

디지털 멀티미디어 도서관 (The Digital Multimedia Library)

차재혁(첨단학술정보센터)

디지털도서관은 기존 도서관과 달리 시간과 공간에 구애받지 않고 원하는 정보를 보다 효율적으로 획득할 수 있도록 정보 입력/저장/출력, 정보 저작권 보호, 정보 전달 등의 핵심 기능을 제공한다. 그런데 실세계의 정보는 멀티미디어 정보이므로 디지털도서관이 이를 지원하기 위해서는 각 핵심기능이 확장되어야 하므로, 현존하는 디지털 멀티미디어 도서관은 각 기능의 멀티미디어 지원 수준에 따라 다양하게 나눌 수 있다. 구축 사례를 중심으로 디지털 멀티미디어 도서관의 현황을 살펴보고 이상적인 디지털 멀티미디어 도서관이 갖추어야 할 필수 기능을 정의하고 이를 토대로 나아갈 방향을 제시한다.

1. 서론

정보 기반구조의 급속한 발전은 대규모 분산 정보자원에 대한 접근을 가능하게 하고 있다. 이러한 정보 기반구조의 대표적인 응용인 디지털 도서관은 보다 효율적인 정보 공유체제를 구축하기 위해서 필수적인 구성요소가 되므로 최근 이에 대한 연구가 활발히 진행되고 있으며 많은 디지털 도서관이 구축되고 있다. 이러한 디지털도서관은 실세계에 존재하는 모든 정보의 디지털화된 형태를 관리하고 이를 컴퓨터 네트워크를 포함하는 정보 기반구조 상에 유통·공유하도록 하는 체제이다. 따라서 이러한 개념은 실세계 정보를 담당하는 기존 도서관이나 박물관, 정보센터 등에서 지향하는 개념을 통합·발전하였으며, 구체적으로 디지털도서관에 대한 확실한 하나의 모델이 없는 실정이다.

이러한 디지털도서관을 기존 도서관 시스템의 관점으로 한정시키더라도 전자도서관(Electronic Library), 벽없는 도서관(Library without Walls), 가상 도서관(Virtual Library) 등과 같이 다양한 용어와 의미를 가지고 있다. 이들의 공통적인 특징을 살펴보면 디지털도서관은 크게 하드웨어, 소프트웨어, 정보의 디지털화된 형태인 컬렉션(Collection), 그리고 인력으로 구성되며, 정보 입력(저작)/저장/접근 기능, 정보 검색 기능, 저작권 보호 기능, 정보 프리젠테이션 및 배달 기능 등을 핵심 기능으로 제공하고 있다. 그런데 디지털도서관이 실세계 정보를 담당하기 위해서는 기존 도서관 시스템이 다루어 왔던 문자나 숫자로 구성된 목록 데이터 뿐만 아니라 멀티미디어 데이터를 다루어야 하므로 제공하는 기능마다 멀티미디어를 지원하기 위한 확장이 필요하다. 본 논문에서는 멀티미디어를 지원하기 위해 확장해야 할 디지털도서관의 기능과 확장방안에 대해 살펴보고, 컬렉션 관점에서 멀티미디어를 지원하는 현존 디지털 멀티미디어 도서관들을 조사·분석하였다.

본 논문의 구성은 다음과 같다. 2장에서는 디지털 도서관의 기능 각각에 대해 멀티미디어 정보를 다루도록 확장한 사례를 중심으로 확장 방안을 살펴본다. 3장에서는 멀티미디어 정보를 서비스하는 실제 디지털 도서관을 살펴보고 이를 디지털 도서관의 기능에 따라 평가한다. 4장에서는 결론과 현존하는 디지털 멀티미디어 도서관이 제공하지 못하나 이상적인 디지털 멀티미디어 도서관으로서 제공해야 할 기능

을 설명한다.

2. 디지털 멀티미디어 도서관의 기능

실제적인 이미지 또는 오디오 디지털도서관을 구축하려면 이러한 미디어 데이터의 입력, 압축, 저장, 접근, 전송 및 프리젠테이션 기술 외에 내용을 기반으로한 검색 및 브라우징, 데이터 표현 및 처리 기술 등이 해결되어야 한다 (Aigrain, 1995).

가. 디지털 멀티미디어 도서관의 기반 구조

다양한 미디어 데이터 타입이 존재하고 레코드, 문서, MARC 등 다양한 종류의 정보가 존재한다. 그리고 데이터베이스는 한 곳에 집중하기 어려우므로 여러 데이터베이스가 분산되어 있는 경우가 많다. 이러한 정보를 일관성있게 다루기 위해서는 분산 디지털 객체 개념이 도입되어야 하는데, Xerox PARC 연구소의 디지털도서관 프로젝트와 스탠포드 대학의 디지털도서관 프로젝트 등에서 채택하고 있다 (Williams Y. Arms, 1997).

예를 들어 Xerox PARC 연구소의 디지털도서관 프로젝트에서는 문서 서비스를 지원하기 위한 분산 환경, 문서관리 시스템 인터페이스 등의 미들웨어로서 언어와 시스템 상호운영을 지원하는 CORBA 확장판인 ILU(Inter-Language Unification System), 저장소에 대한 인터페이스 표준으로 OpenDMA(Document Management Architecture) 기반 인터페이스, 컬렉션에 대한 다양한 서비스를 통합하는 기술 등을 개발하고 있다.

그리고 스탠포드대학 디지털도서관 프로젝트에서는 네트워크 상에 제공되는 여러가지 정보 서비스를 가상적인 하나의 도서관으로서 이용할 수 있도록 독립된 서비스 간의 상호작용을 위해 InfoBus와 DLIP(Digital Library Interoperable Protocol)을 제안하고 이들 상에서 동작할 수 있는 사용자 인터페이스, DLITE(Digital Library Integrated Task Environment), 메타데이터 구조 등의 여러 서비스와 기존 정보 서비스의 인터페이스, 그리고 Z39.58을 지원하는 서버-클라이언트를 개발하고

있다 (Andreas Paepcke, 1996).

나. 정보 입력/저장/접근

(1) 정보 표현 형식 및 인덱싱

멀티미디어 정보를 다루기 위해서는 이러한 정보에 대한 디지털 형태의 표현 형식이 지원되어 하므로, 단일 미디어 데이터 형식과 멀티미디어 데이터 형식이 필요하다. 따라서 특정 응용에서 지원되는 사설 데이터 형식이나 여러 응용에서 널리 쓰이는 공식 표준 또는 실제적인 표준 데이터 형식을 사용하고 있다. 단일 미디어 데이터 형식의 표준으로 JPEG, JBIG, TIFF, MPEG, Photo CD 등과 같은 공식 표준 또는 실제적인 표준이 있으며, 멀티미디어 데이터 형식의 표준으로 MHEG, SGML & HyTime, ODA, OMFI 등과 같은 공식 표준 또는 실제적인 표준이 있다.

한편 비디오, 오디오, 이미지 등과 같은 정보자원의 경우 적절한 인덱스가 없으면 효과적으로 검색하거나 관리하기 어렵다. 이에 데이터 입력 기능의 일부로서 자동 또는 수동으로 인덱스를 생성해야 한다 (M. Christel, 95).

(2) 정보 입력 및 저장

기존 도서관 시스템의 경우 목록 정보만을 생성하나 많은 디지털도서관이 정보를 생성할 수 있는 기능을 제공하고 있다. 멀티미디어 정보는 다양한 데이터 원을 가지고 있으므로 스캐너, 디지털 카메라, 디지털 캠코더, 마이크 등과 같은 다양한 입력장치를 지원하는 미디어별 입력/편집 도구와 멀티미디어 입력/편집 도구가 필요하다.

예를 들어 Xerox PARC 연구소의 디지털 도서관 프로젝트에서 개발하고 있는 제품은 디지털 도서관에서 필요한 모든 구성요소를 포함하는 것으로, 문서 이미지의 입력, 분석, 그리고 복합 프리젠테이션 기능으로 UC Berkeley 환경 디지털 도서관 프로젝트에서 사용되고 있는 문서 이미지 디코딩(Document Image Decoding: DID)과 문서 이미지 요약(Document Image Summarization: DIMSUM) 기술을 개발하고 있다.

한편 멀티미디어 정보는 대용량 데이터이므로 하드디스크와 같은 2차 기억장치 만으로는 감당하기 어려우므로 실제적인 디지털 멀티미디어 도서관을 구축하기 위해서는 계층적 저장 시스템을 사용하는 일반적이다. 예를 들어 버클리소재 캘리포니아주립대학 디지털도서관 프로젝트와 IBM 디지털도서관 제품 모두 계층적 저장 시스템을 사용하고 있는데, 이들은 하드디스크, 광디스크, 광선반, 그리고 테이프 선반 등으로 구성되는데, 각각은 데이터 접근 성능과 비용당 저장용량이 서로 반비례 되는 특성을 가지고 있다 (Virginia Ogle, 1996).

다. 정보 보존

기존 도서관에서는 인쇄물에 대해 중성지를 사용하거나 마이크로 필름으로 만드는 등 여러가지 보존 기법이 사용되고 있다. 또한 영화 산업에서도 비디오 테이프에 대한 주기적인 재녹화, 색상 다시 입히기 등의 다양한 보존 방안을 사용하고 있다. 그러나 현재 디지털화된 데이터의 경우 하드웨어 장치, 소프트웨어, 그리고 사회환경이 급속하게 변화하므로 이러한 데이터의 보존 방안에 대한 연구가 부족하다. 이에 최근 선진 여러나라에서는 이에 대해 연구가 진행되기 시작하였는데 이러한 연구는 정보의 디지털화 방안과 연계하여 추진하고 있다.

예를 들어 미국에서는 The Commission on Preservation and Access과 The Research Libraries Group에서 디지털화된 데이터의 보존 방안에 대해 연구하여 1996년 보고서를 발간하였다. 그리고 영국에서는 HEFCE(Higher Education Funding Councils)의 JISC(Joint Information Systems Committee)와 BL(British Library)가 모여 디지털화된 데이터의 보존 방안에 대한 필요성을 인식하고 1995년부터 연구를 시작하였으며 1996년 NPO(National Preservation Office) 주관으로 이에 대한 학술회의를 개최하였다. 또한 호주에서도 National Library of Australia의 National Preservation Office 주도로 디지털화된 데이터의 보존 방안을 연구하여 보고서를 발간하였다.

라. 데이터 검색 및 브라우징

기존 문자나 숫자와 같은 정형 데이터와 텍스트 데이터에 대한 내용기반 검색

방법에 대해서는 많은 연구가 진행되어 왔으며, 사용가능한 여러 검색 방법이 있다. 멀티미디어 정보를 관리하는 디지털 도서관을 구축하기 위해서는 대용량인 멀티미디어 정보를 효율적으로 입출력, 저장하는 방법과 더불어 효과적으로 원하는 정보를 검색할 수 있는 방법이 제공되어야 한다. 그런데 멀티미디어 정보에 포함된 이미지, 오디오, 비디오와 같은 비정형 데이터에 대해서는 주로 입출력 성능에 대한 연구를 많이 진행하였고, 최근에는 내용기반 검색 방법에 대한 연구를 활발히 진행하고 있다.

이미지, 오디오, 비디오 미디어에 대한 내용기반 검색은 크게 실제 데이터에 대해 검색하는 방법과 실제 데이터로부터 자동 또는 수동으로 추출된 메타데이터에 대해 검색하는 방법이 있다.

이미지 정보자원의 경우 IBM의 QBIC(Query By Image Content), Virage사의 상용제품, Excalibur사의 Visual RetrievalWare 등이 내용기반 검색 기능을 제공하고 있고, 무늬, 색 등과 같은 정보를 실제 데이터로부터 추출하거나 데이터 입력시 해당 키워드를 입력한 메타데이터에 의해 검색하고 있다.

오디오 정보자원의 경우 홍얼거림 또는 실제 오디오의 파형에 의해 찾는 방법과 오디오에 대한 메타데이터를 Dublin Core 메타데이터에 기반하여 정의하고 이를 이용하여 검색하는 방법이 사용되고 있다.

비디오 정보자원의 경우 비디오 전체를 여러 논리적인 단위로 나누고 이에 대해 이미지 기반 검색을 수행하거나 특정 객체의 움직임에 의한 검색을 수행할 수 있고, 키워드를 입력할 때 일반적인 키워드 입력에 의한 메타데이터를 구축하거나 주석기반 검색 방식을 사용하고 있다.

그런데 현재 제안된 내용기반 검색 방법이 불완전하므로 브라우징 방법으로 보완하고 있는 시스템이 대부분이다. 예를 들어 Xerox PARC 연구소의 디지털도서관 제품은 정보 접근과 시각화로 많은 텍스트로 구성된 컬렉션을 탐색, 브라우징, 그리고 시각화하는 기능과 이를 요약할 수 있는 기능, 그리고 구조의 자동 인식 기능으로 Scatter/Gather 문서 클러스터링 기법, 타이틀바 방식과 질문 응답 방식의 세가지 텍스트 검색 및 브라우징 기법과 Cone-Tree 방식, Perspective Wall, Table Lens 방식을 통합한 Butterfly라고 하는 3차원 정보 시각화 기법 등을 개발하고 있다.

마. 저작권 관리 기능

일반적인 디지털도서관에서도 요구하는 저작권 관리 시스템은 멀티미디어 도서관인 경우 문제가 보다 심각해진다. 왜냐하면 멀티미디어 정보란 여러가지 대용량 미디어 데이터를 혼합한 형태이므로 현실적으로 이를 모두 다시 생성하기 어렵다. 따라서 멀티미디어 정보를 만들기 위해서는 기존 미디어 데이터베이스로부터 원하는 내용을 검색, 추출하여 이를 재사용하는 것이 일반적인 작업 수행 방식이 될 것이다. 이러한 멀티미디어 데이터를 효과적으로 사용하기 위해서는 생성부터 소멸까지의 유통 체제 안에서 저작권 관리 기능이 제공되어야 한다. 이 기능은 다시 자동화된 저작권 설정 및 로얄티 교환, 저작권물 안전 보장, 정보사용 인증, 효과적인 청구 및 지불 메커니즘으로 나눌 수 있다.

이러한 저작권을 해결하는 여러 디지털도서관 프로젝트로는 7개 박물관과 7개 대학의 연합으로 디지털화된 박물관 이미지와 관련 정보의 교육적 활용을 위한 MESL(Getty Museum Education Site Licensing Project), IEEE 등 7개 학회 및 출판사와 협력하여 학술지 논문을 SGML에 기반하여 전문 데이터베이스로 구축하고 서비스하는 일리노이주립대 디지털도서관 프로젝트, Elsevier Science 출판사 주도하에 학술논문의 전자적 배포에 관한 문제점을 조사하기 위해 9개 대학과 초기 약 40종, 최종에는 83여 종의 재료학 학술지 원문을 서비스하는 TULIP 프로젝트 등이 있다.

(1) 지불 메커니즘

디지털화된 정보의 전자적인 유통을 위해서는 정보서비스 사용의 대가로 비용을 지불하는 메커니즘이 디지털도서관 시스템에 필요하다. 이를 오라클 사의 지불 서버(Payment Server) 외에 다음과 같이 구현된 메커니즘이 있다.

Stanford 대학의 디지털도서관 프로젝트 중의 하나로 InterPay는 고객과 서비스 간의 상호작용을 수행하는 서비스 계층의 하부로 지불 정책을 담당하는 지불 정책 계층(Payment Policy Layer)과 실제 지불을 수행하는 지불 메커니즘 계층(Payment Mechanism Layer)으로 구성된다. 지불 정책 계층에서는 고객대신 지불 대리인

(Payment Agent), 서비스대신 콜렉션 대리인(Collection Agent)이 상호작용을 수행한다. 지불 메커니즘에서는 고객 대신 지불 능력(Payment Capability), 서비스 대신 콜렉션 능력(Collection Capability)이 명확한 지불 수단으로 상호작용하게 되는데, 예를 들어 이러한 능력은 VISA, Master 등이 된다.

미시간대학교 디지털도서관 프로젝트나 일리노이주립대학교 디지털도서관 프로젝트에서는 카네기멜론대학교 디지털도서관 프로젝트에서 사용되는 NetBill을 기초로 해서 개인 이용자에 대한 과금을 검토하고 있다.

(2) 저작권 등록 및 보호

미국 국회도서관의 미국 저작권 사무소(United States Copyright Office)과 CNRI(Corporation for National Research Initiatives)에서 개발하고 있는 저작권 관리 시스템 테스트베드, CORDS(Copyright Office Registration, Recordation and Deposit System)는 디지털화된 정보자원에 대해서 전자적으로 저작권 등록에 관한 모든 작업을 수행하도록 설계되었다. 이 시스템은 여러 미디어 데이터 형식을 지원하고 있으며 보다 확장하고 있다.

워터마킹 기법은 암호화나 인증(Authentication)과 같은 강제적인 저작권 보호 방식이 아니라 각 정보자원마다 식별가능한 내용이 내재되어 불법적인 복사를 심리적으로 막는 간접적인 보호 방식이므로 저작권 관리 시스템의 보조 수단으로 많이 사용된다. 그런데 기존 워터마킹 기법은 인쇄물 또는 이미지 데이터만을 다루므로 멀티미디어 정보에 포함된 비디오나 오디오 등의 다양한 미디어 데이터를 다룰 수 있는 워터마킹 기법이 활발히 연구되고 있다. 예를 들어 SysCoP(System for Copyright Protection)의 경우 저작권 표시가 각 정보에 내재되나 볼 수는 없도록 하며, 기존 미디어 데이터 표준과 호환을 유지하도록 하는데, 이미지의 경우 PPM(PGM, PBM)을 직접 지원하고, JPEG, TIFF, GIF는 PPM으로 변환된다. 동영상의 경우 MPEG-1을 지원하며, 텍스트 이미지 문서의 경우 Postscript를, 오디오의 경우 wave 형식을 직접 지원하며, SUN의 au와 같은 다른 데이터 형식도 변환가능하다 (Jian Zhao, 1996).

바. 정보 프리젠테이션 및 전자배달

실세계에 존재하는 대다수의 정보는 정형 데이터 뿐만 아니라 텍스트, 이미지, 오디오, 비디오 등과 같은 모노미디어 데이터가 다양한 관계에 따라 다단계로 연관되어 있는 멀티미디어 데이터 객체이다. 그리고 프리젠테이션 명세는 데이터 객체가 실세계에서 프리젠테이션되는 형태를 설명하는 추상적인 정보 개체이다.

멀티미디어 객체를 프리젠테이션하기 위해서는 내재되어 있는 프리젠테이션 명세를 표현, 관리하는 기능과 객체를 해당 프리젠테이션 명세에 따라 프리젠테이션하는 기능이 필요한데, 이러한 두 기능을 제공하는 프레임워크를 멀티미디어 객체를 위한 프리젠테이션 환경이라 한다. 멀티미디어 정보가 시공간 관계와 사용자 상호작용 관계를 가지고 있는 경우 이를 지원해야 한다.

또한 멀티미디어 정보는 다른 출력 장치를 통해 프리젠테이션할 수도 있다. 예를 들면 텍스트 데이터에 대해 음성 합성을 통해 소리로 출력하거나 통계를 그래픽으로 보여주기도 한다.

한편 고객의 원격 주문에 따라 디지털화된 정보의 네트워크를 통한 원격 전달 및 프린팅 시스템으로 일리노이 주립대학에서 사용하고 있는 Ariel 시스템이나 Xerox PARC 연구소의 UPrint1 등이 있다.

3. 디지털 멀티미디어 도서관의 현황

디지털도서관의 구성요소인 하드웨어, 소프트웨어, 컬렉션, 그리고 지원인력 가운데 컬렉션 관점에서 디지털 멀티미디어 도서관을 나누어 설명한다. 즉, 컬렉션을 구성하는 핵심 미디어 타입에 따라 나누고 각 분류에 속하는 대표적인 디지털도서관을 살펴보기로 한다.

가. 단일미디어 도서관

(1) 디지털 텍스트 도서관

Andrew W. Mellon 재단의 지원으로 추진되는 JSTOR(Journal Storage) 프로젝트에서는 정선된 학술지를 과거로 소급해서 스캔하고, OCR을 이용해서 텍스트로 변환하여, 개시 시점에서 75만 페이지 이상의 컬렉션을 구축할 예정이다. 이 프로젝트의 초점은 역사·경제분야의 10개 학술지(19세기부터 1990년까지의 발행분)를 변환입력하는 것이었다. 현재는 소장범위를 다른 학술지로 넓히는 것과 기발행분에 더해서 최신호를 입력하는 것 등으로 JSTOR 프로젝트는 급속히 확대해가고 있다.

Cornell 대학 전산학과 주도로 시작된 NCSTRL(Networked Computer Science Technical Reports Library)은 ARPA에서 지원하는 분산 기술보고서 도서관 시스템이라 할수 있다. NCSTRL 컬렉션은 각 참여 기관의 해당 서버의 연동체제에 분산되어 있다.

SGML 데이터베이스로는 Royal Irish Academy와 University College Cork의 The CURIA 프로젝트가 있는데, 이 프로젝트에서는 초기, 중세, 그리고 현재 아일랜드의 언어로 쓰여진 문헌 및 역사물의 대화적 온라인 검색 가능한 데이터베이스를 SGML을 사용하여 구축하고 있다. 일리노이주립대 디지털도서관 프로젝트에서는 과학기술 분야의 대규모 학술문헌을 많은 이용자에게 제공한다. IEEE 등 7개 학회 및 출판사와 협력하고, 학술지 논문을 SGML에 기반하여 전문 데이터베이스로 구축하여 서비스한다.

TULIP 프로젝트는 Elsevier Science Publication 주도하에 학술논문의 전자적 배포에 관한 문제점을 조사하는 것을 목표로 하고 있으며, 여기에는 9개 대학과 약 초기 40종, 최종에는 83여 종의 재료학 학술지가 관련되어 있다. Elsevier 사는 학술지에 대해서 TIFF 형식의 비트맵 이미지, 서지 사항과 초록을 포함하는 기사 앞부분의 ASCII 텍스트, 그리고 프리젠테이션에는 사용하지 않고 검색에만 사용할 OCR로 생성한 본문의 전문 텍스트를 제공했다. 기사의 처음 부분에 있는 서지 데이터(표제, 저자 등), 키워드, 초록은 사람이 직접 검사하여 철자가 교정된다. 미정정 전문, 즉 OCR로 읽혀진 그대로의 기사 본문(리드아웃 에러가 존재하는 것)이 데이터로서 비축되며, Full-text 검색에 사용된다. 페이지 이미지는 압축되어 1페이지당 70~80Kbytes로 된다. 미정정 전문은 1페이지당 4Kbytes, 서지데이터 등은 1.5Kbytes이다. Elsevier는 Web 서비스를 1994년 9월에 시작했다.

(2) 디지털 그래픽/이미지 도서관

캘리포니아주립대 산타바바라분교의 Alexandria 디지털도서관 프로젝트에서는 지도나 항공사진 등을 디지털화하고 대규모 컬렉션을 구축하고 있다. 이러한 대규모 컬렉션의 구축과 공간 정보의 상호이용성을 높이기 위해서 메타데이터베이스에 관한 연구를 진행하고 있다.

MESL(Getty Museum Education Site Licensing Project)은 7개 박물관과 7개 대학의 연합으로 디지털화된 박물관 이미지와 관련 정보의 교육적 활용을 위한 저작권 해결 방안을 연구하고 있다. 즉 참가 기관들은 크게 다음과 같은 세가지 역할을 수행하고 있다. 첫째, 모범적인 교육기관 라이선스를 제시하고 둘째, 박물관의 디지털화된 이미지 및 정보에 대한 수집, 배분 절차를 시험, 평가하며 셋째, 이러한 배분의 파급효과를 기술적이고 경제적인 기준으로 평가한다.

(3) 디지털 오디오도서관

국회도서관은 6,000,000편 이상의 곡을 소장하고 있고, 프랑스 국립도서관은 300,000시간분의 녹음을 소장하고 있는데, 일반인이나 음악 전문 사서는 종종 곡의 일부를 부르거나 흥얼거린 것을 가지고 원하는 음악을 찾는다. 뉴질랜드의 MELDEX(MELody inDEX) 디지털도서관은 노래 전문 디지털도서관으로 곡을 입력, 저장하는 기능 외에 마이크를 통해 노래의 일부를 부르면 이를 가지고 원하는 노래를 찾아주는 내용기반 검색 기능을 제공하고 있다. 이 도서관의 노래 데이터베이스는 9,400개의 민요를 포함하고 있다 (R.J. McNab, 1997).

Xerox PARC의 디지털 전통민요 데이터베이스는 1988년 봄부터 구축되기 시작하여 1996년 4월 현재 5630개 이상의 노래를 포함하고 있으며 매달 약 300개 노래를 추가하고 있다. 이러한 민요는 인터넷을 통해 접근하여 목록이나 악보를 보거나 곡을 들을 수가 있다. 현재 이러한 민요에 대한 목록 데이터베이스 구축에 대해 연구하고 있다.

(4) 디지털 비디오/애니메이션 도서관

카네기멜론대학의 Infromedia 디지털 비디오 도서관 프로젝트에서는 방송국과 협력하고 대화형 검색과 시청을 할 수 있는 비디오 도서관을 만들고 있다. 원하는 내용을 효율적으로 찾기 위해 음성인식 기술, 영상인식 기술, 자연언어 처리 기술, 비디오 데이터의 자동 색인 기술을 개발하고 있다. 초기에는 WQED/Pittsburgh의 비디오 데이터베이스에서 추출된 1,000시간 분량의 비디오를 가지고 시작하였다.

나. 복합미디어 디지털도서관

미국 국회도서관(Library of Congress)의 국가 디지털도서관 프로그램과 프랑스 국립도서관에서는 소장하고 있는 고유 텍스트, 그림, 오디오, 비디오 자료에 대한 많은 종류의 Collection을 구축하였으며 계속 확장 중에 있다.

Tufts 대학 주도로 1985년부터 추진하고 있는 Perseus 프로젝트는 고대 그리스와 로마 시대에 초점을 둔 유명하고 인정받는 컬렉션을 구축하고 있다. 이 컬렉션에는 문헌, 화병, 조각, 동전, 건물 뿐만아니라 위성 사진으로부터 추출한 그리스 천연색 지도를 포함하고 있다. 현재 300개의 고대 문헌 및 번역 내용과 1,412개의 화병, 366개의 조각, 524개의 동전, 384개의 건물, 179개의 지역에 대한 자세한 목록 정보와 그림으로 구성되어 있다. 이러한 컬렉션은 CD-ROM 버전으로 제공되고 있으며 추후 Perseus 디지털도서관은 인터넷을 통해 제공할 예정이다. 또한 Shakespeare의 원본이나 다른 르네상스 시대 자료를 포함하도록 확장하는 계획을 추진하고 있다.

Berne 대학도서관의 Ryhiner 프로젝트는 1993년부터 1997년까지 지구 전체에 대한 15,000개 이상의 지도, 차트, 그리고 16세기부터 18세기까지의 모습과 계획으로 구성된 컬렉션을 구축하고 있다. Public Records Office의 20,000개 지도와 함께 이 도서관은 지역적이 아니라 전세계적인 지리 데이터베이스를 소유하고 있다. 또한 이 프로젝트에서는 이러한 컬렉션에 대한 보존, 마이크로필름 생성, 일반적으로 접근가능한 목록 구축 등을 추진하고 있다.

일본 NDL(National Diet Library)에서는 국가종합목록 구축 프로젝트 외에 멀티미디어 데이터베이스를 구축하는 전자도서관 실험 프로젝트를 수행하고 있는데 약 10,000,000 페이지의 이미지 데이터베이스 3,000개의 CD-R에 저장되어 서비스되고 있으며, 이미지 데이터 형식은 흑백 이미지의 경우 TIFF 형식으로, 천연색 이미지의

경우 JPEG 형식을 기반으로 한다. 각 컬렉션이 다른 정보검색 인터페이스를 사용하므로 이를 수용할 수 있도록 하고 있으며 Web 브라우저가 TIFF 형식과 JPEG 형식을 지원하지 못하는 경우를 위해 서비스 시스템이 이를 GIF 형식으로 변환하고 전달한다. 이외에도 NDL은 자체 소장한 어린이용 책 약 100,000권의 목록 데이터베이스가 있는데, 이중에서 저작권 협상이 용이한 1950년 이전 책, 즉 6,000권 이상의 어린이용 책을 TIFF 400dpi의 흑백 이미지 또는 칼라 이미지로 디지털화하였다. 특히 130 권의 경우 키보드를 사용하지 않고 소리로 원하는 정보를 얻을 수 있도록 하였다.

캘리포니아주립대 버클리분교에서는 환경정보에 대한 대규모 데이터베이스를 구축한다. 연구 프로젝트에는 공공도서관도 포함되며, 항공사진 등을 포함하는 다양한 데이터를 다양한 이용자에게 제공하는 환경을 구축한다.

미시건대학교의 디지털도서관 프로젝트에서는 우주·지구과학 분야의 다양한 자료를 고교생으로부터 연구자에 이르기까지 폭넓은 다수의 이용자에게 제공하는 것을 목적으로 하고 있다. 에이전트 모델에 기초해서 시스템을 실현한다.

4. 결론: 디지털 멀티미디어 도서관의 나아갈 방향

디지털 멀티미디어 도서관은 크게 하드웨어, 소프트웨어, 컬렉션, 그리고 인적 자원으로 구성되므로, 이러한 구성요소가 멀티미디어를 지원하기 위해 개선되고 있는데, 이 중에서도 다음과 같은 문제를 시급히 해결해야 한다.

일반적인 디지털도서관에서도 요구하는 저작권 관리 시스템은 멀티미디어 도서관인 경우 문제가 보다 심각해진다. 왜냐하면 멀티미디어 정보란 여러가지 대용량 미디어 데이터를 혼합한 형태이므로 현실적으로 이를 모두 다시 생성하기 어렵다. 따라서 멀티미디어 정보를 만들기 위해서는 기존 미디어 데이터베이스로부터 원하는 내용을 검색, 추출하여 이를 재사용하는 것이 일반적인 작업 수행 방식이 될 것이다. 이러한 멀티미디어 데이터를 효과적으로 사용하기 위해서는 생성부터 소멸까지 효율적인 유통 체제가 구축되어야 한다.

멀티미디어 정보에 포함된 다양한 미디어 데이터 간에는 논리적인 관계에 따라 상호연관되었을 뿐만 아니라 시공간 관계와 상호작용 관계도 존재한다. 따라서 이러한 관계에 따라 멀티미디어 정보를 표현, 검색, 프리젠테이션할 수 있어야 한다. 그리고 각 미디어마다 다양한 내용기반 검색 방법이 개발되어야 한다.

이외에도 디지털 멀티미디어 도서관에서 사서는 멀티미디어 목록과 여러 기존 자료를 모아 하나의 멀티미디어 자료를 만드는 역할을 할 수도 있으므로 기존 도서관 시스템에서 추진하고 있는 주제별 사서 뿐만 아니라 미디어별 내용 전문가와 이를 조정할 수 있는 프로그램 디자이너가 필요할 것이므로 이를 육성해야 한다.

참고문헌

- (Philippe Aigrain, 1996) Philippe Aigrain, "Image and Sound Digital Libraries Need More Than Storage and Networked Access," pages 112-118, Proc. of Int'l Sym. on Digital Libraries 1995, Aug. 22-25, 1995, Tsukuba, Japan
- (William Y. Arms, 1996) William Y. Arms, Christophe Blanchi, Edward A. Overly, "An Architecture for Information in Digital Libraries," D-Lib Magazine, Feb. 1997
- (M. Christel, 1995) M. Christel, A.Hauptmann, T.Kanade, M.Maudlin, R.Reddy, M.Smith, S.Stevens, "Technical Challenges for the Informedia Digital Video Library," pages 10-16, Proc. of Int'l Sym. on Digital Libraries 1995, Aug. 22-25, 1995, Tsukuba, Japan
- (CPA, 1996) The Commission on Preservation and Access and The Research Libraries Group, "Preserving Digital Information", Report of the Task Force on Archiving of Digital Information, The Commission on Preservation and Access and The Research Libraries Group, May 1, 1996
- (Rodger J. McNab, 1997) Rodger J. McNab, Lloyd A. Smith, David Bainbridge, and Ian H. Witten, "The New Zealand Digital Library MELody inDEX,"

D-Lib Magazin, May 1997

(Virginia Ogle, 1996) Virginia Ogle, "Testbed Development for the Berkeley Digital Library Project," D-lib magazine, July/August 1996

(Andreas Paepcke, 1996) Andreas Paepcke, "Summary of Stanford's Digital Library Testbed Design and Status," D-lib magazine, July/August 1996

(Jian Zhao, 1997) Jian Zhao, Eckhard Koch, "A Digital Watermarking System for Multimedia Copyright Protection," ACM Multimedia '96, pp. 443-444, Boston, MA, USA, 1996.