

# 디지털 도서관 네트워크 구축에 정부가 주도적으로 참여

미국과 일본 등 해외 각국들은 정부가 주도적으로 나서 각 대학 디지털 라이브러리를 하나의 네트워크로 구축하는 디지털 도서관 망 설치 작업을 진행중이다. 이는 사용자가 어디서나 필요한 자료를 연결된 대학 도서관을 검색해 쉽게 획득할 수 있도록 하기 위해서이다. 미국과 일본의 사례를 살펴봤다. <편집자>

## 미국 - 인터넷과 연동된 도서관 망의 구축

미국에서 디지털 도서관은 인터넷의 발전과 궤를 같이 한다. 1969년 미국의 나사와 국방성이 아르파넷이라는 이름으로 최초의 인터넷망을 구축할 때는 핵전쟁을 대비한 군사적 정보 교류에 목적을 두었지만, 1988년 미국과학재단(NSF)이 미국내 5개대학을 묶는 NSF넷으로 인터넷을 계승 발전시킬 때 가장 중요한 목적은 학술적인 정보의 교류에 있었다.

대학내에서 공동으로 갖고 있는 자원을 교류하여 학술적인 연구의 발전에 도움을 주고 전자메일등을 통해 공간적 제약을 뛰어 넘은 연구자들간의 의견교환이 인터넷 발전의 주된 동력이 되었던 것이다.

이러한 맥락에서 미국에서의 디지털 도서관 구축이 도서관간의 네트워크 구축 계획과 병행해서 시도되고 있는 것은 당연한 귀결이라 할 수 있다. 네트워크를 이용한 미국내 디지털 도서관망의 구축은 따라서 나사와 NSF, 아르파넷이 주도가 된 '디지털 라이브러리기술 프로젝트(DLT)'가 주축을 이루고 있다.

이 프로젝트는 원래 나사의 우주 과학 관련 자료들을 디지털화해 각 도서관이나 대학 등에서 자유롭게 이용할 수 있도록 하기 위해 추진된 프로젝트로 90년대 초반부터 일부 대학도서관에서 시작된 개별적인 디지털 도서관 프로젝트의 표준화를 이끌어내

는 역할을 해주고 있다.

DLT 프로젝트의 지원에 의해 카네기멜론 대학, 미시간 주립대학, 스탠포드 대학, 산타바바라 대학, UC 버클리, 일리노이 대학 등에서 디지털 도서관 구축을 위한 표준 기술을 개발중이다.

미국에서는 최근들어 기존의 서적을 디지털화하는 계획과 병행해 디지털 도서관을 단순한 책의 디지털화라는 차원을 넘어서 멀티미디어 도서관을 구축하는 방향으로 바꾸고 있다. 대표적인 것이 현 클린턴 정부의 전폭적인 지원을 받고 있는 '인터넷 2' 프로젝트이다.

2천년대 초까지 미국내 주요 대학들을 100Mbps 급의 차세대 인터넷네트워크로 묶으려는 이 야심찬 계획에서 디지털 도서관은 모든 연구활동의 중심 센터로 자리잡게 된다.

100Mbps급의 인터넷 망에서는 교수에서 학생에 이르기까지 모든 연구활동 및 강의, 세미나 등을 실시간 화상회의 형태로 진행할 수 있으며, 각종 멀티미디어 데이터의 고속 전송이 가능해지게 된다.

디지털 도서관에서는 텍스트로 된 도서의 저장보다는 멀티미디어에 기반한 새로운 방식의 연구자료의 저장이 더욱 긴요해지게 된다. 따라서 멀티미디어 북의 탄생이 가능하게 된다는 것이다.

이같은 대규모 디지털 도서관 구축 프로젝트와 달

리 미국의 사례에서 주목할 만한 것은 인적, 물적 지원에 의한 무료 디지털 도서관 구축이다. 대표적인 것이 1971년부터 시작된 구텐베르크 프로젝트(<http://www.promo.net/pg/index.html>)다. 일리노이 베네딕트 대학의 마이클 하트교수에 의해 주창된 이 구텐베르크 프로젝트는 2001년 10월까지 1만여권의 책을 완전히 디지털화하는 것을 목표로 자원봉사에 의해 추진되는 디지털 도서관 프로젝트다.

1971년 미국 독립선언문을 텍스트화한 것을 시작으로 다양한 분야의 저작물들을 전자문서(etext)화하는 이 프로젝트는 97년 9월 현재 1,036권의 책이 디지털화됐다. 71년의 경우 단 1개의 텍스트만이 디지털화되었으나 올해들어서는 월 평균 30여권 이상이 디지털화되고 있어 목표량을 무사히 채울 것으로 전망되고 있다.

또한 이 프로젝트의 취지에 따라 전세계에 미러사이트 도서관들이 계속 추가되고 있는 상황이어서 무료 디지털 도서관의 대명사로 자리잡고 있다.

이미 별도의 프로그램을 통해 서지정보를 공개하고 있는 미국의 국회도서관도 아메리테크사로부터 200만달러의 자금지원을 받아 국립디지털 도서관(NDL) 프로그램을 진행하고 있는데 의회 정보 데이터베이스로서 의사록, 법안, 위원회 보고 등의 의회자료를 이용할 수 있는 'THOMAS'(thmas.loc.gov), 법률데이터베이스 네트워크인 GLIN, 미국의 역사를 소개하는 '아메리칸 메모리 도서관'이라는 이름의 무료 디지털 컬렉션을 구축했다.

미국의회도서관은 현재 의회로부터 매년 300만달러의 예산을 받을 것을 포함해 대기업의 투자등을 유도해 2천년까지 6천만달러의 기금을 조성, 1억 1천만권의 장서를 디지털화하는데 박차를 가할 계획이다.

### 일본 - 디지털 스캐닝 프로젝트

미국이 기존 도서의 경우는 서지정보 중심으로 하고 디지털 도서관의 구축을 멀티미디어화하려는데 비해 일본에서는 기존 도서를 이미지화해 이를 처리하려는 계획을 추진하고 있는 것이 뚜렷이 구분되는 점이다.

일본 국립도서관의 전자도서관 프로젝트

(<http://www.ndl.go.jp/ndlelp/index-e.html>)의 일환으로 추진되고 있는 파일럿 전자 도서관 프로젝트가 대표적인 사례. 이 프로젝트는 1993년 일본 정부에 의해 17억엔의 예산이 편성돼 일본 국립 다이어트 도서관과 새로이 구성된 정보처리진흥사업협회(IPA)에 의해 시작됐다.

1994년 8월부터 본격적으로 시작된 이 프로젝트의 목표는 엄선된 자료들로부터 1천만장의 디지털 이미지를 스캐닝해 데이터베이스화하는 것이었다. 약 반년간의 작업 끝에 목표량을 완성하고 시스템을 구성해 서비스를 시작한 것이 1995년 10월이었으니 짧은 기간안에 대단한 성과를 거둔 셈이다.

이 프로젝트는 다시 두 개의 부수적인 프로젝트와 함께 진행됐다. 첫 번째는 '전국 단일카탈로그 네트워크 프로젝트'로 전 일본에 산재해 있는 도서관들의 서지 정보를 통합하는 네트워크 시스템을 구축하는 것이었다.

이 서지 정보에는 MARC 번호와 ISBN, 제목, 출판사, 출판연도 등이 포함되었는데 이 정보들은 각 지역 도서관들에서 일본MARC 포맷으로 작성된 뒤 FTP를 통해 이 프로젝트의 중앙 컴퓨터로 전송되면 자동으로 처리되는 방식으로 구축됐다.

이 작업에 참여한 도서관은 일본 국립도서관을 포함해 총 27개에 이르렀으며 1996년 5월까지 총 720만권의 서지 정보가 모여 이중 공통부분을 제외한 330여만개의 서지 정보가 최종적으로 수록되게 되었다. 이 카탈로그 프로젝트는 현재 도서관간의 도서 대여 시스템을 구축하고 서지정보의 효율적인 인증방법과 인덱스 파일의 구축방법을 모색하는 방향으로 진행되고 있다.

두 번째 프로젝트는 '전자도서관 실험 프로젝트'다. 이것은 말 그대로 언제 어디서나 열람이 가능한 전자도서관을 구축하는 것으로 일본내 도서관들이 보유하고 있는 인쇄매체와 마이크로필름 등의 자료를 디지털화하는 것이다. 이 프로젝트는 다시 '자료의 디지털화', '데이터베이스의 구축', '정보검색' 등 세단계로 나뉘어 진행됐다.

'자료의 디지털화'는 각 분야별 자료를 스캐닝해 표준 이미지 파일로 저장하는 것. 논문, 잡지, 단행본, 마이크로필름 등 다양한 형태의 자료들을 스캐

닝해 표준 포맷의 이미지 파일로 변환시키는 작업이었다.

이 작업을 통해 약 1천만장의 이미지 파일이 만들어졌고 이 이미지는 3천장의 CD-R에 담겨져 각 500장을 담을 수 있는 6대의 CD-롬 체인저에 담겨졌다. 모노 이미지는 TIFF 파일 형식으로, 칼라 이미지는 JPEG 파일 형식으로 저장됐으며 희귀도서의 경우는 먼저 사진을 찍어 필름으로 만든다음 스캐닝하는 작업을 거치기도 했다.

'데이터베이스의 구축' 단계에는 세종류의 데이터베이스가 필요했다. 자료들이 모두 텍스트가 아닌 이미지 파일 형태로 되어 있었기 때문에 이미지 데이터와 서지정보를 연결시키는 데이터베이스를 구축해야 했다.

먼저 '정보검색 데이터베이스'는 각 분야별로 다른 정보검색 인터페이스가 필요해서 구축된 것으로 서지 정보, 목차, 전문 등의 종류에 따라 차별적인 데이터베이스를 구축했다. 모든 데이터를 이미지화했기 때문에 '비트맵 이미지 데이터베이스'가 요구됐는데 CD-R에 접근함으로써 빛어지는 지연현상을 방지하기 위해 빠른 브라우저를 위한 축소이미지를 제작해 이를 마그네틱 디스크에 담아 데이터베이스화했다.

또한 이미지 데이터를 서지정보와 연결하는 정보를 담은 '이미지 링크 데이터베이스'도 별도로 구축됐다. '정보검색' 단계에서는 사용자들이 넷스케이프 내비게이터등의 웹 브라우저를 이용해 이 데이터베이스에 접근할 수 있도록 지원하는 것이 목적이었다. 당시까지는 일부 브라우저들이 TIFF와 JPEG 이미지를 지원하지 않았기 때문에 모든 이미지 파일을 검색시에 GIF이미지로 변환시켜 출력하는 작업이 이뤄졌다.

검색 방법으로는 '서지정보데이터에서의 키워드 검색, 메뉴를 이용한 분야별 디렉토리 검색, 전문 검색등을 제공하도록 개발됐다.

그러나, 이 도서관 시스템은 개념상 전세계 어떠한 사용자에게도 접근이 가능하도록 개발되었으나 현재는 일본 국립도서관과 정보기반센터를 비롯 몇몇 연구소에서만 검색할 수 있도록 제한적으로 열람을 허용하고 있다.

파일럿 프로젝트는 앞으로 3년간 계속될 예정이며 이 기간동안 보다 차원높은 정보검색 기술과 데이터베이스 관리 기술이 적용될 계획이다.

### 전세계 단일 디지털 도서관 구축 계획들 - G7 프로젝트 UNESCO의 계획

위에서 든 예 외에 유럽에서도 각국의 국립도서관을 중심으로 각 대학 도서관들을 연계해 국가적 차원에서 프로젝트가 진행되고 있다.

또 하나 주목할 만한 것은 전세계적인 차원에서 디지털 도서관을 구축하려는 계획도 이미 시작되고 있다는 것이다.

대표적인 프로젝트가 G7 프로젝트와 유네스코에서 추진하고 있는 인류문화자료 보존 프로젝트(Virtual Memories of the World)다. G7 프로젝트는 1995년 2월 브뤼셀에서 열렸던 '정보사회에 관한 G7 각료회의'에서 전세계정보기반(GII)을 구축하는데 디지털 도서관이 중심적 역할을 담당할 것이라는데 G7 국가들이 합의해 추진되고 있는 전세계 디지털 도서관망 구축 계획이다.

이 프로젝트의 공식 명칭은 'Bibliotheca Universalis'로 일본과 프랑스가 공동 간사국이 되어 각국의 국립도서관을 중심으로 진전되고 있는 디지털 도서관 프로젝트간의 협력관계를 강화하고 일단 디지털화한 자료의 공유를 확산시키려는 것이 기본적인 목표이다.

이 프로젝트에서는 각 국가의 도서관이 진전시키고 있는 디지털 도서관을 인터넷으로 묶고 시스템의 구성과 데이터베이스, 전자문서데이터 압축기술 등 다양한 면에서 국제 표준을 따르게 함으로써 향후 G7국가를 비롯 전세계에서 이용가능하도록 할 방침이다.

유네스코의 프로젝트는 다양한 종류의 자료, 도서, 그림 등 인류가 남긴 가치있는 문화 자료를 보존하고 보호하기 위한 계획이다. 1993년부터 시작된 이 프로젝트는 지역에 관계없이 인류문화에 가치있는 정보는 모두 디지털화해 보존한다는 것을 원칙으로 삼아 진행되고 있다. DC