

# 향토문화 콘텐츠를 위한 디지털 편찬 관리시스템 개발에 관한 연구: 『한국향토문화전자대전』의 사례를 중심으로

**A Study on Development of Digital  
Compilation Management System for Local Culture Contents:  
Focusing on the Case of *The Encyclopedia of Korean Local Culture***

김수영(Su-Young Kim)\*

## 초 록

향토문화란 한 지방의 자연환경 속에서 과거로부터 현재까지 면면히 전승되어 온 역사와 전통, 풍물과 생활, 예술과 유물 및 유적 등의 모든 유산을 의미한다. 한국학중앙연구원에서는 이런 향토문화를 디지털 콘텐츠로 제작하고 이를 이용하여 『한국향토문화전자대전』을 편찬하였다. 향토문화 콘텐츠는 기록물의 특성을 가지고 있어 출처주의, 계층목록과 같은 기록물관리체계를 따르고 있으며 이를 편찬·관리하기 위한 시스템은 기초자료, 단편적 정보 및 고급정보 등이 하나의 시스템 안에서 순환하면서 새로운 지식정보를 만들어내도록 도와주는 순환형 지식정보관리시스템을 지향한다. 순환형 지식정보관리시스템의 이용자는 이 시스템을 통해 직접적으로 자료를 수집할 수 있을 뿐만 아니라 다른 곳으로부터 데이터를 수집할 수 있으며, 나아가 수집한 데이터를 가공하여 새로운 지식 정보를 창출할 수 있다. 그러나 다양한 향토문화 콘텐츠의 구조에 포함된 의미적인 특징을 손상시키지 않고 데이터베이스를 구축하기가 매우 어렵고, 또한 이러한 작업은 장시간에 걸쳐 여러 차례의 교정 작업이 수행되어야 하므로 문서편찬, 교정, 서비스가 동시에 수행되는 시스템을 필요로 한다. 따라서 본 논문에서는 『한국향토문화전자대전』의 사례를 중심으로 고문서가 많이 포함된 향토문화 콘텐츠의 의미적인 특징을 손상시키지 않으면서, 문서의 구조정보를 표현할 수 있는 XML 기반의 디지털 편찬 관리시스템을 제시하고 본 연구에서 개발된 관리시스템에서 향토문화 콘텐츠 관리를 위해 확장된 기능을 소개한다.

## ABSTRACT

Local culture is a cultural heritage that has come down from generation to generation in the natural environment of a region. It includes history, tradition, natural features, art, and historic relics. The Academy of Korean Studies has complied "The Encyclopedia of Korean Local Culture" using those local culture contents. Local culture content share the features of documentary, such as authenticating the source, and managing hierarchy structure. Thus, to deal with local culture contents, a "circular knowledge information management system" is sought for that helps basic, fragmentary, and high-level information to circulate to create new knowledge information within the system. A user of this circular knowledge information management system is able not only to collect data directly in it, but also to fetch data from other database. Besides, processing the collected data helps to create new knowledge information. But, it's very difficult to sustain the features of the original hierarchy bearing meaning contained in the various kinds of local culture contents when building a new database. Moreover, this kind of work needs many times of correction over a long period of time. Therefore, a system in which compilation, correction, and service can be done simultaneously is needed. Therefore, in this study, focusing on the case of "The Encyclopedia of Korean Local Culture", I propose a XML-based digital compilation management system that can express hierarchy information and sustain the semantic features of the local culture contents containing lots of ancient documents, and introduce the expanded functions developed to manage contents in the system.

키워드: 한국향토문화전자대전, 디지털 콘텐츠, 편찬 관리시스템, 메타데이터, 데이터베이스  
the encyclopedia of korean local culture, digital contents, compilation management system, metadata, database

\* 한국학중앙연구원 한국학정보센터 연구원(tmt21@aks.ac.kr)

■ 논문접수일자 : 2009년 8월 14일 ■ 최초심사일자 : 2009년 8월 18일 ■ 게재확정일자 : 2009년 8월 25일  
■ 정보관리학회지, 26(3): 213-237, 2009. [DOI:10.3743/KOSIM.2009.26.3.213]

## 1. 머리말

한국학중앙연구원은 전국의 지방자치단체를 하나의 범주로 향토문화를 집적하여 인터넷 서비스하는 '한국향토문화전자대전 편찬사업'을 수행하고 있다. 2003년 착수한 성남문화대전을 시작으로 2009년 현재까지 총 32개 지역을 대상으로 편찬 완료 및 콘텐츠 구축 사업을 진행하고 있다.

『한국향토문화전자대전』은 그 성격과 내용상 사전의 속성을 가지는 것이기는 하나 그 형태가 일반 책과는 다르다. 특히 편찬물의 표출을 디지털 형태로 하고 있다. 이런 이유로 이용자들에게 디지털매체로써 보여주기 위해서는 이것에 적합한 시스템적인 안이 마련되어야 한다.

『한국향토문화전자대전』 시스템의 구축을 위하여 정보편찬 및 기능구현의 기본단계인 콘텐츠를 연동하는 기본구조로서 메타데이터를 어떻게 설계하고, 텍스트를 어떻게 구성하여 집필해야하는가는 서비스 시스템 운영의 효율성과 직결된 것으로, 곧 편찬 실무에 직접 참여하는 개발자와 반대로 제작된 것을 사용하는 이용자의 입장에서 고려해야 하는 중요한 것이다.

그러나 지금까지는 대부분의 디지털 콘텐츠 구축 및 편찬 사업들이 표준화된 프로세스 없이 수행 주체의 특성에 따라 주먹구구식으로 진행되고 있는 것이 현실이다. 이는 콘텐츠라는 것 자체가 저마다의 독창성을 지니고 있음에 기인한다. 콘텐츠 구축의 주체가 그 내용에 대하여 특화된 인력과 조직으로만 구성되어 있을 뿐, 대량의 표준화된 콘텐츠를 구축하는데 필요한 표준화 관리, 공정 관리, 시스템 관리에 대한 고려가 부족하다. 또한 유사한 콘텐츠 구

축 사업에서 실제적으로 비슷하거나 동일한 기초자료들을 필요로 하지만 이러한 자료들은 각 콘텐츠 구축 사업의 결과를 만들기 위한 중간부산물로 인식될 뿐 기초자료 자체의 재사용에 대한 고려나 연구가 미흡하다.

따라서 이런 문제점을 해결하기 위하여 본 연구는 한국학중앙연구원에서 수행하고 있는 32개 지역 중 가장 최근에 편찬 한 디지털진천문화대전과 디지털충주문화대전의 사례를 중심으로 향토문화 콘텐츠 편찬 과정에서 수집되고 생산되는 자료를 수집, 정리하여 문서편찬, 교정 및 서비스에 이르기까지의 표준적인 과정과 구조를 설계하고 시스템으로 구현하여 그 발전방안을 모색하는 것을 목표로 한다.

## 2. 관련연구

국내의 경우, 디지털 콘텐츠 관리시스템은 대부분 아카이브로 관리되고 있으며 그 움직임은 시작단계라 할 수 있다. 현재 국회도서관, KERIS, KISTI 등의 국가기관과 북토피아, DBPIA 등의 개인기업체에서 디지털 콘텐츠를 저장 관리하고 있다. 그리고 국가기록원은 기록물관리법 시행령 제 30조에 의거 자료관리시스템(정보공개청구의 접수 등 기록물의 관리가 전자적으로 수행)과 전문관리기관시스템(중앙기록물관리기관, 특수기록물관리기관 등에서의 기록물의 수집, 평가, 재분류, 보존)의 구축을 진행 중이다. 하지만, 이를 디지털 아카이브는 상기 보존 및 엑세스에 대한 체계적 접근이 없는 상태로서 진정한 의미의 아카이브로 보기 어려운 실정이다.

## 2.1 대학 및 기관 중심 디지털 아카이브의 동향 분석

정부에서 시행한 디지털 아카이브 구축 사업으로는 국가적으로 보존 및 이용가치가 있고 학술, 문화 또는 과학 기술 등에 관한 디지털화 또는 디지털화의 필요성이 인정되는 자료를 지식정보자원이라 정의하고 이를 데이터베이스로 구축하는 사업을 추진하였다.

구축된 지식정보 자원을 보면, 2008년 말 2억 5천만건의 정보 자원이 연계, 서비스되고 있다. 각 분야별로 차지하는 비중은 <표 1>에서 보듯이 교육학술이 46.7%로 가장 높으며, 다음은 과학기술·정보통신 42.1%, 그리고 문화·역사 8.8%, 기타 2.3% 순이다.<sup>1)</sup> 다른 분야에 비해 인문학 범주에 속하는 문화와 역사 분야의 DB구축 비중이 매우 낮게 나타나고 있음을 알 수 있다.

해외에서는 디지털 아카이빙에 관한 연구가 활발하게 진행되고 있으며, 현재까지 진행된 주요 연구 사례를 살펴보면 다음과 같다.

1) Yale 대학교의 디지털 아카이브  
2001년도에 Mellon 재단의 지원을 받아 Yale

대학 도서관과 Elsevier사의 공동연구로 수행된 이 연구에서는 첫째, 주 정보자원의 장기적 손상, 권리보유자 또는 승계자의 책임포기 등으로 인한 이용성의 손상, 합의된 기간의 경과 등 디지털 아카이빙을 촉발할 수 있는 사안들에 대한 연구, 둘째, 아카이브 구축에 비용이 소요되는 주된 요인들을 밝혀내는 비용요인에 관한 연구, 셋째, 이러한 비용조달 모델 및 라이센스 모델에 관한 연구와 기타 아카이브 대상자료 선정과 메타데이터에 관한 연구를 수행하였다.

### 2) Harvard 대학교의 디지털 아카이브

Harvard 대학은 Blackwell, Wiley, University of Chicago Press의 3개 출판사와의 협력을 바탕으로 추진한 디지털 아카이빙 연구를 통해 저널의 구성요소별 아카이브 대상요소의 선정 기준을 만들었고, 비즈니스 모델 및 아키텍처 설계에 관한 연구를 마쳤으며 아카이브 추진 주제별 역할과 임무를 정의하였다.

### 3) 유럽연합 프로젝트 디지털 아카이브

NEDLIB(Networked European Deposit Library) 프로젝트는 유럽 국립도서관(프랑스, 노르웨이, 핀란드, 독일, 포르투칼, 스위스, 이

<표 1> 분야별 지식정보자원 디지털화 현황

연도	과학기술·정보통신	교육학술	문화·역사	기타	합계
DB구축건수	106,011 천 건	117,643천 건	22,132천 건	5,907천 건	251,693천건

※ 과학기술·정보통신 분야: 과학기술, 산업기술, 건설기술, 해양수산정보 등

※ 교육학술 분야: 학술논문, 사회통계, 국방학술정보 등

※ 문화·역사 분야: 국가문화유산, 한국영화자료, 한국역사정보, 유교문화정보 등

※ 기타 분야: 산림정보 탐사용 항공정보, 근대법령 지식정보 등

1) 정보통신부 보도자료, 「지식정보자원관리사업의 향후 방향 및 상업적 활용 방안」, <www.mic.go.kr>.

탈리아 등)들의 협력 프로젝트로서 European Commission's Telematics Application Programme의 자금 지원 하에 네덜란드 국립 도서관이 진행하고 있다. 이 프로젝트의 목적은 현재의 디지털 정보가 현재와 미래에도 사용될 수 있는 네트워크화 된 유럽 보존 도서관을 구축할 수 있는 기본적인 하부구조를 구성하는데 있다. 이 프로젝트는 OAIS(Open Archival Information System) 참고 모델(reference model)을 기준으로 구축되었으며 독립적이고 개별적인 시스템 구조로 구성된 것이 아니라 기준 도서관에서 운영되고 있는 도서관 자동화 시스템과의 상호작용으로 구축되었다.

#### 4) PANDORA 프로젝트

호주 국립도서관이 추진하는 디지털 아카이브 구축 과제로서 호주 정부간행물 및 관련 웹 정보를 수집하여 보존하는 PANDORA Digital Archive System(PANDAS)을 개발 운영하고 있다.

#### 5) CEDARS(Curl Exemplars in Digital ARchiveS) 프로젝트

영국의 Leeds, Oxford, Cambridge 등 3개 대학이 공동으로 수행한 e-Library 일환인 CEDARS 프로젝트에서는 디지털 정보자원 보존정책의 수립과 보존 메카니즘의 개발을 목표로, OAIS(Open Archival Information System) 참조 모델을 토대로 한 디지털 정보자원의 보전에 대한 기술적, 조직적 전략 연구, 콘텐츠 관리정책에 대한 가이드라인 제시, 장기보전에 필요한 메타데이터 연구를 수행한 바 있다.

## 2.2 향토문화관련 디지털 아카이브의 동향 분석

대부분의 향토자료 관리는 각 지역의 문화원에서 이루어지고 있으나 열악한 인력과 재정적인 문제로 향토자료의 체계적인 자료수집과 관리가 이루어지지 못하고 있다. 또한 지자체의 향토자료의 인식의 정도가 달라 문화원마다 향토자료 관리정도의 편차가 심하며, 표준화가 제대로 이루어지지 않아서 지역의 풍부한 문화에 대한 보존과 교육 및 문화콘텐츠 제작이 제대로 활용되지 못하고 있다.

### 1) 문화원형 디지털 콘텐츠 아카이브

국내에서의 문화 콘텐츠 중심의 디지털 아카이브는 정부가 주도하고 있는 국가적인 사업으로 진행되고 있다. 문화관광부는 문화콘텐츠진흥원을 발족시키고 곧 이어서 2002년부터 문화원형을 발굴하여 디지털 콘텐츠화하는 사업을 시작하였다. 문화콘텐츠 진흥원에서는 크게 4가지 소재를 중심으로 문화 콘텐츠를 분류하였다.

- (1) 이야기형 소재: 신화, 전설, 민담, 역사, 문학 등의 문화원형 창작소재
- (2) 예술형 소재: 회화, 서예, 복식, 문양, 음악, 춤 등
- (3) 경영 및 전략형 소재: 전투, 놀이, 외교, 교역 등
- (4) 기술형 소재: 건축, 지도, 농사, 어로, 음식, 의학 등

이 네 가지 소재를 기초로 해서 개발자들은 디지털 콘텐츠를 제작하여 왔고, 그 결과물들은 문화콘텐츠닷컴([www.culturecontent.com](http://www.culturecontent.com))

에 유통되고 있다. 이곳에는 개발자들이 만든 문화관련 각종 자료들과 개발자들의 홈페이지가 연결되어 있다. 문화 콘텐츠의 포털사이트를 표방하면서 그 동안 개발된 자료들을 데이터베이스로 구축해서 이용자들에게 제공되고 있다. 하지만 문화콘텐츠닷컴은 그 구축의 특성이라고 할 수 있는 “각 과제별 사이트와의 연결”이라는 측면이 사이트 내의 통합성을 가로 막는 요인으로 작용하고 있는 것으로 보인다. 문화 전시관을 통해 접근할 수 있는 120여 가지의 과제들은 웹사이트의 주요한 사용자를 콘텐츠 구매 희망자로 설정할 것인지 정보 검색을 위한 학생 및 일반인으로 설정할 것인지를 명확히 결정하지 못하는 결점이 있다.

## 2) 한국독립운동사연구소의 구술 자료 아카이브

독립기념관 산하 한국독립운동사연구소([www.independence.or.kr](http://www.independence.or.kr))에서는 2001년 이래 제일동포 귀환자, 징병자 등에 대한 구술 기록을 수집해 왔으며, 2005년도부터는 독립운동가에 대한 구술 기록 수집을 진행하고 있다. 독립기념관은 구술 자료의 동영상, 음향, 텍스트를 웹사이트를 통하여 멀티미디어로 정보를 제공하고 있으며, 구술 자료의 수집 결과물 즉, 녹음 및 영상 자료, 녹취문이 보관되고 있다. 독립기념관 홈페이지에서 자료마당의 멀티미디어란을 살펴보면, 동영상과 음성으로 크게 구분해 놓고 대분류와 중분류로 나누어져 있다. 그 아래에는 검색조건과 함께 전체목록검색이 가능하다. 독립기념관 홈페이지 검색도구의 장점으로 볼 수 있는 것은 상세화면에서 구술기록의 간략내용과 함께 음성파일을 제공한다는 것이

다. 하지만 구축된 자료들이 구술정보에 국한되는 한계점이 있다.

## 3) 울산시 행정자료관

울산시의 행정자료관은 행정 현장에서 발생하는 각종 행정자료의 수집과 기능만을 해오다 1993년부터 본격적으로 기록물 관리를 시작하였다. 울산시 의사당 지하 2층에 위치하고 있는 자료관에는 울산시에서 발생한 문서류, 시청각 기록물류, 도면류, 간행물류, 행정박물류 등 각종 행정자료들이 보관되어 있다. 2009년 현재 문서 9만 793권, 카드 4887개, 도면 2560개, 시청각기록물 16만 960개, 행정박물 3154권 등 총 11만 8300권의 자료 분류를 완료, 보존하고 있다. 하지만 기록물의 수집과 정리 관리에 대한 표준안이 수립되어 있지 않기 때문에 그에 대한 필요성을 인식하고 있다.

## 3. 디지털 편찬 관리시스템의 메타데이터 설계원칙

『한국향토문화전자대전』에 가장 기본적으로 필요한 요소는 텍스트(본문원고) · 멀티미디어자료라는 콘텐츠와 이것을 운용하는 시스템이다. 이 중에서 텍스트는 양적인 면에서 가장 많을 뿐만 아니라 이용자들이 알고자 하는 욕구를 충족시켜주는 중요한 부분이다. 디지털 편찬 관리시스템에서 텍스트의 정보를 효과적으로 제공하기 위해서는 그 기능을 하는 요소가 필요한데, 이것을 메타데이터(Meta Data)라고 한다. 메타데이터는 일반적으로 ‘data about data’라고 하며, 대상이 되는 정보 자원의 속성

과 특성 및 다른 자원과의 관계를 기술하여, 정보이용자의 관점에서는 검색을 돋고 정보제공자에게는 정보원의 제어와 관리를 돋는 역할을 한다. 메타데이터는 실제로 저장하고자 하는 데이터는 아니지만, 이 데이터와 직접적으로 혹은 간접적으로 연관된 정보를 제공하는 데이터를 일컫는 말이다. 즉 메타데이터는 정보의 자원을 설명하는 것이다. 또 데이터에 관한 구조화된 데이터로서 정보 자체에 대한 데이터이다. 그리고 메타데이터는 데이터의 속성, 내용 등 자원을 표현하는데 필요한 요소들을 포함하고 있다. 그러므로 메타데이터는 그 자체만으로 또 하나의 중요한 자원이다. 본 연구에서 제시한 디지털 편찬 관리시스템에 쓰이는 텍스트 자료와 멀티미디어자료들의 메타데이터 구조 및 표준에 대해 설명해보도록 하겠다.

### 3.1 텍스트자료의 메타데이터 구조

현재 통일된 하나의 메타데이터 표준은 존재하지 않는다. 이는 다양한 환경과 요구에 맞는 적절한 메타데이터 포맷이 요구되며, 이에 따라 각 분야에서 사용되는 메타데이터 구조도 특정 분야의 요구에 맞게 매우 상세하고 전문적으로 개발되고 있기 때문이다. 그 결과 오늘 날 수십 여종의 메타데이터가 사용되고 있으며, 다양한 메타데이터들이 각각의 목적과 용도에 따라 개발되고 있다.

메타데이터는 디지털환경에서 필수적인 것이다. 그러므로 『한국향토문화전자대전』의 메타데이터는 컴퓨터 기능에서의 메타데이터에 디지털사전의 콘텐츠의 특성을 부각시킬 수 있는 요소와 시스템에서 필요로 하는 적합한 용

도로 설계되어야 한다. 향토문화의 단위구조는 기사제목(표제어)인 이른바 '항목'과 '메타데이터' 및 '텍스트(본문과 멀티자료)'로 구성된다.

구체적인 기본정보와 세부정보의 템플릿을 통해서 메타데이터의 설계구조와 서술요소를 비교하고 각 메타데이터의 기술요건을 제시하겠다. 본 연구에서 사용된 메타데이터는 모든 항목의 텍스트에 공통적인 기본정보를 기입하는 템플릿과 19개 유형은 물론 이것들의 하위유형인 모두 53개 세부 유형별로 각각 다른 유형별 상세정보를 기입하는 템플릿을 본문 텍스트 앞에 가지고 있다. 이 템플릿에는 해당 항목과 관련한 메타데이터의 값을 표기한다. 아래 표는 기본정보에 대한 메타데이터이다(표 2 참조).

위의 기본정보 템플릿에서 각 메타데이터 기술요건을 정리하면 다음과 같다.

- (1) 지역 향토문화전자대전에 수록될 전체 항목에서 각 개별항목에 부여된 고유한 이름(ID)을 표기한다.
- (2) 항목의 한글 명칭을 표기한다.
- (3) 항목의 한자 명칭을 표기한다. 다만 없는 경우에는 기입하지 않는다.
- (4) 공식 영문 명칭을 기입한다. 다만 없는 경우에는 표기하지 않는다.
- (5) 이용자가 찾을 가능성이 있는 이칭·약칭·별칭 등을 기입한다.
- (6) 본문 텍스트에서 핵심이 되는 주제어를 명사형으로 기입한다.
- (7) 해당지역 향토문화전자대전의 전체 항목 가운데 실제로 존재하는 항목명을 기입하여, 해당 전자대전 내에서는 물론 다른 한국학 관련 웹사이트와 연계가 가능하도록 한다.

〈표 2〉 기본정보 메타데이터

[기본 정보]				
항목ID	(1)			
항목명(한글)	(2)			
항목명(한자)	(3)			
항목명(영어)	(4)			
이칭 · 별칭	(5)			
키워드	(6)			
관련항목	(7)			
분야1	(8)	분야2	분야3	
유형1	(9)	유형2	유형3	
지역	(10) 도 · 광역시		시 · 군 · 구	
시대	(11)			
멀티미디어	사진	(12)		
	지도	(13)		
	도면	(14)		
	동영상	(15)		
	음향	(16)		
정	기타 자료	(17)		
출처	(18)			
집필자	(19)			
집필자 의견	(20)			

(8) 한국학중앙연구원 향토문화 분야분류 기준에 따라 해당 분야를 기입한다. 다만 중복성을 가진 항목의 경우는 분야1과 분야2 및 분야3이 다르게 표기될 수도 있다.

(9) 한국학중앙연구원 향토문화 유형분류 기준에 따라 해당 유형을 기입한다. 다만 중복성을 가진 항목의 경우는 유형1과 유형2 및 유형3이 다르게 표기될 수도 있다.

(10) (광역)시 · 도를 포함한 해당(중단위) 시 · 군 · 구까지만 표기한다. 보다 자세한 하위명칭은 각 항목의 유형별 상세정보의 값으로 기입한다.

(11) 한국학중앙연구원의 시대분류 기준에 따라 해당 시대를 기입한다.

(12) · (13) · (14) · (15) · (16) · (17) 각 접적인 1차 자료를 원칙으로 하며, 다만 인물의 경우는 2차 내지 3차 자료도 가능하다.

(18) 항목의 추출 대상이 된 원자료를 기입 한다.

(19) 메타데이터와 본문 텍스트를 작성한 사람의 이름을 한글로 기입한다.

(20) 집필자가 각 항목의 텍스트 원고를 집필하는 도중에 발견한 항목의 적합성, 중복, 명칭의 오류 등 여러 가지 사항에 대해 의견을 제시한다.

결국 하나의 통일된 기본정보 템플릿을 모든 항목에 적용함으로써, 비록 때로는 해당항목의 분야와 유형이 중복성을 갖는 경우라 해도 어려움 없이 적용할 수 있다.

또한 『한국향토문화전자대전』은 디지털사전인 까닭에 모든 항목마다 기본정보와 더불어 유형별 상세정보 메타데이터를 가지고 있다. 이것도 결국은 이용자의 편의를 위하여 제작자들이 제공하는 시스템적 요소이므로, 양자의 입장은 동시에 고려해야 한다. 유형별 상세정보의 메타데이터는 〈표 3〉과 같다.

〈표 3〉 유형별 상세정보 메타데이터:  
인물/전통인물

[유형별 상세정보]	
성격	문신
출신지	한양(漢陽)
성별	남
생년	1595년
몰년	1671년
본관	전주
대표관직	이조참판/이조판서(1639)/대사간

본 연구에서는 상세유형을 크게 개념용어, 기관단체, 놀이, 동물, 문현, 인물, 행사 등 총 19개로 구분하였고 나아가 각 유형 내에서 보다 자세한 하위유형별 상세정보(53개)가 제시하였다(표 4 참조).

〈표 4〉 유형 및 세부유형

유형		세부유형(53개)
1	개념용어	개념용어(일반) 개념용어(기획) 개념용어(개관)

2	기관단체	기관단체(일반) 학교 사찰
3	놀이	놀이
4	동물	동물
5	문현	단행본 연속간행물 문서
6	물품도구	물품도구
7	사건	사건사고와 사회운동 조약과 회담
8	성씨	성씨
9	식물	보호수 식물(일반)
10	유물	기명류 불상 서화류 유물(일반)
11	유적	건물 능묘 터 유물산포지 비 탑과 부도 고분 유적(일반)
12	음식물	음식물
13	의례	평생의례와 세시풍속 제
14	의복	의복
15	인물	전통인물 현대인물 가공인물
16	작품	음악, 공연작품 및 영상물 민요와 무가 무용과 민속극 문학작품 미술과 공예 설화
17	제도	법령과 제도 상훈 관직 관청(전통시대)

18	지명	행정지명과 마을
		고지명
		시설
		도로와 교량
		자연지명
19	행사	군락, 서식지 및 철새도래지 행사

### 3.2 멀티미디어자료 아카이브의 메타데이터 구조

본 연구에서는 향토문화와 관련된 각종 멀티미디어 자료들을 조사·수집하고 이를 데이터 베이스화 하여 '디지털 아카이브'로 구축하였다. 이를 통해 지역 향토문화의 전반적인 성격을 파악하여 『한국향토문화전자대전』의 편찬 방향을 설정하는 동시에 수집된 각종 자료를 편찬의 참고자원으로 활용한다.

자료수집의 기본 원칙은 해당 지역을 대상으로 서술한 자료 수집을 원칙으로 하며 전국 또는 여러 지역의 내용을 포함하는 자료라도 해당 지역에 관한 내용이 있으면 수집하도록 한다. 또한 인쇄물뿐만 아니라 음향, 사진, 영상 등의 멀티미디어 콘텐츠들도 수집의 대상에 포함한다.

위에서 설명한 자료수집의 기본 원칙에 따라 아카이브의 대상을 정리하면 아래와 같다.

- (1) 지자체(시·군청, 읍·면·동사무소, 산하 기관 및 사업소 등 모두 포함)에서 만든 공식자료
- (2) 박물관, 자료관(사료관), 미술관, 기념관, 문화원, 예술극장, 예총, 예술회관 등에서 만든 공식자료
- (3) 해당 지역의 입법·사법·행정기관과

농·축·수협, 한국전력 등 모든 정부 투자(산하) 기관에서 만든 공식자료

- (4) 방송국, 신문사 등의 언론기관과 사단법인, 재단법인 등에서 만든 공식자료
- (5) 지역문중, 청년회, 부녀회, 상가변영회, 동호회 등과 같은 지역민들의 자생적인 모임과 단체(NGO) 등에서 지역의 문화 활동과 관련해서 제작한 자료  
- 공식 등록된 단체는 아니더라도 지역 민의 삶과 밀접한 관계에 있는 단체나 모임(예: 철거민세입자 모임) 등에서 생산된 자료들도 최대한 작업 대상에 포함
- (6) 개인이 조사연구 혹은 문화·예술 산업 관련 작업을 통해 만든 자료

위에서 정리한 아카이브 대상 자료들을 디지털 아카이브 시스템에 적재하기 위해 필요한 메타데이터의 설계구조와 기술요건들을 정리하면 다음의 템플릿이 된다(표 5 참조).

여기에서 굵은 글씨는 반드시 입력해야하는 필수 입력사항이 된다.

ⓐ는 자료의 제목을 기입하는 곳으로 필수 입력사항이다. 한자로 기록된 것은 한국어 음으로 바꾸어 입력한다. ⓑ는 기타 제목을 입력하는 곳으로 제목 이외의 부제목이나 별도의 외국어(영어, 일어, 한자 등)로 된 제목이 있을 경우 입력한다. ⓒ는 검색 키워드가 되는 단어를 1개 이상 입력한다. 단, 해당 지역명은 제외하며 내용의 전반을 대표하거나 주제가 될 만한 단어로 입력한다. ⓑ는 저편자가 한 사람 이상일 경우 쉼표(,)로 구분한 후 한 칸 띄어서 입력한다. ⓑ는 발행일 기입은 YYYY(년).MM

〈표 5〉 멀티미디어 콘텐츠 아카이브 목록 템플릿

(월).DD(일)의 형식으로 입력한다. ①의 기입은 국제표준도서번호(ISBN)가 있는 도서자료의 경우 입력한다. ②는 발행처는 발행한 기관의 공식명칭을 입력함을 원칙으로 하며 발행처가 한 곳 이상일 경우는 쉼표(.)로 구분한 후 한 칸 띄어서 입력한다. ③는 도서자료의 전체 쪽 수를 입력하며 ④는 자료가 여러 권으로 이루어져 있을 경우 전체 총 권수를 입력한다. ⑤는 전국 도 단위와 광역시 단위를 입력하며 특정 지역에 한정되지 않을 경우는 대한민국으로 기입한다. ⑥는 전국 도와 광역시의 하위 단위인 시·군·구를 입력한다. ⑦은 전국 시·군·구의 하위 단위인 읍·면·동을 입력하며 특정 읍·면·동에 한정되지 않을 경우는 생략할 수 있다. ⑧은 ⑦에 입력한 지역 이외에 중복으로 해당하는 광역 단위를 입력한다. 특정 지역에 한정되지 않을 경우 대한민국으로 입력한다. ⑨은 ⑧에 입력한 지역 이외에 중복으로 해당하는 시·군·구를 입력한다. 그리고 ⑩은 ⑨에 입력한 지역 이외에 중복으로 해당하는 읍·면·동을 입력한다. 특정 읍·면·동에 한정되지 않을 경우는 생략할 수 있다. ⑪는 ⑩에 입력한 지역 이외에 중복으로 해당하는 시·군·구를 입력하는 곳으로 항목리스트의

‘근거자료목록’에 있는 자료를 아카이브 목록에 입력할 경우에 사용한다. 이때 반드시 항목리스트의 ‘근거자료목록’에서 사용한 약호와 동일하게 입력한다. ④는 매체유형을 입력하는 곳으로 사용자의 편의성을 위해 콤보박스로 선택할 수 있도록 구현하였으며 ‘단행본, 팜플릿, 연속간행물, VHS, 오디오테이프, CD-ROM, DVD-ROM, 기타’ 중에서 해당하는 내용을 선택한다. ⑤은 ‘일반 자료, 통계자료, 관광안내, 보고서, 학술자료, 지도, 도록/사진, 기타’ 중에서 해당하는 내용을 선택한다. 그리고 마지막 요소인 ⑥는 실물자료가 존재할 경우에는 ‘첨부’를, 존재하지 않을 경우는 ‘첨부안함’을 선택한다.

이상과 같이 기본정보에서 필요한 메타데이터의 요소들을 정리하여 시스템에 실제로 쓰이는 실례는 〈표 6〉과 같다.

자원의 특성을 기술하는 구조화된 데이터로서 메타데이터는 아카이브에서 다루게 되는 각종 정보자원들을 기술, 조직하는 역할뿐만 아니라 사용자들이 그 자원들을 검색하도록 하며, 각종 부가적인 서비스를 제공하는 기반요소이다. 즉, 디지털 정보에 대한 통로(pathway) 역할을 한다는 것이다. 따라서 디지털 아카이브 시스템

〈표 6〉 멀티미디어 콘텐츠 아카이브 목록 실례

00시 이전이브 목록													
제목	기타 제목	주 제 어	저 편 자	발 행 일	ISBN	발 행 처	총 존 수	지역1			지역2		
								광역단위	시·군· 구	총 존 수	광역단위	시·군· 구	총 존 수
(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
고령문화시대 제 1 역사학	문화사	고령군 대가야박 물관·경북대학교 제1연구소	고령군 대가야박 물관·경북대학교 제1연구소	2008. 03.26	978-89-55 536-602-4 93910	고령군 대가야박 물관·경북대학교 제1연구소	529	1	경상북도 고령군		(o)	(p)	(q)
고지도와 읍지로 보는 영양(박물관아 카이브 제2집)	古地圖와 邑誌로 보는 英陽	고지도, 읍 지, 생활사 활박물관	영양신촌읍 활박물관	2008. 11.28	978-89-91 263-27-7	영양신촌읍 활박물관	176	1	경상북도 영양군		(r)	(s)	(t)
도시체생 디자인으로 승부하자! (서 유럽 4개국 선진도시 디자인연수 보고서)	구로구, 양천구 디자인	구로구, 양천구 디자인	구로구, 양천구 디자인	2008. 05.00		서울특별시 구청 정협의회 서남권	58	1	서울특별시 구로구	서울특별시 양천구	(u)	(v)	(w)
한국과 일본의 새로운 시각	일제강점 기, 한일관 계	다나카 히로 시, 이타기치 류타, 한국학 중앙연구원 한국문화고 류센터 억	2007. 12.27	978-89-92 037-26-6		대한민국	201	1	대한민국		(x)	(y)	(z)
장월중선의 음악세계	The Music World Jang Woljungseon	국립민속극 악원	국립민속극 악원	2004. 12.00		국립민속 극원			전라남도 곡성군		DVD- ROM	일반 자료	첨부
구로문학	문학	구로문인협 회	구로문인협 회	2007. 12.27		구로문인협 회	375	1	서울특별시 구로구	개봉2동		연속간 행물	일반 자료
충주시지	시지	충주시	충주시	2002. 01.10		충청북도 충주시	2734		충청북도 충주시		CD- ROM	일반 자료	첨부

구축에 있어서, 데이터의 호환성을 유지하고 통합 검색 및 상호운용성(interoperability)을 감안해서 메타데이터의 표준을 따르는 것은 필수적이라고 할 수 있다.

#### 4. XML 기반의 디지털 편찬 관리시스템 설계 및 구현

텍스트 및 멀티미디어를 포함한 콘텐츠의 관리는 협의의 의미로 저작권 보안 측면에서 바라볼 수 있으며, 이와 관련된 연구는 DRM(Digital Right Management) 기술들과 같은 한 분야로 현재 활발히 진행되고 있다. 또한 관리만을 고려할 때, 어떤 정보가 어디에 있는지에 관심을 두는 영상 검색 시스템(Image Retrieval System)과 관련된 멀티미디어 자산관리 시스템(Multimedia Asset Management System)이 한 분야를 이루고 있다. 이러한 시스템은 미디어 자체에 제한된 또는 관리만을 위한 시스템이어서 서비스 제공자가 서비스를 고려한 온라인 시스템으로 적용하기에는 불편한 점들이 있다.

특히, 기존의 시스템은 관계형 데이터베이스에 기반을 두고 있는 인터넷 환경의 콘텐츠 관리에 효율적이지 못한 면이 있다. 즉, 계층적 구조와 객체 지향성을 갖는 콘텐츠를 표현하기 위하여 관계형 데이터베이스에서는 많은 수의 테이블 조인(join)이 요구된다.

따라서 본 연구에서는 향토문화 콘텐츠의 객체 지향성과 계층적 구조를 효율적으로 표현함

과 동시에 분산 환경으로 구성되는 서비스 시스템을 계층적 구조로 표현하고 관리하기 위하여 XML DB를 기반으로 한 시스템을 설계하고 이를 기반으로 구현한 디지털 편찬 관리시스템의 예를 보인다.

〈그림 1〉은 서비스가 중심이 되는 디지털 편찬 관리시스템에 대한 전체 구성도이다. 웹 인터페이스는 디지털 콘텐츠 아카이브를 사용자가 언제 어디서든 원하는 정보를 접할 수 있도록 웹으로 접근이 가능하도록 한 것이며 시스템 구조는 다양한 콘텐츠의 특성을 분류하고 독립적으로 관리하기 위해 카테고리형 계층으로 구성하였다.

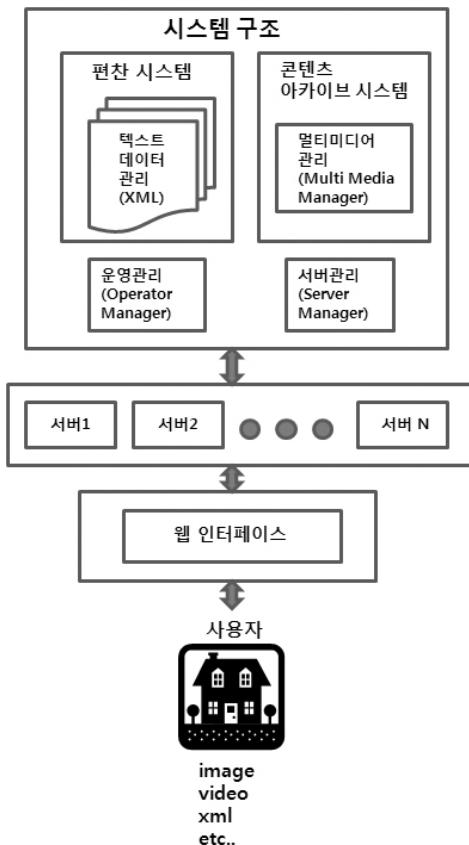
본 연구에서 보이고자 하는 디지털 편찬 관리시스템은 디지털 콘텐츠 데이터를 운영자가 특정 서버에 직접 옮겨서 서비스를 준비하는 기존의 단순한 형태의 것 보다 진보된 것으로서, 분산 환경의 각 서버들에 대한 정보를 기반으로 각각의 다양한 텍스트 및 멀티미디어 콘텐츠들이 해당 서버들 즉, 한 대 이상의 중복된 서버들로 구성되는 각 웹서버, 멀티미디어 서버<sup>2)</sup> 및 텍스트 서버<sup>3)</sup> 등에 적절히 위치하도록 하여 사용자에게 서비스 한다. 여기서의 서비스는 실시간 디지털 편찬 관리와 아카이브의 기능이 모두 실행될 수 있음을 의미한다.

##### 4.1 디지털 편찬 관리시스템의 기능설계

편찬 관리시스템은 업무와 대상 자료의 특성상 디지털 편찬시스템과 디지털 아카이브시스

2) 멀티미디어 서버에서 관리되고 있는 멀티미디어의 종류는 사진(gif, jpg), 음향(swf, mp3), 동영상(avi, flv), VR(mov)로 총 4가지이다.

3) 텍스트 서버에서 관리되고 있는 데이터의 주요파일 형태는 xml이다.



〈그림 1〉 디지털 편찬 관리시스템의 전체 구성도

템이 통합된 형태의 시스템 구축이 목표이므로 업무단계를 기준으로 각 시스템별로 제공해야 할 기능의 설계에 차이가 있다. 〈표 7〉은 디지털 편찬 시스템의 기능 설계 요건을 보여주고 있다.

〈표 7〉에서 제시된 각 기능들은 각기 독립적인 클래스로 구현되었기 때문에 새로운 기능들을 추가하거나 수정하기가 용이하다.

본 연구에서의 디지털 편찬 시스템은 항목조정단계, 집필단계, 교열단계로 크게 세 개의 단계로 정의함을 제안한다. 그리고 집필단계와 교열단계는 각기 의뢰, 검수, 완료 세 가지 상태

가 존재한다. 아래 〈그림 2〉는 편찬시스템의 제작 과정을 정리한 그림이며 〈그림 3〉은 제작 과정을 토대로 본 시스템에서 정의하여 구현한 프로세스의 흐름도이다.

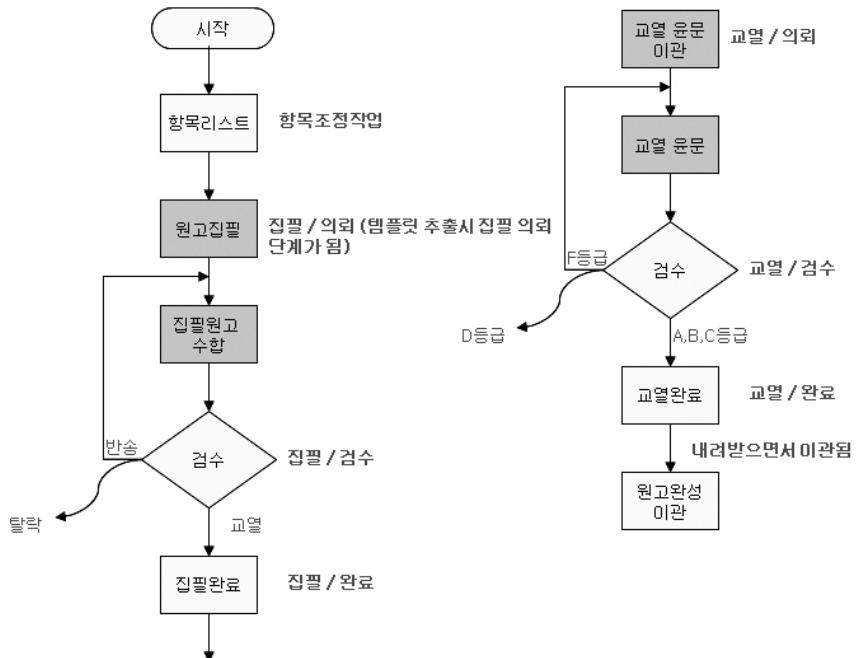
집필의뢰 항목 및 집필자가 확정되면 항목리스트를 만들고 이 리스트를 편찬시스템에 올리게 된다. 이 단계가 '항목조정작업'에 해당하며 이 단계에서는 항목리스트에 있는 항목들의 데이터들을 수정하거나 조정하는 일을 수행하게 된다. 항목조정작업이 끝난 항목들은 템플릿을 자동 추출하여 '집필/의뢰'단계로 넘어가게 된다. 내려진 템플릿들을 이용하여 집필

〈표 7〉 디지털 편찬시스템 기능설계

구분	단계	상태	공정	기능	입력	파일포맷
디지털 편찬시스템	집필	항목 조정 작업	메타데이터 수립	• 파일업로드 기능 • 메타데이터 필드 구성 기능 • 메타데이터 수정 기능 • 지역별 목차 및 분류보기 기능	항목메타데이터	*.xls, *.xml
			원고집필	• 템플릿 내리기 기능 • 원고 및 메타데이터, 상세정보 데이터베이스에 저장하는 기능	원고집필 템플릿	*.hwp
		의뢰	집필원고수합	• 집필의리목록과 집필완료 원고 확인기능 • 원고등록기능 • 집필원고 매수 산정기능	1) 집필원고 2) 원고별 원고지매수, 집필비목록	*.hwp, *.xml *.xls
			원고검수	• 템플릿 준수여부 확인 및 형식오류체크 기능 • 반송, 완료, 탈락 등급 부여기능 • 반송의 경우 집필방향 의견 작성 기능 • 탈락 경우 탈락 사유작성기능 • 교열/의뢰 단계 전환 기능	1) 수합된 집필원고 2) 검수대상 목록 및 원고	*.hwp, *.xml *.xls
	교열	의뢰	교열윤문이관	• 교열검수 작업자 지정 기능	1) 교열윤문대상원고 2) 항목별 교열윤문 작업자목록	*.hwp, *.xml *.xls
			교열윤문	• 원고 교열윤문 기능	교열윤문 완료원고	*.hwp, *.xml
		검수	교열검수	• 교열윤문 결과 확인 및 등급부여 기능	1) 교열윤문 완료원고 2) 등급목록	*.hwp, *.xml *.xls
		완료	교열완료 및 원고완성이관	• 완성원고내려받기 기능	원고완성 항목	*.hwp, *.xml



〈그림 2〉 디지털 편찬시스템 제작 과정



〈그림 3〉 디지털 편찬시스템 프로세스 흐름도

자들은 집필을 수행하게 되고 집필이 끝난 원고들은 집필원고 수합이 이루어지며 수합된 원고들은 편찬 시스템에 다시 올려지게 된다. 이 올려진 원고들을 검수('집필/검수': 1차 검수)를 통해 부적절한 원고는 반송 또는 탈락되고 적절한 원고는 다음 단계인 '집필/완료'단계로 이동된다. 집필/완료가 된 항목들은 자동으로 '교열/의뢰'단계로 변환되며 교열 윤문을 수행하게 된다. 교열 윤문을 거친 항목들은 '교열/검수'단계에서 다시 2차 검수작업을 수행하게 된다. 검수작업의 결과로 A, B, C, D, F의 판정 등급을 부여받게 되며 A, B, C 등급을 받은 항목들은 '교열/완료'상태로 확정되고 이 상태의 원고들을 시스템에서 내려 받게 됨으로써 최종 단계인 '원고완성/이관'이 이루어지게 된다. 판정등급 D를 받은 항목들은 탈락이 되며 F를 받

은 항목들은 다시 '교열/의뢰'단계로 보내지게 된다. 이 모든 과정은 디지털 편찬시스템에서 자동으로 관리되고 수행되게 된다.

다음으로 디지털 아카이브 시스템의 기능 및 구조에 대해 살펴보도록 하겠다.

디지털 아카이브 시스템에 필요한 데이터 형식파일은 위에서 제시한 아카이브 메타데이터 뿐만 아니라 디지털 콘텐츠의 매체유형, 지역 코드가 필요하다.

현재 구현한 디지털 아카이브 시스템에 쓰인 매체유형과 지역코드는 〈표 8, 9〉와 같다.

그리고 디지털 아카이브의 세부기능을 요약하면 다음과 같다.

- 향토문화 관련 새로운 아카이브 자료를 데이터베이스에 올리고 검수할 수 있는 기능

〈표 8〉 디지털 아카이브 매체유형

매체유형	
항문자료	GC
단행본	PO
연속간행물	PS
팸플릿	PP
VHS	RV
오디오테이프	RA
CD-ROM	RC
DVD-ROM	RD
기타	PR

〈표 9〉 디지털 아카이브 서울시에 관한 지역코드

지역(시도)	지역(군)	기호
서울특별시	강남구	001
서울특별시	강동구	002
서울특별시	강북구	003
서울특별시	강서구	004
서울특별시	관악구	005
서울특별시	광진구	006
서울특별시	구로구	007
서울특별시	금천구	008
서울특별시	노원구	009
서울특별시	도봉구	010
서울특별시	동대문구	011
서울특별시	동작구	012
서울특별시	마포구	013
서울특별시	서대문구	014
서울특별시	서초구	015
서울특별시	성동구	016
서울특별시	성북구	017
서울특별시	송파구	018
서울특별시	양천구	019
서울특별시	영등포구	020
서울특별시	용산구	021
서울특별시	은평구	022
서울특별시	종로구	023
서울특별시	종구	024
서울특별시	중랑구	025

- 향토자료 관련 자료들의 형식을 자동 변환하여 새로운 데이터베이스에 적재시키

### 는 기능

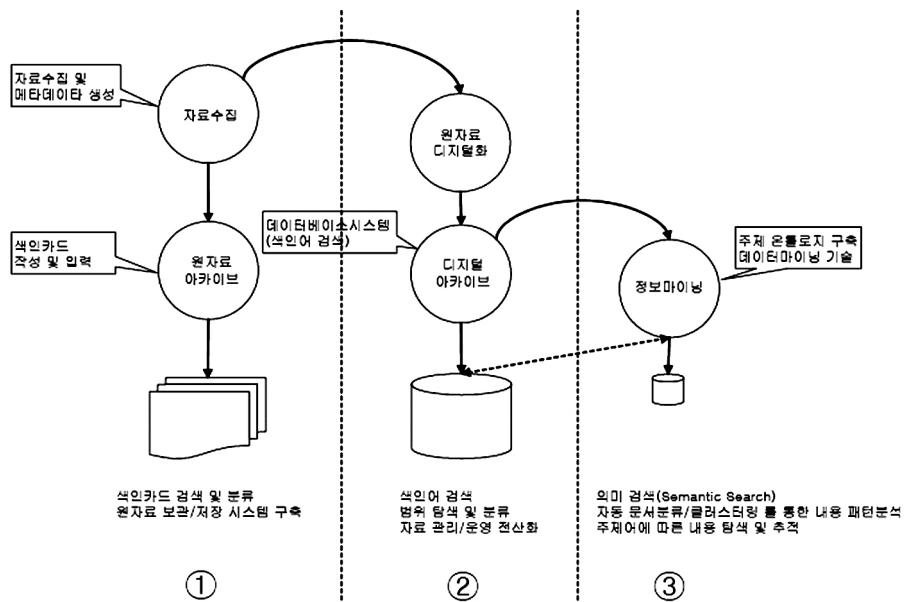
- 아카이브 자료들의 청구기호를 자동으로 부여하는 기능
- 아카이브 자료들을 엑셀 및 XML 파일로 전환해서 볼 수 있게 하는 기능
- 목차 및 표지 등을 바로 검수할 수 있는 기능
- 향토문화자료 관련 원고 데이터베이스(DB)에서 관련된 자료들을 바로 검색 및 확인하고 관련 자료들과 연결하는 기능
- 아카이브 자료들 사이의 연관 있는 자료끼리 연결할 수 있는 기능

전체적인 기능별 구조도를 살펴보면 〈그림 4〉와 같다.

각 단계별 세부내용을 살펴보면, 단계 ①은 원자료아카이브 구축 단계로 자료수집 및 메타데이터 작성 단계를 의미한다. 본 연구에서는 〈차례3〉에서 설명한 메타데이터를 기술하는 부분이 바로 이 단계에 속하게 된다. 단계 ②는 디지털 아카이브 구축 단계로 자료를 컴퓨터가 인식할 수 있도록 데이터를 디지털화한다. 그리고 마지막 단계 ③은 아카이브 활용단계로 정보마이닝을 통해서 사용자들에게 원하는 정보 및 연관 정보들을 빠르고 정확하게 제공함으로써 정보의 확대 및 재생산 과정까지를 포함함을 의미한다.

## 4.2 디지털 편찬 관리시스템 구현

디지털 편찬 관리시스템은 앞서 설명한 바와 같이 필수메타데이터, 편찬시스템 및 검색기, 디지털 아카이브로 구분된다. 디지털 편찬 관



〈그림 4〉 디지털 아카이브 시스템의 기능별 구조도 및 인터페이스 구조도

리시스템은 Microsoft Visual C#으로 구현되었으며 디지털 콘텐츠를 색인하고 그 결과를 원하는 형식의 메타데이터로 출력한다. 본 시스템이 정상적으로 설치되어 실행되기 위해서는, Net Framework 2.0이 설치되어 있어야 하며, 한글 파일을 접근하기 위한 한글 컨트롤러의 정상적인 동작을 위해 C 드라이브에 '한글 2005' (Version 6.5.0.825)가 설치되어 있어야만 한다. 지금부터 디지털 편찬 관리시스템의 구체적인 모습과 주요 기능 부분들을 살펴보도록 하겠다. 우선 본 연구에서 구현한 디지털 편찬 시스템은 향토문화 항목들의 구체적인 세부 요소들을 보여주는 항목관리 부분과 항목에 대한 원고들의 상세부분을 보여주는 원고관리 부분을 텁으로 나눠서 구현하였다.

〈그림 5〉는 충주지역의 향토문화와 관련 항목들의 메타데이터를 보여주고 있다. 항목관리에

서는 기본적으로 항목의 ID, 구분, 한글항목명, 한자항목명, 대분류, 중분류, 소분류, 분야1, 유형1, 하위유형을 보여주며 또한 사용자가 원하는 메타데이터 필드를 설정하여 보여줄 수 있는 기능도 추가하였다. 항목관리 부분에서의 주요기능으로는 항목들의 목록을 내려 받을 수 있는 기능과 템플릿을 내리는 기능이 있다. 항목들의 목록은 xml 및 엑셀파일 형식으로 내려 받을 수 있도록 구현하였고 템플릿 내리기 기능은 hwp파일 형식으로 각 항목에 해당하는 메타데이터에 맞추어 내려받을 수 있다. 이렇게 내려받은 원고는 집필자에게 보내지게 되며 이때의 단계 및 상태는 '집필/의뢰'가 된다. 〈그림 6〉는 원고관리 부분이며 원고관리에서는 항목ID, 한글항목명, 작업단계, 차수, 상태정보, 검수결과, 등급, 검수의견, 검수자, 최종작업일과 같은 디지털 편찬 작업에서의 관리 및 원고



〈그림 5〉 디지털 편찬 시스템의 항목관리 부분

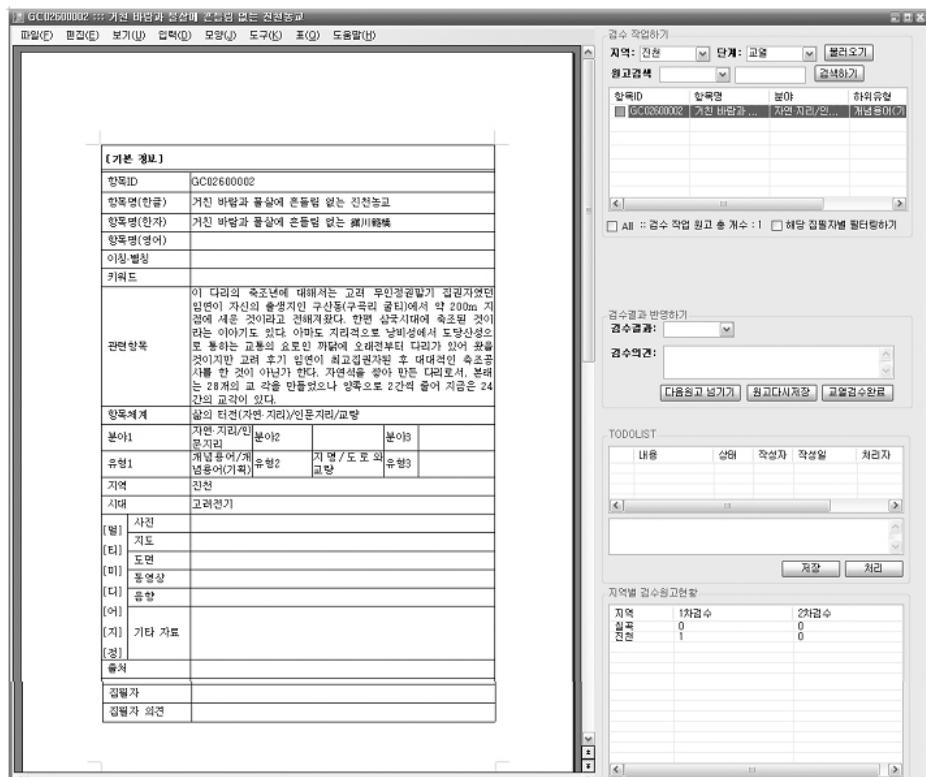


〈그림 6〉 디지털 편찬 시스템의 원고관리 부분

상태정보를 볼 수 있도록 하였다. <그림 6>의 메뉴에서 보이는 [검수 브라우저] 버튼을 클릭하면 작업 지역에 해당하는 아래와 같이 원고를 검수 할 수 있는 창이 뜬다. 아래 <그림 7>과 같이 지역별로 집필, 교열, 원고완성에 해당하는 원고들이 리스트에 뜨게 되고 각 원고들을 선택할 때마다 창의 왼쪽에 해당 원고를 볼 수 있도록 하여 작업자가 실시간으로 원고들을 수정 및 삭제와 같은 교정 작업이 가능하도록 구현하였다. 그리고 지역별, 단계별뿐만 아니라 원하는 집필자별로도 필터링이 가능하도록 하였다. 따라서 작업자는 원하

는 항목 원고의 실시간 검수가 가능하며 검수가 완료되면 자동으로 다음 단계로 상태가 변환된다.(예: '집필/검수'에서 '집필/완료(교열/의뢰)'로 넘어감 혹은 '교열/검수'에서 '교열/완료(원고완성/이관)'로 넘어감), <그림 8>은 디지털 아카이브 시스템의 웹 인터페이스를 나타낸 것으로 사용자가 웹으로 언제든지 편리하게 이용할 수 있도록 설계되었다.

<그림 9>는 디지털 아카이브 시스템<sup>4)</sup>의 인터페이스를 나타낸 것으로 필요한 파일 목록으로는 XML 형식의 아카이브 메타데이터 및 기타 목록과 목차 및 표지 사진 자료들인 jpg파일



<그림 7> 디지털 편찬 시스템의 검수 브라우저

4) 프로그램 직접 사용 및 자세한 매뉴얼은 tmt21@aks.ac.kr로 요청 가능함



〈그림 8〉 디지털 아카이브 시스템의 웹 인터페이스5)



〈그림 9〉 디지털 아카이브 시스템 인터페이스

5) 한국학중앙연구원의 향토문화아카이브는 한국향토문화대전 사업과 관련하여 수집, 생산된 한국학정보센터 소장 자료 1,600여건, 전국 시군단위별 5,000여건의 향토자료 목록 및 서지정보, 기본정보를 포함하는 향토문화자료와 28,700여건의 고문서 목록과 메타정보 뿐만 아니라 7개 가문의 한글 간찰 2,600여건을 포함하고 있다.

그리고 『한국향토문화전자대전』의 디지털 편찬 시스템들의 데이터베이스로 이루어진다. 아카이브 및 편찬 시스템에서 저장된 모든 자료들은 〈그림 8〉의 웹 인터페이스로 연결가능하도록 설계하였기 때문에 웹이 연결된 곳이라면 언제 어디서나 사용자가 쉽게 접근이 가능하도록 하였다.

본 연구에서 정한 기본메타데이터와 기타메타데이터는 XML을 기반으로 하였기 때문에 고정된 것이 아니라 변경이 필요할 시에는 언제든지 재정의를 하여 사용될 수 있다. 따라서 필자와 같은 개발자가 프로그램 코드의 수정 없이도 쉽게 변화된 데이터를 적용 시킬 수 있다. 아카이브 자료들의 XML 구조는 다음과 같다.

XML 기반 방식은 〈그림 10〉에서 보듯이 많은 태그들을 포함하고 있기 때문에 데이터의 용량면에서 다소 성능의 저하를 불러일으킬 수 있으며 XML문서의 저장 관리 측면에서 XML 문서의 효율적인 변경과 검색을 동시에 지원하기는 쉽지 않은 단점이 있다. 또한 XML을 잘 모르는 사용자들은 작업하기에 문제가 있을 수 있다. 하지만 표준화된 DTD<sup>6)</sup>를 통하여 모든 분야에서 사용이 가능하므로 교환 및 재사용에 대한 활용도가 높을 뿐만 아니라 사용자가 규칙에 맞추어 사용하면 컴퓨터나 사람이 모두 해석할 수 있는 메타 언어가 될 수 있다.

디지털 아카이브 시스템은 디지털 편찬 시스템의 자료들인 여러 원고항목들과도 연결하여 검색할 수 있도록 기능을 부여하였다. 이 기능

```
<?xml version="1.0" encoding="utf-8" standalone="yes"?>↓
<향토자료목록>↓
  <향토자료 ID="">↓
    <자료정보>↓
      <자료명>↓
        <한글자료명 />↓
        <외래어자료명 />↓
      </자료명>↓
      <저편자 />↓
      <발행일 />↓
      <발행처 />↓
      <첨부자료 />↓
      <청구기호 />↓
    </자료정보>↓
    <내용정보>↓
      <대분류 />↓
      <중분류 />↓
      <매체유형 />↓
      <세부유형 />↓
      <주제어 />↓
      <개요 />↓
      <ISBN-ISSN />↓
      <지역 />↓
    </내용정보>↓
    <형태정보>↓
      <권수>/<권수>↓
      <쪽수>/<쪽수>↓
      <판형>/<판형>↓
    </형태정보>↓
    <수집정보>↓
      <출처 />↓
    </수집정보>↓
    <기타정보>↓
      <기타 />↓
    </기타정보>↓
  </향토자료>↓
</향토자료목록>
```

〈그림 10〉 아카이브 자료들의 XML 구조

은 디지털 아카이브 시스템의 기능별 구조도의 제③단계에 속하는 정보마이닝 단계에 속한다. 우리의 시스템은 메타데이터를 이용한 메타검색이 가능하다. 이는 한국향토문화대전의 성격상 문서나 문헌의 편차가 많고 서로 다른 분류 방식으로 정리되어있는 것이 많기 때문에 메타검색이 필수적이라고 할 수 있겠다.

디지털 편찬 관리시스템은 크리스탈(KRISTAL)를 정보서버엔진으로 사용하였으며 이를

6) 문서의 정보 모델을 구체적으로 서술한 것을 DTD(Data Type Definition: 문서 형태 정의)라 한다. DTD에는 문서를 구성하는 정보 요소, 정보요소의 구조와 특성 등 문서의 형태를 구조화하여 정의한 것으로, 문서의 조직 규칙을 정의한 것이다. 따라서 DTD는 문서 작성을 위한 템플릿(Template)의 역할을 한다.

통해 향토문화 데이터와 XML 구조문서를 하나의 데이터베이스에서 저장, 관리하고 검색할 수 있도록 설계 및 개발하였다.

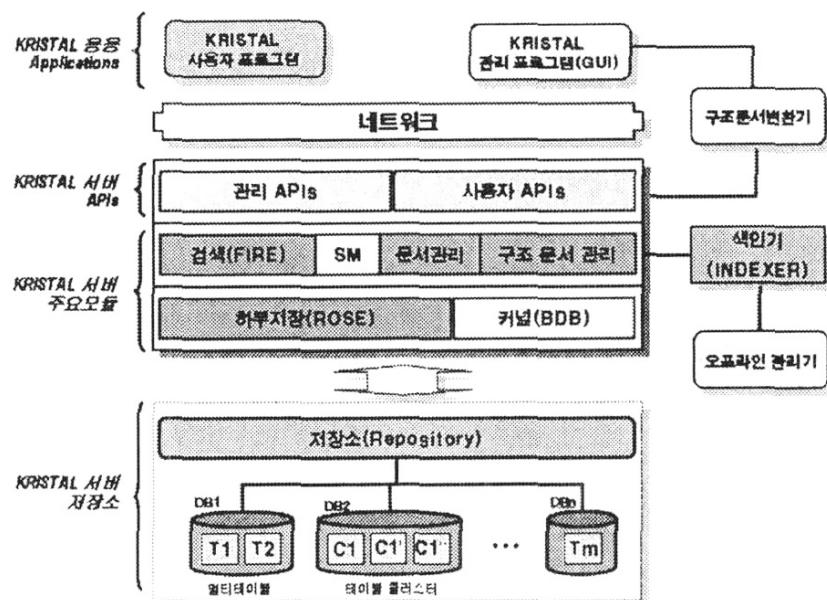
〈그림 11〉은 정보 검색 관리 시스템 크리스탈(KRISTAL)의 전체 구조도이다. 본 연구에서는 향토문화 텍스트를 효율적으로 저장 관리 및 검색을 수행할 수 있도록 정보검색관리 시스템의 기능을 확장하였다.

향토문화 텍스트는 계층적인 구조로 구성되어 있는 XML문서로 작성되어 있으며, 이러한 계층적인 구조를 유지하면서 저장, 관리 및 검색하기 위해 크리스탈에서 제공하는 XML문서 단편화 방법을 이용한다. 아래 〈그림 12〉는 XML문서의 단편화 및 적재 과정을 보여준다.

크리스탈에 XML 문서를 적재하기 위해서

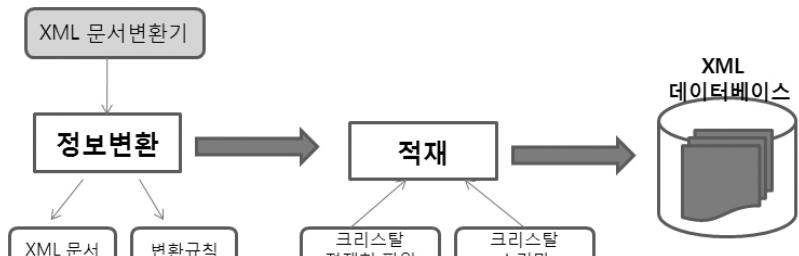
는 원본 XML문서와 XML 문서를 변환하기 위한 변환 규칙을 정의해야 한다. 서비스할 XML 문서에 대한 변환규칙을 정의 한 뒤, XML 문서 변환기를 이용하여 단편화 작업을 수행하는 1단계를 거친 후, 스키마를 이용하여 결과로 생성된 단편화된 데이터를 저장 시스템으로 적재하는 2단계 작업을 수행한다. XML 문서 변환기는 문서의 계층정보를 유지하면서 문서의 단편화 시킨다. 이때 사용자가 입력한 규칙정보를 이용하여 XML 문서를 구조적 특징을 포함한 단편화된 XML 문서 요소들로 변형시킨다.

지금까지 향토문화와 관련된 텍스트자료를 포함한 여러 종류의 디지털 콘텐츠들을 디지털 편찬 관리시스템으로 어떻게 표준화 하고 관리



〈그림 11〉 KRISTAL의 전체 구조<sup>7)</sup>

7) 크리스탈은 한국과학기술정보원(KISTI)이 개발한 검색엔진으로 기관간 협약을 통해 업그레이드가 추진되었고 그 결과 한적자료의 검색에 강점을 보여 역사분야의 많은 기관들이 안정적으로 적용하고 있다.



〈그림 12〉 XML 문서 적재 과정

및 활용 하는지에 대해 주요 기능별로 살펴보았다. 본 연구에서 제시한 디지털 편찬 관리시스템은 단지 자료를 저장하고 관리하는 기능을 넘어 장시간에 걸쳐 여러 차례의 교정 작업이 필수적인 편찬 작업을 자동화하는 동시에 실시간으로 디지털 편찬 서비스를 제공할 뿐만 아니라 하나의 정보를 바탕으로 유추하고 검색하여 더 넓은 지식의 공유로 재사용 및 활용할 수 있도록 하였으며 그에 따른 여러 디지털 콘텐츠 서비스의 제공도 가능하도록 하였다.

## 5. 맺음말

디지털 시대가 도래하면서 문화콘텐츠를 기반으로 한 디지털 서비스 및 아카이브의 활용은 현대 시대 문화산업의 주축을 이루고 있다. 이에 따라 디지털 콘텐츠의 종류도 매우 다양화되고 있으며 그 영역은 시간이 지나고 기술이 발전함에 따라 더욱 넓어질 것이다. 본 연구에서는 이러한 많은 문화콘텐츠 중 향토문화 자료들을 이용하여 자원의 디지털화와 공유에 있어서 분류체계, 콘텐츠 개발, 지역연계시스템을 목표로 한 체계화된 데이터베이스를 기반으로 향토문화 콘텐츠가 좀 더 효율적이고 적극

적으로 활용 될 수 있도록 디지털 편찬 관리시스템을 설계하고 구현하였다. 편찬 관리시스템의 목적은 향토문화 디지털 콘텐츠를 제작하여 지역문화 관련 인적자원을 교육하고 조직하여 전국의 지역문화자료를 총체적으로 발굴·분석하는 것이다. 그리고 편찬 관리시스템을 통한 시·군·구별 디지털 향토문화대전을 통합·구축하여 지역문화에 대한 총체적인 정보를 제공함으로써 지식기반사회의 토대를 마련하고 나아가서는 정부의 지역문화균형발전과 지역경제 활성화에 기여하려는 것이다. 본 시스템에서 제안한 프로세스 정의 및 알고리즘과 응용프로그램들은 지금까지 수동적으로 관리하면서 겪은 많은 시행착오들의 문제점을 분석하고 수정하여 만든 결과물이다. 특히, 기존 관계형 데이터베이스의 사용을 지양하고 텍스트 및 멀티미디어 콘텐츠뿐만 아니라 계층적 구조를 갖는 시스템을 효율적으로 표현하기 위하여 객체지향 데이터베이스 특징을 갖고 있는 XML을 시스템 관리 데이터베이스로 사용하였다.

본 연구에서 구현한 디지털 편찬 관리시스템의 특징은 다음의 여섯 가지로 요약될 수 있다.

- (1) 디지털 편찬을 위한 제작 프로세스 및 공정 단계를 6가지 상태 및 단계로 정의하고 표준화

하여 실시간 디지털 편찬 및 관리가 가능한 시스템을 구축하였다. (2) 향토문화 자료는 문서(텍스트자료), 음향, 사진, 동영상, VR 등으로 구성된 텍스트 및 멀티미디어 데이터베이스 시스템으로 구축되었으며, XML을 기반으로 구현하였기 때문에 데이터의 이식, 공유 및 호환이 용이하다. (3) 아카이브는 웹 인터페이스로도 설계하여 웹이 연결된 곳이라면 언제 어디서나 이용할 수 있다. (4) 원자료아카이브와 디지털아카이브가 한 시스템으로 운영되고 있으며 이 아카이브 시스템을 이용하여 다양한 성과물과 응용프로그램을 만들 수 있다. (5) 디지털 편찬 관리시스템의 구현에 쓰인 여러 가지 응용프로그램들은 각각 독립적인 클래스로 구현하였기 때문에 새로운 기능들을 추가하거나 수정이 필요한 부분이 생기더라도 큰 틀의 변화없이 빠른 시간에 바로 적용시킬 수 있다. (6) 시스템을 통해 구축된 자료는 지역문화 포털사이트의 개설이 가능하며 이를 통해 일반연구자 및 학술연구자에 이르는 광범위한 이용자

의 정보수요를 만족시킬 수 있다.

앞으로의 연구방향은 지금까지의 연구를 바탕으로 다양한 정보마이닝 기술과 DRM(Digital Rights Management) 기술들을 디지털 편찬 관리시스템에 적용하고, XML 문서의 효율적인 변경과 검색이 동시에 지원되는 편찬 및 저장관리 시스템을 구현하는 방안이다. DRM은 미디어 자체에 보안을 설정하기 때문에 본 연구에서 구현한 디지털 편찬 관리시스템과 독립적으로 설계가 가능할 것이다.

또한, XML 문서의 저장 및 관리는 DTD의 타입별 스키마에 대한 고려와 페이징 기법의 도입을 생각해 볼 수 있을 것이다. 그리고 저장된 문서에 대한 온라인 문서 개선에 대한 속도 향상과 문서 관리 목적으로 사용되는 특수 문자에 대한 색인지원 등 다양한 요구사항에 대한 지원을 통해 새로운 디지털 편찬 관리시스템이 정보사회의 요구에 부응하고 미래형 관리시스템의 새로운 방향을 선도해 나갈 수 있기 를 기대한다.

## 참 고 문 헌

김선미. 1997. Dublin Core 메타데이터 형식 확장 설계에 관한 연구. 이화여대 석사논문.

김주관. 생활사이아카이브구축의 의미와 방법. 『지방사와 지방문화』, 8: 1.

김희정. 2003. 전자저널 아카이빙을 위한 OAIS 참조모형의 적용방안에 관한 연구. 연세대학교 석사논문.

문정희. 2008. 문화원의 향토자료 관리 실태분석. 『정보관리학회지』, 25(3): 297-320.

박은정. 2005. 전자문서관리와 디지털아카이브: 성장과 전망. 『한국기록관리학회 학술논문집』, 23-32.

서은경. 2004. 디지털아카이브의 연구적 보존을 위한 개념적 모형 설계에 관한 연구. 『한국문화정보학회지』, 38(1): 13-34.

역사문화학회. 2005. 『한국향토문화전자대전 선

행조사 지침서』. 한국학중앙연구원 한국 학정보센터 문화콘텐츠편찬연구실, 217-249.

윤소영. 2009. 수집사료 관리시스템 개발에 관한 연구. 『정보관리학회지』, 26(1): 371-387.

이선수. 2006. 한국형 디지털 콘텐츠를 구축 방안 연구: 이미지와 동영상 클립을 중심으로. 『인문콘텐츠학회』, 인문콘텐츠, 8: 357-376.

최원태. 2001. A Study on the Status and Component of Digital Archive. 『한국 문헌정보학회지』, 35(2).

최윤수. 2002. 고문서를 위한 문서모델 제안 및 편찬관리 시스템 개발. 『한국인터넷정보학회 2002 추계학술발표대회 논문집』, 3(2): 364-367.

한국과학기술정보연구원. 2002. 『정보검색관리 시스템 KRISTAL-2000 Programmer's Manual for C++ User』, 한국과학기술 정보연구원 정보시스템연구실.

『한국향토문화전자대전』『향토문화아카이브 구축 사업』 관련 문서.

한국학중앙연구원 · 제주대 탐라문화연구소. 2006. 『디지털제주시문화대전 원고집필 방법 설명회』 자료.

한승희. 『메타데이터간의 상호호환성을 위한 RDF 구조 연구』.

한국학중앙연구원. 2009. 『디지털충주문화대전』.

한국학중앙연구원. 2009. 『디지털진천문화대전』.

한국학중앙연구원. 2005. 『디지털한국향토문화 전자대전 작업지침서』.

황상규 · 윤세진 · 오경록. 2001. 디지털 콘텐츠를 위한 메타데이터에 관한 연구. 『제8회 한국정보관리학회 학술대회논문집』.

홍창용. 2003. 지상파 TV 방송의 디지털화에 따른 방송 영상 디지털 아카이브 시스템 방안 연구, 서강대 언론대학원 석사학위논문.

20세기민중생활사연구단. 2003-2006. 『민중생활사연구 학술연구보고서』.

Baeza-Yates, Ribeiro-Neto. 2001. 최신 정보 검색론. 『홍릉과학출판사』.

Buneman, P. 2004. "Archiving scientific data." *ACM Trans. Database Syst.*, 29(1): 2-2.

Caplan, Priscila 지음, 오동근 옮김. 2004. 메타 데이터의 형식과 구조. 『태일사』.

Dublin Core Metadata Initiative. <http://dublincore.org>.

Fernstrom, Mikael and Eoin Brazil. 2001. Sonic Browsing: An Auditory Tool for Multi-media Asset Management, in Proc. 2001 International Conference on Auditory Display, 132-135.

Wang, F. 2003. *Publishing and querying the histories of archived relational databases in XML*, WISE.

Wang, Fusheng. 2008. Temporal queries and version management in XML-based document archives, *Data & Knowledge Engineering*, 65, 304-324.

Zhou, Xiaobo, R. Luling, and Li Xie. 2000. "Solving a media mapping problem in a hierarchical server network with parallel simulated annealing," in Proc. *Parallel Processing*, p.115.