터를 xml schema로 정의하여 구현한 것이다. <그림 7>은 실제 메타데이터 값이 들어간 xml 파일이다.

<그림 5> Namespace를 참조하여 작성되는 XSD



<그림 6> 대학사료 XSD

XML Schema for SNUOL old book and document based on UKOLN dc.xsd
 dcterms_univ.xsd <http://purl.org/dc/terms/>
 vcard_univ.xsd <http://www.w3.org/2001/vcard-rdf/3.0/>

titleType
creatorType
contributorType
descriptionType

XML Schema for SNUOL old book based on UKOLN dcterms.xsd, Written by Jinseok Chae
 docml_univ.xsd <http://purl.org/dc/xml/>
 dc_type_univ.xsd <http://purl.org/dc/type/>
 snterms_univ.xsd <http://www.snu.ac.kr/metadata/terms/>

alternativeType

XML Schema for SNUOL work metadata written by Jinseok Chae,
 dc_type_univ.xsd <http://purl.org/dc/type/>
 vra_univ.xsd <http://www.vraweb.org/>

XML Schema for SNUOL old book based on UKOLN dcterms.xsd, Written by
 dc_type_univ.xsd <http://purl.org/dc/terms/>
 titleRefinementType
 alternativeRefinementType
 classificationType
 /dc_type_univ.xsd <http://purl.org/dc/type/>

IN
N
PRAN

/dc_type_univ.xsd <http://purl.org/dc/type/>

currentSite
formerSite
creationSite
discoverySite
currentRepository
formerRepository

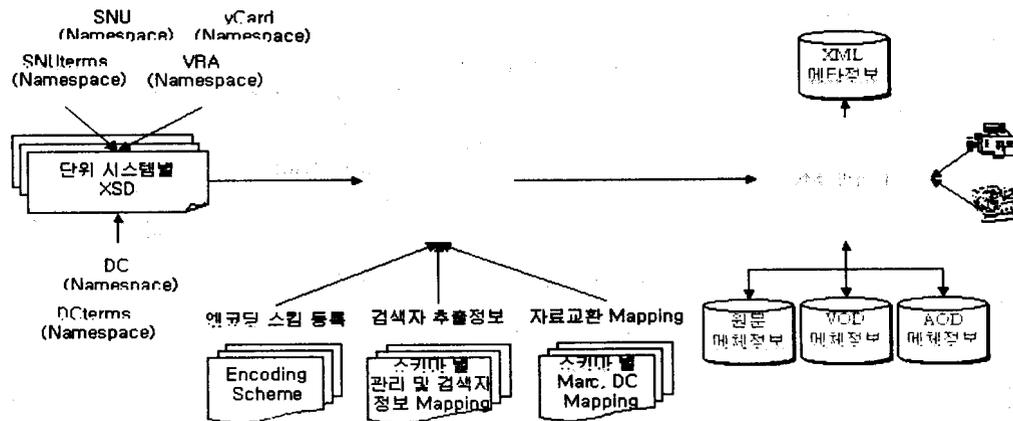
<그림 7> 대학사료 XML 메타데이터

▼ XML	
▲ metadata	
= xmlns:do	http://purl.org/dc/elements/1.1/
= xmlns:doterms	http://purl.org/dc/terms/
= xmlns:doxml	http://purl.org/dc/xml/
= xmlns:snu	http://www.snu.ac.kr/metadata/
= xmlns:snu:terms	http://www.snu.ac.kr/metadata/terms/
= xmlns:xsi	http://www.w3.org/2001/XMLSchema-instance
= xmlns:vCard	http://www.w3.org/2001/vcard-rdf/3.0#
= xmlns:vra	http://www.vraweb.org/
= xsi:noNamespaceSchemaLocation	univ.xsd
▲ do:title	<div style="margin-left: 20px;">▲ snu:terms:mainTitle</div> <div style="margin-left: 40px;">= doxml:refines http://purl.org/dc/elements/1.1/title Abc Text 교원신문 1966년</div> <div style="margin-left: 20px;">▲ do:terms:alternative</div> <div style="margin-left: 40px;">= doxml:refines http://purl.org/dc/elements/1.1/title Abc Text 조선민주주의인민공화국보통교육장 조경교육문화보건의료공무원직업 중앙위원회기관지</div>
▲ do:creator	<div style="margin-left: 20px;">() vCard:FN 교원신문사</div>
▲ do:subject	<div style="margin-left: 20px;">= doxml:refines http://www.snu.ac.kr/metadata/terms/classification Abc Text 951.8</div>
() do:publisher	<div style="margin-left: 20px;">교원 신문사(평양)</div>
▲ do:date	<div style="margin-left: 20px;">() snu:terms:startDate 1966-01-01</div>
() do:type	<div style="margin-left: 20px;">4</div>
▲ do:language	<div style="margin-left: 20px;">= doxml:scheme ISO639-2 Abc Text kor</div>
▲ do:relation	<div style="margin-left: 20px;">▲ do:terms:isPartOf</div> <div style="margin-left: 40px;">= doxml:refines http://purl.org/dc/elements/1.1/relation Abc Text 000000016006</div>
() snu:workType	<div style="margin-left: 20px;">() snu:terms:isCopy y</div>
▲ snu:place	<div style="margin-left: 20px;">() vra:currentRepository 2</div>

그러나, 전자도서관에서 필요한 데이터들을 구축하는 데에는 메타데이터셋을 정하고, XSD를 만드는 것으로 끝나는 것이 아니다. 위와 같은 과정은 실제로 구축할 데이터를 제대로 저장하고 관리할 수 있는 기반을 만들어 주는 것으로, 사서의 관리대상이 될 수는 없다. <그림 8>에서처럼 각 단위시스템에 해당하는 디지털 콘텐츠 메타데이터가 XSD에 맞게 입력되고 입력된 디지털 콘텐츠가 웹에서 최대한 용이하게 검색이 될 수 있도록 관리하기 위해서는 별도의 관리 시스템이 필요하다. 서울대학교 전자도서관에서는 각 단위시스템에 해당하는 데이터를 입력하기 위한 CMS(Content Management System)와 CMS를 통해서 입력된 데이터가 웹에서 용이하게 검색되기 위한 검색 Mapping을 관리할 수 있는 환경관리 시스템이 있다. 환경관리 시스템<그림 9>에서는 단위시스템을 관리·추가할 수 있으며, 웹상에서 디스플레이 되는 항목과 검색 키워드로 쓰일 항목을 선

정하고 관리하는 기능이 있다. CMS(그림 10)에서는 단위시스템과 해당 Schema를 선택하면 디지털 콘텐츠 메타데이터 값을 입력할 수 있는 화면이 나오고, 데이터를 입력하면 해당 XSD에 맞게 XML 파일로 저장되게 한다.

<그림 8> XML Schema 운영



<그림 9> 환경 관리 시스템

시스템(S) 저장소(B) 권한(P) IP(I)

시스템: 신규 저장 삭제 종료

표준시스템명	영문시스템명	시스템명	구분	원문관리	VOD관리	AOD관리	초록입력	목차입력
1 대학사료	History	HIS	GS	<input checked="" type="checkbox"/>				
2 회의	Conference-New	CNF	GS	<input checked="" type="checkbox"/>				
3 솔라즈 서지	Solars Bib	BIB	GS	<input checked="" type="checkbox"/>				
4 솔라즈 기사	Solars Article	ATL	GS	<input checked="" type="checkbox"/>				
5 음악	Music	MUS	GS	<input checked="" type="checkbox"/>				
6 학술지	CMS Journal	JNL	GS	<input checked="" type="checkbox"/>				
7 단행본		MON	WM	<input checked="" type="checkbox"/>				
8 미술	Art	ART	GS	<input checked="" type="checkbox"/>				
9 연극		SER	WM	<input checked="" type="checkbox"/>				
10 E-Book	E-Book	EBK	GS	<input checked="" type="checkbox"/>				
11 의학		MED	GS	<input checked="" type="checkbox"/>				
12 솔라즈	Solars	SOL	WM	<input checked="" type="checkbox"/>				
13 기사	Articles	ATI	WM	<input checked="" type="checkbox"/>				
14								
15								
16								
17								
18								
19								
20								

표준시스템명: 원문관리 초록입력 서명입력
영문시스템명: VOD관리 목차입력 참고문헌입력
시스템이니셜: AOD관리 해제입력 외부링크입력
시스템이니셜: 영문초록 E-BOOK

저장소:

2003-04-28 오전 8:56

<그림 10> 콘텐츠 관리 시스템

음악 > <input type="checkbox"/> 신규 <input checked="" type="checkbox"/> 상세검색 <input type="checkbox"/> 대표자료 <input type="checkbox"/> 환경 <input type="checkbox"/> 통계 <input type="checkbox"/> 종료

[간략검색] 검색항목 <input type="checkbox"/> 전체 검색어 <input type="text" value="" /> <input type="button" value="검색" />

<input type="checkbox"/> 신규 <input type="checkbox"/> 저장 <input type="checkbox"/> 삭제 <input type="checkbox"/> XML보기 <input type="checkbox"/> 메타구조 <input type="checkbox"/> 예 달기

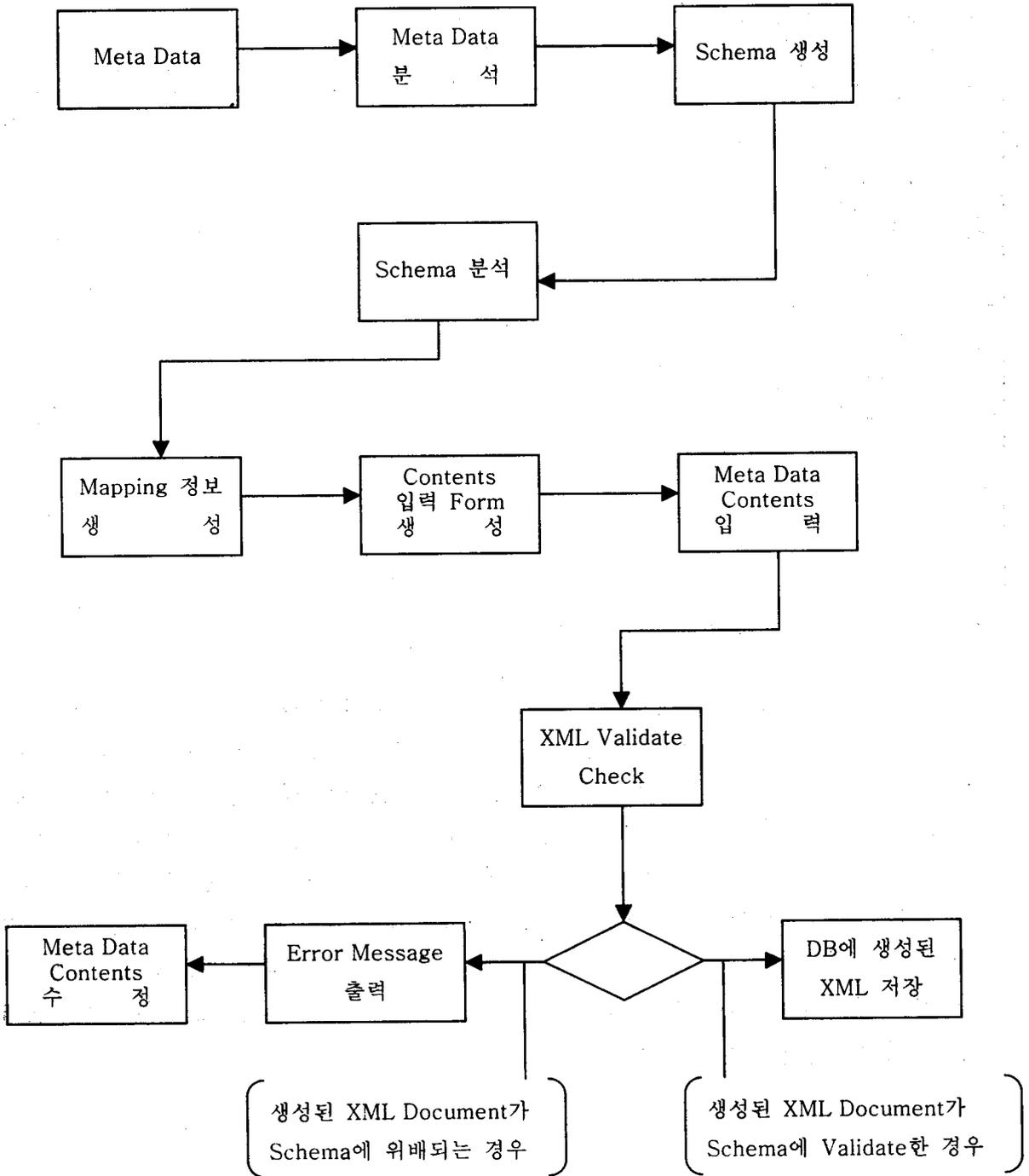
스키마 <input checked="" type="checkbox"/> 작품 <input type="checkbox"/> 재어번호 <input type="text" value="" /> 최초입력 <input type="text" value="///" /> 최종정신 <input type="text" value="///" />

	항목	입력	스킴
2	<input checked="" type="checkbox"/> 작품명		
3	<input checked="" type="checkbox"/> <input type="checkbox"/> L 작품명		AACR2.SNUCR
7	<input type="checkbox"/> <input type="checkbox"/> L 다른 작품명		AACR2.SNUCR
11	<input type="checkbox"/> 음악가정보		
12	<input checked="" type="checkbox"/> <input type="checkbox"/> L 음악가명		
14	<input type="checkbox"/> <input type="checkbox"/> L 음악가의 별칭		
16	<input type="checkbox"/> <input type="checkbox"/> L 생몰년		
18	<input type="checkbox"/> <input type="checkbox"/> L 역할		
20	<input type="checkbox"/> <input type="checkbox"/> L 이-메일		
22	<input type="checkbox"/> <input type="checkbox"/> L 소속기관		
24	<input checked="" type="checkbox"/> <input type="checkbox"/> 주제		DDC
27	<input type="checkbox"/> <input type="checkbox"/> 주기사항		
29	<input type="checkbox"/> <input type="checkbox"/> 작품관련 일자		
30	<input checked="" type="checkbox"/> <input type="checkbox"/> L 작품생성일		DCMI-Period
34	<input type="checkbox"/> <input type="checkbox"/> L 첫 연주/공연 일시		DCMI-Period
38	<input type="checkbox"/> <input type="checkbox"/> L 첫 출판일시		DCMI-Period
42	<input type="checkbox"/> <input type="checkbox"/> 작품형식		SMF
45	<input type="checkbox"/> <input type="checkbox"/> 언어		ISO639-2
48	<input type="checkbox"/> <input type="checkbox"/> 관련작품		
49	<input type="checkbox"/> <input type="checkbox"/> L is Part Of		URI
53	<input type="checkbox"/> <input type="checkbox"/> L Has Part		URI
57	<input type="checkbox"/> <input type="checkbox"/> L 연주/악보/공연 정보		URI
61	<input type="checkbox"/> <input type="checkbox"/> L 음반/매체 정보		URI
65	<input type="checkbox"/> <input type="checkbox"/> 내용범위		

A. 演

지금까지 설명해 온 XML Schema를 기반으로 구축된 전자도서관 시스템에서의 데이터 입력 및 흐름을 다시 한번 요약해 본다면 우선, 새로운 메타데이터의등장이 있고, 그 메타데이터에 대한 연구 및 분석을 한 후 적용할 Schema를 생성하며, 새로 생성된 Schema를 분석하여 필요한 매핑을 위한 정보를 저장한다. 여기까지가 메타데이터 입력과 관리를 위한 기반이 되는 시스템의 구축과정으로서 다음 단계는 실제 데이터 입력 및 관리 단계이다. 클라이언트 프로그램인 CMS를 열어 실제 메타데이터의 입력을 하고 저장을 하면, 우선 데이터의 무결성 검사 즉, Validate 체크를 하게 된다. 유효성 검사에 위배가 될 경우 에러가 날 것이며, 그렇지 않을 경우에는 비로서 그 메타데이터에 대한 XML 파일이 생성되면서 DB에 저장된다.

<그림 11> XML Schema 기반의 메타데이터 흐름



3. 문제점 및 개선방향

현재까지 국내 도서관계에서 XML Schema를 기반으로 구축된 시스템은 서울대학교가 유일하다. 어떤 분야에서든지 처음으로 시도하는 것에는 위험과 시행착오가 따르기 마련이다. 그럼에도 불구하고 서울대학교에서 그 첫 시도를 주저하지 않았던 것은 XML의 가장 큰 장점인 데이터의 호환성 때문이었다. 데이터의 호환은 서로 다른 시스템의 통합뿐만 아니라, 정보의 공유측면에서 매우 중요하다. 처음 연구단계에서는 이러한 XML의 장점을 살려 Schema 기반으로 구현한 의도를 충분히 반영하기 위해, 서울대학교에서 연구개발한 메타데이터 셋을 국내표준화로 만드는 것이 목표가 되었다. 이 표준화가 실현됨으로써 유관기관간의 데이터 호환과 상호운영이 원활하게 되어 국내 도서관계가 보다 전사적인 시스템으로 발전이 되는 청사진을 가지고 출발한 것이었다. 그러나, 스키마나 메타데이터의 공동 연구나 협의가 턱없이 부족하고 어려운 점이 있었기에 서울대학교 독자적인 연구로 진행이 될 수밖에 없었다. 이로 인해 XML Schema 기반의 서울대학교 메타데이터의 상호운영과 교환을 목표로 한 효용성 여부는 현재로서는 평가되기 힘들다. 그러나, 데이터의 공유와 호환이 용이한 시스템을 서울대학교가 국내 최초로 진행하고 개발하였다는 데 큰 의의를 부여할 수 있고, 이러한 시스템의 효용과 발전은 여러 기관의 많은 관심과 협의로 이루어 나갈 수 있을 것이며, 그렇게 되길 기대해본다.

The following information is provided for your reference:

1. The first section of the document contains a list of items.

2. The second section discusses the importance of maintaining accurate records.

3. The third section outlines the procedures for handling sensitive information.

4. The fourth section provides details on the reporting requirements.

5. The fifth section describes the roles and responsibilities of the staff.

6. The sixth section covers the financial aspects of the project.

7. The seventh section addresses the legal and regulatory compliance issues.

8. The eighth section discusses the environmental impact of the activities.

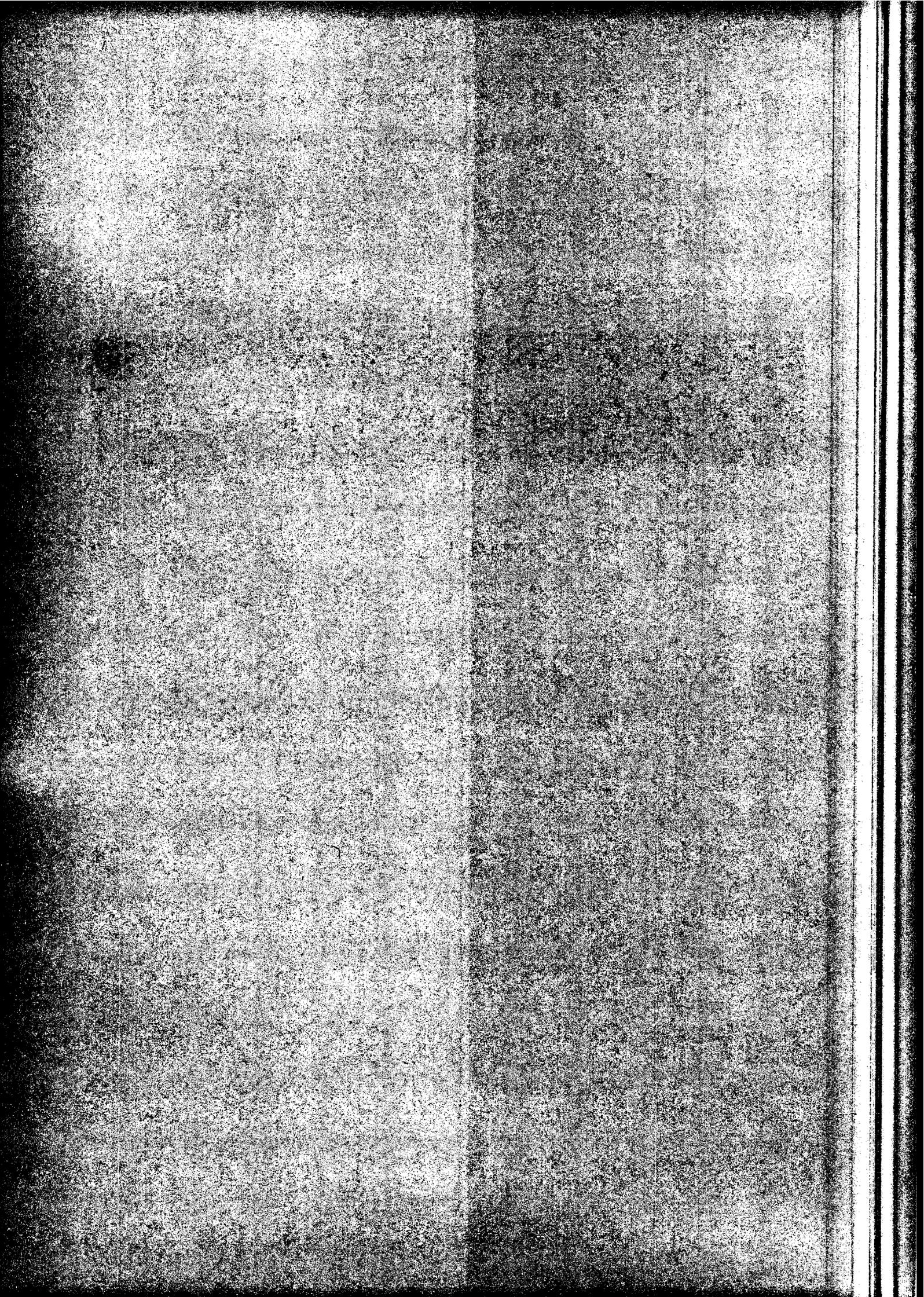
9. The ninth section provides information on the public consultation process.

10. The tenth section concludes with a summary of the key findings.

[서명표]

2003. 5

강원아
quass@gnu.ac.kr



< 사례발표 4 >

서울대학교 메타데이터

2003. 5.

강 연 아
(quasi@snu.ac.kr)

차 례

1. 서론	2
2. 서울대학교 메타데이터 연구	2
2.1. 음악	3
2.2. 고문헌	10
2.3. 학술행사	13
3. 서울대학교 자료 메타데이터	16
3.1. 대학사료	16
3.2. 미술	18
3.3. 의학	21
3.4. 학술지	23
4. 문제점 및 향후 과제	25
5. 결론	26
참 고 문 헌	27

1. 서론

서울대학교에서는 두뇌 한국 21사업의 일환으로 2001년 12월부터 전자도서관 구축사업을 추진하여 2003년 5월 21일 그 가동식을 가질 예정이다.

전자도서관은 전통적인 서지자료 외에 각종 전자적인 형태의 자료를 DB로 구축하여 네트워크를 통해 이용자에게 편리하게 제공하는데 그 목적이 있다. 따라서, 그렇게 구축한 자료를 이용자가 효율적으로 검색하고 제공받을 수 있도록 하는 것은 전자도서관을 구축하는데 있어 가장 우선시 되는 부분 중의 하나이며, 이를 위해 각 소장 자료의 특성을 반영하면서 다양한 검색 접근점을 제시할 수 있는 메타데이터의 마련이 필수적이라 할 수 있을 것이다.

서지 데이터베이스에 있어서 표준 메타데이터 양식으로 쓰이고 있는 MARC는 전통적인 자료를 표현하는데 있어 그 효용성을 널리 인정받고 있으나, 정보통신 기술의 발달로 급증하고 있는 인터넷 자원 및 각종 디지털 형태의 자원을 수용하는 데는 무리가 있다. 그리하여, MARC 중심의 메타데이터 통합이나 MARC의 대체형식 개발, 혹은 MARC와 새로운 메타데이터의 공존 등 세계적으로 여러 가지 양상의 주장이 제기되고 있으며, 이에 따른 다양한 시도 역시 이루어지고 있는 형편이다.

서울대학교는 이중에 MARC와 새로운 메타데이터의 공존을 모색하기로 하고, 서지자료와 같이 MARC로 정리 가능한 정보자원은 MARC로 정리하고, MARC로만 정리하였을 경우, 해당 정보자원의 특성을 표현하는데 문제가 있거나 정보 접근점을 다양하게 제시할 수 없는 자원을 대상으로 새로운 메타데이터를 개발하는 연구를 수행하였다.

서울대학교의 메타데이터 연구는 서울대학교 전자도서관 구축을 위한 메타데이터 표준 정립하며, 더 나아가 장래 국내에서의 대학 전자도서관 구축 사업 전개에 활용될 수 있는 대학 전자도서관 메타데이터 표준 제시하는데 그 연구의 목적이 있다.¹⁾

2. 서울대학교 메타데이터 표준 연구²⁾

서울대학교 전자도서관 구축을 위한 메타데이터 연구는 웹 자원을 기술하는데 주로 사용되고 있는 더블린 코어(Dublin Core element set: DC)의 요소 및 한정어를 기반으로 하였으며, 웹 자원이 아닌 특정 소장 자료들의 특성을 표현하기 위해 필요한 추가 요소를 개발하는 방향으로 진행되었다.

서울대학교는 일반 자료들에 비해 그 특성이 뚜렷한 음악, 고문헌, 학술행사 자료를 중심

1) 최유미. 2002. "서울대학교 전자도서관 구축을 위한 메타데이터 연구", 국립대학도서관보 제20집

2) 서울대학교 전자도서관 구축을 위한 메타데이터 연구는 오삼균 교수(성균관대학교 문헌정보학과)와 채진석 교수(인천대학교 컴퓨터공학과)와의 공동연구로 2001년 12월부터 연구되었다.

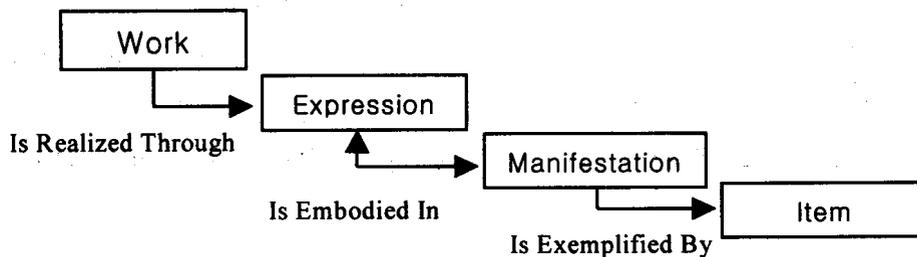
으로 연구를 수행하였다.

2.1. 음악

음악 자료를 메타데이터로 표현하는 데는 원작품(work) 뿐만이 아니라, 그 작품이 연주 되었을 때 각각의 연주가 갖는 정보 및 그 연주 자료가 수록된 매체에 대한 정보 세 가지가 모두 중요하다. 예를 들어, 어떤 음악가가 작곡한 작품은 여러 사람에게 의해 연주될 수 있으며, 한 작품에서 개개의 악장들이 따로따로 연주되는 경우도 많다. 또한 한 작품이 같은 사람에게 의해 여러 번 연주될 수도 있다. 이 때 각각의 연주곡은 서로 다른 특성을 지니므로 개별적인 메타데이터를 작성해 주어야 할 것이다. 또한, 연주곡이 수록된 매체 역시 다양하게 제작될 수 있으며, 이것은 악보의 경우도 마찬가지로, 음악자료는 원작의 파생 자료가 매우 다양함을 알 수 있다. 따라서, 이러한 파생 자료를 개별적으로 검색할 수 있음과 동시에 원작품의 검색 후 개별 작품을 계층적으로 검색할 수 있는 입체적인 구조의 메타데이터가 필요하다.

서울대학교에서는 이러한 음악분야 메타데이터 구현을 위하여 IFLA FRBR (Functional Requirements for Bibliographic Records) 모델과 인디애나 대학의 음악전자도서관 (DML : Digital Music Library)의 데이터 모델을 기본 모델로 적용하고, 서울대학교 자료에 부합하는 자료의 특성을 표현할 수 있도록 메타데이터를 변형, 추가하였다.

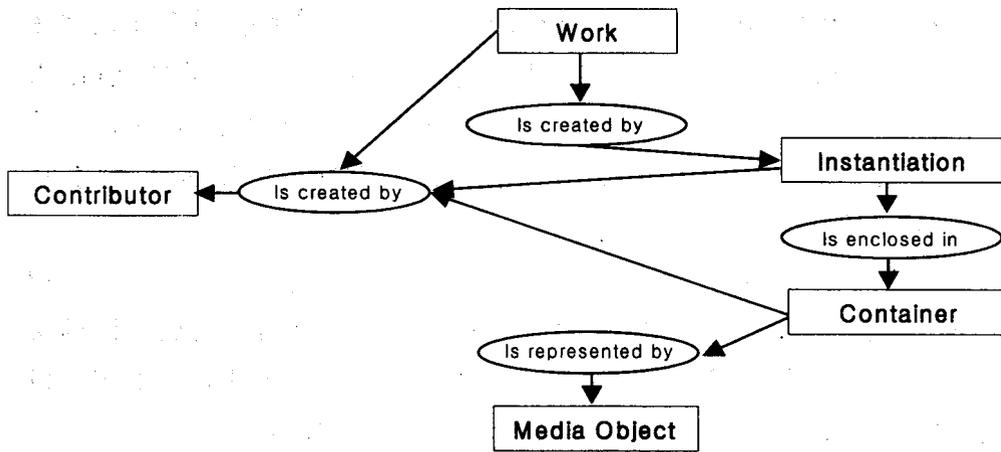
IFLA FRBR 모델을 간단하게 표현하면 다음과 같다.



<그림 1> Works, Expressions, Manifestation and Items

Work로 표현되는 원작품은 특정 공간과 시간을 배경으로 실행되며 이것을 Expression이라 한다. Expression은 Manifestation을 통해 구체화되며, Manifestation이 대량 제작될 때 다수의 Item이 발생한다. 예를 들어, 어떠한 곡(Work)은 연주를 통하여 실행되며(Expression), 연주가 CD와 같은 매체에 수록되는 것을 물리적인 형태의 실현(Manifestation)이라고 볼 수 있다. 마지막으로 음반회사가 대량생산한 개개의 CD를 Item이라 한다.

인디애나 대학(Indiana University)의 DML 메타데이터는 USMARC에서 추출한 주요 기술 메타데이터를 활용하고 있으며, 여기에 관리 메타데이터를 추가하여 적용하고 있다.



<그림 2>인디애나 대학 DLM 메타데이터 모델

IFLA FRBR 모델에 비교하여 볼 때, Expression은 Instantiation으로, Manifestation은 Container라는 개념으로 대치시켜 볼 수 있다. 인디애나 대학의 DML은 이 외에, Media Object라는 엔티티를 두어 디지털화된 음악관련 파일을 이용자에게 직접 서비스 할 수 있도록 하였다. 또한 메타데이터에 작품구조(Work Structure)와 수록매체 구조(Container Structure)를 기술할 수 있도록 하였으며, 각 구조 메타데이터의 요소와 미디어 자원을 연결할 수 있도록 구성되어 있다.

서울대학교 음악 메타데이터는 DC를 기반으로 상기와 같은 두 모델을 참고하고, 서울대학교 음악대학의 음악 전문가의 의견을 참조하여 작품(Work), 표현(Instantiation), 작품구조선언(Work Structure Declaration), 작품구조연결(Work Structural Bindings), 수록매체(Container), 수록매체구조 및 구조연결(Container Structure Declaration & Structural Bindings), 미디어 자원(Media Object) 메타데이터 셋으로 구성되었다.

이와 같은 메타데이터 셋을 통해 양악과 국악의 작품 구조(Section) 정보를 표현하고, 이용자가 음악 자료를 검색했을 때, 서로 연관되어 있는 자료의 구조를 파악하여 원하는 정보로 바로 접근할 수 있도록 관계(relation) 구현에 중점을 두었으며, 특정 작품의 미디어 자원과 직접 연결이 가능하도록 설계하였다. 각각의 메타데이터 셋에서 DC의 요소/한정어와 서울대학교에서 추가한 요소는 각각 요소명 앞의 dc/dcterms와 snu라는 접두어로서 구분하였다.

<표 1>의 작품(Work) 메타데이터에서 dc:creator라는 요소는 vCard의 다양한 요소를 하위요소로 두었는데, 전자명함이라는 개념을 통하여 작품을 생성한 사람에 대해 보다 자세히 기술하고, 앞으로 웹 기반 정보자원 통합검색이 더욱 발달하였을 때 창작자의 정보를 통해 연관된 정보를 얻을 수 있도록 하기 위함이다.

snu:workType은 기술하고 있는 작품을 전집(Collective Work)과 전집에 속한 개별 곡(Single Work)으로 구분하는 역할을 하며, snu:workRelation에서 그 관계유형을 규정하

고 있다. 작품이 내부구조를 가질 때에는 `snu:hasStructure`에서 그 여부를 기술하고, <표 2>의 작품구조(Work Structure Declaration)에서 보다 상세한 정보를 기술하게 된다. 또한, 내부 구조를 갖는 특정 작품에 대한 연주곡과 그 연주곡의 미디어 자원이 사운드 파일로 존재할 경우, 전체 작품연주를 듣지 않고, 원하는 부분만 선택적으로 들을 수 있도록 <표 3>과 나머지 메타데이터 셋을 구성하였다.

서울대학교 메타데이터 셋 연구가 진행되고, 전자도서관 자료를 실제로 적용하여 DB로 구축하면서 여러 요소들이 수정, 추가, 삭제되었다. 그 중 추가된 요소의 예로 `snuterms:requestDept`를 들 수 있는데, 이것은 기관의 특성을 살려 대학도서관 자료의 이용자들이 대학별 자료 검색을 할 수 있도록 하기 위해 추가되었다.

<표 1> 작품(Work) 메타데이터 셋

요소명	제1 하위요소	정의
dc:identifier		작품에 부여된 고유한 번호로 시스템상의 고유한 Record ID
dc:title		작품의 제목
	snuterms:mainTitle(D)	작품의 통일서명(Uniform Title). AACR2 목록규칙 표준을 충족시켜야만 한다
	dcterms:alternative(D)	작품의 대등서명
dc:creator		editor, conductor 등 Work를 생성한 사람
	vCard:FN(D)	editor, conductor 등 Work를 생성한 사람 이름
	vCard:N	
	vCard:BDAY	Work를 생성한 사람의 생몰년
	vCard:ROLE	작품의 작곡/작사가의 작품에서의 역할
	vCard:EMAIL	
	vCard:ORG	
dc:subject		작품의 주제
dc:description		작품에 대한 내용주기
dc:date		자원의 존재기간 동안 어떠한 사건이 발생한 날짜
	snuterms:created(D)	작품의 창작일이나 창작일로 예측되는 일자
	snuterms:dateOfFirstPerformance(D)	작품이 처음으로 연주된 일시
	snuterms:dateOfFirstPublication(D)	작품이 처음으로 출판된 일시
dc:type		작품의 형식
dc:language		작품의 창작에 사용된 언어
dc:relation		Work 와 관련된 자원들과 관계를 기술한다.
	dcterms:isPartOf	작품유형이 Single 일때, Collection으로 기술된 작품의 메타데이터 ID기입
	dcterms:hasPart	작품유형이 Collection일때, 이에 속한 Single 작품의 메타데이터 ID 기입
	snuterms:instantiation	관련된 Instantiation 메타데이터 ID
	snuterms:isContainedBy	관련된 Container 메타데이터 ID
dc:coverage		작품의 내용이 다루고 있는 범위 (시간적, 공간적)
	dcterms:spatial(D)	작품의 내용이 다루고 있는 범위 (공간적)
	dcterms:temporal(D)	작품의 내용이 다루고 있는 범위 (시간적)
dc:rights		작품의 저작권정보를 기술한다.
snu:workType		작품의 유형을 기술한다.
snu:originalTextTitle		음악작품에 사용된 비음악작품의 Title
snu:instrumentation		작품의 연주에 사용될 악기관련 사항(예. 피아노곡, 첼로곡 등)을 기술
snu:key		작품의 키 (C 단조 등)
snu:place		작품과 관련된 장소 정보
	vra:creationSite	작품을 작곡/작사한 장소
	snuterms:placeOfFirstPerformance	작품이 처음 연주된 장소
	snuterms:requestDept	신청기관
snu:hasStructure		내부적인 하부 구조를 가지는지의 유무

<표 2> 작품구조(Work Structure Declaration)

요소명	제 1 하위요소	정의
dc:identifier		내부구조를 갖는 작품의 ID
snu:workLabel		화면출력에 사용되는 작품레이블
snu:section		작품을 구성하는 섹션에 대한 기술
	dc:title	검색의 키가 되는 섹션명으로 공식명칭을 기술
	dc:identifier	섹션의 고유 ID
	snu:sectionLabel	화면출력에 사용되는 섹션레이블
	snu:sectionType	섹션 유형기술(예:악장)

<표 3>작품에 대한 표현물 (Instantiation) 메타데이터 셋

요소명	제1하위요소	정의
dc:identifier		자동적으로 생성되는 Instantiation의 식별번호
dc:title		
	snterms:mainTitle	Container에 나타나지는 실질적인 제목
	dcterms:alternative	Work의 mainTitle에 입력한 작품명(검색의 키로서 활용하기위함)
dc:creator		editor, conductor 등 Instantiation을 생성한 사람
	vCard:FN(D)	editor, conductor 등 Instantiation을 생성한 사람 이름
	vCard:N	
	vCard:BDAY	
	vCard:ROLE	
	vCard:EMAIL	작품의 작곡/작사가의 작품에서의 역할
	vCard:ORG	
dc:subject		Instantiation의 주제
dc:description		Instantiation에 관한 설명
dc:date		Instantiation과 관련된 날짜 정보
	dcterms:created(D)	Instantiation이 만들어진 날짜(Performance가 일어난 날짜)
dc:type		Instantiation의 유형에 대한 정보
dc:language		Instantiation에 사용된 언어
dc:relation		Instantiation과 work/Container와의 관계 기술
	dcterms:isVersionOf(D)	Instantiation과 관련된 work과의 관계를 규정한다.
	snterms:isModifiedFor(D)	이 Instantiation을 수정한 Instantiation들과의 관계를 규정한다.
	snterms:isModifiesFrom(D)	이 Instantiation이 수정한 Instantiation들과의 관계를 규정한다.
	snterms:isContainedBy(D)	기술되고 있는 Instantiation이 속한 Container와 연결
dc:coverage		Instantiation과 관련된 공간/시대정보
	dcterms:spatial(D)	Instantiation의 유형이 Performance인 경우 Performance가 일어난 장소.
	dcterms:temporal	
snu:representationType		Container내에서의 내용의 표현형태를 구분하는 유형(Text, Score, Audio, Visual)
snu:completeness		Work 내용의 전체를 Instantiation이 표현하는지의 여부
snu:key		작품의 키 (C 단조 등)
snu:instrumentation		Instantiation의 연주에 사용될 악기관련 사항(예. 피아노곡, 첼로곡 등)을 기술
snu:place		
	snterms:requestDept	신청기관
snu:hasStructure		
dc:rights		Instantiation의 저작권과 관계된 모든 정보, copyright 날짜와 소유자 정보

<표 4>작품구조연결 (Work Structural Binings) 메타데이터 셋

요소	제1하위요소	제2하위요소	정의
snu:instantiationID			작품을 표현한 표현물 ID
snu:workID			작품 ID
snu:section			작품구조에 속한 섹션과 미디어자원들과의 연결관계구조에 관한 기술
	snu:sectionID		작품구조에서 선언된 작품의 섹션 ID
	snu:binding		해당 섹션을 디지털화한 미디어자원과의 연결에 관한 기술
		snu:mediaObjectID	연결될 미디어 자원의 ID
		snu:begin	미디어 자원의 시작부분
		snu:end	미디어 자원의 종료부분

<표 5>수록매체(Container) 메타데이터 셋

요소명	제1하위요소	정의
dc:identifier		자동적으로 생성되는 Container의 식별번호
dc:title		Container의 대표적인 제목
	snuterms:mainTitle(D)	Container의 시리즈 명
	snuterms:seriesTitle(D)	editor, conductor 등 Container를 생성한 사람
	snuterms:seriesNo	
dc:creator		editor, conductor 등 Container를 생성한 사람
	vCard:FN(D)	editor, conductor 등 Container를 생성한 사람 이름
	vCard:N	
	vCard:BDAY	
	vCard:ROLE	작품의 작곡/작사가의 작품에서의 역할
	vCard:EMAIL	
	vCard:ORG	
dc:description		Container에 관한 설명
dc:publisher		Container를 출판한 사람
dc:date		Container와 관련된 일시정보
	dcterms:issued(D)	Container의 출판일
dc:format		Container의 특징에 대한 정보
	dcterms:extent(D)	Container의 시간, 양
	dcterms:medium(D)	container의 물질적인 매체
dc:language		Container에 사용된 언어
dc:relation		Instantiation와 관련된 자원 기술
	snuterms:containedWork(D)	관련된 작품(Work) ID(s)
	snuterms:containedInstantiation(D)	관계된 Instantiation, container에 담긴 모든 Instantiation .ID
snu:edition		Container의 판차정보
snu:condition		Container 상태 정도, 디지털화 될 때 물리적인 것의 상태
snu:place		
	snuterms:requestDept	신청기관
snu:hasStructure		내부적인 하부 구조를 가지는지의 유무
dc:rights		Container에 대한 저작권정보를 기술한다.

<표 6> 수록매체 구조 선언 및 연결 (Container Structure Declaration & Bindings)

요소	제1하위요소	제2하위요소	제3하위요소	제4하위요소	정의
snu:containerID					수록매체 ID
snu:item					수록매체의 항목
	snu:itemID				항목ID
	snu:itemLabel				화면출력에서 사용할 항목 레이블
	snu:div				수록매체의 구성 부분, 디비전
		snu:divID			디비전 ID
		snu:divLabel			화면출력에서 사용할 디비전 레이블
		snu:chunk			미디어 자원과 연결되는 단위
			snu:chunkID		Chunk의 ID
			snu:chunkLabel		화면출력에 사용될 레이블
			snu:binding		하나이상의 미디어 자원에 저장된 디지털 콘텐츠 섹션과의 연결정보
				snu:mediaObjectID	연결될 미디어 자원의 ID
				snu:begin	미디어 자원의 시작부분
				snu:end	미디어 자원의 종료부분

2.2. 고문헌

고문헌의 메타데이터 연구는 고문헌 연구자를 대상으로 보다 질높은 도서관 서비스를 하기 위해 착수되었다. 전자도서관은 이용자가 도서관에 직접 오지 않고 온라인으로 소장목록을 검색하고, 원하는 자료의 원문을 그 자리에서 직접 열람할 수 있도록 하는데 그 효용성이 있다. 즉, 도서관에 오지 않고도 도서관을 이용하는 것과 같은 효과를 내는 것이 주된 목적 중의 하나인데, 다른 자료와 달리 고문헌 자료의 경우에는 고문헌 연구자들이 실제 자료를 직접 보는 것과 동일한 효과를 구현하는 것은 쉽지 않은 일이다. 이를 위해 고문헌 메타데이터 연구는 MARC에서 고문헌을 기술하는 것보다 상세한 메타데이터를 만들어 다양한 항목이 검색기로 사용될 수 있도록 하는 방향으로 진행되었다.

그러나 이러한 경우, 고서에 대한 배경 지식이 부족한 일반 이용자에게는 다양한 검색키는 검색에 오히려 장애가 될 수 있으며, 이와 같은 상세한 검색을 할 필요성 그 자체 역시 그다지 크지 않다는 문제점이 발견되었다.

또한, 전자도서관을 구축하여 기대되는 효과를 내기 위해서는 고문헌에 대한 연구가 선행되어야 한다는 점 역시 고문헌 메타데이터 연구의 한 장애가 되었다. 예를 들어 어떤 자료의 탁본을 이미지로 서비스하는 경우, 제한된 크기의 모니터에서 탁본은 실제 도서

관에서 탁본을 열람하도록 하였을 경우에 비해 그 유용성이 크게 떨어진다. 이는 시스템이나 개인 컴퓨터와 같은 장비의 성능에 따라 구현 능력이 크게 좌우되기 때문이며, 이미지의 가독성이 고문헌 자료의 특성상 실물이 주는 가독성보다 떨어지기 때문이다. 따라서, 애초에 목적인 서비스를 위해서는 대상 탁본의 내용이 텍스트 파일로 준비되어 참조할 수 있어야 하며, 이 탁본의 내용에 대한 해제 파일 역시 필요하다는 것을 알 수 있다. 이와 같은 부분에서, 고문헌의 메타데이터는 그 연구가 진행되어 완료된다고 하더라도 실제 상용성은 떨어지는 문제가 발생하게 된다. 메타데이터 셋은 잘 정비되어 있더라도, 실제 데이터 값이 없으면 무용지물이기 때문이다. 그리고, 실제 데이터 값을 얻기 위해서는 각각의 고문헌에 대한 주제전문사서와 해당 학과 전문 연구진들의 합동 연구가 선행되어야 함을 알 수 있다. 반대로, 합동 연구가 병행되지 않고서는 메타데이터 셋의 항목 값 역시 정하기 어렵다. 고문헌에 대한 학술 연구를 지속적으로 행하면서 메타데이터 연구도 끊임없이 수정 보완을 해야 하기 때문이다.

따라서 이 연구는 1년이나 2년의 기간만으로는 도저히 수행할 수 없는 방대한 양의 사업이며, 큰 규모의 예산과 각 분야의 전문 인력이 지속적으로 투자되어야 하는 학술 연구 사업이라 할 수 있다. 더구나, 이 연구가 국내대학 도서관 중 서울대에서 처음으로 시도되고 있음을 감안할 때, 앞으로의 파급 효과로 보아 짧은 시간에 표준안을 만들어 낼 수 없음은 뚜렷하다.

이러한 이유로 2003년 전자도서관 가동을 앞두고 초기에 음악, 고문헌, 학술행사의 세 부분으로 나뉘어 진행되어 오던 메타데이터 연구에서 고문헌은 부득이하게 제외되었다. 고문헌에 대한 메타데이터 연구는 앞으로 충분한 시간과 인력을 투자하여 진행되어야 할 것이다.

현재 연구 개발이 중지된 고문헌 메타데이터 셋은 <표 7>과 같다.

<표 7>서울대학교 고문헌 메타데이터 셋

요소명	제1하위요소	제2하위요소	고도서	고문서
dc:title				
	snu:mainTitle		주제목(권수제)	문서명
	dcterms:alternative			
		snu:translatedTitle	한글서명	한글 문서명
		snu:headingTitle	표제	
		snu:plateCentralTitle	판심서명	
		snu:coverTitle	표지서명	
		snu:otherTitle	다른 서명	다른 명칭
dc:creator				
	snu:creatorName			
		snu:creatorOriginalName	저자명	발급자명
		snu:creatorTranslatedName	저자 한글명	발급자 한글명
	snu:recipientName			
		snu:recipientOriginalName		수급자명
		snu:recipientTranslatedName		수급자한글명
	snu:editorName			
		snu:editorOriginalName	편자명	편자명
		snu:editorTranslatedName	편자한글명	편자한글명
snu:subject				
	snu:classification			
		snu:firstCategory	대분류	대분류
		snu:secondCategory	중분류	중분류
dc:description				
	dcterms:tableOfContent		목차	
	dcterms:abstract		초록	발급목적
dc:publisher			발행처/발행자	발행처/발행자
dc:date				
	dcterms:issued		발행일	발행일
dc:type			고도서유형(예:문체)	고문서 유형
dc:format				
	dcterms:medium		매체	매체
	dcterms:extent		페이지수/크기	페이지수/크기
dc:language			언어	언어
dc:relation				
	dcterms:isFormatOf		디지털화된 자료의 원본의 서지레코드 URI	디지털화된 자료의 원본의 서지레코드URI
	snu:hasCommentary		해제의 서지레코드URI	해제의 서지레코드URI
	dcterms:isPartOf			성책고문서 서지레코드URI
	dcterms:hasPart			탈림자료 서지레코드URI
	snu:metaAccess		관련자료의 서지레코드 URI	
dc:coverage			고도서/고문서의 내용이 다루는 범위	
	dcterms:temporal		시대	시대
	dcterms:spatial		지역	사건지역

2.3. 학술행사

다층적인 구조를 갖는 학술행사 메타데이터는 IFLA FRBR 모델의 개념을 참조하여 개발되었다.

규모가 큰 컨퍼런스나 세미나와 같은 학술행사는 <표 9>의 세션 메타데이터 셋을 통해 가지고 있는 각 세션에 대한 정보를 기술하고, 세션이 없는 경우에는 이를 생략하고 <표 10>의 기사 메타데이터 셋을 통해 개개의 발표자료를 기술할 수 있도록 하였다. <표 11>의 미디어자원 메타데이터 셋은 학술행사를 녹음하거나 촬영하여 DB로 구축하고 전자도서관 통합검색에서 학술행사의 정보를 얻은 뒤 바로 해당하는 미디어자원을 연결하여 이용할 수 있도록 서비스하기 위해 개발되었다.

<표 8> 학술행사 메타데이터 셋

요소명	제1하위요소	정의
dc:identifier		학술행사에 부여된 고유번호로 시스템상의 고유 레코드 ID
dc:title(필)		학술행사의 명칭 관련 요소로, 메타데이터의 의미구조연결에 사용
	snuterms:mainTitle	학술행사의 명칭
	dcterms:alternative	학술행사의 이칭
	snuterms:eventNumber	학술행사의 회차
dc:contributor		학술행사의 개최 후원기관
	vCard:FN(D)	
	vCard:N(D)	
	vCard:BDAY	
	vCard:ROLE	
	vCard:EMAIL	
	vCard:ORG	
dc:subject(필)		학술행사에서 발생하는 주제
dc:description		학술행사의 내용주기
dc:publisher(필)		
dc:date(필)		학술행사와 관련된 일시 정보
	snuterms:startDate	학술행사의 시작일과 종료일
	snuterms:endDate	
dc:type(검색제한자)		학술행사의 내용구분
dc:coverage		학술행사의 내용이 다루고 있는 범위 (시간적, 공간적)
	dcterms:spatial	공간적
	dcterms:temporal	시간적
dc:language(검색제한자)		학술행사에서 사용된 언어
dc:relation		학술행사에서 발생하는 자료와의 연결정보 기술
	snuterms:hasArticle	학술행사에서 발표되는 Article과의 연결정보
	snuterms:hasSession	학술행사에 포함되는 Session과의 연결정보
snu:place		
	vra:creationSite	학술행사의 개최지
	snuterms:placeType	서울대 소장여부
	snuterms:requestDept	신청기관

<표 9> 세션 메타데이터 서브 셋

요소명	제1하위요소	정의
dc:identifier		학술행사에 부여된 고유번호로 시스템상의 고유 레코드 ID
dc:title(필)		학술행사에서 다루는 주 내용의 제목
	snuterms:mainTitle	
	dcterms:alternative	
dc:contributor		
	vCard:FN(D)(필)	
	vCard:N(D)(필)	
	vCard:BDAY	
	vCard:ROLE	
	vCard:EMAIL	
	vCard:ORG	
dc:type		Session의 내용구분
dc:relation		Session과 관련된 자료와의 연결정보
	snuterms:isContainedBy	세션이 속한 학술행사와의 연결정보
	snuterms:hasArticle	학술행사에서 발표되는 Article과의 연결정보
snu:place		
	snuterms:placeType	서울대소장여부
	snuterms:requestDept	신청기관
dc:rights		Session의 저작권정보 및 이용제한정보를 기술

<표 11> 미디어 자원(Media Object) 메타데이터 셋

요소명	제1하위요소	정의
dc:title		학술회의/세미나 개별 발표(Article)의 미디어 자원 타이틀(과일명)
dc:creator		미디어 자원 작성에 기여한 사람
	snu:role	미디어 자원작성에 기여한 사람의 역할
dc:description		미디어 자원에 관한 내용주기
dc:date		미디어 자원이 최종 수정된 일시정보
dc:format		미디어 자원의 형태
	snu:formatType	미디어자원 형식의 대범주
	snu:medium	미디어 자원의물리적 매체 정보
	snu:extent	미디어 자원의 시간이나 크기 등의 정보
dc:identifier		미디어 자원 ID

3. 서울대학교 자료 메타데이터

서울대학교의 자료 메타데이터 중 음악, 고문헌, 학술행사를 제외한 나머지 자료의 메타데이터는 이 세 자료의 연구를 진행하면서 얻은 성과를 토대로 개발되었으며, 자료의 특성상 입체적이거나 다층적인 구조를 갖는 대신 평면적인 구조를 갖고 있는 경우가 많다. 기타 메타데이터는 대학사료, 미술자료, 의학자료, 기사색인, 학술지의 경우로 나누어 볼 수 있다.

3.1. 대학사료

대학사료 메타데이터에서는 기록사료만이 갖는 독특한 주제 분야를 한눈에 볼 수 있도록 <표 12> dc:subject 요소의 제1하위요소인 주제분류표를 DDC외에 UAC(대학기록관 분류), KEAC(한국교육사고 분류), snuPress(대학신문사 분류) 등 기관별로 추가 확장하였다.

dc:description 요소에서는 초록이나 키워드 외에 일반 주기사항을 분리할 필요성이 제기되면서 snuterms:general을 추가하였다. 또한, 기록편찬기관을 기술하는 요소인 dc:publisher는 처음에 단일항목이었으나, 단체명의 약칭 등 색인 항목을 잡아주기 위해 이를 공식명칭과 다른 명칭의 둘로 나누어 확장하였다. 기록생산일자를 나타내는 dc:date는 date.issued라는 하위요소로 자료의 생성일을 기술하였으나, 복간년을 갖는 자료로 인하여 snuterms:startDate와 snuterms:endDate로 나누어 자료의 처음 생성일자(또는 단일년도를 갖는 자료의 생성일)와 자료의 마지막 생성일자를 구분하여 기술하는 방법을 취하였다.

<표 10>기사(Article) 메타데이터 셋

요소명	제1하위요소	정의
dc:identifier		Article에 부여된 고유번호로 시스템상의 고유 레코드 ID
dc:title(필)		
	snuterm:mainTitle	Article 명칭
	dcterms:alternative	Article 이칭
dc:creator(필)		학술회의/연구/세미나에서 발표되는 Article의 저자 및 공저자
	vCard:FN(D)(필)	
	vCard:N(D)(필)	
	vCard:BDAY	
	vCard:ROLE	
	vCard:EMAIL	Article에서의 Creator의 역할
	vCard:ORG	
dc:subject(필)		Article의 주제
dc:description		Article에 대한 주기사항
	dcterms:tableOfContents(D)	Article의 목차
	dcterms:abstracts(D)	Article의 초록
	snuterm:keywords	
dc:date(필)		Article에 관련된 날짜 정보
dc:type(검색제한자)		Article의 내용구분
dc:coverage		Article의 내용이 다루고 있는 범위 (시간적, 공간적)
	dcterms:spatial	공간적
	dcterms:temporal	시간적
dc:language(검색제한자)		Article에서 사용된 언어
dc:relation		Article에서 발생하는 자료와의 연결정보 기술
	snuterm:isArticleOfConference	Article이 발표된 학술회의/세미나와의 연결정보 (Is Article of Session과 함께 기술할 수 없음)
	snuterm:isArticleOfSession	Article이 발표된 학술회의/세미나내의 Session과의 연결정보 기술(isArticleOfConference와 함께 기술할 수 없음)
	snuterm:metaAccess	
dc:format		Article의 형태정보
	dcterms:extent	Article의 Pagination등과 같은 크기 정보
snu:place		Article에서 발생하는 자료와의 연결정보 기술
	snuterm:placeType	서울대소장어부
	snuterm:requestDept	신청기관
dc:rights		Article의 저작권정보 및 이용제한정보를 기술

<표 12> 대학사료 메타데이터 셋

요소명	제1하위요소	대학기록물
dc:identifier		레코드 ID
dc:title(필)		
	snuterm:s:mainTitle(D)	타이틀(필검)
	dcterms:alternative(D)	기타 타이틀
dc:creator(필)		사료를 생산한 개인이나 단체
	vCard:FN(D)	생산자명
	vCard:N(D)	생산자 이름 색인어
	vCard:BDAY	
	vCard:ROLE	생산자의 역할 (저자, 편집자, 역자 등을 드롭다운으로 선택)
	vCard:EMAIL	
	vCard:ORG	생산자 소속기관
dc:contributor(필)		사료를 제공한 개인이나 단체
	vCard:FN(D)	제공자명
	vCard:N(D)	제공자 이름 색인어
	vCard:BDAY	
	vCard:ROLE	
	vCard:EMAIL	
	vCard:ORG	제공자 소속기관
dc:subject		주제분류는 DDC분류함
	dcterms:DDC	
	snuterm:s:UAC	
	snuterm:s:KEAC	
	snuterm:s:snuPress	
dc:description		
	dcterms:abstract	사료의 초록이나 내용 요약
	snuterm:s:key words	사료의 키워드(주요 인물명 혹은 주요어)
	snuterm:s:general	일반주기
dc:publisher		
	snuterm:s:publisherOriginal	실질출판사명
	snuterm:s:publisherOthers	출판사명 기타
dc:date		발행일(필검)
	snuterm:s:startDate(D)	
	snuterm:s:endDate(D)	
dc:type(검색제한자)		문서, 도서, 사진, 마이크로자료, 팜플렛, 낱장자료, 기타의 값중 하나를 선택한다
dc:format		
	dcterms:medium(D)	매체(필)
	dcterms:extent(D)	양, 페이지, 크기 등을 기술
dc:language(검색제한자)		언어(선)
dc:relation		
	snuterm:s:metaAccess(D)	대상사료가 다른 포맷으로 존재 시, 해당 포맷의 레코드ID
	dcterms:isPartOf(D)	work type이 Single 일때 해당 RG를 기술한 ID와의 연결정보를 취함.
	dcterms:hasPart(D)	RG에 대한 기술일 경우, 해당 Single의 ID와 연결하거나 혹은 직접 리스트를 입력할 수 있도록 함.
	snuterm:s:isCopy	Y 또는 N 값을 취한다
dc:coverage		
	dcterms:spatial(D)	지역(필)
	dcterms:temporal(D)	시대(선)
snu:workType		Collective/Single 유형정의
snu:place		위치정보
	snuterm:s:callNo	현재소장기관의 청구기호
	vra:formerRepository	이전소장기관
	vra:creationSite	제작기관
	snuterm:s:requestDept	신청기관

이렇게 기록사료의 특징을 나타낼 수 있도록 메타데이터 셋의 개발을 하였으나, 기록사료의 특이성과 기존 자료의 정리 방법상의 차이로 작업과정에 여러 가지 난점이 발견되었다. 다른 기존 자료들의 경우 dc:date에 자료의 출판년/편찬년을 기술하지만, 기록사료는 자료의 생성일을 매우 중시하기 때문에 별도의 필드를 구성해 주어야 하고, 원본의 생산기관과 자료의 보관기관이 다른 자료가 많아 이 필드 역시 별도로 구성해 주어야 하는데, 메타데이터의 개발 도중 문제 제기가 늦어 미처 필드 추가를 할 수 없었다. 또한, 기록사료의 소장 기관에 따라 별도의 구분이 필요한 사항이 많았으나 메타데이터 셋의 통일을 위해 일부러 끼워 맞추거나 특정 항목이 다른 용도로 쓰이는 등의 문제가 있었다.

3.2. 미술 메타데이터

미술 메타데이터는 작품 그룹과 개개 작품에 대한 메타데이터로 나뉜다. 대학도서관 자료의 특성상 졸업 작품집과 같은 학생들의 작품집이 자료의 많은 부분을 차지하므로, 개개의 작품 검색과 동시에 작품집이라는 그룹으로 상위 검색을 할 수 있도록 하였다. 미술 메타데이터 셋은 보다 상세한 검색 접근점을 마련하기 위하여 박물관과 미술관 등에서 쓰이는 메타를 차용하기도 하였는데, 메타데이터 구축 결과 오히려 상용성은 떨어지는 점도 발견되었다. 예를 들어, snuterm:style(작품의 스타일)이나 snuterm:period(작품의 시기) 같은 요소는 헬레니즘 스타일, 르네상스 시기 등을 기술하기 위한 것으로 현대의 학생이나 교수의 작품을 기록하는 데는 쓰이지 않는다. 또한 미술 자료의 경우, 자료는 존재하나 기록된 정보가 불충분하여 실제 메타데이터를 구축하는데 어려움이 많고, 검색 접근점 역시 다양하게 제시하지 못하는 문제점이 있다.

<표 13> 미술 작품 메타데이터 셋

요소명	제1 하위요소	정의
dc:identifier		작품에 부여된 고유한 번호로 시스템상의 고유한 Record ID
dc:title(필)		작품의 제목
	snuterm:s:mainTitle	작품의 제목
	dcterms:alternative	작품의 대등서명
dc:creator(필)		작자
	vCard:FN(D)(필)	작자의 공식명칭
	vCard:N(D)(필)	검색을 위한 다른 작자명 (전거를 활용한다면 전거로 대체가능)
	vCard:BDAY	
	vCard:ROLE(검색제한자)	작자의 역할 (사진사, 회화가, 판화가, 조각가 등등)
	vCard:EMAIL	작가의 -email
	vCard:ORG(검색제한자)	
	vCard:CLASS	학년
dc:subject(필)		작품의 주제
dc:description		작품에 대한 내용주기
dc:date(필)		
	dcterms:created	작품의 제작일
dc:type		
	snuterm:s:style(검색제한자)	작품의 스타일(렐레니즘 스타일 등등)
	snuterm:s:period	해당작품의 시기(르네상스 등등)
	snuterm:s:group	작품그룹
	snuterm:s:school(검색제한자)	작가가 속한 학파(후기인상주의, 야수파 등등)
	snuterm:s:dynasty	작품이 제작된 왕조 (조선왕조)
	snuterm:s:movement(검색제한자)	작품의 movement(예: 바우하우스 운동에 의해 만들어진 작품 등)
	vra:technique	미술작품에 사용된 미술적 기법(에칭기법을 이용한 판화 등등)
dc:format		
	dcterms:medium	미술작품이 표현 매체 (종이에 잉크로 그려진 그림은 medium : ink, support : paper임)
	dcterms:extent	미술작품의 크기, 부피 등의 양적 특성
	snuterm:s:support	
dc:language		작품의 창작에 사용된 언어
dc:relation		Work 와 관련된 자원들과 관계를 기술한다.
	dcterms:isPartOf	작품이 Single Work이면 해당되는 Collection을 기술하고 있는 메타데이터의 ID로 연결
	dcterms:hasPart	작품이 Collection(개개의 작품이 모여 전체 작품을 이룰때)일때, 해당되는 Single Work (개개의 작품)의 ID.
	snuterm:s:isContainedBy	이 미술작품의 이미지를 수록하고 있는 매체(작품집)의 ID (서지데이터 존재시, BID)
	dcterms:hasFormat	미술작품에 대한 다른 형태의 포맷이 존재할때 (예: 이미지파일) 해당 Media Object의 ID
	dcterms:isReferencedBy	이 미술작품을 참고한 다른 미술작품
	dcterms:references	이 미술작품이 참고한 다른 미술작품
dc:coverage		작품의 내용이 다루고 있는 범위 (시간적, 공간적)
	dcterms:spatial	작품의 내용이 다루고 있는 범위 (공간적)
	dcterms:temporal	작품의 내용이 다루고 있는 범위 (시간적)
dc:rights		작품의 저작권정보를 기술한다.
snu:workType		기술하는 작품의 유형이 Single인지 Collective인지를 기술
snu:originalTextTitle		미술작품에 사용된 비미술작품의 Title(구전 설화를 바탕으로 작가가 회화를 그렸다면, 해당 구전설화명칭)
snu:place		
	vra:currentSite	현재 작품이 있는곳
	vra:currentRepository	현재 작품이 소장된 기관명
	vra:formerRepository	이전에 작품을 소장하였던 기관명
	vra:creationSite	작품이 제작된 장소
	snuterm:s:requestDept	신청기관
	snuterm:s:displayType	전시구분
	snuterm:s:buildingType	
	snuterm:s:buildingSite	

<표 14> 미술 작품그룹 메타데이터 셋

요소명	제1하위요소	정 의
dc:identifier		자동적으로 생성되는 Container의 식별번호
dc:title(필)		
	snu:mainTitle	대표적인 서명 (예: 미술대학 졸업작품집 '97)
	snu:seriesTitle	시리즈 명(총서일경우, 총서명)
	snu:seriesNo	총서가 존재할경우, 총서번호를 아라비아 숫자로 기입
dc:creator(필)		기획 /제작/출판에 참가한 모든 사람들
	vCard:FN(D)	
	vCard:N(D)	
	vCard:BDAY	
	vCard:ROLE	
	vCard:EMAIL	
	vCard:ORG	
dc:subject(필)		
dc:description		주기사항
dc:publisher(필)		
	snuterm:publisherName	발행처명
	snuterm:nationality	발행지명
dc:date(필)		일시정보
	dcterms:issued(D)	발행일 및 미술작품전시기간
dc:type		작품구분
dc:format		
	dcterms:medium	물질적인 매체 (책, video 등등)
	dcterms:extent	시간, 양(min, page 등)
dc:language		사용된 언어
dc:relation		
	snuterm:containedWork	이 작품집, 도록에 수록된 작품
	snuterm:hasSession	
	snuterm:metaAccess	이 작품집 등이 MARC 데이터로 존재할때, 해당 BID를 기입
snu:edition		판차정보
snu:condition		상태 정도, 디지털화 될 때 물리적인 것의 상태 (상,중,하)
snu:place		
	snuterm:requestDept	신청기관

3.3. 의학

의학 메타데이터는 의학슬라이드 자료와 VOD 자료 등을 수업 자료로 활용할 수 있도록 부위명, 증세명 등의 의학 자료의 특성을 반영한 접근점을 두어 검색할 수 있도록 구성하였다. 의학메타만의 특수한 항목으로는 해당자료의 내용이 담고 있는 장기의 명칭을 영어 또는 한글로 입력하는 snuterms.organName(장기명)과 장기의 세부 부위 명칭을 입력하는 snuterms . regionName (부위명), 그리고 환자명을 입력하는 snuterms . anotherName (별칭)을 들 수 있다. 또한 주제분류를 DDC 대신 NLMC로 하였다.

<표 15> 의학 그룹 메타데이터 셋

요소명	제1하위요소	정의
dc:identifier		본 그룹을 식별할 수 있는 URL(자동생성)
dc:title(필)		
	snuterms:mainTitle	그룹의 고유제목으로 이는 이 그룹에 속한 개별 슬라이드의 description중 <강의제목>과 일치함
	dcterms:alternative	
dc:creator(필)		제1저자 또는 책임관리자에 대한 정보로 이는 이 그룹에 속한 개별 슬라이드의 Creator 정보와 일치함
	vCard:FN(D)(필)	제1저자 또는 책임관리자의 성명
	vCard:N	
	vCard:ORG(검색제한자)	제1저자 또는 책임관리자의 전공 또는 소속
	vCard:ROLE(검색제한자)	제1저자 또는 책임관리자의 슬라이드 생성에 있어서의 역할
	vCard:TITLE	제1저자 또는 책임관리자의 전공/소속에서의 직책
	vCard:EMAIL	제1저자 또는 책임관리자의 이메일주소
	vCard:TEL	제1저자 또는 책임관리자의 연락 가능한 전화번호
dc:subject(필)		그룹의 주제어: 검색을 위한 항목
	dcterms:MESH	
	NLMC	
dc:description		본 그룹에 대한 설명
dc:publisher(필)		슬라이드의 발행처로 모두 (서울의대)로 자동 표시됨
dc:date(필)		Date
	dcterms:created	
dc:format		자료의 형태: 모두 (slide)로 자동 표시됨
dc:relation		본 슬라이드와 관련된 다른 슬라이드와의 다양한 연계정보
	dcterms:isPartOf	현재의 자료가 다른 자료의 일부분으로서 설명될 때
	dcterms:hasPart	현재의 자료가 다른 여러 개의 슬라이드를 포함할 때
	dcterms:isReferencedBy	현재의 자료가 다른 자료에 의해서 참조되고 있을 때
	dcterms:references	현재의 자료가 다른 자료를 보완하여 설명하거나 참고 자료로 쓰일 때
snu:place		
	snuterms:requestDept	신청기관

<표 16 >의학 슬라이드 메타데이터 셋

요소명	제1하위요소	Definition
dc:title(필)		creator(저자)나 publisher(발행처)가 슬라이드에 부여한 제목
	snuterm:smainTitle	본 slide 한 장의 고유제목울 영어로 기입
	dcterms:alternative	이칭을 영어 또는 한글로 기입
dc:creator(필)		슬라이드의 제1저자 또는 책임관리자에 관한 정보
	vCard:FN(D)(필)	제1저자 또는 책임관리자의 성명
	vCard:ORG(검색제한자)	제1저자 또는 책임관리자의 전공 또는 소속
	vCard:ROLE(검색제한자)	제1저자 또는 책임관리자의 슬라이드 생성에 있어서의 역할
	vCard:TITLE	제1저자 또는 책임관리자의 전공/소속에서의 직책
	vCard:EMAIL	제1저자 또는 책임관리자의 이메일주소
	vCard:TEL	제1저자 또는 책임관리자의 연락 가능한 전화번호
dc:contributor		슬라이드 생성에 관계한 사람. 편집자, 번역자, 공저자 등
	vCard:FN(D)(필)	슬라이드 생성에 관여한 사람의 성명
	vCard:ORG(검색제한자)	슬라이드 생성에 관여한 사람의 전공 또는 소속
dc:subject(필)		슬라이드의 주제어로 검색을 위한 항목
dc:description		슬라이드에 대한 자세한 설명
	snuterm:sdeptName	슬라이드 생성에 관여한 전공교실명(2개 이내)
	snuterm:skeywords	강의제목 이외에 본 슬라이드 검색에 필요한 단어
	snuterm:slectuerTitle	슬라이드가 속하는 그룹의 강의제목
	snuterm:sorganName	슬라이드 내용 또는 사진이 담고 있는 장기의 명칭
	snuterm:sregionName	슬라이드 내용 또는 사진이 담고 있는 인체부위명칭 또는 장기의 세부 명칭
	snuterm:sanotherName	환자이름, 별칭, 병록번호, 관련번호 등 본 슬라이드 식별을 위해 필요한 항목
dc:publisher		슬라이드의 발행처로 모두 (서울의대)로 자동 표시됨
dc:date(필)		슬라이드의 생성일
	dcterms:created	슬라이드의 생성일로 YYYY-MM-DD 형식으로 기입함
dc:type		슬라이드화 된 원자료의 유형
dc:format		자료의 형태: 모두 (slide)로 자동 표시됨
dc:identifier		슬라이드를 식별할 수 있는 URL(자동생성)
dc:language(검색제한자)		슬라이드에서 사용된 언어
dc:relation		본 슬라이드와 관련된 다른 슬라이드와의 다양한 연계정보
	dcterms:isPartOf	현재의 자료가 다른 자료의 일부분으로서 설명될 때
	dcterms:hasPart	현재의 자료가 다른 여러 개의 슬라이드를 포함할 때
	dcterms:isReferencedBy	현재의 자료가 다른 자료에 의해서 참조되고 있을 때
	dcterms:references	현재의 자료가 다른 자료를 보완하여 설명하거나 참고 자료로 쓰일 때
dc:source		슬라이드의 출처가 되는 정보자원에 대한 정보
snu:place		
	snuterm:srequestDept	신청기관
dc:rights		작품의 저작권정보를 기술한다.

3.4. 학술지 메타데이터 셋

학술지와 기사의 메타데이터는 서울대학교 전자도서관 구축을 위한 메타데이터 셋 구성에 맞게 DC를 기반으로 전체적인 메타데이터 셋의 기준에 맞추어 서울대학교만의 한정 요소를 두었다.

<표 17> 학술지 메타데이터 셋

요소명	제1 하위요소	학술지
dc:title(필)		
	snuterm:s:mainTitle(D)	저널명
	snuterm:s:subTitle(D)	부서명
	dcterms:alternative(D)	대등서명
	snuterm:s:seriesTitle(D)	총서명
	snuterm:s:OtherTitle(D)	약어명
	snuterm:s:preceedingTitle(D)	변경전 저널명
	snuterm:s:succeedingTitle(D)	변경후 저널명
dc:creator(필)		
	vCard:FN(D)	발행자명
	vCard:N(D)	발행자명 색인어
	vCard:ROLE	발행자 역할
	vCard:UID	이용자 ID
	vCard:KEY	패스워드
dc:contributor		
	snuterm:s:sourceTitle(D)	학술지가 속한 DB 명
dc:identifier		
	snuterm:s:issn	학술지 ISSN
	snuterm:s:eIssn	전자저널의 EISSN
	snuterm:s:CODEN	CODEN 부호
dc:publisher		
	snuterm:s:publisherName(D)	출판사명
	snuterm:s:nationality	출판국명
dc:subject		학술지 주제
dc:source		원문링크정보
dc:date		
	dcterms:Issued	발행년도
	dcterms:valid	학술지의 유효기간
	dcterms:available	학술지의 이용할수 있게된 날짜(즉, 자판제공연도)
	snuterm:s:dateOfFirstPublication	최초 출판년도
	snuterm:s:dateOfPublication	최초 전자출판년
dc:type		01:Online Only, 02:Online+Paper 중 택일
dc:format		
	dcterms:medium	매체 Type
	dcterms:extent	가격
dc:language(검색제한자)		언어(선)
dc:relation		
	snuterm:s:metaAccess	술라즈 연동정보(제어번호)
snu:bibInfo		
	snuterm:s:frequency	간기
	snuterm:s:price	학술지 계약가격이나 개개의 학술지 가격
	snuterm:s:havingabstracts	있으면 Y, 없으면 N
	snuterm:s:havingindex	있으면 Y, 없으면 N
	snuterm:s:charged	유료일경우 Y, 무료일경우 N
	snuterm:s:isCopy	동시이용자수
	snuterm:s:dsourc	01:E-Journal, 02:WebDB
snu:place		
	vra:currentRepository	현재구독기관(분관코드등록)
	snuterm:s:callNo	현재소장기관의 청구기호
	snuterm:s:requestDept	신청기관(1000 : 코드값)

<표 18> 기사 메타데이터 셋

요소	제1하위요소	정의
dc:title(필)		
	snuterms:mainTitle(D)	기사명 (Main title)
	snuterms:subTitle(D)	부서명
	dcterms:alternative(D)	대등서명
	snuterms:otherTitle(D)	다른서명(Romanized된 서명 혹은 색인을 위한 다른 서명등등)
dc:creator(필)		기사의 저자관련정보
	vcards:FN(D)	저자명
	vcards:N(D)	검색을 위한 저자 색인어
	vcards:role	기사에서의 저자의 역할 (예: 저자, 공저자 등)
	vcards:eMAIL	저자의 e-mail
	vcards:ORG	저자의 소속기관
dc:identifier		기사번호
	snuterms:issn(D)	기사가 속한 학술지의 ISSN
	snuterms:eIssn(D)	기사가 속한 전자저널의 EISSN
	snuterms:recordID(D)	기사 레코드 ID
dc:subject(필)		해당 기사의 주제분야
dc:date(필)		
	dcterms:created(D)(X)	기사 저작일
	dcterms:issued(D)(필)	기사의 발행일
dc:publisher		출판사 관련 정보
	snuterms:publisherName(D)(필)	출판사명
	snuterms:nationality(검색)	출판국명
dc:source		
	snuterms:sourceTitle	기사가 속한 저널이나, 스크랩북 서명
	snuterms:PaperName(D)	저널 article일때 해당되지 않음, 신문스크랩의 경우, 해당 신문명 정보를 기술
	snuterms:IssueDate	신문스크랩의 경우, 해당 신문발행일 정보를 기술 YYYY-MM-DD형식으로 기입
	snuterms:IssuePage	신문스크랩의 경우 해당신문의 면수기재
dc:type	(검색)	언어
dc:relation	(검색제한자)	기사유형
	dcterms:isPartOf	권호정보 레코드의 ID
	dcterms:hasFormatOf	기사가 다양한 형태의 매체에 수록되어있을 경우, 형태별 자원의 연결정보(ID) 혹은 URI를 제공 (전자저널 등).
	dcterms:isReferencedBy	기사를 인용하고 있는 다른 자원들의 URI
	dcterms:References	이 기사가 참고하고 있는 참고문헌들의 URI
	snuterms:metaAccess	이 기사를 담고 있는 서지정보의 BID
dc:description		
	tableOfContents(D)	기사의 목차
	dcterms:abstracts(D)	기사의 초록문
snu:bibInfo	snuterms:keywords(D)(필)	기사의 키워드
		기사의 서지사항
	snuterms:frequency	기사가 속한 학술지의 간기
	snuterms:volume(검색제한자)	기사가 속한 학술지의 권
	snuterms:number(검색제한자)	기사가 속한 학술지의 호
	snuterms:part(검색제한자)	기사가 속한 학술지의 파트
dc:format	snuterms:total(검색제한자)	기사가 속한 학술지의 통권
	dcterms:extent(D)	기사의 시작페이지-마지막 페이지, 총페이지수

4. 문제점 및 향후 과제

이상에서 살펴본 바와 같이, 서울대학교 메타데이터 연구는 음악, 고문헌, 학술행사 분야를 1단계로 하였고, 대학사료, 미술, 기사, 의학, 학술지 분야의 연구로 확대 진행되었다. 그리고, 이렇게 서울대학교 전자도서관 메타데이터 셋을 개발하여 실제 전자도서관 시스템에 적용을 시키면서 예상과는 달리 쓸모없거나 추가적으로 필요한 여러 요소들이 수정되거나 삭제, 혹은 추가 되었다.

그러나, 각 세부 메타데이터에서 살펴보았듯이 여전히 문제점이 존재하고 있음을 알 수 있다. 메타데이터의 존재 목적 중의 하나인 다양한 검색이 접근점을 제시하기 위해 상세한 하위요소가 개발되었으나, 메타데이터 셋이 복잡해진 것에 비해 상용성이 떨어지는 경우가 발생한 것도 그 중 하나이다. 예를 들어, 학술행사의 경우, 자료에 따라 다층적인 구조를 표현하기 위해 만들어진 메타데이터가 실제 검색했을 때에는 오히려 이용자를 혼란시킬 수 있는 여지를 만들기도 한다. 가령 규모가 작은 학술행사는 세션이 존재하지 않고 단 1개의 발표 자료(Article)만을 가지는 경우가 있다. 만일 학술행사의 제목과 발표 자료의 제목까지 일치한다면, 검색 실행 후 간략서지에서 동일하게 보이는 2건의 검색 결과는 이용자의 혼란만 초래한다.

또 다른 문제점으로 서울대학교 메타데이터 연구가 독자적으로 이루어졌다는 것을 지적할 수 있다. 앞으로 다른 기관과의 메타데이터 상호 호환이 가능하도록 하기 위해서는, 의미 매핑 작업의 결과에 근거한 핵심 메타데이터 요소에 대한 RDF/XML 기반 방식의 메타데이터 표현 방법에 대한 연구 역시 이루어져야 한다.

마지막으로, 메타데이터 셋을 만드는 데까지 진행되었으나 결국 중단된 고문헌 메타데이터의 연구처럼 자료에 대한 연구가 선행되어야 하는 경우, 보다 많은 시간과 전문 인력을 투자하여 연구 사업을 수행해야 할 것이다.

5. 결론

서론에서 언급했듯이, 급증하는 디지털 정보 자료와 이를 구현하는 환경은 점점 다양해지고 급속하게 변하고 있다. 따라서, 전통적인 서지 자료와는 달리 다양한 특성을 지니는 디지털 정보 자원을 명확히 기술하고 이를 검색에서 이용하기 위하여, MARC의 단점을 보완하고 단순히 정보 자원의 전반적인 내용뿐만이 아니라, 구조적인 측면을 고려한 메타데이터의 연구는 당연한 결과라고 할 수 있을 것이다.

서울대학교 메타데이터 연구는 앞으로 국내 다른 대학에서 이루어질 전자도서관 구축과 이에 따른 메타데이터 연구에 기초 자료로서 활용될 수 있을 것이며, 나아가 후일 이용자들이 전자도서관을 통해 기관을 넘나들 수 있는 자유로운 검색의 바탕이 되기를 기대해 본다.

참 고 문 헌

최유미. 2002. “서울대학교 전자도서관 구축을 위한 메타데이터 연구”, 국립대학도서관보 제 20집

오삼균. 2002. “디지털 도서관 구축을 위한 XML 스키마 메타데이터 연구계획서”.
서울대학교 중앙도서관

김태수. 1998. “MARC 데이터의 메타데이터 변환 연구”, 국회도서관보 1998년 11/12호

이재윤. 1998. “메타데이터 환경과 과제”, 한국정보검색위원회(KISEC) 1998년 10월
연구발표회 발표논문

Ahmed, Kal et al. Professional XML Meta Data. Birmingham : Wrox Press., 25-60
p. Chap.2

Creating the Digital Library (1999, May)

<<http://www.dml.indiana.edu/overview/proposal.html>>

Digital Music Library Metadata Model

<<http://dml.indiana.edu/pdf/DML-metachart.pdf>>

IU Digital Music Library Data Model Specification V1

<<http://dml.indiana.edu/pdf/DML-DataModel-V1.pdf>>

Dear Sir,

I have the pleasure to acknowledge the receipt of your letter of the 15th inst.

in relation to the above matter and in reply to inform you that the same has been forwarded to the relevant authorities for their consideration.

I am sure that you will understand the need for a thorough and careful examination of all matters of this nature.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

I am sure that you will be satisfied with the outcome of the process.

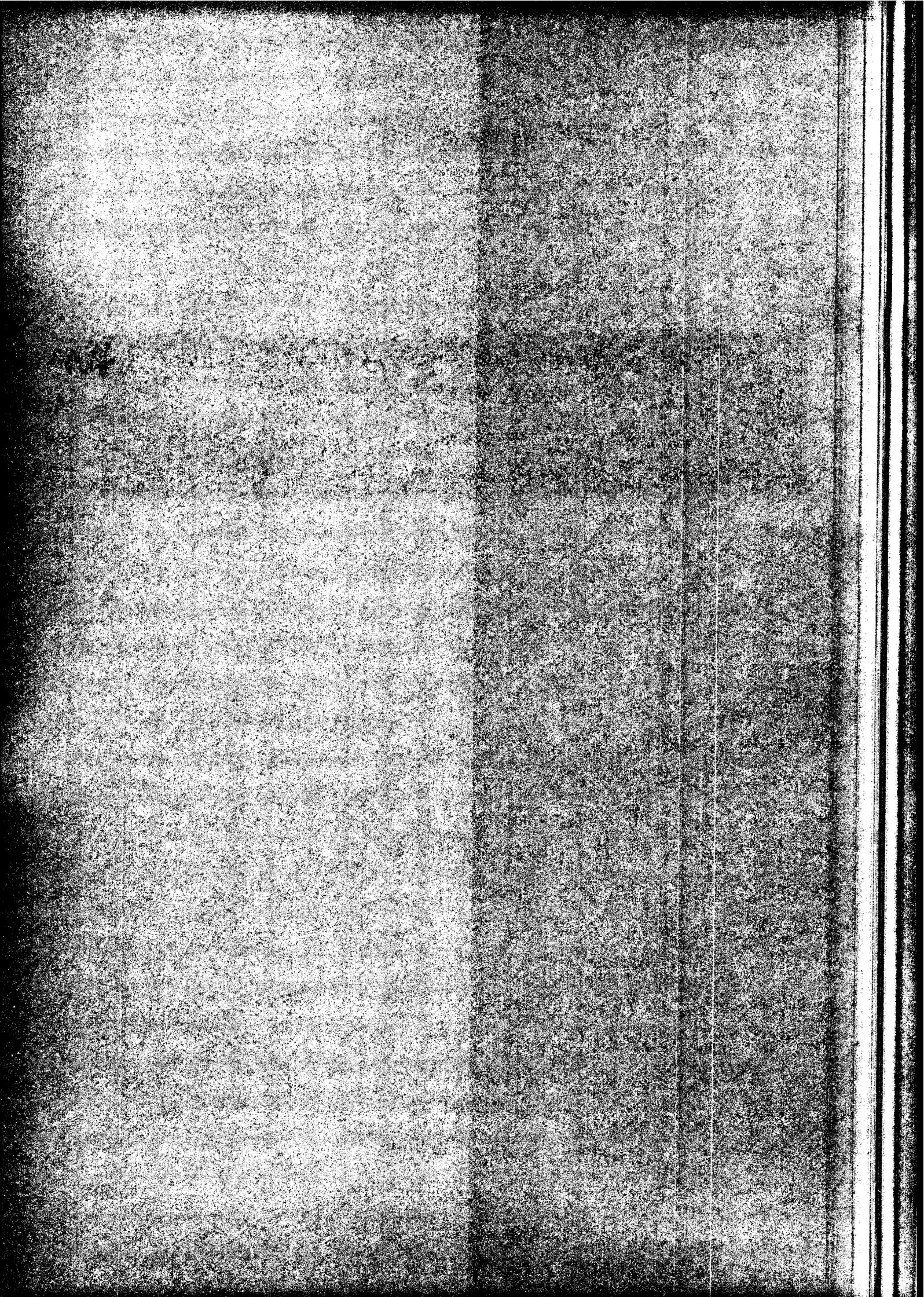
I am sure that you will be satisfied with the outcome of the process.

[사서번호 53]

2003. 5

황영숙

youngsak@snu.ac.kr



< 사례발표 5 >

『MARC & XML Schema』

서지데이터와 전자도서관 메타데이터 통합검색 운용

2003. 5.

황 영 숙
(youngsuk@snu.ac.kr)

차 례

1. 개요	1
2. 서울대학교 통합검색 사례	1
3. DB 통합 운영 사례	4
4. 문제점 해결방안 및 개선 사항	6

1. 개요

통합검색은 검색항목간의 통합검색에서부터 매체 및 단위시스템별, 신규 개발시스템과 기존 시스템 간의 통합검색 등이 있으며 더 나아가서는 학내 다른 시스템 간이나, 물리적으로 거리가 먼 시스템, 개발환경이 다른 타기관 시스템 등을 통합하여 하나의 쿼리를 이용해 정보자원에 접근하기 위한 접근점을 제공하는 것이라고 할 수 있다. 통합검색을 통해 이용자들은 여러번 검색하는 번거로움을 줄이고 원하는 정보를 매체 중심이 아닌 내용적 동일성을 함께 묶어 검색결과를 보여줌으로써 검색의 차원을 한 단계 상승시킨 것이라고 할 수 있다. 검색항목은 각 시스템별, 매체형태별로 다양하고 전문성 있는 검색어 중의 핵심 메타데이터요소(Core Metadata Element)를 중심으로 이루어진다.

2. 서울대학교 통합검색 사례

2.1. 통합검색 대상

- 도서정보 시스템
단행본, 학술지, 기사색인, 학위논문, 고문헌, 비도서, 웹학술정보
- 디지털정보 시스템
전시자료(미술전, 건축전), 음악연주, 전자저널, E-BOOK, 대학사료,
학술회의, 의학, 개인학술연구물, 강의자료
- Legacy 시스템
학내 학술정보기관의 시스템(규장각, 박물관, 전문센터 등)

2.2. 통합검색 개발 시 중점 사항

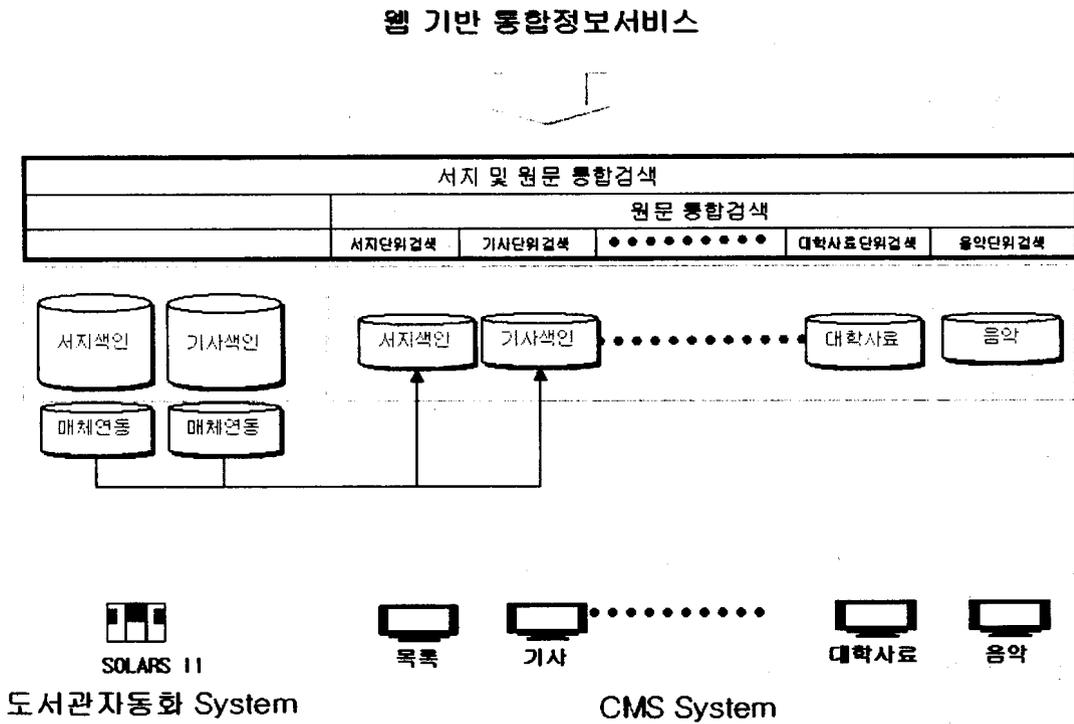
2.2.1. 검색

- 기존시스템의 IR 엔진(BRS)을 수용
- 디지털자료와 서지정보 및 Legacy 시스템 통합
- 동일한 자료가 여러건 단위시스템에 존재하는 경우 중복자료 연동
- 통합검색시 응답시간 최소화
- 각 단위시스템별, 자료형태별로 다양하고 적합한 검색항목지원
- 통합검색시 단위시스템 제한검색기능 제공
- 한번의 로그인으로 도서관과 전자도서관 홈페이지 자유롭게이용

2.2.2. 콘텐츠 관리 시스템(CMS)

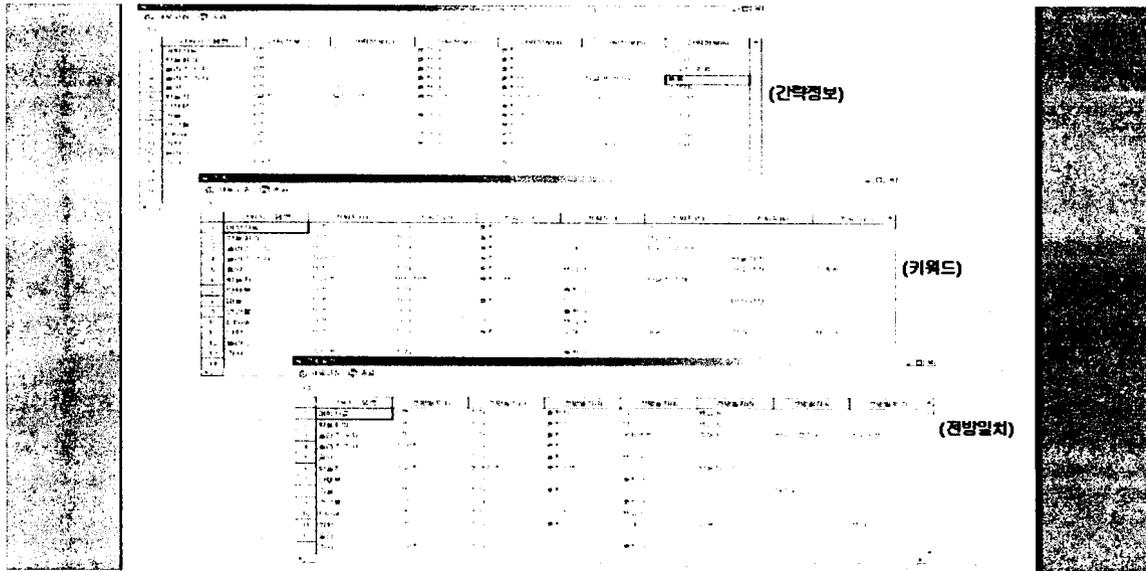
- 콘텐츠별로 신규구축되는 다양한 종류의 단위시스템을 통합 적용하는 시스템을 구성
- 관리자가 검색대상, 제한검색 및 간략정보 출력 항목 설정 기능
- 음악자료(IFLA Model) 등 구조 메타데이터 연계
- 단위시스템별 색인 운영
- 서지자료(MARC) 매체 관리 기능 제공

2.3. 시스템 구성도



2.4. 관리 메타데이터

솔라스 서지정보에 매체를 연동하기 위해 솔라스 자료를 검색 후 매체를 연동 하거나 메타데이터를 신규 입력 후 매체연동, 으로 이미지나 동영상을 구축한다. 또한 매체를 관리하기 위한 관리정보 입력과, 저작권 보호를 위한 저작권 정보를 입력한다.



2.5. 홈페이지 통합검색

서울대학교 전자도서관 login

홈 | English | 사이트맵 | 서울대학교 | 서울대학교 중앙도서관

전자도서관소개 DL Contents 전자저널 개인학술연구실 My DL 도서관서비스 게시판

간략검색 | 도움말

우측열단
 키워드
 완전일치

AND
 OR
 NOT

AND
 OR
 NOT

다국어 입력

전체
 단행본
 연감
 학회논문
 기사
 학술잡지
 기록자료

E-Journals
 과학자료
 미술작품
 음악자료
 개인학술연구물
 E-Book

비도서
 웹문서

전체
 원문
 PDF
 Audio
 스트림
 음악
 해설
 비문

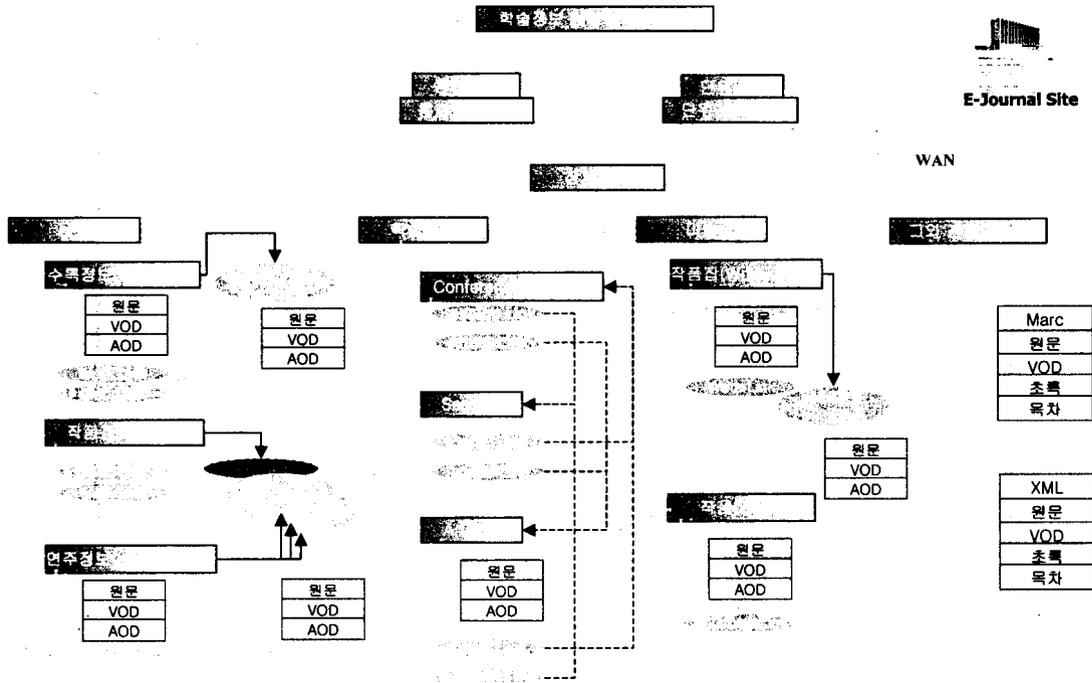
참고문헌
 비로그인
 E-Book

연대 : 전체 발행년 : ~

해당 총력 건수 : 15 최대 검색 건수 : 전체

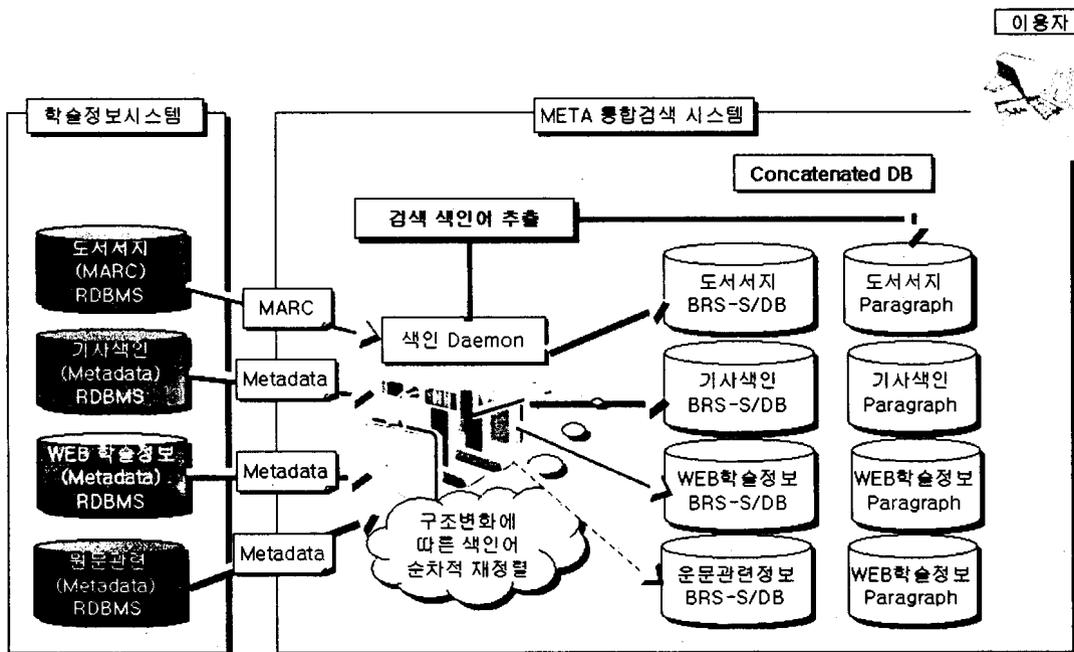
[검색](#)
[검색 History](#)
[다시쓰기](#)

검색 서비스 순서도

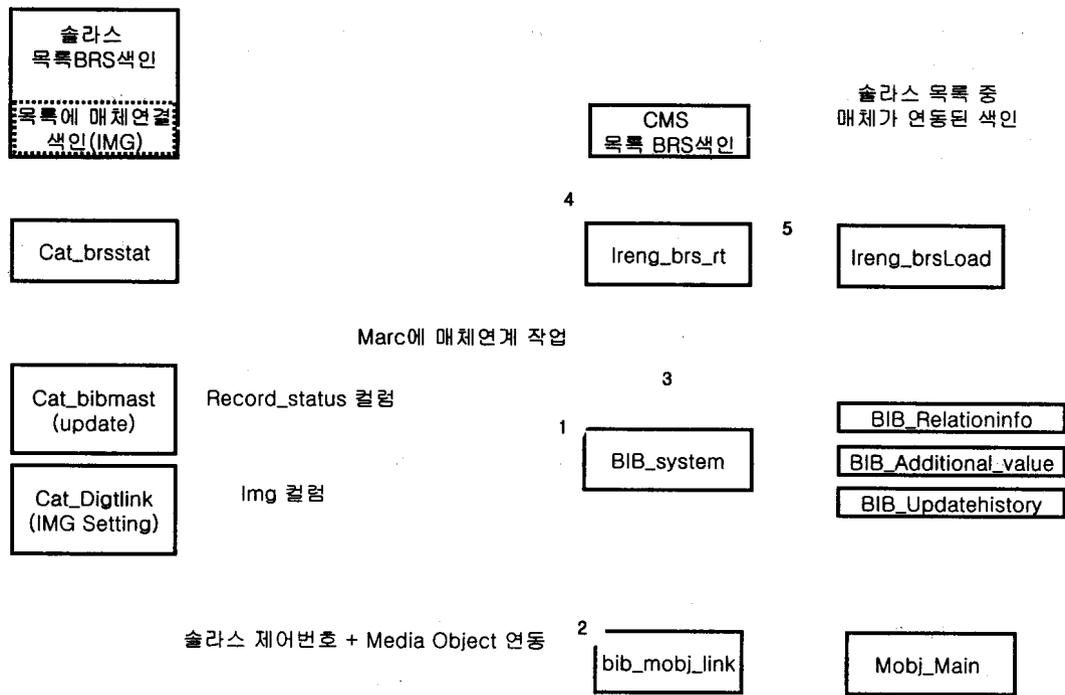


3. DB 통합 운영 사례

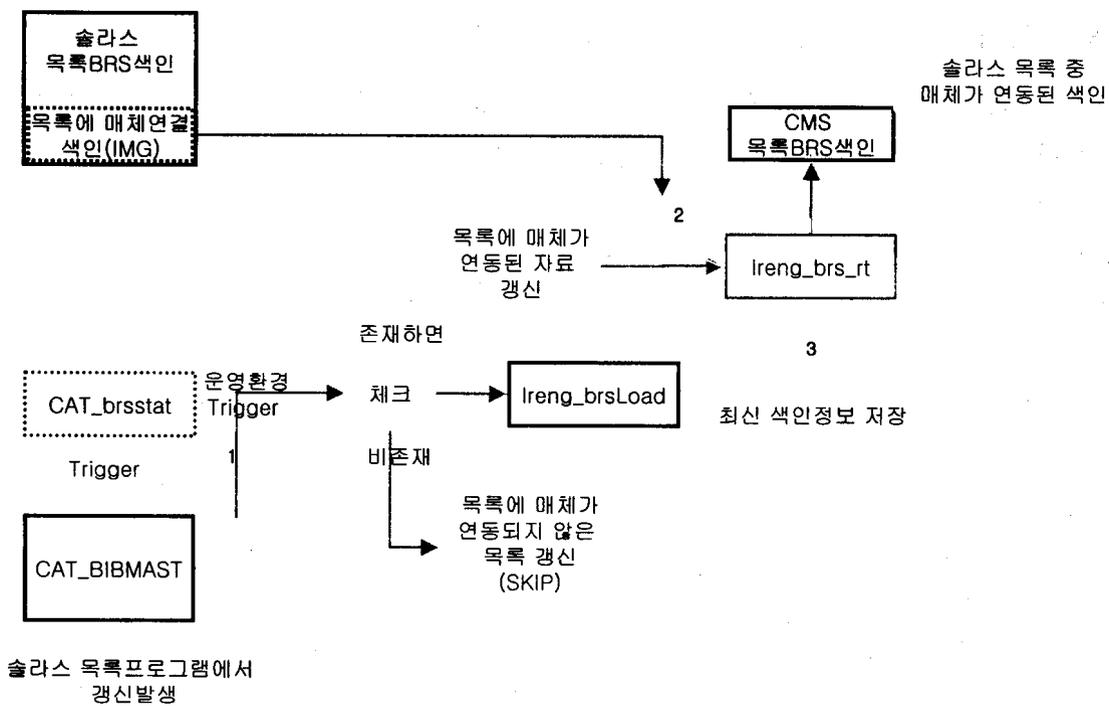
3.1. 검색을 위한 색인 운영방안



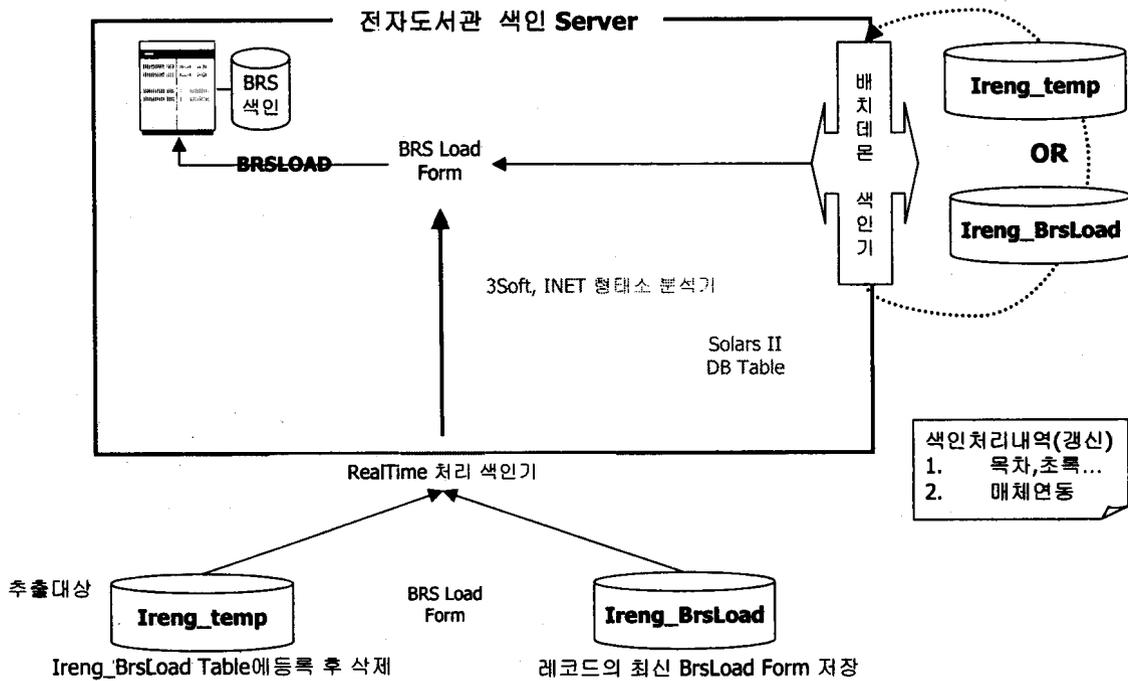
3.2. 솔라스 목록 시스템에 매체 연계방안



3.3. 솔라스 색인과 통일된 색인 유지 방안



3.4. 색인 추출 방식



4. 문제점 해결방안 및 개선 사항

4.1. 통합검색에서 제한검색이 어려운 단위시스템

- 고문헌, 연구보고서 자료들은 서지정보의 MARC 포맷 단행본으로 구축되어 있어 원문 검색과 달리 단행본 자료에 포함되어 검색된다.
- KORMARC 고도서 포맷 등이 개발되기 이전에 고도서 자료를 단행본 포맷으로 구축 하였으므로 일괄 수정 등이 고려되어야 할 부분이다.
- 연구보고서 자료의 경우 대학에서 생산되는 중요한 보고 자료로서 단행본이 아닌 자료로 구분하고자 하지만 입력 형태인 MARC 단행본 포맷에 입력되어 있으므로 통합검색에서는 제한검색을 할 수 없고, 단행본으로 통합검색 되어 학위논문과 같이 마크 유형을 별도 추가하여 수정 보완 하여야 한다.
- 비도서 자료 (단품CD 등)는 전문구축(VOD, AOD)을 하였지만 각 전자도서관 단위시

시스템과 레코드 유형이 동일하지 않아 비도서 에서만 검색 된다. 향후 비도서 자료와 전자도서관 단위시스템 매핑을 통해 수정 보완이 되어야 한다.

4.2. 목록시스템과 콘텐츠관리시스템(CMS) 운영방안

- 목록시스템이 소장 정보 중심이라면 CMS는 콘텐츠 관리 시스템으로서 매체관리 중심이라고 할 수 있다.
- 이원화 되어 있는 두 시스템은 각각 특성은 강하면서도 업무의 중복성이 있어 사용상에 불편함을 느끼게 한다. 목록시스템에서 서지정보를 입력하고 다시 CMS에서 매체등록을 하거나 또는 매체등록을 하다가 서지사항을 수정하기 위해서는 목록시스템으로 이동 등이 필요하게 되어 있어 향후 목록시스템에서는 매체를 연동할 수 있고 CMS에서도 목록을 수정 할 수 있는 인터페이스가 제공해야 할 것이다.
- 매체가 연동된 자료인 경우에도 불구하고 서지만 삭제가 가능 하여 서지가 연결되지 않은 원문이 서비스 되는 경우가 있다. 삭제자료에 대해서는 연동 자료가 있는 경우 삭제가 되지 않도록 처리해야 할 것이다.
- 목록시스템은 MARC에 익숙한 사서들이 사용하는 시스템이지만, CMS는 학내 학술자원을 통합관리 하는 시스템인 만큼 사용자들이 다양하고 메타데이터 항목들이 전문화 되어 있어 오히려 사서들이 입력하기에는 한계가 있다. 따라서 단과대학 교수 및 연구자들이 직접 콘텐츠에 대한 메타데이터를 입력하거나 수정하여야 하는 경우가 많은 점을 고려하여 두 시스템의 중복성을 최대한 배제하고 전문성을 강조하는 시스템으로 발전되어야 한다고 본다.

4.3. IR 검색(BRS) 엔진

- 빠른 검색 속도와 다양한 연산까지도 최소이용시간을 제시하지만 색인 생성에 대한 병렬작업이 불가능함으로 작업 후 즉시 확인이 불가능하고 Online 환경에 최적화 되지 않아 색인 중 일부 생성이 안 되는 현상들이 발생되어 주단위 일괄 색인 생성 등으로 보완하고 있다.
- 관리환경(목록)시스템에서는 생성 후 즉시 확인되는 경우가 많음으로 향후 DBMS와 병행방법도 고려 사항이다.
- 논리적 DB 색인 생성이 가능하여 단위시스템별 색인 생성 후 통합이 가능하다.

4.4. 전거통제

- IR 검색엔진에서는 다음과 같이 4가지 타입의 색인을 갖고 있다
 - Master 데이터 : 자료 관리를 위해서 만든 메터데이터 원본
 - Display용 데이터: 간략정보등 보여주기 위해서 추출한 정보
 - Sort용 데이터: 소트를 위해 정규화 된 데이터
 - 검색용 데이터: Master에서 추출하여 이용자의 검색 데이터를 정규화한 데이터
- IR 검색엔진을 사용할 경우 DBMS와 같이 현재로는 대표표목 개념을 도입하기가 어렵다. 각 레코드 마다 검색용 데이터를 포함하고 있기 때문에 실상은 무전거와 같은 개념이고 현재는 전거레코드 수정시 각 레코드를 수정할 수 없는 상황이다
- XML Schema 기반인 전자도서관에서는 각 레코드 마다 반복적으로 읽기 항목을 색인으로 입력했다. (전자 자료의 희소성으로 인해 전거통제 도입의 무의미로 판단)
- 향후 IR 검색엔진 도입에 맞는 전거통제 재검토가 필요하고, 전자도서관 XML Schema 자료에도 절충안이 필요하다.

4.5. 색인에 대한 형태소 분석 방식

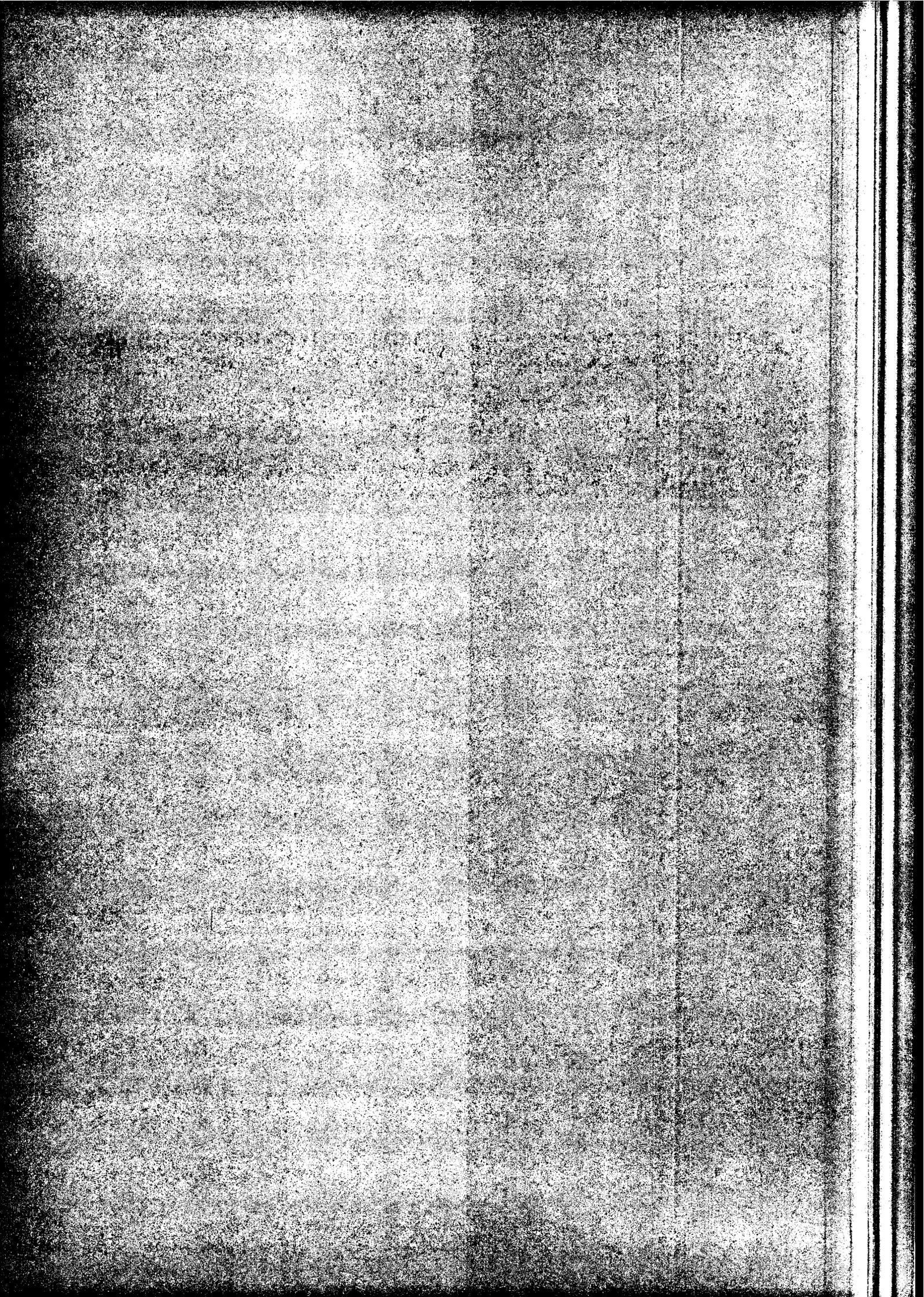
- IR검색엔진은 형태소분석방식과 앵그램방식 두 가지가 있다 그중 서울대학교에서는 형태소분석방식을 도입했다.
- 96년에 도입한 3Soft사 제품을 사용하고 있지만 단어의 최소단위로 색인을 만들고 있어 도서관에 맞게 수정보완하고 있다.
 - ex) 서울대학교중앙도서관 =>서울, 대학교, 중앙, 도서관
 - [저자] 이용성 => 이용 (형용사처리)
- 3Soft사 형태소 분석기를 보완해서 서울대학교 자체 형태소분석기를 개발했다(현재 8,600 단어 등록) 복합명사를 인정해주고, 신조어(처음 들어오는 문자)를 관리할 수 있다.
 - ex) 프로그래머 =>프로, 게이머, 프로그래머
 - 이용성 =>이용, 이용성

【서명표 8】

2003. 5

임영희

chamy@snu.ac.kr



< 사례발표 6 >

서울대학교전자도서관
전자저널, Web DB 검색관련 시스템 구현 사례 발표

2003. 5.

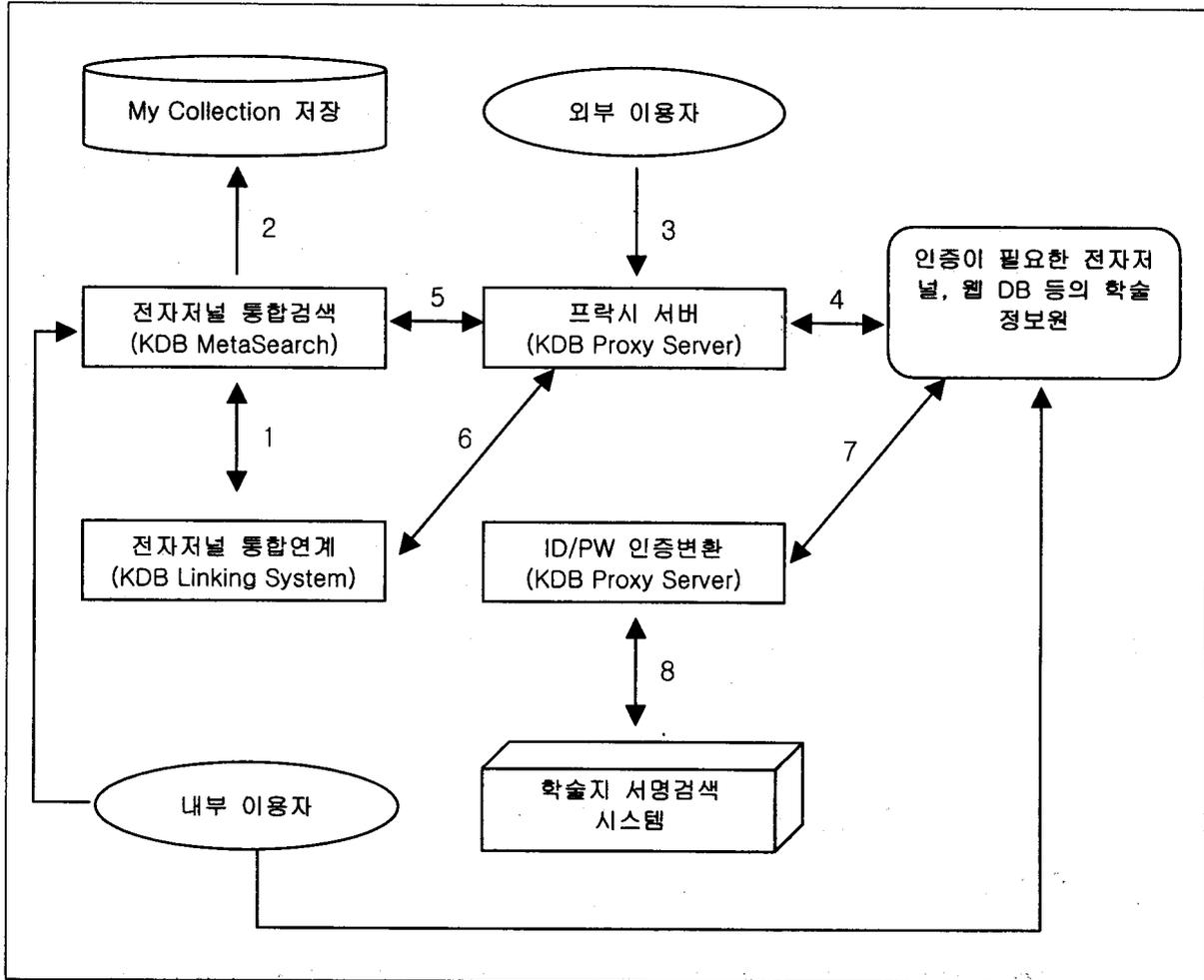
임 영 희
(charmmy@snu.ac.kr)

차 례

1. 시스템 개요	1
1.1. 시스템구성	1
2. 서울대 전자저널 웹디비 검색 시스템 구축 사례	2
2.1. 웹디비 통합검색	2
2.2. 프락시 서버	12
2.3 ID/PWD 인증변환시스템	18
2.4 시스SDI 서비스.....	19

1. 시스템 개요

1.1. 전자저널 통합검색 시스템 구성



1. 전자저널 통합검색 시스템, 전자저널 통합연계 시스템간 연동을 통해 전자저널 통합검색 결과에 대한 연계정보 출력
2. 전자저널 통합검색 결과에 대한 My Collection 저장
3. 외부이용자의 전자저널, 웹 DB 등 학술정보 이용을 위한 프락시 서버 이용
4. 인증이 필요한 학술정보에 대한 프락시 서버 등록 및 이용
5. 전자저널 통합검색 결과 이용시 인증이 필요한 학술정보원일 경우 프락시 서버 이용
6. 전자저널 통합연계 정보 이용시 인증이 필요한 학술정보원일 경우 프락시 서버 이용
7. 학술정보원 이용시 ID/PW에 의한 인증이 필요한 전자저널, 웹 DB에 대하여 자동 접속
8. 학술지 서명검색 시스템 이용시 ID/PW에 의한 인증이 필요한 전자저널에 대하여 자동 접속

2. 서울대학교 구축 사례

2.1. 웹디비 통합검색

2.1.1. 시스템 소개

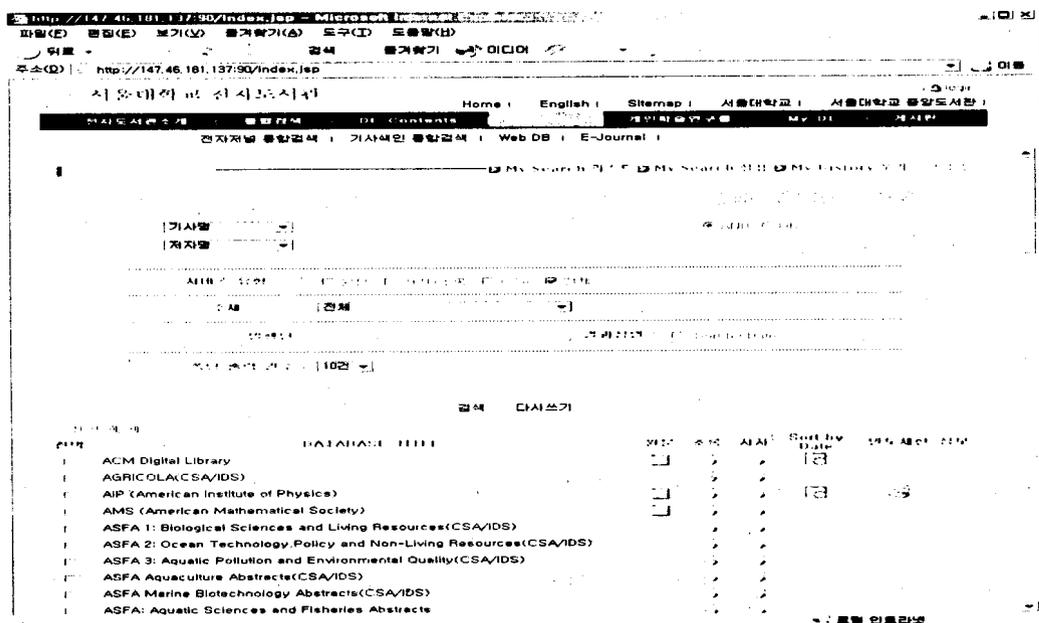
2.1.1.1. 시스템 필요성

전자저널 통합검색 시스템은 서울대학교 중앙도서관에서 라이선스 계약에 의해 구독중인 국내외 전자저널, 웹 학술 DB에 수록되어 있는 논문단위 기사에 대한 통합 검색을 목적으로 한다. 이는 다양하고 전문화된 Web-DB, 전자저널의 홍수 속에서 실제 자신이 원하는 기사나 정보가 어느DB에 속해있는지를 알 수 없는 이용자들이 각 DB로 접속하여 건건이 검색해야 하는 어려움을 해결해 주기 위함이다. 즉 적절한 주제분야의 웹디비 전자저널을 통합하여 이용자가 원하는 검색결과를 한번에 얻을 수 있는 이점을 제공한다.

2.1.1.2. 시스템 특징

국내외 전자저널 출판사, 웹 학술 DB 제공사별로 분산된 시스템에 구축되어 있는 기사 데이터를 이용자가 입력한 검색어를 각 시스템별 표준 검색어로 번역하여 쿼리를 전송, 각 사이트에서 검색된 데이터를 분석하고 분석된 데이터의 포맷을 변환하여 서울대 중앙도서관에서 지정한 통일된 포맷으로 검색 결과를 제공한다.

2.1.2. 서울대 웹디비 통합검색 주요 기능 및 특징



<초기화면>

◎ 검색기능

키워드, 기사명, 저자명, 저널명을 의 4개자 검색 필드를 이용하여 검색을 수행할 수 있으며 AND, OR 연산자의 사용이 가능하다

◎ 서비스 유형제한

원문, 서지+초록, CSA, 전체 중 이용자가 원하는 검색 서비스 유형을 지정한다.

◎ 주제별 DB 선택

각 Database별로 서울대도서관에서 자체적으로 만든 45여개의 특정 주제분류를 등록하여 이용자가 원하는 주제분야 Database만을 선택하여 검색할 수 있도록 지정한다.

◎ 발행년 제한 검색

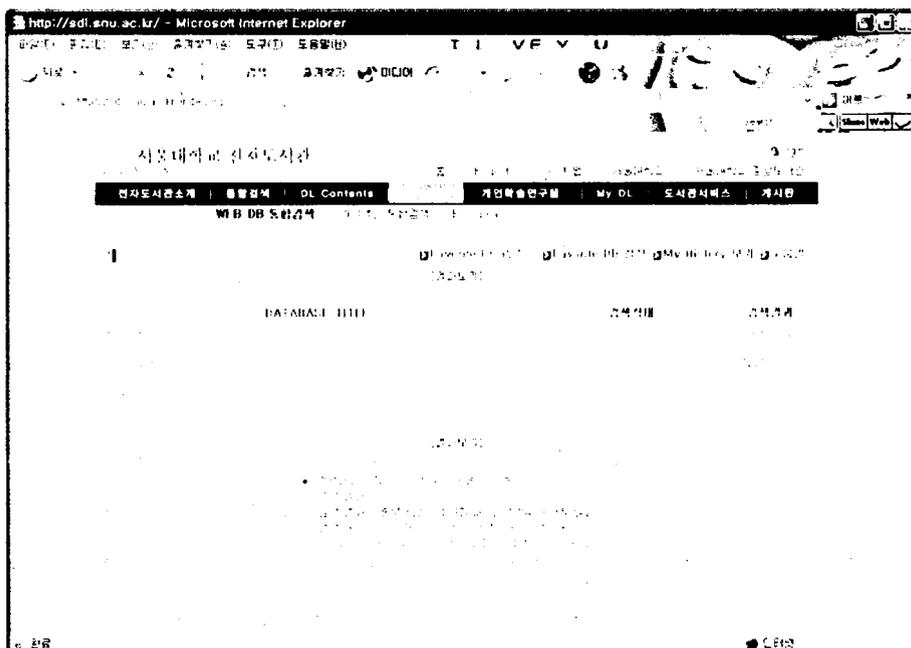
검색 대상 DB중 출판년도 제한에 의한 검색이 지원될 경우 이 옵션을 이용하여 검색 결과를 제한한다. (단, 출판년도 제한이 지원되지 않는 DB와 중복 선택할 경우 이 옵션은 적용되지않는다.)

◎ 결과정렬 (Sort By Date)

검색 대상 DB중 Sort By Date 제한에 의한 검색이 지원될 경우 이 옵션을 이용하여 검색 결과를 제한한다. (단, Sort By Date 제한이 지원되지 않는 DB와 중복 선택할 경우 이 옵션은 적용되지 않는다.)

◎ 검색 상태 보기 기능

다수의 DB를 선택한 후 선택된 DB의 검색 진행을 확인할 수 있다.



◎ 검색결과 간략목록

◎ 검색결과를 가지는 DB선택

* 선택된 DB에 대한 검색이 진행중인 경우

* 선택된 DB 모두에 대한 검색이 완료된 경우

◎ 저장/출력

간략검색 결과중 저장하고자 하는 정보를 선택하여 이용자가 원하는 형태로 저장 또는 출력하는 기능으로 서울대학의 경우 이메일 전송 기능은 물론, Text파일 형태 뿐만 아니라, Bibliography S/W 포맷을 제공함으로써 이용자는 Procite, Refwork 등의 인용문헌 시스템 이용포맷으로 저장이 가능하다.

◎ 검색결과 상세정보

The screenshot shows a web browser window with the following content:

- Source:** Phys. Med. Biol., 46 (April 2001), 1263-1295
- Title:** Receptor dose in digital fluorography: a comparison between theory and practice
- Author:** N W Marshall, C J Kotre, K J Robson and A R Lecomber
Regional Medical Physics Department Newcastle General Hospital, Westgate Road, Newcastle upon Tyne, NE4 6BE, UK
- Publish Information:** Received 17 August 2000 in final form 4 January 2001
- Abstract:** A method of identifying the dose per image when quantum noise no longer dominates the image statistics is presented as a first step towards quantitative optimization in native and subtracted digital fluorography. The method is based on measurements of threshold contrast over a range of receptor doses and the application of a simple model of the threshold contrast detection task to estimate the magnitude of system noise sources. The point at which system and quantum noise sources are equal in magnitude is proposed as the practical upper limit for dose per image. The method is applied to a typical digital fluorography system and the results are placed into the context of the range of dose per image values found from a regional survey of digital fluorography units. While there is broad agreement between the dose per image values in the survey with values predicted from the experimental method, the considerable spread in survey doses suggests there are instances where the use of a high dose per image is unjustified.
- DOI:** URL: stacks.iop.org/0031-9155/46/1263
DOI: 10.1088/0031-9155/46/1263
Epub: 2001-01-16 14:51:00
- Link:**
 - Acrobat PDF (361 KB)
 - References - including HyperCite® links
 - Information about filing cabinet
 - Information about Phys. Med. Biol.
- Extended Linking:**
 - Link to Issue List of Institute of Physics
 - Link to Issue List of EBSCO EJS
 - 전자저널 - Ulrichs
 - Web Database - Web Of Science

◎ 전자정보 통합연계

서울대학교 중앙도서관 및 분관, 학내 연구소에서 구독하고 있는 다양한 전자저널, 웹 학술DB에서 제공하는 원문/초록/목차정보/권호정보를 도서관 전자도서관 시스템에서 자동으로 연계할 수 있는 통합 연계 시스템이다

◎ 연계정보 이용

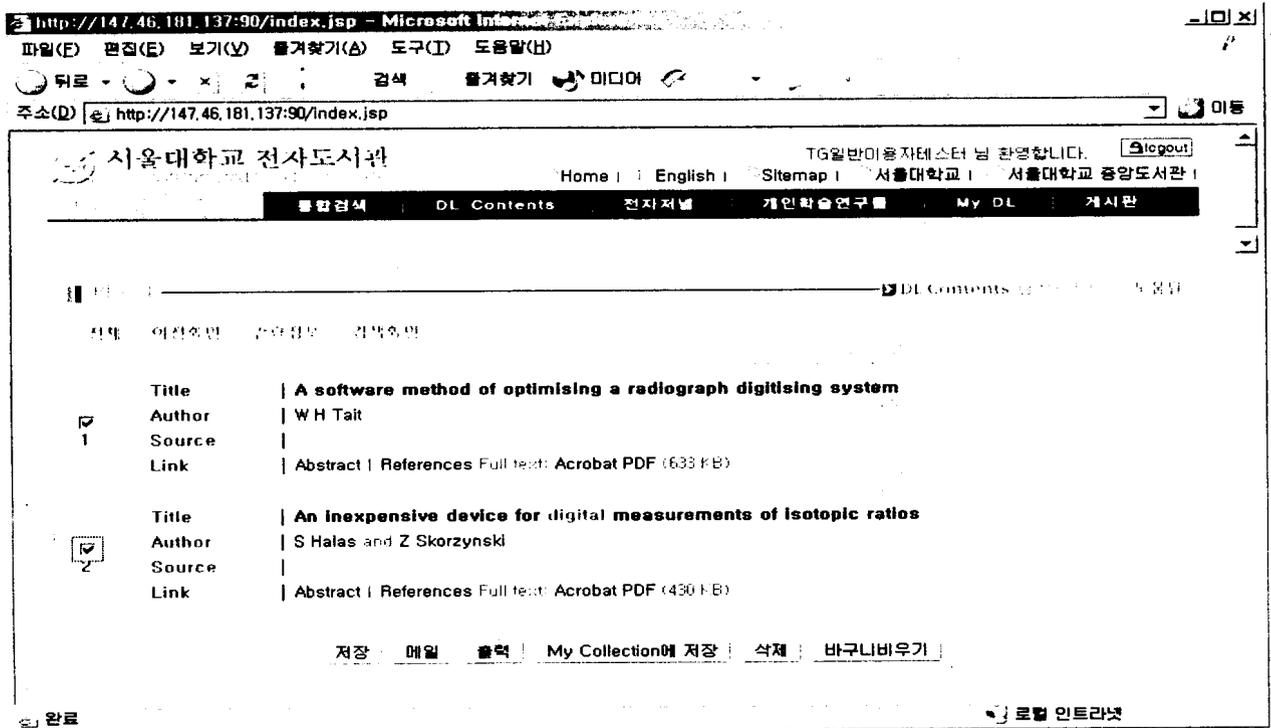
전자저널 통합검색 검색결과 상세정보 중 ISSN 정보가 출력되는 경우 통합연계 시스템에서는 ISSN 정보를 검색하여 추가적인 연계정보를 이용할 수 있도록 출력한다.

The screenshot shows the 'Extended Linking' section with the following links:

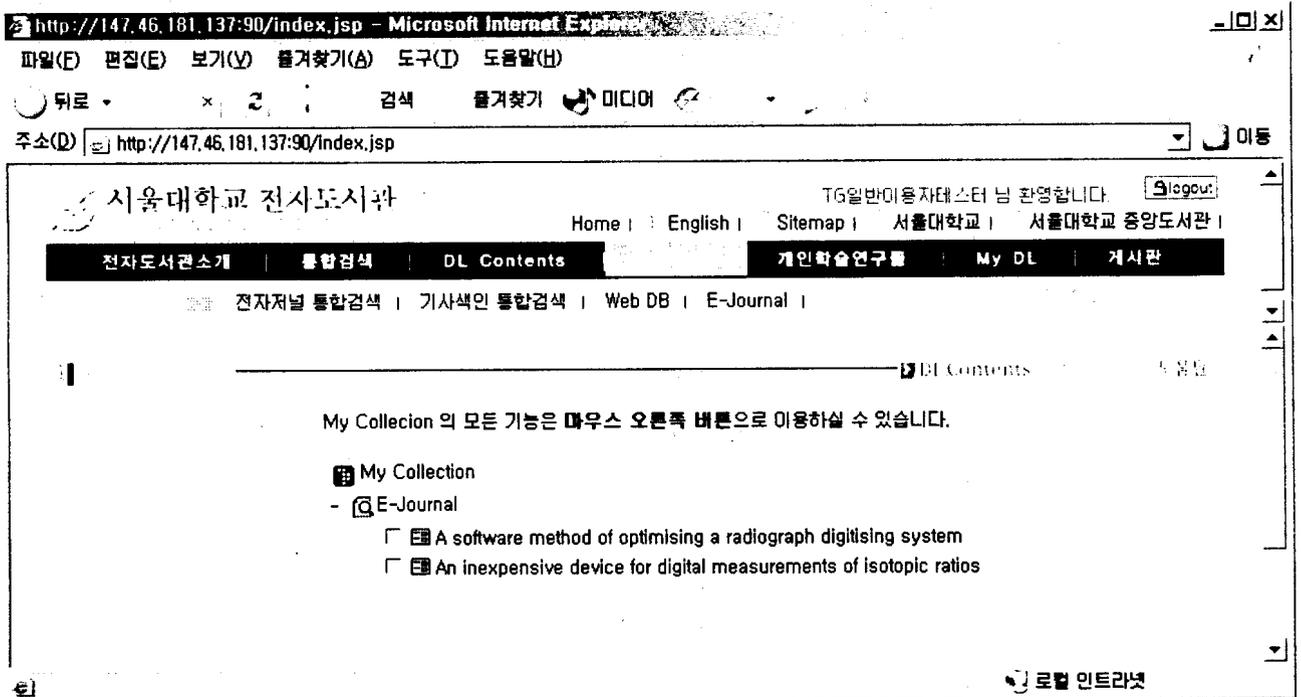
- Information about Smart Mater. Struct.
- Link to Issue List of Institute of Physics
- Link to Issue List of EBSCO EJS
- 전자저널 - Ulrichs
- Web Database - Web Of Science

◎ 바구니 기능

간략검색 결과 중 저장하고자 하는 정보를 선택하여 바구니로 전송후 개인 PC로 텍스트 저장 출력 혹은 My collection에 담기가 가능하다.



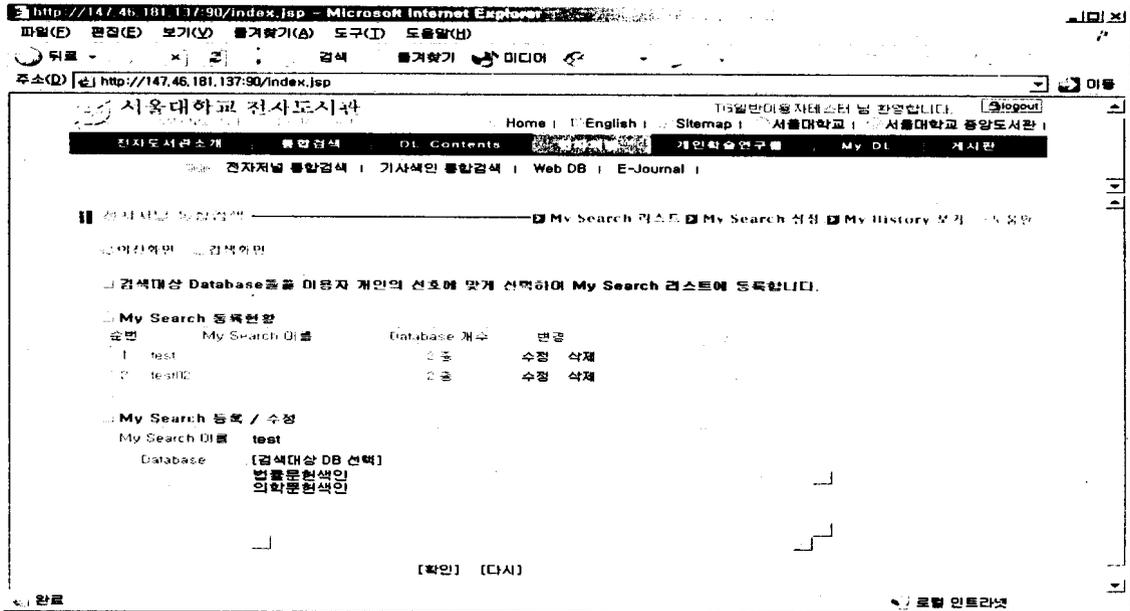
◎ My collection



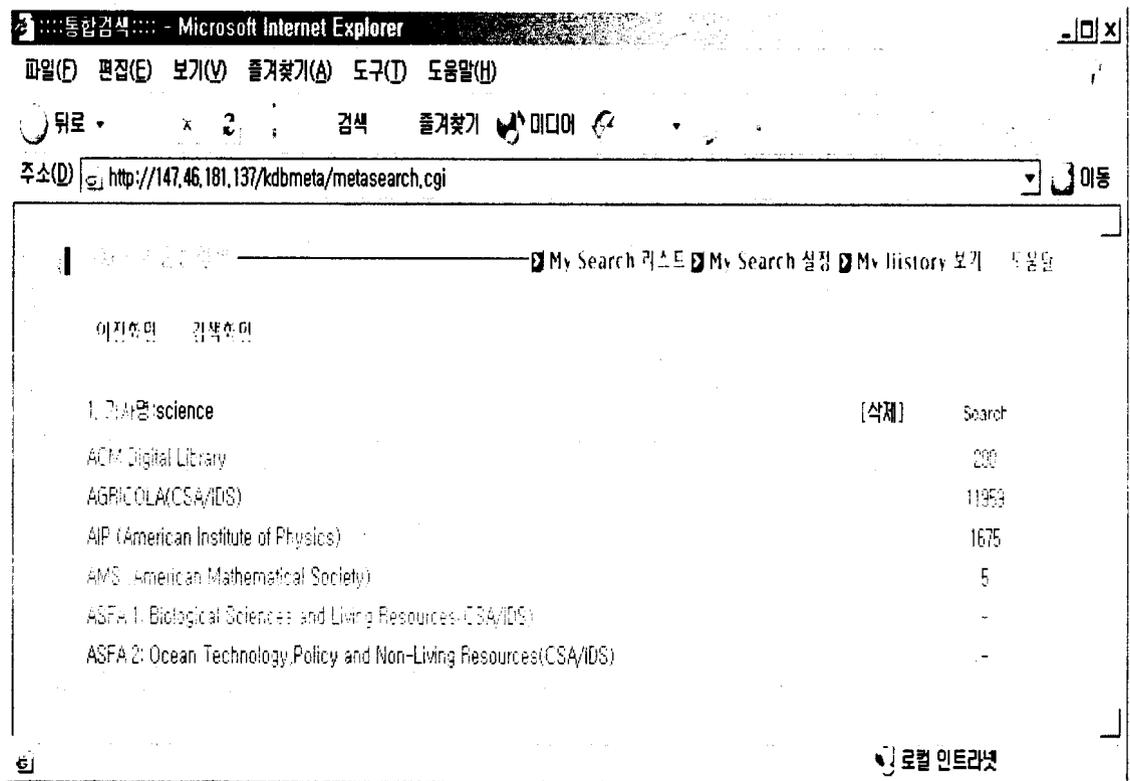
◎ My Search

My Search 기능은 전자저널 통합검색 시스템을 이용자별로 적용하여, 이용자가 자주 검색하는 DB를 등록하여 원하는 검색 결과를 보다 빠르게 얻을 수 있도록 한다.

< My Search 설정화면 >

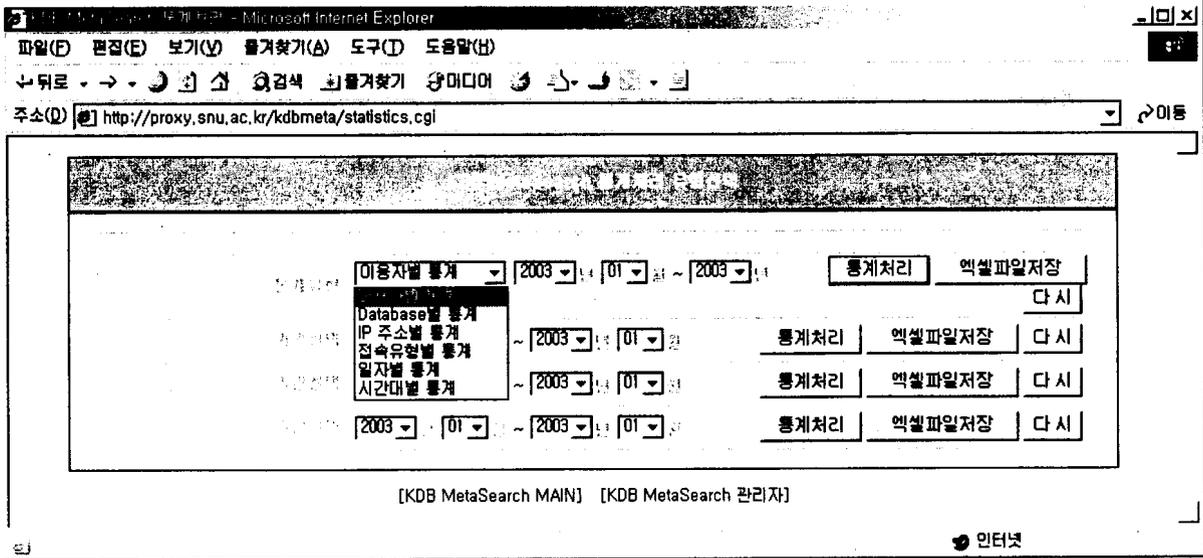


◎ My History 보기



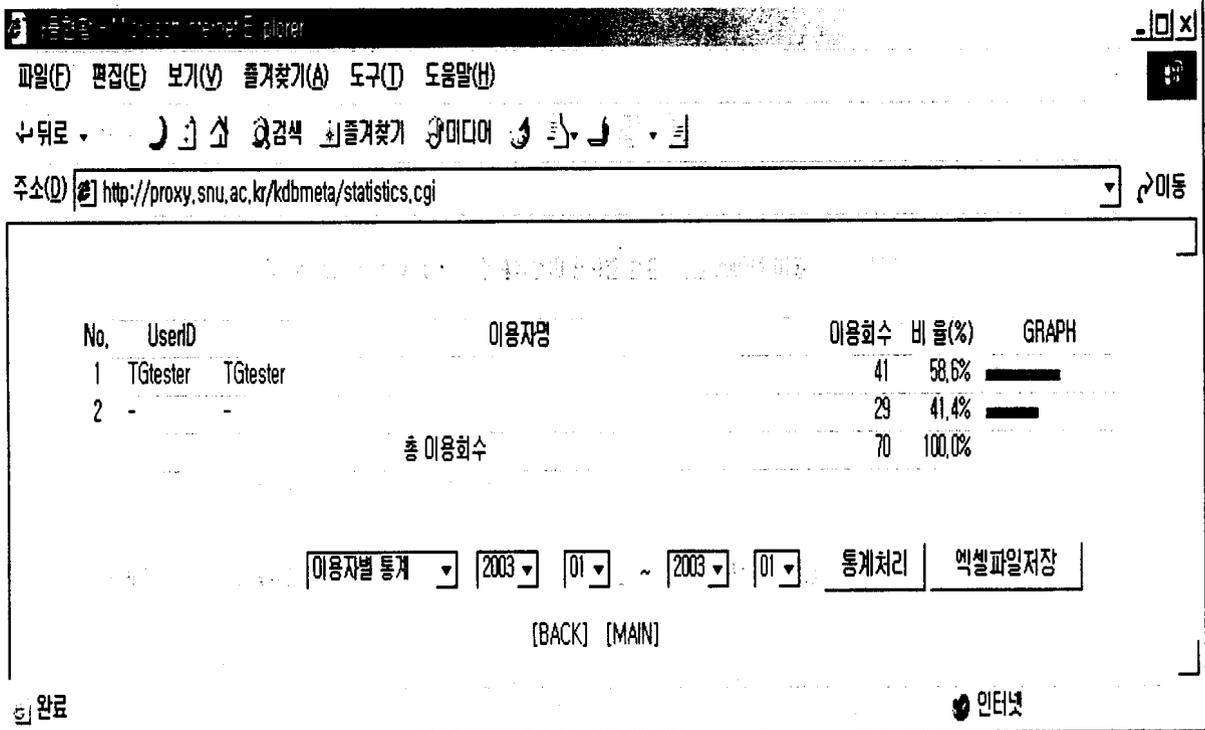
◎ 이용자 통계

메타서치 이용자 통계는 메타서치를 이용한 이용자별 통계, 메타서치를 통해 검색한 Database 통계, IP 주소별 통계, 접속 유형별 통계, 일자별 통계, 시간대별 통계, 월 별 이용현황, 웹 브라우저별 이용현황, OS별 이용현황을 처리하여 관리자가 메타서치의 활용도 및 응용도를 확인할 수 있다.

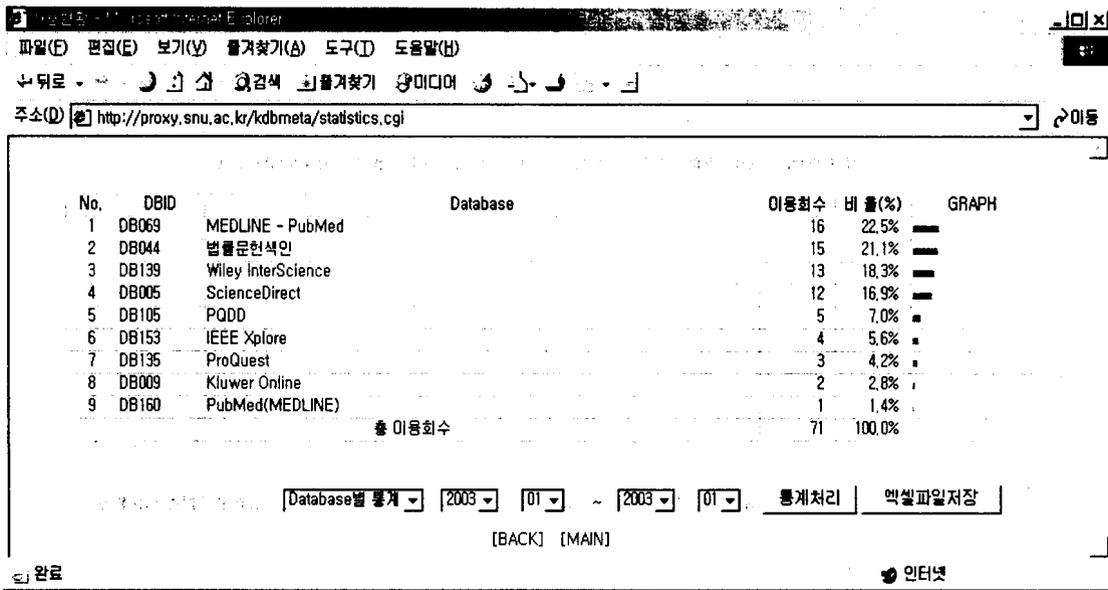


<이용자별 통계>

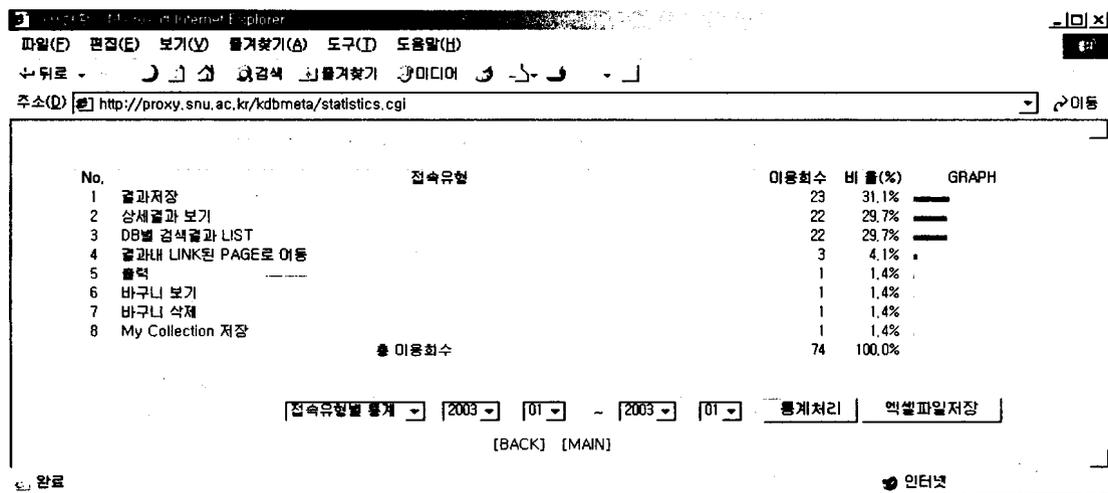
통계처리 선택에서는 이전화면으로 이동하지 않고 상세이용현황 통계를 확인할 수 있다.



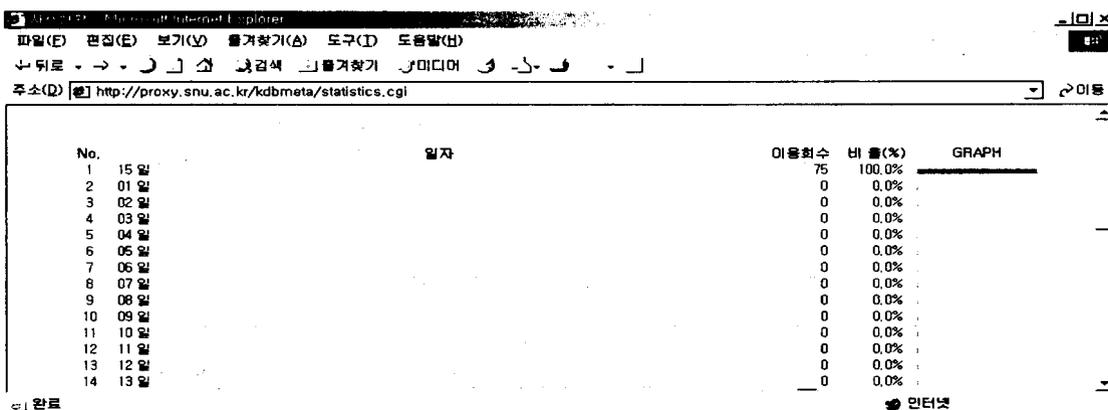
<데이터베이스별 통계>



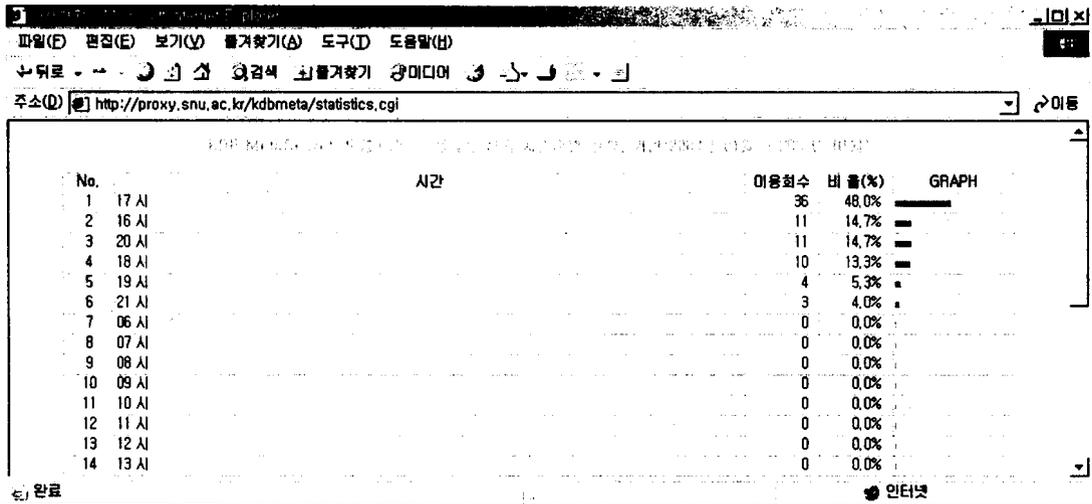
<IP 주소별 통계>



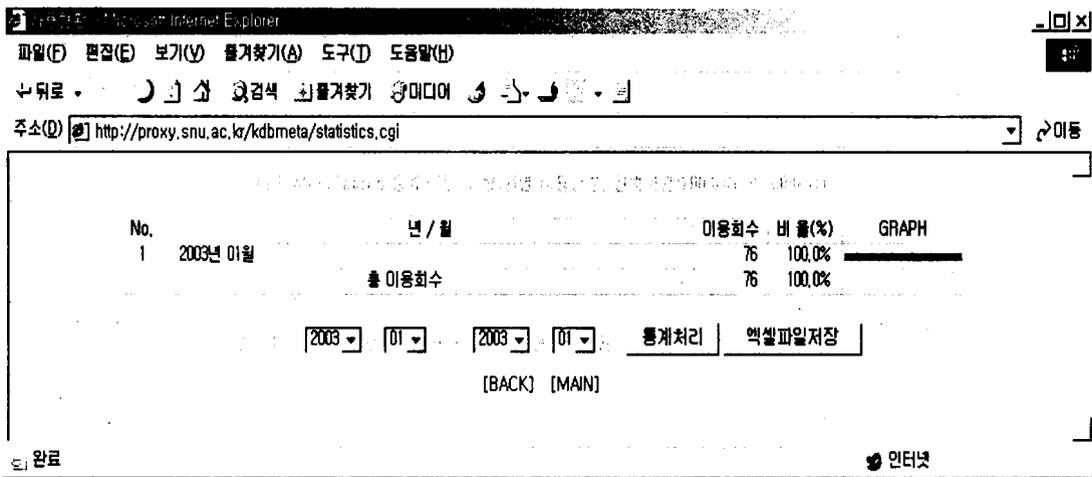
<일자별 통계>



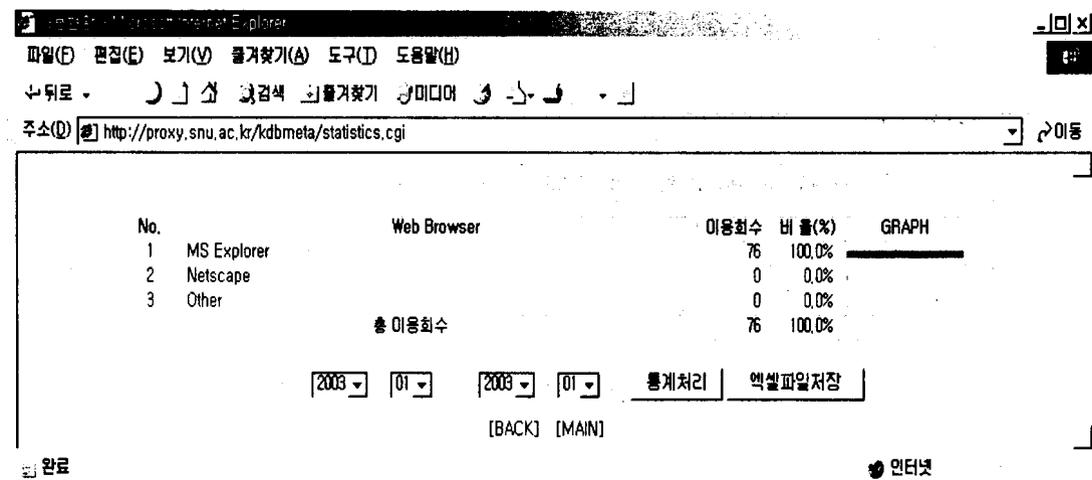
<시간대별 통계>



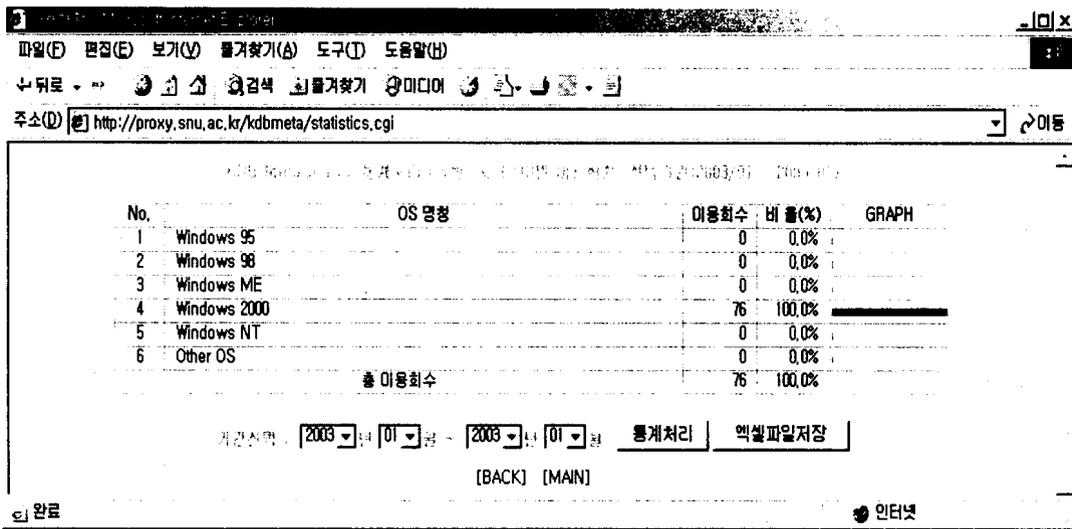
<월별 이용현황>



<웹 브라우저별 이용현황>



< OS별 이용현황 >



2.1.3. 웹디비 통합 검색 시스템 문제점

2.1.3.1. 사이트별 검색 유형 및 표시항목이 상이

웹디비 통합검색 시스템은 각각의 Web DB 회사에서 제공하는 사이트의 정보를 parsing하여 분석된 정보를 불러와 서울대에서 원하는 항목들을 display해주고 있으나 검색방식, 제한 검색자, 검색 후 뿌려지는 간략항목 및 상세 항목 등이 각 DB 회사별로 상이하므로 이에 대한 표준을 마련하기 어렵다.

2.1.3.2. 저자명, 서명, 기사명, 키워드 검색

위 항목에 대한 검색이 원래 사이트에서 우측절단으로 검색어를 인식할 경우도 있을 수 있고 다른 사이트의 경우는 키워드 검색으로 세팅 되어 있을 수 있으나 이 정보를 DB별로 구분하여 표시해 주긴 어렵다.

2.1.3.3. 구독 되는 모든 DB에 대한 통합검색 불가

현재 서울대학의 경우 40여 개가 넘는 DB를 구독하고 있으나 동시에 통합검색을 해올 수 있는 최대의 DB개수는 20여 개 이므로 원하는 주제 분야에 해당하는 DB가 20여 개가 넘을 경우 이용자들이 혼란스러울 수 있다.

2.1.3.4. DB 유형이 다른 경우 통합검색이 불가

Westlaw나, Britannica 백과사전의 경우 전자는 기사별 검색기능을 제공하지 않으므로 후자는 기사유형이 아니므로 통합검색이 불가하여 이러저러한 사유로 일부DB는 통합검색 대상에서 제외되었다.

2.1.3.5. 구독 사이트 추가 시 통합 검색 대상 사이트 분석을 전문가에게 의뢰해야 함. 매년 새로운 DB구독이 증가하고 있는 반면 각 DB사의 사이트 정보에 대한 분석이 자동으로 이루어 질 수가 없으므로 이럴 때 마다 매번 사이트 분석을 의뢰 해야 하는 번거로움이 있다.

2.2. 프락시 서버

2.2.1. Proxy Server 시스템 개요

도서관에서 도입하여 서비스하고 있는 대부분의 해외학술정보(전자저널, 웹 데이터베이스)들은 서비스 제공사와의 계약조건에 의해 허가받은 IP주소범위 내에서만 이용가능하다.

즉, 서울대학교 이용자들이 교내를 벗어나 외부에서 본 서비스를 이용하고자 할때 접속 허용이 안되어 해외학술정보를 이용할수 없는 제약이 따른다.

본 "인증방식변환서버"는 이러한 지역적 제약에서 벗어나 외부에서도 서울대학교 소속이란 신분확인이 된 이용자에 한해 해외학술정보를 자유롭게 이용할 수 있도록 지원해주는 시스템이다.

외부에서 해외학술정보를 이용하려면 아래 설명대로 최초 한 번 이용자의 웹 브라우저에 프록시 자동구성 스크립트 설정을 해 주어야 한다.

자동구성 스크립트에 의한 프록시 서버 경유는 도서관 담당자에 의해 등록된 해외학술정보 제공사의 서비스 URL에 접속할 경우에만 자동으로 프록시 서버를 통해 접속되며 기타 다른 사이트의 접속은 프록시 서버를 경유하지 않고 직접 접속된다.

****프락시 서버 이용 대상으로 설정된 해외학술정보 서비스 URL로 접속 시 아래와 같은 이용자 인증창이 나타나면 ID와 암호를 입력한 후 확인을 클릭하면 해당 서비스를 이용할 수 있다.**

네트워크 암호 입력
? | X

사용자 이름 및 암호를 입력하십시오.

방화벽: proxy.snu.ac.kr

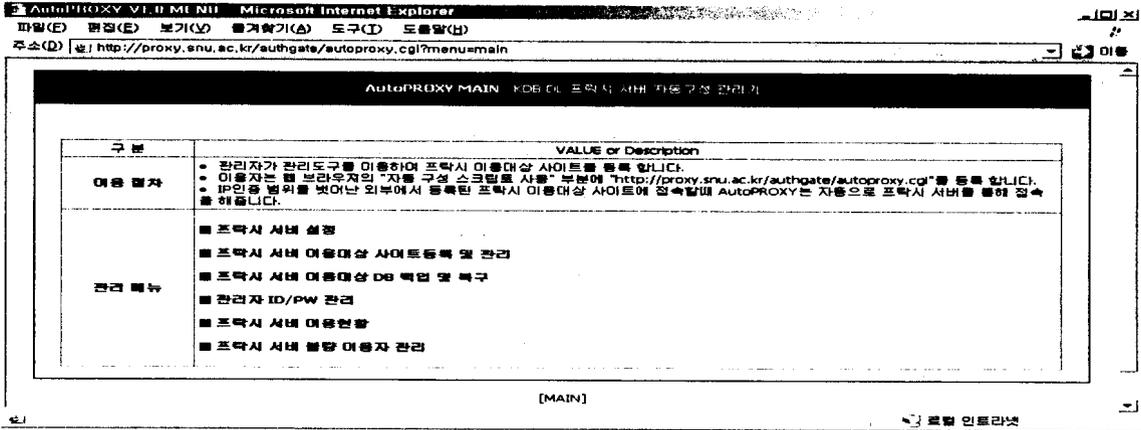
영역: 해외학술정보 GATEWAY(외부사용자용)

사용자 이름(U):

암호(P):

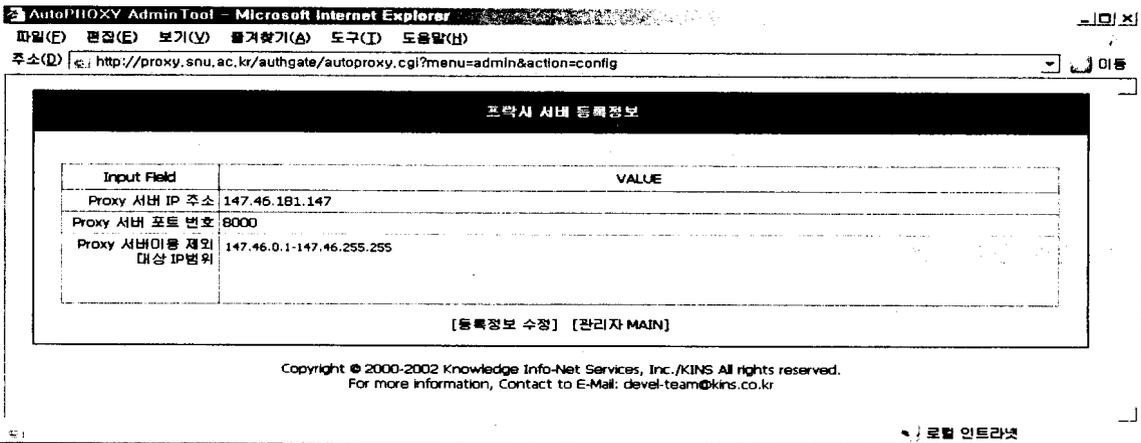
이 암호를 암호 목록에 저장(S)

2.2.2. 시스템 운영 및 관리



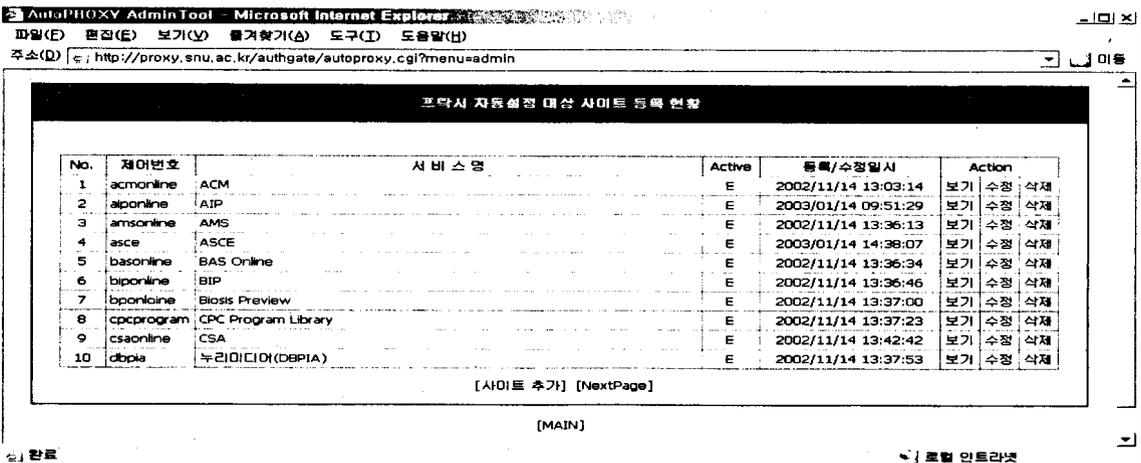
2.2.2.1. 프락시 서버 설정

프락시 서버 이용 대상 IP 대역을 설정한다.



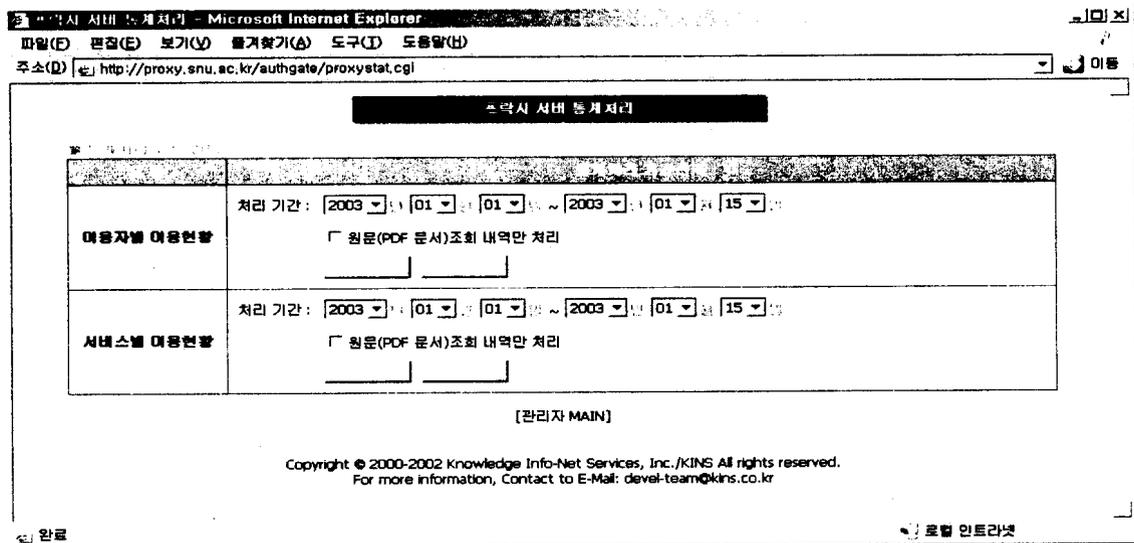
2.2.2.2. 프락시 서버 이용대상 사이트 등록 및 관리

프락시 서버를 통해 이용할 대상 전자저널, 웹 DB를 등록한다.



2.2.2.3. 프락시 서버 이용현황

프락시 서버 이용에 대한 이용자별, 서비스별 통계를 산출한다.



<이용자별 이용현황>

이용자별 이용횟수, 비율, 상세내역을 확인할 수 있다.

No.	이용자 ID	이용자 명	이용회수	비율(%)	GRAPH	상세내역
1	TGtster	TG일반이용자데이터	4	0.0%		
2	admin	admin	4	0.0%		
3	6306301030213	강용구	70	0.4%		
4	7605302030023	고속경	42	0.2%		
5	7111051052713	고영상	28	0.1%		
6	7708201063715	고원규	339	1.8%		
7	7806071079629	공수철	30	0.2%		
8	7603081802410	구자균	20	0.1%		
9	7607251080512	권경성	121	0.6%		
10	7212291779111	권영배	300	1.6%		
11	6601051701710	김경현	12	0.1%		
12	7102211011130	김광선	86	0.5%		
13	8306152012113	김근희	6	0.0%		
14	6510181046416	김남수	10	0.1%		
15	7806282056215	김미담	24	0.1%		
16	8012262168313	김미순	125	0.7%		
17	7509301155521	김민영	9	0.0%		
18	5203111068315	김상중	2	0.0%		
19	6801181057022	김상철	77	0.4%		
20	7307071495318	김석	46	0.2%		
21	6403251018711	김성수	15	0.1%		

<서비스별 이용현황>

서비스별 이용횟수, 비율, 상세내역을 확인할 수 있다.

프록시 서버 이용 현황

No.	서비스 ID	서비스 명	이용횟수	비율(%)	GRAPH	상세내역
1	acmonline	ACM	35	0.2%		<input type="checkbox"/>
2	alponline	AIP	660	3.7%		<input type="checkbox"/>
3	amsonline	AMS	7	0.0%		<input type="checkbox"/>
4	asce	ASCE	16	0.1%		<input type="checkbox"/>
5	basonline	BAS Online	9	0.0%		<input type="checkbox"/>
6	biponline	BIP	90	0.5%		<input type="checkbox"/>
7	bponline	Bios Preview	1014	5.6%		<input type="checkbox"/>
8	cpcprogram	CPC Program Library	27	0.1%		<input type="checkbox"/>
9	csaonline	CSA	153	0.8%		<input type="checkbox"/>
10	ebasco	EBSCO Online	1784	9.9%		<input type="checkbox"/>
11	engnetbase	ENGNETBASE	22	0.1%		<input type="checkbox"/>
12	eric	ERIC Document	148	0.8%		<input type="checkbox"/>
13	ecorlit	Ecorlit	6239	34.6%		<input type="checkbox"/>
14	firstsearc	FirstSearch	146	0.8%		<input type="checkbox"/>
15	frantext	Frantext	2	0.0%		<input type="checkbox"/>
16	hanson	Hanson Online	143	0.8%		<input type="checkbox"/>
17	historical	Historical Abstract	8	0.0%		<input type="checkbox"/>
18	ideal	IDEAL	32	0.2%		<input type="checkbox"/>
19	ieeexplore	IEEExplore	449	2.5%		<input type="checkbox"/>
20	iponline	IOP	76	0.4%		<input type="checkbox"/>
21	jstor	JSTOR	149	0.8%		<input type="checkbox"/>

<원문(PDF 문서) 조회내역만 처리>

통계처리시 '원문(PDF 문서) 조회내역만 처리' 항목을 체크하면 실질적인 원문 사용인 PDF 접속건수를 count 하여 불량이용자 판별 정보를 제공한다.

프록시 서버 이용 현황

No.	이용자 ID	이용자 명	이용횟수	비율(%)	GRAPH	상세내역
1	6306301030213	강용구	35	0.9%		<input type="checkbox"/>
2	7708201063715	고원규	30	0.8%		<input type="checkbox"/>
3	7806071079629	공수철	11	0.3%		<input type="checkbox"/>
4	7603081802410	구자근	4	0.1%		<input type="checkbox"/>
5	7607251080512	권경성	5	0.1%		<input type="checkbox"/>
6	7212291779111	권영배	107	2.8%		<input type="checkbox"/>
7	7102211011130	김광선	24	0.6%		<input type="checkbox"/>
8	7806282056215	김미담	2	0.1%		<input type="checkbox"/>
9	8012262168313	김미순	11	0.3%		<input type="checkbox"/>
10	7509301155521	김민영	4	0.1%		<input type="checkbox"/>

보기를 선택하면 이용자별로 접근시간, 접속IP 주소, 접속 PDF 정보가 출력되며 이 정보를 이용하여 불량 이용자 여부를 판단하고 이 화면에서 직접 불량 이용자 등록처리를 할 수 있다.

<이용자별 프락시 서비스 상세 내역>

프락시 서버 관리자 - Microsoft Internet Explorer

주소 http://proxy.snu.ac.kr/authgate/proxystat.cgi

프락시 서버 이용 현황

전체서버 옵션
 한국과학기술원 프락시 서버 이용 현황 (2003/1/10 10:29:06 ~ 2003/1/10 10:29:06)
 한국과학기술원 프락시 서버 이용 현황 (2003/1/10 10:29:06 ~ 2003/1/10 10:29:06)
 한국과학기술원 프락시 서버 이용 현황 (2003/1/10 10:29:06 ~ 2003/1/10 10:29:06)
 한국과학기술원 프락시 서버 이용 현황 (2003/1/10 10:29:06 ~ 2003/1/10 10:29:06)

불량 사용자 등록 :
 - 제한기간 : 2003 01 15 ~ 2003 02 14
 - 등록사유 :

접수시간	IP	URL
2003/1/09 18:32:33	203.229.137.1	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/09 18:32:33	203.229.137.1	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/09 18:32:39	203.229.137.1	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/09 18:32:40	203.229.137.1	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:29:06	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:29:06	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:29:16	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:29:17	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:29:46	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:30:25	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:31:14	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:31:25	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:31:25	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:31:25	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:32:27	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:32:28	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:32:31	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef
2003/1/10 10:32:36	211.253.87.34	http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6TOC-473M6RM-2-G&_cdi=4859&_orig=browsef

완료 인터넷

2.2.2.4. 프락시 서버 불량 사용자 관리

다량의 다운로드등 이용에 문제가 있는 불량 이용자를 등록 관리한다.

AutoPROXY AdminTool - Microsoft Internet Explorer

주소 http://proxy.snu.ac.kr/authgate/proxybuser.cgi

프락시 서버 불량 사용자 등록 현황

No.	이용자 ID	이용자 이름	이용제한 기간	이전 제한 이력	등록/수정일시	Action
1	12	홍길동		1 회	2003/01/12 22:42:42	보기 수정 삭제

[불량 사용자 추가]

[MAIN]

Copyright © 2000-2002 Knowledge Info-Net Services, Inc./KINS All rights reserved.
For more information, Contact to E-Mail: devel-team@kins.co.kr

완료 로컬 인트라넷

<불량 이용자 등록>

ID 검색, 이름 검색에 의해 불량 이용자 등록 대상을 검색하고 이용제한 기간, 이용 제한 사유를 등록한다.

<불량 이용자 등록II>

관리자는 프락시 이용통계 상세 내역을 확인후 해당 이용자에 대한 불량 이용자 등록을 처리할 수 있다.

날짜	IP	URL
2003/1/08 22:32:18	203.248.194.101	http://content.ebsco.com/fulltext.asp?wasp=524cuptrw15ac159dcb&ext=.pdf
2003/1/08 22:34:29	203.248.194.101	http://content.ebsco.com/fulltext.asp?wasp=524cuptrw15ac159dcb&ext=.pdf

2.2.3. Proxy Server 이용 시 문제점

2.2.3.1. One ID, Password 방식 이용에 어려움.

현재 서울대 도서관에서는 솔라즈 II의 개인정보 관리나, 전자도서관에서의 개인정보 관리 등을 위하여 같은 Id, Password를 사용할 뿐만 아니라 한번 로그인 했을 경우 별도의 로그인 작업이 필요 없이 필요한 기능을 수행 할 수 있으나, 프락시 서버 사용을 위해 로그인 할 경우는 이미 솔라즈나 전자도서관에 로그인이 된 상태라 할지라도 별도의 로그인 작업이 요구된다.

2.2.3.2. 프락시를 경유한 웹디비 이용은 사이트 등록으로만 가능

현재 학내에서 IP등록으로 가능한 웹디비 대상 사이트의 경우 학내IP를 등록해주므로써 학내 이외에서 접속할 경우 프락시 서버를 경유하도록 하는 방법을 택하고 있다. 그러나 구독방법이 다양화되어 분관과 같은 IP대를 공유함과 동시에 다른 구독 DB를 가지게 될 경우 대상 사이트를 벗어난 비구독 캠퍼스 이용자들의 이용 통제방법이 필요하다 -> 현재 이용자들의 소속 코드를 등록하여 로그인 했을 당시 소속기관별로 프락시 이용을 제한하는 방법을 연구하고 있다.

2.2.3.3. 통계 정형화의 어려움

현재 어떤 DB의 어떤 원문을 이용자가 몇번 이용했는가를 분석함에 있어서 같은 사이트에서 하나의 PDF를 열어 보는 동안에도 사이트 정보를 계속 주고 받는 방식을 택하는 DB사의 PDF를 이용할 경우 실제 이용자는 하나의 원문을 열어보았으나 같은 url정보가 계속 상세 내역으로 시간대별로 등록이 되어 원문이용별 정확한 정보를 분석하기 위해 수작업이 필요하다.

2.3. ID/PW 인증변환 시스템

2.3.1. Authgate 시스템 개요

AuthGATE은 기관에서 구독하는 ID/PW 인증방식의 웹 데이터베이스를 IP방식의 인증으로 변환, 제공하는 도구로 학내 IP로 등록된 DB를 학외에서 로그인정보를 통해 서비스를 제공하는 프락시시스템과는 반대의 개념이라 볼 수 있으며 기관에 속한 이용자들이 학술 데이터베이스를 등록된 사이트내에서 보다 간편하게 이용할 수 있도록 해준다. 이 과정은 ID/PW 이용 대상을 관리자가 Authgate 시스템에 등록하면 이용자는 별도의 작업 없이 사용할 수 있다.

2.3.2. 시스템 운영시 문제점

ID/PWD인증방식을 IP인증방식으로 변환하기 위해서는 전문가가 인증 변환 스크립트를 구성하여야 하는 번거로움이 있다.

2.4. SDI 서비스 (KDB AutoAlert SDI)

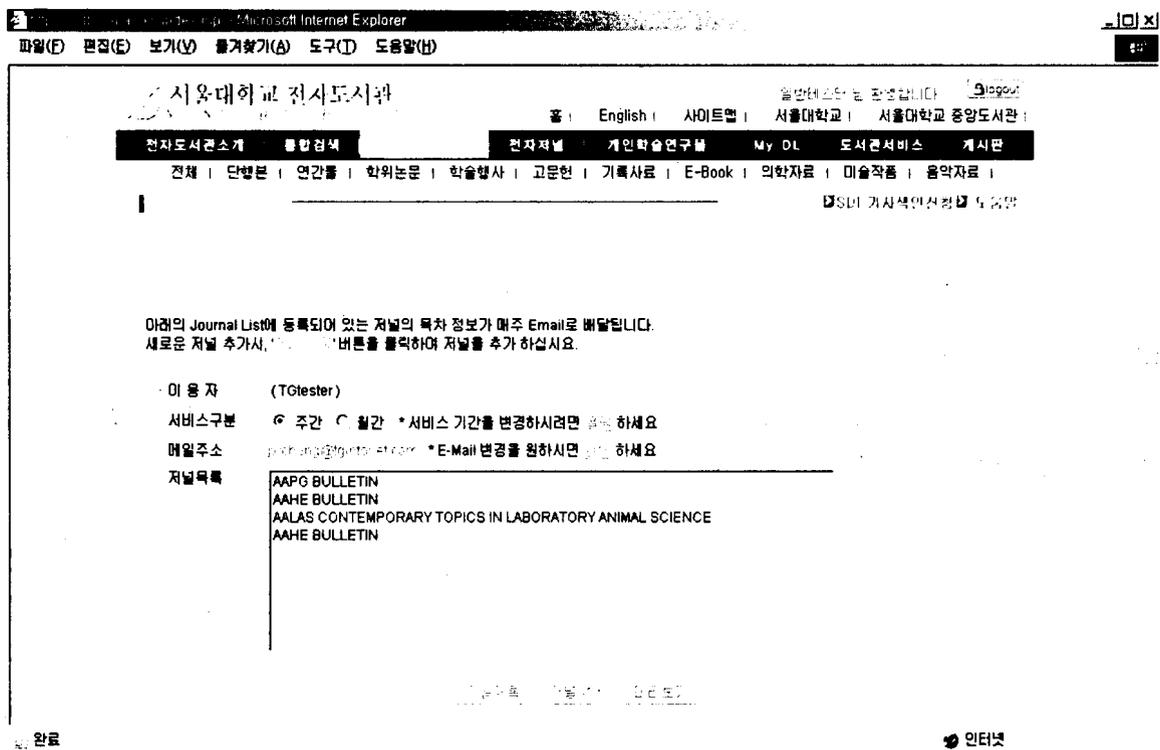
2.4.1. SDI 서비스 소개

British Library와 한국학술정보에 수록된 데이터를 기반으로 하여 26,000 여종의 국내외 학술지 수록 기사 및 목차 리스트를 제공한다. SDI 서비스 기간은 이용자가 주간 또는 월간을 선택하여 이용할 수 있으며 저널 단위로 혹은 원하는 주제분야의 기사를 서비스 받아 볼 수 있다.

2.4.2. 기능

◎ 학술지명으로 신청

Journal List에 등록되어 있는 저널의 목차 정보가 이용자가 등록한 간격에 따라 (주간, 혹은 월간) Email로 배달된다.



◎ 저널 등록 및 삭제

저널 등록을 누르면 등록된 저널 리스트가 출력되며 이용하고자 하는 저널을 검색하여

선택한 후 추가 버튼을 누른다. 삭제를 원할 경우는 저널목록에서 삭제할 저널을 클릭한 후 저널 삭제를 누르면 삭제 확인 창이 출력되며 확인을 누르면 삭제된다

서울대학교 전자도서관

일반테스터님 환영합니다. [logout]

홈 | English | 사이트맵 | 서울대학교 | 서울대학교 중앙도서관

전자도서관소개 | 통합검색 | DL Contents | 전자저널 | 개인학습연구실 | 도서관서비스 | 게시판

My Collection | My Query | 도서관 소식지 | My Search Preference | 매체보기 신청조회 | SDI 서비스 신청

SDI 가사색인신청 | 도움말

저널검색

저널검색 입력란에 저널 이름을 입력한 후, 검색버튼을 클릭하시면 저널목록이 나타납니다.
 검색된 저널 목록에서 알맞은 저널을 선택하시고 후 추가 버튼을 선택하시면 저널이 새롭게 등록됩니다.

서비스구분: 주간 월간 *서비스 기간을 변경하시려면 클릭 하세요

이용자: (TGtester)

저널검색 검색

총저널수: 26332 건 현재 페이지: 1 / 1317 페이지

전체선택 전체해제 Goto Page: 1

NO.	저널명	선택
1	17 TO 18	<input type="checkbox"/>
2	19TH CENTURY MUSIC	<input type="checkbox"/>
3	21ST CENTURY SCIENCE AND TECHNOLOGY	<input type="checkbox"/>
4	3 R INTERNATIONAL	<input type="checkbox"/>
5	33 METAL PRODUCING	<input type="checkbox"/>
19	AAPG EXPLORER	<input type="checkbox"/>
20	AARC TIMES	<input type="checkbox"/>

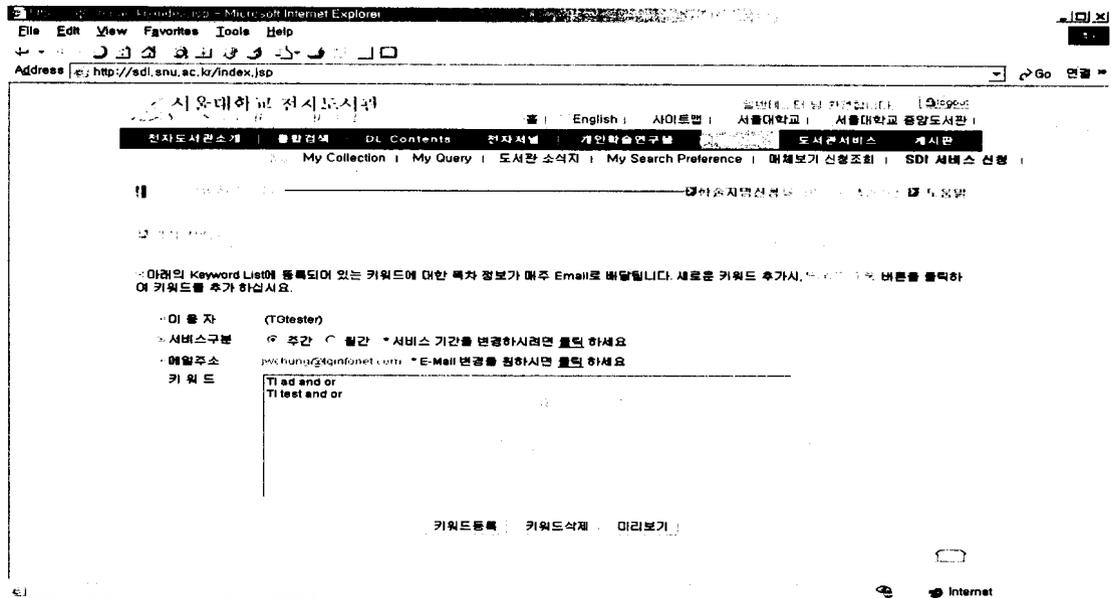
추가 취소

◎ 미리보기

저널 목록에 등록된 저널을 클릭후 미리보기 버튼을 실행하면 해당 저널명으로 등록된 SDI 서비스 정보를 확인할 수 있다.

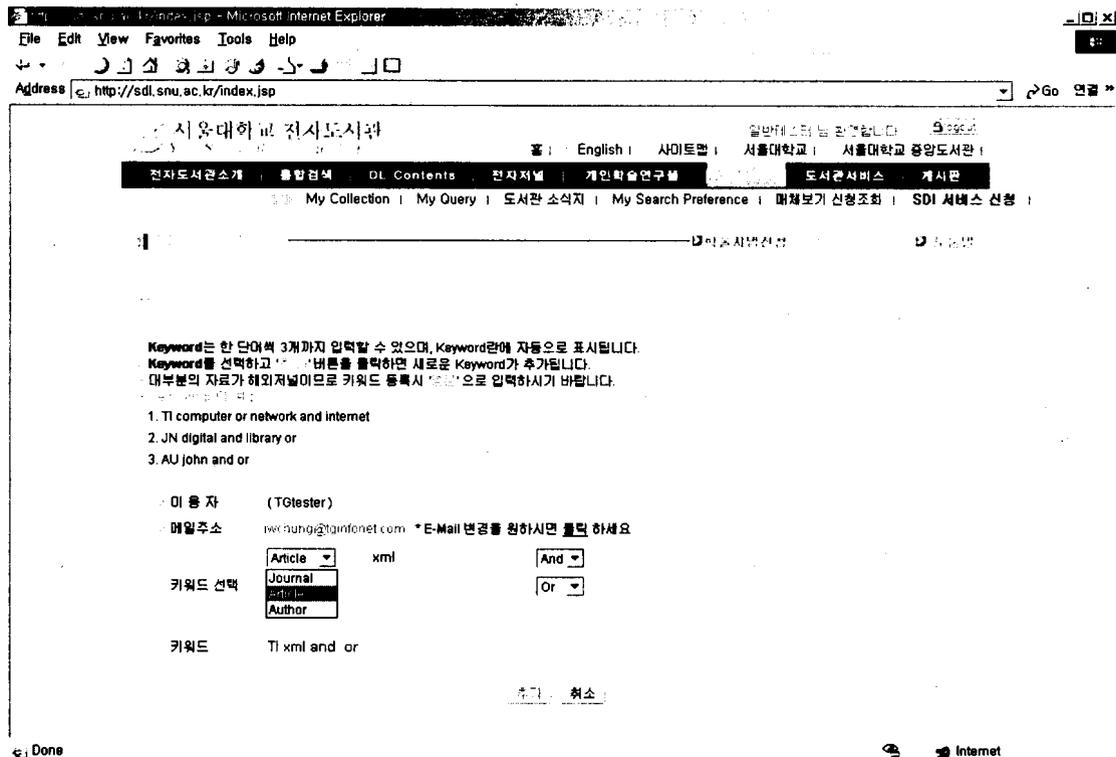
◎ 기사색인 신청

Keyword List에 등록되어 있는 키워드에 대한 목차 정보가 매주 Email로 배달된다.



◎ 키워드 등록 및 삭제

Journal, Article, Author 항목에 해당하는 키워드를 AND, OR 조합에 의해 등록 할 수 있고 삭제할 수 있다.

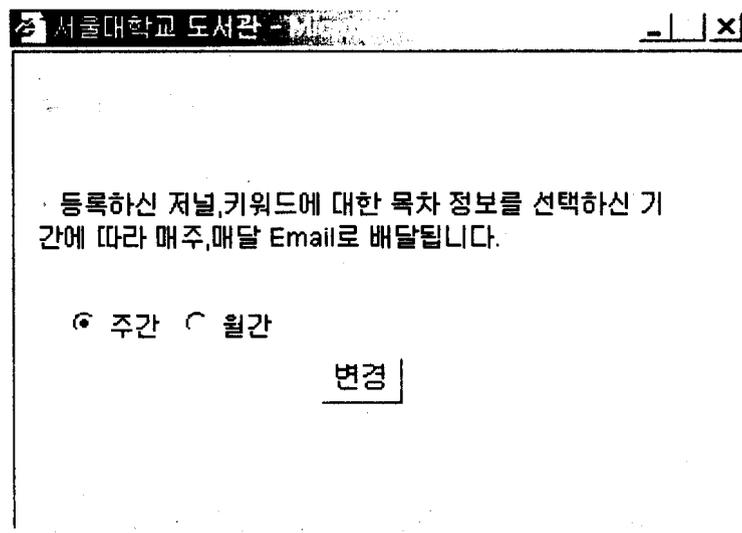


◎ 미리보기

저널 목록에 등록된 저널을 클릭후 미리보기 버튼을 실행하면 해당 저널명으로 등록된 SDI 서비스 정보를 확인할 수 있다.

◎ 서비스 구분

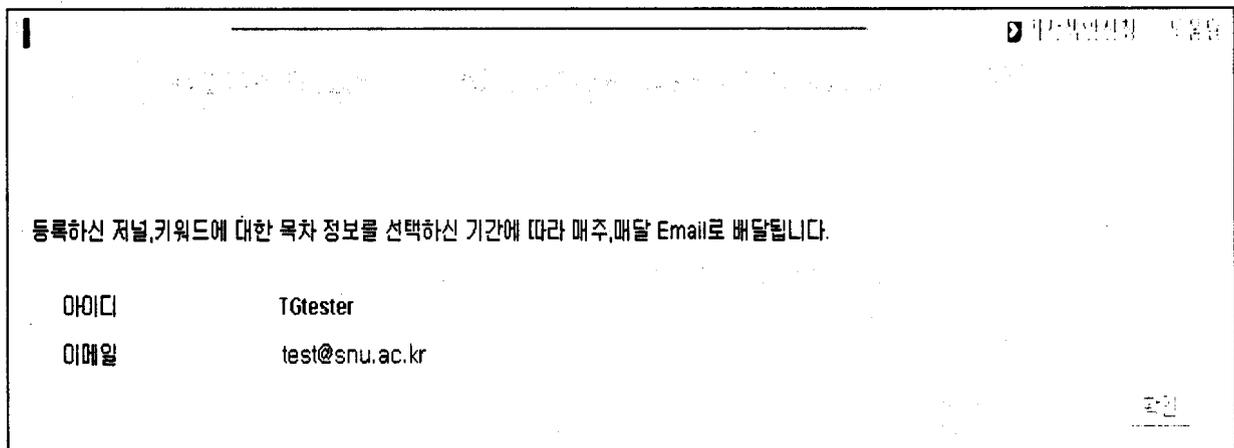
SDI 서비스 기간을 주간 또는 월간으로 설정하여 이용자의 편의에 맞춰 제공한다.



The screenshot shows a web browser window titled "서울대학교 도서관" (Seoul National University Library). The main content area contains the following text: "등록하신 저널, 키워드에 대한 목차 정보를 선택하신 기간에 따라 매주,매달 Email로 배달됩니다." (The table of contents information for the journals and keywords you registered will be delivered to you via email weekly or monthly, depending on the period you selected). Below this text are two radio buttons: "주간" (Weekly) with a selected dot, and "월간" (Monthly) with an unselected dot. At the bottom right of the form is a button labeled "변경" (Change).

◎ 메일주소 변경

SDI 서비스 이용을 위해 등록하는 E-Mail 주소를 입력, 수정할 수 있다.



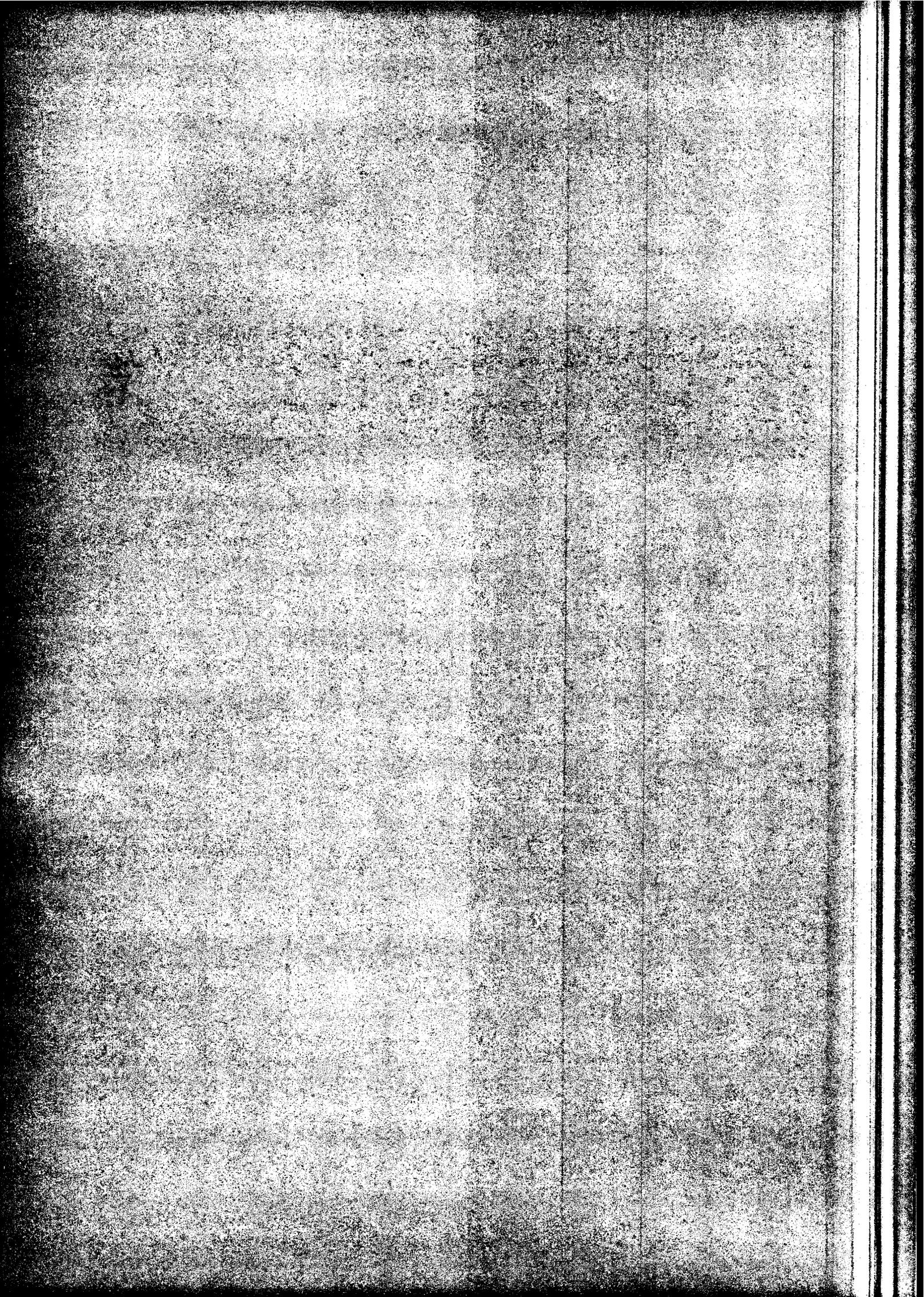
The screenshot shows a web browser window titled "서울대학교 도서관" (Seoul National University Library). The main content area contains the following text: "등록하신 저널, 키워드에 대한 목차 정보를 선택하신 기간에 따라 매주,매달 Email로 배달됩니다." (The table of contents information for the journals and keywords you registered will be delivered to you via email weekly or monthly, depending on the period you selected). Below this text are two input fields: "아이디" (ID) with the value "TGtester" and "이메일" (Email) with the value "test@snu.ac.kr". At the bottom right of the form is a button labeled "확인" (Confirm).

[1994.7]

2003. 5

류운주

woonju@nhan.ac.kr



< 사례발표 7 >

DjVu 포맷의 적용사례

2003. 5.

류운주
(woonju@snu.ac.kr)

차 례

1. 머리말	1
2. DjVu 포맷의 도입 배경	1
3. DjVu 포맷	2
4. 원문 이미지 포맷과 DjVu 포맷 비교	6
5. 원문 이미지 구축의 기본방침	12
6. 맺음말	16

1. 머리말

대학이나 주요 기관들이 전자도서관 서비스를 위해 원문컨텐츠 구축 사업을 추진하고 있거나 추진할 계획을 세우고 있다. 서울대학교에서는 학내 모든 기관이 소장하고 있는 다양한 유형의 자료를 고품질의 원문 DB를 구축하여 서비스 하고자 하였다. 이 과정에서 고화질의 원문 DB를 보관하고 웹을 통해서 서비스하고자 할 때 현재 표준 포맷들의 한계점이 노출 되었다.

이러한 한계점을 극복하기 위해 원문 컨텐츠 구축 및 서비스 기술인 DjVu 포맷을 국내에서는 처음으로 서울대학교 전자도서관 서비스에 적용하게 되었다. PDF 포맷이 디지털 문서 중심인 반면 DjVu 포맷은 스캐닝한 원문 이미지 중심으로 특히 칼라 원문 이미지 서비스 기술이 우수하다. 이러한 DjVu 포맷을 소개하고, 도입배경, 적용대상, 서울대학교의 실제 적용사례 그리고 DjVu 기술을 처음 적용함으로써 발생된 기술 및 운영상의 문제점을 분석하여 전자도서관 컨텐츠 구축 사업을 추진하는 대학 및 주요 기관들에게 도움이 되고자 한다.

전자도서관 컨텐츠 구축 시 서울대학교의 적용 사례를 참고하여 다양한 자료에 대한 원문 파일 포맷을 검토한다면 구축 과정의 시행착오를 줄일 수 있는 방법이 될 것이다.

2. DjVu 포맷의 도입 배경

서울대학교 중앙도서관은 1999년부터 학위논문 4만여 건의 원문을 구축하여 PDF로 원문을 서비스해 왔으며 2001년부터 3년간 다량의 컨텐츠를 구축하여 2단계 전자도서관 사업을 추진하고 있다. 컨텐츠 구축은 도서관 자료와 캠퍼스 내 전 기관에 컨텐츠 구축 수요 조사를 실시하여 컨텐츠 구축 선정 위원회에서 구축 대상을 선정하였다. 컨텐츠 구축 대상 자료는 책자형 자료, 슬라이드, 마이크로필름, 사진, VOD, AOD 등으로 자료의 유형이 다양하다.

전자도서관 관련 소프트웨어의 개발과 고문헌, 슬라이드 및 사진자료를 고해상도의 컬러 이미지로 다량의 컨텐츠 구축 사업을 추진하다 보니 원문의 파일 크기가 커지게 되고 서비스에 대한 다음과 같은 사업추진의 문제점이 발생하게 되었다.

○ PDF 파일의 한계

그동안 책자형 이미지는 PDF 포맷으로 원문을 구축하였는데 파일크기가 10MB 이상이면 PDF파일이 열리지 않으며 10MB 이하라도 원문을 보는데 시간이 많이 소요된다.

○ 네트워크 트래픽

파일 사이즈가 큰 PDF 파일, 컬러 이미지로 구축된 슬라이드 및 사진 등은 이용자가 많을 경우 네트워크의 트래픽을 유발할 수 있다.

○ 고해상도 유지를 위한 적정 해상도 유지

고해상도의 컬러 이미지로 구축되는 이미지가 적절한 해상도를 유지하면서 서비스에 장애가 되지 않도록 하는 것은 어렵다. 따라서 고해상도의 보관용 이미지와 저해상도의 서비스용 이미지를 이중으로 구축하는 문제점이 발생하였다.

○ 저장장비 확장비용 및 Back-up

텍스트 위주의 문서와는 달리 이미지 원문 파일은 사이즈가 커서 콘텐츠구축에서 저장장비 확장 및 Back-up 비용을 최소화 할 수 있어야 하고 용이하여야 한다.

위와 같은 문제점을 해결하기 위하여 서울대학교에서는 DjVu를 분석하여 부분적으로 도입하게 되었다.

3. DjVu 포맷

3.1. DjVu란?

DjVu라는 단어는 불어 'déjà vu' 에서 온 것으로 '어디서 본 듯한 느낌 또는 착각' 이란 뜻을 갖고 있다고 한다. 우리말로로는 '데자뷰'로 발음한다. DjVu는 웹상에서 통용되는 TIFF, PDF, JPEG과 같이 또 다른 하나의 파일 포맷이기도 하다. DjVu로 만들어진 문서는 확장자가 'djvu' 또는 'djv'로 표시된다. DjVu는 원문 이미지 구축 시 원본자료의 그 품질 그대로 유지하고 검색, 배포, 압축, 저장할 수 있게 한 압축 기술이다. 이 기술은 1990년대 후반에 미국 AT&T Lab에서 연구 개발된 것으로 2000년 초에 미국 시애틀 소재의 리자드텍사에서 관련 기술을 인수하였다.

DjVu 기술이 주는 주요 사상은 "Scan-to-Web" 으로서, 그 동안 스캐닝을 통해서 디지털화 되는 일반 문서나 고화질의 사진 등이 그 파일의 크기가 커서 웹을 통해서 서비스가

불가능하던 문제점을 해소하는 기술인 것이다. 최근에는 고품질의 컬러 스캐너와 디지털 카메라의 보급으로 오프라인상의 자료들에 대한 디지털 화 요구가 증대되고 있는 시점에서 이러한 기술은 진가를 발휘할 수 있을 것이다.

3.2. DjVu의 특징

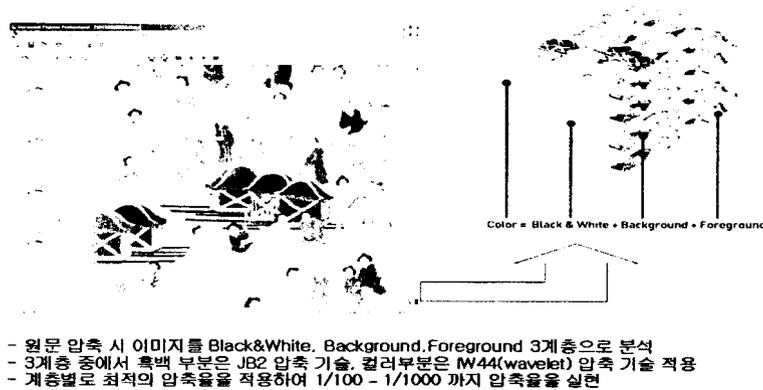
3.2.1. 압축방식에 있다.

○흑백 이미지 압축 방식

흑백 이미지 압축방식인 TIFF G4와 비교해서 DjVu는 TIFF G4보다 3-5배 정도의 압축 효율을 갖는다. 흑백 이미지 내에서 텍스트 영역을 분리하여 텍스트의 윤곽선을 최대한 보존하여 확대를 할 경우도 선명도를 최대한 유지하게 하며, 반복적으로 나타나는 글자의 공통적인 패턴을 추출하여 대표가 되는 패턴만을 압축 저장하고 나머지 반복되는 패턴은 위치 정보만을 저장하는 기법을 적용하고 있기 때문에, 글씨가 많은 흑백 이미지 문서를 압축할 때 매우 효과가 있다.

○컬러 이미지 압축 방식

컬러 원문 이미지는 텍스트 부분의 흑백 계층과, 배경 부분의 컬러 계층, 전경 부분의 컬러 계층으로 분리하여 각각 가장 좋은 효율을 갖는 압축 방식을 적용하여 기존의 JPEG 압축 방식보다 5배에서 20배정도까지 압축률을 높일 수 있는 것이 특징이다. 또한 원문 이미지가 일반 배경 사진과 같은 이미지는 계층별로 분리하지 않고 전체 이미지를 한 통으로 압축하는 방식을 선택적으로 적용할 수도 있기 때문에 원문 이미지의 특성을 최대한 고려하여 최상의 효율을 갖는 압축 방식을 적용하여 가장 빠르게 전송 및 디스플레이 할 수 있게 하는 것이 특징이다. [그림-1] 참조



[그림-1]

3.2.2. 스캐닝한 이미지 문서도 문서의 속성을 최대한 유지한다.

기존의 압축 방식들(흑백인 경우는 TIFF G4, 컬러인 경우는 JPEG이나 JPEG2000)과 같은 경우는 스캐닝한 문서나 디지털 사진이든 모두를 문서의 내용과 상관없이 전체를 픽셀들로 구성된 이미지로 보고 압축을 하거나 표현하는 방식이다. 이러한 방식들은 스캐닝한 문서는 디지털 카메라로 찍은 사진이든 모두를 같은 이미지로 보고 일괄적으로 압축 표현하는 방식이다. 이러한 방식들이 갖는 대표적인 문제점은 여러 페이지로 구성되어 있는 문서나 책자의 정보를 가질 수가 없고 메타 정보 등을 저장할 수도 없는 단순한 순차적인 이미지의 파일로 표현할 뿐이다. DjVu는 JPEG이나 TIFF와 같이 낱장 단위의 파일로 구성할 수도 있지만, 페이지들로 구성되는 책자를 표현하는 수단으로 DjVu 포맷 안에 페이지 정보, 메타 정보를 XML 형태, 목차 정보 등을 삽입하여 문서의 형태를 갖고 배포나 검색 서비스를 제공할 수 있다.

3.2.3. DjVu 문서의 웹 배포 방식의 특징

○번들 방식

PDF 문서와 같이 모든 페이지를 갖는 문서를 하나의 파일로 저장하여 전송할 수 있게 하는 방식이다. 사용자가 웹을 통해서 이 방식의 문서를 볼 경우 PDF 문서와 같이 파일 전체를 다운로드 받은 후 읽게 하는 것이다. 주로 적은 페이지를 갖는 작은 크기의 문서를 e-mail이나 FTP로 전송할 때 유용한 방식이다.

○인다이렉트 방식

DjVu 포맷에서만 제공하는 방식으로 문서의 페이지 수와 상관없이 페이지 단위의 스트리밍을 제공하기 위한 방식이다. 귀중도서와 같은 고문헌들은 500페이지가 넘는 자료들이 많은데 이러한 자료를 컬러 JPEG이나 PDF로 표현했을 경우 그 크기가 100MB 이상 될 수가 있으며 이 정도 크기면 한 파일로 인터넷을 통해서 배포나 검색을 하는 것은 거의 불가능 하다. 일단 한 파일의 크기가 페이지 수와 상관없이 10MB 이상 되면 사용자는 일단 다운로드 받는 시간으로 인하여 기피 대상이 될 수 있다. DjVu는 문서의 페이지 수와 상관없이 그리고 파일 크기에 상관없이 전체 파일을 모두 다운로드 받은 후 읽는 방식이 아닌 페이지 단위로 필요한 내용만 선택적으로 전송하는 기술을 사용하고 있으므로 읽으면서 동시에 주변의 페이지를 자동으로 다운로드 할 수 있게 하는 방식이다. 그러므로 500페이지 이상의 문서를 인터넷을 통해서 읽을 경우 첫 페이지를 액세스하나 마지막 페이지를 액세스하나 일정한 검색 속도(1-3초)를 항상 유지한다.

3.3. DjVu의 장. 단점

DjVu 포맷의 장. 단점을 간단히 요약하면 아래와 같다.

3.3.1. 장점

○ 일반적으로 많이 사용하는 PDF는 디지털 문서전용 포맷에 장점이 있고 DjVu는 스캔된 원문 이미지 전용 포맷에 특히 이점이 있다.

○ 고해상도 컬러 원문 이미지의 경우 고품질 해상도를 유지하면서 압축률이 높아 파일 사이즈를 최소화 할 수 있고 원문 저장 장비의 확장비용을 절감할 수 있다.

○ 페이지 스트리밍이 지원되어 큰 용량의 파일이라도 필요한 페이지만 가져와서 보여주기 때문에 1-3초 이내 원문을 볼 수 있어 원문 서비스의 속도를 개선할 수 있고 네트워크 트래픽을 줄일 수 있다.

○ 사이즈가 큰 파일은 PDF 뷰어에서 열리는데 한계가 있지만 DjVu는 파일 사이즈에 관계없이 서비스 할 수 있다.

○ 원문보기에서 DjVu는 압축을 하여도 파일크기가 압축이 되고 해상도는 원문과 같이 유지하기 때문에 원문을 확대하여 보아도 이미지의 손실없이 확대/축소를 자유롭게 하여 볼 수 있다.

(2) 단점

○ 별도의 DjVu 전용 뷰어가 필요하다. PC에 DjVu 전용 뷰어가 설치되어 있지 않았을 때 몇 십초 이내에 자동으로 설치되지만 원문 컨텐츠에서 일반적으로 많이 사용하는 PDF 뷰어 이외의 또 하나의 뷰어를 사용하게 된다.

○ PDF 뷰어에 익숙한 사용자들은 DjVu 뷰어가 다소 불편하며, 표준화가 되어 있지 않다.

○ PDF-to-DjVu 변환 유틸리티는 PDF문서를 DjVu로 압축 변환하여 주지만 PDF 문서 내의 목차 정보를 추출하지 못하는 번거로움이 있다. (다음 업그레이드 버전에서는 지원 예정이라고 함.)

○ DjVu 포맷으로 서비스 시스템을 구축하기 위해서는 최소 100,000면 단위로 라이선스 비용을 지불하여야 한다.

○ DjVu 파일은 PDF와 같이 원문 파일을 사용자가 수정 편집할 수 없다. 책자 자료는 문제가 없지만 사진이나 슬라이드의 낱장자료는 필요에 따라 편집을 하여 사용할 수 있어야 한다. 이는 낱장 사진자료 적용에 있어 단점이 되기도 하지만 저작권의 보호측면에서 장점이 될 수도 있다.

4. 원문 이미지 포맷과 DjVu 포맷 비교

4.1. TIFF 포맷과 DjVu와의 비교

1980년에 CCITT 표준화 그룹에서는 소위 Bi-level 이미지(흑백 이미지)를 표현하기 위한 Group 3 표준 포맷을 정했다. 1984년에는 G3 표준을 좀더 향상 시킨 압축 권고안 Group 4가 발표되었다. 팩스 전송뿐만 아니라 디지털 도서관 분야에서 흑백 원문을 스캐닝 하여 이미지화하는데 가장 많이 사용하고 있는 포맷이 바로 G4 표준 압축 방식으로 표현되는 TIFF G4 타입이다.

TIFF G4는 단순히 흑백 이미지를 압축하는 표준 포맷일 뿐이다. TIFF 포맷 안에 PDF 문서와 같이 목차를 보여주는 책갈피 정보를 삽입하여 활용할 수 있게 하는 기능, 메타데이터를 이미지 안에 삽입하여 검색 시 활용하는 기능, 원문의 본문 내용을 검색할 수 있도록 텍스트 정보를 원문이미지와 같이 표현할 수 있는 기능 등을 지원하지 않기 때문에 이러한 기능을 위해서는 별도의 응용프로그램을 만들어서 제공해야 하는 문제점들을 갖고 있다.

다른 이미지 포맷과는 다르게 여러 장의 페이지를 하나의 파일에 번들로 저장할 수 있는 Multi-TIFF 기능을 제공하기도 하지만 위에서 언급한 기능들을 위한 응용프로그램에서는 이러한 Multi-TIFF 기능과 함께 활용할 수 없기 때문에 별로 효용성이 없는 것이 되어 버렸다.

TIFF는 표준 포맷으로 인식되어 있음에도 불구하고 익스플로러와 같은 표준 브라우저에서 직접 TIFF 문서를 볼 수 없기 때문에 별도의 TIFF 뷰어를 설치하여 이용하여야 하는데 TIFF 뷰어 기능도 표준화 되어 있지 않아 서비스 하는 기관이나 업체에 따라서

그 사용법과 서로 호환이 안 되는 문제점을 갖고 있기도 하다.

G4 표준이 발표된 이후 1993년에 Joint Bi-level Images Experts Group(JBIG)에서는 새로운 흑백 이미지 코딩 표준을 JBIG1 이름으로 발표하였으나 G4 표준보다 좋은 압축률을 갖고 있음에도 불구하고 G4 표준만큼 널리 보급되지는 못하였다. 2000년에는 G4 표준보다 약 3-4배 정도의 압축률이 좋은 JBIG2가 발표되었다. AT&T 는 당시에 JBIG2 표준에 근거하여 흑백 문서 이미지 압축 포맷을 개발 하였는데 이것이 DjVu JB2이다.

4.2. JPEG 포맷과 DjVu와의 비교

컬러 이미지를 저장하는 방식 중에 가장 보편화된 포맷이 JPEG이다. JPEG은 나름대로 압축 방식을 제공하기 때문에 컬러 이미지 파일을 표현하고 저장하는데 매우 적당한 포맷으로 인식되고 있다.

JPEG보다 압축률을 향상 시킨 JPEG2000 표준안이 2000년에 발표가 되었다. JPEG 압축 방식과 다른 Wavelet 알고리즘을 기반으로 하여 만들어진 표준안이다. 효율성이나 성능, 이미지 품질 면에서 JPEG보다 훨씬 좋은 것으로 평가되고 있으나 아직까지 JPEG2000이 보편화되지 못하고 있는 실정이다.

JPEG은 일반적으로 한 장으로 표현되는 사진과 같은 데이터를 표현하고 저장하는데 적당하다고 할 수 있으나, 컬러로 자료를 스캐닝 하여 여러 장을 하나의 책으로 묶어서 저장하고 서비스하는 구조에는 잘 맞지 않은 단점이 있다. TIFF 포맷과 같이 여러 장의 이미지를 한 개의 번들로 묶을 수 있는 Multi-TIFF와 같은 방식이 지원되지 않기 때문이다.

DjVu는 컬러 이미지 압축을 위해서 JPEG 2000과 같은 Wavelet 기반의 압축 방식인 IW44 방식을 이용한다. 압축률은 JPEG과 비교하여 같은 질을 보장하면서 그 크기는 JPEG보다 1/10 - 1/20 정도로 압축이 될 수 있다.

원본의 이미지가 고해상도의 이미지인 경우 DjVu로 압축된 이미지라고 할지라도 파일 크기가 클 수 있다. 그러나 웹에서 일반 브라우저에 플러그인 된 DjVu 전용 뷰어를 통해서 보게 될 경우 TIFF나 JPEG과 같이 전체의 파일을 모두 다운로드 받은 후 디스플레이 하는 방식을 이용하지 않고 점진적인 이미지 전송 기술을 이용하기 때문에 사용자는 원하는 이미지를 지연되는 시간없이 즉시 볼 수 있게 된다.

DjVu는 원문의 해상도와 상관없이 압축이 가능하고 서비스 시 DjVu 전용 뷰어는 사용자의 모니터의 해상도와 보는 화면의 크기에 맞게 자동으로 맞추어 주는 기능을 제공하기 때문에 고해상도의 원문 이미지도 서비스에는 전혀 문제가 되지 않는다.

○ JPEG(컬러 원문) 을 DjVu로 적용한 경우 [서울대학교 적용사례]

컬러로 스캐닝된 대부분의 원문들은(고문헌, 조선 근대 신문, 창간호 잡지, 의학 슬라이드 자료, 사진 및 필름 자료, 미술 작품집) 원본이 모두 JPEG 이미지로 압축, 생성되어졌고 이 JPEG 이미지를 DjVu로 다시 압축, 변환하였다.(이미지 상당부분 원본을 JPEG 포맷으로 압축한 것을 다시 DjVu로 압축하였기 때문에 평균 압축률은 1/3-1/4 정도 소요).



자료 유형	총페이지수	압축형식	용량	페이지당 평균크기
창간호잡지	824	JPEG	277 MB	336 KB
		DjVu	176 MB	213 KB
귀중본도서	1,228	JPEG	2,008 MB	1,635 KB
		DjVu	248 MB	201 KB

- JPEG 포맷으로 압축된 것을 다시 DjVu로 압축하여 평균 압축률은 1/3 ~ 1/4정도 소요됨

[그림-2]

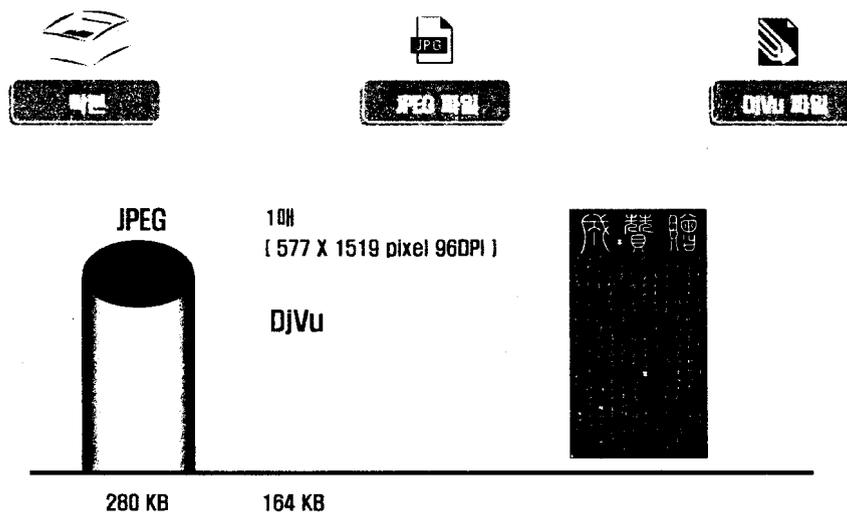
[그림-2]에서 창간호 잡지를 보면 총 824페이지 되는 고화질의 원본을 JPEG으로 표현하면 총 277 MB가 소요되나 DjVu로 압축을 하면 176MB가 된다. 또한 귀중본 도서를 보면 총 1,228 페이지의 JPEG으로 표현하면 2,008 MB가 되나 DjVu로 압축을 하면 248 MB가 된다. 웹에서 248 MB나 되는 DjVu 문서를 검색하여 보면 전체 248 Mb를 모두 다운로드 받아서 디스플레이 하지 않고 필요한 페이지만 선택적으로 스트리밍을 해서 보여지기 때문에 이렇게 큰 용량의 칼라 이미지를 지연 시간 없이 볼 수 있게 된다.

위의 자료들이 대부분이 DjVu가 소개되기 이전에 컬러로 스캐닝한 원문들로 고해상도로 원본을 스캐닝 했음에도 불구하고 이미 상당부분 웹 서비스를 위해서 DPI를 다운사이징한 상태였다. DjVu는 그 이후에 적용됨으로써, 이미 다운사이징 된 JPEG 이미지를

다시 DjVu로 압축한 결과를 만들었기 때문에 상대적으로 압축률(일반적으로 고해상도의 JPEG과 비교하여 1/10 1/20까지 압축 가능)은 높지 않았고 조금 더 고품질의 원문 서비스 제공의 기회를 얻지 못한 아쉬움이 있었다. [그림-3] 참조

○ JPEG(탁본 원문)을 DjVu로 적용한 경우 [서울대학교 적용사례]

우리 도서관에서는 탁본을 일반 카메라로 찍어서 JPEG으로 보관 및 서비스를 해왔으며, 이번에 DjVu로 압축 적용하였다. 원문을 저해상도의 일반 카메라로 찍은 것을 JPEG 이미지(흑백 그레이스케일 컬러)로 서비스 하고 있다. DjVu 로 압축 적용을 하여보니 원문의 JPEG 이미지 해상도가 좋지 않았기 때문에 , 압축 효율은 그리 높지는 않았다. (JPEG 보다 약 1/1.5 정도 압축) [그림-3] 참조



[그림-3]

처음부터 해상도가 좋은 장비로 이미지 작업을 하여 JPEG으로 서비스를 하였다면 고해상도를 유지하면서 압축 효율도 높아 DjVu 변환 시 좋은 방법이 되었을 것이다.

또 탁본은 다른 원문과 달리 목차 정보를 갖고 있지 않았기 때문에, 사용자들이 탁본의 원문을 보다 쉽게 볼 수 있게 하기 위해 DjVu 뷰어 안에 자동으로 썸네일 이미지를 자동 추출하여 좌측에 표시하는 방법을 이용하였다. 이러한 기능을 위해서 DjVu 전용 뷰어에 탁본 이미지를 보기 위한 고유 기능이(원문 서비스 시 디폴트로 좌측에 썸네일 이미지가 보여짐) 추가되어 개발 되었다.

4.3. PDF 포맷과 DjVu포맷 비교

PDF 포맷은 본래 디지털 문서들을 웹을 통해서 Publishing 하기 위한 솔루션으로 이용되어 왔다. 워드나 아래아한글 등으로 작성된 텍스트 기반의 문서를 웹을 통해서 배포하기 위해서는 PDF 포맷만큼 좋은 틀은 없을 것이다. PDF는 전용 Acrobat Reader를 무료로 사용하면서 누구나 쉽게 PDF 문서를 읽을 수 있는 장점이 있다. 뿐만 아니라 PDF 문서는 내부 문서를 수정할 수 없고 읽을 수만 있게 되어 있기 때문에 더욱 문서 배포 용으로는 매우 적당한 솔루션이다. 그러나 스캐닝한 이미지 문서를 PDF 포맷으로 저장하고 배포하는 데는 그 한계점이 있다.

스캐닝한 문서는 그 자체가 이미지 포맷으로 디지털화 된 텍스트 문서보다는 파일 크기가 몇 배나 클 수 밖에 없다. PDF는 기본적으로 압축을 기반으로 하지 않고 있기 때문에 스캐닝 된 이미지 문서를 PDF로 담기 위해서는 그 크기 이상의 파일을 요구하게 된다.

우리 도서관에서 컬러 PDF로 서비스를 하고 있는 원문은 없지만 이번 전자도서관에서 컬러 원문을 구축한 JPEG 이미지를 가지고 PDF로 변환 작업하여 DjVu와 비교하여 보았다.

○ PDF(흑백) -to- DjVu 변환 적용 [서울대학교 적용사례]

우리 도서관에서는 그동안 귀중본 중 일부를 마이크로필름으로 찍어 보관되어 왔던 것을 스캐닝하여 흑백 PDF로 서비스하고 있었다. PDF 포맷으로 저장된 귀중본은 모두 저해상도의 흑백 이미지로 되어 있었고, DjVu 포맷이 적용되기 이전에 PDF 문서 형태로 서비스 하기 위해서 만들어진 것들이었다. 대부분 한 개 파일의 크기가 평균 50MB가 넘었고 페이지 수도 200페이지 이상 되는 것이 대부분이었고, 파일의 크기가 200MB가 되는 것도 있었으며 페이지 수가 600 페이지가 넘는 귀중본도 많이 있었다. 한 볼륨의 크기가 커서 대부분 임의로 파일을 나누었음에도 불구하고 서비스는 사실 제대로 되지 못했다.

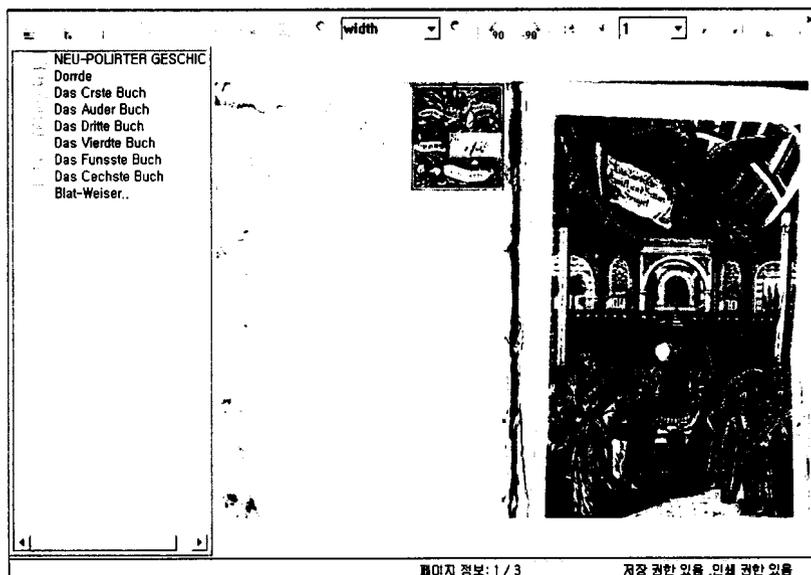


구 분	총페이지수	파일형식	용량	페이지당 평균크기
1건 용량	620	PDF	112 MB	185 KB
		DjVu	105 MB	173 KB
전체 용량	61,315	PDF	10.1 GB	173 KB
		DjVu	7.33 GB	125 KB

[그림-4]

[그림-4]에서와 같이 PDF-to-DjVu 변환 유틸리티에 의하여 모두 압축, 변환, 적용하였다. 원문이 상당 부분 저해상도의 흑백 이미지 형태를 갖고 있었기 때문에 DjVu로 압축되었을 때는 평균 약 1/1.5 ~ 1/2 정도 압축이 될 정도였다.

현재의 PDF-to-DjVu 변환 유틸리티는 PDF 문서 내의 목차정보를 추출하지 못하였기 때문에 (다음 업그레이드 버전에서는 지원 예정) PDF내의 순수 원문 이미지만을 DjVu로 변환하였고, 변환된 DjVu 문서를 보고 목차 정보를 재입력하는 방식을 취하였다. DjVu는 언어의 종속적이지 않기 때문에 일본어, 중국어, 러시아어 등의 PDF 문서를 DjVu로 변환하는데 전혀 문제는 없었다. [그림-5] 참조



[그림-5]

그 동안 페이지 수도 많고 파일 크기가 커서 웹 서비스가 거의 불가능하였던 흑백 귀중본의 PDF 문서를 다시 활성화 할 수 있었다.

5. 원문이미지 구축의 기본방침

5.1. 원문이미지 구축의 기본방침

위와 같이 PDF, JPEG를 DjVu 포맷으로 변환하면서 서울대학교는 원문 컨텐츠 구축에 있어 다음과 같은 기본 방침 아래 구축을 하고 있다.

○ 중요자료의 원본 보존

고문헌 같은 중요자료는 보통 마이크로필름으로 촬영하여 보존하지만, 한번 스캔한 자료를 원본상태와 거의 유사한 해상도로 스캔하여 원본 보존을 이미지로 대체한다.

○ 고품질 원문 구축

한번 구축한 원문의 재구축을 피하기 위하여 원문구축은 현재의 장비나 기술 중 가장 우수한 품질로 구축한다.

○ 원본 포맷과 서비스 포맷 동시 보존

이미지, AOD, VOD 등 구축되는 모든 컨텐츠는 고품질로 일차 원본을 구축하고, 서비스는 압축한 포맷으로 서비스한다. 따라서 원본 포맷과 서비스용 포맷이 달라지기도 한다.

○ 서비스 포맷의 유동적 대처

원본 파일을 고품질로 구축하고 보존하여 서비스 포맷의 기술이 발전되면 저장된 원본 파일을 새로운 서비스 포맷으로 변환하여 서비스한다. 예를 들면 컬러 원문 이미지의 경우 현재는 PDF보다 DjVu 포맷이 우수하여 DjVu 포맷으로 서비스하지만 PDF 포맷이 DjVu보다 더 우수하거나 새로운 우수한 포맷으로 업그레이드 될 경우 저장된 원본을 가지고 PDF 포맷이나 새로 개발된 우수한 다른 포맷으로 변환하여 서비스한다.

5.2. DjVu 포맷의 적용대상

원문 컨텐츠 구축에 있어 다음의 경우는 서비스 포맷에 한하여 DjVu로 변환하여 서비스한다.

5.2.1. 컬러 이미지 원문구축자료

컬러 이미지로 구축하는 원문의(귀중본, 조선근대신문, 창간호 잡지, 의학 및 미술 슬라이드 자료, 사진 및 필름자료, 미술작품집) 원본은 모두 JPEG 이미지로 보관하고 서비스 포맷은 DjVu로 한다.

- 대상자료

- 고문헌
- 조선 근대 신문
- 창간호 잡지
- 탁본 (Grayscale)
- 대학사료 (Grayscale)
- 슬라이드 자료
- 사진 및 필름자료
- 미술 작품집

5.2.2. 흑백 이미지 원문구축자료

흑백 이미지로 구축한 콘텐츠 중 파일 사이즈가 큰 다음의 콘텐츠는 DjVu 포맷으로 서비스한다.

- PDF 파일로 구축한 자료 중 사이즈가 큰자료

5.3. DjVu 포맷 적용방법

5.3.1. 원본 보정작업 후 일괄 변환

컬러 이미지를 600DPI JPEG로 스캔하면 사이즈가 커서 이미지 보정 작업에 어려움이 있다. 따라서 스캔은 원본 대상 크기에 따라 (600, 400, 300, 200 DPI) JPEG(압축율 10%)으로 이미지 보정이 필요한 것은 이미지를 보정한다. 이미지 보정이 완료되면 JPEG 자료의 종류별로 일괄 DjVu로 변환한다. DjVu로 변환 하는 방법은 구축 시 하나하나 변환 하는 방법이 있고, 대량의 데이터를 일괄로 변환하는 방법이 있지만 서울대학교에서는 고문헌, 슬라이드, 미술작품집 등 매체 구축별로 일괄 변환 하였다.

5.3.2. 메타데이터와 연계 및 TOC 북마크 생성

DjVu 변환 시 결여되었던 TOC 북마크 기능은 보완하여 생성하고 이용을 위해 DjVu로 변환 된 파일들과 별도로 구축된 각종 메타데이터(서지, 파일관리정보, 저작권관리정보)

와 연계는 구축자료 업로드 시 서로 제어번호로 연동 관계를 맺고 서비스한다.

5.3.3. DjVu 적용 해상도

○ 고문헌

원본대상 크기에 따라 원본을 (600, 400, 300, 200DPI) JPEG으로 스캔하여 보관하고, DjVu로 압축하여 서비스한다. 이때 해상도는 원래의 해상도를 유지하면서 파일크기만 압축이 된다.

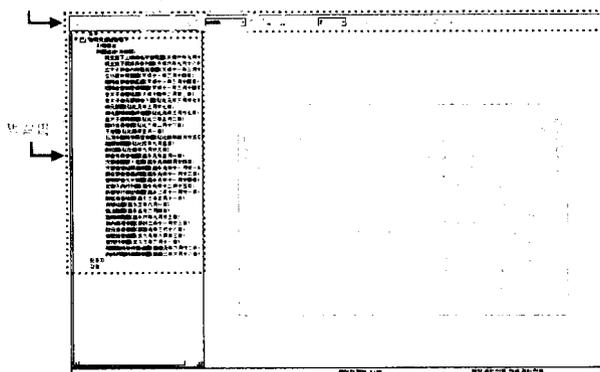
○ 슬라이드

슬라이드는 원본을 2400DPI JPEG으로 스캔 하여 600DPI로 다운사이징하여 원본을 보관하고 DjVu로 서비스한다.

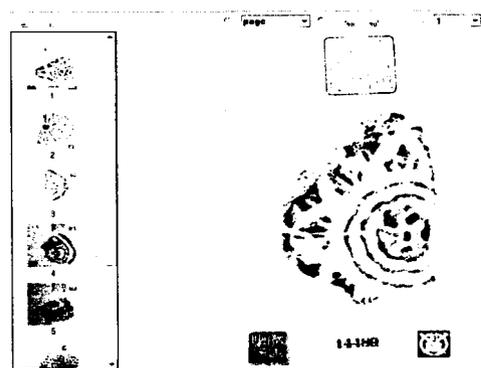
원본 보관용을 600DPI로 다운사이징한 이유는 보관용을 2400DPI로 저장하게 되면 파일 사이즈가 커지므로 저장장비의 무리가 되기 때문이다.

5.3.4. DjVu 서비스 화면

[그림 6-1]은 서울대학교 중앙도서관의 고문헌 중 웹에서 익스플로러 브라우저에 플러그인 된 DjVu 전용 viewer를 통해서 검색한 결과 샘플 화면이다. 상단에는 메뉴 툴 바가 있고 좌측에는 목차 정보를 책갈피 형태로 나타나게 하였다. 축소 확대를 자유롭게 할 수 있어서 원하는 부분을 자세히 볼 수도 있다. 이외에도 목차 정보가 없는 탁본과 같은 경우는 [그림 6-2]와 같이 viewer에서 좌측에 자동으로 썸네일 이미지로 출력하도록 하였다.



[그림 6-1]



[그림 6-2]

5.4. DjVu 포맷 적용 시 제기되었던 문제점

국내에서 처음으로 원문 서비스 포맷으로 DjVu를 적용하면서 DjVu의 표준 기능에서 제공하지 못하였던 기능들을 추가로 개발하게 되었다.

5.4.1. TOC Manager

TOC 파일의 북마크가 생성되지 않아 DjVu 문서 안에 자동으로 목차 정보를 삽입할 수 있는 유틸리티, 서울대학교 중앙도서관에서 서비스하는 TOC 파일을 읽어서 DjVu 문서에 목차를 삽입하는 기능이다.

5.4.2. 서울대학교 DjVu 뷰어

DjVu 표준 뷰어에는 목차 정보를 볼 수 있는 기능이 제공되지 않기 때문에 표준뷰어 기능 위에 목차 정보를 트리 형태로 볼 수 있도록 하는 서울대학교 DjVu 뷰어를 개발하였다.

5.4.3. 워터마크 기능

사용자가 DjVu 문서를 프린트 시 자동으로 서울대학교 로고 및 출력 일시를 삽입하며, 원문에는 손상을 가하지 않고 출력시 워터마크를 삽입하는 방식이다.

5.4.4. 프린트, 저장통제 및 스크린 캡처 방지 기능

프린트/저장 기능을 사용자의 권한에 따라서 통제할 수 있는 기능, 프린트 시 페이지의 범위를 지정하여 출력하는 기능, 스크린 캡처 방지 기능 등이 있다.

5.4.5. CMS와 연동 자동 변환기

서버에서 운영되는 CMS(Content Management Server)와 연동하여 등록되는 문서를 자동으로 DjVu 문서로 변환하고 모니터링 할 수 있는 기능이다.

5.4.6. 탁본을 위한 썸네일 이미지 출력 기능

목차 정보가 없는 탁본 이미지를 화면에 보여줄 때는 썸네일 이미지를 기본으로 보여주는 기능이다

위와 같은 문제들을 개발하여 문제점을 해결하였으나 저작권 보호를 위한 DRM은 아직 적용되지 않는다.

5.5. 기대효과

컬러 원문 이미지와 흑백 원문 이미지 의 일부를 DjVu로 적용하여 고품질 해상도로 원문을 서비스 할 수 있다. 원문보기에서 파일 사이즈에 관계없이 원문을 빠른 속도로 볼 수 있고 다수의 이용자가 원문을 이용하더라도 네트워크의 트래픽을 최소화 할 수 있을 것으로 기대된다. 또한 대량의 원문 콘텐츠를 구축하면서 원문 저장장비를 1/3로 축소할 수 있었다.

6. 맺음말

일부 국내 도서관에서는 1990년 중반부터 원문 이미지를 구축하기 시작하였고, 1990년대에 구축한 원문 이미지 포맷은 주로 TIFF였으나 익스플로러와 같은 표준 브라우저에서 직접 TIFF 문서를 볼 수 없기 때문에 서비스 하는 기관이나 업체에 따라서 그 사용법과 서로 호환이 안되는 문제점을 갖고 있다.

그 후 PDF가 보편화 되면서 PDF 포맷으로 원문 콘텐츠를 구축하여 서비스하고 있지만 문서 중심의 PDF로는 고품질로 다양한 유형의 원문 콘텐츠를 서비스 하기에는 현재로서는 한계가 있다. 서울대학교는 이러한 한계를 극복하기 위하여 컬러 이미지 서비스 포맷을 DjVu로 적용하였다. 국내에서 처음으로 컬러 원문 이미지 서비스에 적용하다 보니 여러 가지 시행착오를 겪은 부분도 있었다.

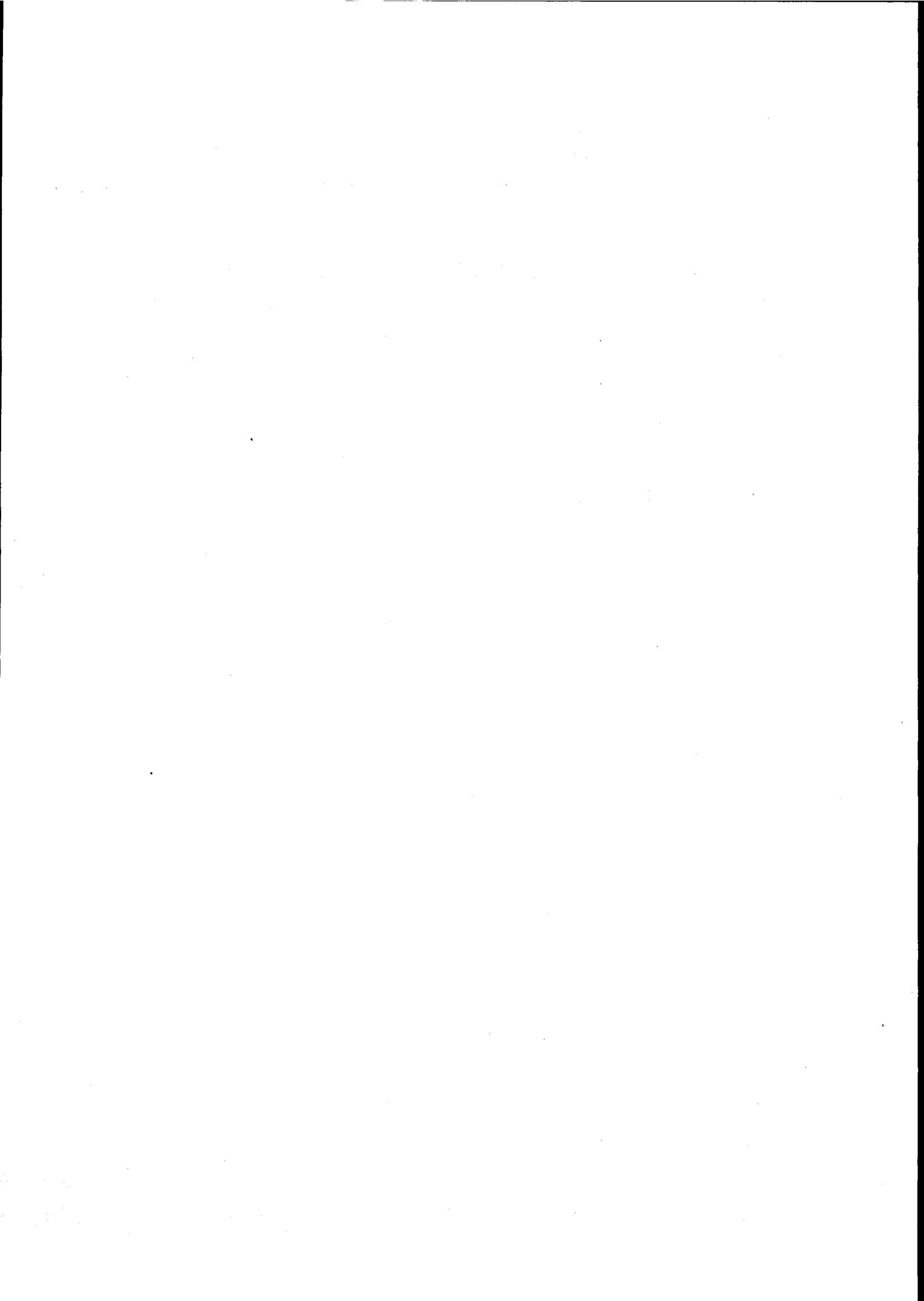
특히 사업 초기부터 DjVu를 선택한게 아니기 때문에 원문 보관을 위해서 해상도가 높은 버전과, 서비스를 위한 낮은 해상도 버전 따로 이원화하여 저장하였던 것을 DjVu로 적용하다 보니 DjVu의 특성을 최대한 활용하지 못한 부분이 있었고 원문 컬러 이미지를 고품질 해상도로 서비스에 그대로 적용하지 못한 아쉬움도 있었다. 2003년에 구축하는 원문들은 초기 단계에서부터 고품질의 원문 이미지를 생성 할 수 있도록 할 것이다.

앞으로 발표될 Acrobat Reader 6.0에서는 어떤 기술이 업그레이드 되어 출시될지 모르겠지만, 그동안 큰 사이즈의 원문 파일들이 PDF에서 열리지 않는 문제점을 해결할 수

있을지 는 지켜보아야 할 것이며, 컬러 원문 이미지의 고품질 서비스는 당분간 DjVu가 우수하여 이 포맷을 적용할 수밖에 없었으나 좀더 사용하여 보아야 할 것이다.

DjVu도 향후 DRM과 PDF 문서내의 목차정보를 추출할 수 있는 기능, DjVu 뷰어 기능의 표준화, 서비스하는 기관이나 업체에 따라서 서로 호환이 안 되는 문제점을 해결할 수 있다면 좋은 포맷으로 사용될 수 있을거라고 기대한다. 다만 DjVu 포맷 적용 시 원본 파일과 서비스 파일 포맷을 따로 하여 원본 파일을 보관하는 것이 기술발전예 적응하는 방법일 것이다. DjVu나 또 다른 우수한 포맷이 나온다면 원본에 파일포맷을 적용하여 변환하는 것이 바람직 할 것이다.

아울러 서울대학교에서 DjVu 기술 적용을 통해 시행착오를 겪었던 것들을 바탕으로 현재 전자도서관을 구상하고 있거나 추진하고 있는 많은 대학도서관과 주요 기관들에게 조금이나마 도움이 되었으면 하는 바람이다.

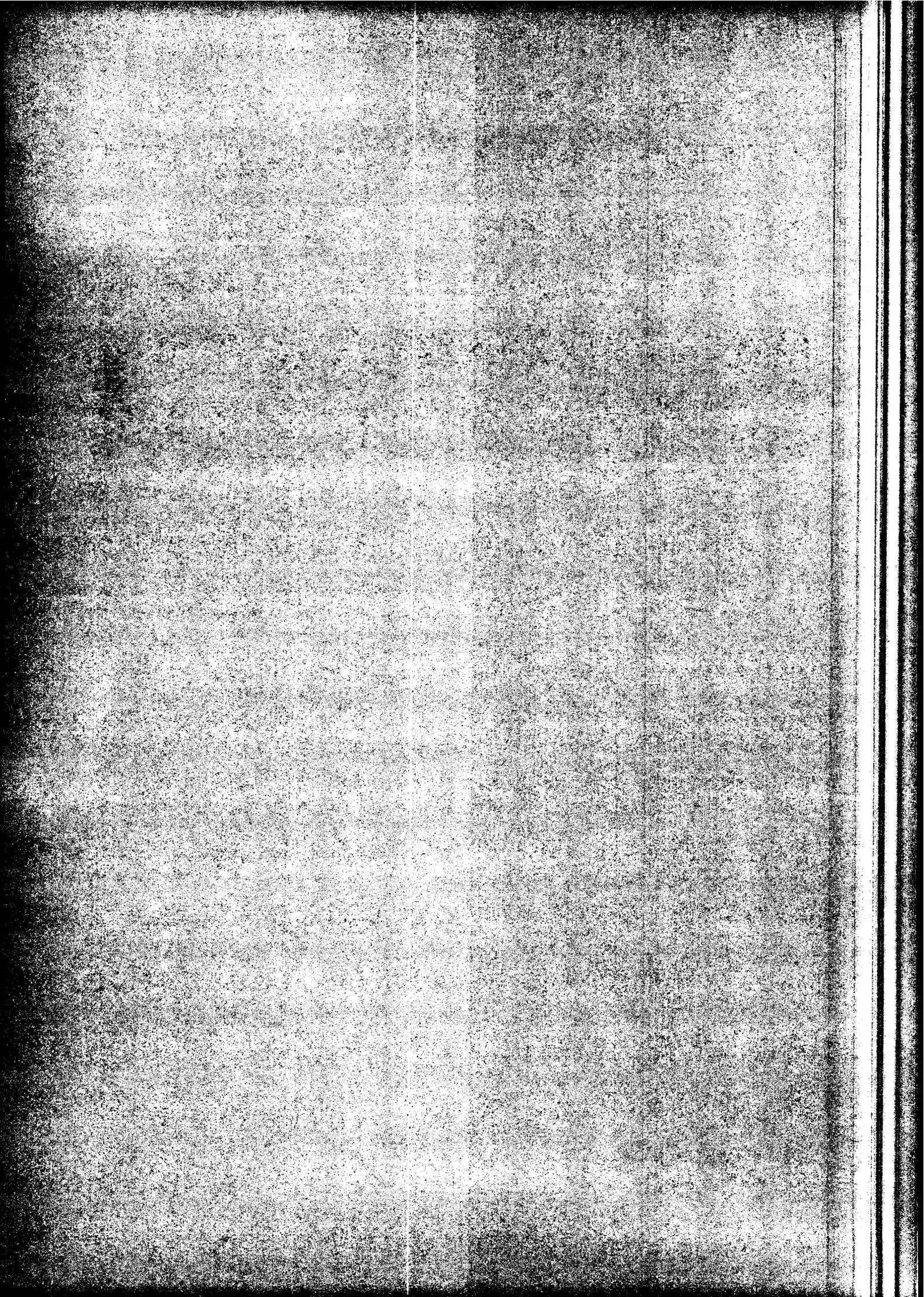


[400000]

2003. 5

박 전 민

prais@snu.ac.kr



< 사례발표 8 >

서울대학교전자도서관의
저작권 보호 적용범위 및 정책

2003. 5.

박진만
(praise@snu.ac.kr)

차 례

1. 저작권 보호 개요	1
2. 저작권 보호 시스템 적용 사례 (서울대학교 중앙도서관)	2
2.1 저작권 보호 방법	2
2.2 저작권 정책 정	4
2.3 저작권 관리 시스템 구성도.....	7
2.4 시스템 운영환경 / 하드웨어 구성도.....	13
2.5 기타사항.....	13
2.6 문제점 및 보완사항	14
3. 맺음말 (워터마킹 및 DRM 업체 선정시 유의사항).....	15
3.1 워터마킹	15
3.2 DRM	15

1. 저작권 보호 개요

오늘날 온라인 서비스 이용자의 폭발적인 증가와 멀티미디어 및 네트워크 기술의 발전으로 다양한 디지털 자원에 대한 서비스 환경이 요구되고 있습니다. 디지털 형태의 음악, 동영상, 출판물, 문서, 정보, 도면, 소프트웨어 등의 디지털 자료에 대한 수요와 활용이 급격하게 증가하고 있으나, 이러한 디지털 콘텐츠는 품질의 손상 없이 복제 및 활용이 매우 용이하기 때문에 저작권의 불법 활용과 무단 복제의 확산으로 이용되고 있으며 이에 따라 디지털 저작권 및 소유권 보호의 중요성이 증대되고 있습니다.

그 단적인 예로서 최근 국내 및 해외 음반 산업의 사건들이 디지털 정보와 관련된 저작권 관리 및 보호의 심각성을 잘 보여주고 있습니다.

심지어는 내용에 인용한 참고문헌을 기록하지 않았다고 학위가 취소되는 경우도 발생하고 있으니 개인의 저작물에 대한 보호 마인드가 얼마나 확산되고 법제화 되어 가고 있는지 알 수 있습니다.

이러한 추세에 따라 디지털 저작물에 대한 체계적인 보호와 관리를 위한 정책 수립 및 기술개발을 필요로 하고 있습니다. 기존의 서비스 방식에서는 디지털 저작물의 검색과 다양한 서비스에 초점이 맞추어져 실제 저작권과 관련된 관리 및 보호 시스템이 준비되어 있지 않았습니다. 보다 많은 저작물들이 디지털 과정을 거쳐 이용자들에게 서비스되고 있는 가운데 체계적인 디지털 저작권 관리 시스템 및 저작물들에 대한 보호 기술의 필요성이 세계적으로 부각되고 있습니다.

* 저작권 관리 시스템(DRM : Digital Rights management)

- 음악공유서비스로 유명한 냅스터가 2001년 MP3 저작권 보호를 위하여 채택한 것이 시초. 온라인 콘텐츠가 유료화되며 중요한 기술로 떠올라 미국 매사추세츠 공과대학에서는 미래10대 핵심정보기술로 선정
- 저작물 제공 기관의 서비스 방법에 따라 다양한 운영 원칙으로 설정하여 저작물에 대한 체계적인 관리와 보호를 제공
- 디지털 콘텐츠의 무단 사용을 막아 저작권 관련 당사자들의 이익과 권리를 보호해 주는 기술과 서비스

2. 저작권 보호 시스템 적용 사례 (서울대학교 중앙도서관)

서울대학교 중앙도서관은 불법복제와 변조방지를 위해 워터마킹 기술과 보안인증시스템을 뒷받침으로 하고 있다.

- 워터마킹 : 기밀정보를 디지털 데이터에 숨긴 후 저작권 분쟁이 발생했을 때 디지털 저작권자가 누구인가를 확인
- 보안인증시스템 : 권한이 있는 사용자만 콘텐츠를 이용

2.1. 저작권 보호 방법

2.1.1. 원문 이미지(학위논문, 대학간행물 등 PDF 파일)

2.1.1.1. Watermark

- 암호화 : 저작권 분쟁 발생시 사용, Invisible 워터마킹
- Stamp : 가시적인 저작권 정보 제공(프린터 종속성 배제)
인쇄시에 서울대학교 로고가 보이도록 처리

2.1.1.2. Application 으로 제어

- 보기/저장/출력 기능
- 사용기간 제한(년/월/일)
- 이용장소 제한(관내/학내/학외/기타 지정장소)
- 신분별 제한(교수/대학원/학부생...)
- 개인별 제한(보기/저장/출력)

2.1.1.3. DRM(Digital Rights Management)

- 콘텐츠 자체 암호화 (복사 방지)
- 콘텐츠 접근권한 인증

2.1.2. 원문 이미지(고문헌 자료 등 Djvu 파일)

2.1.2.1. Watermark

- Stamp : 가시적인 저작권 정보 제공(프린터 종속성 배제)
인쇄시에 서울대학교 로고가 보이도록 처리

2.1.2.2. Application 으로 제어

- 보기/저장/출력 기능
- 사용기간 제한(년/월/일)
- 이용장소 제한(관내/학내/학외/기타 지정장소)
- 신분별 제한(교수/대학원/학부생...)
- 개인별 제한(보기/저장/출력)

2.1.2.3. DRM(Digital Rights Management) : 적용 불가

2.1.3. 멀티미디어 콘텐츠(학술행사, 음악자료 등 비디오/오디오 파일)

2.1.3.1. 스트리밍 서비스(asf 파일로 서비스) : 다운방지

2.1.3.2. 버퍼링 캡처 방지 : 서비스 중 캡처 방지

2.1.3.3. Application 으로 제어

- 보기만 가능(필요시 고품질, 고음질 원본 파일 제공)
- 사용기간 제한(년/월/일)
- 이용장소 제한(관내/학내/학외/기타 지정장소)
- 신분별 제한(교수/대학원/학부생...)
- 개인별 제한(보기/저장/출력)

2.1.3.4. DRM(Digital Rights Management)

- 콘텐츠 자체 암호화 (복사 방지)
- 콘텐츠 접근권한 인증

※ 워터마킹 : Windows Media Encoder 9 버전에서는 콘텐츠를 인코딩시 처리 가능

2.1.4. 워터마킹 및 DRM 적용 시기

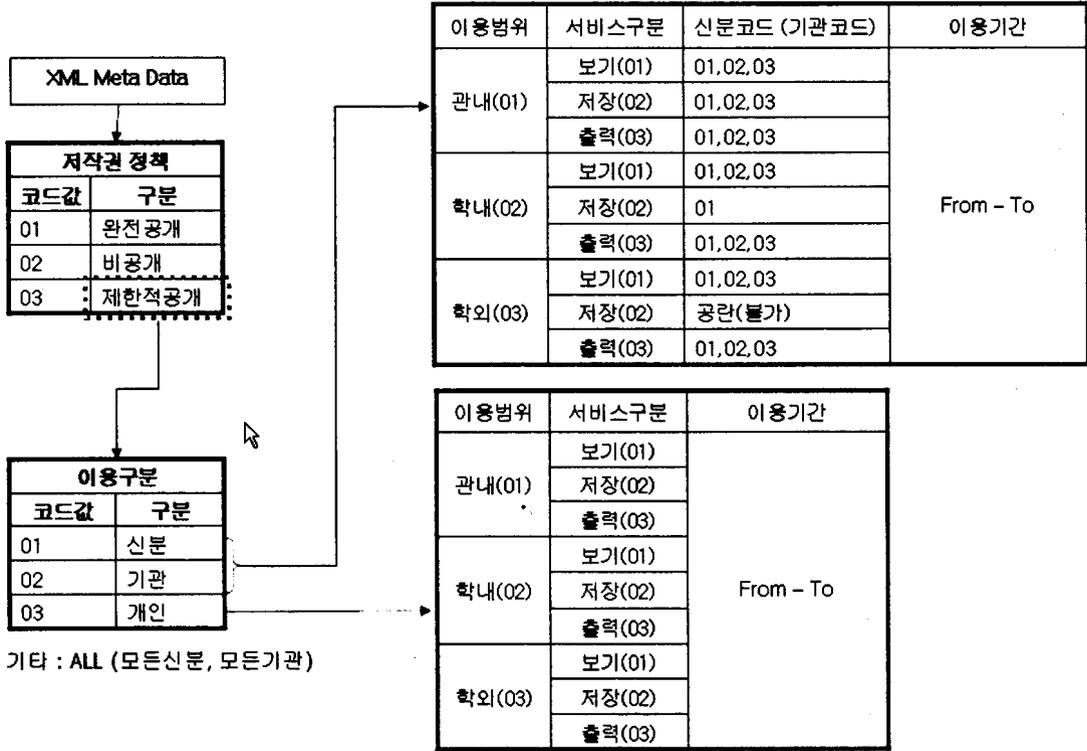
2.1.4.1. 워터마킹, DRM을 적용 후 서비스 (업체에 따라 방법적 차이)

단 데자뷰(Djvu)의 워터마킹 처리는 서비스 시점에 적용

2.1.4.2. 일괄 배치(Batch) 처리로 대량의 자료에 작업 가능

2.2. 저작권 정책 결정

저작권 정책



저작권.xls

A	B	C	D	E	F	G	H	I	J	K	L	M	N
구분번호	서비스구분	신문코드	기관코드	이용범위	서비스구분	신문코드	기관코드	이용범위	서비스구분	신문코드	기관코드	이용기간	구분번호
0000000001	01	01,02,03	01,02,03	01	01	01,02,03	01,02,03	01	01	01	01	20020101	20021231
0000000002	01	01,02,03	01,02,03	01	01	01,02,03	01,02,03	01	01	01	01	20020101	20021231
0000000003	01	01,02,03	01,02,03	01	01	01,02,03	01,02,03	01	01	01	01	20020101	20021231
0000000004	01	01,02,03	01,02,03	01	01	01,02,03	01,02,03	01	01	01	01	20020101	20021231
0000000005	01	01,02,03	01,02,03	01	01	01,02,03	01,02,03	01	01	01	01	20020101	20021231

이용구분이 개인(03)인 경우

술라스 신문코드값 (clr_user_type.user_type)

술라스 기관코드값 (clr_user_dept.user_code)

8자리 Ex) 20020202

저작권_개인.xls

A	B	C	D	E	F	G	H	I	J	K	L	M
매체제어번호	학번(주민등록번호)	관내	저장	출력	학내	저장	출력	학외	저장	출력	이용기간	종료일
000000000003	1234	1	1	1	1	0	1	1	0	1	20020101	20021212

<예시 1> - 콘텐츠별 권한 설정 요청 양식

이용제한	이용자그룹(가)	이용서비스(나)	이용기간(다)	이용기관(라)
완전공개(A)	<input checked="" type="checkbox"/> 모든그룹	<input checked="" type="checkbox"/> 전 부	<input checked="" type="checkbox"/> 영 구	<input checked="" type="checkbox"/> 모든기관
비 공개(F)	X	X	X	X
제한공개 (밑줄 부분은 자세히 기술)	<input checked="" type="checkbox"/> 학부생(1)	<input checked="" type="checkbox"/> 보 기(1)	<input checked="" type="checkbox"/> 기 간(1) (yyymmdd- yyymmdd)	<input checked="" type="checkbox"/> 소속기관(1)
	<input checked="" type="checkbox"/> 석사과정(2)	<input checked="" type="checkbox"/> 저 장(2)		<input checked="" type="checkbox"/> 타기관(2) (복수선택가능)
	<input checked="" type="checkbox"/> 박사과정(3)	<input checked="" type="checkbox"/> 출 력(3)		
	<input checked="" type="checkbox"/> 졸업예정자(4)			
	<input checked="" type="checkbox"/> 휴학생(5)			
	<input checked="" type="checkbox"/> 교 수(6)			
	<input checked="" type="checkbox"/> 강 사(7)			
	<input checked="" type="checkbox"/> 연구원(8)			
	<input checked="" type="checkbox"/> 직 원(9)			
	<input checked="" type="checkbox"/> 법인소속(B)			
<input checked="" type="checkbox"/> 일반이용자(C)				

* 아래의 예시를 참고하여 표기부분과 기술 할 것. (콘텐츠명-표기)

컨 텐 츠 명		이용자그룹(가)	이용서비스(나)	이용기간(다)	이용기관(라)
스포츠댄스 동영상	예시	<input checked="" type="checkbox"/> 모든그룹	<input checked="" type="checkbox"/> 보 기	<input checked="" type="checkbox"/> 기 간	<input checked="" type="checkbox"/> 소속기관
	표기	(가A, 나1, 다1<020808-020908>, 라1)			
애국가 동영상	예시	<input checked="" type="checkbox"/> 모든그룹	<input checked="" type="checkbox"/> 전 부	<input checked="" type="checkbox"/> 영 구	<input checked="" type="checkbox"/> 모든기관
	표기	(A) : 모두 완전공개일때			
보안모듈 성능분석 세미나	예시	<input checked="" type="checkbox"/> 교 수	<input checked="" type="checkbox"/> 보 기	<input checked="" type="checkbox"/> 기 간	<input checked="" type="checkbox"/> 모든기관
		<input checked="" type="checkbox"/> 강 사	<input checked="" type="checkbox"/> 저 장		
		<input checked="" type="checkbox"/> 박사과정	<input checked="" type="checkbox"/> 출 력		
		<input checked="" type="checkbox"/> 석사과정			
		<input checked="" type="checkbox"/> 학부생	<input checked="" type="checkbox"/> 보 기	<input checked="" type="checkbox"/> 기 간	<input checked="" type="checkbox"/> 소속기관
				<input checked="" type="checkbox"/> 타기관	
표기	(가2367, 나123, 다1<020808-020821>, 라A), (가1, 나1, 다1<020808-020821>, 라12<공과대학>)				

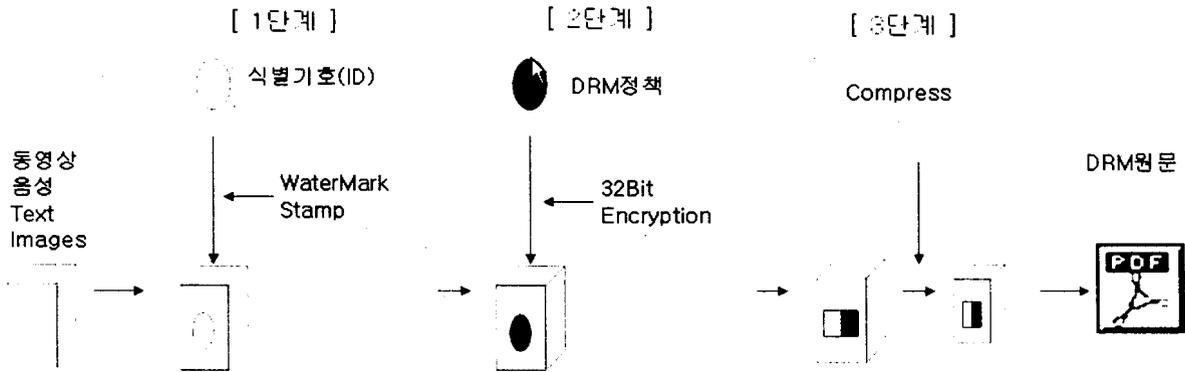
<예시 2> - 기관별 콘텐츠 권한 신청 현황

기 관	공개여부	범 위	권 한	종 류	
음악대학	완전공개	학내, 학외	보기	음악자료	VOD /AOD
미술대학	"	학내, 학외	보기	미대자료	
인문학연구소	"	학내, 학외	보기	인문학연구소	
간호대학	"	학내, 학외	보기	간호학과	
건축학과	"	학내, 학외	보기	건축학과	
의류학과	"	학내, 학외	보기	의류학과	
대학사료	"	학내, 학외	보기	대학사료	
기 관	공개여부	범 위	권 한	종 류	
종교학과	제한공개	학내	보기	종교학과	대학간행물
비뇨과	"	의대, 병원	보기	의학자료	
미술대학	"	학외	보기	학위논문	
해양시스템공학	"	학외	보기	연구보고서	
지진공학연구센터	"	학외	보기	단행본	
나머지기관	"	학내	보기, 저장, 출력	학술지	
대학기록관(2,997)	"	대학기록관	보기, 저장, 출력	대학기록관	서울대학교
한국교육사고(399)	"	한국교육사고	보기, 저장, 출력	한국교육사고	기록사료

<예시 3> - 저작권 정책 결정(특별한 언급이 없는 한 학내는 관내 포함)

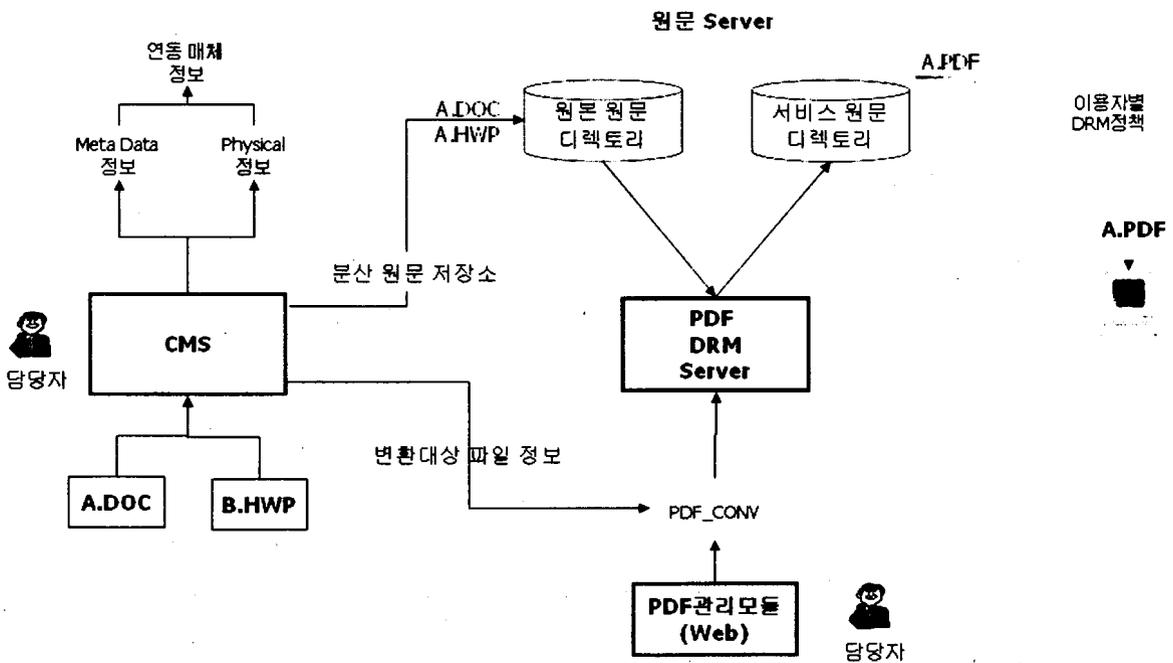
종 류	건 수	공개여부	범 위	권 한	파일포맷	
학 위 논문	69,000	제한공개	학내 학외	보기, 저장, 출력 보기, 출력	pdf	
		제한비공개	특허, 출판 관련된 자료			
학술지(연속간행물)	280	제한공개	학내 학외	보기, 저장, 출력 보기, 출력	pdf	
단행본(연구보고서포함)	1,845	"	"	"	pdf	
음영 M/F 기사색인	67	"	관내	보기	pdf	
교수연구업적보고서 (기사색인)	4,446	완전공개			pdf	
고 문 헌 이시	귀중본	514	제한공개	학내 학외	보기 보기	djvu
	귀급도서					
	창간호 잡지		"	"	"	djvu
	조선근대신문	83,623	"	"	"	djvu
	탁본	670	"	"	"	djvu
	고지도		"	"	"	djvu
의 대 자료		제한공개	학내	보기	djvu	
		개인별공개	개인별 권한부여			

2.3. 저작권 관리 시스템 구성도

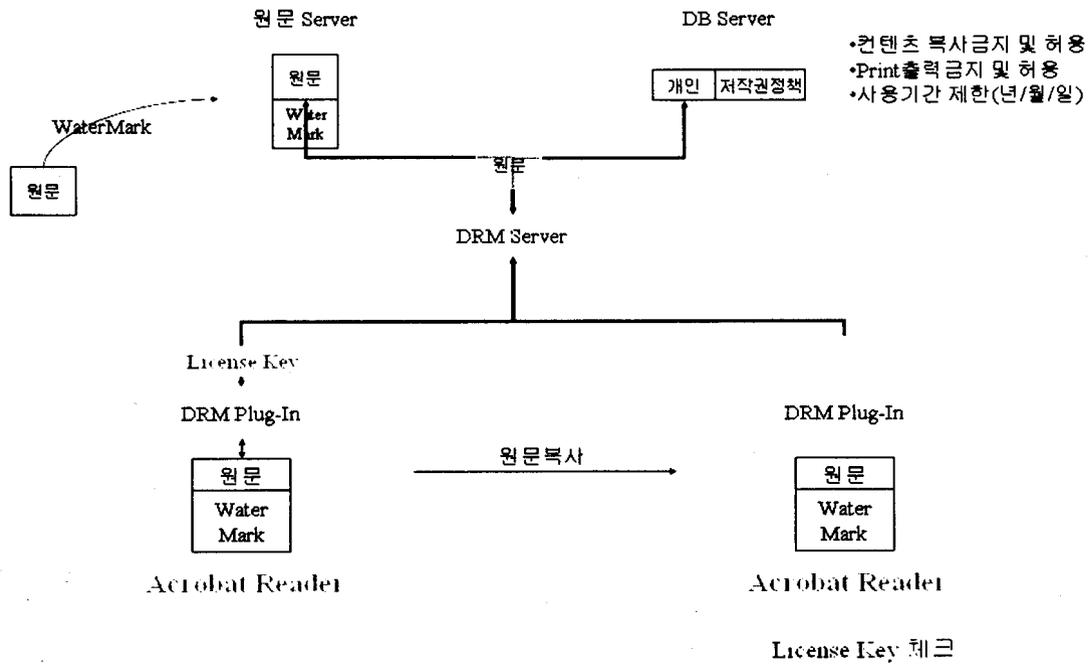


* 콘텐츠에 워터마크 삽입 및 암호화 방법

원문 작업 흐름도(PDF)



원문 DRM 서비스 구성도



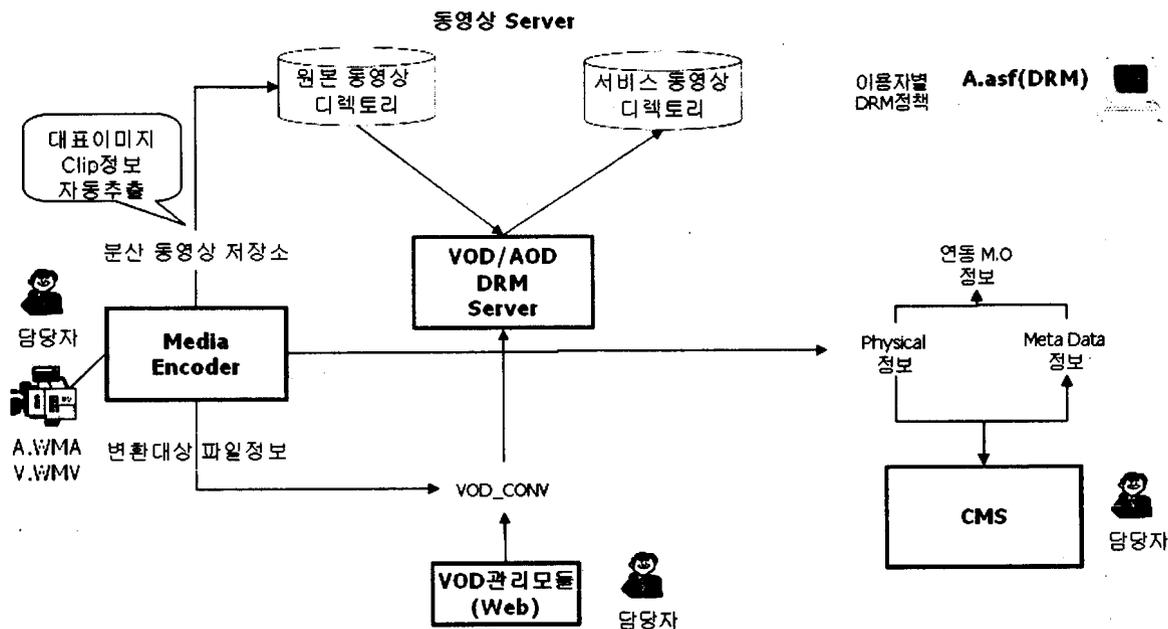
* 로그인 인증 과정

1. 암호화 된 PDF를 사용자의 PC에서 열면 자동으로 브라우저를 통하여 인증 받을 URL로 연결되어 인증 시작
2. 이미 로그인 인증을 거친 경우: 연결된 사용자 연결정보(아이디)를 사용하여 인증 시도
인증을 거치지 않은 경우: 사용자 로그인 화면을 열어 사용자 인증
3. 이용자의 해당 컨텐츠 이용 권한을 DBMS에서 확인
4. 우선권한리스트의 이용자나 그룹에 속했는지 먼저 찾아 확인 후 다른 권한 리스트를 검색하지 않고 설정된 이용권한을 적용
5. 우선권한리스트 내에 이용자가 속하지 않았을 경우, 권한리스트 내에서 이용자가 속한 그룹을 모두 찾아내고 OR연산을 통하여 가능한 이용 권한을 획득
6. 이용자 PC의 CPU-ID 값을 통해 이용권한에 맞는 인증키를 생성 후 이용자에게 전달
7. 인증 완료 후 인증 관련 로그 저장
(이용자 ID, 컨텐츠 ID, 이용권한, CPU ID, IP, 인증시간 등을 저장)
8. 완료

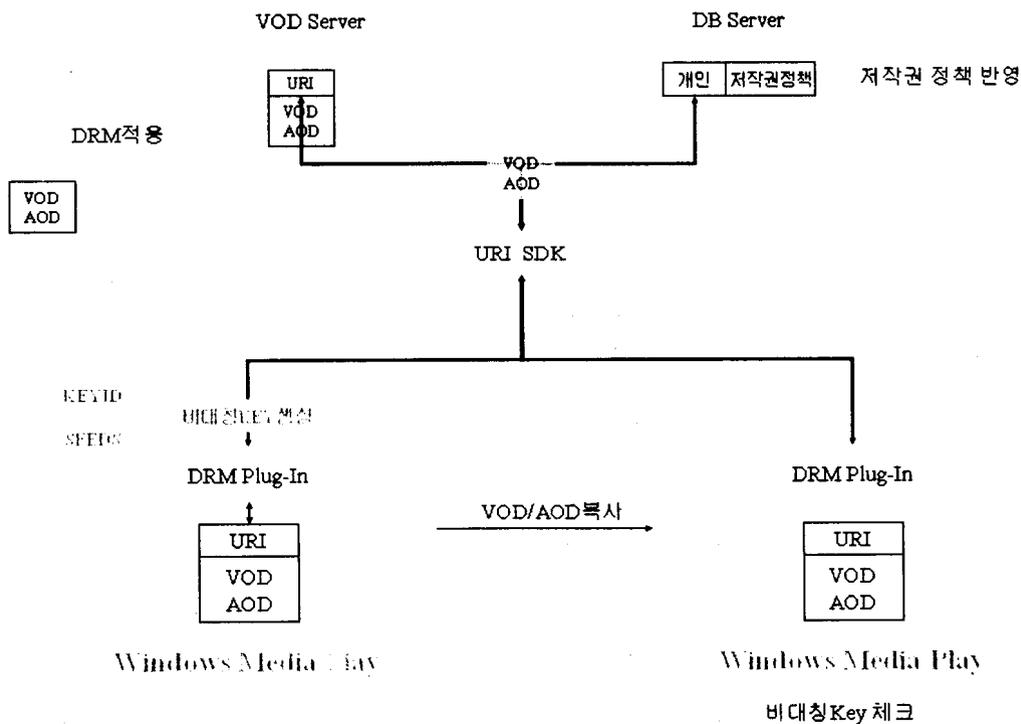
<참고>

- 이용자의 PC에는 Adobe Acrobat Reader(v5.0)가 설치되어 있어야 함.
- CPU-ID를 인증키 생성시 활용하므로 다른 PC에서는 사용 불가
- 암호화 된 PDF파일과 인증키(Rmf) 파일이 동시에 있으면 정상적으로 내용을 볼 수 있으며 양쪽에 모두 없는 경우에는 재인증을 받아야 함
(단 자동으로 My Documents 폴더에 백업을 생성하여 됨)

멀티미디어 DRM 적용



멀티미디어 DRM 서비스 구성도



Microsoft Internet Explorer
 주소 http://sdl.snu.ac.kr/index.jsp
 서울대학교 전자도서관

로그인

전자도서관소개 통합검색 DL Contents 전자저널 개인학술연구실 My DL 도서관서비스 게시판

검색 방법 : 연속질문 키워드 완전일치

전체 digital convergence 결과내검색

자료 유형 : 전체 단행본 연간물 학위논문 기사 학술행사 기록자료
 E-Journals 의학자료 미술작품 음악자료 개인학술연구물 E-Book
 지도서 웹문서

검색어: digital convergence

전체선택 이전화면 1건중1건 출력 1/1

서명	저자	출판사	출판년
<input type="checkbox"/>	Digital Convergence - Your Dreams & Challenges/머수관/	2003-03-28/	

라이선스 업데이트중

페이지보기

보기 바꾸니보기 다시쓰기

- 미디어 플레이어 보안 모듈을 연
- 개발할 대화상자가 나타지면 확인
- 최소버전을 누르면 라이선스동
- 네트워크 상태에 따라 20초정도

라이선스 발급완료

Microsoft Internet Explorer

SOC 신화외 System 설계

• SOC 신화
 • SOC는 무엇인가?

00:34:10.25 - 00:36:50.25 - 00:39:00.25

© Title : Digital Convergence - Your Dre...
 © Performer : 머수관
 © Subject :
 © Place : 서울대학교 신공학관
 © Date : 2003년 03월 28일

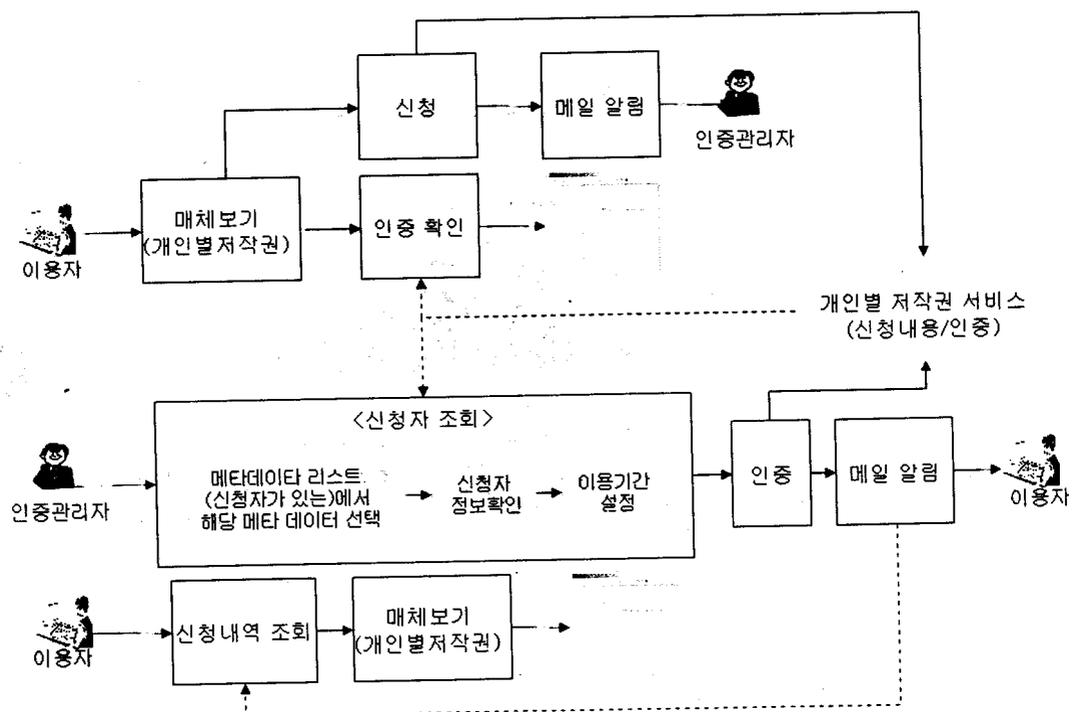
* 멀티미디어 DRM 적용 예

AOD/VOD 저작권관리시스템

관리기능 선택		DRM 패키징 관리					
131	00000071105	중세의 미술람	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
132	00000071106	중세 힌두교	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
133	00000071108	탄트리즘	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
134	00000071109	중세불교의 복합성	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
135	00000071111	중세 도교1	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
136	00000071112	송대의 성리학	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
137	00000071114	명청대의 유교	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
138	00000071116	중세 도교2	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
139	00000071117	스텝과 비단길	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
140	00000071118	중세 종교의 회교	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
141	00000071120	한국종교의 역사적 개관	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
142	00000071122	한국종교2	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
143	00000071124	과학지식과 종교의 신념	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
144	00000071125	기복과 사회보장 제도	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
145	00000071126	지구촌의 종교인들	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
146	00000071128	세속화 이후의 지구촌	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
147	00000071129	세계 종교사를 돌아보며	VOD	VOD 유관자료 없음	science_of_religion...	SUCC	┘
148	00000071986	Retroperitoneoscopic..	VOD	VOD 유관자료 없음	urology_operation_00..	SUCC	┘
149	00000071988	Molecular Mechanisms..	VOD	VOD 유관자료 없음	urology_operation_00..	SUCC	┘
150	00000072099	98-국대안-고병익	AOD	AOD 유관자료 없음	98_Gobyungik.asf	FTPFail	┘

* 웹 관리자용 화면 제공

개인별 저작권 업무 흐름도



2.4. 시스템 운영환경 / 하드웨어 구성도

- 운영체제: Windows 2000 Server
- DB 서버: Oracle 9i
- 웹서버: IIS 5.0
- 필요 소프트웨어: MS WMRM SDK7.1

구분		장비명	O S	C P U	RAM	Disk	서비스명	설치모듈
전자도서관	전자보안	Dell 6400	Windows 2000	900MHz × 4개	4GB	72GB	VOD서버	AOD/VOD DRM 저작권 관리기
	멀티미디어	Dell 6400	Windows 2000	900MHz × 4개	4GB	72GB	VOD 변환서버	AOD/VOD DRM 변환기 & AOD

2.5. 기타 사항

2.5.1. 플러그인 및 구동 프로그램

- 이미지 파일(PDF) : Adobe Acrobat Reader 5.0 이상 버전
 - ※보안 플러그인이 포함되지 않은 프로그램에서는 정상적으로 내용을 볼 수 없음
 - ex) Adobe Acrobat Reader 5.1 Standard 버전의 경우
- 이미지 파일(Djvu) : 작은 크기의 전용 플러그인 설치
- 비디오/오디오 파일 : Windows Media Player 이용
 - ※ Windows Media Player 라이선스 획득시
 - 라이선스 비용 : 무료
 - 단 기관 인증 필요 : 약 400\$ 정도
 - 절차가 상당히 까다로움
 - 이점 : 앞으로의 Microsoft사의 가격정책에 유연하게 대처 가능
신뢰성 있는 사이트로 인증
 - 해외에서는 S/W 개발비용에 보통 포함됨

2.5.2. 통계

- 기간별 인증로그
- 사용자별 이용 횟수 및 콘텐츠 분류
- 콘텐츠별 이용 횟수 내역
- 시간대별 인증 횟수
- IP별 인증 횟수 및 내역

2.6. 문제점 및 보완 사항

2.6.1. 워터마킹 처리된 경우

- 프린터에 따라 인쇄가 제대로 안되는 문제
(특히 오래된 행망프린터의 경우)
- 인쇄 속도가 약간 감소 (파일 크기가 큰 경우 문제 발생 가능성)
- 페이지 수가 많은 PDF 파일의 경우 워터마킹 처리시 많은 시간 소요

2.6.2. PDF 파일의 경우 Acrobat Reader 5.0 이상 버전 설치 요망

- 보안 플러그인이 포함되지 않은 프로그램에서는 정상적으로 내용을 볼 수 없음
ex) Acrobat Reader 5.1 Standard 버전의 경우
- Acrobat Reader의 설치 및 버전을 체크하는 OCX 보완 필요

2.6.3. PDF 및 Djvu 파일 서비스 시 화면 캡처 방지 기능

- 윈도우 기본 키의 경우만 방지 가능 (Alt + Print Screen)
- 상용 캡처 프로그램의 경우 캡처 가능

2.6.4. Djvu 파일의 경우

- 고용량 콘텐츠(고문헌자료 등)에 사용시 속도 향상 (고압축률 제공)
- DRM 적용 불가
- 전용 뷰어 설치

2.6.5. 원본 파일을 서비스 해야 할 경우

- JPG, Wav, MP3 등은 DRM 적용하기 어려움

2.6.6. 학외에서 로그인 사용자의 경우 학내이용자의 권한으로 처리

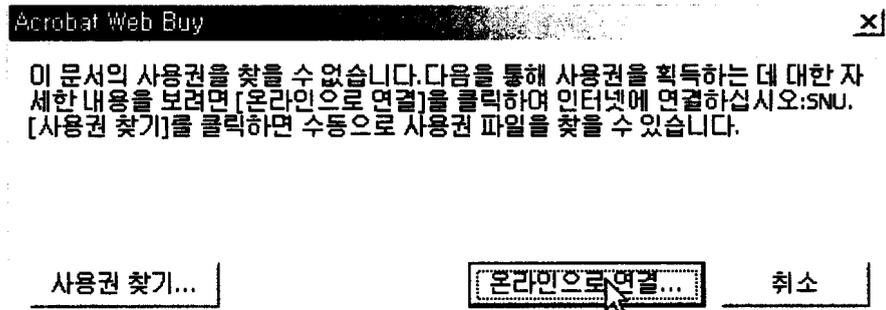
- Proxy 서버 참고 구현

2.6.7. PDF 파일 이용시 시스템 시간과 이용자 PC 시간의 차이가 크면 에러 발생

- Acrobat 6.0의 경우에 처리 가능 (여름 출시 예정)

2.6.8. 인증시 나타나는 화면의 메시지 수정 불가능

- Acrobat 프로그램 자체 내장



3. 맺음말 (워터마킹 및 DRM 도입시 유의사항)

3.1. 워터마킹

- 워터마킹 처리 시점 유의(업체마다 방식 상이)
 - * 워터마킹 일괄 처리 후 서비스
 - : 서비스 시간 단축
 - : 워터마킹 위치, 크기, 로고 변경시 일괄 작업 필요
 - * 서비스 시점에서 워터마킹 처리
 - : 약간의 시간 지연
- 파일 크기 증가 여부(저장 공간의 충분한 확보)
- 프린트 소요 시간의 증가 여부
- Visible / Invisible Watermarking 선택 (가격 차이)
- 원문내용 판독에의 영향 여부 검토 (로고 위치, 크기, 투명도 조정)
- 워터마킹 처리시 소요 시간 고려 (페이지 수에 따른 소요 시간 증가)

3.2. DRM

- DRM 업체에 따라 지원하는 포맷형식이 상이 (지원 포맷 확인 필요)
- Application으로 제어 가능한 기능 파악
- 오프라인 상태에서 인증 여부
- PDF 파일의 경우 Acrobat 지원 여부

- PDF 파일의 경우 Fast Web View 적용 불가
- 비디오/오디오 경우 Windows Media Player 지원 여부
- 비디오/오디오 경우 버퍼링 캡처 방지 여부
- 비디오/오디오 경우 경우에 따라 스트리밍 서비스로 대체 가능
- 암호화 방식 (암호화 키 확인)
- P2P(Peer To Peer : 인터넷에서 이루어지는 개인 대 개인의 파일 공유 기술 및 행위), B2C(Business To Customers : 기업 대 소비자의 전자상거래) 지원 여부

VOD/AOD 파일 구축형태

Audio

구분	WAV	MP3	ASF(WMA)
용도	보관용	보관용	서비스용
Bit rate	256Kbps	128Kbps	64Kbps

Video

구분	MPEG2	ASF(WMV)
용도	보관용	인터넷 스트리밍 서비스 용
Bit Rate	4Mbps	300Kbps/500Kbps
초당 프레임 수	30 Frame	24 frame
화면 Size	640 * 480	320 * 240

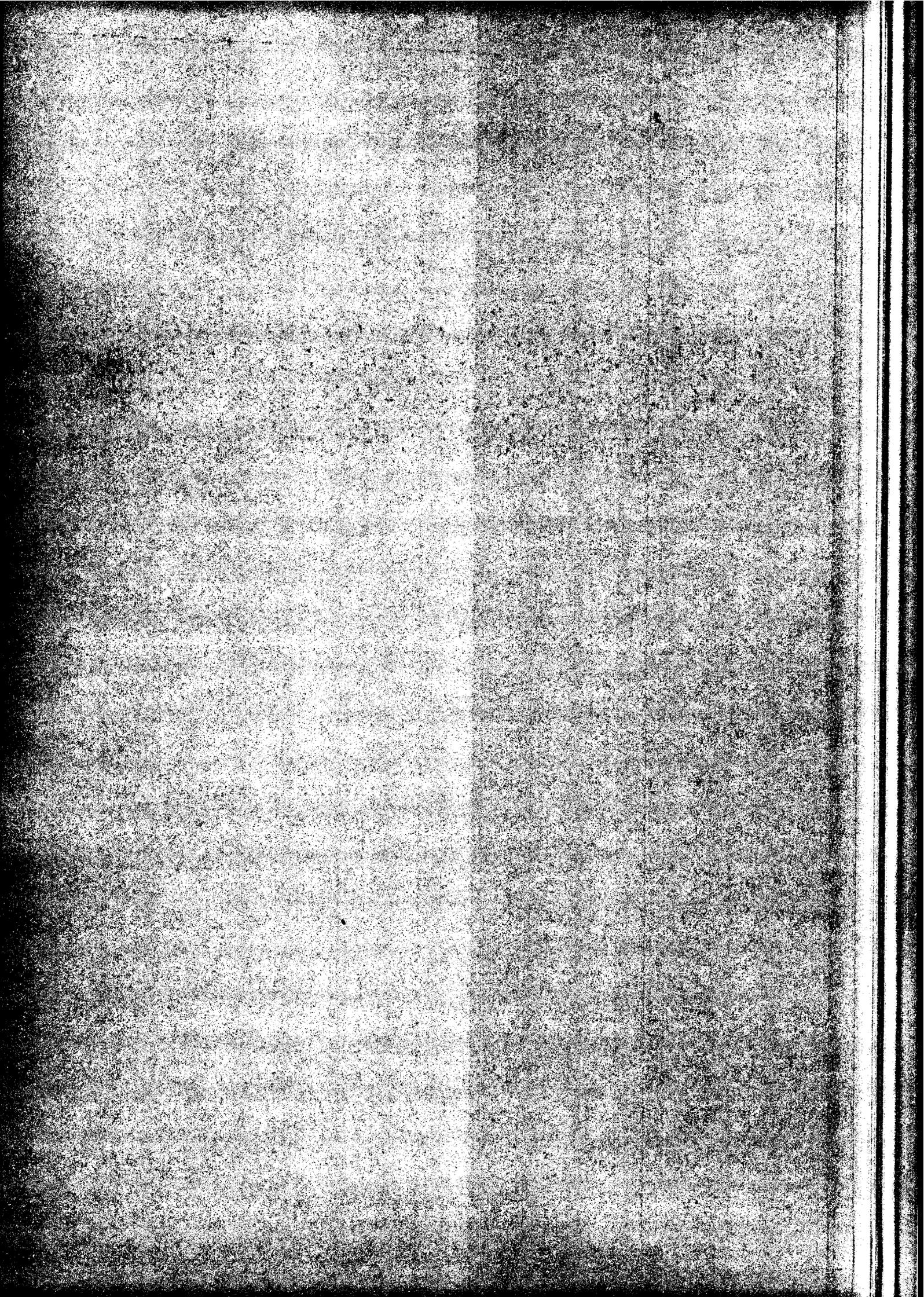
3.3. 검토 업체

한빛소프트(구 EasySolution, AcroSolution), 메타라이츠(MetaRights),
엔피아(Enpia), 아르파(ARPA), MarkAny, 대우정보시스템, 파수닷컴,
유니닥스(UniDocs)

[400000]

2003. 5

김희배
khi@nu.ac.kr



< 사례발표 9 >

서울대학교전자도서관
학위논문 자동변환 및 원문 링크정보 공유

2003. 5.

김 화 택
(kht@snu.ac.kr)

차 례

1. 학위논문 온라인 제출시스템 개발 배경.....	1
2. 학위논문 온라인 제출시스템 소개	1
3. 학위논문 온라인 제출시스템의 문제점	7
4. 학위논문 온라인 제출시스템 보완점	8
5. 학위논문 원문정보 공동 이용체제 현황.....	8
6. 서울대학교 학위논문 원문링크 서지데이터 제공	9
7. 학위논문 원문정보 링크 서지데이터 상호이용 공동협력	12
8. 학위논문 원문 링크 서지데이터 교환 방법.....	12
9. 원문 링크 서지 교환 포맷 및 상세 안내.....	13
10. 추진배경 및 목적.....	16
11. 추진방향.....	16
12. 세부 추진계획	16
13. 대상기관 및 추진일정	18
14. 서비스 구성도	19
15. 기대효과	19

1. 학위논문 온라인 제출시스템 개발 배경

1.1. 학위논문 PDF 서비스 속도 개선

1) 1998년도부터 2002년까지 외부업체를 통해서 학위논문 PDF변환작업을 하였기 때문에 논문파일 및 논문책자 제출 이후 PDF서비스까지 약 3~4개월정도의 시간이 소요되는 것을 학위논문 온라인 제출시스템 개발 이후 1개월 이내 웹을 통한 학위논문 PDF서비스가 가능해짐

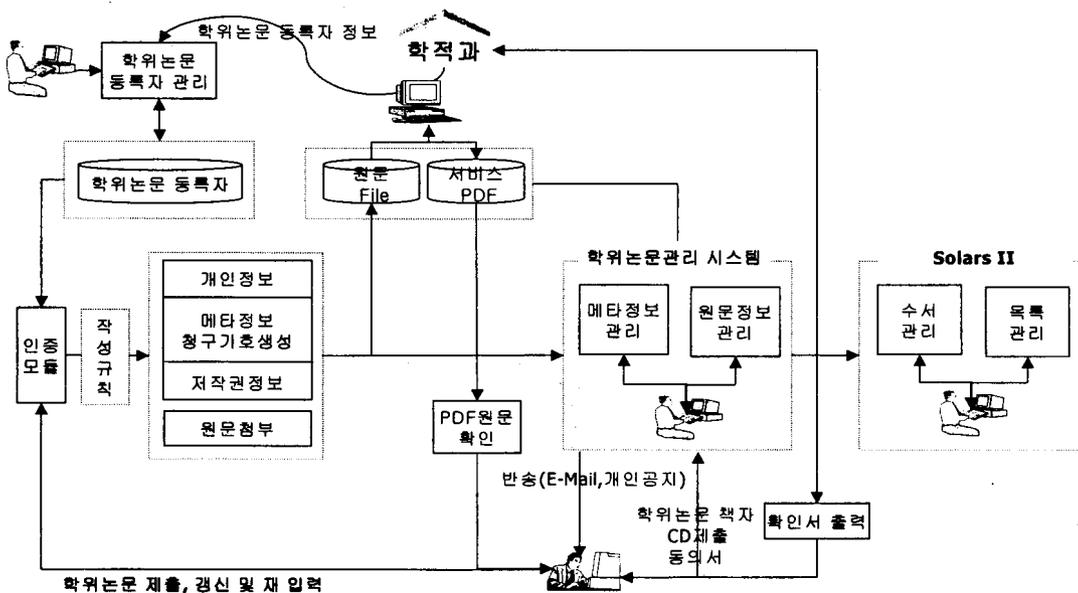
2) PDF서비스 중 오류(그림 누락)가 있거나, 논문제출자가 내용을 수정하여 PDF를 다시 변환하고자 할 경우 PDF구축업체에게 요청하여 수정해서 다시 서비스하는데 시간이 많이 소요되나 자체 PDF 변환시스템이 설치되면, 바로 수정한 뒤 PDF 변환이 가능하여 이용자에게 빠른 서비스가 가능

1.2. 학위논문 PDF 구축 비용 절감

- 매년 2번(1학기, 2학기)씩 3500여건의 학위논문 PDF 변환을 위해 구축업체에게 지불되는 비용이 절감

2. 학위논문 온라인 제출시스템 소개

2.1. 업무 흐름도



2.2. 업무 순서

2.2.1. 학위논문 제출대상자 등록

- 학적과에서 학위논문제출대상자를 엑셀파일로 받아 등록자관리 시스템에서 등록 처리하며 수정 및 삭제, 추가가 가능함

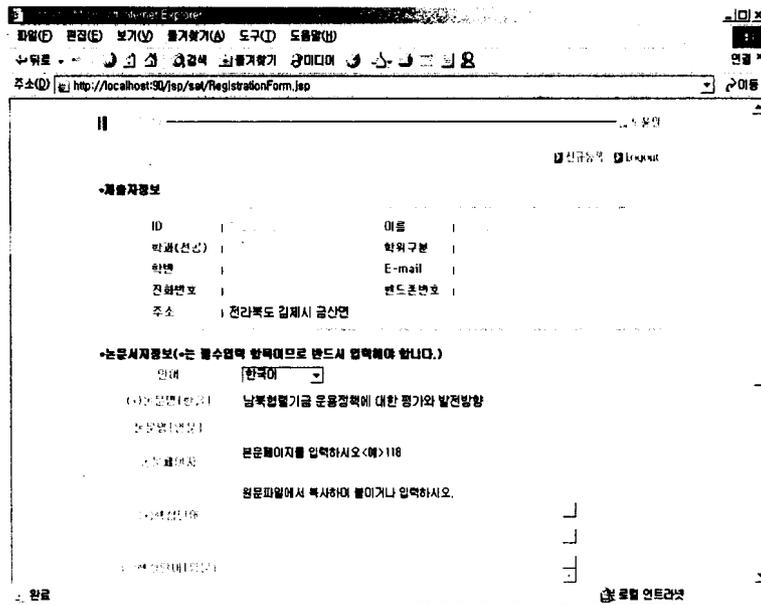
구분	학번	성명	주민등록번호	지도교수명	논문명	대학명	대학코드	학과명	학과코드	학위구분
1	99921-645	조상문	7403062140811	장정길	한국 지방정부 행정대학원	920		행정학과 정책학	921B	석사
2	97921-651	최형민	7008241816618	정종덕	한국 노인장위 행정대학원	920		행정학과 정책학	921B	석사
3	2001-22793	장혁교	7712021274214	김준기	정부와 NGO의 행정대학원	920		행정학과 정책학	921B	석사
4	99921-636	이승훈	7110041110313	이달준	인종 다양성 행정대학원	920		행정학과 정책학	921B	석사
5	2001-22763	서정미	7104161930416	임도빈	정책자원의 행정대학원	920		행정학과 정책학	921B	석사
6	2000-23061	생시영	7926181080218	김광종	제16대 국회의 행정대학원	920		행정학과 정책학	921B	석사
7	2001-22789	임재진	7408011350111	최병선	미국 부시행정 행정대학원	920		행정학과 정책학	921B	석사
8	2000-23112	장지은	6308281786013	정종덕	정책의재행정 행정대학원	920		행정학과 정책학	921B	석사
9	97921-611	김종화	6711041126522	최종철	정책미래대상 행정대학원	920		행정학과 정책학	921B	석사
10	2001-22751	김정민	7806112636112	최종철	북한이탈주민 행정대학원	920		행정학과 정책학	921B	석사
11	2001-22749	김종태	7007131406310	이달준	주5일 근무제 행정대학원	920		행정학과 정책학	921B	석사
12	99921-642	장광조	7305151534138	김병섭	책임운영기관 행정대학원	920		행정학과 정책학	921B	석사
13	2001-22770	오영민	7808111939413	오영민	정책수용해 행정대학원	920		행정학과 정책학	921B	석사
14	99921-788	이은영	7404162647918	김병섭	정부위원회 행정대학원	920		행정학과 정책학	921B	석사
15	2000-23102	박동주	6811201988627	최종철	정보통신시장 행정대학원	920		행정학과 정책학	921B	석사
16	2000-23065	권승석	7102061037919	김종민	도표교를향미 행정대학원	920		행정학과 정책학	921B	석사
17	94921-601	권용택	6806241269610	김준기	BS 2TV 민영 행정대학원	920		행정학과 정책학	921B	석사
18	99921-752	고승호	6410151953411	김동진	지방자치단체 행정대학원	920		행정학과 정책학	921B	석사
19	2000-22980	김혜정	7402072094017	김병섭	정부출연연구 행정대학원	920		행정학과 행정학	921A	석사
20	2001-22719	정성훈	7902271622511	이달준	마동록 이주 행정대학원	920		행정학과 행정학	921A	석사
21	2001-22731	허준영	7403311898816	정정길	개발형 자위 행정대학원	920		행정학과 행정학	921A	석사
22	2001-22691	문지은	7612022636726	방석현	정보화사업 행정대학원	920		행정학과 행정학	921A	석사
23	2001-22722	장혜진	7603062691415	김광종	한국과 일본의 행정대학원	920		행정학과 행정학	921A	석사
24	99921-703	김광종	7008101396626	김준기	기초자치단체 행정대학원	920		행정학과 행정학	921A	석사
25	2000-23002	이만수	7902071654212	임도빈	주민자치센터 행정대학원	920		행정학과 행정학	921A	석사

2.2.2. 학위논문 온라인 제출시스템 로그인

- 학위논문 제출대상자만 주민등록번호, 학번, 이름을 입력하여 로그인 가능

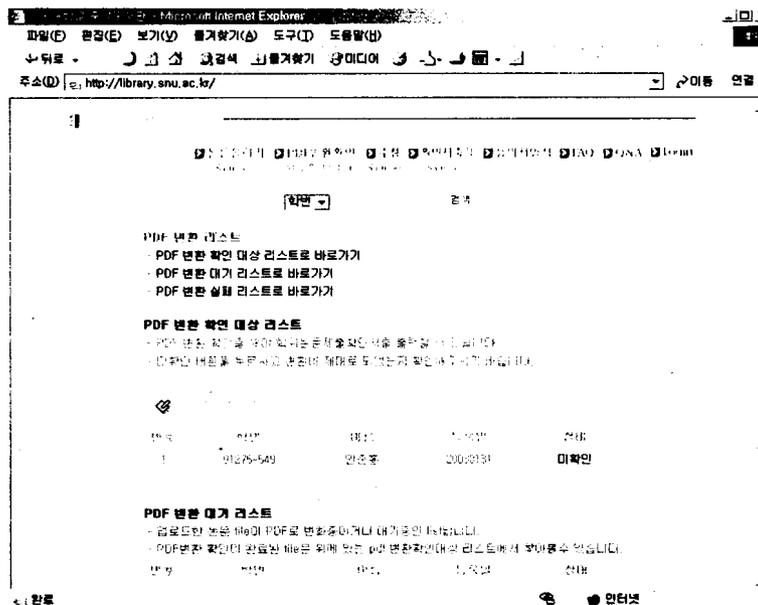
2.2.3. 학위논문 정보 입력 및 논문파일 올리기

- 제출자 정보 및 논문 정보(서명, 저자, 초록, 주요어, 목차 등)를 입력하고 원문 파일을 올리면 시스템이 자동으로 PDF파일로 변환



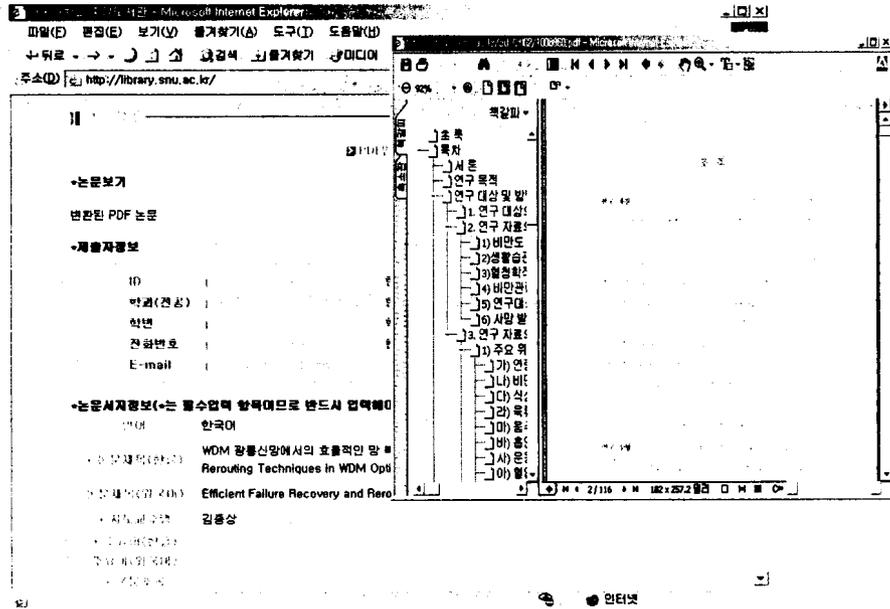
2.2.4. PDF 변환대기리스트 확인

- 온라인제출시스템은 PDF 변환이 완료되면 제출자의 메일로 변환완료 안내메일이 발송되며, 그 전에는 제출자가 직접 변환대기리스트에서 순서를 확인할 수 있음



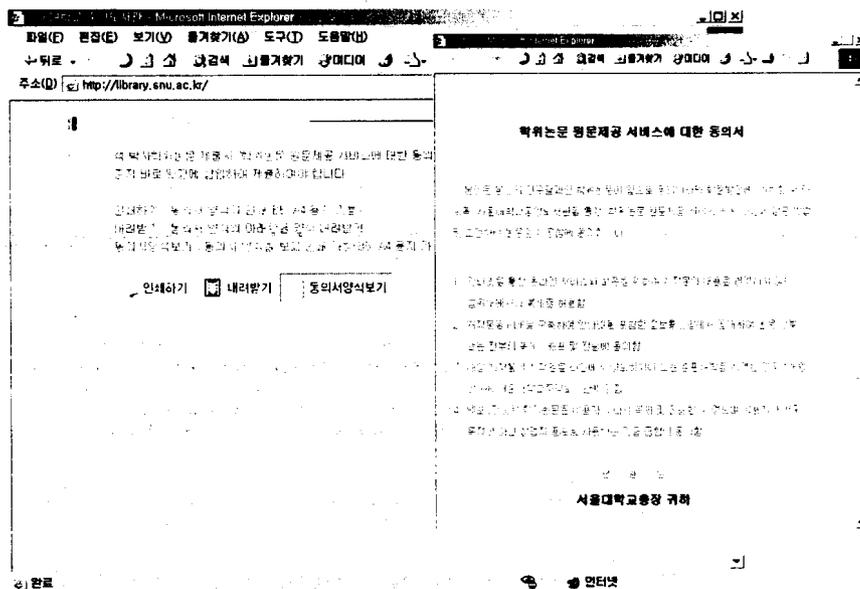
2.2.5. PDF 변환 확인 및 수정

- PDF변환이 완료되면 제출자는 논문파일이 PDF로 정확히 변환되었는지를 확인하고, 잘못된 경우 논문파일 정보 수정 및 논문파일을 다시 올림



2.2.6. 학위논문 원문 제공 서비스에 대한 동의서 출력

- 학위논문 원문제공 서비스에 대한 동의서 1부를 출력하여 논문책자 인준지 바로 뒷장에 삽입하여 논문책자를 제출하여야 함



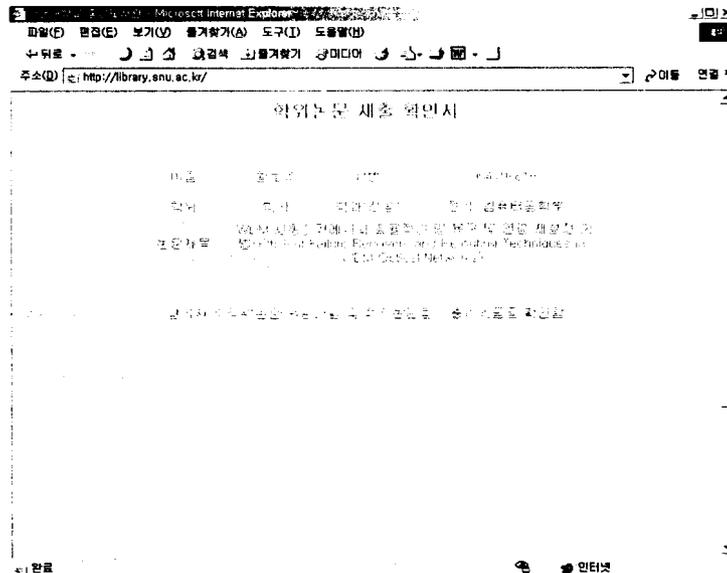
2.2.7. 논문책자 제출

- 학위논문 책자 제출시 담당자는 학위논문관리자시스템에서 제출자의 PDF파일과 대출정보를 확인함

제출자(학부)	제출자	논문제목	학위구분	학과	학번	지도교수	출발	접수	발행	인출	대출	대출학인
1	1392 최형일	위 P27수용체가 발현된 HB2294 세포	석사	화학부	95305-578	최영민	✓	✓	✓	✓	✓	✓
2	1960 박지현	지급순환포(Flow of Funds)를 이용한 석사	석사	경제학부	2001-20263	이합용	✓	✓	✓	✓	✓	✓
3	2269 오주환	동종간기 부속 광디스크 드라이브의 석사	석사	기계항공공학부	2001-21104	신호철	✓	✓	✓	✓	✓	✓
4	2051 서동현	특이성 단백질 효소의 전기적 특성의 석사	석사	응용화학부	2001-21301	이종희	✓	✓	✓	✓	✓	✓
5	2272 윤세연	이연방인쇄 구상요소와 거시신경망의 석사	석사	경영학과	2000-20912	곽수근	✓	✓	✓	✓	✓	✓
6	2055 이강현	카로신 비드를 이용한 공유결합법과 석사	석사	응용화학부	2001-21321	김병기	✓	✓	✓	✓	✓	✓
7	2373 이관용	외환위기에 후 경제학 부가치(CVA) 석사	석사	경영학과	2000-20917	곽수근	✓	✓	✓	✓	✓	✓
8	3044 이은희	한강철강의 경제인용성해 관한 연구 석사	석사	외국어	2001-22287	윤영규	✓	✓	✓	✓	✓	✓
9	1755 장우주	개에서 salactose 초음파 조영제를 석사	석사	수의학과	2001-22225	윤영희	✓	✓	✓	✓	✓	✓
10	1528 최지은	결합성 결핵균에서 TSC1 및 TSC2 유전자	석사	의학과	9591-499	황용순	✓	✓	✓	✓	✓	✓
11	3147 박한배	IWA 고분 초전도 변압기의 권선 설 석사	석사	컴퓨터공학부	2001-21461	한승민	✓	✓	✓	✓	✓	✓
12	2520 박경민	QoS를 제공하는 디스크 스케줄링 프 석사	석사	컴퓨터공학부	2001-21446	고건	✓	✓	✓	✓	✓	✓
13	2076 하현규	미리코로 핵융합 동향에 관한 석사	석사	응용화학부	2001-21343	장태석	✓	✓	✓	✓	✓	✓
14	2581 장진호	음식물쓰레기 전처리 방법이 산 발 석사	석사	지구환경시스템공학부	2001-21256	정태학	✓	✓	✓	✓	✓	✓
15	3093 조영문	갈루아 필 심에서 정의된 다항 곡선 석사	석사	컴퓨터공학부	2001-21574	노종선	✓	✓	✓	✓	✓	✓
16	2659 이정우	황산이 함유된 지르코니아에서 1-부 석사	석사	응용화학부	2001-21322	이현규	✓	✓	✓	✓	✓	✓
17	1282 김기건	저금양분함유강의 품질개선을 위한 석사	석사	응용화학부	2001-21690	이무하	✓	✓	✓	✓	✓	✓
18	2155 최호진	개에서 Prionin Isomericity Surfs 석사	석사	수의학과	2000-30966	윤영희	✓	✓	✓	✓	✓	✓
19	2941 허도훈	기하와 삼각의 논리학적 연구 - 한국 석사	석사	경영학과	2001-20914	조태호	✓	✓	✓	✓	✓	✓
20	2036 김동식	디이산 구의 배선 구조에서 논 석사	석사	응용화학부	2001-21278	김병기	✓	✓	✓	✓	✓	✓
21	1220 반문우	생인 기준지형식 개발의 위험인자 석사	석사	의학과	2001-22410	조상현	✓	✓	✓	✓	✓	✓
22	1205 장수미	새로 진단된 다발성 골수종에서 고 석사	석사	의학과	2001-22411	박선영	✓	✓	✓	✓	✓	✓
23	1211 장성안	한국의인 우물장애 진단법주의 특 석사	석사	의학과	2001-22405	조영훈	✓	✓	✓	✓	✓	✓
24	1142 임우식	소아에서 기내류의 유변학적 특 석사	석사	의학과	97801-563	김용삼	✓	✓	✓	✓	✓	✓

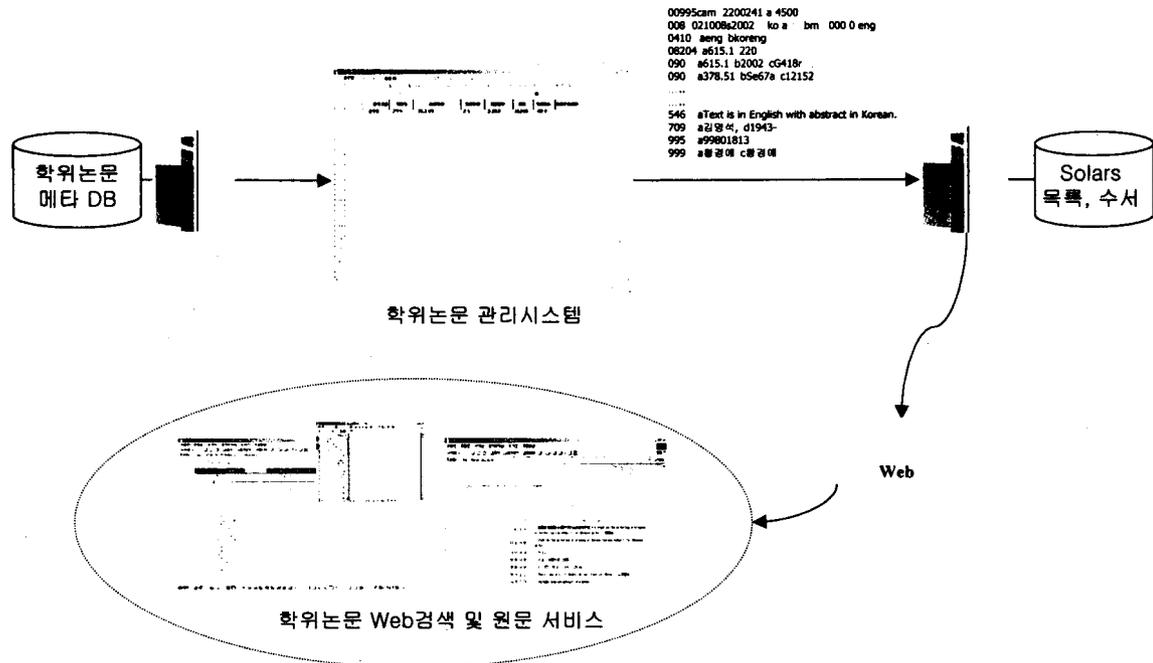
2.2.8. 학위논문 제출 확인서 출력

- 제출자는 논문책자를 제출한 후 학위논문 온라인 제출시스템에서 개인별로 확인서를 출력하여 각 대학 교무과에 제출하여야 하며, 대출중인 도서가 있거나 연체료 미납자는 확인서 출력이 불가능



2.2.9. 도서정보시스템(SOLARS)으로 이관

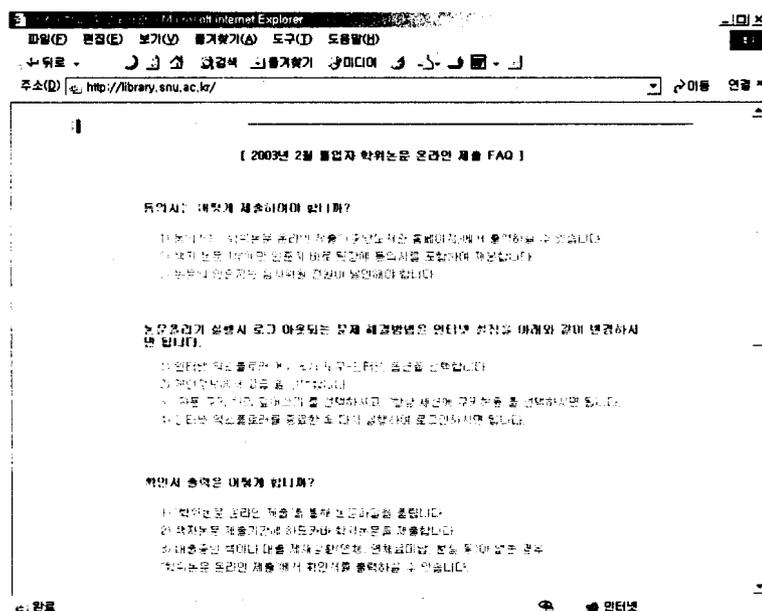
- 관리시스템에서 학위논문 Meta정보를 수서 및 목록테이블로 이관하여 웹에서 학위논문 PDF서비스가 가능함



2.3. 학위논문 온라인 제출시스템 기타 기능

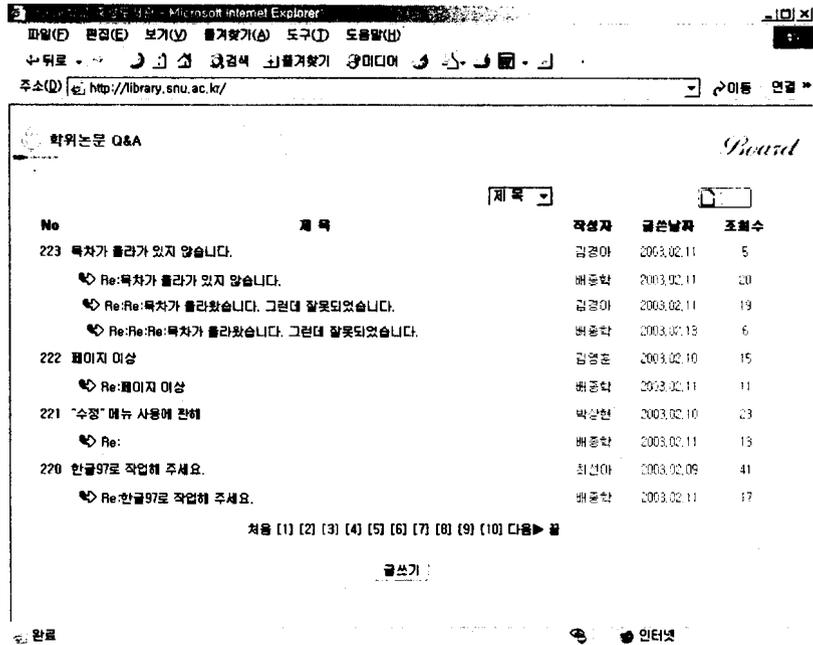
2.3.1. 학위논문 온라인 제출 FAQ

- 온라인 제출시스템 사용시 많이 문의되는 사항을 정리하여 제출자에게 안내



2.3.2. 학위논문 Q&A

- 제출자가 온라인 제출시스템 이용시 문의사항이 있을 경우 관리자가 답변하도록 하였음



3. 학위논문 온라인 제출시스템의 문제점

3.1. 학위논문 제출자의 협조가 필요

- 기존에는 석박사 논문 제출자가 디스켓 제출만 하면 되었지만, 지금은 웹을 통해 논문정보(초록, 주요어, 목차 등)를 입력하고, 논문파일을 올려야 하며, 자신의 논문이 PDF 변환이 잘 되었는지를 확인하여야 하기 때문에 학위논문 제출자의 협조가 절대적으로 필요함

3.2. 온라인제출시스템 속도 문제 발생

- 1) 네트워크 장애시 또는 논문파일 업로드 기한 마지막 시점에 제출자가 동시에 올릴 경우 업로드 시간이 많이 소요되는 문제가 발생하였음
- 2) 업로드 속도 개선을 위해 파일 용량이 큰 경우(20MB으로 제한)는 제출자가 직접 CD-ROM을 제출해야 하는 불편이 있음

3.3. PDF변환시스템 문제 발생

- 1) PDF변환 이후 제출자가 직접 PDF 확인하고, 논문 책자 제출 이후 확인서 출력

이 가능하기 때문에 PDF변환시스템의 오류로 인해 PDF변환에 어려가 생길 경우 제출자가 확인서 출력하는데 시간이 많이 걸릴 수 있음

- 2) 한글97, 2002인 경우 PDF 변환시간이 많이 걸리고, 한글2002가 한글97을 지원하지 못하는 부분이 있어 97과 2002를 분리하여 PDF를 변환해야 함

3.4. PDF변환 중 오류가 있는 논문 발생

- 1) PDF변환이 안되는 논문파일이나 부분적 오류(그림, 표, 폰트 누락)가 있는 일부 논문은 외부 구축업체를 통해 PDF파일을 구축하여야 하기 때문에 그에 따른 구축 비용 발생
- 2) PDF변환에 오류가 있는 논문들은 수정작업을 거쳐 PDF서비스가 이루어지기 때문에 일부 논문의 PDF서비스 시작이 늦어지는 문제가 있음

4. 학위논문 온라인 제출시스템 보완점

- 4.1. 논문 제출자에게 논문표준형식(파일 순서 및 목차형식)을 제공하여 PDF변환시 오류가 최소화 되도록 해야 함
- 4.2. 제출자가 웹에서 논문파일을 올릴 때 그림 누락, 폰트 누락을 점검할 수 있도록 하여 웹에서 제출자에게 바로 통보할 수 있도록 하여 논문파일을 수정하여 올리는 시간을 단축할 있는 기능이 필요함
- 4.3. PDF 변환 대기시간을 안내하여 제출자가 PDF 변환완료 때까지 기다려야 하는 불편을 조금이나마 감소할 수 있도록 해야 함
- 4.4. HWP파일은 PDF 변환작업이 오래 걸리기 때문에 PDF 변환서버를 여러 대 설치하여 다른 논문파일(MS-word 등)과 구분하여 변환하게 하여 PDF 변환 대기시간을 단축해야 함

5. 학위논문 원문정보 공동 이용체제 현황

5.1. KERIS 종합목록 이용 원문 링크정보 (원문 10만건)

종합목록에서 각 대학에서 제공한 원문 링크 정보에 의하여 원문 이용

5.2. 국회도서관: 인문·사회분야 (원문18만건)

상호협력에 의하여 원문이 구축된 인문사회 분야 학위논문을 도서관에 지정된 PC에서 이용

5.3. 현행 학위논문 공동이용 체제의 문제점

- 국회도서관 학위논문 원문 정보

- 인문사회 분야에 한정되며 도서관의 지정된 PC에서만 이용 가능

- KERIS

- KERIS에 접속하여 학위논문 원문을 이용하며 10만건에 한함

5.4. 학위논문 원문 정보 공동이용체제의 개선방안

- 학위논문의 원문링크 서지데이터를 상호교환하여 자관의 도서관 시스템에 올려놓고 이용함으로써

- 써 원문 이용 대상을 확대하고 이용자의 편의를 제공함

- 학위논문의 생산자인 대학이 원문 이용체제의 주체가 됨

6. 서울대학교 학위논문 원문링크 서지데이터 제공

6.1. 목적

열린 대학과 대학생산 학술자료의 공유를 위하여 서울대학교 원문 링크 서지데이터를 제공함

6.2. 서울대학교 학위논문 구축 현황

- 원문 구축 학위논문: 총 71,832건

- 개교 이래 서울대 수여 학위논문 원문 구축 완료

- 신규학위논문 : 연 3,300여건

- 매년 2월 및 8월에 수여 되는 학위논문 3,300여건의 원문을 도서관 제출 후 1개월 이내 원문 구축

6.3. 기본방침

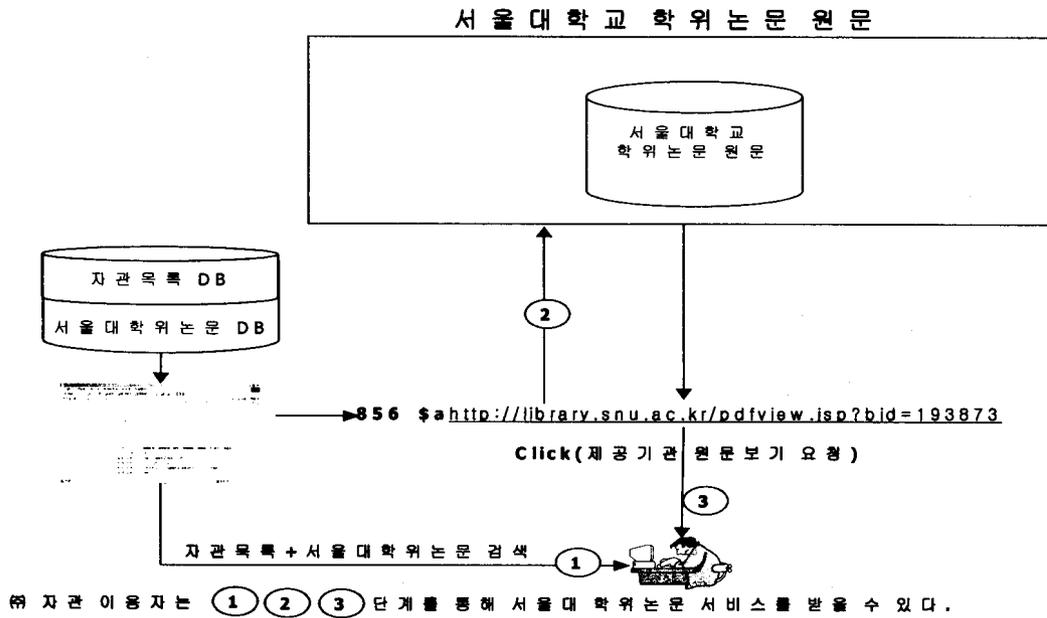
- 자관 시스템에서 원문 이용
- 서울대학교에서 제공하는 학위논문 원문링크 서지데이터를 도서관 시스템에 통합 업로드 하고 자료 검색 시 원문 보기 아이콘을 클릭하면 서울대 학위논문 원문서버로 접속하여 이용
- 논문전체 또는 특정주제분야(해양학, 교육학)만 교환가능
- 상호협력 체결에 의하여 제공
상호협력 체결 후 서울대학교 학위논문 원문 링크 서지를 먼저 제공하며 상호 교환
- 캠퍼스 내에서 이용
캠퍼스 내에서 자유롭게 이용할 수 있으며 상업적 이용은 배제함
- 학위논문 원문 미 구축 대학
학위논문 원문이 구축되지 않은 대학도 상호협력 체결만으로 원문 링크 서지 데이터를 제공하며 차후 원문이 구축될 시 서울대에 제공 함
- 제공 대상 기관
 - 대학: 국·공·사립대학
 - 비상업적 연구 및 전문 도서관
 - 공공기관

6.4. 추진일정

- 2002. 11 사업계획 수립 및 착수
- 2002. 11. 대학본부 처실장 회의에서 심의
- 2002. 12 시범운영을 위한 관련 대학 협의
- 2003. 1. 시범운영대학 : 안동대학교
 - 안동대학교 도서관 시스템에 서울대 학위논문 원문 이용 방법 및 속도 등 테스트가 가능
 - 안동대학교 주소: <http://lib.andong.ac.kr/solarsweb4/>
- 2003. 3 전국대학 참여 신청 여부 조사

- 2003. 4 전국대학 참여 신청 결과 보고 회신
- 2003. 4 공동협의체 구성 관련 협의회 개최 예정 등

6.5. 서비스 구성도



1.3. 원문 링크 서지데이터의 교환 방법

- 서울대 학위논문 원문 링크 서지데이터 용량 : 47MB
- CD-ROM 으로 제공
- FTP로 다운로드
 - 상호협력 체결 후 FTP 서버의 ID, password 부여

1.4. 원문링크 서지데이터 자관 다운로드 및 서울대 업로드 방법

- 각 도서관에서 사용하는 도서관 시스템 패키지 별로 다운로드 업로드 소프트웨어 개발 지원함

7. 학위논문 원문정보 링크 서지데이터 상호이용 공동협력

7.1. 목적

학위논문의 생산자인 대학이 원문 정보 상호이용 체제의 주체가 됨

7.2. 공동 협약 체결

○ 협약방법

대학간 협약 체결의 불편을 해소하기 위하여 공동 협약 체결을 원칙으로 하여 필요에 따라 대학간 개별 협약도 가능

○ 원문 미 구축 대학

학위논문의 원문이 구축되지 않은 대학은 공동 협약에 의하여 제공 받으며 원문 구축 후 원문 링크 서지 데이터를 제공

○ 공동 협약서 【별첨1】

7.3. 공동협약 협의체 구성

학위논문 원문정보 링크서지 상호이용 공동협약 기관간의 원활한 사업 추진과 표준화 등을 위하여 협의체를 구성 함.

8. 학위논문 원문 링크 서지데이터 교환 방법

8.1. CD-ROM으로 교환

- 원문이 링크된 서지데이터를 CD-ROM으로 받아 공동 협약 기관에 배포

8.2. FTP 서버 이용 (서울대에서 FTP 서버 제공)

- FTP 서버에 올려놓고 공동 협약 기관이 직접 다운로드

8.3. 온라인으로 자동 업데이트

- 추가 수정 및 변경 서지를 온라인으로 자동 갱신하는 방법은 공동 협약 협의체 회의에서 추진함

8.4. 공동 협약 관련 게시판 운영

- 공동협약에 대한 문의 및 의견 게시판 운영
- URL: <http://libraryboard.snu.ac.kr/cgi-bin/CrazyWWWBoard.cgi?db=solars2>

9. 원문 링크 서지 교환 포맷 및 상세 안내

9.1. MARC 서지자료 교환 포맷

- 로컬 정보 삭제 (등록번호 등)
- 청구기호 (090 tag)은 \$a에 대학명 표기
ex) 서울대 원문
- 자관 제어번호는 035tag로 대학부호를 포함해서 이동
ex) 035 \$a(211032)000000193873
- 원문정보 주소 856(URL) 필수
- KORMARC, MARC21 자료 구분

9.2. MARC 제공 예시 DATA

```
005199990306195923▲
008 960903s1996 ulka AK 000a kor ▲
035 ▼a(211032)0000193873▲
0410 ▼akor▼beng▲
090 ▼a서울대 원문
```

- 1001 ▼a정평기,▼d1969-▲
- 24510 ▼a韓國 經濟發展에서 技術集約的 産業의 役割 /▼d鄭平基.▲
- 260 ▼a서울 :▼b서울大學校 大學院,▼c1996.▲
- 300 ▼aiii, 40장 :▼b삽도 ;▼c26 cm.▲
- 5020 ▼a학위논문(석사)--b서울大學校大學院:▼c電氣工學科,▼c1996.▼e오영석▲
- 504 ▼a참고문헌이 포함되어 있음▲
- 856 ▼uhttp://library.snu.ac.kr/pdfview.jsp?bid=193873 \$y이 자료의 원문과 서지 정보는 서울대학교에서 제공합니다.

9.3. 검색 안내 및 홈페이지 서비스 안내

안동대학교 시연 url: <http://lib.andong.ac.kr/solarsweb4/>

기관

검색항목:

제한항목: 연어 검색수

자료유형 페이지당 건수

AND 기사포함

검색

- 자관 목록 종합검색에서 저자 “김진곤” 입력한다

순번	제목	저자	발행년도	청구기호	자료유형	원문
1	한국의 경제발전에서 기술집약적 산업의 역할 / 김진곤	김진곤	1996	서울대 원문	<	
2	국립중앙도서관	김진곤	1996	서울대 원문	<	
3	국립중앙도서관	김진곤	1993	서울대 원문	<	
4	한국의 경제발전에서 기술집약적 산업의 역할 / 김진곤	김진곤	1969	서울대 원문	<	
5	美山谷詩論 研究 / 최일영	최일영	19	620.4062 종63 C-18	<	

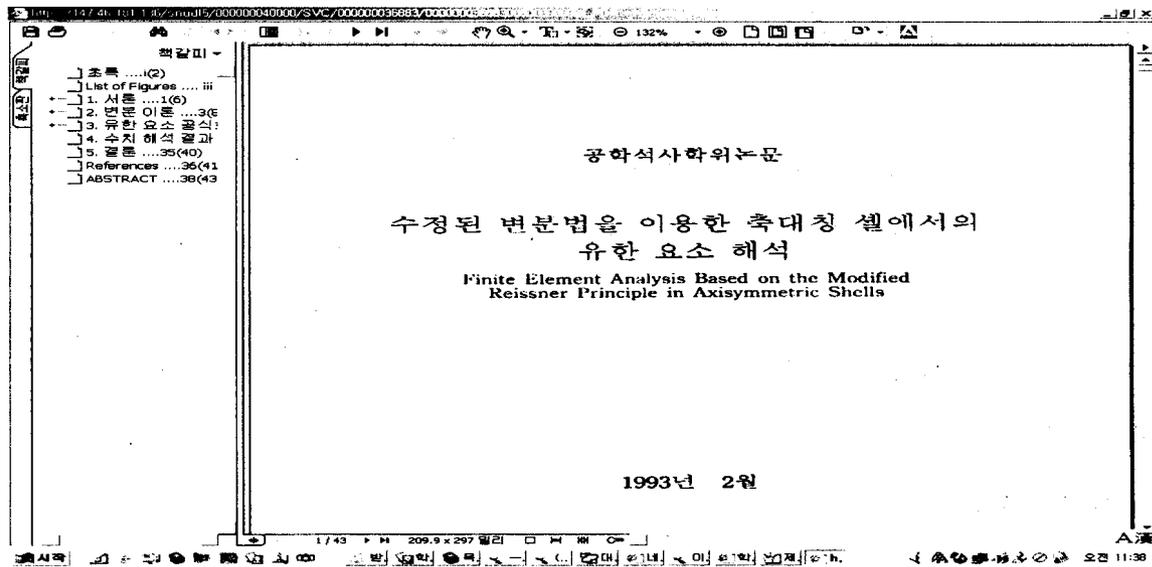
- 검색 결과 5건 중 서울대학교 원문 4건, 자관 소장자료 1건 검색 됨

서지번호 : 55719

MARC

항 목	내 용
서 명 / 저 자	수정된 변분법을 이용한 축대칭 셀에서의 유한 요소 해석 / 김진곤.
원 저 자 명	김진곤.
발 행 사 항	서울 : 서울대학교 대학원, 1993.
형 태 사 항	35 장 : 삽도 : 26 cm.
학 위 논 문	학위논문(석사)-- 서울대학교 대학원 : 기계설계학과, 1993.
서 지 주 기	참고문헌이 포함되어 있음.
웹 사 이 트	http://library.snu.ac.kr/pdfview.jsp?bid=172431 : ★ 자료의 원문(PDF)과 서지정보는 서울대학교에서 제공합니다.

- 수정된 변분법을 이용한.. 간략서지에서 더블클릭 상세화면
- 웹사이트 url 클릭으로 아래와 같이 원문 보기



10. 추진배경 및 목적

- 대학마다 매년 발행되는 학위논문을 한 대학에서 모두 소장하는 것은 어려운 일이며 중복사업으로 인한 예산 낭비
- 자료 검색시 각 대학의 홈페이지를 따로 드나들어야 하는 불편을 해소하고 한번의 접속으로 모든 대학의 소장 정보를 검색 가능
- 검색한 자료를 바로 원문 복사 신청 가능하므로 빠른 정보 입수가 가능

11. 추진방향

- 대학에서 발행된 학위논문을 PDF로 구축하여 원문 정보의 링크 데이터를 타 대학에 제공함
- 대학간 구축된 학위논문의 원문 정보 공유 및 추후 자료유형 확장
- 공동협약 구성등의 방법으로 학술 자원의 공동 활용

서울대학교 학위논문 69,650건

○ 석사 51,075건

○ 박사 13,308건

○ 구축중 5,267건

※ 서울대학교 논문은 매년 3,300건 이상 신규로 증가 됨

12. 세부 추진계획

12.1. 기본사항

- 학술교류 상호협정을 체결하여 제공
- 상업적으로 사용할 수 없으며 다른 기관에 양도 금지
- 관리기능은(저장/출력) 해당대학에 기준을 우선

- 지속적인 관리가 될 수 있도록 신규 자료를 주기적으로 교환
- 교환 정보가 수정될 경우 최신성 유지를 위한 상호협력
- 서버 변경, OpenURL변경, 서지사항 변경 등

12.2. MARC 자료 제공

- 학위논문 서지데이터는 상호 호환성이 있는 MARC 포맷이고 원문을 볼 수 있는 주소가 856 Tag(URL)에 포함
- 로컬정보(090,093,9XX TAG)는 삭제
- 제어번호(001tag)은 자관에서 수정 및 삭제된 정보 연동을 위해 035tag로 대학부호를 포함해서 이동 후 입력
- KORMARC 과 MARC 21(USMARC) 자료 구분

[제공 예시자료]

- 005 199990306195923▲
- 008 960903s1996 ulka AK 000a kor ▲
- 035 ▼a(211032)0000193873▲
- 041 0 ▼akor▼beng▲
- 100 1 ▼a정평기,▼d1969-▲
- 245 10 ▼a韓國經濟發展에서 技術 集約的 産業의 役割 /▼d鄭平基.▲
- 260 ▼a서울 :▼b서울大學校 大學院,▼c1996.▲
- 300 ▼aiii, 40장 :▼b삽도 ;▼c26 cm.▲
- 502 0 ▼a학위논문 (석사)--▼b서울大學校 大學院:▼c電氣工學科,▼c1996.▼e오영석▲
- 504 ▼a참고문헌이 포함되어 있음.▲
- 856 ▼uhttp://library.snu.ac.kr/pdfview.jsp?bid=193873\$y이 자료의 원문과 서지 정보는 서울대학교에서 제공합니다

- DB 전송방법은 CD-ROM, E-mail, FTP 등을 활용
- DB 운용방안은 자관 DB와 통합하거나 분리하여 별도 DB로 운용

12.3. 저작권 분쟁 문제

- 서울대학교 경우 제출자의 동의를 받지 않은 논문(1999년 이전 논문)인 경우 문제점이 발생 될 수 있으나, 원문 link 정보만을 제공하는 것으로 문제점 해소

라. 이용자 증가로 인한 네트워크 부하

- 서울대학교 경우 동시 사용자 수 제한으로 조정

12.4. 통계기능 강화

13. 대상기관 및 추진일정

13.1. 대상기관

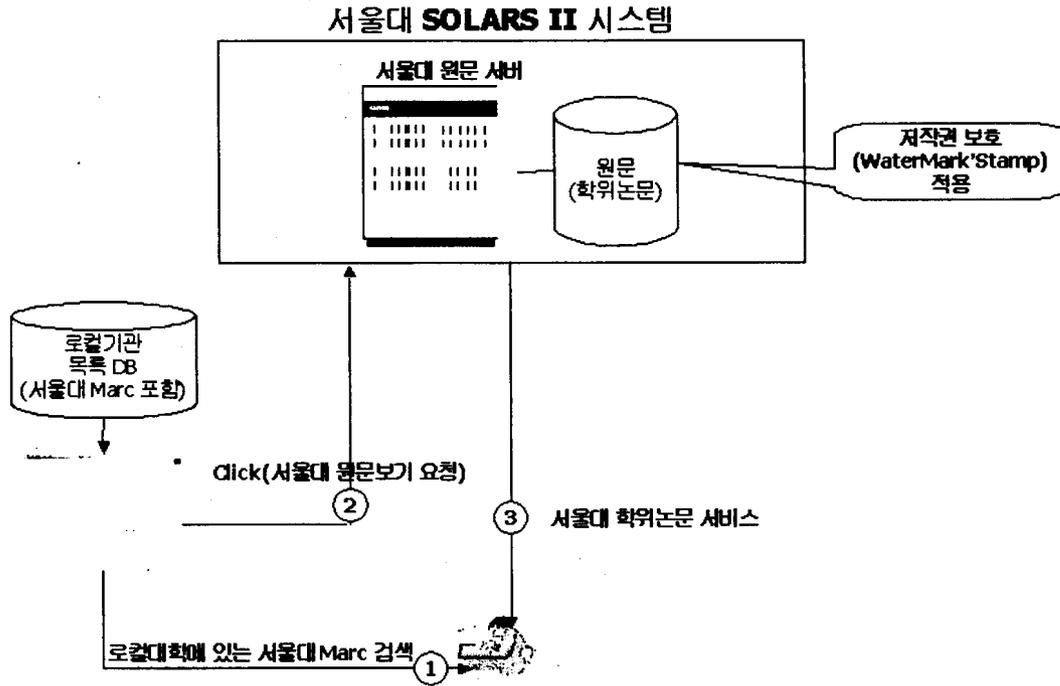
- SOLARS 사용 대학, 국·공·사립대학
- 국립중앙도서관, 국회도서관
- 한국교육학술정보원(KERIS)
- 공공기관 및 기타

13.2. 추진일정

- 2002. 11. 11. 서울대학교 확대간부회의 심의 제출 및 통과
- 2002. 11. 11. SOLARS 사용대학 시범운영 회의 개최
- 2002. 11. 20. 경북대 시범 가동을 위한 서버 프로그램 개발
- 2002. 11. 22. 경북대학교 시범 가동
- 2002. 11. 23- 프로그램 개발
- 2002. 12. 4 1차 SOLARS 사용 대학 관장 협의회 개최
- 2003. 4. 24 2차 국. 공립 대학 전체 회의

- 2003. 5. 3차 사립대학 전체 회의
- 2003. 12. 가동

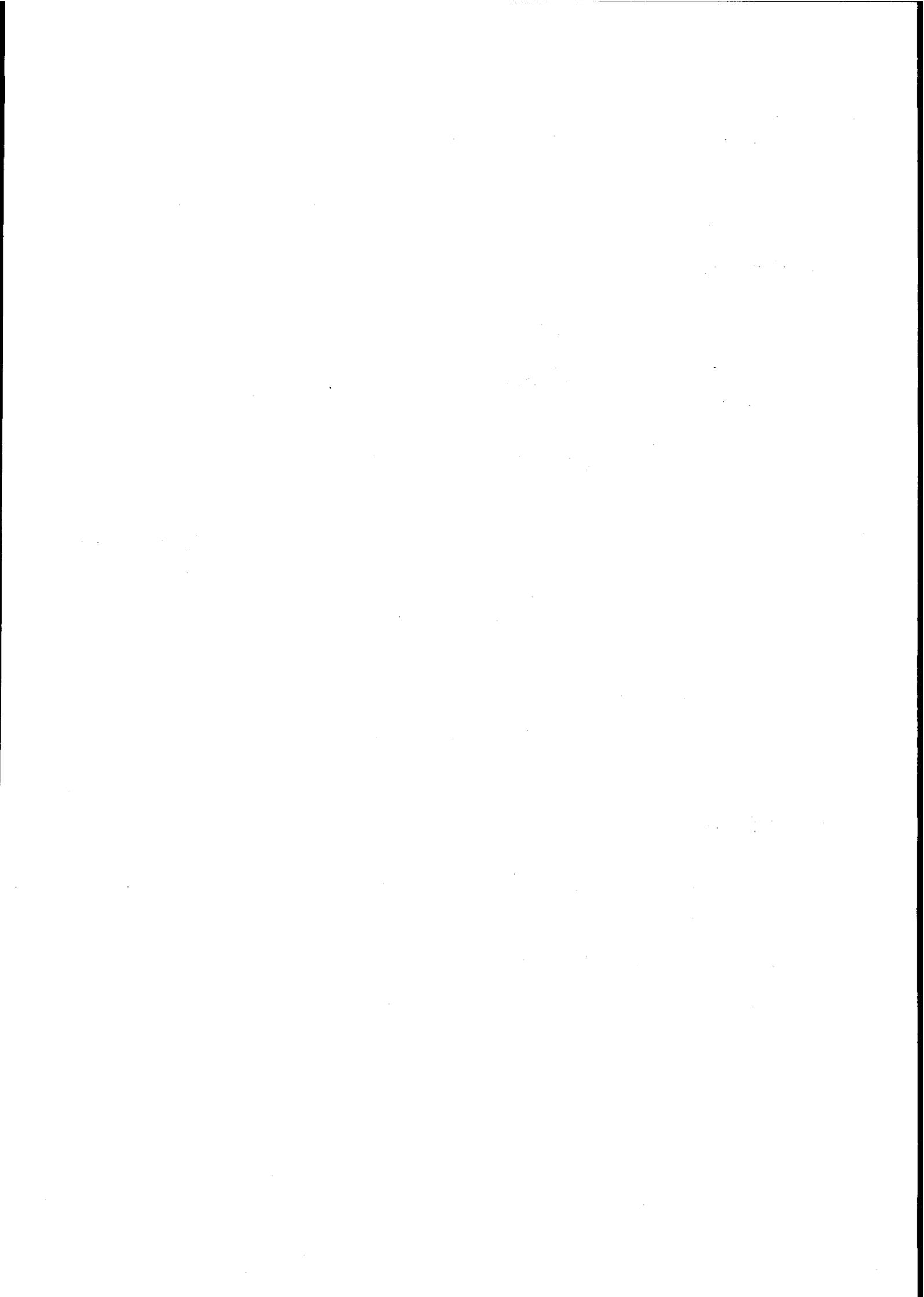
14. 서비스 구성도



로컬 이용자는 ①②③ 단계를 통해 서울대 학위논문 서비스를 받을 수 있다.

15. 기대효과

- 타 대학 학위논문을 자관에서 구축한 DB와 동일한 형태로 서비스 받을 수 있음
- 기존의 원문 복사 신청 방법에 의해 걸리던 자료 입수 시간등이 대폭 단축되고 자관이 소장하고 있지 않은 문헌도 온라인 제공 가능
- 자료를 검색하기 위하여 여러 사이트를 돌아다니는 불편을 해소하고 원클릭으로 모든 자료를 검색 가능하게 됨으로 정보를 찾는데 걸리는 시간 단축

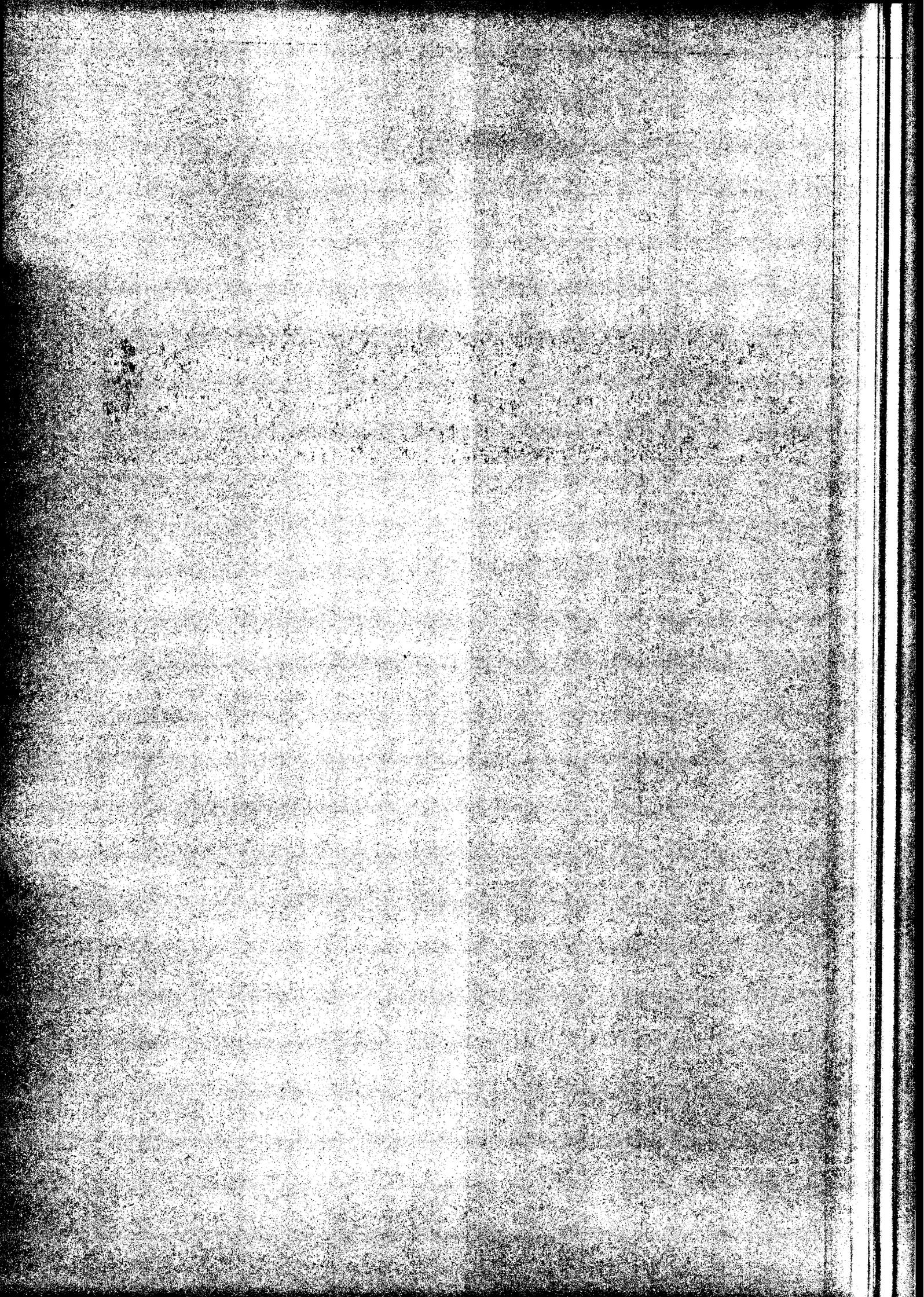


【서명표 10】

2003. 5

이태순

taesoon@snu.ac.kr



< 사례발표 10 >

서울대학교전자도서관의
학내학술 정보 통합사례

2003. 5.

이 태 순
(taesoon@snu.ac.kr)

차 례

1. 학내학술정보 통합 추진배경 및 방향.....	1
2. 학술정보시스템의 개발 및 캠퍼스 분담목록 추진과정.....	1
3. SOLARS 시스템 개발에 따른 통합작업.....	2
4. 서울대학교 학과 및 연구소 서지DB통합	6
5. 서지통합식 운영에 따른 문제 및 해결점.....	5
6. 대학도서관 정보화 사업 소개	6

1. 학내학술정보 통합 추진배경 및 방향

1.1. 추진배경

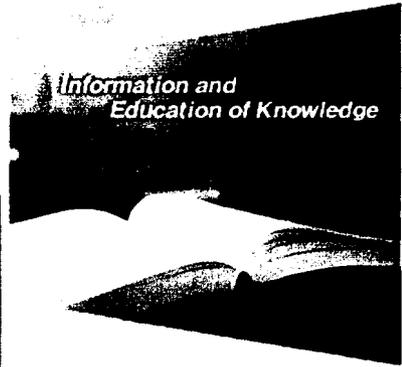
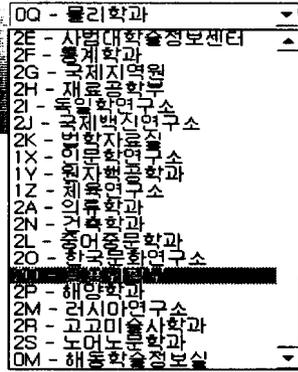
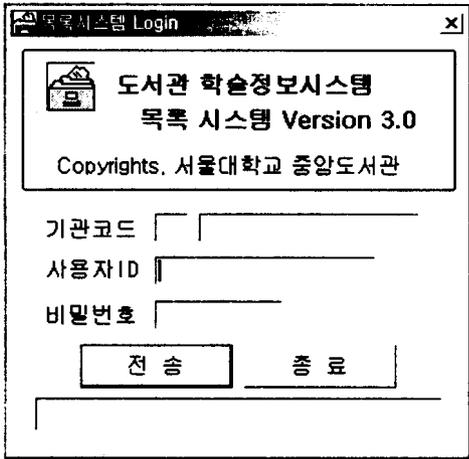
1992년부터 시작한 중앙도서관의 학술정보시스템(SOLARS)의 개발은 각 학과 및 연구소 자료실에 산재되어 있는 대학 학문의 발달에 가치가 있는 중요한 자료의 공동 활용 및 공유가 가능하게 되었다.

1.2. 추진방향

- 중앙도서관 학술정보시스템(SOLARS)의 목록시스템을 통한 공동서지데이터를 구축 자료정리에 신속함과 편리함을 제공
- 각 학과 및 연구소 자체의 특성에 맞는 대출단위시스템을 제공 원활한 학내간 상호 대차의 기반을 조성
- 각 학과 및 연구소의 홈페이지에 검색시스템을 통한 DB를 연결 학내간의 신속한 학술정보자료를 제공 학문의 연구·발달에 기여

2. 학술정보시스템의 개발 및 캠퍼스 분담목록 추진과정

- 1994년 3월 중앙도서관 학술정보시스템(SOLARS) 1.0 version (Tandom) 개발
- 1997년 3월 중앙도서관 학술정보시스템(SOLARS) 2.0 version (UNIX-WINDOWS version) 개발
- 1997년 6월 중앙도서관 학술정보시스템(SOLARS) 웹버전 개발
- 1998년 3월 캠퍼스 분담목록 52개 학과 및 연구소 오픈
- 2001년 8월 중앙도서관 학술정보시스템(SOLARS II) 개발
- 2002년 12월 SOLARS II 시스템에 의한 60여개 학과 및 연구소 소장자료의 학내 학술정보통합작업완료
- 2003년 현재 미구축 학과 및 연구소 소장자료의 통합작업 진행중



[그림] SOLARS 목록시스템 로그인 화면 변화

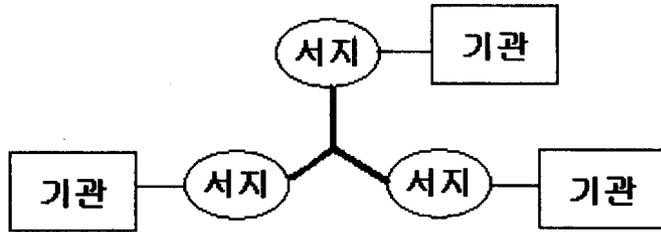
3. SOLARS 시스템 개발에 따른 통합작업

- SOLARS시스템개발의 환경변화는 학내 통합자료를 구축하는데 많은 영향을 끼쳤고 그 결과는 이용자들의 서비스에 많은 변화를 주었다.

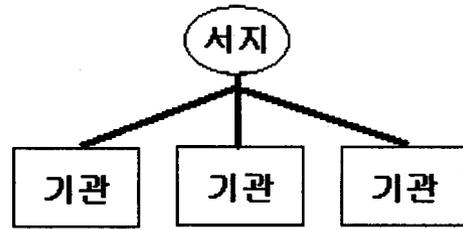
구 분	목록 시스템	대출 시스템	검색 시스템
SOLARS	서지병렬형 DB구축	수작업 및 개별 대출시스템	기관별 개별 검색시스템
SOLARS II	서지통합형 DB구축	통합시스템이용 및 개별 대출시스템	기관별 통합 및 개별 검색시스템

- 1998년 3월에 오픈한 학과 및 연구소의 서지DB구축형식은 소장자료에 따라서 각 기관마다 같은 서지라도 따로이 서지DB 구축하고 등록자료를 기입하는 서지병렬형 형식으로 데이터의 검색도 중앙도서관과 학과 및 연구소 자료를 따로이 검색해야 하는 불편함이 있었으나 2002년 8월에 오픈한 SOLARS II는 이러한 불편함을 제거 함으로써 하나의 서지를 기본으로 각 학과 및 연구소의 소장자료의 등록번호만을 기입하는 서지통합형으로 DB를 구축 업무시간의 효율적인 활용이 가능하게 되었으며 이러한 서지DB의 통합구축은 곧 이용자의 검색서비스에 큰 영향을 끼쳤다.

[표1] 서지병렬식 DB구축



[표2] 서지통합식 DB구축



3.1. 서지DB통합 구축에 따른 목록시스템의 변화

- 서지통합 DB를 구축함으로써 기존 중앙도서관에서 사용하던 목록의 참고자료 (OCLC, KERIS)의 공유가 가능, 각 학과 및 연구소에서는 신규서지작성에 큰 도움을 받을 수 있게 되었으며 통일적인 규칙을 제공 받으므로써 효율적인 업무관리를 할 수 있게 되었다

목록시스템 Ver 0.9.43
 시스템 [C] 출력업무 [B] 인수인계 [I] 서지데이터 [R] 전거데이터 [A] 편집 [E] 마크처리 [P] 도구 [T] 윈도우 [W]
 [서울대학교] [서명 + 저자] [화학공학] [서명 저자] 검색 [X]

서지 간략 화면

상세보기 | 파일저장 | 프린트 | 가리기 | 닫기 검색건수 : 52 건

순번	선택	서지번호	서명	저자	청구기호	년도	유형
41		124518	化學工學 術語集	한국화학;		1987	BK
42		124047	化學工學實驗法 / 2 : 工業化學	한국화학;		1975	BK
43		120440	化學工學 / 上, 下卷	Badger, W		1961	BK
44		113551	化學工學	김주봉		1976	BK
45		113549	化學工學概論 : 量論篇	이희철		1971	BK
46		113547	化學工學 熱力學	Van Ness,		1977	BK
47		113545	화학공학 열역학	Smith, J.		1988	BK
48		113543	化學工學特論 1	한국화학;		1987	BK
49		90866	화학 공학 통론 : 공업계 고등	한국		1973	BK
50		11708	化學工學	Kagaku Ke			Cont SE
51		10940	化學工學論 1	Yaguchi K			Cont SE
52		4129	화학 공학	한국화학;			Cont SE

소장기관

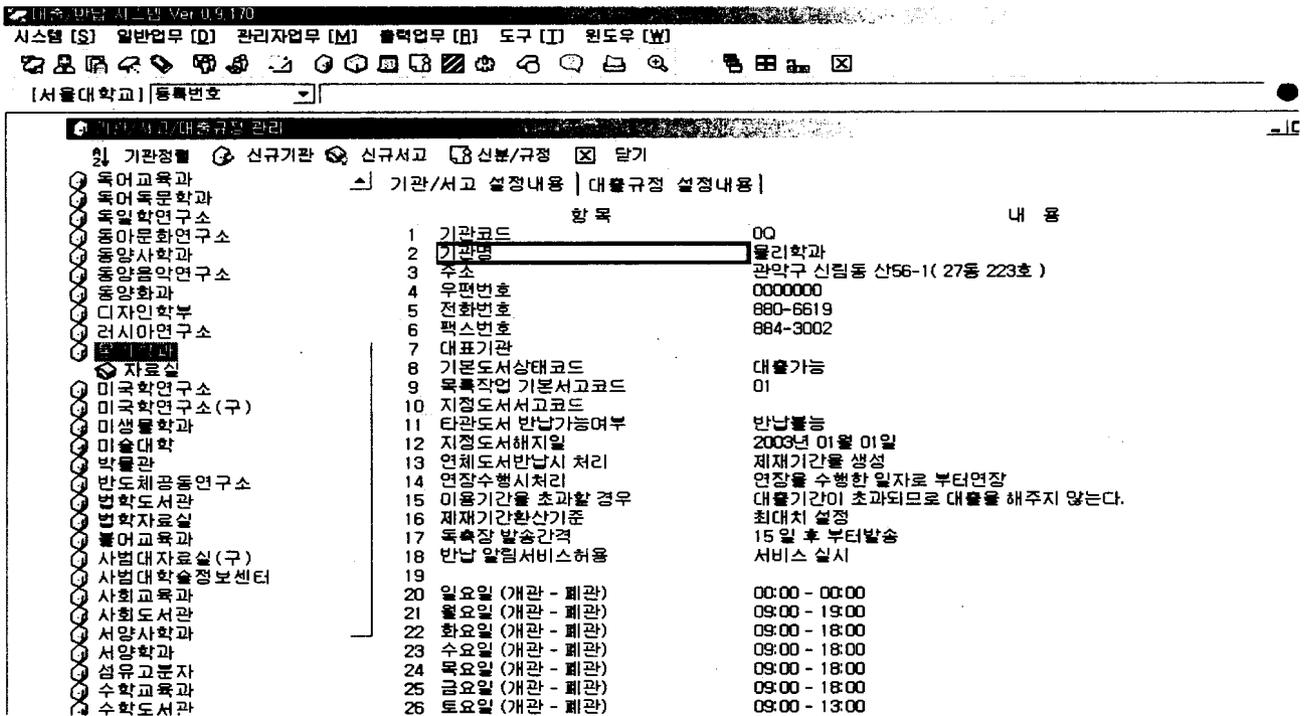
순번	소장처	청구기호
1	중앙도서관 - 01	P 660 K119r
2	중앙도서관 - 01	P 660 K119r

서명 + 저자 = '화학공학 +'

3.2. 서지DB통합 구축에 따른 대출시스템의 변화

- 목록시스템의 서지통합DB구축과 더불어 대출시스템의 이용자파일등의 공유는 개별 시스템 및 수작업으로 이루어지던 학과 및 연구소의 대출정책에 큰 변화를 가져 왔다 단순한 대출과 반납의 기능에서 좀더 다양한 통계의 작성이 가능하게 되므로 각 학과 및 연구소 자료실의 수서 및 정책결정에 영향을 줄 것으로 예상 된다

- 대출시스템 이용자파일의 공유는 이용자들에게 학과 및 연구소 자료의 대출시 신분의 검증은 따로이 받아야 하는 번거러움을 줄여주었으며 대출자료에 대한 반납 메일등 통합서비스를 받을 수 있게 되었다



3.3. 서지DB통합 구축에 따른 검색시스템의 변화

- 서지자료의 통합 검색은 필요한 자료의 소장유무에 대한 이용자들의 만족도에 큰 영향을 끼치게 될 것으로 학과 및 연구소의 자료는 중앙도서관의 자료보다 부분적으로는 좀더 세밀하게 주제에 접근한 자료가 많아 학문적연구를 위한 자료에 대한 검색 효과가 높다고 할 수 있으며 이는 도서관정보화사업의 기대효과인 “학술 연구 정보의 신속한 제공으로 연구자들의 연구경쟁력 강화” 결과로 나타나게 될 것이다.

3.3.1. 통합검색의 효과

- * 이용자가 원하는 자료 : v. 11, no. 1 의 논문내용
- * 검색서명 : 화학공학논문집
- * 소장기관 : 중앙도서관, 응용화학부
- * 소장자료 : 중앙도서관 --> v.17 no. 3-부터 소장
< 응용화학부 -- > v.11 no. 1 소장 >

서울대학교 중앙도서관
SEOUL NATIONAL UNIVERSITY LIBRARY

자유검색 전자도서관 도서관 서비스 My Library 이용안내 도서관 안내 도서관 통계 게시판

Home Sitemap E English S 서울대학교 i 정보광장

통합검색

전체 유속출판 키워드 완전일치

서명 [검색] [결과내림]

사양유형 : 전체 단행본 연속간행물 기사색인 학위논문 지도서 원문서

전체 단행본 연속간행물 기사색인 학위논문 지도서 원문서

검색어 : > YBTI=국문학 이야기/A/ > YBTI=국문학/A/ > YBTI=행정학/A/ > YBTI=건축/A/ > YBTI=건축/A/ > YBTI=화학공학논문집/A/

1980-1999 2000-2009 2010-2019 2020-2029 2030-2039 2040-2049 2050-2059 2060-2069 2070-2079 2080-2089 2090-2099

서명 저자 출판사 발행년도 소장위치 자료유형

화학공학논문집
Kagaku Kagakukai (화학공학부)
출판사 : 중앙도서관 [P 660 K119r] 응용화학부 [P 660 K119r]

초 검색수 [11] GO

인쇄/인출

번호	표제	출판사명	서지번호
1	화학공학논문집	화학공학부	10000010040
2	화학공학논문집	화학공학부	
3	화학공학논문집	화학공학부	
4	화학공학논문집	화학공학부	
5	화학공학논문집	화학공학부	
6	화학공학논문집	화학공학부	
7	화학공학논문집	화학공학부	
8	화학공학논문집	화학공학부	
9	화학공학논문집	화학공학부	

도서상태 최신입수정보 : [전체] 중앙도서관

도서상태

번호	청구기호	소장위치	도서상태	등록번호
1	10000010040	응용화학부 지도서	대출가능	10000010040
2	10000010040	응용화학부 지도서	대출가능	10000010040
3	10000010040	응용화학부 지도서	대출가능	10000010040
4	10000010040	응용화학부 지도서	대출가능	10000010040
5	10000010040	응용화학부 지도서	대출가능	10000010040
6	10000010040	응용화학부 지도서	대출가능	10000010040
7	10000010040	응용화학부 지도서	대출가능	10000010040
8	10000010040	응용화학부 지도서	대출가능	10000010040
9	10000010040	응용화학부 지도서	대출가능	10000010040

4. 서울대학교 학과 및 연구소 서지DB통합

4.1. 서울대학교의 학과 및 연구소의 서지DB구축을 크게 SOLARS 오픈시점인 1998년과 2002년 통합서지에 대한 비교

4.1.1. 연도별 참여기관 및 구축자료

년 도	참 여 기 관	구 축 자 료
1998년	41개 기관	182,671책
2002년	60개 기관	345,807책

* 참여기관과 구축자료의 수가 많이 증감 하였으며 2003년에는 보다 더 많은 기관이 참여 할 것으로 예상된다.

4.1.2. 연도별 자료실 담당자

년 도	사 서	조 교	일 반	계
1998	7	25	9	41
2002	11	40	9	60

* 참여기관의 증가에 비해 각 학과 및 연구소의 자료실에 대한 관리는 효율적으로 이루어지지 못하고 있음을 알 수 있는 통계로 이는 통합에 따른 여러 가지 해결해야 할 문제를 제시하고 있다.

4.2. 학과 및 연구소 자료실의 장서현황

장서수 (책)	3000-5000	5000-10,000	10,000-15,000	15,000-20,000	20,000-
기관수	34	15	5	3	3
평균장서(책)	1,800	6,600	12,900	17,000	25,000

* 2003년 4월 현재 자료실 장서현황으로 학과 및 연구소 자료실의 운용정책 결정에 기초자료를 제시 하고 있다.

* 이 표는 서지DB구축자료를 기본으로 작성하였으며 연간물자료의 경우 미제본도서는 서지DB구축 대상에서 제외 되어 있다.

5. 서지통합식 운영에 따른 문제 및 해결점

5.1. 문제점

- 각 학과 및 연구소기관의 서지 구축 담당자에 대한 능력의 차이
 - 관리자 대부분이 조교로 이루어져 있어 도서관 고유 업무인 목록에 대한 숙지가 낮은 편이다.
- 각 학과 및 연구소기관의 다양한 요구 조건 수렴
 - 각 기관마다 검색에 필요한 조건이 다양하다. (일본어 한글표기식 검색...)
- 각 학과 및 연구소기관의 대출정책에 따른 이용자들의 인식
 - 폐가제 및 신분별 대출 제한을 요구하는 기관들이 있다.
- 중앙도서관과 각 학과 및 연구소의 자료실 정책변화에 대한 인식
 - DB통합에 참여한 다른 기관에 대한 서로의 인식이 약하다.

5.2. 해결점

- 각 학과 및 연구소의 서지업무 담당자에게 주기적인 교육을 실시 업무내용의 차이 점을 줄여감
- 각 학과 및 연구소의 검색 조건등 요구 조건을 최대한 수렴하며 통일된 규칙에 대한 이해도를 높이 도록함
- 통합검색을 이용하는 이용자들에게 각 학과 및 연구소 대출정책의 차이는 서비스에 혼란을 줄 수 있으나 상호대차의 원활한 지원, 홈페이지를 통한 자세한 안내를 제공하여 이용자들의 불편을 최소화 시키도록 노력함
- 중앙도서관과 각 학과 및 연구소의 자료실 정책결정은 통합검색서비스를 기본으로 하는 이용자들에게 불편함을 초래 할 수 있으므로 자료실 정책결정시에는 자료실 담당자들간의 원활한 의사 교환을 충분히 논의 후 대처방안을 마련하여 이용자의 자료이용에 혼란이 없도록 해야함

6. 대학도서관 정보화 사업 소개

6.1. 2002년도 교육인적자원부 지원 서울대학교 서지DB구축 사업

구 분	참여기관	구축책수	소요액	비고
서지DB통합작업	50개 기관	217,000책	1억1천만원	부분적 기관 통폐합 및 신규자료생성작업
서지신규생성작업	10개 기관	129,000책	1억3천만원	
합	60개 기관	약346,000책	약2억4천만원	

6.2. 2003년도 교육인적자원부 지원 서울대학교 서지DB구축사업 계획

구 분	예상참여기관	구축책수	소요액
서지신규생성작업	35여개 기관	360,000책	2억5천만원
사업기간	2002년 5월 - 2002년 12월		

* 2002년도 서지DB통합 사업에 참여했던 각 학과 및 연구소의 이용자 및 관리자들의 적극적인 참여가 예상되며 서지DB구축사업과 더불어 교육 및 기타 서비스에 대한 지원이 더 많이 요구 되고 있다.