

# 정보자원의 디지털화에 관한 고찰

## A Study on digitization of information resource

최홍식, 이두영  
중앙대학교 대학원 문헌정보학과

Choi Hung-Sik, Lee Doo-young  
Dept. of Library & Information Science, Chung-Ang University

본 연구는 정보자원을 디지털화하는 과정에서 기본적으로 고려해야 할 과제를 고찰하였다. 구축에 관련된 과제를 대상자료 선정 기준, 구축준비, 보존, 접근 및 이용, 상호연계, 비용문제로 구분하여 디지털화 작업시 고려해야 할 사항을 고찰하였고, 향후 실제적인 디지털화 지침 작정을 목적으로 연구를 수행하고자 한다.

### 1. 서론

#### 1.1 연구의 필요성 및 목적

정보자원 디지털화의 기본 목적은 가능한 많은 이용자들이 쉽게 접근하여 이용할 수 있도록 자료를 재생산하는 것으로서 비디지털 및 아날로그 정보자원을 디지털 형태로 변환하는 것이다. 즉 인쇄 및 기타 형태의 정보를 디지털 형태로 재생산하는 것을 의미한다. 이것은 도서관의 기본목적인 보존과 이용의 측면에서 상당한 유용성을 꾀할 수 있다. 고서와 같은 희귀본과 귀중자료, 파손 및 훼손의 위험이 존재하는 자료 등의 영구보존 및 활용을 가능하게 하고, 필요한 정보를 광범위한 범위에서 보다 쉽게 접근하고 활용할 수 있도록 하여 이용효과를 증대할 수 있다. 이러한 정보자료를 디지털 형태로 변환하는 일은 사람과 기술, 기계라는 인적, 기술적 측면과 여기에 소요되는 비용과 관련된 예산측면을 총체적으로 고려하여 종합적인 계획과 면밀한 검토과정을 수행한 이후에 추진해야 한다.

현재 국가 정보화 사업의 일환으로 추진하고 있는 디지털화 사업은 저작권에 저촉되지 않는 고서와 학위논문, 연구보고서와 같은 자료를 대상으로 추진되고 있고, 그 밖에 많은 도서관에서 개별적으로 추진하고 있거나 계획하고 있다. 이러한 개별적인 디지털화 작업은 추진과정의 중복투자로 예산, 시간, 인적낭비를 초래하고, 통일된 기준과 표준 결여로 분산된 정보

자원의 상호연계 및 공동활용에 심각한 부정적 영향을 유발할 수 있다. 따라서 정보자원의 디지털화에 대한 모든 과정에서 통합된 기준과 실행지침이 마련되어 체계적인 수행이 필요하다.

본 연구의 목적은 정보자료의 디지털화에 대한 효과적인 구축방안을 모색하는데 있다. 즉 디지털화의 모든 과정에서 요구되는 사항과 수행기관의 임무와 목표, 이동목적 등을 조사, 분석하여 통합된 기준과 실행지침을 제공하는데 목적이 있다. 본고는 그 이전 단계로 디지털화의 각 과정에서 나타날 수 있는 과제를 살펴보았다.

#### 1.2 연구방법

본 연구의 방법은 국내외 정보자원의 디지털화와 관련된 선행연구를 조사, 분석하여 효율적인 자료로 활용한다. 현재의 상황과 지침개발 요소를 파악하기 위하여 실제적으로 디지털화를 하고 있는 기관을 대상으로 조사하고, 관련 전문가의 인터뷰 등을 통하여 정보자원디지털화에 관한 실제사항을 분석하고, 양자를 비교, 분석하여 효율적인 구축기준과 실행지침을 제시한다.

### 2. 이론적 배경

#### 2.1 디지털정보자료의 경제성

디지털 정보자료란 컴퓨터에 의해 축적되고

조직, 가공되어 네트워크를 통하여 송수신할 수 있는 정보자료로서 최초로 디지털 형태로 발행된 자료와 물리적 작용을 통하여 디지털화한 자료, 인터넷을 통하여 원거리 디지털 DB에 접근할 수 있는 자료 등이 포함된다.

디지털 정보자원의 경제성은 기존 인쇄자원에 비교하여 많은 잇점이 있다. 자원구축 및 축적, 보존, 이용 등에 관한 모든 비용문제는 현재 더 많이 소요되고 있지만 점점 더 저렴해지고 있다는 점에 주목해야 하고, 자원의 활용에 관한 유용성은 접근성, 편의성, 활용성, 신속성, 적시성, 안정성, 영구 보존성 등에서 많은 잇점을 지니고 있다.

## 2.2 디지털화 방법

### ① 스캐너 활용방식

사진 영상 또는 종이위의 도형정보를 읽어 들여 컴퓨터의 그래픽 정보로 바꾸는 입력장치인 스캐너를 활용하는 방식이다. 이는 디지털화에 소요되는 많은 인력 및 예산을 절감할 수 있어 일반적으로 가장 많이 사용하는 방법이다. 이방법의 가장 큰 장점은 원자료의 형태를 그대로 유지할 수 있고, 복잡한 수식이나 도형 등을 원문대로 표현할 수 있다는 점이다.

이 방식에는 단면 플랫폼 스캐너 방식과 양면을 동시에 스캐닝하는 양면 플랫폼 스캐너가 있다.

### ② 키인 방식(key-in)

이 방식은 디지털화 대상으로 선정된 데이터를 사람이 직접 키보드를 사용하여 입력하는 방식이다. 디지털화 방법 가운데 가장 원시적인 방안으로서 장점은 학술문헌에 나타난 특수기호 혹은 한자, 외국어 원문을 입력자가 그대로 입력할 수 있기 때문에 입력의 질이 가장 안정적이다.

### ③ 전자자료입수방식

전자자료입수방법은 디지털 자료구축에 가장 쉽게 접근할 수 있는 방법 가운데 하나이다. 이것은 동일한 인쇄물로 존재하는 것과 존재하지 않는 것이 있고, 무료로 이용할 수 있는 것과 별도의 요금을 지불하는 것도 있다. 현재 저널의 경우를 보면 인쇄본과 전자형태를 병행하는 경우가 빠르게 증가하고 있다. 이들은 목차만 제공하는 것에서부터 초록까지 제공하는 것, 일부 논문만 선택적으로 제공하는 것, 레터나 리뷰 등을 제외한 모든 논문을 제공하는 것, 인쇄물과 동일하게 제공하는 것과 논문에 대한 의견 및 관련된 소스에 대한 하이퍼링크 등을 추가한 것에 이르기까지 다양한 형태를 취하고 있다. 이들은 인쇄물 구독자에게 전자

버전을 무료로 이용할 수 있도록 하기도 하지만(어떤 경우에는 일정 시험기간동안만), 일부 출판사는 처음부터 추가비용을 부담하게 하기도 한다. 이러한 모든 전자형태의 저작물은 축적과 저작권에 관련해서 많은 논란이 있다. 따라서 여러 가지 상황을 종합적으로 고려하여 가장 효율적이고 경제적인 디지털화를 추진해야 한다.

## 3. 정보자료 디지털화의 단계별 과제

### 3.1 대상자료 선정기준문제

자료선정에 대한 기본은 정보 데이터의 영원한 가치에 있다. 이점은 일반적인 자료선정의 핵심내용으로서 디지털화 작업에서 대상자료 선정기준으로 적용할 수 있다. 이러한 선정에 대한 관점을 정의하는 데에는 많은 어려움이 존재한다. 즉 미래에 유용할 것이라는 것을 예측하는 방법, 그 데이터의 보존과 접근을 위한 시기를 예상하는 방법, 계속적인 갱신을 요구하는 자료의 완전성과 책임성을 보장하는 방법 등이 여기에 해당된다.

Janet Gertz는 디지털화를 위한 자료선정기준으로 1)그 아이탬과 장서가 충분한 가치가 있는지, 2)현 이용자들이 디지털화를 정당화하고 있는지, 3)디지털화를 할 수 있는 법적 권리가 있는지, 4)그 자료를 안전하고 성공적으로 디지털화를 할 수 있는지, 5)디지털화를 수행할 수 있는 인프라를 갖추고 있는지, 6)비용이 적당하고 지원이 가능한지를 기준으로 제시하고 있다.

캘리포니아 대학은 전통적인 자료선정과 보존의 고려사항을 기초로 하여 변환장서에 적용할 수 있는 기준과 디지털 비전에 적용할 수 있는 기준으로 대별하고 전자를 장서개발, 보존과 축적, 조직화와 자금으로 후자를 장서개발, 접근, 조직화와 자금으로 소구분하고, 다시 이들을 전체적으로 20개의 항목으로 세분하고 있다.

콜럼비아 대학은 디지털 이미지화를 위한 자료선정기준을 3가지로 구분하고 있다. 1)장서개발 측면에서 콜럼비아 대학 공동체와 구성원들의 이용을 확신하고 충분한 가치가 존재하는지, 이러한 요구를 만족시키는 동일한 디지털 형태의 산출물이 없고 유용성에 책임을 질 수 있는지, 2)핸들링과 이용 측면에서 이용자의 요구를 만족시킬 수 있는 상태로 캡춰하여 저장될 수 있고, 파손 가능성이 있는 자료를 통제하는데 무리가 없는지, 쉽게 이용할 수 없는 자료에 접근을 제공할 수 있고 원자료에 손

상됨이 없이 스캐닝되는지, 네트워크 환경하에서 접근이 가능하고 신속하게 전달될 수 있는지, 이용자가 위치한 장소에 전달을 안정하게 할 수 있는지 3) 부가가치 측면에서 새로운 탐색지원을 창출(서지 레코드 링크, 인덱스 개발, 탐색도구 개발등)하고 분산된 자원과 통합되어 강화된 정보자원이 제공됨으로서 학술적 이용을 고무시킬 수 있는지를 기준으로 제시하고 있다.

그 밖에 많은 기관에서 자체적으로 선정기준을 설정하고 있다. 이러한 여러 가지 기준을 종합적으로 고려하여 디지털화의 대상자료 선정기준과 관련된 통합지침이 제시되어야 한다.

### 3.2 구축준비문제

정보자원의 디지털화는 선정기준에 따라 대상자료가 선정되면 본격적인 디지털화를 위해 준비를 해야 한다. 즉 디지털화를 수행하기 이전에 가능한한 위험을 제거하고, 허용될 수 있는 수준에서 잔존하는 위험을 감소하는 방안을 적용하여 그 위험요소를 찾아야만 한다. 이를 간략하게 살펴보면 다음과 같다.

1) 상태평가이다. 이것은 물리적으로 취약한 자료를 구분하는 것으로 선정된 자료가 디지털화의 모든 과정을 견딜 수 있는지를 초기에 판단하는 일이다. 즉 물리적 상태가 취약한 자료, 제본 구조가 캡취하기 용이하지 않은 자료, 슬라이드나 사진, 화학 실험용지와 같이 빛과 열에 예민한 자료, 크기가 동일하지 않은 자료, 지도와 같이 사이즈가 큰 자료 등이 여기에 포함된다. 이러한 사항은 준비단계에서 기록으로 유지하여 참고자료로 활용되어야 한다.

2) 핸들링 절차이다. 이것을 이루기 위해서는 기본적인 핸들링 절차가 형성되어야 하고 관련자들은 필요한 교육을 받아야 하며, 특별한 요구사항은 기록으로 유지되어 관련자들이 이 요구사항을 이해하고 서로 의사소통하여 정확하게 인지해야 한다. 이를 구현하기 위해 정책과 절차가 공식화되어야 하며, 구현에 따르는 사항을 평가하기 위하여 절차를 먼저 평가해야 한다. 또한 일정한 수준에서 지속적으로 모니터링을 해야 한다.

3) 제본해체 문제이다. 일반적으로 스캐닝을 하기 위해 제본을 해체하는 일이 발생된다. 이 제본해체 문제와 관련된 핵심문제는 컷째, 방법문제로서 대상자료가 국가적 혹은 역사적으로 중요한 가치가 있는 제본인가, 그것을 원상복귀하기 위하여 신중하게 분해할 필요가 있는가, 둘째, 제단문제로서 대상자료가 제단을 허용하고 있는가, 셋째, 재건문제로서 단순하게

재제본할 것인가, 완전무결하게 원상복귀할 것인가, 단순하게 포장해 놓을 것인가, 넷째, 처분문제로서 디지털화한 자료를 버릴 것인가, 보존의 대용물로 유지할 것인가, 아니면 다른 장소에 보관할 것인가 하는 문제이다.

4) 스캐닝 문제이다. 현재 디지털화 구현방법 가운데 가장 일반적인 방법은 스캐너를 활용하는 방식이다. 이것을 작업과정의 문제와 시스템과 관련된 빛의 문제로 구분해 보면 다음과 같다.

McIntyre JE.는 스캐닝 작업을 향상시키기 위하여 제본도서, 낱장 및 2절 자료, 사이즈가 큰 자료, 설계도 및 지도, 사진, 사진의 원판, 판화, 유리 슬라이드 등 자료형태를 다양하게 구분하여 처리하고, 디지털과정을 물리적으로 견디지 못하거나 변질 및 훼손의 위험성을 내포하고 있는 자료들은 중간단계(마이크로 필름, 디지털 카메라등)를 이용할 것을 제시하고 있다. 또한 빛과 관련된 문제로 열과 방사능 에너지로 인한 위험을 언급하고 있다. 그는 열을 받게 됨으로서 기질(基質)과 안료(顔料)에 영향을 받을 수 있고, 방사능 에너지에 노출되어 정보 내용과 텍스트, 이미지에 손상을 줄 수 있다고 한다. 컬러 이미지로 캡취할 경우 약 3500 Kelvin이 필요하고 이 온도에 도달하기 위해 보통 할로겐 빛을 이용하고 있지만 이때 자료표면에 높은 온도의 열을 받기 때문에 특별한 자료인 경우 손상의 위험이 있다고 한다. 빛과 관련하여 시스템 선정시 고려사항은 빛의 형태, UV 방사, 온도, 조명도 수준, 빛에 노출되는 시간과 양등을 고려해야 한다.

한편, 이러한 디지털화를 외부에 용역을 주는 경우 업체선정기준으로서 관련경험, 유무, 특별한 요구의 인지 및 이해, 안전성, 특수 장비, 작업환경 등을 고려해야 한다. 또한 이러한 업체 선정에 대한 현실적인 문제로서 해당업자의 기술 능력을 평가하여 서류상의 숫자로 표기하기가 쉽지 않다는 문제와 저렴한 비용을 제시하는 업체에 우호적인 예산담당자를 설득하는 문제도 고려해야 한다.

### 3.3 보존문제

보존의 대상은 근본적으로 위험이 존재하거나 이미 손상된 자료들을 대상으로 원자료를 정확하게 재생산하여 이용할 수 있도록 하는 것이다. 전통적인 자료보존의 선정모델에서 보면 다음과 같이 4가지로 구분하고 있다. 1) 그 아이템이나 장서가 파손될 수 있거나 위험에 놓일 수 있는지 2) 보존을 정당화할 수 있는 충분한 영구적 가치가 있는지 3) 원자료를 다른

형태(마이크로화, CD 등)로 변환을 성공적으로 실행할 수 있는지의 여부와 불충분한 결여요소가 있는지 4) 보존의 비용이 그 아이템의 가치와 비교하여 적합한지로 구분하고 있다.

디지털정보자원의 보존은 기존 정보자원의 보존 문제와 함께 디지털 정보의 속성상 추가, 변경, 삭제 등과 관련하여 장기적인 보존과 접근, 이용 등과 관련된 사항을 종합적으로 고려하여 보존할 방안이 제시되어야 한다.

### 3.4 접근이용 및 상호연계문제

디지털정보의 접근 및 이용은 과거 수준 이상으로 강화될 것이다. 이를 위해서는 도서관 정보서비스 프로그램과 통합될 수 있는지, 기관에서 지원된 컴퓨터 플랫폼으로부터 접근할 수 있고 네트워크환경에서 접근할 수 있는지, 적당한 속도로 전달될 수 있는지, 원자료의 접근과 관련된 기술적 문제를 해결할 수 있고 다른 종류의 포맷을 통합할 수 있는지, 일반 이용자들이 쉽게 접근할 수 있는지, 문헌확인 및 기술적 캡처정보, 출처, 다중접근점, 색인, 전문검색, 문헌구조 등과 관련된 적합한 메타데이터를 지니고 있는지, 내용의 신뢰성을 지니는지 등이 고려되어야 한다. 또한 분산된 정보자원을 효율적으로 상호연계하여 공동활용할 수 있는 방안을 모색해야 한다.

### 3.5 비용문제

정보자료의 디지털화는 디지털도서관을 구축하는데 있어서 필연적인 일이지만 쉽지 않은 일이다. 왜냐하면 정보량의 증가와 예산의 절감이라는 현실 속에서 기존의 인쇄자료를 유지해야 하고, 인터넷과 전자매체를 통하여 이용할 수 있는 전자자료도 관리해야 하는 비용이 적지 않은 상황에서 추가적으로 많은 비용이 예상되는 사업을 수행하기란 결코 쉬운 일이 아니기 때문이다. 그러나 이 사업은 디지털 도서관 구축에 있어서 가장 기본적인 작업으로 구축과 실제적으로 관련된 여러 핵심문제 가운데 중요영역이다.

디지털화계획이 가능하다고 판단되면 디지털화에 소요되는 비용을 다른 옵션과 비교해야 하고 노력의 중복을 피해야 하며, 재정이 유용한지를 결정해야 한다. 즉 이따지나 텍스트로 저장하는 비용, 자료를 준비하는 비용, 통제활동 등에 소요되는 모든 비용을 고려해야 한다. 그러나 디지털화에 소요되는 비용을 정확하게 산정하는 것은 대단히 어렵다. 디지털화의 초기단계에서(1997년) LC는 300페이지 분량의 도서에 \$500의 비용이 소요될거라고 평가하였고, 같은 해에 Cornell은 동일한 분량의 도서에 약

\$60의 비용을 예상하였다. 이렇게 동일한 연도에 똑같은 분량의 자료를 디지털화하는데 소요되는 비용을 산정하였음에도 커다란 차이를 나타내고 있는 것을 보면 비용산정은 결코 쉽지 않다는 것을 알 수 있다. 즉 비용산정에는 너무 많은 변수를 포함하고 있다고 볼 수 있다.

일반적으로 새로운 사업을 추진하는 단계에서 비용문제는 철저한 조사와 분석이 필수적이다. 특히 기존의 일들을 유지, 관리하면서 새로운 사업을 추가적으로 착수하는데에 소요되는 비용문제는 더욱 그렇다. 도서관에서 기존자료를 디지털화하는 일은 분명 새로운 사업이다. 따라서 수행기관의 목적과 디지털화에 대한 비용과 효과를 분명하게 조사, 분석하여 이에 따른 사업수행여부를 비롯한 비용관련 문제를 구체적으로 고려해야 한다.

## 4. 결론 및 향후과제

컴퓨터 통신망과 인터넷의 발전으로 향후 정보이용은 보다 광범위한 자원을 대상으로 적합정보를 적시에 신속하게 이용할 수 있도록 요구될 것이다. 정보자원의 디지털화는 현재 이런 목적을 달성할 수 있는 효율적인 방안으로 인식되어 많은 정보관련기관에 중요관심 영역에 포함되어 있거나 추진, 구축하고 있다. 이러한 디지털화 작업은 디지털도서관 구축 및 운영과 관련하여 종합적이고 체계적인 관리가 필요하다. 따라서 구축작업이 본격적으로 실행되기 이전에 각 작업단계를 세분화하고 여기에 따른 방향과 지침을 상세하게 설계하여야 하고, 이를 토대로 구축방안, 준비, 보존, 접근 및 이용, 상호연계 등과 관련된 활동방안을 종합적으로 탐구하여 실제적인 구축방안을 모색해야 할 것이다.

### <참고문헌>

- UKOLN(1996), Long term preservation of electronic materials, A JISC/BRITISH LIBRARY Workshop as part of the Electronic Libraries programme.
- <http://www.rlg.org/preserv/joint/confpapers.html>
- <http://homes.ukoln.ac.uk/~lismd/preservation.html>
- [http://www.library.ualberta.ca/library\\_html/libraries/law/digit1.html](http://www.library.ualberta.ca/library_html/libraries/law/digit1.html)
- <http://www.ukoln.ac.uk/services/papers/follett/lesk/paper.html>