

# 개화기 방한 서양인 기록물의 디지털 아카이브 구축에 관한 연구\*

## Developing Digital Archives from the Records of Westerners who visited Korea during the Enlightenment Period of Chosun

정희선 (Heesun Chung)\*\*, 김희순 (Heesoon Kim)\*\*\*  
송현숙 (Hyun-Sook Song)\*\*\*\*, 이명희 (Myeong-Hee Lee)\*\*\*\*\*

### 목 차

- |                                       |                                   |
|---------------------------------------|-----------------------------------|
| 1. 서론                                 | 4. 개화기 방한 서양인 기록물의<br>디지털 아카이브 구축 |
| 2. 선행 연구                              | 5. 결론 및 제언                        |
| 3. 개화기 방한 서양인 기록물의<br>디지털 아카이브 구축의 의의 |                                   |

### 초 록

본 연구는 개화기 방한 서양인 기록물의 디지털 아카이브를 구축하여 지역문화콘텐츠로 활용하기 위하여 수행되었다. 수집대상 자료는 서양인 기록물 11권이고, 주제분류체계로부터 10개의 대주제와 120개의 소주제를 추출하였다. 37개 메타데이터 항목에 맞추어 텍스트 내용의 분석과 입력자료 유형을 분류하고 엑셀로 된 데이터베이스를 구축하였다. 그리고 디지털 아카이브 시스템 모델과 5개의 메뉴로 구성된 홈페이지 구축(안)을 제시하였다. 추후연구를 위해 아카이브 구축을 위한 데이터의 지속적인 발굴과 아카이브 확대 방안, 아카이브 시스템간의 활발한 연계, 시스템 표준화, 시스템 호환성과 사용자 편의성을 고려한 시스템 설계를 제안하였다.

### ABSTRACT

This study was conducted to create a digital archive for local cultural contents compiled from the records of westerners who visited Korea during the Enlightenment Period of Chosun. The compiled information were gathered from 11 records, and 10 main subjects and 120 sub-subjects were derived through the subject classification scheme. Item analysis was conducted through 37 metadata, and input data types were classified and databased in Excel. Finally, a model of the digital archive system was simulated, and a webpage consisting of five menus was presented. Suggestions for future research were extensive aggregation of new data for archive expansion, active connections between archive systems, standardization of systems, and improved system design for compatibility and user-friendliness.

키워드: 서양인 기록물, 디지털 아카이브, 지역문화콘텐츠, 분류체계, 메타데이터  
Records of Westerners, Digital Archives, Local Cultural Contents,  
Classification Scheme, Metadata

\* 이 논문은 2014년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임  
(NRF-2014S1A5A2A03064904).

\*\* 상명대학교 지리학과 교수(hchung@smu.ac.kr) (제1저자)

\*\*\* 고려대학교 지리교육과 강사(pritikim@gmail.com) (공동저자)

\*\*\*\* 충신대학교 역사교육과 강사(shs454@naver.com) (공동저자)

\*\*\*\*\* 상명대학교 문헌정보학과 교수(mehelee@smu.ac.kr) (교신저자)

논문접수일자: 2015년 7월 16일 최초심사일자: 2015년 7월 21일 게재확정일자: 2015년 8월 3일  
한국문헌정보학회지, 49(3): 135-154, 2015. [http://dx.doi.org/10.4275/KSLIS.2015.49.3.135]

## 1. 서론

### 1.1 연구의 필요성

조선시대의 일반적인 대외정책은 쇄국정책을 취하여 외국인들의 출입이 제한되었지만 개화기에 미국, 영국, 프랑스를 비롯한 다양한 국가의 외교관, 선교사, 군인, 상인 등이 정치적, 경제적, 종교적, 학술적 목적으로 조선에 입국하여 수개월 내지 수년간 동안 체류하며 생활하였다. 그들 중의 대부분은 일정한 시간이 지난 후 본국으로 귀국하여 서구사회에 거의 알려지지 않았던 변방국가인 조선에서의 삶을 기록으로 남겼다. 그들의 동양에 대한 호기심어린 실상을 견문록, 회고록, 기행문 등의 저널리즘의 형태로 저술하거나 조선에 대한 학술적 연구논문으로 작성한 경우도 있어서 동아시아의 끝에 존재하는 조선이라는 국가에 대한 실상을 서구사회에 다양한 방법으로 소개하였다.

서양인들이 남긴 기록물은 대개 1876년 개항 전후부터 일본의 강제 합병 시기에 저술된 것으로서, 19세기 중반에서 20세기 초 서양인들에 의해 조선의 지리적 특성과 생활상을 견문록 등의 형태로 남긴 작품들은 개화기 조선시대 생활사 연구를 위해 매우 중요한 가치를 지니고 있다. 특히 서민 일상사를 중심으로 한 당시 조선의 사회·문화·제도 등을 생생하게 기술할 뿐 아니라 기록과 함께 제공되는 사진자료나 인물화, 사실화 등의 시각적 자료의 가치는 국내에서 찾아보기 드문 귀중한 사료적 가치를 가지고 있다. 이들의 활동 및 기록에 대한 자료의 수집과 분석은 체류자 개인의 신분상 기록물로서의 개념을 넘어 우리나라 근대 역사를

재조명할 수 있는 중요한 사료로 평가될 수 있을 것이다.

지금까지 개화기 조선 체류 서양인의 기록물에 대한 분석은 주로 역사학적으로 이루어졌는데 대부분의 연구는 서양인 입장에서 조선 사회의 제도적·사상적 특징을 살펴보고 서양인의 조선 인식에 투사된 타자화의 양상과 오리엔탈리즘의 요소들을 찾거나 풍속, 신앙, 사회구조 등과 같은 특정 주제 정보를 중심으로 분석하는데 집중되어 있다(박재영 2007). 따라서 기존연구에서는 서양인의 기록 속에 담긴 지역정보를 수집하여 개화기의 도시, 촌락, 지역, 경관, 장소 등에 대해 역사적·지리적인 관점에서 접근한 경우는 거의 찾아보기 어렵다는 점에서 본 연구의 필요성이 제기되고 있다.

이와 같은 배경에서 본 연구에서는 개화기 조선을 방문한 서양인 기록물의 텍스트를 분석하여 자료의 속성 유형을 구조화하고, 지역정보와 주제정보를 추출하여 디지털 아카이브 구축을 시도할 것이다. 이를 통해 지역문화콘텐츠 자료로서의 활용에 기여하며, 관련 정보를 저자 생애사, 시간적·공간적 측면에서 비교·종합하게 되면 역사·지리적 가치와 함께 디지털콘텐츠로서의 가치도 부가할 수 있을 것이다. 또한, 지리정보를 바탕으로 디지털 지도기반 정보 서비스 플랫폼을 작성하고 해당 정보를 시각화시키는 후속연구의 기초연구로 사용하고자 한다.

### 1.2 연구내용과 방법

본 연구에서 개화기 서양인의 조선 체류와 여행에 대한 체험을 서술한 기록물의 디지털 아카이브를 구축하고자 하는 이유는 19세기에서 20세

기 초 한반도의 자연지리적·인문지리적 정보를 분류하여 과거의 지리를 복원하고 역사지리적 측면에서 지역에 관한 이해를 심화시키며 지역문화 콘텐츠로서 활용하기 위한 것이다. 연구의 시공간적 범위는 서양인의 한반도 출입과 국내 행보가 비교적 자유스러웠던 1870년대에서 1910년대까지의 개화기 한반도로 한정하였다.

구체적인 연구내용은 서양인 기록물의 디지털 아카이브 구축을 위한 속성정보 파악 및 분류체계 작성, 메타데이터의 구조화, 지역정보(자연지리·인문지리 정보)의 추출과 분석, 기초적인 데이터베이스 구축이다. 그리고 본 연구를 통해 분석·개발된 기준을 근거로 하여 향후 데이터베이스의 관리와 서비스를 위한 모델로 관리시스템과 홈페이지 구성안을 제시하였다.

첫째, 관련 정보를 효율적으로 검색하고 이용할 수 있도록 서양인 기록물의 샘플 텍스트 분석을 통해 속성과 의미 관계를 고려하여 개화기 방한 서양인 기록물 분류체계를 작성한다. 연구방법으로는 연구대상인 11권의 기록물 분석과 한국향토문화전자대전, 조선왕조실록 분류체계를 참고하여 서양인 기록물 정보의 속성 유형에 맞게 수정·보완한다.

둘째, 문화유산정보의 특성과 메타데이터의 필요성에 대하여 문헌과 인터넷을 통해 조사하고, 서양인 기록물 정보의 데이터베이스 구축에 적합한 최소 수준의 메타데이터 항목을 구성한다. 구체적인 연구방법으로 더블린 코어(Dublin Core) 메타데이터 항목을 분석하여 서양인 기록물 정보에 적합한 데이터베이스와 저자정보 데이터베이스 구축에 필요한 메타데이터 항목을 추출하고 이를 바탕으로 워크시트 폼을 개발한다.

셋째, 역사지리적 복원을 통한 지역연구를 위해 서양인 기록물에 나타난 ① 지형, 토양, 기후, 생태 등과 관련된 자연지리 정보, ② 인구, 행정, 도시·촌락의 기능, 분포, 연계 등과 관련된 지역적 정보, ③ 정치, 경제, 사회, 문화 등의 계통적 주제와 관련된 인문지리 정보를 추출·분석하여 자연지리 정보, 지역적 정보, 인문지리 정보를 종합하는 데이터베이스를 구축한다. 구체적인 연구방법으로는 연구자가 텍스트를 직접 읽고 내용을 분석한 후 관련 정보를 추출하여 엑셀파일로 데이터베이스를 구축한다.

넷째, 엑셀로 구축된 데이터베이스에 근거하여 디지털 아카이브 시스템 구축을 제안하였다. 디지털 아카이브 시스템은 자료등록시스템, 디지털 자원관리시스템, 서비스 제공시스템으로 구성된다.

다섯째, 구축된 데이터베이스의 편리한 검색과 활용을 위해 디지털 아카이브 홈페이지(안)를 제시하였다. 홈페이지 메뉴는 5개의 구조로 구성되어 있으며, 중요한 테마별 콘텐츠와 유형별의 색인을 중앙에 배치하여 검색의 접근성을 높였다.

## 2. 선행 연구

선행연구는 디지털 아카이브 구축 필요성과 활용성 제고 방안을 제시한 연구와, 기 구축된 데이터베이스의 효율성과 활용도를 높이기 위한 방안을 제시한 연구로 나누어진다. 먼저 아카이브 구축 필요성을 제시한 논문으로 백영주와 김수자(2005)는 여성사 관련 기록물 보존상태의 취약점을 발견하고, 자료담당자의 부재,

공간의 협소함, 재정의 부족 등의 문제점을 지적하였다. 문제 해결을 위하여 자료수집 관리 전담기구로서의 자료관 설치, 자료의 체계적 DB 구축으로 웹을 통한 한국여성사 관련 정보의 제공을 주장하였다. 장윤금(2013) 역시 1800년대부터 1910년까지 국내 활동 개신교 외국인 선교사 자료의 수집과 디지털 아카이브 구축의 필요성을 제시하고, 선교사 기록물은 종교적인 가치를 넘어 근대 역사, 정치, 경제, 사회를 새롭게 조명하는 사료적 가치가 크다고 지적하였다. 또한 개별적으로 보관된 선교사 자료의 수집, 보존, 접근정책, 방안에 대한 체계적인 시도가 필요함을 보고하였다. 김주관(2005)은 20세기 민중생활사연구단 사례를 중심으로 문화자료와 디지털 아카이브 구축에서 생활사 아카이브 구축의 유용성과 방법론에 대하여 논하였다. 아카이브 구축시 해결과제로 참여 연구자의 아키비스트 역할 수행 교육훈련, 구술자료 관리 단위의 세분화, 새로운 민중생활사 분류체계 연구 등을 제안하였다. 서해숙(2013)은 '진도아리랑'의 디지털 아카이브 구축 필요성, 세부적 설계 과정과 방안에 대해 연구하였다. 디지털 아카이브의 표준화 모형과 디지털 정보의 선택 및 컬렉션 개발, 구분 체계, 기술 방법에 관한 현황과 문제점을 살펴보고, 진도아리랑 디지털아카이브 구축을 위해 기록물 문서, 이미지, 동영상, 음성 사운드 등의 표준포맷으로 HWP, TXT, JPEG, AVI, WA을, 메타데이터 표준으로 더블링크어 요소와 세부속성을 제안하였다.

이미 구축된 문화역사자료 데이터베이스의 효율성과 활용도를 높이기 위한 방안 제시 연구로서 함한희와 박순철(2006)은 문화원형콘

텐츠 DB 활용도 제고, 문화산업의 성장, 이용자의 편리한 시스템 사용과 자료 활용, 새로운 아이디어 창출, 적극적 생산자로 전환하기 위한 시스템 구축을 제안하였다. 기존 주제검색 시스템의 어려움을 극복하기 위해 온톨로지 기능과 의미검색, 문서요약, 문서클러스터링, 문서분류, 주제추적 영역의 데이터마이닝 기능을 포함한 디지털 아카이브 시스템을 제안하였다. 심재석(2008)은 '한국향토문화전자대전'의 홈페이지 수록 자료들의 디지털 아카이브 활용성을 인식하고, 하이퍼텍스트 전환과정과 구축된 홈페이지의 활용가능성을 제시하였다. 그리고 기존 한국학 관련 시스템들과의 자료공유와 연계서비스를 제공하는 아카이브 게이트웨이로서 한국향토문화전자대전의 활용가능성을 언급하였다. 강순애(2008)는 국가기록원, 국사편찬위원회, 한국학중앙연구원의 데이터베이스 구축과정과 이들이 운영하는 중앙기록관리시스템, 나라기록포털시스템, 한국사 데이터베이스시스템, 한국향토문화대전시스템의 특성에 관하여 기술하였다. 특히 세 기관 사이의 공동협의회 구성, 근·현대사료의 디지털화 자원 공유를 위한 기술개발, 콘텐츠 이용과 개선, 공공 이용과 참여, 국제적 이용을 위한 기관간의 역할 분담을 강조하였다. 김수영과 임준근(2009)은 한국학중앙연구원의 한국향토문화전자대전의 편찬과정에서 산출된 텍스트 자료를 포함한 디지털 콘텐츠들을 효과적으로 기록 보존하기 위해 XML 기반의 디지털 아카이브 시스템을 설계하였다. 그리고 객체지향형 멀티미디어 데이터베이스를 구축하여 디지털 아카이브 시스템의 호환, 공유 가능한 알고리즘과 응용프로그램의 수동관리의 시행착오 보완·수정을 주장하였다.

2005년도 연구의 후속연구로 김주관(2006)은 디지털 아카이브 구축이 역사자료와 문화자료의 정보 공유와 중복연구의 방지, 문화자료의 문화원형으로서 문화산업에 미치는 파급효과를 추정하고 구체적인 방법론의 고려 사항을 점검하였다. 특히 자료 저장매체의 선택, 자료 형식의 선택, 메타데이터의 항목 설정 등의 표준화 전제와, 20세기민중생활사연구단 아카이브 구축을 보고하였다.

이상에서 디지털 아카이브 구축 필요성이 제기되고, 기 구축된 각종 문화·역사자료의 아카이브 활용성을 높이기 위해 메타데이터와 저장 방법의 표준화, 온톨로지 구축을 통한 의미검색, 관련 데이터베이스간 연계 필요성을 제시하고 있다.

### 3. 개화기 방한 서양인 기록물의 디지털 아카이브 구축의 의의

디지털 콘텐츠는 디지털 형태로 제작·유통·소비되는 각종 동영상 파일, 이미지 파일, 멀티미디어 자료 등의 정보자원을 말한다. 디지털 아카이브는 미래의 접근을 위해 디지털 콘텐츠를 효과적으로 저장·보존하기 위한 시스템이며, 디지털 콘텐츠의 선택, 접근, 처리, 보존, 관리, 검색, 활용 기능을 체계적이고 효율적으로 구축하고 지원하는 관리시스템이다. 디지털 아카이브는 정보의 생산과 전달, 공유 속도를 극대화하고 지식에 대한 접근성을 향상시키며, 디지털로 전환된 콘텐츠는 디지털 복제로 수많은 사람의 용도에 맞게 사용될 수 있다. 그리고 디지털 아카이빙은 디지털 콘텐츠를 처

리, 보존하고 상호작용 전달기능, 원문 재생기능에 의해 미래 이용자들이 정보를 활용할 수 있도록 하는 전반적인 작업을 의미한다.

디지털 아카이브 구축의 필요성은 현실적 측면과 경제적 측면에서 모두 필요하다. 현실적 측면에서 최근 많은 자료들이 디지털 카메라, 디지털 캠코더 등의 디지털 기기에 의해 생산단계에서부터 디지털 자료로 제작되므로 전통적인 아날로그 형태 자료에 비해 저장성 및 보존성이 뛰어나다는 장점이 있다. 경제적 측면에서는 문서나 사진, 녹음테이프, 영상테이프 등의 아날로그 자료를 디지털화 하여 원자료 훼손없이 영구보존할 수 있으며, 검색시스템을 통해 손쉽게 자료를 검색·열람·활용할 수 있다.

인터넷 환경이 발달함에 따라 디지털 아카이브 자료의 공유, 활용이 과거보다 훨씬 용이해졌다. 디지털 아카이브를 통해 OSMU(One Source Multi Use) 활용과 COPE(Create Once, Publish Everywhere) 기능 구현이 가능하고, 다면적 정보 공유와 정보 활용이 가능하다. 특히 최근에 정부와 지자체, 기업 등의 지역문화 콘텐츠 개발이나 문화원형에 대한 관심이 고조됨에 따라 학계 뿐 아니라 민간 부분에서도 지역문화 콘텐츠의 디지털화에 대한 관심이 높으며 응용 가능성 또한 확대되고 있다. 이처럼 디지털 아카이브는 개방성, 접근성, 복제성, 융복합의 용이성의 속성을 가지고 있어서 학술적 가치 뿐 아니라 게임, 애니메이션, 드라마, 공연 등의 다양한 문화산업에서 활용되는 추세이다.

본 연구에서 개화기 방한 서양인의 조선 체류와 여행 체험을 서술한 기록물의 디지털 아카이브 구축 의의는 19세기에서 20세기 초 한반도의 자연지리적·인문지리적 정보를 분류하여 과거

의 지리를 복원하고 역사지리적 측면에서 지역에 관한 이해를 심화시키며 지역문화 콘텐츠로서의 활용에 기여하는 데 있다.

구체적으로 첫째, 관련 정보를 효율적으로 검색하고 이용할 수 있도록 외국인 기록물에 포함된 지역정보의 내용과 속성 유형을 구조화하고 데이터베이스 구축 방안을 마련할 것이다. 데이터베이스에 포함되는 내용으로는 자연지리 정보(지형, 토양, 기후, 식생 등의 계통적 분류에 따른 정보와 자연경관), 인문지리 정보(정치, 경제, 사회, 문화 등 계통적 분류에 따른 정보와 인문경관), 저자정보(생애사 및 조선 여행 및 체류 정보) 등이다.

둘째, 역사지리적 복원을 통한 지역연구를 위해 서양인 기록물에 나타난 ① 지형, 토양, 기후, 생태 등과 관련된 자연지리 정보, ② 인구, 행정, 그리고 도시·촌락의 기능, 분포, 연계 등과 관련된 지역적 정보, ③ 정치, 경제, 사회, 문화 등의 계통적 주제와 관련된 인문지리 정보를 추출하여 입력하고 분석할 것이다. 개화기 서양인의 방한 기록이 지리적 분석의 가치를 갖는 이유는 한반도의 여러 지역에서 수집된 원자료가 다양한 주제를 중심으로 기록되어 있어서 19세기 말 도시의 중심성, 배후지로서의 촌락의 역할, 지역 간 연계 등의 파악이 가능하기 때문이다. 그러므로 데이터베이스 구축을 통해 19세기 말에서 20세기 초의 주요 도시와 지역에 관한 기록 속의 자연지리적, 인문지리적 특성에 대한 분석이 가능해지며, 저자의 국적·성별·직업·체류시기·여행 장소 등의 맥락에 따른 차이도 살펴볼 수 있다.

셋째, 정보 공유와 지역문화콘텐츠로서의 활용을 위해 디지털 아카이브와 지도를 구축한다.

데이터베이스는 지역명(행정구역명, 지방명)에 따라 검색이 가능하도록 구성되고 위치 정보에 기반한 디지털 지도로 구축될 예정이다. 또한 매쉬업(mashup) 기능을 활용하여 해당 지역에 대한 정보를 제공하도록 설계되는데 지도를 매개로 이루어지는 지역 중심의 검색과 계통적 주제 정보 중심의 검색이 가능하여 검색의 편이와 효율성 뿐만 아니라 지역에 관한 교육효과도 높일 수 있다. 즉 시간과 공간을 디지털 지도를 통해 융합시키고 인터넷 상에서 공유함으로써 지역교육 및 지역문화 콘텐츠로서의 활용 가치를 부가시킬 수 있다.

넷째, 개화기에 서양인들이 방문했던 곳으로 북한의 지역들이 다수 포함되어 있어 북한에 관한 지역정보를 얻을 수 있다. 영국 영사 칼스(W. R. Carles), 영국인 여행가 비숍(I. B. Bishop), 독일인 기자 겐테(S. Genthe) 등은 1880년대~1890년대에 서울·평양·제물포·원산·의주·금강산·북한산 등의 관광지와 조선과 만주의 국경지대 등을 여행하여 방문지의 지리적 정보를 상세하게 기술하였다. 따라서 개화기 북한 주요 도시와 촌락을 포함하여 관광지, 국경지대에 관한 정보를 얻을 수 있다.

#### 4. 개화기 방한 서양인 기록물의 디지털 아카이브 구축

##### 4.1 구축 대상 자료

디지털 아카이브 구축 대상 자료는 개화기 조선에 체류했던 서양인이 저술한 기록물 중 1876년부터 갑오개혁 사이에 저작된 11권의 단

행본으로 구성되어 있다. 일반적으로 개화기는 조일수호조규가 이루어진 1876년부터 한일합방이 된 1910년까지를 의미하며, 본 연구의 구축 대상 범위도 이 시기를 망라하고 있다. 그러나 자료량의 방대함으로 인해 1차년도의 구축 대상 범위를 1876년부터 갑오경장이 시행된 1895년까지 기술된 자료 11권으로 한정하고 나머지는 2차년도에 포함시켰다. 대상 기록물의 국적, 저자명, 직업, 서명(원서명과 번역서명), 출판

년, 주요 내용은 <표 1>에 나타나 있다. 본 연구의 자료 입력범위를 조일수호조규가 이루어진 1876년부터 1894년까지로 정하였지만 뒤 알드의 『조선전』(1741년), B. 홀의 『조선 서해 탐험기』(1818년), 쥐베르의 『프랑스 군인 쥐베르가 기록한 병인양요』(1873)는 연도의 범위를 벗어났음에도 불구하고 서양인의 조선에 대한 기록적 가치를 고려하여 본 연구의 범위에 포함시켰다.

<표 1> 디지털 아카이브 구축 대상 자료

국적	직업	저자	번역서명 (원서명)	출판년	주요 내용
프랑스	신부	뒤 알드* (Jean-Baptiste Du Halde)	조선전 (Kingdom of Korea)	1741	파리 의방전교회 신부로 방한한 뒤 알드가 기록한 조선 견문록
	해군	H. 쥐베르* (Henry Zuber)	프랑스 군인 쥐베르가 기록한 병인양요 (Le Tour de Monde)	1873	병인양요에 군인으로 참전한 앙리 쥐베르의 기록
	여행가, 지리학자	C. 바라 (Charles Louis Varat)	조선종단기 (Le Tour du Monde)	1892	프랑스의 지리학자, 민속학자였던 C. 바라가 작성한 『세계일주(Le Tour du Monde)』 속에 포함된 한국 여행과 민속에 관한 기록
	군인, 외교관	C. 롱 베 (Chaille Long Bey)	조선기행 (La Coree ou Tchosen: La terre du calme mainal)	1894	1887~1889년까지 미국의 한성 주재 총영사이자 공사관의 서기관으로 부임했던 롱 베의 기록
	문헌학자	M. 쿠랑 (Maurice Courant)	한국의 역사와 문화	1894	프랑스 문헌학자였던 M. 쿠랑이 저술한 한국에 관한 논문 11편 수록
미국	목사, 동양학자	W. E. 그리피스 (William Elliot Griffis)	은자의 나라 한국 (Corea: the Hermit Nation)	1882	미국의 목사이며, 동양학자였던 W. E. 그리피스가 한국을 2회 방문하고 저술한 조선의 역사와 풍속
	천문학자	P. 로웰 (Percival Lowell)	내 기억 속의 조선, 조선 사람들 (Choson, The Land of the Morning Calm)	1885	미국 천문학자 P. 로웰의 한국 방문기
	교사, 선교사	G. W. 길모어 (George W Gilmore)	서울풍물지 (Korea from It's Capital)	1892	육영공원의 교사로 근무하다 미국으로 귀국한 후 조선에서의 생활을 회상하며 쓴 회고록이자 견문록
영국	외교관	W. R. 칼스 (William Richard Carles)	조선풍물지 (Life in Corea)	1888	조선의 영사였던 칼스가 서울, 경기도, 평양, 황해도 등지를 여행하고 저술
	해군, 지리학자	B. 홀* (Basil Hall)	조선 서해 탐사기 (Account of a voyage of discovery to the west coast of Corea)	1818	영국 해군이며 왕립지리학회 회원 B. 홀이 기록한 서해5도의 풍물
독일	무역상	E. 오페르트 (Ernest Oppert)	금단의 나라 조선 (A Forbidden Land)	1880	남연군의 묘를 도굴한 오페르트의 조선의 지리, 쇄국정책, 풍습에 대한 기록

## 4.2 분류체계 구성

대표적인 문헌분류체계인 한국십진분류법과 듀이십진분류법은 학문분야의 지식분류에 기반한 분류체계인데 반해 웹기반 분류체계는 기존의 웹 분류체계나 문헌분류체계를 변용 또는 확장하여 사용하는 실용적 관점의 분류체계이다. 이 연구에서 제시하는 웹문서 분류체계 모형은 실용적이면서도 학술정보용에 맞는 분류체계의 장점을 보완하는 방식으로 전개하였다. 본 연구에서 사용된 분류체계는 한국학중앙연구원에서 개발한 한국향토문화전자대전 콘텐츠 분류체계와 조선왕조실록 분류체계를 참고하여 서양인의 기록물에 담긴 지역 및 생활 정보의 속성 유형에 맞게 수정·보완하였다.

한국향토문화전자대전 콘텐츠 분류체계는 한국학중앙연구원에서 편찬된 '향토문화대전'에 사용된 분류체계이다. 한국향토문화전자대전의 분류체계는 콘텐츠 목차와 디렉토리 분류로 나누어진다. 콘텐츠 목차에서는 인간의 삶을 중심으로 대분류를 구성하였는데, 삶의 터전(자연과 지리), 삶의 내력(지방의 역사), 삶의 자취(문화유산), 삶의 주체(성씨 인물), 삶의 틀1(정치와 행정), 삶의 틀2(경제와 산업), 삶의 내용(종교와 문화), 삶의 방식(생활과 민속), 삶의 이야기(구비전승과 어문학) 등 9개의 대주제 하에 28개의 중주제, 160개의 하위주제를 가진 계층분류체계이다. 디렉토리 분류체계는 분야, 유형, 시대의 세가지 하위 메뉴로 제공되는데, 분야별 내용은 콘텐츠목차의 내용과 유사하며, 자연·지리(자연지리, 인문지리, 동식물), 역사(전통시대, 근현대), 문화유산(기록유산, 무형유산, 유형유산), 성씨·인물(성씨·

세거지, 전통시대인물, 근현대인물), 정치·경제·사회(정치·행정, 경제·산업, 사회·복지, 과학기술), 종교(불교, 유교, 기독교, 신종교), 문화·교육(예술, 체육, 교육, 언론·출판), 생활·민속(생활, 민속), 구비전승·어문학(구비전승, 언어, 문학)의 9개 하위주제, 29개 중주제로 구성된다. 유형별 내용은 개념용어, 기관단체, 놀이 등의 18개 하위주제를, 시대별 내용은 선사, 고대, 고려를 포함한 6개 하위주제를 가지고 있다.

조선왕조실록은 표점교감 조선왕조실록 데이터베이스, 국역 조선왕조실록데이터베이스, 국역원전 고종순종실록 데이터베이스의 세 종류로 구성되어 있으며, 현재 CD-ROM과 온라인 데이터베이스([http://sillok.history.go.kr/intro/intro\\_rule.jsp](http://sillok.history.go.kr/intro/intro_rule.jsp))로 제공되고 있다. 조선왕조실록 분류체계는 정치, 경제, 사회, 문화의 4개의 대주제 하에 40개의 중주제와 174개의 하위주제를 가진 정사 위주의 분류체계이다. 구체적인 대주제는 정치(왕실, 정론, 행정, 인사, 사법, 군사, 외교, 변란, 인물), 경제(재정, 금융, 물가, 상업, 무역, 교통, 도량형, 농업, 수산업, 광업, 공업, 건설), 사회(가족, 호구, 신분, 향촌, 의생활, 식생활, 주생활, 윤리, 풍속, 구휼, 보건), 문화(사상, 역사, 과학, 의학, 어문학, 예술, 교육, 출판)로 구성되어 있다.

위의 두 분류체계를 검토한 후 본 연구에서 구성된 분류체계는 개화기 방한 서양인이 기록한 기록물에 나타난 생활사 중심의 분류체계이다. 구체적으로 10개의 대주제(자연환경, 지역구분 및 경관, 역사, 인구, 인종, 정치, 경제, 문화, 사회, 언어·문학)와 40개의 중주제, 120개의 소주제로 구성되어 있다(〈표 2〉 참조). 분류



〈표 2〉 개화기 방한 서양인 기록물 정보의 분류체계 구성

대분류	중분류	소분류
자연환경	자연지리	위치, 지형, 지질, 토양, 기상·기후, 동물상, 식물상, 자연재해, 풍토, 수문, 하천, 산지
	자연경관	지역경관, 하천경관, 수변경관, 산지경관, 해안경관, 평야경관, 도서경관(섬)
지역구분 및 경관	지역구분	도시·촌락 구분, 행정구역, 지명, 지리지·고지도
	인문경관	도시경관, 촌락경관, 건조환경, 주거경관, 상업경관
역사	전통시대	선사시대, 낙랑, 고조선, 기자조선, 가야, 부여, 옥저, 삼한(마한, 진한, 변한), 신라, 백제, 고구려, 통일신라, 고려, 조선, 대한제국, 일제강점기(각 시대별 특정한 주제는 주제에 분류)
	유형유산	유적(건축물, 터, 왕릉 등)과 유물(기록물, 공예품, 조각, 서화 등)
	무형유산	음악, 무용, 놀이, 의식, 의례, 무예
인구	성씨	본관, 성씨집단
	인물	왕족, 행정가, 외교관, 사절단, 관리, 노비, 학자, 교육자, 독립운동가, 종교인, 문인, 예술인, 경제인(사업인), 의료인, 군인/경찰, 종교인(승려, 목사, 신부, 선교사), 외관, 기녀, 관리, 궁인(상궁, 내시, 시녀)
	인구구성	인구분포, 젠더, 아동, 노인, 인구수, 인구밀도
인종	민족	신체특성, 민족성, 지방색, 민족구성
	정치현황	정파, 정치조직, 정책
정치	행정	중앙행정, 지방행정
	인사	선발, 임면, 관리
	사법	법제, 치안, 재판, 탄핵
	군사	군정, 중앙군, 지방군, 통신, 병법, 전쟁, 무기
	외교	원, 명, 청, 일본, 여진, 러시아, 미국, 영국, 독일, 프랑스, 네델란드, 벨기에, 이태리
경제	경제현황	기반시설, 생산물, 특산물, 무역품
	기타산업	도매 및 소매업, 금융업, 운수업, 제조업, 숙박·음식점업, 교육, 오락·문화, 비공식 경제활동
	과학	과학기술, 의학, 약학, 수학, 물리, 생물, 천기, 역법, 지학, 보검
	재정·금융	재정일반, 전세, 공물, 상공, 전매, 화폐, 계, 물가, 임금, 수수료
	상업	상인, 시장, 상품, 무역, 교통
	농업	농작, 임업, 축산, 양잠, 과수원, 특용작물, 면직
	수산업	어업, 염업
	광공업	광산, 채광, 제련, 수공업품, 장인
	건설	건축, 토목
문화	종교	유교, 불교, 천주교, 기독교, 대종교, 도교, 천도교, 원불교, 무속신앙, 기타 종교
	언론	신문, 잡지, 출판
	예술	음악, 미술, 사진, 무용, 연극, 조경
	교육	초등교육, 중등교육, 고등교육, 여성교육, 교육수준, 인문교육, 기술교육, 특수교육
사회	민속	생업, 풍속, 세시풍속, 평생의례(결혼, 장례, 돌잔치, 회갑 등), 민간신앙(무속, 굿), 민속놀이, 마을축제
	공동체	조직, 활동
	가족	가족, 친족, 호적
	신분	양반, 중인, 양인, 천인, 신분변동
	의생활	관복, 상복, 예복, 평상복, 군복, 장신구
	식생활	주·부식, 주류, 기호식품, 식기
	주거생활	가옥, 가구, 택지, 취락, 생활용품
	여가생활	여행(관광, 순례), 오락(놀이)
언어·문학	언어	방언, 속담, 민요, 언어체계
	문학	고전시가, 고전산문, 한문학, 근대문학, 설화/전설

체계 구성은 문화지리, 역사지리, 종교지리 전문 연구자들과 문헌정보학 전문 연구자가 함께 모여 확정함으로써 전문성을 유지하도록 노력하였다. 주제분류 과정에서 주제, 사건, 시대, 공간상 중복되거나 누락되는 경우가 많아서 연구자들이 모여 추가적인 원칙을 부가하면서 진행하였다.

#### 4.3 메타데이터 항목 구성

디지털 아카이브 구축을 위해서는 메타데이터의 항목 구성이 필요하며, 메타데이터는 대상 정보자원의 속성과 특성 및 다른 자원과의 관계를 기술하여 정보 이용자의 검색에 도움을 주고 정보 제공자의 정보원 제어와 관리를 돕는 역할을 한다. 본 연구에서 웹 서비스 제공을 위한 데이터베이스 구축용 메타데이터를 구성하기 위해 더블린 코어 메타데이터를 표준 메타데이터로 사용하기에는 스키마가 부족하여 표현할 수 있는 정보 양이 제한적이므로 더블린 코어 기반 메타데이터 항목에 개화기 방한 서양인 기록물 내용을 분석한 후 내용에 나타난 생활정보를 보완한 37개 수정 메타데이터를 사용하였다. 특히 내용 메타항목의 대주제, 중주제, 소주제는 <표 2>의 주제를 그대로 사용하였다.

<표 3>은 본문정보 데이터베이스 구축에 사용되는 입력 메타데이터 항목명과 기능, 검색 가능 여부를 나타낸 것이다. 메타데이터 항목은 내용 메타항목, 사진자료 메타항목, 저자사항 메타항목, 출처 메타항목, 번역 메타항목의 5부분으로 나누어진다. 각 항목 기능에 대한 설명이 주어졌으며, '검색요소 가능여부'에서 '가

능' 표시 항목은 검색에서 접근점으로 사용 가능한 것을 의미한다.

내용 메타항목은 데이터베이스의 내용을 구성하는 항목들로서 표 번호, 해당자료 페이지 수, 대주제, 중주제, 소주제, 내용주제어(통제언어), 키워드(본문언어), 내용 항목이 포함된다. 표번호는 내용 분석과 데이터베이스 구축시 분석자가 매긴 일련의 번호를 말하는 것이며, 해당자료의 페이지 수 역시 같은 맥락에서 해당 페이지의 데이터베이스 내용을 지칭한다. 대주제, 중주제, 소주제는 <표 2>에 나타난 분류표상의 대·중·소주제에서 추출하였다. 자료를 3단계로 분류하는 이유는 각 데이터 베이스에 입력된 자료(예를 들어 문단) 당 검색어가 대·중·소분류의 위계 작성으로 진행되어 데이터 베이스 작성 시 분류자의 임의적인 분류로 인한 오류를 막을 수 있고, 검색의 정확성을 높일 수 있기 때문이다. 내용주제어(통제언어)는 원문 정보를 읽고 분석하여 제목으로 작성한 주제로서 내부 색인어 사전에 내장되어 통제언어 역할을 담당한다. 키워드(본문언어)는 원문정보에 나타난 중요한 단어로서 이들은 색인어인 동시에 검색의 접근점으로 사용되어 검색 효율성을 향상시킬 수 있다.

사진자료 메타항목은 사진파일명, 사진 설명, 사진 출처 항목으로 구성되며 사진은 첨부파일 형태로 부가된다. 사진파일명은 본문 중에 나오거나 인터넷에서 찾은 사진에 대한 설명으로 주로 jpg 파일 형태로 저장하였으며, 본문 중에 사진 설명이 있으면 그대로 사용하고, 인터넷 사진을 활용한 경우에는 설명과 함께 URL을 수록하였다.

저자사항 메타항목은 개인 저자명, 직업, 출

〈표 3〉 메타데이터 항목 구성

구분	항목명	항목 기능 설명	검색요소 가능여부
내용 메타 항목	표 번호	저자명과 함께 해당자료의 쪽번호 내에서 주어진 일련번호	
	해당자료 페이지수	해당자료의 쪽번호 내에서 주어진 일련번호	
	대주제	분류체계에 사용된 대분류의 주제임	가능
	중주제	분류체계에 사용된 중분류의 주제임	가능
	소주제	분류체계에 사용된 소분류의 주제임	가능
	내용주제어(통제언어)	원문 독해 후 연구자가 부여하는 주제	가능
	키워드(본문언어)	연구자가 본문 중에서 선택하는 자연어 주제	가능
사진 자료 메타 항목	내용	데이터베이스 내용으로 입력하는 본문 정보, 주로 문단형태로 입력됨	
	사진파일명	본문 중에 나오거나 인터넷에서 찾은 사진에 대한 설명. 주로 jpg 파일 형태임	가능
	사진 설명	본문 중에 사진 설명이 있으면 그대로 사용하고, 인터넷 사진을 활용한 경우에는 설명을 붙임	
	사진 출처	인터넷 사진을 활용한 경우에 URL을 밝힘	
저자 사항 메타 항목	사진	첨부파일 형태로 저장됨	
	저자명	저자 이름, 한글명과 원어명으로 병기	가능
	직업	저자의 직업	가능
	출신국	저자의 국적	가능
	출생 연도	출생 연도 표기	
	출생지	출생 장소	
	가족	부모, 형제, 배우자, 자녀 등	
	교육	학교, 전공	
	체류 기간	입국 일시, 체류 장소, 출국 일시 등	
	조선에서의 여행	여행 목적, 연도, 장소(경로) 등	가능
	출간자료	해당자료의 제목, 출간 연도, 출간지	가능
	사망연도	사망한 연도	
사망장소	사망한 장소		
사진 파일	인물의 사진을 파일로 제시(인물명, 연도, 출처)		
출처 메타 항목	서명	해당자료의 한국어 서명	가능
	원서명	해당자료의 원서명	가능
	장 제목	해당자료 안에 포함된 장 제목	가능
	원작 출판년	원작품이 출판된 연도	가능
	문헌 형태	작품의 성격(회고록, 기행문, 견문록, 여행서 등 다양함)	가능
	지역명	해당자료에 나타난 지명. 고대명인 경우에는 현대명으로 변환가능	가능
	GIS 변환용 지역	원작 지명을 GIS로 변환한 현대지명 사용	가능
	시대명	고조선부터 일제시대까지 포함됨	가능
원문링크	해당자료의 원문정보가 인터넷에서 확인되는 경우에 링크에 의해 연결		
번역 메타 항목	번역자	해당자료의 번역자	가능
	출판사	번역자료를 출판한 출판사	가능
	출판연도	번역자료를 출판한 연도	가능

신국, 출생 연도를 비롯한 13개 메타항목으로서 출생지와 가족, 교육, 직업, 사망 연도, 사망 장소 등의 부가적인 정보까지 수록하여 저자정보 데이터베이스 구축이 가능하도록 하였으며, 저자 사진으로 구성된 인물(전기)자료 데이터베이스 내용의 구성항목도 수록하였다. 따라서 데이터베이스 구축이 완료된 후 저자정보를 검색하면 기존의 포털사이트나 백과사전 데이터베이스에 나타난 저자정보보다 훨씬 깊이 있고 포괄적인 내용을 검색할 수 있다. 출처 메타항목에는 서명, 원서명, 장 제목, 원작 출판년, 문헌 형태, 지역명, GIS 변환용 지역, 시대명, 원문 링크 메타항목이 포함된다. 문헌 형태는 회고록, 기행문, 견문록, 여행서 등 다양하며, 지역명이 고대명인 경우에는 내장된 지역명 변환사전을 제공하여 현대명으로 변환이 가능하도록 하고, 원작 지역명을 GIS 변환 프로그램에 의해 현대명으로 변환이 가능하게 하였다. 또한 번역서의 원서 확인이 가능한 경우에는 URL 링크를 제공하여 원문을 확인해 볼 수 있게 하였다. 번역 메타항목에는 번역자, 출판사, 출판 연도가 포함된다.


#### 4.4 데이터베이스 구축

1차년도 연구과제인 본 연구의 데이터베이스 구축 과정 중 주제 확인과 분류, 텍스트 입력 과정에서 많은 예외 사항이 발생하였는데 주제 분류 과정에서 주제, 사건, 시대, 공간상 중복되거나 누락되는 경우가 많아서 추가적인 원칙을 부가하면서 업무를 진행하였다. 예를 들면, 추가로 어떠한 물건이나 행위, 대상의 장점이거나 기능에 대해 말할 경우에는 그 대상에 분류하


고, 조선 이외의 나라(고려, 신라, 백제, 고구려) 등의 신화 등은 언어문학>문학>설화에 분류하였다. 또한 분류가 이중으로 겹치는 경우에는 포괄적 대상을 기준으로 하며, 나머지는 통제언어에 삽입하는 등의 원칙을 정하였다. 또한, 역주자의 머릿말과 헌사를 DB에 포함시키는 과정에서 기존의 주제분류표로 구분하기가 어려워 '인구-인물-저자명 또는 역주자명'으로 입력하였다. 책에 있는 색인어 경우에는 DB 키워드에 반드시 포함시켰고, 해당 색인어로 검색되지 않아도 내용상 색인어에 해당되는 경우에는 색인어에 표기된 단어를 추가하였다. 또한, 색인어 단어와 다르게 본문에 패러프레이징(paraphrasing)되어 있는 경우에 색인어에 표기된 단어를 키워드에 추가하였다. 이러한 다양한 예외적 경우를 매뉴얼로 작성하여 2차년도에도 일관성을 유지할 수 있도록 하였다.

<표 3>의 메타데이터 항목에 의해 로웰(Percival Lowell)의 기록물인 <내 기억속의 조선, 조선 사람들>의 데이터베이스 사례는 <표 4>에 나타나 있다. 내용 메타항목 중 표 목차의 '로웰 74 2'는 로웰이 지은 저서의 74페이지 2번째 문단을 의미하고, '인구', '인물', '관리'는 <표 2>의 분류표에 나타난 대주제, 중주제, 소주제이다. '두 관리의 뛰어난 능력'은 본 연구자가 내용을 읽고 제목으로 부여한 내용주제어이며, '친절함', '온정', '사관', '능숙함', '군대지휘관', '사관', '민속학자', '외무어문'은 본문에 나타난 의미있는 키워드이다. 한편, 검색적인 측면에서 살펴보면, <표 3>의 검색요소 가능여부 항목에 '가능'으로 나타난 모든 항목들이 검색요소로 사용될 수 있는데, 지역명인 '서울', 'korea', 원작 출판년인 '1885년', 시대명인 '1883년', 문헌형태인 '여행

〈표 4〉 개화기 방한 서양인 기록물의 데이터베이스 구축 사례

구분	메타항목	내용
내용 메타 항목	표번호	로웰 74_2
	해당자료의 페이지수	74_2
	대주제	인구
	중주제	인물
	소주제	관리
	내용주제어(통제언어)	두 관리의 뛰어난 능력
	키워드(본문언어)	친절함, 온정, 사관, 능숙함, 군대지휘관, 사관, 민속학자, 외무아문
내용	두 사람의 관리가 나를 보살펴 주었는데 그 중 한 사람은 바로 옆집에 사는 군대 지휘관으로 집안일을 전담했다. 그는 어느 누구 못지않게 착하고 친절한 마음씨를 지닌 조용하고 사려 깊은 사람으로 그의 태도에는 언제나 온정이 담겨 있었다. 다른 관리는 외무아문의 비서 가운데 한 사람으로 보통 하루에 한 번 정도 방문해서 내가 무엇을 원하는가를 파악하는 임무를 띠고 있었다. 그 史官-그의 직명이다-은 고명한 민속학자로서, 접대하는 일에 능숙해 임시로 맡은 사무임에도 불구하고 놀라우리만치 잘해냈다. 만일 다른 나라에서 그의 위치라면 남의 초대를 받아 밖에서 식사하는 기회를 자주 가졌을지도 모르겠다.	
사진 자료 메타 항목	사진	 <p>조선의 지방 행정관(좌)과 공수(우), 태생 Zuber.</p>
	사진 파일	조선의 지방행정관 모습
	사진설명	
	사진출처	
출처 메타 항목	서명	내 기억속의 조선, 조선 사람들
	원서명	Choson, The land of the morning calm: a sketch of Korea
	장 제목	장안 풍경
	원작 출판년	1885년
	문헌 형태	여행서
	지역명	서울
	GIS변환용지역	서울
	시대명	1883년
원문링크	<a href="https://archive.org/details/chosnlandmornin00lowegoog">https://archive.org/details/chosnlandmornin00lowegoog</a>	
번역 메타 항목	번역자	조경철
	출판사	예담
	출판 연도	2001년

〈표 5〉 로웰(Percival Lowell)의 저자정보 데이터베이스 구축 사례 예시

항목	내용
저자명	Percival Lowell, 퍼시발 로웰
출생 연도	1855년
출생지	보스턴
가족	부모 캐서린 비겔로 로웰, 형제 애보트 로렌스 로웰은 하버드대학교 총장 역임. 여동생 에이미 로웰은 시인임. 미국의 부호가문업
교육	하버드대학 졸업
직업	사업가, 카운슬러, 천문학자. 주일 외교대표로 일본 10년 체류. 조선수호사절단 안내
기간	1883. 12. 20-1884. 3.17. 서울, 조미수호통신 사절대의 일환으로 조선 방문
조선 여행	고종 초청으로 부산, 제물포, 남산, 서울 등 한국 전역 방문
출간자료	『Choson, the Land of the Morning Calm: A Sketch of Korea』(내 기억속의 조선, 조선사람들), 1885년 하버드대학교 출판부 발간
사망 연도	1916년
사망 장소	아리조나주 플래그스태프
사진	 <p><a href="https://www.google.com/webhp?sourceid=chrome-instant&amp;ion=1&amp;espv=2&amp;ie=UTF-8#q=percival%20lowell">https://www.google.com/webhp?sourceid=chrome-instant&amp;ion=1&amp;espv=2&amp;ie=UTF-8#q=percival%20lowell</a></p>

서'로도 검색 가능하다. 본 연구에서는 기존의 기록물을 단순히 아카이브화 하는 것이 아니라 이에 대한 새로운 속성을 부여함으로써 자료의 검색 가능성과 활용성을 높이는 것을 목적으로 하였다. 〈표 3〉의 저자정보 메타항목을 가지고 로웰(Percival Lowell)의 저자정보 데이터베이스를 구축한 사례의 예시는 〈표 5〉와 같다.

#### 4.5 디지털 아카이브 시스템 구축

디지털 아카이빙은 협의로 컴퓨터 파일의 백업 컬렉션이나 데이터를 정리 통합하여 처리

가능한 형태로 정보파일 또는 집합체를 만드는 것을 의미하지만, 광의로는 디지털자원을 체계적으로 수집, 관리, 보존하여 활용할 수 있도록 자원의 평가, 입수, 소장 및 보존, 접근과 같은 처리기능과 네트워크 환경, 정보기술 환경, 이용자 환경 등을 관리하는 전반적인 작업을 말한다(두산동아대백과). 이러한 체계를 바탕으로 디지털 아카이빙 프로세스를 수집·등록, 관리, 서비스 제공 단계로 나누었다. 자료등록 단계 이전에 선행되어야 하는 작업은 인쇄형태의 대상자료를 수집하고 디지털화할 것인지 평가하는 작업이 선행되어야 한다. 이 시기에는 제출된 자료가 자료수집 정책에 부합하는 내용

인지 등의 평가와 검증이 행해진다. 또한, 주제 분류 체계를 설계하고 메타데이터 항목 설정이 이루어진다. 디지털 변환작업에서 저작권은 자료 선택에 영향을 미치는 중요한 요소 중의 하나이지만 본 연구에서 사용되는 자료는 출판된 지 백년이 지나 저작권 영향에서 비교적 자유롭다. 다만 번역본을 변환작업의 저본으로 삼기 때문에 번역본 사용 허가 등을 상세히 검토해야 한다.

자료등록 단계는 아날로그 문헌자료를 스캐닝하거나 이미지화 작업을 하는 디지털 변환작업이 이루어지는 단계이다. 디지털 아카이브 구축단계로 자료를 컴퓨터가 인식할 수 있도록 데이터를 디지털화 하는 단계이다. 유형별 메타데이터 정의 및 항목 설정, 분류기호 부여, 검색용 색인 데이터 추출, 메타데이터 입력, 원문 데이터 입력으로 등록시스템이 구축된다. 내장 통제어사전을 시스템에 구축하여 내용주제어를 통제하고, 지역명이 고대명인 경우에 현대명으로 변환하는 지역명 변환작업도 이루어진다. 또한 GIS 변환용 프로그램을 사용하여 원작에 나오는 지명을 현대지명으로 변환하는 작업도 행해진다.

디지털자원 관리 단계는 자료등록 단계에서 추진된 각종 업무가 제대로 수행되고 있는지 관리·감독하는 단계이며, 이용자의 정보검색에 필요한 각종 데이터 처리 작업을 수행하는 단계이다. 주제분류시스템의 관리, 메타데이터 처리과정, 색인데이터를 추출하고 관리하는 기능, 원문데이터 처리과정에서의 오류와 누락, 중복 등이 점검된다. 디렉토리 관리, 내장 통제어사전 관리, 웹사이트 구축이 이루어지며 웹사이트 상에서의 검색 및 화면출력 등과 같은

이용자 서비스 요청에 대한 처리도 고려된다. 이 단계에서 영구적 디지털 아카이브 환경을 구축하기 위한 시스템 보안과 통제기능도 수행된다.

서비스 제공 단계는 체계적인 서비스 시스템을 마련하여 보존에 의해 유지되는 디지털 콘텐츠를 이용자의 요청에 따라 원하는 정보 및 관련정보를 빠르고 정확하게 제공함으로써 정보를 확대 및 재생산하는 과정까지 포함한다. 이 단계에서는 웹사이트 구축으로 일반검색과 통합검색 등의 다양한 검색이 가능하며, 주제별, 테마별, 지역별, 시대별, 사진자료 검색 등도 가능하다. 검색된 디지털 콘텐츠에 접근하기 위해서 콘텐츠의 식별자에 대한 해석과 다운로드 작업이 필요하다. 또한, 위치 기반 디지털 지도 구축과 지역별 스토리텔링 자료 제공이 가능하고 아카이브의 연관 자료끼리 연결할 수 있는 기능으로 서비스 범위 확대도 가능하다. 자료등록시스템, 디지털자원 관리시스템, 서비스 제공시스템으로 구성된 아카이브 시스템을 <그림 1>로 표현하였다.

<그림 2>는 디지털 아카이브 홈페이지 화면(안)이다. 홈페이지 메뉴는 주제분류, 디렉토리 분류, 내용색인, 사진 데이터베이스, 아카이브 소개의 5개의 메뉴로 구성되었다. 주제분류 메뉴는 <표 2>의 대주제와 중주제를 계층분류 체계에 의해 조직한 것이고, 디렉토리 분류 메뉴는 테마분류, 시대분류, 지역분류로 나누어진다. 테마분류는 <표 2>의 소주제 중 의미있는 테마를 추출하여 가나다순으로 배열한 것으로 검색하려는 키워드가 어떤 주제에 속하는지 알지 못해도 쉽게 검색할 수 있도록 지원하고 있다. 시대분류는 선사시대, 고대, 삼한, 삼국, 고



〈그림 1〉 개화기 방한 서양인 기록물 디지털 아카이브 시스템 구조도



〈그림 2〉 개화기 방한 서양인 기록물의 디지털 아카이브 홈페이지 화면(안)

려, 조선, 대한제국, 일제강점기로 나누어지고, 지역분류는 경기도, 강원도, 충청도, 전라도, 경상도, 황해도, 평안도, 함경도의 8도로 구분하였다. 내용색인 메뉴는 원하는 항목을 가나다순으로 배열하여 이용자들이 쉽게 접근할 수 있게 한

색인으로서 인명색인과 지명/기관명 색인이 있다. 사진 데이터베이스 메뉴도 인명, 테마, 지명으로 검색 가능하다. 또한, 본문 중에 배치된 테마 콘텐츠는 자주 활용되는 콘텐츠를 5개 테마로 구성하여 제공하고 있으며, 사진·그림 갤러



리도 자주 사용되는 사진 콘텐츠를 5개의 주제로 구분하여 검색 가능하다. 또한, 북한지역 콘텐츠를 란을 따로 두어 평양, 원산, 개성, 백두산, 황해도 등의 자료를 찾아볼 수 있다. 그리고 메인 메뉴 속에서 자주 사용되는 콘텐츠인 사진 데이터베이스, 인명색인, 지역별 콘텐츠, 시대별 콘텐츠, 교육테마별 콘텐츠, 원작 내용보기 란을 홈페이지 밑에 따로 배치하여 접근성을 높였다.

검색시스템은 이용자가 원하는 다양한 검색 방법과 검색요소로 검색할 수 있도록 시스템을 구현해야 한다. 홈페이지를 구축하여 개화기 방한 서양인의 기록물 소개와 검색, 출력물 등을 제공하기 위해 검색시스템의 검색요소로 저자명, 서명, 원서명 등의 원작에 관한 사항뿐 아니라 장 제목, 키워드, 대주제, 중주제, 소주제, 지역명, 내용주제어(통제언어), 키워드(본문언어) 등의 다양한 접근점을 제공하고 있다. 특히 검색엔진은 내장된 검색어 사전 기능을 포함하여 내용주제어를 통제언어로 바꾸어 주는 기능과 과거 지명과 현재 지명을 매치시키는 매칭 기능도 수행해야 한다. 내용물을 웹상에서 제공하기 위하여 PDF 뷰어와 HWP 뷰어 기능이 제공되어야 한다. 또한 검색시스템은 스토리텔링으로 구성된 교육테마별 콘텐츠를 제공하여 학생들에게 교수학습용으로 활용할 수 있게 하고, 국내외 관광객을 위한 관광자원용의 다국어 지원 서비스와 모바일 서비스 제공도 고려되어야 한다.

## 5. 결론 및 제언

본 연구는 개화기 방한 서양인 기록물의 디지털 아카이브를 구축함으로써 19세기에서 20세

기 초 한반도의 자연지리적·인문지리적 정보를 분류하고 과거의 지리를 복원하여 역사지리적 측면에서 지역에 관한 이해를 심화시키며 지역문화콘텐츠로서 활용하기 위하여 시도되었다. 구체적인 내용으로는 조일수호조규가 이루어진 1876년부터 갑오개혁이 전개되었던 1894년까지 저술된 서양인 방한 기록물 11권의 내용항목을 분석하여 속성정보를 부여하였다. 10개의 대주제와 40개의 중주제, 120개의 소주제로 구성된 분류체계를 작성하고, 내용 메타항목, 사진자료 메타항목, 저자 메타항목, 출처 메타항목, 번역 메타항목으로 구성된 37개 메타데이터 항목을 추출하였다. 메타데이터 항목에 맞추어 텍스트 내용을 분석하고 엑셀로 된 데이터베이스를 구성하였으며, 디지털 아카이브 시스템 모델과 홈페이지 구축(안)을 제시하였다.

데이터베이스 구성은 크게 텍스트 자료와 사진자료로 나누어지는데, 텍스트 자료에는 자연지리 정보, 인문지리 정보, 저자정보가 포함되며, 사진자료에는 역사 기록물, 선교사 기록 자료, 신문자료 등에 포함된 사진 등이 포함된다. 디지털 아카이브 시스템 모델은 원자료 아카이브 구축단계인 자료수집시스템, 데이터를 디지털화 하는 단계인 디지털자원 관리시스템, 디지털 콘텐츠를 이용자의 요청에 따라 빠르고 정확하게 제공하는 서비스 제공시스템으로 구성된다. 그리고, 주제분류, 디렉토리 분류, 내용색인, 사진 데이터베이스, 아카이브 소개의 5개 메뉴로 구성된 디지털 아카이브 시스템의 홈페이지(안)를 제시하였다. 5개 메뉴 이외에 테마 콘텐츠, 사진·그림 갤러리, 북한지역 콘텐츠 란을 중앙에 배치하여 검색의 접근성을 향상시켰다.

본 연구는 2년차로 구성된 연구과제의 1차 연도 분량으로서 텍스트 자료의 방대함 때문에 1차 연도 자료구축의 범위는 조일수호조규가 이루어진 1876년부터 갑오개혁이 전개되었던 1894년까지로 제한하였다. 2차 연도에서는 1895년 갑오경장 이후부터 1910년 한일합방까지 기록물에 대한 텍스트 자료 분류, 데이터베이스 구축, 디지털 지도와 정보 서비스 플랫폼의 구축, 지역정보의 종합적 분석이 이루어질 것이다. 특히 2차 연도 연구에서는 관련 정보를 인터넷 상의 지도를 통해 확인·검색하여 과거지리 복원과 지역성에 대해 종합적·경험적 이해가 가능하고 지역문화콘텐츠로 활용할 수 있을 것이다. 디지털 지도의 구성은 지역 중심, 인물(저자) 중심, 주제정보 중심, 경관 중심 등의 여러 가지 테마를 개발하여 이용자들이 다양한 방식으로 콘텐츠를 활용할 수 있도록 할 예정이다.

추후 연구를 위한 제언은 다음과 같다. 첫째, 아카이빙 구축을 위한 데이터의 지속적인 발굴을 통한 아카이브의 확대방안을 모색해야 한다. 이와 함께 기존 구축된 데이터베이스를 바탕으로 콘텐츠 단위의 활용 방안을 모색하여 단순히 축적을 중심으로 하는 아카이빙이 아닌 콘텐츠의 조합을 통한 새로운 콘텐츠의 창조가 가능한 아카이브 시스템이 구축 되어야 한다. 다양한 콘텐츠와 서비스를 매쉬업을 통해 융합하면 창의적인 지역 문화콘텐츠를 풍부하게 할 수 있다.

둘째, 아카이브 시스템간의 연계를 통한 종합적인 확인과 활용 방안을 모색해야 한다. 역사·지리분야 디지털 아카이브는 국·공립기관, 민간연구소, 대학기관 등 다양한 기관에서 한국고전적 종합목록시스템, 한국역사정보 통합시스템, 국립중앙도서관 전자도서관, 한국고전종합 DB 등으로 활발하게 구축되고 있다. 다양한 역사·지리 관련 홈페이지를 연결하고 원하는 자료에 접근 가능하게 하는 게이트웨이 역할이 필요하며, 다양한 원문서비스가 가능한 통합정보시스템과 아카이브가 연계된다면 디지털 한국학 아카이브 역할을 충실히 수행할 수 있을 것이다.

셋째, 향후 아카이브 콘텐츠 공유와 상호운용성을 고려한 시스템의 표준화가 준수되어야 한다. 디지털 콘텐츠 제작 과정에서 정보자원의 특성을 반영한 표준 정보기술, 표준 메타데이터 개발, 분류체계의 표준화 등으로 디지털화가 이루어지면, 표준화를 통해서 이루어진 결과물은 서로 연계 활용되는 상호운용성이 가능하며 구축된 아카이브의 활용도는 극대화될 것이다.

넷째, 디지털 아카이브 시스템은 호환성과 사용자 편의성을 고려한 시스템 구축이 이루어져야 한다. 디지털 아카이브 시스템의 자료 유형과 내용 기술형식, 검색과 색인, 웹 인터페이스 및 각종 인프라 등의 호환성과 안정성, 확장 가능성이 고려되어야 한다.

## 참 고 문 헌

- [1] 강순애. 2008. 근·현대 지역사료 수집, DB 구축 및 공동 활용: 국가기록원, 국사편찬위원회, 한국학중앙연구원을 중심으로. 『한국기록관리학회지』, 8(2): 39-60.
- [2] 김수영, 임준근. 2009. 디지털 향토문화콘텐츠를 위한 XML 기반의 아카이브 시스템에 관한 연구: 『한국향토문화전자대전』의 사례를 중심으로. 『인문학콘텐츠』, 15: 147-172.
- [3] 김주관. 2005. 생활사 아카이브 구축의 의미와 방법. 『지방사와 지방문화』, 8(1): 217-249.
- [4] 김주관. 2006. 문화자료와 디지털 아카이브의 구축-20세기민중생활사연구단의 사례를 중심으로. 『지방사와 지방문화』, 9(2): 447-483.
- [5] 박재영. 2007. 역사적 스테레오타입 사례연구: 서세동점기 오페르트와 뮐렌도르프의 조선이미지. 『한·독사회과학논총』, 17(1): 351-380.
- [6] 백영주, 김수자. 2005. 여성사 관련 자료의 보존현황과 관리방안: 해방이후~1980년대 여성정책 및 여성단체 자료를 중심으로. 『한국기록관리학회지』, 5(1): 5-21.
- [7] 서해숙. 2013. 진도아리랑 디지털 아카이브 구축의 필요성과 설계. 『인문콘텐츠』, 29: 214-240.
- [8] 심재석. 2008. 디지털 한국학 아카이브의 활용방안에 관한 연구: 한국향토문화전자 대전을 중심으로. 『제4회 세계한국학대회 논문집』.
- [9] 이남희. 2008. 생활사 아카이브 구축을 위한 문서자료의 조직: 고문헌의 전자문서 편찬과 표준화. 『영남학』, 14: 63-100.
- [10] 장윤금. 2013. 우리나라 초기 외국인 선교사 자료의 디지털 아카이브 구축 필요성 연구(1800-1910). 『정보관리학회지』, 30(4): 265-281.
- [11] 함한희, 박순철. 2006. 디지털 아카이브즈의 문제점과 방향: 문화원형 콘텐츠를 중심으로. 『한국비블리아학회지』, 17(2): 23-42.

**• 국문 참고자료의 영어 표기**

(English translation / romanization of references originally written in Korean)

- [1] Kang, Soon-Ae. 2008. "The Acquisition, Construction and Common Use of Modern and Post-modern Document DB in the NAK, NIKH, and AKS." *Journal of Korean Society of Archives and Records Management*, 8(2): 39-60.
- [2] Kim, Su-Young and Lim, Jun-Keun. 2009. "A Study for Digital Contents in XML based Archive System." *Korea Humanities Content Society*, 15: 147-172.
- [3] Kim, Joo-Kwan. 2005. "The Meanings and Methods of the Construction of People's Everyday

- Life Histories Archives.” *Local History and Local Culture*, 8(1): 217-249.
- [4] Kim, Joo-Kwan. 2006. “Cultural Resources and the Construction of Digital Archives: A Case Study of Group for the People without History.” *Local History and Local Culture*, 9(2): 447-483.
- [5] Park, Jae-Young. 2007. “A Study of historical Stereotype: Images of Chosun by Oppert, E. J. and Moellendorff, P. G. in times of Occupancy of Western Countries.” *Zeitschrift der Koreanisch-Deutschen Gesellschaft fur Sozialwissenschaften*, 17(1): 351-380.
- [6] Paik, Young-Joo and Kim, Soo-Ja. 2005. “The Preservation and Management of Materials of Women’s History: On the Materials of Women’s Policies and of Women’s Organizations from 1945 to 1980’s.” *Journal of Korean Society of Archives and Records Management*, 5(1): 5-21.
- [7] Seo, Hae-Sug. 2013. “Necessity and Design of Construction of Digital Archive of Jindo-Arirang.” *Humanities Content*, 29: 214-240.
- [8] Shim, Jae-Suk. 2008. A Study on a Practical Use of Digital Archive for Korean Studies: Focusing on the “Grandculture Project” of the AKS. *Proceedings of the 4th International Conference of Korean Studies*.
- [9] Lee, Nam-Hee. 2008. “The Organization of Old Documents for the Lifestyle History Archives: With the Standardization of Compilation for Digital Documents.” *Yeongnam Area Studies*, 14: 63-100.
- [10] Chang, Yunkeum. 2013. “The Needs of Digital Archive Development for the Records of Early Foreign Missionaries in Korea (1800-1910).” *Journal of the Korean society for Information Management*, 30(4): 265-281.
- [11] Hahm, Han-Hee and Park, Soon-Cheol. 2006. “Digital Archives of Cultural Archetype Contents: Its Problems and Direction.” *Journal of Korean Biblia Society for Library and Information Science*, 17(2): 23-42.