

EAD 기반 영구기록물 통합검색시스템 설계 및 구현

Design and Implementation of an EAD-based Integrated Retrieval System for the Archives in Korea

박 상 현 (Sang Hyun Park)*

김 희 섭 (Heesop Kim)**

목 차

- | | |
|----------------------------|-----------------------------|
| 1. 서 론 | 3.2 영구기록물의 EAD 필수요소 |
| 1.1 연구의 필요성 및 목적 | 3.3 영구기록물 통합검색시스템(KAIRS) 설계 |
| 1.2 연구 방법 및 제한점 | 3.4 영구기록물 통합검색시스템(KAIRS) 구현 |
| 1.3 선행 연구 | 4. 시스템 평가 |
| 2. 국가기록원과 대통령기록관의 검색시스템 개요 | 4.1 시스템 구동 검증 |
| 2.1 검색시스템 현황 | 4.2 사용성 평가 |
| 2.2 Open API | 5. 결 론 |
| 3. 영구기록물 통합검색시스템 설계 및 구현 | |
| 3.1 EAD 요소 선정 | |

<초 록>

본 연구의 목적은 기존 국내 영구기록물관리기관의 검색시스템을 하나의 통합된 검색과 상호 데이터 교환으로 그 사용성을 높이하고자 새로운 통합검색시스템을 설계하고 구현하는 것이다. 이를 위하여 국내 영구기록물관리기관의 검색 시스템 현황을 분석하여 국내 환경에 적합한 EAD의 요소를 선정하였다. 선정된 EAD 요소를 기반으로 새로운 검색시스템(가칭 한국영구기록물 통합검색시스템, KAIRS)을 설계하였고, 기존의 두 영구기록물관리기관의 시스템을 에뮬레이션을 통하여 새로운 검색시스템을 구현하였다. 사용성 평가결과 새롭게 구현된 KAIRS가 기존 시스템보다 높은 만족도를 보였다.

주제어: EAD, 영구기록물관리시스템, 기록물 통합 검색, 기록물 검색 시스템 설계 및 구현

<ABSTRACT>

This study aims to design and implement an integrated information retrieval system for the archives in Korea. Particularly, this study is focused on the improvement of the usability of the system. To achieve the goal, this study analyzed the current retrieval systems of national archives and presidential archives in Korea. A new integrated retrieval system, called KAIRS, is designed based on the EAD elements, which is implemented through emulation technic between two retrieval systems. KAIRS showed higher of user satisfaction compared to the existing two retrieval systems in their usability test.

Keywords: EAD, Encoded Archival Description, Archives, Integrated Archival Retrieval, Archival Retrieval System Design and Implementation

* 경북대학교 대학원 기록학과(manai3@naver.com) (제1저자)

** 경북대학교 문헌정보학과 교수(heesop@knu.ac.kr) (교신저자)

■ 접수일: 2014년 1월 31일 ■ 최초심사일: 2014년 2월 4일 ■ 게재확정일: 2014년 2월 14일

■ 한국기록관리학회지 14(1), 101-124, 2014. <<http://dx.doi.org/10.14404/JKSARM.2014.14.1.101>>

1. 서론

1.1 연구의 필요성 및 목적

정보서비스의 요구 증대에 따라 2013년 현재 정부는 '정부3.0'을 정부 운영의 핵심가치로 설정하고 공공기관의 정보공개 확대를 정책 기조로 하여, 원천데이터의 민간 개방, 범정부 단일 정보공유플랫폼 구축, PC·스마트기기의 활용 등으로, 언제 어디서나 협업 가능한 정부통합의 사소통신시스템을 구축하여 대국민 정보서비스의 제공을 추진하고 있다.

이와 같은 정부의 대국민 정보서비스의 일환으로 기록물 관리측면에서도 새로운 전환이 요구된다. 기록물은 「공공기록물 관리에 관한 법률」 제5조에 따라 생산부터 활용까지의 전 과정에 걸쳐 진본성·무결성·신뢰성 및 이용가능성이 보장될 수 있도록 관리되어야 한다. 이 중 이용가능성을 높이는 방안으로는 기록정보서비스의 확대를 들 수 있다. 이러한 서비스 확대의 배경에는 자동화 기법의 발전, 전산화, 인터넷 등으로 네트워크의 발전에 따라 기록관리 분야에도 기록물의 관리와 보관뿐만 아니라 이용자를 위한 기록정보서비스에 대한 관심이 커지고 있는 실정이다.

실제로 국가기록원은 국민 맞춤형 서비스 증진을 목표로 정부 3.0 사업의 조기정착과 실질적 성과 창출을 위해 기록물을 통한 공공정보 활용기반 확대를 적극 추진할 계획이라고 밝히고 있다. 구체적으로는 행정기관 기록물과 국가 기록원 보유 기록물의 통합 검색, 활용 체계 마련을 위해 정보공개시스템 등에 각 기관 기록 관리 시스템(RMS)의 기록물 제공 기능을 구현하

여 2015년부터 서비스 제공 예정이며, 현재 4%인 기록물 원문서비스 비율을 2017년까지 20%로 높이고, 또한 매년 약 15만권의 비공개 기록물을 적극적 공개 재분류하여 비공개 기록물을 최소화하고 공개전환 기록물 목록을 서비스할 예정이라고 한다.

이와 같이 기록물의 정보서비스를 확대하기 위해서는 이용자가 쉽게 기록물을 찾을 수 있는 검색도구가 중요하다. 기록물은 도서와 달리 유일본인 경우가 대다수인 까닭에 이용자들은 원하는 기록물이 어느 기록관에 소장되어 있는지를 파악하는 것이 매우 중요하지만 그 소장처를 알아내기가 여간 어렵지 않다. 이를 해결하기 위한 방안으로 기록물 관리기관들이 소장하고 있는 기록물에 대한 통합적인 검색이 하나의 대안이 될 수 있다.

외국의 경우, 많은 기록관들이 EAD(Encoded Archival Description)를 표준으로 채택하고, 검색도구의 데이터와 구조를 표준화하여 국제적으로 기술정보를 교환하기 위한 EAD DTD(Encoded Archival Description Document Type Definition)를 이용하고 있는 상황이다. EAD는 미국에서 개발되어 현재 영국, 캐나다, 호주 등에서 각국의 기록물 기술 규칙에 적용하여 폭넓게 사용하고 있다. EAD를 이용해 기술된 기록들은 XML(eXtensible Markup Language)을 표준으로 기록관간의 기록물 정보공유 시스템 구축에 적용하기도 하고, 관련기관인 박물관, 도서관과의 정보공유를 꾀하고 있다.

이와 같은 장점을 가진 EAD의 적용으로 기록물 검색과 이용을 위한 통합검색시스템이 우선적으로 필요시 되는 곳은 영구기록물관리기관일 것으로 판단된다. 국내 영구기록물관리기

관은 국가기록원, 대통령기록관, 지방기록물관리기관, 헌법기관에 속하는 국회기록보존소, 대법원기록보존소, 헌법재판소 심판사무2과, 선거기록보존소가 있다. 이 중 현재 온라인 검색을 제공하고 있는 기관은 국가기록원과 대통령기록관 뿐이며, 대다수의 기록관들은 기록물 검색을 위한 온라인 검색도구와 기록물 데이터베이스조차 마련하지 못하고 있는 실정이며 이들 기록관에서 소장하고 있는 기록물의 공유나 통합검색은 아직 실현되지 못하고 있는 실정이다.

이에 본 연구의 목적은 기존 국내 영구기록물관리기관의 검색시스템을 하나의 통합된 검색과 상호 데이터 교환으로 그 사용성을 높이고자 통합검색시스템을 설계하고 구현하는 것이다.

1.2 연구 방법 및 제한점

본 연구의 목적을 달성하기 위하여 채택한 연구방법은 다음과 같다.

첫째, 국가기록원과 대통령기록관의 검색시스템 구조를 파악하고, 온라인 조사를 통하여 이들의 기록물 검색도구를 조사하였다. 또한 두 기관의 외부 기관과의 연동을 위한 Open API의 구조를 비교하고 이에 대한 문제점을 도출하였다.

둘째, EAD 요소를 선정하기 위하여 「영구기록물기술규칙」과 EAD를 매핑하였고, 현재 국가기록원과 대통령기록관에서 사용하고 있는 요소들을 이에 적용시켰으며 더불어 필수요소도 선정하였다. 이때 대통령기록관의 기록물 계층인 레코드그룹/시리즈를 국가기록원의 군/계

열과 동일한 계층으로 다루었다.

셋째, 통합검색시스템 구현을 위하여 실험환경을 구축하였고, 구현된 시스템은 기록물 검색 중 대부분을 차지하는 철과 건으로 통합검색시스템의 기록물 검색을 한정하였다. 이에 EAD의 구성요소 중 <archdesc>과 <dsc>에 중점을 두었다.

넷째, 새롭게 구현된 통합검색시스템의 성능을 평가하기 위하여 기존 국가기록원의 나라포털 검색시스템과 비교하였다. 또한 기록물 검색에 경험이 있는 기록학 전공자들과 다른 분야의 석·박사 과정 연구자들을 평가 집단으로 선정하여 사용성 평가 실험을 실시하였다.

본 연구의 한계는 아래와 같다.

첫째, EAD 요소 선정 시 국내 영구기록물관리기관 전체를 대상으로 분석하려고 의도하였으나 현재 기록물 기술과 온라인 검색을 제공하고 있는 곳이 국가기록원과 대통령기록관이어서 이들 두 기관만을 대상으로 하였다.

둘째, 국가기록원과 대통령기록관의 내부 시스템은 외부에 공개하지 않고 있으며, 또한 기록물 DB 역시 내부 시스템관리자만이 수정 권한을 가지고 있기 때문에 영구기록물관리기관의 실제 기록물 DB를 설계하기는 불가능하였다. 따라서 기록물 검색을 통한 출력 정보를 바탕으로 새롭게 설계된 데이터베이스를 통해 검색이 가능한 기록물 데이터는 다소 제한적일 수밖에 없었다.

셋째, 통합검색시스템의 프로토타입은 기록물의 통합검색 기능에 초점을 맞추어 구현하였으며 이용자인터페이스(UI: User Interface)는 크게 고려하지 않았다.

넷째, 새롭게 구현한 통합검색시스템은 성능

평가에 초점을 둔 것이 아니라 연구 목적에 따라 사용성 평가에 초점을 두고, 기존의 국가기록원 나라포털 검색시스템과의 만족도 비교에 그쳤다는 점이다.

이러한 제한 요소는 향후 실험적 프로토타입 수준의 설계를 벗어난 보다 실제적인 설계 및 구현에서는 시간적 편의성을 배제함으로써 최대한 반영될 수 있을 것이다.

1.3 선행 연구

지금까지 연구된 EAD 관련 연구를 EAD의 소개에 관한 연구, EAD 요소 선정에 관한 연구, EAD를 토대로 구축된 검색도구에 관한 연구, 그리고 통합검색에 관한 연구를 국내·외로 구분하여 그 연구 동향을 개략적으로 살펴보면 다음과 같다.

강소연(2002)은 USMARC AMC의 한계와 EAD 1.0의 출현 배경, EAD의 구성과 요소를 개괄적으로 다루었다. 외국 기관들의 EAD 이용 사례를 들며 각 기관의 EAD 필수요소와 검색시스템의 체계에 대하여 설명하였다.

박은경(2002)은 EAD를 이용한 OAC 프로젝트를 통하여 기록 자료의 통합데이터베이스 구축 효과를 제시하였다. 또한 OAC의 구체적 입력템플릿과 화면디스플레이에 관하여 소개하고 있다. 또한 외국의 검색도구나 EAD의 필요성을 언급하였다.

박지영과 김태수(2007)는 EAD 2002의 구성과 특징을 분석하고, 연세대학교 기록보존소 소장 4.19컬렉션에 EAD를 적용하였다.

박현이(2011)는 EAD의 관리 현황과 사례 분석으로 국외의 RLG, OAC, LC에서 사용하

는 EAD 요소와 국내 영구기록물관리기관에서 사용하는 EAD 요소를 비교하고, 각 기관 기록 연구사와의 면담을 통하여 검색도구 공유의 필요성과 EAD 필수요소를 도출하였다.

한편, 국외의 EAD 관련 선행 연구를 개략적으로 살펴보면 아래와 같다.

Stocking(2004)은 “The RLG Best Practice Guidelines for Encoded Archival Description”을 실무에 적용한 사례를 소개하였다. 또한 검색도구의 이용자들은 더욱 더 자세하게 기술되는 메타데이터를 요구하고, 이러한 요구를 반영하여 하위 계층까지 필수요소의 구체화할 필요가 있음을 강조하였다.

Thurman(2005)은 EAD의 <eadheader>부터 <archdesc>의 하위요소 <did>까지 구체적으로 설명하고, EAC(Encoded Archival Context)의 구조와 기술요소를 살펴보았다. 또한 온라인 환경에서 EAD와 생산자 기술 표준인 EAC의 연계를 시도하였다.

Brown과 Harvey(2007)는 DUASCSC(Dalhousie University Archives and Special Collections)의 2003년에 구축된 EAD 검색도구를 Novanet library의 MARC 21의 형태로 변환하고, EAD에서 MARC 21 변환표를 설명하였다.

이상의 선행연구에서 살펴본바와 같이 EAD를 이용한 사례소개와 그 구축효과나 필요성에 대한 연구는 이루어졌지만, 실제로 EAD를 기반으로 영구기록물관리기관의 통합검색시스템을 설계하고 구현한 연구는 아직 찾아볼 수 없다는데 본 연구의 의의가 크다고 할 수 있다.

2. 국가기록원과 대통령기록관의 검색시스템 개요

2.1 검색시스템 현황

2.1.1 국가기록원

국가기록원의 중앙영구기록관리시스템(Central Archival Management System, CAMS)의 내부망과 외부망은 정부의 정보보안정책에 의해 망 연계가 단절되어 있다. 내부에서는 CAMS의 검색기능을 통해 내부 직원을 대상으로 기록물 기술 정보를 제공하고 있고, 외부에서는 나라기록포털을 통해 일반 이용자에게 기술서와 목록을 제공하고 있다. CAMS의 기록물 철, 건 목록DB에서 주기적으로 인터넷 서비스용 기록물 철, 건 목록 데이터를 나라기록포털DB로 전송한다. 검색서버는 포털DB와 콘텐츠DB의 신규 자료에 대해 자동적으로 색인 작업을 수행하고, 기록물 통합 검색 시에는 색인DB를 먼저 검색하는데 이는 기록물 철·건의 제목, 생산년도, 생산기관 등의 기본적인 메타데이터를 중심으로 구성되어 있다. 이들 데이터는 국가기록원의 공개서비스과에서 이용자들에게 제공하고 있다(한국. 안전행정부, 2013a).

나라기록포털(<http://contents.archives.go.kr>)의 기록물 검색은 '기록물검색', '기술계층검색', '주제별검색', '통합검색'으로 구분되어 있다. 나라기록포털의 검색페이지에서 디폴트로 제공되는 통합검색은 키워드 검색만 가능하며, 검색에 대한 조건 필드가 없다. 검색결과는 기록물철, 기록물건, 군기술서, 계열기술서, 기록정보컨텐츠, 주제별검색, 웹페이지, 대통령기록물, 대통령 온라인콘텐츠에서 검색된 내용을 제공한다. 이

중 대통령기록물, 대통령 온라인콘텐츠는 대통령 기록관의 온라인 검색부분을 Open API를 이용하여 탐색하여 연결해 준다.

'기록물검색'은 기본검색, 상세검색, 연관검색, 생산기관별 검색이 있으나 대통령기록관의 기록물들은 검색되지 않고, 국가기록원 DB의 기록물에 관해서 검색할 수 있다. 상세검색은 조건필드로 생산기관, 생산년도, 기록물형태, 관리기관, 공개구분, 철/건구분, 원문서비스 여부로 나누어져 있다. 생산기관별 검색은 각 조직에 대한 기록물을 철, 건 형태로 찾을 수 있다.

'기술계층검색'은 현재 서비스대상을 중앙행정기관 총 355군과 2,347개의 계열로 하고 있고, 기술항목의 효율적 관리와 서비스를 위해 기록물군은 기록물 수집·보존 등 기록물 관리이력 중심으로, 계열은 기록물 열람 및 활용 등 이용자 서비스 중심으로 기술하고 있다. 현재 계층별 검색에서 제공되는 철/건 계층의 기록물은 중앙행정기관 기술시 군/계열이 분류되어 있는 기록물에 대해서만 제공하고 있고 지방자치단체 기록물 등은 계층별 검색에서는 검색이 불가능하며 나라기록포털 전체 통합검색을 통해 이용해야한다 밝히고 있다.

'주제별검색'은 국정분야별, 주제유형별, 주제별 상세검색으로 구분되어 있다.

2.1.2 대통령기록관

대통령기록관의 온라인 검색은 대통령기록포털(<http://www.pa.go.kr>)에서 서비스하고 있다. 대통령기록관 검색시스템은 국가기록원의 나라기록포털과 마찬가지로 내부망과 외부망으로 분리되어 있어서 기록물검색 입력창을 통해 들어온 검색 요청은 내부망의 검색서버에

전달되며, 검색서버는 웹기록물DB, 포털DB, 콘텐츠DB에 접속하여 입력된 검색어 관련 검색 수행을 한다. 검색서버는 검색 요청에 대한 결과를 대통령기록포털의 검색 결과 화면에 출력한다. 웹기록물DB는 제14대~제17대 대통령 웹 기록물에 대해 색인 작업이 완료된 DB이고, 포털DB는 대통령기록포털의 게시판, 간행물 발간, 기록관리시스템, 기록물 분류체계검색 등의 색인 작업이 완료된 DB이다. 콘텐츠DB는 대통령기록물 온라인 콘텐츠 관련 색인 작업이 완료된 자료를 저장하고 있다.

대통령기록관의 온라인 검색 유형은 '통합검색', '기본검색', '상세검색', '분류체계검색', '빠른검색'으로 나누어진다. '통합검색'은 웹기록물DB, 포털DB, 콘텐츠DB의 통합검색으로 외부의 기록물과의 통합검색이 아닌 대통령기록관이 가진 여러 가지의 DB의 내용을 통합적으로 검색하는 것이다. '기본검색'은 검색조건을 기록물 유형과 온라인콘텐츠, 웹 기록물, 사이트로 주고 있고 검색어를 전송하는 방식이다. '상세검색'은 조건필드로 대통령명, 생산기관, 생산년도, 기록물형태, 철/건 구분, 공개구분을 조건필드로 하고 있다. '분류체계검색'은 이승만, 윤보선, 박정희, 최규하, 전두환, 노태우, 김영삼 등의 대통령을 기준으로 각 생산기관을 레코드그룹으로 하여 정리되어 있고, 이에 대한 계층별 검색이 가능하도록 설계되어 있다.

2.2 Open API

국가기록원의 나라기록포털과 대통령기록포털은 국가기록물에 대해 공개 및 활용의 활성화와 데이터 제공을 목적으로 Open API를 제

공하고 있다. 이는 실질적인 외부 연동을 위한 방식으로 이를 이용하기 위해서는 각 기관에서 제공하는 인증키를 받아 API 서비스 이용을 허가받은 이용자임을 식별하도록 한다. 기록물 검색에 대한 요청 방식은 아래와 같이 인증키와 검색어를 넣는 방식으로 되어있다.

국가기록원과 대통령기록관의 Open API를 통한 검색의 자세한 입력변수인 요청변수 및 조건변수, 출력변수를 살펴보면 <표 1>과 같다.

두 기관 모두 요청 필수요소로 인증키, 질의어, 철/건에 대한 구분을 가지고 있다. 두 기관에서 공통적으로 지니고 있는 조건변수로는 검색 방식에 관련하여 다음의 문구 정확하게 포함하는(<query_phrase>), 단어를 적어도 하나 포함(<query_or>)시키거나, query에서 연산자 'NOT' 사용과 동일시 사용되는 다음의 단어를 제외하는(<query_exclude>)가 있다. 그리고 나머지 공통적으로 가지는 태그들은 기록물 형태(<doc_type>), 출력값에 대한 정렬방식과 정렬순서를 나타내는(<sort>, <order>)와 한 페이지에 출력되는 결과 수와 조회 페이지를 나타내는(<display>, <pageno>) 태그들이 있다. 또한 공개여부를 나타내는 태그(<is_open>)이 공개:1, 비공개:2, 부분공개:3으로 동일한 값을 가진다.

그러나 이 태그들 중에는 이름과 정의는 같지만 값에 대한 정의가 다른 태그들이 있다. 이에 해당되는 예로는 기록물 형태(<doc_type>)의 경우 나라기록포털에서는 1: 행정박물(관인류), 2: 행정박물(상징기념물), 3: 행정박물(사무집기류), 4: 행정박물(기타), A: 일반문서류, B: 도면류, C: 사진, 필름류, D: 녹음, 동영상류, E: 카드류, F: 대장류, G: 국무회의록, H: 지도, I: 대통령전자문서, M: 정부간행물, O: 일반도서,

〈표 1〉 국가기록원과 대통령기록관의 Open API 변수 비교

| 입력 변수 | | 출력 변수 | | 변수설명 |
|---------------------------------|----------------------------------|----------------|-----------|-----------------|
| 국가기록원 | 대통령기록관 | 국가기록원 | 대통령기록관 | |
| key | key | - | - | Open API 사용 인증키 |
| query | query | - | - | 검색을 원하는 질의 |
| rc_type | rc_type | rc_type | rc_type | 철건 구분 |
| query_type | - | - | - | 검색방식 |
| query_field | - | - | - | 검색영역 |
| query_phrase | query_phrase | - | - | 다음 문구 정확하게 포함 |
| query_or | query_or | - | - | 다음 단어 적어도 하나 포함 |
| query_exclude | query_exclude | - | - | 다음 단어 제외 |
| - | - | rc_rfile_no | - | 기록물철ID |
| - | - | rc_ritem_no | - | 기록물건ID |
| - | - | title | title | 기록물 제목 |
| - | - | rc_code | - | 기관구분값 |
| prod_code | prod_nm | prod_name | prod_nm | 생산기관코드, 생산기관이름 |
| prod_dir_code | - | - | - | 생산기관코드(하위기관포함) |
| prod_year_from~ prod_year_to | prod_year_start prod_year_end | prod_year | prod_year | 생산년도 |
| is_open | - | is_open | is_open | 공개여부 |
| doc_type | doc_type | doc_type | - | 기록물 형태 |
| | | arcave_type | - | 기록물 구분 |
| online_reading | - | online_reading | - | 원문서비스 유무 |
| sort | sort | - | | 정렬 |
| order | order | - | | 정렬순서 |
| mgt_organ | - | mgt_org_name | - | 관리기관 |
| display | display | - | - | 한 페이지에 출력될 결과수 |
| pageno | pageno | - | - | 조회페이지 |
| version | - | - | - | 버전 |
| - | pre_type | - | pre_type | 대통령구분 |
| - | - | link | ink | 기록물 상세정보페이지 링크 |

P: 총독부간행물로 사용되고 있지만, 대통령기록포털에서는 A: 전자기록물, B: 종이문서, C: 시청각기록물, D: 간행물, E: 행정박물, F: 선물, G: 기타로 사용되고 있다.

태그의 이름은 다르지만 같은 정의의 변수와 어느 한 기관에만 있는 나머지 태그들을 정리해보면 생산기관에 대해서는 나라기록포털과 대통령기록포털에서 각각 <prod_code>, <prod_nm>으로 사용하고 있다. 생산년도는 나라기록포털

의 경우 <prod_year_from>~<prod_year_to>의 형태로 대통령기록포털은 <prod_year_start>~<prod_year_end>의 형태로 주어진다. 기록물구분과 관리기관에 대한 태그 <record_type>, <mgt_organ>는 나라기록포털에서만 주어진다. 대통령기록포털에서만 가지는 태그의 경우 대통령구분으로 <pre_type>이 있다. 대통령 기록관은 대통령별 레코드그룹으로 정리·기술하기 때문에 이와 같은 태그를 가진다.

출력필드를 살펴보면 두 기관 공통으로 기록물 구분값(<rc_type>)을 가지고 구체적인 구분값으로 rfile, ritem을 동일하게 가진다. 기록물 철, 건에 대한 출력 필드로 <title>을 가지고, 생산년도에 대하여 입력의 요청변수 형태는 달랐지만 출력필드로는 (<prod_year>)을 가진다. 또 공개 구분을 나타내는 출력변수로 동일하게 <is_open>을 가지며, 이에 대한 구분값으로 1:공개, 2:비공개, 3:부분공개로 동일값을 가진다. 마지막 공통 필드는 상세정보에 대한 링크로 url주소를 링크해주는 <link>를 가지고 있다.

앞서 살펴본바와 같이 현재 국가기록원과 대통령기록관은 Open API의 입력 변수 및 출력 변수들이 같지 않아 영구기록물의 통합검색을 지원하기가 어렵다. 또한 기록물의 메타데이터를 충분히 표현하고 있지 않고 있어 통합검색 시스템의 설계에서 많은 시간과 노력을 요하는 부분이라고 할 수 있다.

3. 영구기록물 통합검색시스템 설계 및 구현

3.1 EAD 요소 선정

EAD를 기반으로 영구기록물 통합검색시스템을 설계하기 위하여, 본 연구에서 기본적으로 국내의 「영구기록물 정리·기술 지침」의 요소를 중심으로 국가기록원과 대통령기록관에서 실제로 사용되고 있는 기술요소를 분석하고, 이를 바탕으로 적합한 EAD 요소를 선정하는 과정을 거쳤다.

3.1.1 식별영역

1) 참조코드

식별영역에서 참조코드의 경우 영구기록물 정리·기술 지침의 필수요소에 해당한다. 국가기록원에서는 군, 계열에서 '참조코드'를 사용하여 각 군과 계열은 고유코드로 표시된다. 그러나 철과 건의 경우 두 개가 같은 '관리번호'가 주어지며 건에서는 '문서번호'라는 새로운 번호가 주어진다. 홈페이지 검색상에서 건의 경우 '관리번호'만 주어지고 '문서번호'는 주어지지 않은 경우도 있다. 그러나 시스템 내부에는 철과 다른 식별 번호가 부여 되어 있다. Open API를 통한 출력 값을 보면 <rc_code>1310377</rc_code>, <rc_rfile_no>201104884117</rc_rfile_no> 이와 값이 다른 일련의 값들로 되어 있음을 확인할 수 있다.

대통령기록관은 레코드그룹과 시리즈는 '참조코드'라는 형태로 되어 있고, 철과 건은 '기록물철ID', '기록물건ID'가 주어져 있다. 이들은 각각의 계층에서 사용되는 용어가 다르게 되어 있지만 하나의 일련번호로 표시된다. '참조코드'에 해당하는 EAD요소로 <unitid>를 사용할 수 있다.

2) 일자

식별영역의 '일자'의 경우, 영구기록물 정리·기술 지침의 필수요소에 해당한다. 국가기록원에서는 군, 계열에서는 생산시기로 철에서는 '생산년도'와 '종료년도'로 표기하며, 건에서는 '생산년도'로만 나타내고 있다. 대통령기록관에서는 레코드그룹과 시리즈에서는 '포괄일자'로 년도, 월, 일까지 8개의 숫자들로 기간을 나타내며, 철과 건에서는 '생산일자'를 8자리 숫자

로 나타내고 있다. 이들은 EAD의 <unitdate>에 해당하는 부분이며 EAD에서는 생산년도~종료형태, 생산년도/종료년도 등의 다양한 형태를 사용할 수 있게 하고 있다.

3) 기술계층

식별영역의 '기술계층'은 <c level="series">와 같은 형태로 나타낸다. 국가기록원에서는 군과 계열로, 대통령기록관은 레코드그룹과 시리즈의 단위로 나타내고 있다. 군은 레코드그룹으로, 계열은 시리즈로 매핑이 가능하다.

4) 기술단위의 규모와 유형

식별영역 중 '규모와 유형'은 국가기록원에서는 기록물군, 계열 단위에서는 '기록물 유형별 수량'으로 나타내며, 철과 건단위에서는 '기록물 유형', '기록물 형태'로 구분하여 나타내고 있다. 이에 대한 기술은 EAD의 <physdesc>에 해당하며 <physdesc>는 하위요소로 <extent>, <dimensions>, <genreform>, <physfacet>를 가진다. 국가기록원과 대통령기록관에서의 '기록물 형태와 유형'이 각 기관에 맞추어져 있어 기관별로 기록물의 형태와 유형이 다르기 때문에, 기호가 아닌 서술형으로 장르와 물리적 특성을 나타내는 <genreform>에 포괄적인 기술을 할 수 있다. 또한 기록물의 수량에 페이지나 수량은 <extent>를 통하여 서술하여 나타낼 수 있다.

3.1.2 배경영역

배경영역의 생산자명은 국가기록원과 대통령기록관의 군, 계열, 철, 건 단위 모두에서 '생산기관명'으로 동일하게 사용하며 이는 EAD의 <origination>에 해당한다. '행정연혁'은 각

기관의 군, 계열 단위에서 사용되며 국가기록원은 전거를 이용하여 링크되는 형식임에 비하여, 대통령기록관에서는 '행정연혁'을 다른 기술요소들과 함께 관리하고 있다.

국가기록원의 경우 EAD로 변환시 <bioghist>내부에 <dao>를 통해 행정연혁 링크를 전송하고, 대통령기록관과 같이 서술형으로 다른 기술요소와 함께 있는 경우 <bioghist>안에 나타낼 수 있다. 기록물 이력은 대통령기록관에서는 기술요소로 사용하지 않으나, 국가기록원에서는 군, 계열에서 '보존이력', '수집이관이력', '수집이관기관', '수집이관방법'으로 나누어 기술하고 있다. 이를 EAD요소에 매핑시키면 '기록물 이력'을 <custodhist>으로 하여 '보존 이력'과 '수집이관 이력'은 <custodhist>에 직접적으로 전달되고, '수집이관기관'은 <custodhist>의 하위요소로 '수집이관의 직접적 출처'를 나타내는 <acqinfo>로 나타낼 수 있다.

3.1.3 내용과 구조영역

내용과 구조영역은 '범위와 내용', '평가·폐기·정리일정 정보', '추가수집 예상기록물', '정리체계'로 4개의 요소로 구성되어 있다. 국가기록원에서는 이 중 '범위와 내용', '평가·폐기·정리일정 정보'에 해당하는 '보존기간'을 사용하고 있고, 대통령기록관은 '범위와 내용'과 '정리체계'를 사용하고 있다. '범위와 내용'은 서술형의 형태로 EAD의 <scopecontent>에 해당하고, '평가·폐기·정리일정 정보'는 <appraisal>에 해당한다. '추가수집 예상기록물'은 현재 두 기관에서 모두 사용하고 있지는 않지만 향후 기술요소로 사용할 경우 EAD의 <accruals>를 통해 전달한다. 대통령기록관에서 사용하는 '정리체계'

에 대한 설명은 <arrangement>에 해당하는 영역이고, '색인어/기능어'는 ISAD(G)에는 사용되지 않는 영역이지만 EAD의 <controlaccess> 내부에서 기술이 가능하다.

3.1.4 접근과 이용 환경영역

접근과 이용 환경영역에 해당하는 것으로서 국가기록원과 대통령기록관에서 사용되는 하위 영역은 '접근환경', '이용환경', '자료의 언어'이다. '접근환경'과 '이용환경'은 국가기록원에서만 기술하고 있고 이들은 각각 <accessrestrict> 과 <userrestrict>로 나타낼 수 있다. '자료의 언어'는 국가기록원에서는 '언어'로, 대통령기록관에서는 '언어와 문자' 요소로 나타내고 있고, 이는 EAD의 <langmaterial>에 해당한다. 물리적 특성과 기술적 요구조건, 검색도구는 현재 두 기관 모두 기술요소로 사용하고 있지만, 향후 기술요소로 사용할 경우 <phystech>와 <otherfindaid>에 각각 매핑할 수 있다.

3.1.5 관련자료 영역

관련자료 영역은 국가기록원에서는 원본의 존재와 위치를 표시할 때 사용하고 있고, 이는 <originalsloc>에 해당한다. '사본의 존재와 위치'는 현재 두기관 모두 사용하고 있지는 않다. 그리고 '관련 기술 단위'와 '출판물 설명'은 국가기록원에서 각각 '관련기록', '출판정보'라는 요소로 사용하지만 실질적인 기술은 하고 있지 않다. 이 요소들은 각각 <relatedmaterial>와 <bibliography>에 해당한다.

3.1.6 추가설명 영역

추가설명 영역은 현재 두 기관에서 모두 사

용하고 있지 않은 영역으로 EAD의 <note>에 해당하며, <note>에서 <p>를 이용하여 서술형의 형태로 나타낼 수 있다.

3.1.7 기술통제 영역

기술통제 영역 또한 현재 사용되고 있지 않은 영역으로 요소들로는 '기술일자', '기술담당자', '규칙과 협약'이 있으며 이는 각각 <processinfo> <date>, <processinfo>, <descrules>로 표현 가능하다.

3.2 영구기록물의 EAD 필수요소

앞서 설명한 현재 국내 영구기록물관리기관에서 사용하는 기술요소를 EAD에 매핑하여 정리하면 <표 2>와 같다.

<표 2>는 국가기록원과 대통령기록관에서 실제 사용하고 있는 기술요소들을 군, 계열, 철, 건에 따라 표기하였다. 또한 표의 필수요소는 선행연구(박현이, 정은경, 2012)에서 영구기록물 담당자와의 면담을 통하여 선정된 필수요소와 『영구기록물 정리·기술 지침』의 필수요소를 토대로 구성하였다. 이 중 음영 처리된 부분은 선행연구에 의해 선정된 각 기록물 기술 단위에서 필수요소를 나타낸다. 선행연구의 필수요소는 5개의 영구기록물관리기관 국가기록원, 국회기록보존소, 대법원기록보존소, 헌법재판소 심판사무2과, 대통령기록관의 담당자를 대상으로 한 EAD 필수요소 면담조사 중 3개 기관 이상이 공통으로 필요하다고 응답한 요소이다.

〈표 2〉 국내 영구기록물관리기관에서 사용하는 기술요소와 EAD의 매핑표

| 기술요소 | EAD | 필수구분 | 군 | 계열 | 철 | 건 |
|-----------------|-----------------------------|------|---|----|---|---|
| 기술계층 | <c> LEVEL attribute | 필수 | ○ | ○ | ○ | ○ |
| 기술적 식별 | <did> | 선택 | ○ | ○ | ○ | ○ |
| 참조코드 | <unitid> | 필수 | ○ | ○ | ○ | ○ |
| 제목 | <unittitle> | 필수 | ○ | ○ | ○ | ○ |
| 일자 | <unitdate> | 필수 | ○ | ○ | ○ | ○ |
| 생산자명 | <origination> <corpname> | 필수 | ○ | ○ | ○ | ○ |
| 기록물 유형 | <physdesc> <genreform> | 선택 | ○ | ○ | ○ | ○ |
| 기록물 수량, 페이지 정보 | <physdesc> <extent> | 선택 | ○ | ○ | ○ | ○ |
| 자료의 언어 | <langmaterial> | 선택 | ○ | ○ | | |
| 행정연혁 | <bioghist> | 선택 | ○ | ○ | | |
| 수집이관이력 | <custodhist> | 선택 | ○ | ○ | | |
| 수집/이관의 직접적 출처 | <acqinfo> | 선택 | ○ | ○ | | |
| 범위와 내용 | <scopecontent> | 필수 | ○ | ○ | | |
| 평가, 폐기, 처리일정 정보 | <appraisal> | 선택 | | | ○ | ○ |
| 색인어 | <controlaccess> | 선택 | | ○ | | |
| 접근환경 | <accessrestrict> | 선택 | | ○ | | ○ |
| 이용환경 | <userrestrict> | 선택 | | ○ | | |
| 원본의 존재와 위치 | <originalsloc> | 선택 | | ○ | | |
| 관련 기술단위 | <relatedmaterial> | 선택 | | ○ | | |
| 출판물설명 | <bibliography> | 선택 | | ○ | | |

필수요소: 「영구기록물 정리·기술 지침」의 필수/선택요소.
선행연구(박현이, 정은경, 2012)에 의한 필수요소는 음영 처리하였음.

3.3 영구기록물 통합검색시스템(KAIRS) 설계

본 연구의 목적을 위한 통합검색시스템은 국가기록원의 CAMS와 대통령기록관의 PAMS, 그리고 현재는 존재하지 않지만 향후 설립될 기관인 가상의 지방기록물관리기관의 시스템(Regional Archives Management System: RAMS라 지칭함)의 영구기록물에 대한 통합 검색을 목표로 하여 설계하였다. 현재 헌법기록물관리기관들은 설치되어 있지만 기록물에 대한 온라인 검색시스템을 구축하지 않아 소장

기록물에 대하여 기술 정보를 획득하는 것이 불가능하고, 지방기록물관리기관은 현재 설립되지는 않았지만 이 기관에서 소장하고 있어야 할 기록물들을 국가기록원에서 보유하며 기술하고 있어 기록물의 기술정보를 획득할 수 있기 때문에 본 연구에서 구현될 통합검색시스템은 가상의 기관인 지방기록물관리기관을 예물레이션 하였다. 이러한 요구사항을 반영하여 구현될 시스템의 명칭을 편의상 한국영구기록물 통합검색시스템(Korea Archives Integrated Retrieval System: KAIRS)이라 지칭한다.

3.3.1 KAIRS의 구조

KAIRS에서는 <그림 1>과 같이 통합 검색 서버와 각 영구기록물관리기관 서버 사이에 표준화된 EAD 프로토콜을 이용하여 모든 영구기록물관리기관과 동일한 메시지 형태로 기록물 정보를 검색할 수 있도록 하였다. 또한 이용자는 어디에서나 웹을 통해 KAIRS으로 접속하여 필요한 기록물을 검색할 수 있도록 구현하였다.

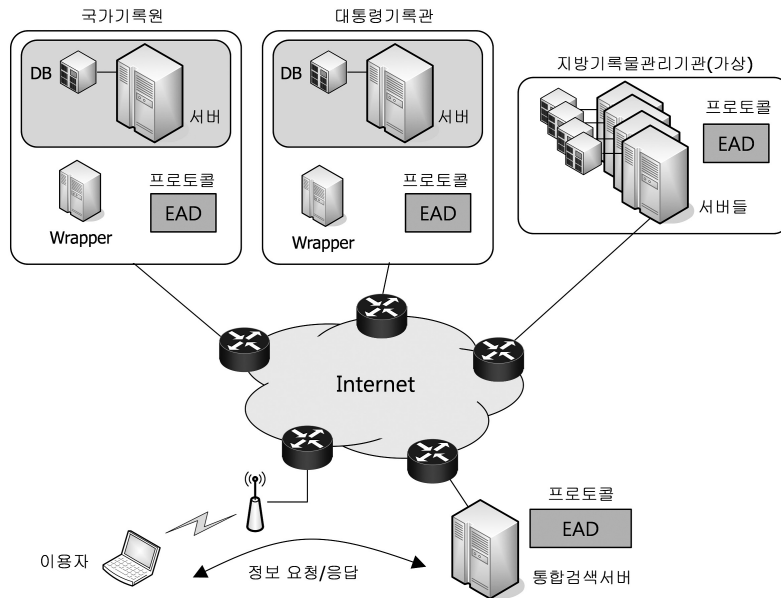
KAIRS에서는 기본적으로 규격화된 EAD 프로토콜을 사용하여 기록물 정보를 검색/교환하기 때문에 시스템 확장에 용이한 특징이 있다. 기존의 영구기록물관리기관은 검색 시스템을 쉽게 수정할 수 없는 경우도 있어, Wrapping 시스템을 이용하여 기존의 기록물 검색 메시지 형태를 EAD 프로토콜로 변환하여 KAIRS에 연동할 수 있도록 하였다. Wrapping 시스템 구조와 검색 절차는 다음과 같다. 통합검색 서버로부터

EAD 프로토콜 기반의 기록물 검색을 요청받은 Wrapper는 기록물 검색 요청 메시지를 분석하여 기존 시스템(Legacy System)에서 사용하는 형식으로 기록물 검색을 요청한다. 기존 시스템에서 검색된 기록물 정보는 다시 Wrapper를 통해 EAD 프로토콜로 변환되고 통합 검색 서버로 검색 결과가 전달된다(<그림 2> 참조).

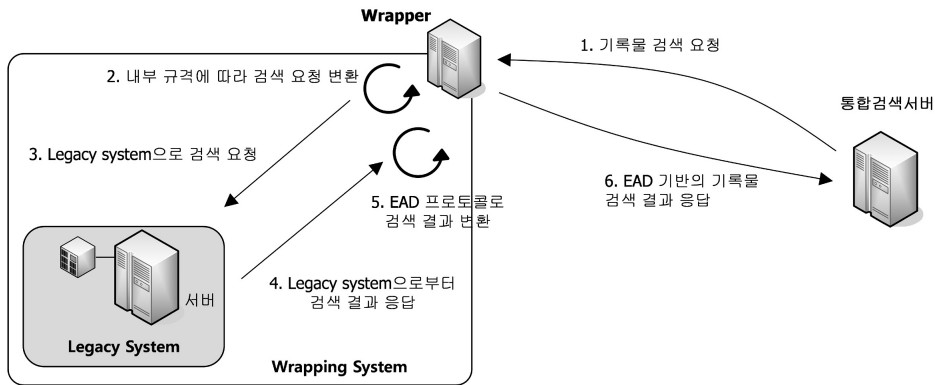
3.3.2 통합검색 서버의 소프트웨어 구조

통합검색 서버는 <그림 3>의 클래스 다이어그램과 같이 세부 기능별로 모듈화하여 다양한 클래스로 구성하고, 각 클래스간의 관계를 설계하여 통합검색 기능을 지원하도록 하였다. 각 클래스의 역할은 다음과 같다.

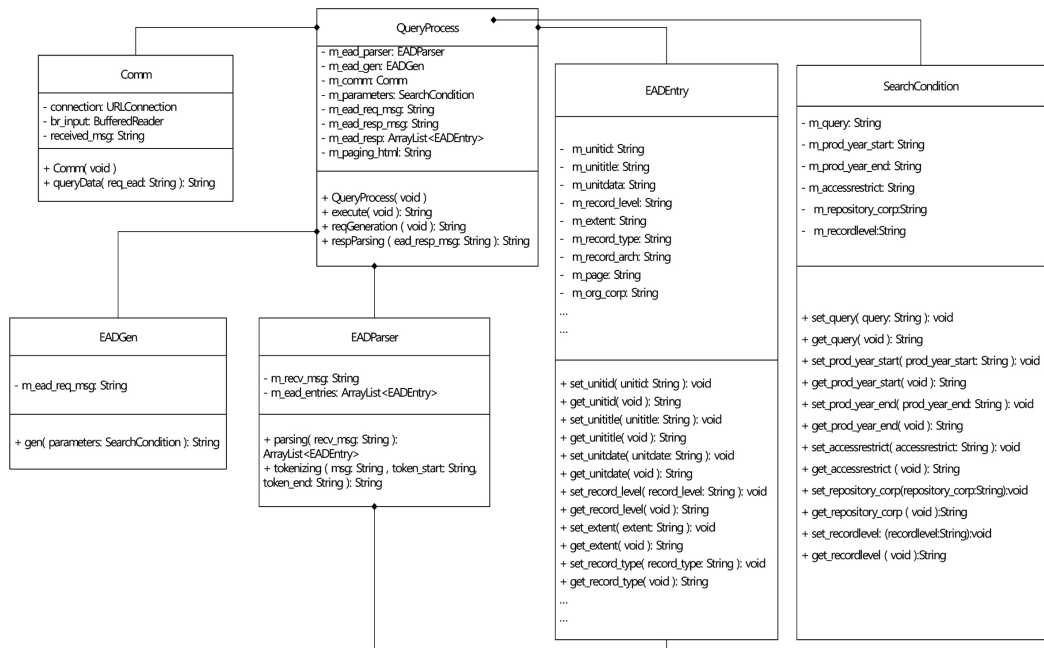
- ① QueryProcess - 이용자로부터 검색 요청을 받고 EAD 검색 요청 메시지를 이용하여 각 영구기록물관리기관의 기록물 검색을 수행



<그림 1> KAIRS 구조



〈그림 2〉 기존 영구기록물관리기관의 검색시스템을 위한 통합검색시스템 구조와 검색 절차



〈그림 3〉 통합검색 서버의 클래스다이어그램

- 한다. 또한 웹 응답페이지를 구성하여 이용
자에게 전달하는 역할을 담당한다.
- ② EADGen - 이 클래스는 이용자로부터 수
신된 검색 조건을 바탕으로 String 타입의
EAD 검색 요청 메시지를 생성하는 역할을

- 수행한다.
- ③ EADParser - 이 클래스는 영구기록물관
리기관으로부터 수신된 EAD 응답 메시
지를 분석하는 역할을 수행한다.
 - ④ Comm - 이 클래스는 영구기록물관리기

관과 통신을 수행하며 EAD 검색 요청 메시지와 EAD 응답 메시지를 교환하는 역할을 수행한다.

- ⑤ SearchCondition - 웹을 통한 이용자의 정보 검색 조건을 저장하고 관리하는 역할을 수행한다.
- ⑥ EADEntry - 각 기록물의 관련 정보를 저장, 관리하는 역할을 수행한다.

3.3.3 영구기록물관리기관 서버의 구조

영구기록물관리기관 서버는 <그림 4>의 클래스 다이어그램과 같이 세부 기능별로 모듈화하여 다양한 클래스로 구성하고, 각 클래스간의 관계를 설계하여 통합 검색 서버로부터의 검색 요청에 대한 응답 기능을 지원하도록 하였다. 각 클래스의 역할은 다음과 같다.

- ① QueryProcess - 통합검색 서버부터 EAD 기반의 기록물 검색 요청 메시지를 수신하고 데이터베이스의 기록물을 검색하는

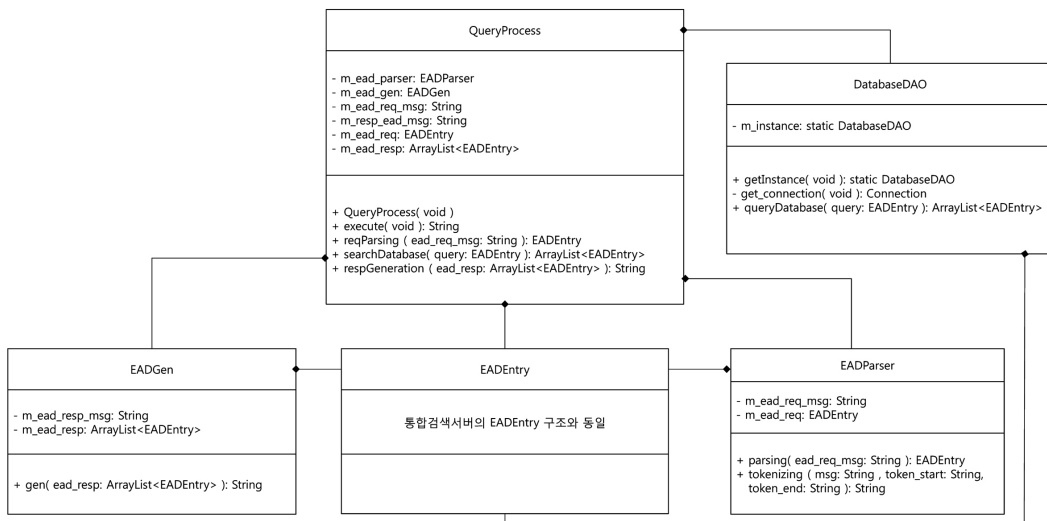
역할을 수행한다. 또한 검색 결과를 바탕으로 EAD 응답 메시지를 생성하여 통합 검색 서버로 송신하는 역할을 담당한다.

- ② EADGen - 데이터베이스에서 검색된 기록물 정보를 String 타입의 EAD 응답 메시지로 생성하는 역할을 수행한다.
- ③ EADParser - 통합 검색 서버로부터 수신된 "EAD 검색 요청 메시지"를 분석하는 역할을 수행한다.
- ④ DatabaseDAO - 통합 검색 서버로부터 수신된 검색 조건에 따라 데이터베이스에 접근하여 기록물들의 정보를 수집하는 역할을 담당한다.

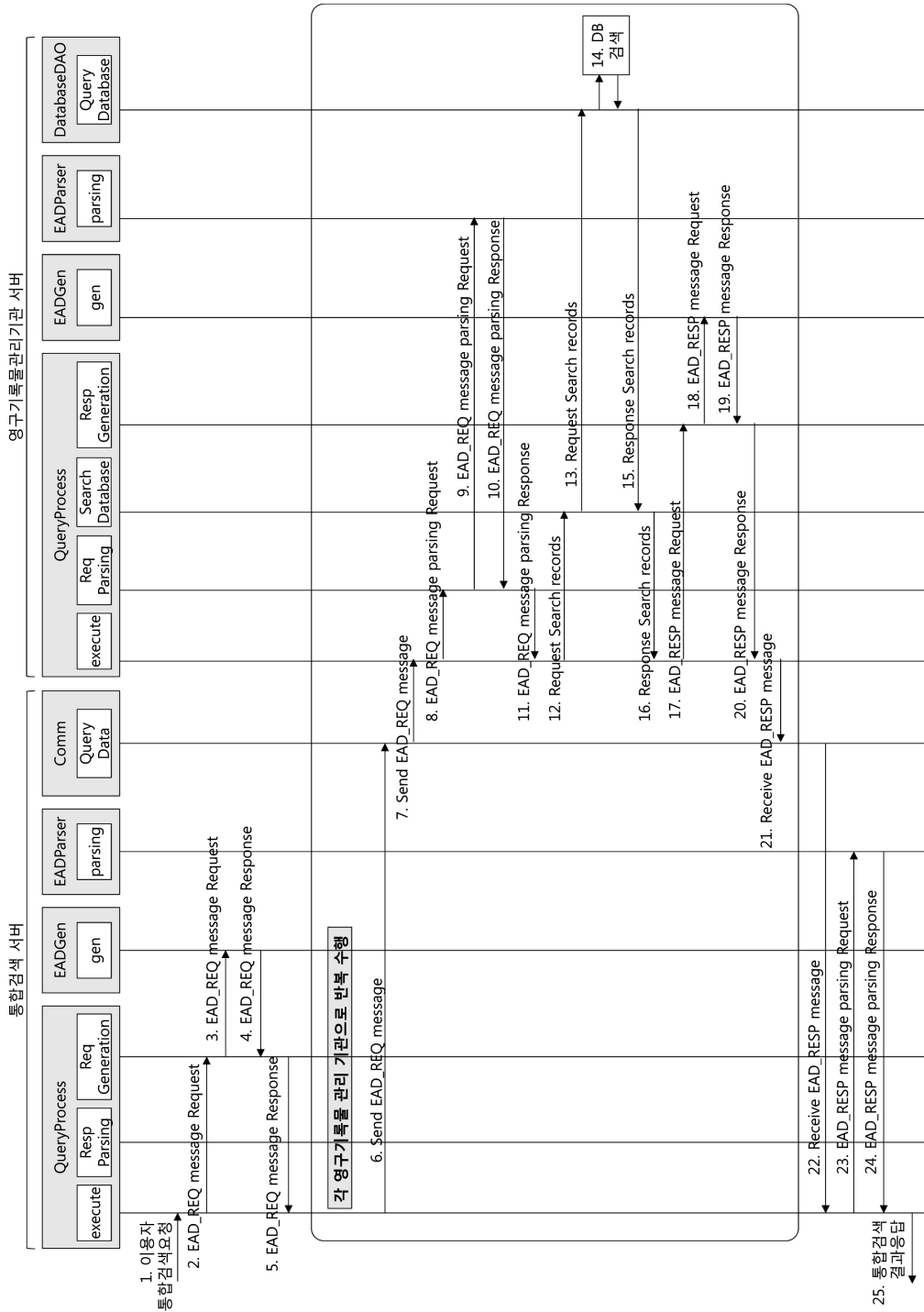
3.3.4 KAIRS의 검색 절차

KAIRS에서 이용자와 통합검색 서버, 영구 기록물관리기관 서버 사이의 전체 메시지 교환 절차를 정리하면 <그림 5>와 같다.

- ① 기록물의 검색 조건을 포함한 이용자 통



<그림 4> 영구기록물관리기관 서버의 클래스 다이어그램



<그림 5> KAIRS의 전체 검색 절차

- 합 검색 요청 받음으로써 QueryProcess 클래스의 execute 멤버함수가 호출된다.
- ②~⑤ 이용자로부터 전달된 검색 조건을 바탕으로 'EAD 검색 요청 메시지'를 String 형태로 변형한다.
 - ⑥ 통합 검색 서버는 영구기록물관리기관으로 'EAD 검색 요청 메시지'를 전송한다.
 - ⑦ 통합검색 서버로부터 'EAD 검색 요청 메시지'를 수신함으로써 QueryProcess 클래스의 execute 멤버함수가 호출된다.
 - ⑧~⑩ 통합검색 서버로부터 String으로 형성된 'EAD 검색 요청 메시지'의 검색 조건을 분석을 요청한다. 분석 결과를 EADEntry 형태로 구성하여 전달한다.
 - ⑪~⑬ KAIRS 통합검색 서버로부터 전달되어진 검색 조건을 바탕으로 DB의 기록물 정보 검색을 요청한다. 검색 결과를 이용하여 EADEntry 리스트 생성 및 응답한다.
 - ⑭~⑯ EADEntry 리스트로 생성된 기록물 검색 결과를 통합 검색 서버로 전달하기 위하여 String 형태의 'EAD 응답 메시지'로 변환한다.
 - ⑰ 'EAD 응답 메시지'를 포함하여 통합검색

서버로 검색 결과를 전송한다.

--- ⑥~⑰까지의 절차를 각 영구기록물관리기관으로 반복 수행된다. ---

⑱~㉔ 분석된 EADEntry의 리스트를 바탕으로 이용자의 웹으로 표시할 HTML 메시지 생성하여 이용자에게 검색 결과 페이지를 출력한다.

3.4 영구기록물 통합검색시스템(KAIRS) 구현

본 연구의 분석 대상인 국가기록원(CAMS 서버), 대통령기록관(PAMS 서버), 그리고 가상의 지방기록물관리기관(RAMS 서버)와 같은 기존 정부기관의 시스템은 수정/확장하기에는 어려움이 있어 테스트베드 구축을 위하여 3개의 영구기록물관리 기관 서버를 에뮬레이션하고 통합검색 서버를 개발하였다. 각 서버는 Ubuntu 12.04 운영체제를 사용하고 통합검색 서버와 영구기록물관리 기관 사이에서는 규격화된 EAD 프로토콜을 사용하여 정보를 교환하도록 하였다. 또한 통합 검색 서버는 apache tomcat과 struts MVC 모델을 이용한 웹 서버 기능을 지원하도록 하여 인터넷 환경에서 정보 검색이 가능하도록 하였다(〈표 3〉 참조).

〈표 3〉 테스트베드 구성 환경

| 파라미터 | | 속성 |
|--------------------|-----|----------------------------|
| 서버 하드웨어 구성 | CPU | Intel Dual Core |
| | 메모리 | 4Gbytes |
| 통합 검색 서버의 운영체제 | | Ubuntu 12.04 |
| 영구기록물관리기관 서버의 운영체제 | | Ubuntu 12.04 |
| 데이터베이스 | | CUBRID 2008 R14 |
| 통합 검색 웹 서버 구성 | | Apache Tomcat + struts 2.0 |

3.4.1 데이터베이스 구조

영구기록물관리기관에서는 기록물들의 저장 및 관리를 위한 데이터베이스를 운영하고 있으므로, 테스트베드에서는 CUBRID 2008 R14를 이용하여 국가기록원과 대통령기록관, 그리고 가상의 지방기록물관리기관의 데이터베이스를 독립적으로 구성하고 에뮬레이션하였다.

각 영구기록물관리기관의 기록물 데이터베이스는 내부시스템으로 외부에 공개되어 있지 않으며, 시스템 관리자만이 직접 접근 및 수정이 가능하기 때문에 국가기록원과 대통령기록관의 온라인 검색에서 제공되는 기술요소들을 바탕으로 각 영구기록물관리기관의 데이터베이스를 구성하였다. 또한 아직까지 존재하지 않는 지방기록물관리기관의 기록물 정보들은 국가기록원에서 관리하고, 지방기록물관리기관의 데이터베이스도 국가기록원의 기록물 정보를 바탕으로 구성하였다.

특히, 구현 시스템의 국가기록원 데이터베이스는 '대구 지하철 참사', '노무현 전화통화', '대구 육상 선수권 대회'라는 키워드로 검색했을 때 나오는 철/건 문서의 기술요소와 상위의 군/계열의 기술을 바탕으로 총 100개의 데이터를 입력하였다.

대통령기록관 데이터베이스는 '대구 지하철 참사', '노무현 전화통화', '대구 육상 선수권 대회'와 '대통령 선거'라는 키워드를 사용하여 임의의 데이터 50개를 입력하였다.

지방기록물관리기관의 기록물들은 아직까지 국가기록원에 소장되어 있고 2013년에 기술을 완료할 예정인 까닭에 구체적인 데이터를 제시할 수 없으므로 국가기록원의 기록물 검색 결과를 바탕으로 지방기록물관리기관의 데이터베

이스를 가상적으로 구성하였다. 국가기록원에서 기록물을 검색할 때 중앙부처 기록의 경우 소장기관이 국가기록원으로 나타나는 반면에 지방기록물의 경우에는 소장기관이 영구기록물관리기관으로 나타난다. 따라서 지방기록물관리기관의 기록물 데이터는 대구지방기록물관리기관에 한정하여 '대구 지하철 참사', '노무현 전화통화', '대구 육상 선수권 대회'와 '대구'라는 추가적인 키워드를 통하여 50개의 DB 데이터를 입력하였다.

4. 시스템 평가

4.1 시스템 구동 검증

시스템의 구동 검증은 기본적으로 이용자들이 접속하는 웹 환경에서 다양한 검색 조건을 입력하고, 이에 따른 결과 출력이 정확히 이루어지는지를 확인함으로써 검색시스템의 올바른 구동을 확인하였다.

기록물의 통합 검색을 위한 웹 환경에서의 이용자의 검색화면의 구성은 <그림 6>과 같다. 검색화면은 기존 시스템과는 달리 단순한 검색어 뿐 아니라, 관리기관, 생산년도, 공개여부, 철건구분의 추가적인 검색 조건을 제공하였다. 이는 국가기록원과 대통령기록관의 Open API에 공통적으로 사용되는 요소이면서 EAD 요소로 고려되고 있는 항목을 선택하였다. 관리기관은 전체, 국가기록원, 대통령기록관 그리고 지방기록물관리기관으로 선택할 수 있도록 하였으며, 공개여부는 전체, 공개, 비공개, 부분공개 중 선택하여 사용할 수 있도록 하였다.

록원과 대통령기록관의 기록물 통합검색 기능을 지원하고 있으나, 다양한 조건 검색 기능이 단순한 키워드 검색만을 지원하고 있다. 반면 통합검색시스템에서는 키워드와 함께, '관리기관, 생산년도, 공개여부'를 조건으로 하여 검색을 할 수 있다.

둘째, 기록물 검색결과 페이지에 대한 비교이다. <그림 7>과 <그림 8>은 각각 본 연구에서 구현된 시스템과 나라기록포털의 통합검색에 키워드 '노무현 전화통화'라는 키워드로 검색된 결과이다. 나라기록포털에서 검색된 화면의 기록물철과 기록물건은 국가기록원에서 지닌 기록물의 정보로서 기록물철, 건의 제목과 하단에 소장기관, 생산기관, 생산년도, 기록물 유형에 대해 나타낸다. 그리고 철, 건의 제목을 클릭하면 상세정보를 팝업하고, 팝업된 상세정보 페이지에서는 검색된 화면에서 나타나지 않은 보존기간, 문서번호 등의 정보가 나타난다. 나라기록포털에서 검색되는 대통령 기록물은 대통령기록관이 제공하는 Open API의 출력필드 값인 대통령명, 생산기관, 생산년도, 공개여부, 검색 결과 문서의 하이퍼텍스트 링크를 이용하고 있다. 검색된 화면 중 대통령 기록물을 보면 철, 건의 제목이 나타나고 하단에 대통령명, 생산기관, 생산년도, 공개여부가 나타난다. 이 중 생산년도는 대통령기록관에서는 생산일자로 '년도, 월, 일'을 이용한 8자리 숫자로 나타내지만, 나라기록포털에서는 년도만을 나타내기 때문에 검색화면에 나타나지 못하고 있다. 검색된 건 제목을 클릭하면 상세정보를 보여주기 위하여 Open API를 통하여 받은 대통령기록관의 검색 결과 페이지로 연결된다. 이 페이지에서는 상세정보로 검색페이지에서 제공하지 않은 기록

물의 생산일자, 기록물 페이지 정보, 기록물철, 건 ID를 제공하고 있다. 이에 비해 KAIRS에서는 검색페이지 화면에서 기록물의 철, 건 제목과 함께 상세정보를 함께 제공한다. KAIRS에서는 EAD를 이용하여 모든 기술요소를 받기 때문에, <그림 7>과 같이 각 기록관의 철, 건 제목과 하단에 다양한 기록물들의 정보를 제공할 수 있다. 이 시스템에서는 기록물 정보로 국가기록원, 대통령기록관의 철, 건 하단에 '소장기관, 소장기관, 생산년도, 기록물 유형, 기록물 형태, 공개여부, 보존기간, 관리번호'를 부여하여 이용자가 필요로 하는 기록물을 정확히 찾을 수 있도록 제공하였다. 그리고 각 기록관으로의 연결은 EAD의 <dao>를 통해 링크가 되도록 구성하였다.

셋째, 타 기관과의 연계를 지원하는 기능으로 국가기록원과 대통령기록관은 각각 Open API를 제공하고 있으며, 국가기록원의 나라기록포털은 대통령기록관의 Open API를 사용하여 통합 검색을 하고 있다. 그러나 이와 같은 통합검색을 위해서는 향후 설립될 지방기록물관리기관과 현재 온라인 검색을 지원하지 않는 헌법기록물관리기관들이 각각의 Open API를 정의하여 제공하여야 하며, 나라기록포털에서는 모든 기관들의 Open API를 사용하여 기록물의 통합 검색을 지원하는 구조가 되어야 한다. 반면, 본 연구에서 설계 구현한 KAIRS와 같이 EAD 표준의 데이터 구조를 사용하는 경우, 기존의 기관들은 Wrapping 시스템을 이용하여 기존의 기록물 검색 메시지 형태를 EAD 메시지 형태로 변환하고, 향후 설립되는 영구기록물관리기관의 서버들은 EAD 프로토콜을 지원하여 연동되는 형태가 될 수 있다.

4.2 사용성 평가

본 연구에서 구현된 KAIRS의 평가는 연구의 목적에 따라 전통적인 정보검색시스템의 성능 평가 척도인 정확도와 재현율을 사용하지 않고, 대신 이용자의 사용성 평가를 통하여 이루어졌다.

4.2.1 사용성 평가 방법

사용성 평가에 참여한 집단은 기록물 검색에 경험이 있는 기록학 전공자들과 다른 분야의 석·박사 과정 연구자들이었고, 이들은 현장 혹은 전자우편을 통하여 실험에 참여하였다.

현장 평가는 기록학 전공의 석사 과정 학생들과 다른 분야의 석·박사 과정 연구자로 이루어진 그룹으로 이들은 기존시스템(나라기록포털) 및 KAIRS의 연결이 가능한 컴퓨터 실습실 현장에서 기존시스템과 KAIRS에 대하여 설명을 듣고, 두 검색시스템을 각각 5분간 사용한 후, 설문지에 응답하는 방식을 택하였다.

전자우편 평가는 기록연구사들로 이루어진 그룹으로 이들은 전자메일을 통하여 현장 평가에서 제공된 기존시스템과 KAIRS에 대하여 설명 안내문을 숙독한 후, 기존시스템(나라기록포털) 및 KAIRS 두 시스템에 대한 인터넷 주소를 링크를 통하여 각각 5분간 사용한 후, 설문지에 응답하는 방식을 택하였다.

평가 기간은 2013년 11월 2일부터 11월 13일까지 12일간이며, 배포된 설문지 총 60부 중 50부를 회수하여 83%의 회수율을 보였다.

설문문항은 KAIRS의 기능, 검색방법, 검색결과, 이용자 인터페이스에 대한 이용자의 전반적인 만족도를 조사하는 4개의 문항과 전반적인

만족도를 세분화하고 기존시스템과 KAIRS에 대하여 비교하여 질문하는 문항들로 구성되어 있다. 이 중 이용자 인터페이스는 화면디스플레이에 관한 것으로 제한하였다. 설문문항 중 비교평가문항은 기존 시스템과 KAIRS의 검색 방법에 대한 3문항, 검색 결과에 대한 6문항, 이용자 인터페이스에 대한 3문항으로 구성되어 있고, KAIRS의 기능에 대한 비교평가문항은 검색 방법과 검색 결과의 문항에서 포괄적으로 다루고 있다.

4.2.2 설문조사 결과

KAIRS의 사용성에 관한 만족도 조사에서는 전반적인 검색 기능, 검색 방법, 검색 결과에 대하여 만족하였고, 이용자 인터페이스에 대하여 불만족하였다.

세부적인 만족도에 관한 질문 중 검색 방법에 대한 세부 만족도에 대한 설문에서는 대체적인 만족을 보였다. 또한 검색 결과에 대한 세부 만족도 질문 중 검색 결과 중에서 원하는 정보를 찾아내기에 얼마나 편리한지에 관한 설문 조사와 검색된 기록물 하단의 설명이 정확한 기록물을 찾는데 얼마나 도움을 주는지에 관한 설문 조사, 원하는 기록물을 찾는데 소요되는 시간에 대한 만족도 조사에서는 EAD를 기반의 KAIRS이 검색 기록물에 대한 더 많은 정보를 보여주고 있어 높은 만족도를 나타내었다. 반면 KAIRS의 설계 시 기능적인 측면만 고려하여 이용자 인터페이스에 대한 세부 만족도는 전반적으로 불만족으로 나타났다. 설문조사의 내용을 요약 정리하면 <표 4>와 같다.

〈표 4〉 KAIRS 사용성 평가 설문 결과

| 구분 | 문항 | 응답내용 | | | | |
|-----------------------|--|--|-----------|-----------|-----------|-----------|
| | | 매우 그렇다 | 그렇다 | 그저 그렇다 | 그렇지 않다 | 매우 그렇지 않다 |
| KAIRS 사용성에 관한 만족도 | KAIRS의 전반적 검색 기능에 대한 만족도 | 13명 (26%) | 25명 (50%) | 4명 (8%) | 5명 (10%) | 3명 (6%) |
| | KAIRS의 검색 방법에 대한 만족도 | 36명 (72%) | 7명 (14%) | 5명 (10%) | 2명 (4%) | 0명 (0%) |
| | KAIRS의 검색 결과에 대한 만족도 | 21명 (42%) | 23명 (46%) | 6명 (12%) | 0명 (0%) | 0명 (0%) |
| | KAIRS의 이용자 인터페이스(UI)에 대한 만족도 | 3명 (6%) | 6명 (12%) | 21명 (42%) | 18명 (36%) | 2명 (4%) |
| 검색 방법에 대한 만족도 | 기존시스템(나라기록포털)에 대비 검색 조건 이해의용이성 | 19명 (38%) | 24명 (48%) | 4명 (8%) | 2명 (4%) | 1명 (2%) |
| | 기존시스템 대비 검색 조건 필드의 만족도 | 32명 (64%) | 12명 (24%) | 3명 (6%) | 2명 (4%) | 1명 (2%) |
| | 검색 조건 필드의 불만족에 대한 의견 | ① 더 많은 검색 조건을 주어진다면 좋겠음 ② 날짜 선택하는 인터페이스가 불편함 | | | | |
| 검색 결과의 만족도 | 기존시스템 대비 검색 결과 정확성에 대한 만족도 | 20명 (40%) | 23명 (46%) | 3명 (6%) | 3명 (6%) | 1명 (2%) |
| | 기존시스템 대비 검색 결과 중에서 원하는 정보를 찾기의 편리성 | 41명 (82%) | 5명 (10%) | 2명 (4%) | 1명 (2%) | 1명 (2%) |
| | 기존시스템 대비 검색결과 배열에 따른 정보 찾기의 편리성 | 2명 (4%) | 16명 (32%) | 18명 (36%) | 10명 (20%) | 4명 (8%) |
| | 기존시스템 대비 검색된 기록물 하단의 설명이 정확한 기록물을 찾는 데 주는 도움정도 | 31명 (62%) | 10명 (20%) | 4명 (8%) | 3명 (6%) | 2명 (4%) |
| | KAIRS에서 원하는 기록물을 찾는 데 소요되는 시간에 대한 만족도 | 6명 (12%) | 30명 (60%) | 8명 (16%) | 2명 (4%) | 4명 (8%) |
| | KAIRS의 기록물 검색 결과에 대한 전반적인 의견 | ① 검색결과내 기존 시스템보다 세부적인 정보가 보여져서 기록물을 이해하기에 편리하였음 ② 검색결과와 재검색 기능이 추가되었으면 함 ③ 검색결과내에서 기관별 선택되면 좋겠음 ④ 기존시스템보다 자료 찾기가 편함 ⑤ 통합검색이 되어 하나의 사이트에서 이용할 수 있어 편리하였음 | | | | |
| 이용자 인터페이스(UI)에 대한 만족도 | 기존시스템 대비 검색을 위한 입력 인터페이스의 만족도 | 1명 (2%) | 10명 (20%) | 28명 (56%) | 8명 (16%) | 3명 (6%) |
| | 기존시스템 대비 검색 결과 출력 인터페이스의 만족도 | 2명 (4%) | 5명 (10%) | 21명 (42%) | 18명 (36%) | 4명 (8%) |
| | 통합검색시스템 이용자 인터페이스에 대한 전반적인 의견 | ① 생산년도 선택에 있어 마우스로 날짜를 클릭할 수 있었으면 좋겠음 ② 공개여부, 관리기관 선택을 선택하지 않아도 세부항목이 모두 보일 수 있도록 라디오박스 형태로 수정하였으면 함 ③ 검색어 입력 후, 검색 버튼을 마우스로 클릭하지 않고 엔터키로 검색이 수행되면 편리할 것 같음 ④ 검색페이지와 결과페이지에 이미지를 추가하여 조금 보기 좋게 구성하였으면 함 | | | | |

5. 결 론

본 연구의 목적은 기존 국내 영구기록물관리기관의 검색시스템을 하나의 통합된 검색과 상호 데이터 교환으로 그 사용성을 높이고자 새로운 통합검색시스템을 설계하고 구현하는 것이다. 연구의 목적을 달성하기 위하여 국가기록원과 대통령기록관의 기록물 기술요소와 온라인 검색 시스템을 분석하여 국내 실정에 적합한 EAD 요소를 선정하였고, 이를 바탕으로 새로운 영구기록물 통합검색시스템인 KAIRS를 설계하여 구현하였고, 시스템의 평가는 구동 검증과 사용성 평가를 통하여 이루어졌다.

본 연구의 결과를 요약하면 다음과 같다.

첫째, 영구기록물관리기관들의 기록물 정보 교환을 위한 EAD 데이터 구조 표준을 정립하고자 국가기록원과 대통령기록관의 외부 연동 현황을 분석한 결과 두 기관 모두 Open API를 제공하고 있으나, 프로토콜이 서로 다르게 구성되어 있어 모든 기술 정보의 전달에는 한계가 있다는 것을 확인하였다. 이에 따라 본 연구에서는 EAD를 기반으로 새로운 통합검색시스템을 설계하였다.

둘째, 통합검색시스템의 구현을 위하여 테스트

베드로 3개의 영구기록물관리기관과 가칭 한국영구기록물 통합검색시스템(KAIRS)이라는 시스템의 서버를 구성하였다. 또한 각 서버는 통합검색 서버와 영구기록물관리 기관 사이에서는 규격화된 EAD 프로토콜을 사용하여 정보를 교환하도록 하였다.

셋째, KAIRS의 구동 검증과 사용성 평가는 현장 평가와 전자메일 평가를 통하여 이루어졌다. 평가자들은 본 연구에서 새롭게 설계 구현된 KAIRS가 EAD 요소를 이용한 검색 조건, EAD 요소를 이용하여 풍부해진 검색된 기록물에 메타데이터, 검색 기능이 기존 시스템에 비하여 전반적으로 만족하였고, 이용자 인터페이스에 대해서는 기존 시스템에 비하여 다소 불만족스러운 것으로 나타났다.

앞으로 이와 유사한 영구기록물 통합검색시스템을 구축할 경우 고려되어야 할 점으로는, 기록물의 기술이 철과 건 단위까지 이루어져서 필수 요소를 충족시켜야 하고, 기록물 정보교환을 위하여 각 기관은 별도 프로토콜이 아닌 EAD를 이용한 프로토콜로 표준화가 전제되어야 한다는 것을 들 수 있다. 아울러 본 연구결과를 바탕으로 앞으로 영구기록물 통합검색시스템에 대한 연구가 더욱 활발하게 진행되길 기대해 본다.

참 고 문 헌

장소연 (2003). EAD의 구조와 적용에 관한 연구. 기록학연구, 8, 181-211.
공공데이터포털. 검색일자: 2013. 12. 17. <http://www.data.go.kr>
국가기록원 (2011). 영구기록물 기술규칙. 대전: 국가기록원.
국가기록원 (2013). 영구기록물 정리·기술 지침. 대전: 국가기록원.
국가기록원. 나라기록포털. 검색일자: 2013. 12. 20. <http://contents.archives.go.kr>

- 국가기록원. 대통령기록관 홈페이지. 검색일자: 2013. 12. 20. <http://www.pa.go.kr>
- 국가기록원 홈페이지. 검색일자: 2013. 12. 20. <http://www.archives.go.kr>
- 김지현 (2013). 대통령기록관의 검색, 열람 및 연구지원 서비스에 관한 연구. 한국기록관리학회지, 13(2), 201-225.
- 박은경 (2002). EAD를 이용한 기록자료 자동화방안: Online Archive of California 프로젝트를 중심으로. 한국기록관리학회지, 2(2), 93-106.
- 박지영, 김태수 (2007). EAD를 이용한 기록정보의 기술 및 활용: 4월 혁명 연구반 컬렉션을 중심으로. 지식처리연구, 8(1/2), 17-57.
- 박현이, 정은경 (2011). 국내 영구기록물의 EAD 기술요소에 관한 연구. 한국기록관리학회지, 11(2), 33-55.
- 설문원 (2010). 기록 검색도구의 발전과 전망. 기록학연구, 23, 3-34.
- 장희정 (2012). 국가기록관 웹사이트의 평가에 관한 연구. 한국기록관리학회지, 12(2), 51-70.
- 조윤희, 이두영 (2001). XML/KORMARC 시스템 구현에 관한 연구. 정보관리학회지, 18(4), 217-233.
- Brown, G. & Harvey, K. (2007). Adding archival finding aids to the library catalogue: Simple crosswalk or data traffic jam? Partnership: the Canadian Journal of Library and Information Practice and Research, 2(2), 1-18.
- Pitti, D. V. (1997). Encoded archival description: The development of an encoding standard for archival finding aids. American Archivist, 60(3), 268-283.
- SAA (2003). Encoded Archival Description Tag Library Version 2002. Chicago: The Society of American Archivists.
- Stocking, B. (2004). Time to settle down? EAD encoding principles in the access to archives programme (A2A) and the research libraries group's best practice guidelines. Journal of Archival Organization, 2(3), 7-24.
- Thurman, A. C. (2005). Metadata standards for archival control: An introduction to EAD and EAC. Cataloging & Classification Quarterly, 40(3), 183-212.
- Hensen, S. L. (1997). NISTF II and EAD: The evolution of archival description. American Archivist, 60(3), 284-296.

[정보공개청구 자료]

- 한국. 안전행정부 (2013a). 나라기록포털의 검색시스템의 기술적(technical)사항.
- 한국. 안전행정부 (2013b). 국가기록원 계층별 검색의 기술요소.
- 한국. 안전행정부 (2013c). 대통령기록관의 기록물 기술요소.
- 한국. 안전행정부 (2013d). 대통령기록관의 검색시스템의 기술적(technical)사항.

• 국문 참고자료의 영어 표기

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

- Cho, Yoon-Hee & Lee, Too-Young (2001). A study on implementation of XML/KORMARC system. *Journal of Korean Society for Information Management*, 18(4), 217-233.
- Jang, Hee-Jung (2012). A study on evaluation of national archives websites. *Journal of Korean Society of Archives and Records Management*, 12(2), 51-70.
- Kang, So-Youn (2003). The study of structure and application of EAD. *The Korean Journal of Archival Studies*, 8, 181-211.
- Kim, Ji-hyun (2013). A study on research services of presidential archives in Korea. *Journal of Korean Society of Archives and Records Management*, 13(2), 201-225.
- NAK (2011). *Archival Description Rules, Version 2.0. NAK/S 14:2011(v2.0)*. Daejon: NAK.
- NAK (2013). *Guidelines for Archives Arrangement & Description*. Daejon: NAK.
- Park, Eun-Kyung (2002). Constructing a prototype union database of encoded archival descriptions: a case of the online archive of california project. *Journal of Korean Society of Archives and Records Management*, 2(2), 93-106.
- Park, Hyun-Yi & Chung Yeon-Kyung (2011). A study on encoded archival description (EAD) elements for the archival institutions in Korea. *Journal of Korean Society of Archives and Records Management*, 11(2), 33-55.
- Park, Zi-young & Kim, Tae-soo (2007). Implementation of EAD 2002 for 4.19 collection. *Journal of Knowledge Processing and Management*, 8(1/2), 17-57.
- Seol, Moon-Won (2010). A study on development and prospects of archival finding aids. *The Korean Journal of Archival Studies*, 23, 3-34.