

전자도서관 구축에 따른 장서의 디지털 우선순위에 관한 연구 *

남 영 준 **

〈목 차〉

- | | |
|----------------|------------------|
| I. 서론 | 3) 디지털화 기준설계 |
| 1. 전자도서관의 현황 | 4) 디지털화 우선순위의 제안 |
| 1) 외국의 전자도서관 | II. 결론 |
| 2) 국내의 전자도서관 | 참고문헌 |
| 2. 국가디지털화 사업 | Abstract |
| 1) 소장 자료의 디지털화 | |
| 2) 디지털화의 기준 | |

I. 서 론

우리나라의 도서관계는 과거에 전혀 경험하지 못한 상황에 직면하게 되었다. 인터넷 자료라는 새로운 차원의 정보원이 구축되면서, 또한 인터넷이라는 새로운 정보전달수단이 개발되면서 원격지에서 웹브라우저를 통해 자료의 원문을 볼 수 있게 되었다. 이러한 상황변화에 따라 이용자들은 기존 도서관의 장서까지도 웹브라우저를 통해 열람할 수 있도록 요구하기에 이르렀고 일부 도서관들은 이

* 이 연구는 1998년도 전주대학교 인문과학연구소 지원에 의해 연구되었음.

** 전주대학교 문헌정보학과 조교수

2 圖書館學論集(第 29輯)

러한 욕구를 만족시키기 위해 실제로 서지데이터베이스와 전문데이터베이스를 구축하여 이용자에게 제공하기에 이르렀다. 대부분의 도서관들은 이러한 서비스를 제공하고 싶어하면서도 예산상의 문제와 인터넷상에서 이를 구현하는 기술의 부족과 저작권에 관한 사항 등 여러 제반 문제에 직면하여 선뜻 전문데이터베이스를 구축하지 못하고 있다. 그러나 각 도서관들이 디지털화(full text) 사업을 시작하지 못하는 것은 전문화에 따른 인력과 경비, 시간문제가 실제적인 이유이다.

그럼에도 불구하고 많은 도서관들이 전자도서관으로의 발전을 모색하고 있으며, 1차적으로 자관 목록을 데이터베이스화하였으며, 국내 대부분의 대학들은 목록 데이터베이스와 같은 이차정보를 인터넷을 통해 제공하고 있다. 한편, 1997년 말부터 국가적으로 고학력 실직자의 재취업을 위해 국가정보화 사업의 일환으로 일부 도서관에 소장하고 있는 한정된 책자를 디지털화하는 사업을 추진하고 있다. 이때 집행되는 예산은 극히 제한적일 수밖에 없으며, 따라서 특정 장서를 대상으로 디지털화 사업을 추진하고 있다. 현재 진행되고 있는 국가디지털화사업의 대상자료의 형태는 고서와 같이 실질적인 이용도가 떨어지고 있는 자료에 집중되고 있는 실정이다. 그 이유는 저작권이 가장 큰 원인으로 지적되고 있으나 이는 국가정책적으로 해결할 사항으로 판단되기 때문에, 실제적인 문제로는 어떤 형태의 어떤 주제의 자료를 우선 디지털화하여 국가경쟁력을 높일 수 있는지가 주된 관심사라 할 수 있다.

본 연구는 이점에 착안하여 국가전자도서관을 구축할 경우 어떤 주제와 어떤 형태의 자료부터 디지털화를 추진하는 것이 가장 바람직한 것인지를 제시하고자 한다. 이를 위해 국내외 전자도서관에서 추진하고 있는 디지털 대상의 자료를 조사하고, 기존 도서관의 장서구성, 이용자들의 정보요구현상을 조사한다. 이는 향후 본격화될 국가디지털화사업에 투자 우선 순위의 가이드라인으로 활용될 수 있을 것이다.

1. 전자도서관의 현황

전자도서관의 도래는 정보환경의 변화, 즉 정보기록(정보생산)매체의 변화, 정보관리 시스템의 변화, 정보서비스(정보유통)의 변화와 함께 컴퓨터의 성능발전에서 비롯되었다. 전자도서관의 정의에 대해 정영미는 전자도서관을 인쇄물 형태의 정보뿐만 아니라 디지털 형태의 원문정보를 제공하는 새로운 도서관이라는 의미와 네트워크를 통해 디지털 형태의 원문정보를 제공하는 분산형 컴퓨터 시스템이라는 의미로 설명하고 있다.¹⁾ 즉, 전자도서관이란 디지털 형태의 자료를 웹 브라우저와 같은 도구를 사용하여 언제든지 원하는 자료의 원문을 열람하는 기능을 가진 도서관을 의미한다. 이러한 특성 때문에 대부분의 자료에서 전자도서관을 전자도서관(Electronic Library), 디지털도서관(Digital Library), 가상도서관(Virtual Library), 논리적 도서관(Logical Library), 벽없는 도서관(Wallless Library) 등과 같이 다양한 용어로 표현되고 있으나, 실제적으로 전자도서관의 명확한 개념이 정확하게 정립되어 있지 않은 상황이다.

이러한 다양한 관점의 개념을 종합하여 보면, 전자도서관이란 전통적인 도서관의 기능을 근간으로 하여 전자적인 형태의 정보, 즉 디지털 형태의 멀티미디어 정보의 수집, 가공, 유통에 있어 컴퓨터와 통신망을 기반으로 하여 이용자로 하여금 공간적, 시간적 제약 없이 양질의 정보를 획득할 수 있게 하는 정보서비스 체계라고 정의할 수 있다.

한편, 국가전자도서관 구축이라는 차원에서 전자도서관에 대한 정의는 앞서 언급한 전자도서관의 정의와는 다소 차별성을 지니게 된다고 본다. 물론 원론적 의미에서의 '전자도서관'의 개념에는 변함이 없으나 국가전자도서관 구축사업에서는 다음과 같은 다소 협의적 의미를 지닌다.

국가전자도서관의 개념은 완성단계에서는 포괄적 의미의 전자도서관과 같은 의미를 갖게 될 것이지만, 초기 또는 발전단계에서 국가전자도서관은 주로 어문 저작물을 디지털화하여 원문을 시간적 공간적 제약없이 이용할 수 있도록 하여

1) 정영미, 안현수. **전자도서관구축론**. 구미무역(주) 출판부. 1998. pp.155-156.

필요한 최종 정보를 이용자가 손쉽게 획득할 수 있는 새로운 도서관 시스템이라 할 수 있을 것이다.²⁾

이러한 편리성 때문에 전세계적으로 자료의 디지털화를 통한 전자도서관의 구축을 계획하고 실행하고 있는 실정이다. 전자도서관을 국내외로 구분하여 각 도서관들이 구축하고 있는 디지털화 대상자료를 살펴보면 다음과 같다.

1) 외국의 전자도서관

전자도서관에 대한 관심과 시도는 외국에서부터 시작되었다. 미국은 1990년대 초 IBM사가 디지털 도서관 시스템을 개발한 이후 이에 대한 연구가 지속적으로 발전되었으며 1994년 9월부터 미국과학재단(NSF : National Science Foundation) 과 의회도서관(LC : Library of Congress) 그리고 중요 명문대학이 중심이 되어 국가 전자도서관 사업부터 원문검색을 위한 자료의 디지털화 사업이 본격적으로 이루어졌다. 그 후 캘리포니아주립 버클리대학교를 비롯한 6개 대학교가 선정되어 본격적인 전자도서관사업이 1994년부터 1998년 8월까지 4년동안 총 2,440 만불의 연구비로 이 사업이 진행되었다. 이를 디지털화 대상주제나 형태를 중심으로 자세히 조사하면 다음과 같다.

(1) 카네기멜론대학교³⁾

Carnegie-Mellon University 도서관은 미국정부의 지원으로 전자도서관을 구현하였으며, 이를 머큐리프로젝트라 명명하였다. 이 프로젝트의 기본 개념은 1988년말부터 시작되어 교수나 학생들이 캠퍼스 어느 장소에서라도 어떤 컴퓨터를 이용하든 상관없이 전자문서를 찾을 수 있도록 하자는 것이었다. 초기 디지털화 대상자료로는 이미지 형태로 38개의 논문과 7개의 컴퓨터 관련 논문지, 미국 내 주요 여섯 개 대학의 컴퓨터학분야의 기술논문, 미국 전자공학회의 컴퓨터관련 논문지 등을 선정하여 통신망을 통해 온라인으로 정보를 제공하고 있다. 향후

2) (주)쓰리소프트. 국가전자도서관 구축 기본계획. 보고서. 국회도서관. 1997. pp.8-9.

3) <http://www.informedia.cs.cmu.edu/>

이 대학도서관은 소프트웨어 분야의 논문지를 주로 디지털화할 계획이다.

(2) 일리노이주 어바너-삼페인대학교⁴⁾

University of Illinois at Urbana-Champaign 도서관은 디지털화에 대한 주요 대상으로서 공학분야의 학술지와 잡지를 선정하였으며 이를 SGML을 이용하여 구조화하였다. 디지털화 대상 주제로는 일반공학잡지, 전문공학저널, 전문과학저널이며 상업 출판사나 전문출판사를 통하여 입수된 자료도 전자화 대상으로 선정하였으며 디지털화 작업은 상업출판사등에서 구축된 것을 임차하는 형식으로 활용하고 있다. 특징적인 것으로는 학술지와 논문지에서 추출한 그림, 수식, 테이블도 SGML문헌으로 생성·처리하는 연구에 주력하고 있다.

(3) 캘리포니아주립 산타바바라대학교⁵⁾

UC-Santa Barbara 도서관은 지리정보전자화프로젝트의 명칭으로 The Alexandria Project라는 프로젝트명을 사용하였다. 이 프로젝트는 지리정보시스템용 소프트웨어를 이용한 미국 캘리포니아주의 주요지역 가운데 1차적으로 SantaBarbara, Ventura, LA 등의 도시 관련 지도정보와 이미지 및 공중사진들을 디지털화하였다. 이 프로젝트의 주요 결과물은 지역적으로 분산되어 있는 도서관의 장서에 역시 지역적으로 분산되어 있는 이용자들이 접근할 수 있는 분산형 테스트베드시스템으로 운영되고 있다. 특징적인 것으로는 일반 문헌의 디지털화에 대한 구체적인 연구결과는 제시되지 않고 있다는 점이다.

(4) 스탠포드 대학교⁶⁾

Stanford 대학교의 전자도서관 프로젝트는 미래 디지털 도서관에서 구현될 인터페이스와 시스템에 관련된 소프트웨어 기술을 중점적으로 개발하는 것이다.

4) <http://www.uic.edu/depts/lib/>

5) <http://alexandria.sdc.ucsb.edu/>

6) <http://digilib.stanford.edu/>

6 圖書館學論集(第 29輯)

즉, 이용자가 쉽게 정보를 이용할 수 있고 텍스트, 화상, 오디오, 비디오, 3차원 모델을 비롯한 메타데이터까지 취급할 수 있는 통합환경을 수용할 수 있는 시스템을 개발하는 것이다. 따라서 이 대학의 전자도서관계획도 일정한 문헌자료를 디지털화하기 보다는 이용자인터페이스에 연구의 초점이 집중되고 있다. 이 도서관에서 실험을 위해 구축한 디지털화 자료는 다음과 같다.

- CNRI CS 기술보고서 : 동대학을 포함한 5개 미대학 전자도서관 구축 콘소시움에 참가한 대학이 최근 발표한 컴퓨터공학기술보고서
- Dialog사 소장 전자자료중 일부 : Dialog사가 소장하고 있는 전자자료가운데 경영, 과학 및 공학자료중 일부
- O'Reilly Associates출판사의 전자 단행본 및 저널중 일부
- 인터넷에 있는 학술자료 : WAIS상에서 입수한 자료
- 스탠포드대학에서 제공하는 전자자료 : INSPEC(전기전자, 물리, 컴퓨터공학 초록데이터베이스), ABI(경영학분야 정기간행물 초록), LEXIS-NEXIS(법학관련 자료), EI Compendex(기술 및 과학분야 초록집), Oxford 사전
- ACM(The Association for Computing Machinery) 출판물중 일부
- 미항공우주국 소장자료중 일부
- 지역자료(local document) 및 이미지정보

(5) 미시간주립 앤아버대학교⁷⁾

University of Michigan의 전자도서관은 멀티미디어 디지털 도서관의 연구개발을 위해 기타 연구기관들과 공동으로 협력하는 방안을 모색하는 연구가 중점을 이루고 있다. 특히, 정보서비스에 있어 에이전트 기술의 적용에 대해 중점적으로 연구하고 있으며, 이를 실험하기 위한 디지털화 대상 데이터베이스로 지구학과 우주공학 분야를 선정하였다. 1차적으로 디지털화된 자료는 CORE, MIRLYN 프로젝트에서 이미 구축한 단행본과 잡지를 대상으로 하였다. 구축의

7) <http://www.umich.edu/UMDL/>

목적은 대상자료를 이미지 형태로 하여 전문데이터베이스를 구축하고, 이 DB를 대학 구성원들로 하여금 대학내에서만 제한적으로 이용하게 하는 것이었다. 이 데이터베이스에는 박물관 圖錄을 사진파일로 구축하였으며, 인문과학분야의 참고자료와 단행본들을 SGML형태로 구축하였다. 또한, 사회과학분야의 기초통계 자료도 봉사대상에 포함되었으며, UMI의 Dissertation Abstracts도 봉사대상에 포함되어 있다.⁸⁾ 동대학에서 테스트베드용 실험용 디지털 장서를 구축하였으며, 아스키 형태와 TeX형식으로 마킹된 형태로 구축하였다. 그 내용은 다음과 같다.

· IEEE(미국전기전자공학회)을 비롯한 6개 학회에서 발행한 정기간행물중 일부와 John Wiley & Sons사, US News & World Report와 같은 상업출판사에서 기증한 디지털화자료

- 미항공우주국과 연결되어 대학에서 공유하고 있는 자료 가운데 비트맵형태로 된 학술지 일부
- UMI와 IEEE/IEE 자료가운데 일부
- Mosaic과 Gopher, WWW에 등재된 자료가운데 가치가 있는 자료 등.

(6) 버지니아주립 대학교

VIRGO는 버지니아 대학교의 온라인 정보시스템으로서 1992년 8월 이후 전자문헌센터(The Electronic Text Center)를 운용하고 있다. 전자문헌센터는 전자문헌 작성 표준인 SGML을 사용하여 온라인 데이터 베이스를 제공하고 있다. 디지털화의 대상은 크게 희귀본이나 유일본 자료들을 수집·정리하였으며(1단계), 2단계부터는 백만페이지가 넘는 원문, 사진 등 기타의 자료를 이미지 형태로 축적해 인터넷을 통해 원문정보를 제공한다. 2단계에 포함되는 디지털화 대상자료는 다음과 같다.

- 버지니아 식민지자료와 연관된 50만 건의 인명자료 및 2만8천 페이지에 달하는 보고서 원문

8) 임영선. 전문데이터베이스의 이미지 화일과 텍스트 화일의 검색 효율성 비교 연구. 중앙대학교 대학원. 미간행 석사학위논문. 1997. pp.23-24.

8 圖書館學論集(第 29輯)

- 식민지시대에 작성한 85만장의 舊목록카드
- 버지니아주의 신문에 대한 서지정보
- 마이크로필름에 저장된 40만 건의 특허정보
- 마이크로필름형태의 버지니아주와 관련시의 법정자료
- 1853년부터 버지니아주민들의 가족성경 기록일부

(7) 벨연구소

The RightPage™ Service 전자도서관 프로젝트는 AT & T Bell Lab.에서 개발한 것으로서 잡지를 이미지화하여 미국 전역의 50-100개 Bell Lab. 연구원들에게 제공하는 것이다. 잡지의 주제분야는 일반 과학기술분야의 신착자료이며 잡지의 목차와 기사의 해당 페이지를 이미지 형태로 저장하고 있다. 저자와 출판 정보와 같은 일반 서지정보와 목차, 서문은 웹브라우저로 볼 수 있으며 본문내용은 Gopher로만 연결하여 다운받을 수 있다.

(8) 조지아주 도서관망

GALILEO(Georgia Library Learning Online)는 미국 조지아주에 있는 34개 대학교들의 도서관을 연결한 전자도서관 시스템으로서 다양한 유형의 전자 정보 자원을 제공하고 있다. GALILEO를 통해 온라인으로 제공되는 정보자원의 유형은 다음과 같다:

- 2,500여 개의 핵심 학술지(academic journals)에 대한 색인초록 서비스
- 1,100여 개의 핵심 학술지에 수록된 논문기사 전문
- 연구잡지(research journals)
- 조지아주 센서스 데이터
- 조지아주의 주 출판물(state publication)
- OCLC의 FirstSearch 정보서비스
- Britannica Online
- 온라인 종합목록

핵심 학술지에 대한 색인초록 서비스는 UMI사의 색인초록 데이터베이스인 Periodical Abstracts가 있으며, Research II는 전분야를 망라하여 1,600여개의 잡지를 수록하고 있으며 이 가운데 650여개의 잡지에 대해서는 전문을 포함하는 데이터베이스이다. ABI Inform은 경제 및 비즈니스 분야의 1,000여개의 잡지를 수록하며, 500여개에 대해서는 전문을 포함하고 있다. GALILEO는 두 데이터베이스를 통해 2,600개 이상의 잡지에 대한 이차정보와 1,100개 이상의 잡지 논문 기사에 대한 원문을 제공하고 있다.⁹⁾

(9) 미화학회

CORE 프로젝트(Chemistry Online Retrieval Experiment Project)는 화학관련 학술지를 대학이나 연구소에 전자형태로 배포하기 위한 프로젝트이다. 학술지 기사는 텍스트 형태와 비트맵 형태로 저장되며 다양한 인터페이스 기능을 사용하고 있다. 디지털화 대상자료로는 미화학회(American Chemical Society : ACS)에서 발간하는 20여종 잡지의 5년분 자료(약 425,000페이지)이고, 이미지 파일과 SGML로 표현된 아스키 버전형태로 제공하고 있다. 그 밖의 제공되는 디지털 문헌으로 Britannica Online, E-Journals등이 포함되어 있다.

(10) 일본 학술정보센터

일본의 학술정보센터(National Center for Science System : NACSIS)는 일본의 학술기관 및 대학이 소장하고 있는 자료들을 디지털 형태로 국민들에게 제공하는 것을 목적으로 하고 있다. 주된 디지털화 대상으로는 일본 학·협회에서 발행하는 학회지와 논문집을 대상으로 전문데이터베이스를 구축하는 것이며, 1997년 12월 현재 인문, 경제, 순수과학, 공학, 농학, 의학 분야의 총 33개 학·협회에서 발간하는 52종의 학술지가 디지털화 대상으로 선정되어 구축되고 있다.

9) 안현수, 정영미. 전자도서관구축론, 구미무역(주) 출판부, 1998, pp.183-184.

(11) 일본 국회도서관

일본국회에서 추진하고 있는 전자도서관화 사업은 자료의 급증에 따른 서고부족문제, 도서관 서비스의 새로운 패러다임의 제시 등을 위해 Kansaikan프로젝트를 추진하고 있다. 이는 일본국립국회도서관을 비롯한 16대학 도서관에 대한 소장자료사항과 일본국회도서관 소장의 국회심사용 자료, 명치시대의 사회과학분야의 자료, 현대일본정사문헌 등 에도시대의 민속자료, 훼손으로 보존의 필요성이 높은 자료와 연속간행물, 신문 등을 우선 디지털화할 대상으로 선정하였다. 이 프로젝트는 디지털화의 목적이 對국민봉사부분보다는 史料에 대한 보존차원의 성격이 상대적으로 높은 사업이며, 특히, 서고공간 확대 및 자료의 분산보관에 의한 훼손위험을 방지하는데 중점을 두고 있다.¹⁰⁾

(12) 영국 국립도서관

영국국립도서관(The British Library)은 Initiatives of Access Projects계획에 의거하여 20여개의 하부 프로젝트를 수행하고 있다. 영국에서의 디지털화 사업은 대체적으로 보관적 관점에 중점을 두고 있다. 특히, 현대자료에 대한 디지털화 사업에 대한 구체적인 데이터는 아직 제시되지 못하고 있다. 현재 이루어지고 있는 디지털화 사업을 정리하면 다음과 같다.

① The Patent Express Jukebox : 영국국립도서관 소장 일부 특허자료(130만 건)를 16개로 이루어진 CD-ROM Jukebox연결을 통하여 이용자에게 서비스하고 있다.

② Electronic Beowulf : 영국국립도서관 소장 11세기 필사본 자료를 디지털 형태로 변환한다. 이는 역사적 사료를 보관하기 위해 디지털화를 시도하고 있다. 특히, 원본이 손상된 자료나 고대 희귀자료의 필사본을 디지털 카메라로 스캐닝하여 선명한 이미지를 갖는 디지털 형태의 자료로 저장한다.

10) Nakano, S. "The Construction Plan of the Kansaikan of the National Diet Library and Digital Library," *In Proceedings of International Symposium on Digital Library*, 1995. Tsukuba. University of Library and Information Science. p.85

③ Electronic Photo Viewing System : 약 1만 건의 사진 자료를 디지털화하고 하이퍼링크를 통한 자료와 해당자료에 대한 간략한 설명기사를 제공한다. 600만 건 서지레코드를 국제적 수준으로 확장을 계획하고 있다.¹¹⁾ 즉, 책자형 자료의 디지털화보다는 사진과 도서관이 소장하고 있는 서지레코드에 대한 서지데이터 베이스를 구축하는데 중점을 두고 있다.

(13) 프랑스 국립도서관

프랑스국립도서관(Bibliothèque nationale de France : 이하 BNF)은 1994년 1월 프랑스도서관과 파리국립도서관이 합쳐지면서 설립된 프랑스 대표도서관으로서 전자도서관 프로젝트를 관리하고 있다. BNF에서는 여러 기관에서 분산 소장되어 있는 도서, 화상, 음성 등 다양한 형태의 자료를 디지털화하고, 네트워크를 통해 이를 이용자에게 제공하고자 1990년부터 디지털 장서(digital collection)를 구축하기 위한 작업을 시작하였다. 1998년까지 10만권의 도서와 30만 건의 이미지에 대한 디지털화를 목표로 하는 Tolbiac 프로젝트에서 대상 자료로서 우선 선정된 것은 과학, 프랑스문학, 언어학 분야의 참고자료와 정기간행물, 그리고 고대에서 현대까지의 역사 분야자료이다. BNF는 자료의 디지털화에 있어, 최근에 출판된 40%의 자료는 텍스트형태의 자료들을 대상으로 제한하였으며, 60%의 자료는 마이크로 필름형태를 디지털화하였다. 이 가운데 일부는 출판사로부터 구입한 것이 포함되어 있다. 한편 디지털화하는 형태는 이미지 형식과 텍스트 형식 중에서 비용과 현재 기술을 고려하여 이미지 형식을 채택하고 있다.¹²⁾

2) 국내의 전자도서관

(1) 국립중앙도서관

국립중앙도서관은 1400여종의 CD-ROM의 주제 및 형태 국립중앙도서관에 납본되는 출판물의 특성을 분석하여 납본된 자료 가운데 인문과학 분야가 가장 많

11) (주)쓰리소프트. 국가전자도서관 구축 기본계획. 보고서. 국회도서관. 1997. pp.58-59.

12) (주)쓰리소프트. 국가전자도서관 구축 기본계획. 보고서. 국회도서관. 1997. pp.56-57.

은 부분을 차지하고 있으므로, 그 특성에 따라 1차 디지털화 분야는 인문과학분야로 하였다. 또한, 그 외 장서의 구성의 특징에 따라 고전적(古典籍)과 고신문을 대상으로 디지털화를 수행하고 있다. 시범 서비스되는 내용은 국립중앙도서관이 보유한 약 65만 건의 도서목록 자료에 대한 서지 사항 및 200종의 주요 학술자료 즉, 학위논문 107종, 학술지 43종, 고서 50종에 대한 원문화상정보를 디지털형태로 완료하여, 인터넷을 통해 원문정보를 제공하고 있다.

(2) 연구개발정보센터(KORDIC)

동 센터는 1978년부터 국가가 지원하는 연구개발사업대안 연구보고서 및 관련정보를 데이터베이스로 구축하였다. 구축의 목적은 적시에 산·학·연의 수요자들에게 제공할 수 있는 체제를 구축하기 위해 과기부 산하 과제관리기관인 과학기술정책연구소와 한국과학재단으로부터 지원받는 연구개발정보에 대한 서지사항(초록포함) 및 원문정보 DB를 구축하는 것이다. 납본협력부처는 정보통신부와 농림수산부, 환경부, 교육부, 보건부 등이 있으며, 이들 부처로부터 납본받은 자료에 대한 서지사항 및 원문정보를 DB로 구축하고 있다. 기술보고서의 경우는 전문으로 제공하는 건수는 18,741건이며, 서지정보를 제공하는 형태는 19,529건이다.

학술지의 경우는 컴퓨터 및 정보통신(334종)을 비롯하여 기계 및 선박(362종), 신소재(264종), 기초과학(366종), 항공우주(16종), 생명공학(205종), 에너지(72종), 화학(74종), 전기(75종), 자원(105종), 원자력(277종), 표준 및 측정(48종), 해양 환경(176종)으로 총 2,149종에 대한 디지털화 사업이 완료되었다.

석·박사학위논문은 국내 이공계 석·박사 학위논문으로 제한하고 대상기관은 한국과학기술원, 포항공과대학교, 광주과학기술원, 충남대학교의 학위논문을 제공하고 있다. 4개 대학이 보유하고 있는 석·박사 학위논문 약 10,000편에 대해 서지사항은 텍스트 형태로, 원문은 이미지 및 전자형태로 DB를 구축하였다. 원문의 경우 현재 각 대학은 종이로 인쇄된 책자 형태 학위논문과 아래아한글, MS WORD, TeX, 기타 워드프로세서 등으로 작성된 디지털 형태로 학위논문을 보유하고 있다

(3) 한국과학기술원(KAIST)

한국과학기술원은 1995년부터 국내 과학기술분야 학위논문 및 연구보고서의 초록/원문정보를 웹상에서 서비스하기 해당 자료를 디지털하고 있다. 현재 구축된 디지털장서의 현황은 다음과 같다.

- 석/박사 학위논문의 서지 및 초록정보 전체수록(약13,000여책)
- 전문정보(SCML) 시범데이터: 학위논문(82책), 교수연구논문(115편)

이상에서 살펴본 국내 전자도서관과 기타 도서관의 구축사례를 표로 조사하면 <표 1>과 같다. <표 1>에 나타난 바와 같이 국내 디지털화사업을 추진하는 기관에서 구축한 전문자료는 대부분 특수자료로서 단행본보다는 상대적으로 입수하기 어려운 귀중본, 학위논문, 보고서 등이 포함된다. 특히, 대부분의 국내 기관들은 연구용 자료(학회지, 보고서, 석·박사 학위논문)를 대상으로 디지털화 사업을 추진하고 있는 행태를 확인할 수 있다.

<표 2> 각 기관별 DB 구축 정보

기 관 명	이용가능 DB 정보
국립중앙도서관 (http://www.nl.or.kr)	국가문헌종합목록정보와 고서, 귀중본 전문DB
국회도서관 (http://www.nanet.go.kr)	입법부 발간자료와 국정감사 의원요구자료 전문DB
한국과학기술원 과학도서관 (http://libr0.kaist.ac.kr)	KAIST 석·박사 학위논문 목록·초록정보와 전문DB, 교수연구논문목록정보와 본문 DB
한국과학기술원 연구개발정보센터 (http://www.kordic.re.kr)	과학기술처 지원 연구보고서 목록정보와 전문DB, 정보과학회지 본문 DB
한국학술진흥재단 (http://www.kric.ac.kr)	외국취득 박사학위논문 목록정보와 전문 DB
LG 상남도서관 (http://www.lg.or.kr)	화학, 전기, 전자분야의 학술지 1,000여종, Proceeding 자료 200건내외의 이미지 DB

Ⅱ. 국가디지털화 사업

우리나라는 정보와 지식이 제반 경제활동의 중심이 되는 새로운 경제체제에서 국가경쟁력을 고양하고 국민들의 삶의 질을 높임과 아울러 궁극적으로 선진국으로 도약하기 위해 1993년 이후 초고속 정보통신기반 구축사업을 시작하여 21세기 정보화사회로의 연착륙을 준비하고 있다. 이를 위해서는 정보공유가 필연적이며 국가차원에서 생산된 정보들을 공유할 수 있도록 하는 새로운 형태의 전자도서관 구축을 시도하고 있다. 그러나 전자도서관 구축은 소급자료의 디지털화를 전제로 하고 있으며, 여기에는 많은 양의 인력과 재원이 투입되어야 한다. 한편, IMF라는 외부환경의 변화에 따라 사상최대의 실업률을 기록하고 있음에 따라 정부에서는 실직자 혹은 고학력 미취업자를 해결하고 국가디지털화 사업을 동시에 충족시킬 수 있는 정보화 근로사업을 1997년말부터 시작하고 있다.

정보화근로사업은 국가사회의 구조조정과정에서 발생하는 고급 실업인력의 재취업을 지원하고 지식·정보사회를 주도할 인적자원의 육성과 전업기반을 조성함으로써 산업구조조정을 촉진하고 정보사회를 조기에 실현하기 위한 사업이다. 사업의 주체는 국가이면서 이 사업이 목표로 하고 있는 것은 종이형태를 디지털 형태의 자료로 변환하여 국가에서 생성된 정보를 인터넷과 같은 통신망상에서 시간과 장소에 구애없이 손쉽게 정보를 입수할 수 있도록 하는 것이다. 이 취지에 따라 각 부처에서 수행하고 있는 사업내용과 시행기관을 분석하면 <표 2>와 같다. 사업의 형태는 부처별로 차이가 있지만 대부분 부처별로 소장하고 있는 기록자료를 아스키나 혹은 이미지 형태로 변환하는 것이 주를 이룬다. 대법원을 비롯한 몇몇 기관에서는 대상자료를 이미지 형태의 파일을 구축하고 있으며, 일반 텍스트 파일에 대한 디지털화 작업은 국립중앙도서관과 국회도서관과 같이 전통적으로 책자형 정보를 소장하고 있는 기관에서 이루어지고 있다.

〈표 2〉 정보근로화사업 과제 및 시행주체

시행기관	사업내용	시행기관	사업내용
대법원	부동산등기 전산화	정보통신부	정보화지원사업 영상자료디지털화 Y2K 교육 및 해외취업
국 회	전자도서관		
농촌진흥청	농업기술정보 DB화		
문화관광부	전자도서관	건설교통부	주제도 수치지도화 지하매설물 수치지도화 지형도 전산화
통계청	고용통계인프라 구축	행정자치부	건축물대장 DB화

1. 소장 자료의 디지털화

소장 자료의 디지털화는 국가전자도서관 구축사업 중에서 자원의 중복투자 소지가 가장 큰 부분이다. 그러므로 이 사업을 하나의 기관에서 전담하는 것은 디지털 형태로 구축해야 할 자료자체가 많고 방대하고, 또한 이에 소요되는 비용 및 인력 문제 때문에 현실적으로 불가능한 일일 것이다. 또한 한 기관에서 국가정보자원의 디지털화를 전담하기에는 디지털화 된 정보의 질적인 측면에서도 문제가 발생할 수 있다. 따라서 국가자원의 디지털화는 내용 및 자료의 특성을 고려해서 각 주요추진기관이 분야별 또는 형태별 등으로 분담하여 디지털화하는 것이 비용 효과를 극대화할 수 있으며, 디지털자료의 질을 높일 수 있을 것이다.

따라서 가능한 한 정보생산을 담당했던 기관에서 자신의 정보를 디지털화하는 것이 바람직할 것이다. 한편 앞에서 조사한 바와 같이 자신이 소속된 기관에서 생산된 정보와 무관하게 일부 기관에서는 이미 디지털화를 완료하거나 혹은 추진 중인 자료도 있으므로 이에 대한 면밀한 조사없이 디지털화를 시도하는 것은 막대한 예산 낭비를 초래할 수 있다. 디지털화의 중복입력 등과 같은 사항이 문제로 지적되는 이유를 파악하기 위해 디지털화 방법에 대해 조사하면 다음과 같다.

1) 디지털화 방법¹³⁾

자료를 디지털화하는 방법은 크게 이미지 파일형태로 저장하는 것과 텍스트 파일형태로 저장하는 것으로 크게 구분될 수 있다. 이를 구체적으로 살펴보면 다음과 같다.

① 키인 방식

디지털화 방법 가운데 가장 원시적이고 정확한 입력의 질을 확보할 수 있는 방안 가운데 하나가 키인(key-in) 방식이다. 이 방식은 디지털화 대상으로 선정된 데이터를 사람이 직접 키보드를 사용하여 입력하는 방식이다. 이 때 소요되는 경비는 자료 입력비와 입력자료 검수비로 구분된다. 이 때 소요되는 경비를 분석하면 다음과 같다. 일반적으로 시중에서 판매되는 학술문헌의 본문 양은 일반적으로 약 300페이지의 규모로 구성되어 있다. 한 페이지는 200자 원고지 45매 정도의 분량이다. 이를 원고지 4매로 환산할 경우 한 권의 학술문헌의 양은 200자 원고지 2,400매 정도의 분량을 유지한다. 이를 문자수로 환산하면 대략 480,000문자에 해당한다. 이를 전산입력한다고 가정할 경우에, 1분에 400타(문자)를 입력하는 전문입력자의 경우에 한 권의 문헌을 입력하는데 대략 1,200분이 소요되며, 시간단위로 환산하면 학술문헌을 대상으로 한 권을 입력하는데 약 20시간이 소요된다. 이 디지털화의 방법의 장점은 학술문헌에 나타난 특수기호 혹은 한자, 외국어원문을 입력자가 그대로 입력할 수 있기 때문에 입력의 질이 가장 안정적이다.

② 스캐너 활용방식(이미지 파일)

디지털화에 인력 및 예산을 절감하기 위해 하드웨어를 사용하는데, 그 방안 가운데 가장 일반적인 것이 스캐너를 활용하여 문자 인식을 하는 것이다. 이 방식은 스캐너와 문자인식소프트웨어에 따라 효율이 크게 차이가 발생한다.

스캐너 활용방식 가운데 수작업으로 하는 부분은 ① 자료의 선정을 비롯하여 ② 스캐닝, ⑤ 결과 편집, ⑥ 구축자료검수이며 이에 대한 수작업은 정확히 구분

13) 남영준. 국어정보처리기반구축사업 저작권 해결을 위한 연구. 보고서. 문화체육부, 1997. pp.33-38.

된 데이터는 없으며 통계적으로 ①~④ 과정의 경우 하루에 약 2권 정도를 처리할 수 있다. 특히, 이 작업은 기계의 성능에 크게 좌우된다.

i) 단면 플랫폼형 스캐너를 활용할 경우

스캐닝 방식에 소요되는 경비를 수작업 방식과의 비교를 위해 10권의 데이터를 디지털화하는데 소요되는 시간을 산출하면 다음과 같다. 스캐너로 한 면을 인쇄에 소요되는 시간은 약 30-60초이다. 300페이지 분량의 책 한 권을 스캐닝하는데 소요되는 시간은 약 2시간 30분이 소요된다. 그러므로 책 2권을 스캐닝(흑백 300dpi 수준)하는데 소요되는 시간은 대략 5시간이 소요된다 (150면(2페이지) X 60초 = 9,000초 = 150분 = 2시간 30분)

문자인식에 소요되는 시간은 150면의 데이터(그래픽)를 인식하는데 대략 2시간 정도가 소요된다. 이 방식의 장점은 전문 입력가의 수고를 덜어줄 수 있을뿐더러 문자인식과정에는 인간의 노력이 최소화됨으로서 인력 낭비를 최소화할 수 있다는 점이다.

일반적으로 소요되는 경비는 키인 방식의 2/5정도의 비용과 노력이 투입된다.

ii) 양면 스캐너를 활용할 경우

현재 국외에서 시판되고 있는 가장 일반적인 양면 스캐너는 1분에 약 200매(200dpi 수준)를 스캐닝하며, 하루에 8권 이상을 인식할 수 있다. 이 방식의 단점으로는 양면 스캐닝을 위해서는 반드시 분석대상이 되는 책등(書背 : book spine)을 잘라야 하는 책의 파손이라는 결정적인 문제점을 갖고 있다. 또한, 양면 스캐너의 가격이 플랫폼형의 4배가 넘는 하드웨어의 가격이 높다는 단점도 있다. 이미지 파일구축방식의 가장 큰 장점은 원자료의 형태를 그대로 유지할 수 있는 점이며, 또한 복잡한 수식도 원문대로 표현할 수 있는 점이다.

iii) 전자자료입수방법

전자자료입수방법은 디지털 자료구축에 가장 순위우변서도 가장 경제적인 방법 가운데 하나이다. 전자자료는 크게 영리를 목적으로 하는 전자도서와 비영리를 목적으로 일반 통신선상에 공개된 자료이다. 통신선상에서 열람할 수 있는 디지털자료는 소설류와 만화, 수필류로 크게 구분된다. 소설의 경우는 무협소설,

SF소설, 일반소설류로 구어체의 자료와 픽션적 성격이 강한 자료로 구성되어 있다. 이에 비해 수필류는 논픽션적인 자료가 많으며, 만화는 국내에서 생산된 자료가 대부분이다.

각 기관에서 수집한 전자도서는 상업용이기 때문에 최신간보다는 10년전 자료부터 최근의 자료로 구성되었으며, 또한 많이 읽히는 것이 목적이기 때문에 대부분 대중성이 강한 자료로 구성되어 있는 특징을 보이고 있었다. 온라인으로 입수 및 열람할 수 있는 자료는 회사마다 차이가 있지만 약 600여종으로서 2,400권정도 수준이다.

(2) 디지털화 계획의 기본방향

도서관 자료를 디지털화하는 것은 기존의 도서관에서 사용되고 있는 자료의 선택과정과 동일하다. 문헌선택의 일반적 원칙은 첫째, 도서관의 설립목적에 부합되는 자료, 둘째, 도서의 내용적 특성을 고려할 것, 셋째, 일반인들이 읽어서 도움이 되는 것, 넷째, 양질의 책을 중복해서 소장한다. 다섯째, 적극적으로 이용하는 책을 우선한다. 여섯째, 국가적 책임이 있는 자료는 이용도가 떨어지더라도 선택할 것. 일곱째, 도서관의 특색이 되는 전문적 취미적 장서도 고려한다. 여덟째, 특정 이론에 치중된 자료는 제외한다. 아홉째, 장서구성에 있어 일정부분은 창작 예술을 위해 소설이 일정 부분에 포함되도록 한다. 열 번째, 도서관의 목적에 합치하는 도서로서 보관할 만한 가치가 있는 자료로 한다. 열한번째, 이용자 계층에 대한 정확한 해석이 있어야 한다.¹⁴⁾ 이러한 기존 도서관의 장서구성원칙 가운데 '양질의 책을 중복해서 소장할 것'이라는 취지의 경우는 전자도서관의 특성을 고려할 경우 근본적으로 해결될 수 있기 때문에 이점을 제외하고는 전자도서관 구축과정에도 이 원칙은 그대로 적용될 수 있다. 단, 저작권에 관한 사항은 기존 도서관에서 소장한 자료들은 저작권에 대해 법적으로 면책이 되었지만 전자도서관에 구축된 자료는 영원히 원하는 누구에게라도 무한정 제공될 수 있

14) 임태삼. **도서선택의 이론**. 학문사. 1983. pp.79-82

기 때문에 저작권의 직접적 침해 요인을 갖고 있는 것이다. 따라서 이 조항의 검토가 필요하다.

한편, 이러한 전통적인 기준외에 현대 국가전자도서관 구축사업에서 디지털화해야 할 자료의 선정기준은 우선 다음과 같이 제시하고 있다.¹⁵⁾

첫째, 공공성을 지닌 자료이어야 한다. 국가전자도서관 구축의 초기에는 다수의 이용자가 관심을 갖는 공공성 있는 자료들을 우선 디지털화하여 제공하여야 한다.

둘째, 효용성을 가진 자료이어야 한다. 다수의 사용자가 관심을 가지고 있더라도 가치가 떨어지거나 단순히 오락을 위한 자료라면 국가전자도서관에서 제공할 필요는 없을 것이다.

셋째, 보존가치성이 있는 자료이어야 한다. 자료의 디지털화는 많은 시간과 비용을 필요로 하는 작업이므로 일시적으로는 가치가 있지만 시간이 조금 흐르면 별로 필요치 않을 자료의 디지털화는 자원의 낭비라고 할 것이다.

넷째, 전문성이 있는 자료이어야 한다. 이용자가 손쉽게 구할 수 있는 자료보다는 비교적 구하기가 어려운 전문성을 갖고 있는 자료들을 우선 디지털화하여 제공하는 것이 보다 효과적이라 할 것이다.

위와 같이 전통적인 장서구성의 원리와 일반 전자도서관의 원리에 따라 소급 자료디지털화의 기본 방향을 정리하면 아래와 같다.

- ① 이용자들이 필요로 하는 자료를 대상으로 한다.
- ② 도서관의 특성을 지닐 수 있는 분야의 자료를 대상으로 한다.
- ③ 전문성이 있는 자료를 대상으로 한다.
- ④ 일부이지만 보존적 가치가 있는 자료를 대상으로 한다.
- ⑤ 기대효과가 큰 자료를 대상으로 한다.
- ⑥ 정부가 국민에게 제공하여야 하는 공공적 자료를 대상으로 한다.

이러한 기본적 방향과 디지털화 대상 자료의 난이도에 따라 보다 세부적인 기준

15) (주)쓰리소프트. 국가전자도서관 구축 기본계획. 보고서. 국회도서관. 1997, pp.78-81.

을 세워 객관적인 기준을 측정한 후에 디지털화 대상자료를 선정해야 할 것이다.

2) 디지털화의 기준

디지털화의 기준가운데 가장 합리적인 것은 이용자들이 어떤 자료를 얼마나 원하는지를 계량화하는 것이다. 계량화방법가운데 하나는 필요한 자료의 열람 횟수를 기준으로 필요한 것과 필요하지 않은 것으로 구분할 수 있다. 이러한 기준설정에 앞서 반드시 이루어져야 하는 것은 전자도서관 이용자의 계층과 수준이다. 예를 들면, 전자도서관(예상)이용자들의 연령이나 학력에 따라 그들이 원하는 정보의 종류는 많은 차이를 보일 수밖에 없기 때문이다. 본 연구에서는 디지털화 사업의 주체가 국가이기 때문에 전자도서관 이용자들은 성인남녀로 정하며, 학력은 인터넷과 같은 정보망상에서 전문(full text)데이터를 필요로 하는 수준으로 설정한다. 즉, 일련의 고등교육을 받은 사람으로 간주한다. 따라서 일반 이용자들이 필요로 하는 정보욕구의 기준을 측정하기 위해서는 인터넷이나 PC통신과 같은 정보네트워크를 통해서 주로 입수하는 정보의 행태를 분석한다.

(1) 매체별 이용자 분석

매체별 이용자 분석은 전자화된 장서가 어떤 매체를 통해 어떤 순위로 이용자에게 열람되는지에 따라 구분하는 것이다. 그러나 매체적 특성에 따라 이용자의 특성에 차이가 있을 수 있다. 예를 들면, 인터넷은 상업적 사이트도 있으나, 기본적으로 회원등록만 할 경우 소속기관이 보유하고 있는 자료에 대해 무료 열람을 전제로 하고 있다. 한편, PC통신은 회원등록에 따른 비용과 자료열람에 따른 비용을 청구하는 등 기본적으로 유료서비스를 목적으로 하고 있다. 따라서 전자의 경우는 저작권이 해결된 자료(저작자 자신의 자료 등)가 주로 업로드(upload)되며, 후자의 경우는 상업적 성격을 갖는 업자가 전문데이터를 업로드하고 있다. 즉, 각각의 이용자 계층이 비용적 측면이라는 제한점 때문에 이용수준에 차이가 있을 수 있다.

① 인터넷

현재 인터넷을 통해 입수할 수 있는 대부분의 자료 형태는 간단한 서지정보부터 사진, 원문, 소리 및 동영상자료까지 매우 다양한 수준의 정보들이다. 한편, 주제도 특별한 제한이 없기 때문에 기존의 도서관에서 다루지 않는 사상적으로 편중된 자료와 음란물 등 전통적인 도서관에서는 보관할 만한 가치가 없는 자료들도 상당수를 차지하고 있다. 따라서 이용자의 이용형태를 측정하고 관심주제를 접속순위만으로 결정하는 것은 많은 문제가 있을 수 있다. 단, 이 데이터는 구축될 전자도서관의 이용률을 극대화하기 위해서는 어느 주제를 중점적으로 디지털화할지에 대한 자료로 활용될 수 있다. 예는 사진을 주로 제공하는 사이트와 혹은 원문전체이거나 극히 일부분으로서 전통적인 목록수준의 서지사항과는 차이가 있다. 따라서 인터넷상에서 이용자가 접속하는 사용자 빈도에는 많은 차이가 있다.

i) 인터넷접속횟수

다음은 1996년도 미국에서 한 달간 인터넷을 통해 이용자들이 선호하는 사이트에 대한 자료이다. 이 데이터를 분석하면 전문(full text)데이터의 입수순위보다는 이용자의 관심영역을 파악할 수 있다. 그 당시 인터넷이용자가 가장 많이 접속하는 주제별 영역은 컴퓨터관련분야와 인터넷관련 분야, 무역 및 마케팅, 뉴스와 스포츠, 정부정책, 상식·백과·참고문헌류, 교육 및 구인/구직, 잡지류(성인용 잡지 포함), 예술 및 음악, 웹검색엔진 및 도구, 레저 분야이다.

〈표 3〉 인터넷 베스트 사이트(미국내 접속수기준)

주 제	사이트명과 수
컴퓨터관련분야	Apple사의 8개
인터넷관련 사이트	Java사의 10개
무역 및 마케팅	Amazon.com Books사의 9개
뉴스와 스포츠	CNN Interactive사의 8개
정부정책	Central Intelligence Agency사의 6개
상식·백과·참고문헌류	Britannica On-line사의 9개

주 제	사이트명과 수
교육 및 구인/구직	America's Job Bank사의 8개
잡지류(성인용 잡지 포함)	The Biz사의 9개
예술 및 음악	MTV Online사의 8개
레저	The Electric Postcard사의 9개
웹검색엔진 및 도구	Alta Vista사의 9개

이상의 자료는 미국의 경우이고 다음은 국내 검색엔진가운데 하나인 네이버(<http://www.naver.com>)에서 제공하고 있는 디렉토리 서비스를 통해 이용자들이 해당 디렉토리에 접속한 상위 20위까지의 결과이다.

〈표 4〉 국내 검색엔진(네이버) 사용자 선호사이트 접속 현황

순위	주 제
1	뉴스/신문
2	연예
3	연예/연예인
4	연예/연예인/탈렌트
5	뉴스
6	경영/고용
7	연예/사람
8	컴퓨터
9	컴퓨터/인터넷
10	연예/유머 및 우스개
11	연예/음악/작곡가/한국가요
12	경영/재정 및 투자
13	연예/영화
14	경영/회사/미디어/신문/중앙일보
15	경영/고용/직장
16	경영/회사/미디어/신문/조선일보
17	경영
18	컴퓨터/멀티미디어/비디오/MPEG/MP3/
19	연예/음악
20	뉴스/잡지

<표 4>에서 알 수 있듯이 국내 인터넷 이용자들이 선호하는 주제분야는 연예(오락)분야와 신문·잡지, 경영, 컴퓨터분야로 구분될 수 있다. 이 분야 가운데 연예분야는 전통적으로 장서구성비율이 낮기 때문에 디지털화 대상에서 제외한다면 소장자료의 소급디지털입력은 신문과 경영학, 컴퓨터 분야자료가 상대적으로 우선 순위가 될 수 있을 것이다.

한편, 인터넷과 같은 통신망을 통해 상업적으로 전문정보제공서비스를 제공하는 디지털 자료는 대체적으로 잡지류와 소설류에 집중하는 현상을 보이고 있다. 예를 들면, 미국의 (주)전자도서관은 잡지와 비소설류, 소설을 디지털화하여 인터넷상에서 유료서비스를 하고 있으며, 그 현황은 다음과 같다.¹⁶⁾

- 잡지 (Magazines) : Journal of Consumer Affairs 외 598권
- 비소설 (Non-fiction Books) : Hong Kong Business 외 144권
- 소설 (Fiction Books) : Monarch Notes 외 3권

국내의 상업사이트를 이용하여 제공되는 정보는 주로 PC통신망을 통해 이루어진다. 이 자료의 경우는 정보제공회사와 저자간에 별도의 계약을 통해 디지털화 자료에 대한 저작권을 해결하여 열람과 다운로드가 매우 자유롭다. 현재 PC통신망 위에서 열람할 수 있는 전자책(digital book)의 유형으로는 단행본이 대부분이며 형태는 일반소설, 시/수필, 컴퓨터/교육/실용, 추리/오락 등이다. PC통신(천리안)을 통해 실제 서비스를 실시하는 기관은 4개 기관으로서 전체적으로 약 2,400여종의 자료가 디지털 형태로 제공하고 있다.¹⁷⁾

ii) 웹검색엔진의 분류체계

국내에서 사용되고 있는 웹검색엔진 가운데 일부는 분류체계를 활용하고 있다. 대표적인 것으로는 정보탐정과 심마니, 야후, 네이버가 있으며 각 분류체계는 엔진별 특성을 보이고 있다. 이 메뉴체계는 웹검색엔진 이용자를 위해 고안된 것이기 때문에 웹분류항에 배정된 웹문서의 수로서도 정보의 생산성과 이용자의 요

16) <http://www.eibrary.com/>

17) 남영준, **국어정보처리기반구축사업 저작권 해결을 위한 연구**, 보고서, 문화체육부, 1997, pp. 118-126.

구도를 유추할 수 있다. 남영준은 이 네 개의 검색엔진에서 제공하는 분류항목 가운데 류항목을 비교하여 '경제/사업'항에 배정된 웹정보의 양이 가장 많은 것과 '컴퓨터/인터넷', '교육', '사회/문화', '예술(혹은 연예)'에 관련된 정보가 타 분야에 비해 상대적으로 많았다는 것을 실험으로서 입증하였다.¹⁸⁾

② 상업용 데이터베이스 및 CD-ROM

CD-ROM은 다른 저장매체에 비해 저장용량당 단가가 낮기 때문에 많은 양의 자료를 축적하여 배포하기에 적합한 특성을 갖고 있다. 따라서 상업적으로 가치가 있는 데이터를 대상으로 CD-ROM을 개발하고 있다. 한편, DIALOG나 BRS 등과 같은 대규모 데이터베이스 판매업자들은 1980년대부터 온라인으로 자신의 호스트 컴퓨터에 일련의 데이터를 보관하여 유료로 전세계 정보요구자들에게 판매하고 있다. 또한, 적정한 데이터를 수집하여 이를 제공받으려는 많은 이용자들을 위해 CD-ROM의 형태로 데이터의 판매를 시작하였다. 이러한 상업용 데이터베이스의 개발원칙은 이익창출을 우선으로 하기 때문에 가장 효과가 높은 자료위주로 데이터베이스를 개발하거나 CD-ROM을 개발하고 있다. 따라서 데이터베이스와 CD-ROM의 개발행태를 분석하면 전자도서관의 디지털화의 우선 순위를 예측할 수 있을 것이다.

i) CD-ROM

<표 5>는 서울대학교 소장 CD-ROM리스트로서 현재 인터넷을 통해 열람이 가능하며 전문데이터베이스 뿐만 아니라 서지데이터베이스도 포함되어 있다. 대학 도서관은 제한된 예산으로 CD-ROM을 구입하기 때문에 이용자의 활용도가 높으면서, 해당 도서관의 예산내에서 구입이 허락하는 범위내에서 구입하는 것이 관례이다. 이러한 형태는 국가전산화 작업의 우선 순위를 결정하는 것과 매우 유사하여 학술정보위주로 디지털화가 이루어진다면 하나의 중요한 데이터로 활용될 수 있다. 즉, 이용자의 정보선호도를 유추할 수 있다. 서울대학교의 도서관의 경우는 소장하고 있는 CD-ROM의 주제는 사회과학 분야의 데이터가 가장 많

18) 남영준. "웹문서분류체계의 분석 및 새로운 설계", 한국헌정정보학회지, 제32권 제3호, 1998, pp.223-224.

으며 기술과학(의학포함)과 자연과학순으로 데이터를 제공하고 있다.

〈표 5〉 서울대학교 소장 CD-ROM 목록 : DOS와 WINDOWS 버전은 하나로 계수함.

주 제	종류 및 수	비고
총 류	Dissertation Abstracts Ondisc 포함 6개	
인 문 학	Art & Humanities Citation Index 포함 4개	
사회과학	Journal Citation Reports/SSCI 포함 12개	
자연과학	Biological Abstract 포함 7개	
기술과학	INSPEC 포함 6개	
의료분야	Medline 포함 2개	

ii) 데이터베이스

서지데이터베이스는 CD-ROM과 저장과 이용방법상으로는 큰 차이가 없으나, 전문데이터베이스(full text database)의 경우는 CD-ROM으로 제공되기 보다는 온라인데이터베이스의 형태로 제공된다. 전문데이터베이스의 형태로는 다음과 같이 구분할 수 있다.¹⁹⁾

- 법령류(Statutes, court decisions, and other primary legal documents) : 주 이용자는 법률전문가를 포함하여 일반 국민까지 확대될 수 있다.

- 기타 정부보고서 및 공식서류(Other government documents, patents, regulations, and other official publications) : 특허관련 정보등

- 보도자료 및 미간행정보(News releases and other unpublished information) : 주된 이용자는 주제전문가와 사업가, 저널리스트 등이 포함된다.

- 신문(Newspapers) : 주된 이용자는 최근 정보에 관심이 있는 일반 국민이 포함된다.

- 뉴스레터(Newsletters) : 주된 이용자는 주제전문가 혹은 정보분석가 등이 포

19) Tenopir, Carol, Ro, Jungsoon. FULL TEXT DATABASES. GREENWOOD PRESS. 1990. pp.4-12.

함된다.

- 참고자료(Reference books): 주된 이용자는 초등학생을 비롯한 연구가에 이르기까지 다양하다.

- 학술 및 기술자료(Scholarly or technical journals) : 주된 이용자는 주제전문가나 연구자들이다. 그러나 해당 학문을 연구하는 학생들이나 일반 국민도 포함된다.

- 통속 및 일반 잡지(Nonspecialist or general magazines) : 주된 이용자는 일반 독자층이다.

실제적으로 상업용 데이터베이스 벤더들이 온라인으로 전문형태로 제공하는 주요 데이터베이스들을 살펴보면 다음과 같다. <표 6>은 BRS社에서 제공하는 것이며, <표 7>은 DIALOG社에서 제공하는 것이다.²⁰⁾ 이 두개의 데이터베이스의 특징은 경영학 분야와 의학분야의 자료가 주를 이루며, 상대적으로 인문학 분야의 자료는 극히 적은 수준이다. 이상과 같은 이유는 전문데이터베이스의 경우는 대부분 유료로 제공되기 때문에 이용자들의 수입이 다른 직종에 비해 높은 분야위주로 개발되었기 때문으로 판단된다.

<표 6> BRS에서 제공하는 주요 전문데이터베이스 파일명

데이터베이스 파일명	주제분야
Academic American Encyclopedia Database	백과사전
Arts and Humanities Citation Index	인용색인지 인문학관련 색인지
ADS Journal Online	화학학회지
Combined Health Information Database	의학 및 건강관련 학술지
Comprehensive Core Medical Library	의학관련 학술지
Consumer drug information database	미국약품정보
Directory of graduate research	미국내 화학관련 대학원정보
Disclosure online database	미국내 회사예산연감
Disclosure/spectrum ownership database	미국내 회사예산연감

20) 남영준. 전문데이터베이스의 검색효율성 분석. 중앙대학교 대학원, 미간행석사학위논문. 1988. p. 15.

데이터베이스 파일명	주제분야
Drug information full text	약품정보
Havard Business Review	경영 및 무역학회지
Health lawyer	의학 및 건강관련 법률정보
Hazardline	독극물관련 학술지 정보
IRCS Medical Science Database	약학관련 학회지 정보
Kirk-Othmer Encyclopedia of Chemical Technology	화학관련 보고서
Magazine ASAP III	미국내 주요기업정보
Mental Measurement Yearbook	표준화 검사자료집
Peterson's national college book	미대학교 프로파일
Online Microcomputer Software Guide and Directory	컴퓨터소프트웨어 정보
Superindex	과학, 약학, 공학, 기술관련 색인지
Trade and Industry ASAP III	주요기업정보
IHS Vendor Information	상품정보

〈표 7〉 Dialog에서 제공하는 주요 전문데이터베이스 파일명

데이터베이스 파일명	주제분야
Havard Business Review	경영학 분야 학술지
Washington Post Electronic Edition	일간신문
Commerce Business Daily	경제관련 신문
McGraw-Hill Backgrounder	경영분야단행본일부
American Banker	금융분야 정간물
Business Dateline	경제관련 데이터
Magazine Asap	광고관련 정간물
Trade & Industry Asap	주요기업정보지

(2) 도서관 소장자료 분석

도서관 소장자료의 분석은 전자장서구축에 커다란 도움이 되고 있다. 즉, 전자장서도 자체적으로 하나의 도서관이기 때문에 일반 도서관장서의 기준을 전자도서관에 그대로 적용하는 것도 나름대로 유의하다고 할 수 있다. 다만, 이 기준은 주제별 장서기준은 될 수 있으나 자료의 형태별 기준으로 활용될 수는 없을 것

이다. 왜냐하면, 모든 도서관이 소장장서가운데 신문류 혹은 보고서, 단행본, 사전류의 분포를 확인하는 것이 용이하지 않기 때문이다.

따라서 도서관 소장자료를 조사하는 것은 디지털화의 주제별 우선 순위를 확인할 수 있는 자료로만 활용될 수 있다.

아래에 제시된 표는 학교도서관 설치기준령에 따라 우리 나라 학교 도서관의 설치 기준이 되는 것을 요약한 것이다. <표 8>을 분석하면 초중고등학교에서도 사회과학을 비롯하여 순수과학, 기술과학류는 일정 비율을 차지하고 있으며 학년이 올라갈수록 그 비율을 늘리도록 권장하고 있음을 확인할 수 있다. 한편, 역사와 문학류의 비율은 상대적으로 높게 책정되어 있으며, 학년이 올라갈수록 점차 감소하도록 기준이 설정되어 있다.

대학의 경우는 학교도서관 설치기준령과는 무관한 수치로서 현재 서울대학교 중앙도서관의 장서분포율을 조사한 것이다. 총계가 68.1%로서 전체 합이 100%에 미달되는 이유는 이 수치안에는 재래식 분류에 따라 통계수치에 포함되지 않은 자료와 정리중 자료, 규장각 도서가 포함되어 있기 때문이다.²¹⁾ 이 표에 따르면 초중고등학교에서 높은 장서비율을 보이고 있는 문학 및 역사분야는 급격히 줄어들고 실제 장서의 대부분을 차지하는 것은 사회과학과 순수과학, 기술과학분야임을 확인할 수 있다. 이러한 기준은 전자도서관 이용자가 대학이상의 학력소지자일 경우에 세 개의 주제분야위주로 구축되는 것이 바람직하다는 것을 보여주고 있다.

<표 8> 교육부 학교도서관 설치 기준 및 서울대학교 장서구정비

학교수준	총류	철학	종교	사회과학	순수과학	기술과학	예술	어학	문학	역사	그림책	계	비고
초등	4	2	2	8	13	7	5	3	35	16	15	100%	
중등	5	5	3	10	15	8	6	4	27	17	-	100%	
고등	6	6	4	10	15	9	7	8	25	10	-	100%	
대학	5.4	2.1	0.9	18.2	10.3	17.8	2.5	1.6	5.0	4.3	-	68.1%	

21) 서울대학교 중앙도서관. 서울대학교 중앙도서관 전자도서관의 설계에 관한 연구. 보고서. 동도서관. 1996. p.106.

또한, <표 9>에 보이는 바와 같이 실제 서울대학교에서 조사한 주제별 이용현황을 보면 주제별 장서비율은 사회과학, 응용과학, 문학, 순수과학, 역사 분야 순서대로 많았고, 대출비율도 사회과학, 문학, 순수과학순으로 많았다. 특징적인 것으로는 사회과학 분야는 장서비율은 가장 높고 대출량도 가장 많지만 이용요인(user factor)은 0.73으로 상대적으로 낮은 수치를 보이고 있다. 이러한 이유는 장서의 노화로 인하여 장서비율에 출판년도를 고려하여 이에 따른 대출비율을 조정하였기 때문으로 판단된다.

<표 9> 서울대학교 소장자료의 주제별 이용현황

구분	총류	철학	종교	사회과학	어학	순수과학	응용과학	예술	문학	역사	기타
장서비율	7.3	5.28	2.14	26.09	2.81	12.65	14.21	4.28	13.28	10.64	1.28
대출비율	10.16	7.89	2.23	19.02	2.16	14.36	10.7	6.67	18.74	7.68	0.39
이용요인	1.39	1.49	1.04	0.73	0.77	1.13	0.75	1.56	1.41	0.72	0.3

3) 디지털화 기준설계

전자도서관 구축에 있어 디지털화의 기준설계는 제한된 인원과 제한된 예산, 정해진 시간내에서 가장 효율적인 장서구성을 고려해야 한다. 고려항목으로는 크게 전자도서관 이용총과 전자도서관 장서의 주제별 구성, 전자도서관 장서의 형태별 구성이라는 측면들이 고려되어야 한다. 그 가운데에서 전자도서관 이용계층은 우리나라 전국민을 대상으로 하기보다는 고등교육을 받은 사람으로서 최소한의 새로운 정보를 창출하는 계층으로 제한한다. 이러한 제한의 이유는 전국민 모든 계층을 대상으로 할 경우 모든 변수요인들을 고려할 수 없기 때문이다.

(1) 주제별 기준

전자도서관에 있어 우선 디지털화 순위는 앞절의 분석결과에 따라 다음과 같이 정의될 수 있다. 우선 이용도와 대학도서관의 장서구성, 데이터베이스 개발현

황 등을 종합할 경우에 주제별 분포는 다음과 같이 요약될 수 있다.

1차적으로 디지털화 대상이 되는 것으로는 일반 대학도서관의 장서구성비율과 이용률등을 고려할 경우에 사회과학분야의 장서가 최우선순위가 된다. 한편, 국내외에서 추진되고 있는 전자도서관의 디지털화분야는 대부분 응용과학분야의 장서이며, 상업용 전문데이터베이스의 경우는 화학분야를 비롯한 화학응용분야이며, 일반 및 경제 신문이 주를 이루고 있다. 이러한 특성을 고려하여 하나의 표를 구성하면 <표 10>과 같다.

<표 10> 주제별 디지털화 현황 비교표

분야	인터넷 이용률	CD 및 DB	NSF지원 6개 대학	인터넷 분류체계	도서관 장서구성	도서관 이용요인	비고
신 문	0			0			
화학관련		0					
응용과학		0	0	0	0	0	
사회과학		0		0	0	0	
법 률 학		0					
경영/경제학		0		0	0		
참고자료					0		
오 락	0			0			

<표 10>을 분석하면 우선 디지털화 분야는 응용과학분야를 비롯하여 사회과학, 경영/경제학 분야로 제한할 수 있다. 신문의 경우는 정보의 최신성이 가장 큰 생명이기 때문에 CD-ROM과 같이 정보입수에서 구축 및 배포에 일정시간이 소요되는 매체형태에는 적합하지 않을 수 있는 분야이다.

즉, 사회과학 장서의 경우는 타 분야에 비해 장서의 양은 많지만 문헌자체적인 수명이 타 주제분야에 비해 매우 길기 때문에 전자화이후 갱신하는데 추가로 소요되는 비용이 가장 적게 소요된다는 특성을 갖고 있다. 반면에 과학기술과학 문헌의 경우는 학문의 발전속도가 매우 빠르며 이전 연구가 현재와 미래의 연구

에 커다란 연계성을 갖지 못하기 때문에 기본적인 이론서나 혹은 기술서만을 대상으로 전자화할 필요가 있다. 왜냐하면, 특정 자료의 경우 새로운 이론과 기술의 발달에 따라 자료로서의 생명이 상대적으로 짧기 때문이다. 예를 들면, 윈도 98에 대한 자료가 출간되는 시기에 과거 DOS 3.X버전에 관한 자료는 실질적으로 이용도가 극히 저조할 것이다. 단, 실시간적인 전자화 효과를 얻기 위해서는 이용빈도가 가장 높은 분야의 장서위주로 디지털화하는 것이 효과적일 수 있다. 문학관련 주제는 모든 대학과 모든 통신망에서 가장 이용도가 높은 주제이며, 또한, 전문을 전자화할 횟수도 가장 많은 분야에 속한다. 이 주제에 대한 전자화는 저작권적인 문제에 가장 민감한 부분이지만, 도서관의 기능가운데 오락적 기능보다는 학술·연구적인 기능이 강하기 때문에 문학류에 대한 전자화는 국가사업과는 차이가 있다. 즉, 문학에 대한 전자화는 상업사이트나 별도의 기관(저작권협회 혹은 도서관협회)에서 디지털화하는 것이 바람직하다고 판단된다.

역사관련분야는 문헌의 수명주기도 상대적으로 길기 때문에 전자화에 대한 유지보수비용이 상대적으로 적게 소요되는 특징을 갖는다. 역사관련분야의 경우는 학문 특성상 한자와 같은 특수문자나 도록(圖錄)과 같은 그림정보가 많이 사용되기 때문에 입력자의 입력단가가 상대적으로 사회과학분야와 같은 단순 입력형태 위주의 데이터입력보다 많이 소요되는 단점을 갖는다.

(2) 형태별 기준

국가디지털화사업에서 대상데이터를 형태별로 구분하는 것은 주제별로 구분하는 것과 차이가 있다. 이는 실질적으로 이용자의 관심분야보다는 디지털화 사업의 용이도와 대상자료의 정보적 가치에 따라 디지털화를 결정짓는 것이다. 대체적으로 문헌자료가운데 디지털화 대상이 되는 정보의 형태는 일차문헌과 이차문헌으로 크게 구분이 된다. 일차문헌에 속하는 것은 학술잡지, 기술보고서, 학위논문, 정부간행물, 특허 등이 있다. 한편, 전통적으로 도서관에서의 데이터베이스는 이차문헌위주로 구축하며, 이차문헌으로는 서지와 색인지, 초록지로 크게 구분된다. 전자도서관의 디지털화는 실제적으로 일차문헌을 대상으로 이루어지며 이를

좀 더 세분하면 다음과 같다.

- 보고서²²⁾
- 단행본
- 고서
- 학위논문
- 학술지 (대학논문집 포함)
- 연감
- 특허자료
- 법령 및 규칙

이 가운데 고서의 경우는 국립중앙도서관을 비롯한 많은 연구기관에서 이미지 형태를 포함한 각종 파일형태로 디지털화를 시도하고 있다. 이는 역사적 가치를 인정하는 원인도 있지만 고서의 경우는 대부분 저작권에 따른 제반 갈등을 회피할 수 있기 때문이다. 따라서 현재 국내에서 이루어지고 있는 고서의 디지털화는 이용자의 흥미나 관심도 혹은 홍보적 가치보다는 테스트베드적 성격이 우선하고 있다. 한편, 국립중앙도서관과 연구개발정보센터, 한국과학기술원등 대부분의 전자화를 시도하고 있는 기관은 디지털화 대상으로 크게 학위논문과 기술보고서, 학술지를 선정하고 있다. 이 가운데 학술지는 연구개발정보센터가 중점적으로 이미지 형태의 전문데이터베이스를 구축하고 있으며, 구축은 학술지를 생산하는 학회차원에서 외부전문업자와 연계하여 학술지 디지털화(사진파일 등)를 시도하고 있다. 학위논문의 경우는 서울대학교 도서관의 경우 1995년도 일반 대학원 졸업생이후부터 졸업생의 논문을 아스키파일로 제출할 것을 중용하고 제출된 학위논문은 웹상에서 다운을 허락하고 있다. 그러나 별도의 웹뷰어를 제공하기 않기 때문에 통신선상에서 해당 자료의 직접 열람은 불가능한 실정이고 자신의 개인용 컴퓨터로 해당 논문파일을 전송받을 수 있다. 그 밖에도 대학 차원에서 자신

22) 국가나 공공단체에서 생산되는 보고서 외에 각 기관 및 단체에서 생성되는 보고서

이 생성한 학위논문이나 보고서, 대학 논문집을 개인적 차원에서 인터넷에 업로드하고 있는 실정이다. 따라서 국내에서 인터넷을 통해 이루어지고 있는 실질적인 전문데이터베이스봉사는 주로 학위논문과 보고서 위주로 제공되고 있다.

4) 디지털화 우선순위의 제안

이상에서 조사한 자료를 근거로 디지털화의 우선 순위를 주제별·매체별 기준을 종합하면 국가 디지털화 사업은 사회과학과 기술과학분야의 자료들로서 각 대학 및 연구기관들이 갖고 있는 학위논문과 정부가 소유하고 있는 각종 보고서를 대상으로 디지털화를 시작하는 것이 바람직하다고 판단된다. 그 가운데에서 하나의 주제를 선정하면 사회과학 분야의 장서(보고서, 학위논문)를 대상으로 시작하는 것을 추천한다. 왜냐하면, 학위논문은 소속대학의 특성에 따라 저작권과 입력에 따른 경비가 동시에 해결될 수 있으며, 소급입력의 경우도 사회과학장서의 경우 처리해야할 한자나 특수문자, 외래어가 다른 분야에 비해 상대적으로 적기 때문이다. 또한, 자관에 자신의 대학출신의 논문을 보관한다는 것은 교육적 학술적 측면등 여러 가지면에서 학술적 가치가 있는 작업이기 때문에 디지털화 사업을 추진하는 대학에 소속된 도서관입장에서도 적극적인 협조가 있을 수 있기 때문이다. 기술보고서의 경우는 많은 연구자들이 자신의 연구가 중복연구가 이루어지고 있는지 또한 선행유사연구가 있는지를 알고 싶어하나 국가차원의 연구보고서가 통합적으로 관리되지 않기 때문에 연구자의 의도와는 무관하게 많은 분야에서 중복연구가 있는 것도 사실이다. 따라서 국가차원에서 중복연구가 이루어지고 있음에 따라 예산의 불필요한 투자가 이루어지고 있는 실정이다. 특히, 기업과 같은 영리단체에서 보유하고 있는 각종 보고서(특히, 기술보고서)의 경우는 일반 이용자들이 접근하기가 현실적으로 어려운 실정이다.

이러한 것을 구체화하면 다음과 같이 국가 디지털화 사업의 기준을 제시할 수 있다.

첫째, 대학도서관들이 소장하고 있는 자료 가운데 학위논문과 같이 해당 대학에서 생산되는 연구정보를 디지털화한다. 이러한 사업이 본격적으로 시작하기보

다는 모든 대학을 대상으로 동시에 디지털화 사업을 시작하기보다는 적정규모의 디지털화 대상자료를 보유한 대학을 선정하여 디지털화 사업에 대한 전반적인 흐름도를 파악하도록 한다. 즉, 디지털화에 따른 기술적 문제나 저작권적인 문제 등 디지털화과정에서 야기될 수 있는 모든 문제를 확인하고 대처방안을 얻을 수 있도록 하는 테스트 베드 성격의 디지털화 사업을 운영할 필요가 있다.

둘째, 정부가 소유하고 있는 각종 보고서를 디지털화한다. 현재 과학재단에서는 유사 사업을 추진하고 있으나 대부분 사진파일(스캐너 이용)형태로 구축되고 있으며 국내에서 생산되는 보고서 가운데 특히 일부에 해당한다. 또한, 궁극적으로 국가디지털화사업은 아스키형태(키인 혹은 스캐너후 OCR이용)로 구축하는 것이 바람직하다. 왜냐하면, 실제적으로 연구자들이 필요로 하는 것은 아스키형태의 파일을 필요로 하고 있기 때문이다.

셋째, 기타 통계자료와 수치정보로서 국가와 연구기관에서 생산하고 있는 각종 수치형 데이터베이스는 국가 공공DB사업 장려차원에서 이를 필요로 하는 집단과 연계하여(아웃소싱) 상업적 차원에서 이루어지는 것이 바람직하다.

II. 결 론

인터넷의 도래와 정보통신기술의 발달로 인해 국내에서도 전자도서관의 완성에 대해 많은 연구와 투자가 이루어지고 있다. 특히, 전자도서관의 구축은 고학력 실업자 대책 사업의 일환으로 국가 차원에서 추진되게 되었다. 이러한 투자는 제한된 예산내에서 이루어지기 때문에 투자대 효과를 극대화하기 위해 디지털화 대상을 엄선할 필요가 있다. 본 연구는 디지털화 대상을 선정할 수 있는 기준을 도출하기 위해 국내외에서 이루어지고 있는 디지털화사업을 조사하고 실제 이용자의 정보요구, 학교도서관의 기준, 도서관 장서의 현황 등을 조사하였다. 조사

결과에 따라 다음과 같은 디지털화사업의 기준을 선정할 수 있었다.

첫째, 외국의 경우 디지털화 사업은 주제별 기준보다는 역사적 자료나 보고서 등 자료의 형태별기준에 따라 이루어지고 있었다. 한편, 전자도서관 사업에서 원문의 디지털화사업보다는 구축시스템 구축과 인터페이스개발 등과 같은 하드웨어적인 사업이 주를 이루고 있다.

둘째, 국내의 경우 디지털화 사업은 국립중앙도서관과 같은 공공기관에서 이루어지며, 저작권의 제한을 면할 수 있는 주제나 형태를 디지털화 대상으로 선정하고 있었다. 문헌류와 같은 비학술적 자료는 거의 이루어지지 않고 있으며, 이는 상업기관을 통해 전문제공서비스가 이루어지고 있으나 디지털데이터의 수는 극히 미미한 수준이다.

셋째, 인터넷과 같은 통신망을 통해 이용자들이 관심을 보이고 있는 주제분야는 컴퓨터(인터넷)관련 자료와 뉴스, 정부정책, 예술분야의 자료였다.

넷째, 대학 도서관에서 소장한 CD-ROM과 같은 專門데이터베이스의 경우는 대부분 학술적인 자료였으며, 주제분야는 사회과학의 주제가 상대적으로 많았으나, 제공되는 정보의 수준은 서지 및 초록위주이다. 상업용 데이터베이스의 경우는 대부분 학술적 자료이며, 전문형태로 제공하는 것은 뉴스·시사정보와 화학, 약학, 경영학 분야의 자료로 요약될 수 있다.

다섯째, 대학도서관장서구성과 주제별 이용현황은 제한된 정보이지만 사회과학을 비롯한 순수과학, 응용과학, 문헌류가 장서의 많은 부분을 차지하였으며, 이용요인은 예술, 철학, 총류, 문학순이었다.

이러한 조사결과에 따라 현재 우리나라가 국가적으로 구축하는 자료의 주제나 형태는 이를 필요로 하는 이용자계층에 따라 차이가 발생할 수 있다. 이용자의 요구를 주체의 관점에서 대학도서관장서이상의 정보를 필요로 하는 수준으로 기준을 정할 경우, 주제분야로는 사회과학, 기술과학, 순수과학을 1차 디지털화 대상주제로 선정하는 것이 바람직할 것이다. 문헌류에 대한 높은 이용자 관심과 장서구성비는 저작권과 같은 문제, 최신성을 고려하여 상업적 관점에서 구축하는 것이 바람직하다. 한편, 이용자의 요구를 형태의 관점에서 기준을 정할 경우, 보

고서와 학술지, 학위논문은 1차 디지털화 대상주제로 선정하는 것이 바람직하다.

따라서 제한된 예산으로 정보화 근로사업을 수행한다면 구축할 1차 디지털화 대상은 사회과학 및 기술과학, 순수과학 부분이며, 형태는 보고서와 학술지, 학위논문으로 하는 것이 비용대 효과적 측면에서 가장 경쟁력을 가질 것으로 판단된다.

향후, 국가 전자도서관 사업은 모든 주제와 모든 비전자화 자료를 대상으로 디지털화가 이루어져야 할 것이다. 그러나 디지털화사업에 많은 비용과 인력이 소요되기 때문에 많은 기관에서 디지털화사업을 동시에 투입하여 완료한 뒤에 문제점을 발견하고 보완책을 제시하기보다는 실제적으로 하나의 실험대상기관을 선정하여 디지털화사업을 완성하고 실제 이용자들의 요구태를 분석할 필요가 있다. 이 자료는 이러한 시행착오를 최소화하고 국가전자도서관사업집행에 필요한 최소한의 지침서로 활용될 수 있다.

참 고 문 헌

- Bunker, G. R., Horgan, B. "Strategic Collaborative Models; Information Technology and the Library", *Advances in Library Administration and Organization*, vol.14. 1996.
- Nakano, S. "The Construction Plan of the Kansaikan of the National Diet Library and Digital Library", *In Proceedings of International Symposium on Digital Library*, 1995. Tsukuba. University of Library and Information Science.
- Tenopir, Carol, Ro, Jungsoon. *FULL TEXT DATABASES*. GREENWOOD PRESS. 1990.
- (주)쓰리소프트. **국가전자도서관 구축 기본계획**. 보고서. 국회도서관. 1997.
- 남영준. "웹문서분류체계의 분석 및 새로운 설계", **한국문헌정보학회지**, 제32권 제3호. 1998.

- 남영준. **국어정보처리 기반구축사업 저작권 해결을 위한 연구**. 보고서. 문화체육부. 1997.
- 남영준. **전문데이터베이스의 검색효율성 분석**. 중앙대학교 대학원, 미간행석사학위논문. 1988.
- 서울대학교 중앙도서관. **서울대학교 중앙도서관 전자도서관의 설계에 관한 연구**. 보고서. 동도서관. 1996.
- 임영선. **전문데이터베이스의 이미지 화일과 텍스트 화일의 검색 효율성 비교 연구**. 중앙대학교 대학원, 미간행석사학위논문. 1997.
- 임태삼. **도서선택의 이론**. 학문사. 1983. pp.79-82
- 정영미, 안현수. **전자도서관구축론**. 구미무역(주) 출판부. 1998.
- 첨단학술정보센터. **전국 대학 디지털도서관 구축방안에 관한 연구**. 보고서. 동센터. 1997.
- 최기선. **분야별 연구정보센터의 효율적 연계방안에 대한 연구**. 보고서. 한국과학재단. 1996.

A Study on Digital Priority for Digital Library Construction

Nam, Young-Joon *

〈Abstract〉

The principal aim of this study is to design the priority guideline for construction of digital collection. For this study, at first, type of the Korean and foreign digital library collection are investigated. After user's need on internet and list of commercial database are compared. Upon the result, this study suggests priority order of digital subject and media : the journal, the thesis and the report in social science, natural and technical science.

* Assistant Professor, Dept. of Library and Information Science, Jeonju University