

디지털 연구성과물의 장기보존을 위한 메타데이터에 관한 연구*

A Study of Metadata for Long-Term Preservation of Digital Research Outcome

장 보 성(Bo-Seong Jang)**

남 영 준(Young-Joon Nam)***

〈 목 차 〉

I. 서론	IV. 디지털 연구성과물 장기보존을 위한 메타데이터 요소 설계
II. 이론적 배경	V. 결론 및 제언
III. 국내외 연구성과물 장기보존 관련 메타데이터 사례 분석	

초 록

이 연구는 전자기록물의 첨부물 형태로 수집·관리되는 디지털 연구성과물의 장기보존을 위한 메타데이터요소를 설계하여 제안하였다. 설계는 국내외 연구성과물을 관리하는 기관의 메타데이터 구조와 CERIF와 PREMIS 보존 메타데이터를 참고하였다. 연구결과, 우리나라 디지털 연구성과물의 장기보존을 위한 메타데이터 구조를 연구정보, 서지정보, 기록정보로 구분하였다. 연구정보는 과제, 인명, 기관으로 구분하였으며 8개 상위요소, 29개 하위요소로 구성하였다. 서지정보는 연구성과물인 연구보고서의 서지사항에 대한 요소로써 하위요소 없이 13개의 상위요소로 구성하였다. 기록정보는 연구성과물의 생산을 보고한 전자기록물의 맥락, 구조, 내용, 관리과정에 대한 요소로써, 18개 상위요소와 62개 하위요소로 제안되었다.

키워드: 디지털, 연구성과물, 장기보존, 아카이빙, 메타데이터

ABSTRACT

This research designed and proposed metadata elements for the long-term preservation of digital research outcomes collected and managed in electronic file form. Metadata structure of both domestic and foreign agencies that manage research outcomes and CERIF and PREMIS-preserved preservation metadata were referred to design the metadata elements. The metadata design was divided into research information, bibliographic information and archival information. Research information is divided into project, name and institution, and composed of eight top-level elements and 29 sub-level elements. Bibliographic information is composed of 13 top-level elements as the source of research reports. Archival information is composed of 18 top-level elements and 62 sub-level elements as the elements for context, structure, contents and management process of the electronic records.

Keywords: Digital, Research Outcome, Long-Term Preservation, Archiveing, Metadata

* 본 연구는 중앙대학교 대학원 박사학위 논문을 축약한 내용임.

** 중앙대학교 대학원 기록관리학과(club301@korea.kr) (제1저자)

*** 중앙대학교 문헌정보학과 교수(namyj@cau.ac.kr) (교신저자)

• 접수일: 2011년 11월 25일 • 최종심사일: 2011년 12월 6일 • 최종심사일: 2011년 12월 28일

I. 서론

1. 연구의 필요성과 목적

우리나라는 과학기술분야의 국가경쟁력 강화를 위해 국가 연구개발(이하 R&D)사업을 지속적으로 확대하고 있다. 국가 R&D투자는 2000년 4.2조원에서 2005년 7.8조원, 2010년 13.7조원으로 증가하였다.¹⁾ 이와 같은 정부의 막대한 예산 투자에도 불구하고 국가차원에서 국가 R&D사업의 결과인 연구성과물에 대한 수집, 관리 및 보존이 제대로 이루어지지 않고 있다. 2010년 기준으로 정부차원에서 지급한 국가 R&D과제 중 수집 대상 연구성과물의 20%만이 NTIS(국가과학기술지식정보서비스, National Science & Technology Information Service)에 수집되어 관리되고 있다.²⁾

한편, 연구성과물은 기록관리학 관점에서 중요한 기록물 및 지식정보자원으로 관리의 필요성과 함께 국가기록원에서 반드시 수집·관리되어야 할 공공기록물이다. 왜냐하면, 연구성과물은 법체계 내에 공적 재원에 의하여 형성되는 공공정보에 해당하기 때문이다. 「공공기록물 관리에 관한 법률」(이하 기록물관리법) 시행령 제17조(조사, 연구서 또는 검토서의 작성)에서 법령 제·개정, 주요 정책의 결정, 국가가 수행하는 대규모 사업 등을 시행하는 경우 사전에 조사·연구·검토서를 작성하도록 규정하고 있다. 조사·연구·검토서는 국가가 생산하는 조사서, 연구서, 검토서, 연구용역보고서 및 정책연구용역 수행과정에서 발간되는 모든 결과물 및 최종결과물 등을 의미한다.³⁾ 이러한 당위성과 법적 근거에도 불구하고 국가기록원을 비롯하여 관련 유관기관(기록관)에서 연구성과물에 대한 체계적인 수집·관리·보존이 이루어지지 않고 있다. 주요 원인으로는 정부차원에서 연구과제의 발주와 그 연구성과물에 대한 관리지침의 부재, 특히 관리 및 보존용 메타데이터 요소의 부재가 커다란 원인이다.

현재 국가기록원 차원에서 운용중인 기록관리 메타데이터 요소들은 시청각기록물, 웹기록물에 대하여 일부 별도로 운영하고 있으며, 전자기록물에 대한 메타데이터는 기술적(descriptive) 측면이 강조되어 있다. 즉, 관리적(management) 측면과 보존적(archival) 측면에 대한 것은 상대적으로 취약한 수준이다. 예를 들어, 연구성과물이 지니고 있는 기본적인 속성정보를 국가기록원의 '연구기록물 기술규칙(NAK/S 14:2011(v2.0))'으로 표현하기에는 한계가 있다. 즉 과제번호, 주관기관명, 연구비 등의 요소 값을 기술할 수 있는 메타데이터 요소가 부족하며, 간행물로 취급하여 분리 등록할 경우 책자형태의 간행물과 문서형태의 전자기록물과의 연관관계를 찾기가 어렵다.

1) 한국법제연구원, 수요자 중심의 국가 R&D사업 관리제도 구축을 위한 법제연구(서울 : 교육과학기술부, 2010), p.17.

2) 우지숙, 공공기금에 의한 연구성과물의 공공접근정책 연구(대전 : 한국과학기술정보연구원, 2011), p.62.

3) 국가기록원, 2010 기록관 기록관리업무(대전 : 국가기록원, 2010), p.54.

따라서 본 연구는 전자기록물의 첨부물 형태로 관리되는 디지털 연구성과물의 장기보존을 위한 메타데이터 요소를 도출하는 것이다. 장기적인 관점에서 디지털 연구성과물의 안정적이고, 체계적인 관리를 위해 디지털 연구성과물을 관리·기술·보존할 수 있는 메타데이터 요소를 설계하고자 한다.

2. 연구방법

본 연구는 디지털 연구성과물 장기보존을 위한 메타데이터 요소를 설계하고자 국내외 연구성과물 관리 시스템 및 보존을 위한 메타데이터 표준을 비교·분석하고, 1차적으로 공통된 메타데이터 요소를 도출하였다. 이후 문헌연구와 현장의 사례를 반영하여 추가 및 삭제 대상 요소를 제안하였다. 이후 전문가 집단을 통한 델파이조사를 실시하여 디지털 연구성과물 장기보존을 위한 메타데이터 최종(안)을 제안하였다.

II. 이론적 배경

1. 연구성과물의 개념

법률, 국가기관, 및 학술 연구 분야의 ‘연구성과물’에 대한 정의는 다음과 같다.

교육과학기술부 훈령 제111호, 『인문사회분야 학술연구지원사업 처리규정』에 따르면 ‘연구성과물’이란 연구성과 중 연구자가 수집·작성한 일체의 자료로서 연구결과보고와 연구결과물 및 원자료와 중간산출물 등을 의미한다.⁴⁾ 한국연구재단 『학술연구과제관리규칙』에 따르면 ‘연구성과물’이란 연구성과 중 연구자가 수집·작성한 일체의 자료를 말하는데 원자료와 중간산출물, 연구결과보고서, 연구결과발표물 등으로 구성된다. 한국과학기술정보연구원 R&D보고서원문성과물관리시스템의 등록대상이 되는 ‘보고서원문’은 국가연구개발사업의 과제수행 후, 발생하는 성과물 중 보고서 파일(연차보고서, 중간보고서, 최종보고서 등)을 의미한다.⁵⁾ 교육과학기술부 『2009 연구성과 메뉴얼』에서 ‘연구성과물’이란 연구개발의 최종목표 혹은 연구수행 과정에서 얻어지는 결과물을 의미한다.⁶⁾

4) 교육과학기술부 훈령 제111호, 인문사회분야 학술연구지원사업 처리규정, 제2조.

5) 한국과학기술정보연구원 홈페이지, <<http://report.kisti.re.kr/main.jsp>> [인용 2011. 4. 20].

6) 교육과학기술부, 2009 연구관리 메뉴얼(서울 : 교육과학기술부, 2009), p.10.

장덕현 등(2008)은 ‘연구성과물’이란 연구자의 학술활동 및 연구개발로 산출된 모든 형태의 정보자원이며, 연구보고서를 생산, 납본하는 조건으로 기금을 지원받아 연구 활동을 하고, 그 과정에서 생산된 연구물을 의미한다.⁷⁾

Michael Nentwich(2002)는 ‘연구성과물’이란 연구형태와 상관없이 새로운 지식관점의 연구결과로써, 일반적인 용어상 개념은 연구와 연구성과 두 가지를 함께 지니고 있다.⁸⁾

이와 같이 ‘연구성과물’은 연구과정과 결과로부터 생산되는 창의적이고 가치 있는 연구물로 정의하고 있다. 따라서, 본 연구에서는 ‘연구성과물’이란 연구과정과 결과로부터 생산되는 독자적이고 가치 있는 지식의 산출물이다. 이때 연구성과물은 공공기금의 예산 출처, 수행 주체에 따라 구분되는 것이 아닌, 모든 연구 수행 과정에서 또는 결과에서 얻어지는 산출물을 의미한다. 이러한 연구성과물은 정책연구용역, R&D사업 등을 준비하는 연구과제책임자나 연구기관이 참고하는 실질적인 수요대상정보이며, 연구의 과제정보 및 연구보고서, 논문, 특허 등의 내용이 포함되는 종합성과정보이다.

한편, 교육과학기술부는 2009년 발간한 『연구성과 관리메뉴얼』에서 연구성과물 유형을 8개로 확대·개편하였다.⁹⁾ 그 유형은 <표 1>과 같다.

<표 1> 연구성과물 유형

유형	내용
논문	국내외 학술단체 및 출판사에서 발간하는 학술지 및 학술대회지에 수록된 학술 논문
특허	국내외에 출원 또는 등록된 특허정보
보고서원문	연구개발 종료 시 제출하는 최종보고서 및 연차보고서
연구기자재	국가연구개발사업의 수행 시 취득한 장비 중 취득가격이 3천만원 이상인 장비 또는 취득가격이 3천만원 미만이라도 공동활용이 가능한 장비
기술요약정보	기초·응용·개발단계 등의 최종보고 및 연차보고가 완료된 결과물의 기술정보를 요약하여 공유·활용할 수 있도록 작성된 기록정보
생명자원	유전체정보, 단백질체정보, 발현체 정보, 미생물 자원, 동물자원 식물자원 등 관련정보
소프트웨어	창작된 소프트웨어 및 등록에 필요한 관련정보
화합물	합성 또는 천연물에서 추출한 유기화합물 및 관련 정보

본 연구에서는 <표 1>의 8가지 연구성과물의 유형 중 ‘보고서원문(이하 연구성과물은 연구보고서를 의미함)’을 중심으로 연구를 진행하였다. 보고서원문은 산업계, 학계, 연구계에서 생산되는 각

7) 장덕현, 노류하, 이수상, “R&D 성과물 공동활용에 대한 연구자들의 인식에 관한 연구,” 정보관리연구, 제39권, 제4호(2008), pp.5-6.

8) Michael Nentwich, “Methods and research outcome in the age of cyberscience,” *The Impact of ICTs on Research Practice*, 2002, p.3, <http://www.oeaw.ac.at/ita/ebene5/Nentwich_AoIR2002.pdf> [cited 2011. 6. 3].

9) 교육과학기술부, 전계서, p.14.

중 R&D사업의 최신 연구경향과 성과를 알 수 있는 유용한 수단이다. 또한 관련분야의 전문가에게 가치 있는 정보를 제공하는 유용한 정보자원이며, 서지정보를 확보하기 어려운, 유통구조가 공식화 되지 않은 대표적인 회색문헌이다.

2. 연구성과물의 특성

기록관리적 관점에서 디지털 연구성과물은 일종의 전자기록물이며, 기록관리법에서 명시하고 있는 기록물이다.¹⁰⁾ 디지털 연구성과물은 연구주관기관에서 전자기록물을 통해 그 생산현황을 연구 발주기관으로 보고하며, 발주기관은 전자기록물을 접수하여 전자문서관리시스템(예, 통합 온-나라 시스템) 등을 통해 관리한다. 기록물 등록은 전자 또는 비전자기록물의 구분 없이 전자기록생산시스템으로 등록·관리하는 것을 원칙으로 하고 있다. 기록물 등록에는 전자기록물 자동등록, 비전자 기록물 수기등록, 첨부물 분리등록으로 구분하고 있다. 연구성과물과 같이 첨부물 형태의 기록물은 본문과 형태나 규격 등이 달라 문서와 같이 등록하지 않고 분리 등록하고 있다.¹¹⁾

첨부물 형태의 디지털 연구성과물은 연구성과물만의 고유한 속성정보를 지니고 있다. 『국가연구개발사업의 관리 등에 관한 규정 시행규칙』 [별지 제6호 서식]에서 국가 연구개발사업의 최종보고서 및 요약서에 포함된 기본정보 및 관련 정보에 대한 기본 서식을 제공하고 있다.

〈표 2〉 국가연구개발사업의 관리 등에 관한 규정 시행규칙 [별지 제6호서식]

표지	과제번호
	세부사업명 또는 중과제명
	세부과제명
	세부과제 연구기관명
	주관기관명
제출문	주관연구기관명
	주관연구책임자
	연구원
	협동연구기관명
	협동연구책임자

- 10) 공공기록물 관리에 관한 법률 제2조 '기록물'이란 공공기관이 업무와 관련하여 생산 또는 접수한 문서, 도서, 대장, 카드, 도면, 시청각물, 전자기록물 등 모든 형태의 기록정보 자료와 행정박물을 말한다. 동법 시행령 제2조. '전자기록물'이란 일반적으로 컴퓨터를 이용하여 생산하고, 저장 및 관리되는 기록물이다.
- 11) 현재 각급기관에서 사용하고 있는 전자기록물 생산을 위한 정보시스템은 '통합 온-나라 시스템', '전자기록물시스템', '행정정보시스템' 등이 있음. 중앙행정기관은 '통합 온-나라시스템'을 사용하고 있고, 지방자치단체 등은 단계적으로 '전자기록물시스템'에서 '통합 온-나라 시스템'으로 전환하고 있음(참고: 국가기록원, 2010 기록관 기록관리실무(대전 : 국가기록원, 2010), p.16).

6 한국도서관·정보학회지(제42권 제4호)

보고서 요약서	과제고유번호	
	해당단계 연구기간	
	단계구분(해당단계/총단계)	
	연구사업명	중사업명
		세부사업명
	연구과제명	대과제명
		세부과제명
	연구책임자	해당단계 참여연구원수
		해당단계 연구비
		총연구기간 참여연구원수
		총 연구비
	연구기관명 및 소속부서명	
	참여기업명	
	국제공동연구	상대국 명
		상대국 연구기관명
	위탁연구	연구기관명
연구책임자		
요약		
보고서 면수		
색인어	한글	
	영어	

〈표 2〉와 같이 연구성과물은 과제번호, 연구과제명, 연구사업명, 연구책임자, 연구기관명 등의 연구 수행과제에 대한 기본정보를 담고 있다. 이러한 정보들은 연구성과물을 관리, 식별, 계층화하는데 있어서 중요한 정보원이자, 보존을 위한 메타데이터 요소가 된다.¹²⁾ 그러나 현재 국가기록원은 연구성과물을 간행물로 취급하여 ‘연구기록물 기술규칙’에 따라 기록물관리시스템에 등록하고 있다. 이때 등록을 위한 기본목록은 간행물명, 발간기관, 발간일자, 권/호, 쪽수, 입수방법(제출, 구입, 기증, 기타), 형태구분(책자형태, 파일형태, CD-Title, CD+책자, 책자+파일, CD+파일)이다. 세부목록은 저자, 공저자, 원저자, 편저자, 역자 등의 저자사항, 대등서명, 부서명, 원서명, 총서명, 잡제 등의 서명사항, 유형구분, 내용언어, 형태사항 등의 상세사항과 간행물의 목차를 입력하고 있다. 이처럼 연구성과물이 지니고 있는 기본적인 속성정보를 ‘연구기록물 기술규칙’으로 표현하기에는 한계가 있으며, 간행물로 취급하여 분리등록 할 경우, 간행물과 전자기록물과의 연관관계 설정 등의 메타데이터 요소 설계가 필요하다.

12) 『국가연구개발사업의 관리 등에 관한 규정』(제12조 1항) 주관연구기관의 장은 연구개발이 종료된 때에는 연구개발 최종보고서·요약서 및 주관연구기관의 자체평가의견서와 그 전자기록물을 중앙행정기관의 장에게 제출하여야 한다. (제12조제4항) 제1항에 따른 최종보고서·요약서, 주관연구기관의 자체평가의견서의 서식 및 제출방법에 관하여 필요한 사항은 교육과학기술부령으로 정한다.

3. 선행연구

전자기록물로서 연구성과물에 대한 보존 관련연구는 공공부분의 정부문서나 대학의 학술연구 논문, 역사적 가치가 있는 사진, 이미지 등을 중심으로 이루어지고, 국가 단위의 연구성과물을 대상으로 보존에 관한 연구는 상대적으로 미흡하다. 이 절에서는 관련된 연구를 국내와 국외로 구분하여 조사·분석하였다.

가. 국내

초기 국내 연구성과물에 대한 연구는 효율적인 관리와 그 효용성을 제고하는 방안이 대부분이다. 특히 연구성과물의 검색·활용을 위한 표준 메타데이터 구축 및 시스템 개발 관한 연구들이다.

박동진(2005)은 관리대상이 되는 연구성과물의 범위를 확장하고 이들을 통합한 표준 메타데이터를 제시하였다. 그는 RDF/RDFS를 기반으로 연구성과물 검색과 유통에 있어서 지능적인 정보 서비스가 가능하도록 설계하였다.¹³⁾ 또한 후속연구에서 과학기술분야의 연구성과물 정보 표준 메타데이터를 개발하였다.¹⁴⁾ 하지만 연구성과물의 표준 메타데이터 요소들이 해당 연구성과물의 관리를 위한 관리적 메타데이터라는 한계점을 지니고 있다.

한국과학기술정보연구원(2007)은 2006년 정보화전략기획사업 추진 후 NTIS 및 응용시스템을 개발하였다.¹⁵⁾ 이는 국내에서 처음으로 연구성과물 및 연구과정 전반에 대한 종합 관리시스템을 구축한 것이다. 하지만, 국가 R&D 사업으로 그 범위를 제한하였다는 점과 초기 시스템 구축과정에서 보존을 제외한 수집과 관리 부분만 설계되었다는 한계점을 지니고 있다.

다음으로 연구성과물에 대한 기록관리적 관점에서의 보존 연구는 최근에 이루어지고 있다. 정세영(2007), 여상아(2007), 천누리(2010)의 연구는 정부출연연구기관의 연구개발 기록물(사업 계획서부터 최종 성과물까지 모두 포함)의 관리 실태를 조사·분석하고 그 해결방안으로 ‘기록관리 포털 시스템’ 도입, 기록관리 절차에 따른 기록물 관리 및 관련 메타데이터 요소 개발을 제안하였다. 하지만 대상 연구 모두 단순한 실태 조사와 이를 관리하기 위한 시스템 도입 및 기록물 관리, 기록물 기술규칙에 기반 한 메타데이터 요소를 제안하였기에 연구의 한계점을 지니고 있다.¹⁶⁾

13) 박동진, “RDF/RDFS를 이용한 연구성과물정보 메타데이터 모델링에 관한 연구,” 한국디지털정보학회 학술대회 발표논문집, 제2005권, 제1호(2005), pp.383-389.

14) 박동진, “과학기술분야의 연구성과물정보 표준 메타데이터 개발에 관한 연구,” 산업경영시스템학회지, 제29권, 제4호(2006), pp.83-90.

15) 한국과학기술정보연구원, 국가과학기술종합정보시스템 구축 사업(대전 : 한국과학기술정보연구원, 2007).

16) 정세영, 연구개발기록물 관리체계 분석 및 발전방안 연구(석사학위논문, 충남대학교 대학원 기록보존학과, 2007). ; 여상아, 과학기술분야 정부출연연구기관의 연구기록관리 개선 방안 - K연구소를 중심으로(석사학위논문, 서울대학교 대학원 기록관리학과, 2007). ; 천누리, 연구개발기록물의 메타데이터 요소에 관한 연구(석사학위논문, 중앙대학교 대학원, 기록관리학과, 2010).

즉 연구성과물의 유형별 특성과 구체적인 시스템 설계 방안 및 세부 절차의 논리적 근거 등이 부족하다.

김관준(2010)은 ‘연구정보’를 위한 보존 메타데이터 요소를 개발하고자, CERIF(The Common European Research Information Format, 이하 CERIF)와 PREMIS 데이터 사전의 요소들을 비교·분석한 다음, 양자의 특성을 반영한 보존 메타데이터를 요소를 개발하였다.¹⁷⁾ 연구과제에 대한 보존 메타데이터 요소를 도출하였다는 점에서 의의가 있으나, CERIF의 과제, 산출물을 PREMIS의 객체 개체와, 인명과 기관을 PREMIS의 행위자 개체와 대응관계를 설정한 논리적 개연성이 부족하고, 산출물(발간물)의 특정 상위요소 메타데이터 하나만을 비교대상으로 설정하였기에 메타데이터 전개에도 한계가 있다.

이와 같이 국내 선행연구들은 연구성과물의 효율적인 관리를 위한 메타데이터 요소 개발 연구가 이루어졌으며, 주로 인쇄형태의 연구성과물과 하나의 과제에서 생산된 다양한 연구개발 기록물을 대상으로 하고 있다.

이에 본 연구에서는 기록관리학 측면에서 인쇄자료가 아닌 디지털 연구성과물을 연구대상으로 삼았으며, 국가 기관에서 연구성과물을 생산·보고한 전자문서와 그 문서에 별첨으로 포함되어 있는 디지털 연구성과물의 장기보존을 위한 메타데이터에 대한 연구를 수행한다는 점에 있어서, 선행 연구와의 차별성을 지니고 있다.

나. 국외

국외 디지털자원에 대한 장기보존 연구는 활발하게 진행되고 있으나, 연구성과물이라는 특정 객체에 대한 연구는 상대적으로 미비한 실정이다.

Oxnam(2010)은 애리조나 주립대학교의 주도하에 연구보고서의 완전한 수집 및 보존을 위해 설립된 TRAIL(Technical Report Archive & Image Library)의 역할과 구조, 미래지향점에 대해서 기술하였다. 그는 보존을 위한 망라적 수집, 예산 지원, 기관 협력 및 기록물 평가 등을 주요 내용으로 기술하였다.¹⁸⁾

Caplan(2010)는 플로리다 주립대학교의 장기보존을 위한 리포지토리 인 플로리다 디지털아카이브(FDA)와 FDA에서 사용 중인 DAITSS(Dark Archive In The Sunshine State) 응용프로그램을 기반 한 가능성과 편이성에 대해 설명하였다. 이것은 대학교에서 자체적으로 발간한 연구성과물에 대한 장기보존 모델이다. DAITSS에 의해 구현된 보존 프로토콜은 비트 수준의 보존, 형식 정형화, 미래를 위한 마이그레이션 전략, 광범위한 보존 및 특정 메타데이터 지원, 진본성 유지로

17) 김관준, “연구정보를 위한 보존 메타데이터 요소 개발에 관한 연구: 경제·인문사회연구회 연구관리시스템을 중심으로,” 정보관리학회지, 제27권, 제4호(2010), pp.169-191.

18) Maliaca A. Oxnam, “Multi-Institutional Approach to Technical Report Literature: Development of the Technical Report Archive & Image Library,” *Grey Journal(TGJ)*, Vol.6, No.1(2010), pp.5-7.

구분할 수 있다.¹⁹⁾

한편 PREMIS를 적용한 해외 아카이빙 관련 선행연구의 대상 컬렉션으로 대학이나 연구기관의 논문이나 연구물,²⁰⁾ 연구기관의 연구결과 발생한 데이터 세트²¹⁾ 등이 있다.

이와 같이 해외 선행연구 또한 기록관리측면에서의 연구성과물 장기보존에 대한 연구는 미흡한 실정이다. 다만 연구성과물의 완전한 수집과 보존을 위해 설립된 TRAIL의 사례에서, 또는 대학교에서 자체적으로 연구성과물을 장기보존 한 모델에서 연구성과물의 보존을 위한 필수 조건들을 참고할 수 있다.

Ⅲ. 국내외 연구성과물 장기보존 관련 메타데이터 사례 분석

이 장에서는 국내 연구성과물을 관리하기 위한 관련 시스템의 메타데이터 구조와 미국 NTIS 산하 NTRL(National Technical Reports Library) 시스템, 유럽 연합 기록관리 메타데이터 모형, 유럽의 연구정보 교환과 기술을 위해 개발된 CERIF와 보존 메타데이터의 표준으로 활용되고 있는 PREMIS 보존 메타데이터를 비교·분석하였다.

또한 본 연구에서는 비교 대상 시스템의 메타데이터를 ‘연구정보’와 ‘서지정보’, ‘기록정보’로 카테고리화 구분하였다. ‘연구정보’는 관리적 메타데이터 요소로서, 연구성과물이 지니고 있는 연구과제 정보를 기술하는 요소이다. 즉 연구성과물의 배경정보를 식별할 수 있는 기본 정보로서, 과제코드, 과제수행기관명, 연구비 지원 기관명, 연구기관 등의 정보를 포함하고 있다. ‘서지정보’는 기술적 메타데이터 요소로서, 출판물 형태의 연구성과물이 지닌 서지학적 정보를 기술하는 요소이다. 즉 연구성과물의 검색과 활용을 위한 연구성과물의 서지사항(제목, 저자명, 발행연도 등) 등이 포함된다. ‘기록정보’는 보존적 메타데이터 요소로서, 연구성과물을 생산한 전자기록물에 대한 기록관리 및 보존 정보를 기술하는 요소이다.

이상 세 가지 카테고리 구분은 기록관리적 관점에서 전자기록물의 생산과 별첨형태의 디지털 연구성과물 자체에 대한 관리와 보존적 요소를 모두 포함하고 있다. 이는 해당 전자기록물의 기술적인 측면을 상세히 기술하여 해당 기록물의 관리·보존의 기능을 가능하게 한다.

19) Priscilla Caplan, "The Florida Digital Archive and DAITSS: a Model for Digital Preservation," *Library Hi Tech*, Vol.28, No.2(2010), pp.224-234.

20) Stanford Digital Repository, APSR 프로젝트, DAITSS시스템, MathArc 프로젝트, SHERPA DP프로젝트.

21) National Snow and Ice Data Center Repository, Statistics New Zealand Data Archive.

1. 국내

국내 디지털 연구성과물을 관리하는 대표적인 시스템은 한국과학기술정보연구원의 NTIS시스템, 행정안전부의 정책연구용역관리시스템(Policy Research Information Service & Management 이하 PRISM), 기록관리기관의 표준 기록관리시스템이다.

가. NTIS 시스템

NTIS 시스템은 '과제', '성과', '참여인력', '장비·기자재', '평가위원'에 대한 정보를 관리하는 국가 R&D분야 종합관리시스템이다. 이 중 '국가 R&D보고서 성과물 종합관리시스템'은 국가 R&D 사업을 통해 산출된 보고서 성과물을 체계적으로 등록하고 관리하는 시스템이다. NTIS의 『기관수 집용 NTIS 통합스키마 v.1.09』²²⁾은 연구성과물 관리를 위한 메타데이터 요소이다. 통합스키마에서 연구성과물에 해당하는 요소는 <표 3>과 같다.

<표 3> NTIS시스템 연구성과물 메타데이터

분야	NTIS 개체 및 속성	
	상위요소	하위요소
과제	12개 개체(ex. 사업정보)	93개 속성(ex. 사업명)
성과	1개 개체(ex. 연구보고서)	10개 속성(ex. 보고서제목)
참여인력	8개 개체(ex. 경력사항)	42개 속성(ex. 근무처명)
평가위원	3개 개체(ex. 지역서실적)	10개 속성(ex. 국문저서명)

<표 3>과 같이 NTIS 연구성과물 메타데이터는 총 24개 개체, 155개 속성이다.²³⁾ 성과 분야의 연구보고서 개체는 연구성과물의 '서지정보'를 포함하고 있으며, 과제, 참여인력, 평가위원 분야는 '연구정보'와 '기록정보'를 수록하고 있다.

NTIS시스템 연구성과물 메타데이터 요소의 특징은 연구과제에 대한 상세한 기본 정보를 제공하고 있다. 상세한 메타데이터 요소는 대상 객체를 표현하기에는 장점이 있지만, 장기보존을 위한 메타데이터 요소로서는 보존 정보의 중복과 시스템적 과부하의 원인으로 작용할 수 있다. 즉, 자원의 특성을 상세하고 구체적으로 표현하기 위해 요소 수를 늘리면 검색의 정확성을 향상시킬 수 있으나, 메타데이터를 생성하는데 비용과 시간이 많이 들고 메타데이터 비전문가가 입력하기 어려우며, 메타데이터를 생성하는데 있어 일관성이 떨어질 수 있다는 단점이 있다.

22) NTIS 자료실 홈페이지, <<http://www.ntis.go.kr/ThPds0701.do>> [인용 2011. 5. 30].

23) 성과정보 중 연구보고서 개체를 제외한 생물자원, 생명정보, 논문, 화합물 등 12개 개체의 정보는 본 연구의 연구 성과물 범주에 포함되지 않으며, 장비·기자재 정보 역시 해당되지 않으므로 제외하였음.

나. 정책연구용역관리시스템(PRISM)

행정안전부는 중앙부처의 연구용역 발주에서 완료까지 전 과정을 전산시스템으로 관리하기 위하여 2006년 1월 PRISM을 구축·운영하고 있다.²⁴⁾ PRISM의 메타데이터 요소는 <표 4>와 같다.²⁵⁾

<표 4> PRISM 연구성과물 관리를 위한 메타데이터

분야	PRIMS 개체 및 속성	
	상위요소	하위요소
사업 계획	1개 개체(ex. 사업계획)	20개 속성(ex. 기관명)
용역 계약	1개 개체(ex. 용역계약)	15개 속성(ex. 계약일자)
연구 진행	1개 개체(ex. 연구진행)	5개 속성(ex. 중간점검보고서)
연구 완료	2개 개체(ex. 연구완료)	13개 속성(ex. 제목)

<표 4>와 같이 PRISM의 '연구정보'는 사업계획, 용역계약, 연구진행(평가), 연구완료의 4개 분야이며, 53개의 하위요소로는 구성된다.²⁶⁾ '서지정보'는 연구보고서 등록의 10개 하위속성이 해당된다. PRISM 요소의 특징은 NTIS시스템의 '연구정보'보다는 단순하며 용역계약에 대한 특정 요소 값을 포함하고 있으나, 요소간의 중복이 있다. 또한 '서지정보'는 필수적인 요소만을 포함하고 있으며, 연구성과물의 발행기관명, 보고서 유형과 같은 사항이 누락되어 있다. 한편, PRISM 요소는 기록정보를 포함하고 있지 않고 있다.

다. 표준 기록관리시스템

표준 기록관리시스템은 기록물관리법에 근거하여 기록물의 생산, 보존, 활용에 이르는 기록관리 업무를 전자적으로 수행하는 시스템으로서, 국가기록원에서 개발하여 공공기관을 대상으로 도입·확산되고 있다.²⁷⁾

기록물 관리를 위한 메타데이터는 <표 5>와 같이 현용·준현용 단계 기록물 관리를 위한 '기록관리 메타데이터 : 현용·준현용 기록물용(NAK/S 8:2007(v1.1))', '전자기록물 장기보존포맷 기술규격(NAK/TS 3:2008(v1.0))', '영구기록물 기술규칙'이 있다.

24) 정책연구용역종합관리시스템(PRIMS) 홈페이지, <<https://www.prism.go.kr>> [인용 2011. 5. 1].

25) 중복 데이터 및 2차 결과물이기 때문에 PRISM의 활용상황관리 개체는 제외.

26) 활용상황관리 분야의 메타데이터 요소는 제외(2차적인 메타데이터이며, 기존 분야와 중복임).

27) 국가기록원, 기록관리 메타데이터 국가표준화 연구(대전 : 국가기록원, 2010), p.2.

〈표 5〉 기록물 관리를 위한 메타데이터

분야	개체 및 속성	
	개체(상위요소)	속성(하위요소)
기록관리 메타데이터	21개 개체(ex.행위자)	78개 속성(ex.기관명)
전자기록물 장기보존포맷 기술규격	16개 개체(ex.객체메타데이터)	120개 속성(ex.객체유형)
영구기록물기술규칙	7개 개체(ex.배경)	27개 속성(ex.생산자명)

〈표 5〉와 기록물 관리를 위한 메타데이터에서 국내 표준 기록관리시스템은 ‘기록관리 메타데이터 표준 : 현용·준현용 기록물용(공공표준)’보다 ‘전자기록물 장기보존포맷 기술규격’내의 5.3 장기보존포맷 메타데이터 요소를 적용하여 데이터를 설계하였다. 이에 따라 메타데이터가 축적되고 있어, 현행 기록관리 메타데이터는 문서상의 표준에서 벗어나지 못하고 있다.²⁸⁾ 장기보존포맷은 10년 이상의 기록만을 대상으로 기록과 메타데이터 요소를 함께 축적하고 있을 뿐, 활용적 측면보다 보존을 우선적으로 고려한 관리 방식이다. 따라서 장기보존포맷은 메타데이터 표준이라고 할 수 없으며 시스템 상에서 활용 가능하고 각 시스템 간의 상호운용성이 보장된 표준으로 고도화하고 이를 적용할 수 있도록 구조화되어야 한다. 또한 기록관리 메타데이터, 전자기록물 장기보존포맷 기술규격, 영구기록물 기술규칙은 전자기록물에 대한 철, 건 단위 보존과 관리적 측면을 중심으로 기술된 메타데이터이기 때문에, 연구성과물의 고유한 속성을 반영하여 기술하기에는 제한이 있다. 즉 전자기록물에 대한 관리와 보존의 요소이지, 그 첨부 파일인 연구성과물에 대한 상세한 서지 정보를 기술하기에는 제한이 있다.

2. 국외

국외 사례 비교·분석 대상은 미국 NTIS산하 NTRL과 유럽연합 기록관리시스템 메타데이터 모형, 유럽 CERIF, PREMIS 보존 메타데이터이다. 해당 메타데이터 요소를 분석하여 국내 디지털 연구성과물의 장기보존을 위한 메타데이터 요소로서 활용 가능성과 그 한계를 알아보았다.

가. 미국 NTRL

NTRL는 2010년 기준으로 약 600,000종의 디지털화된 원문 연구보고서의 서지 데이터베이스 메타데이터 연계 및 접근 서비스를 하고 있다.

NTRL Database의 레코드는 21가지의 데이터 요소들로 구성되어 있다.²⁹⁾ 데이터 요소들은 총

28) 국가기록원, 전계서, p.59.

29) NTIS Homepage, 〈<http://www.cas.org/ASSETS/106346A3BB7142FB2E164A178C2F67/ntis.pdf>〉 [cited 2011. 4. 25].

21개이며, 상세 구성요소는 등록번호(AN), 서명(TI), 저자(AU), 출처(CS), 기관코드(CC), 보고서번호(RN), 연구비 지원기관(SP), 계약번호(CN), 개요(NT), 생산년월일(PY), 페이지수(PG), 가격정보(PC), 표기언어(LA), 발행국(CP), 문서유형(DT), 공개학술지정보(JA), 초록(AB), 디스크립터(DE), 식별자(ID), CAS Registry 번호(RG), 주제코드(SH)이다. 이와 같이 NTRL는 자료에 대한 검색과 서비스를 목적으로 연구성과물을 기술하고 있다. 특히 출처(CS), 기관코드(CC), 보고서번호(RN), 연구비 지원기관(SP), 계약번호(CN), 공개학술지정보(JA) 등은 연구성과물을 식별하는 고유한 값으로 활용되고 있다.

나. 유럽연합(EU) 기록관리시스템 메타데이터 모형

유럽연합 기록관리시스템 메타데이터 모형은 2008년 발표된 기록관리시스템 요건 모형 2(MoReq2 : Modular Requirements for Records Systems 2)의 일부로서 모범적 기록 관리를 가능하게 하면서 MoReq2에 부합하는데 필요한 최소 메타데이터 모형이다. 각 기관 및 응용프로그램에 필수적으로 적용할 수 있는 메타데이터 요소로 개발되었다. ISO 15489, ISO 23081-1, 및 ISO 15836 등의 국제 표준의 개념을 기반으로 하고 있으며 기록관리시스템 환경으로 한정하여 적용한 메타데이터 요소이다.³⁰⁾

MoReq2는 9개 구분(ex. 분류 스킴)과 158개 메타데이터 요소로 구성되어 있다.³¹⁾ 이 모형은 필수요소와 선택요소의 2개 유형으로 요소의 의무 사용 여부를 제시하고 있다. MoReq2는 최종 사용자가 기록관리시스템에 대한 입력의 필요 메타데이터를 줄이고 분류 시스템을 단순화했다는 평가를 받고 있다. 또한 MoReq2의 메타데이터 스키마는 MoReq2 호환 시스템 사이의 모든 메타데이터 기능(예 : 액세스 권한, 보존)과 문서의 교환이 가능하다. 단 아직까지 전자기록관리시스템 관련 업체에서 MoReq2 기반의 광범위한 시스템 구현을 이루어지 않고 있다.³²⁾

다. CERIF

CERIF는 유럽전역의 R&D정보를 포괄적으로 표현하는 공통의 언어로서 메타데이터, 데이터베이스 스키마 및 시스템 개발에 필요한 각종 가이드라인을 담고 있다. CERIF 2008 1.2는 발간물 개체의 범위를 확장 및 개선하였고, 발간물 유형과 역할을 갖는 새로운 CERIF 의미 체계를 도입하였다. CERIF 2008 1-2는 데이터모델의 범용성과 데이터 요소의 명료함에 있어서 연구성과물을

30) Appendix 9, *To The MoReq2 Specification : Metadata Model*, 2008, http://www.dlmforum.eu/index.php?option=com_jotloader§ion=files&task=download&cid=374_693631f8cbc94e0cb4ff0597e6ab974c&Itemid=36&lang=en [cited 2011. 5. 20].

31) 개별 기록의 일부분이며, 단일한 물리적 객체를 구성하는 단계.

32) Philipp Wilhelm, "An Evaluation of MoReq2 in the Context of National EDRMS Standard Developments in the UK and Europe," *Records Management Journal*, Vol.19, No.2(2009), pp.117-134.

수집, 관리, 보존하는 기관에서 참조할 수 있는 대표적인 개념적 모형이다.³³⁾ CERIF 2008 1-2의 주요 개체 유형으로는 <표 6>과 같다.

<표 6> CERIF 2008 1-2의 주요 개체 유형

분야	하위 개체
기본 개체(Base Entities)	과제(cfProject), 인명(cfPerson), 기관(cfOrgUnit)
산출물 개체(Result Entities)	발간물(cfResultPublication), 특허(cfResultPatent), 시제품(cfResultProduct)
2차 개체(2ndLevel Entities)	인용(cfCitation), 국가(cfCountry), 통화(cfCurrency), 장비(cfEquipment), 이력서(cfCurriculumvitae), 이메일(cfElectronicAddress), 이벤트(cfEvent), 시설(cfFacility), 예산(cfFunding), 언어(cfLanguage), 구조(cfMetrics), 우편주소(cfPostalAddress), 수상(cfPrizeward), 자격(cfQualification), 서비스(cfServices)
링크 개체(Link Entities)	2개의 개체를 연결하여 양자 간의 관계를 설정

<표 6>과 같이 CERIF 2008 1-2는 기본 개체, 산출물 개체, 2차 개체, 링크 개체로 구성되어 있으며, 이들은 여러 개의 하위 개체로 세분된다. 총 195개의 하위 개체, 1,176개의 하위 세부속성을 지니고 있다. 기본 개체와 2차 개체 분야는 '연구정보'의 메타데이터 요소를 포함하고 있으며, 산출물의 발간물 개체는 '서지정보'에 포함되는 요소들이다. 하지만 CERIF의 메타데이터 요소를 국내 연구성과물의 장기보존을 위한 메타데이터로 몇 가지 한계를 지니고 있다.

첫째, CERIF는 필수적인 보존 메타데이터 요소가 존재하지 않는다. CERIF의 산출물 개체 중 발간물은 총 18개의 하위 개체를 지니고 있지만, 산출물에 대한 직접적인 기술을 위한 요소는 서명, 부서명, 키워드, 초록, 약어, 서지사항(Bibliographic Note) 6개뿐이다. 나머지 11개 하위개체는 발간물이라는 산출물과 2차 개체와의 관계를 설정하는 요소들이다.³⁴⁾ 또한, 연구성과물 하위 개체 내 13개 하위 세부 속성에는 식별자, 출판년일, 권, 호, 판사항, 시리즈, 호, 시작 쪽, 끝 쪽, 전체 페이지, ISBN, ISSN, URI 정보를 담고 있다.³⁵⁾

둘째, CERIF는 너무 엄격한 포맷과 반복되는 정보의 처리가 불가능하고 복수 언어와 문자를 지원하지 않으며, 연구성과물 보존을 위한 풍부한 표현에 한계를 지니고 있다. 다만, CERIF에서 제시한 관계 모형의 구성 요소는 기술적 측면에서 모든 개체들 간의 링크를 중심으로 관계를 설정하였다는 점에서는 의의가 있다.

33) euroCRIS, *CERIF 2008-1.2 Full Data Model(FDM)*, <http://www.eurocris.org/Uploads/Web%20pages/CERIF2008/Release_1.2/CERIF2008_1.2_FDM.pdf> [cited 2011. 2. 9].

34) ex. cfResult Publication_Citation, cfResult Publication_Funding 등.

35) CERIF 2008-1.2 Full Data Model(FDM), ResultPublication 개체.

라. PREMIS

PREMIS 데이터 사전의 데이터 모형은 지적 개체를 비롯하여 객체 개체, 이벤트 개체, 행위주체 개체, 권한 개체로 구성되며 각각의 개체들은 서로 관계를 갖는다. 이 관계를 이어주기 위해 각 개체들은 식별자(identifier)를 갖게 되며 상호연관적인 방향을 갖는다. PREMIS 데이터 사전 2.1의 의미 단위의 계층적 목록은 <표 7>과 같다.

<표 7> PREMIS Data Dictionary의 의미단위

개체	의미단위	개체	의미단위
1. Object	1.1 objectIdentifier	2. Event	2.1 eventIdentifier
	1.2 objectCategory		2.2 eventType
	1.3 preservationLevel		2.3 eventDateTime
	1.4 significantProperties		2.4 eventDetail
	1.5 objectCharacteristics		2.5 eventOutcomeInformation
	1.6 originalName		2.6 linkingAgentIdentifier
	1.7 storage		2.7 linkingObjectIdentifier
	1.8 environment	3. Agent	3.1 agentIdentifier
	1.9 signatureInformation		3.2 agentName
	1.10 relationship		3.3 agentType
	1.11 linkingEventIdentifier		3.4 agentNote*
	1.12 linkingIntellectualEntityIdentifier		3.5 agentExtension*
	1.13 linkingRightsStatementIdentifier		3.6 linkingEventIdentifier*
4. Rights	4.1 rightsStatement	3.7 linkingRightsStatementIdentifier*	
	4.2 rightsExtension	* ver. 2.1 신규 차고 요소	

<표 7>과 같이 PREMIS 데이터 사전과 이것을 지원하는 도큐멘테이션은 디지털 아카이빙 시스템에서 보존 메타데이터를 구현하기 위한 종합적이고 실용적인 리소스이다.³⁶⁾ 또한 디지털 기록의 장기보존을 담당하는 디지털 아카이브가 알아야 할 필요에 초점을 두어 개발되었다. 이러한 PREMIS 데이터 사전 역시 디지털 연구성과물을 장기보존하기 위한 메타데이터로서 몇 가지 한계점을 지니고 있다.

첫째, Paradigm Project(2005)에 따르면, PREMIS 데이터 모델을 연구성과물에 적용하기 위해서는 PREMIS 모델 내의 지적개체에 대한 구조 분석 및 기술이 선행되어야 한다. 이는 PREMIS가 ‘연구성과물’에 최적화된 데이터 모델이 아니라, 보존을 필요로 하는 모든 디지털 객체를 위한 데이터 모델이기 때문이다. 또한 PREMIS는 데이터 모델의 구현을 위해 준수해야 할 구

36) PREMIS Data Dictionary Version 2.1: Hierarchical Listing of Semantic Units, <<http://www.loc.gov/standards/premis/v2/premis-dd-Hierarchical-Listing-2-1.html>> [cited 2011. 6. 4].

축상의 지침이나 최소한의 규정이 마련되어 있지 않다.³⁷⁾

둘째, PREMIS의 5가지 개체 중에서 연구성과물과 직접적으로 관련 있는 개체는 지적 개체와 객체 개체이다. 이 중 객체 개체는 표현, 파일, 비트 스트림의 3가지 서브타임을 가지는 개체로 연구성과물에 직접적으로 적용가능하다. 그러나 PREMIS의 지적 개체를 별도로 생성해야하는 문제점을 지니고 있다.

따라서 연구성과물에서 무엇을 관리와 기술을 위한 하나의 지적 단위로 볼 것인가의 문제와 연구성과물을 생산하는 과정에서 산출된 또 다른 정보단위와의 집합적인 성격을 수용하는 방법을 모색해야 한다.

3. 소결

지금까지 사례 조사한 국내외 연구성과물 메타데이터 요소들은 ① 연구성과물의 연구과제에 대한 정보, ② 연구성과물의 서지사항에 대한 정보, ③ 연구성과물의 관리를 위한 정보로 그룹화할 수 있다. 디지털 연구성과물을 장기보존하기 위한 보존적인 메타데이터는 국내외 사례 시스템 및 메타데이터 모형에서 누락되어 있다. 국내 디지털 연구성과물 객체에 대한 메타데이터는 기술적인 메타데이터만 포함하고 있다. 이것은 별도의 보존 메타데이터를 포함하지 않고 있고, 보존을 위한 메타데이터로 사용하기에는 상대적으로 미흡하다. 특히 보존과 관련된 관리적이고 구조적인 메타데이터 요소는 포함이 되어 있지 않아 전반적인 메타데이터 스키마의 수정과 보완 작업이 요구된다.

한편, 국내 연구성과물에 대한 정보는 연구기관별, 관리기관별로 서로 상이한 체제로 관리하고 있어서 연구성과물의 보존을 위한 호환성 및 상호운영성에 문제가 있다. 연구성과물을 기술하기 위한 메타데이터로써 그 규모가 방대하며 실제 업무에서 수집되어야 할 정보를 망라하고 있기 때문에 이를 연구성과물의 기록관리적 특성을 고려하여 정비할 필요가 있다.

IV. 디지털 연구성과물 장기보존을 위한 메타데이터 요소 설계

이 장은 연구성과물의 장기보존을 위한 메타데이터 요소를 설계하고자 한다. 이를 위해 설계원칙 정립, 조사 대상기관 및 표준에 기반 한 공통 요소를 도출하고, 도출된 요소를 '연구정보', '서지정보', '기록정보' 카테고리에 따라 그룹화하였다. 이후 도출된 메타데이터 요소를 전문가 집단을

37) Paradigm Project, *Workbook on Digital Private Papers*,

〈<http://www.paradigm.ac.uk/workbook/metadata/preservation-metadata.html>〉 [cited 2011. 3. 2].

통한 델파이조사를 실시하여 검증받았다.

1. 설계 원칙

연구성과물을 장기보존하기 위한 메타데이터 요소들은 다양한 요소들이 사용된다. 이 중에는 연구성과물 기술에 사용되는 필수적인 요소들이 있으며, 특정 유형의 연구성과물에만 선택적으로 적용되는 요소들도 존재한다. 본 연구는 제3장의 국내외 비교대상 시스템 및 기관의 메타데이터 요소를 상호 매핑하여 공통적으로 사용되는 요소를 추출하였다. 도출된 공통 요소들은 연구성과물에 공통적으로 적용 가능하며, 세부적인 연구성과물 생성뿐만 아니라, 간략한 형태의 연구성과물 생성에도 기술된다.

따라서, 디지털 연구성과물 장기보존을 위한 메타데이터 구조 설계 절차는 메타데이터의 핵심 모델을 개발하고 기타 관련 데이터 요소들을 그룹화한 다음 전체적 메타데이터 구조 및 상세정도를 결정한 후 최종적으로 각 범주별 메타데이터 요소를 기술하였다.

2. 메타데이터 요소 설계

비교 대상인 국내외 연구성과물 관리 시스템과 표준 메타데이터 간의 공통되는 메타데이터 요소를 중심으로 총 42개 상위요소와 117개 하위요소를 도출하였다. 공통된 메타데이터 요소 중 일부는 연구성과물을 장기보존하기에 한계와 제한이 있다. 그 이유는 다음과 같다.

첫째, 국내 연구보고서의 표준 형식에 포함되어 있는 사항 중 누락된 요소 및 전자기록물을 장기보존하기 위한 필수 메타데이터 요소 추가가 필요하다. 앞서 언급한 것과 같이, 『국가연구개발사업의 관리 등에 관한 규정 시행규칙』에서 국내 R&D사업에서 국가 간의 공동연구 부분을 중요한 메타데이터 요소(국제공동연구, 상대국 명, 상대국 연구기관명)로 선정하고 있으며, 주관연구기관에서 연구과제의 일부를 위탁하여 수행한 위탁연구기관의 기관 및 인명정보(위탁연구, 위탁연구기관명, 연구책임자)와 참여연구원의 수에 대한 메타데이터 요소 추가가 필요하다.

둘째, 3개의 카테고리 간의 메타데이터 중복이 존재하며, 개념상 유사한 사항에 대해서 삭제 및 재그룹화가 필요하다. 왜냐하면, 전자기록물에 포함된 연구성과물을 장기보존하기 위해서는 기록관리시스템 내 메타데이터에서 연구성과물을 기술하고, 전자기록물과 연구성과물을 패키징하여 장기보존하는 전략이 필요하기 때문이다. 따라서 기준이 되는 기록정보와 중복이 되는 연구정보 및 서지정보의 일부 요소를 삭제하였다. <표 8>은 삭제 처리된 공통 메타데이터 요소이다.

〈표 8〉 삭제처리 된 메타데이터 요소

메타데이터 요소명	사유
사업부처명(기관명)	기록정보의 행위주체와 중복, 행위주체에서 관련정보를 확인할 수 있음
과제담당자 부서(명)	
과제담당자 이름	
과제담당자 전화번호	
과제담당자 이메일	
과제담당자 주소	
사업기간	총연구기간, 당해연구기간과 중복
계약금액	당해연도사업비와 중복
국가구분	기록정보에서 해당 내용을 확인가능
지역구분	
대과제명	실질적으로 사업명과 중복
책임연구원-주소, 전화번호, 이메일	기록정보와 일부 주소, 전화번호, 이메일 중복 * 연구자의 주소, 전화번호, 이메일은 수시로 변경됨. 보존행위의 식별자로서 사업자등록번호, 기관명, 부서명, 성명으로 식별 가능함.
연구원-주소, 전화번호, 이메일	
평가위원-주소, 전화번호, 이메일	
주관기관-주소, 전화번호, 이메일	
참여기관-주소, 전화번호, 이메일	
발주기관명	발주기관명은 연구정보의 부서명, 기록정보의 행위주체와 중복, 발행기관명은 주관연구기관명과 중복
발행기관명	
저자명	발행기관명 또는 책임연구원 성명과 중복 됨
(연구보고서)분류	연구과제 분류와 중복 ³⁸⁾

〈표 8〉과 같이 총 30개의 하위요소가 메타데이터 요소간의 중복 및 재그룹화로 삭제하였다. 과다한 요소 기술은 효용의 관점에서 장기보존 시 불필요한 정보를 관리해야 하는 비효율이 발생될 수 있으며, 또한 이메일, 주소, 전화번호는 검색의 접근성으로 활용되지 않을 뿐만 아니라, 기록을 생산 근접 시점에서만 의미 있는 정보로 활용된다. 또한 실제 전자문서에서 확인하는 방법으로도 대체가 가능하다.

따라서 공통된 메타데이터 요소 도출과 추가 및 삭제 등을 통한 1차 디지털 연구성과물 장기보존을 위한 메타데이터 요소는 〈표 9〉와 같이 40개 상위요소와 94개 하위요소를 도출하였다. 이렇게 도출된 메타데이터 1차안을 디지털 아카이빙 전문가들에게 델파이조사를 실시하여, 그 타당성 및 추가 요소를 도출하였다.

38) 연구성과물 자체에 대한 분류는 앞서 과제에 대한 분류로 처리할 수 있기 때문에 삭제 또한 기관별로 상이한 분류체계를 적용하기 때문에 분류에 대한 하나의 고정된 값을 표현하기에는 한계가 있음.

〈표 9〉 디지털형태 연구성과물 장기보존을 위한 메타데이터 요소(안)

구분	상위요소	하위요소	간략설명	
연구 정보	사업정보	사업명(프로그램명)	연구과제에 대한 사업명	
		프로그램코드	부처별 프로그램코드	
		단위사업코드	부처별 단위사업코드	
		세부사업코드	부처별 세부사업코드	
		계약일자	사업 계약 체결 일자	
		계약방식	사업 계약 방식	
	기본정보	과제번호	과제별로 부여하는 고유번호	
		(세부) 과제명	협약을 맺은 과제의 명칭	
		총연구기간	총 사업의 연구 수행기간	
		당해연도 연구기간	연차사업의 당해연도 연구 수행기간	
		연구과제 분류	연구과제의 기준이 되는 분류코드	
		과제개요	연구과제에 대한 개요	
	연구비 정보	키워드	전체 연구과제에 대한 핵심 키워드	
		총 사업예산	연구과제의 총 사업예산(금액)	
	평가정보	당해연도사업비	연차사업의 당해연도 사업비	
		종합의견	연구과제의 평가에 대한 종합의견	
		평가구분	연구과제의 평가구분 코드 값	
	인명	참여인력	평가일자	연구과제의 평가시행 일자
			책임연구원-성명	연구과제에 참여하는 책임인력의 이름
			책임연구원-소속부서명	연구과제에 참여하는 책임인력의 소속 부서명
			연구원-성명	연구과제에 참여하는 연구인력의 성명
		평가위원	연구원-소속	연구과제에 참여하는 연구인력의 소속명
			참여연구원 수	연구과제에 참여하는 연구인력의 수
			평가위원-성명	연구과제 결과평가를 위한 평가위원의이름
			평가위원-소속	연구과제 결과평가를 위한 평가위원의소속
	기관	주관기관	주관연구기관명	연구과제를 수행하는 기관의 명칭
			주관연구기관사업자등록번호	연구과제를 수행하는 기관의 사업자등록번호
		참여기관	참여연구기관명	연구과제에 참여하는 기관의 명칭
			참여연구기관사업자등록번호	연구과제에 참여하는 기관의 사업자등록번호
		국가공동연구	상대국 국가명	연구과제를 공동수행한 국가명(추가)
상대국 연구기관명			연구과제를 공동수행한 국외기관명(추가)	
위탁기관	위탁연구기관명	연구과제를 위탁하여 수행한 기관의 명칭		
	위탁연구기관사업자등록번호	연구과제를 위탁하여 수행한 과제 책임자 명		
서지 정보	연구 보고서	보고서제목	연구보고서의 제목	
		영문제목	연구보고서의 영문제목	
		부제목	연구보고서의 부제목	
		권차	연구보고서의 권차정보	
		권차서명	연구보고서의 권차서명 정보	
		발행년월일	연구보고서의 발행년월일	
		언어	연구보고서에 기술된 언어	
		목차	연구보고서의 목차 정보	

구분	상위요소	하위요소	간략설명	
서지 정보	연구 보고서	초록	하위속성 없음 연구보고서의 초록 정보	
		키워드	하위속성 없음 연구보고서의 핵심 키워드	
		고유식별자	하위속성 없음 연구보고서를 식별하기위한 식별자	
		원문주소	하위속성 없음 연구보고서의 원문자료에 대한 연결 URL	
		페이지	하위속성 없음 연구보고서의 페이지 정보	
기록 정보	기록물 맥락정보	행위주체	개인명	전자기록물의 행위주체 유형의 코드값
			개인ID	생산시스템으로부터 입수된 개인ID
			기관명	전자기록물의 행위자의 기관명
			기관코드	행정표준코드의 기관코드
			부서명	전자기록물의 행위자의 소속 부서명
			부서코드	행정표준코드의 처리과코드
			직위명	전자기록물의 행위자의 직위명
			직위코드	행위자의 직급코드
			이메일	행위자의 이메일 주소
			주소	행위자의 주소
	생산이력	생산자정보	연구성과물 전자기록물 생산자 유형	
		생산시스템	기록을 생산한 시스템	
		부가번호유형	처리과에서 기록에 부여된 부가번호 유형	
		부가번호	부여된 부가번호	
		인수인계일시	인수인계가 발생한 일시	
		인계자	기록을 인계하는 기관명(처리과명.개인명)	
		인수자	기록을 인수하는 기관명(처리과명.개인명)	
	인계인수관련법규	인계인수관련 법규사항		
	관계	관계유형	기록과 타 정보와의 관계유형	
		관계대상식별자	해당 관계 대상 식별자	
		관계설명	관계에 대한 추가설명	
	분류	분류체계구분	기록물의 분류체계	
		분류식별자(ID)	해당 분류체계의 고유식별자	
		분류체계명	해당 분류체계의 이름	
		분류계층	분류체계 레벨	
	기록물 내용정보	객체식별자	기본식별자(값)	기록의 고유한 식별자
			참조식별자(값)	추가적으로 생성되는 식별자
		표제	공식표제	기록물철, 건.문서에 주어진 공식 표제
			기타표제유형	공식표제 이외 부가 제목이 유형 구분
			기타표제명	공식표제 이외 부가 제목
기술		기술유형	기록의 내용, 목적에 대한 서술형식 유형	
		기술내용	기록 내용에 대한 요약 설명	
권한		보호기간	보안업무규정에 따라 비밀로 보호되는 기간	
		보안관련법규	보안과 관련되는 법규	
		열람범위	기록물을 기관내 직원이 접근하는 권한	
	공개여부	기록물건을 대국민에게 제공하는 유형		
	비공개사유	기록물을 비공개하는 사유		
	공개일시	비공개기록물을 공개로 전환된(될) 시점		
공개관련법규	공개와 관련되는 법규			

구분	상위요소	하위요소	간략설명	
기록 정보	기록물 내용정보	일시	(문서) 생산일시	기록이 생산시스템에 생산등록된 일시
			(문서) 접수일시	기록이 생산시스템으로 접수등록된 일시
			(문서) 시작일시	기록물철이 시작된 일시
			(문서) 종료일시	기록물철이 종료되거나 분철이 종결된 일시
	기록물 구조정보	포맷	저장매체포맷	기록의 매체나 물리적인 저장 매체의 종류
			데이터 포맷	첨부파일의 파일 포맷 및 보존 포맷
			저장 매체	기록이 저장되는 매체(공간)
		유형	기록 유형(객체 범주)	기록의 유형
		크기	용량 등록건수	기록의 용량, 문서의 쪽수 기록물철에 포함되는 첨부 건수
	기록계층	하위요소 없음	기록의 계층구조를 설명하는 단위	
	기록관리 과정	저장	내용(소장) 위치	기록이 저장되거나 상태에 대한 정보
			보존	보존처리유형
		보존처리일시		기록물 보존 처리 행위가 취해진 일시
		보존행위주체		기록물에 보존행위를 행한 사람
		보존처리설명		보존 처리 행위에 대한 설명
		보존기간	보존기간	기관에서 책정, 승인 또는 외부에서 재조정된 보존기간
			보존기간책정사유	보존기간을 책정한 사유
		관리이력	관리유형	기록이 상태에 영향을 미치는 모든 관리 행위의 유형
			관리내용	관리행위의 내용
			관리발생일시	해당 관리행위를 수행한 일시
관리행위주체			해당 관리행위를 수행한 사람	
관련법규	해당 관리행위에 대한 근거 법규			

3. 델파이조사 및 검증

델파이기법은 전문가들이 합의점을 찾을 때까지 반복되는 과정으로 진행되며, 단계에 대한 원칙을 고수하지 않지만 전통적으로 3라운드 과정을 거친다.³⁹⁾ 본 연구에서는 1라운드부터 구조화된 설문지가 구성되었으며, 참여자들의 추가의견을 개방형으로 응답하게 함으로써 메타데이터 요소를 수집하는 단계로 진행하였다. 만약, 추가 요인이 제기되지 않았다면 2라운드만으로 델파이 조사를 멈출 수 있었지만, 새로운 요인에 대한 응답의 안정성⁴⁰⁾을 보기 위해 3라운드 단계로 진행하였다.

델파이조사 참여자는 기록관 및 도서관 등에서 디지털 아카이빙, 장기보존, 디지털 콘텐츠 관리 등에 경험 및 전문적 지식을 제공할 수 있는 전문가로 하였으며, 각 단계별 15명을 선정하였다. 델파이조사 1라운드 설문지는 <표 9>의 도출한 메타데이터 요소로, 2라운드는 1라운드 결과에 의

39) Bernice B. Brown, *Delphi Proces: A Methodology Used for the Elicitation of Opinions of Experts*, (California : the RAND Corporation, 1968), pp.7-9, <<http://www.rand.org/pubs/papers/2006/P3925.pdf>> [cited 2011. 4. 30].

40) 응답의 안정성이란 이전 라운드에 비해 의견이 모아지는 정도, 즉 의견수렴을 의미함.

해 추가 및 삭제된 메타데이터 요소로 구성되었다. 3라운드는 응답의 안정성 및 일관성을 위해서 2라운드 설문과 같은 내용으로 구성하고 2라운드의 응답결과를 제시하였으며, 각 라운드별 리커트 척도로 질문하였다.

델파이조사는 2011년 5월 16일부터 2011년 6월 15일까지 총 3라운드가 진행되었으며, 모든 조사과정은 이메일로 이루어졌다. 델파이조사를 실시하기에 앞서 예비조사로 요소 설명 및 응답요령에 대한 이해, 응답 소요시간 등을 확인한 후, 이를 본 조사에 적용하였다.

가. 델파이조사 분석 결과

델파이조사의 자료는 SPSS 12.0 프로그램으로 분석하였다. 조사 참여자에 대한 일반적 사항 및 메타데이터 각각에 대한 요소의 적합성을 평가하기 위해 빈도분석, 평균 및 표준편차 등의 기술통계를 실시하였다. 델파이조사에 대한 안정성을 확인하기 위해 라운드 간 대응표준 t-검증을 실시하였고, 각 라운드별 요인에 대한 표준편차와 사분편차를 산출하였다.

(1) 델파이조사 라운드별 결과

(가) 1라운드 결과

1라운드는 도출된 메타데이터 요소들에 대한 적합정도를 파악하고, 추가요소를 도출하기 위해 실시하였다. 1라운드는 총 40개 상위요소와 94개 하위요소에 대하여 5점 리커트 방식으로 측정하였다.

총 40개 상위요소와 94개 하위요소 중에 평점 3점미만의 평가를 받은 요소는 연구정보에서 '평가위원', '국가공동연구'이다. 이 두 요소 모두 평균 2.5의 평가를 받았다. '서지정보'에서는 목차를 제외한 모든 요소들이 평균 4점 이상의 평가를 받았다. '기록정보'에서는 12개의 하위요소⁴¹⁾에 대해서 평균 3점미만의 평가를 받았다.

한편 제시된 요소 이외에 연구성과물의 장기보존하기 위한 메타데이터 요소를 개방형으로 직접 기재하도록 하였다. 이 중 최소 7명 이상이 언급한 요인 중에 제시된 요인과 중복되는 개념을 제외한 요소를 선정하였다. 선정된 요소는 '기록정보'의 포맷과 관련된 사항으로 델파이조사 1라운드의 '포맷' 상위요소 내 매체포맷, 데이터포맷, 저장매체의 하위요소 명을 포맷 명, 포맷 버전, 생성어플리케이션, 생성 어플리케이션 버전으로 변경하였다.⁴²⁾ 현재 기록관리 메타데이터의 '포맷' 요소는 기록물의 포맷이나 매체를 직관적으로 파악하는데 어려움이 있다는 의견이 있었다. 예를 들어, CD에 담

41) 이메일, 주소, 생산자정보, 인수인계일시, 인계자, 인수자, 인계인수관련법규, 기술유형, 보호기간, 보안관련법규, 열람범위, 접수일시.

42) 관리항목이 많지만, 업무담당자의 포맷 요소에 대한 이해도가 낮으며, 포맷 관련사항은 생산단계에서 이관되는 값이 없음. 생산 소프트웨어에 대한 정보를 관리하지 않으므로 향후 장기보존 측면, 접근가능성 측면에서 문제점을 제기함.

긴 연구성과물의 경우 매체포맷은 간행물(연구보고서), 데이터 포맷은 hwp, 저장매체는 CD로 등록되어야 한다. 그러나 현재의 요소 명칭을 통해 이를 쉽게 이해하고 적용하기에는 어려움이 있다는 전문가의 의견이 있었다. 추가된 생성어플리케이션명과 생성어플리케이션 정보 요소는 PREMIS에서 규정하고 있는 요소로써 기록을 생산한 어플리케이션 명의 정보와 버전의 정보를 의미한다. 다음으로 연구성과물의 장기보존은 전자기록물의 장기보존을 의미하는 것으로, 전자기록물 장기보존 포맷의 핵심요소 추가를 제안하였다. 추가된 요소는 '인증정보'⁴³⁾, '잠김인증정보'⁴⁴⁾의 2개 상위요소, 12개 하위요소가 추가되었다.

이와 같이 텔파이 1라운드를 통해서 적합정도가 평균 3점미만으로 평가된 상위요소 2개, 병합된 1개 상위요소 및 하위요소 16개의 요소를 제외하고, 추가된 2개 상위요소, 13개 하위요소를 포함하여 텔파이 2라운드 평가를 실시하였다.

(나) 2라운드 결과

2라운드는 1라운드에서 삭제 및 추가된 상위요소와 하위요소를 포함한 총 39개 상위요소와 91개 하위요소에 대하여 5점 리커트 방식으로 측정하였다. 가장 높은 적합성을 평가받은 요소는 1라운드와 같이 요소 간의 고유한 식별 정보를 제공하는 요소들이었다. 1라운드에서 추가된 상위요소와 하위요소를 포함한 모든 요인에서 3점 이상으로 나타남에 따라 3라운드에 원안과 같이 적용하였다.

(다) 텔파이 3라운드 결과

3라운드는 2라운드와 같은 요인에 대하여 실시되었다. 모든 요인들이 평균 3점 이상으로 평가되어 연구성과물을 장기보존하기 위한 메타데이터 요소로서 적합한 것으로 나타났다. 가장 높은 요소는 평균 5.0을 기록한 '과제명', '보고서 제목', '책임연구원 성명' 등 이다.

(2) 응답의 안정성과 의견수렴 및 합의도

의견합의는 이전 라운드와의 비교가 아닌 해당 라운드 내에서 발생하는 것으로 응답의 안정성이 확보되어야 한다.⁴⁵⁾ 따라서 본 텔파이조사에서도 각 라운드간의 응답의 안정성을 검증하고, 의견수렴 및 합의도를 분석하였다.

43) 기록물철, 건의 보존포맷 인증을 위한 서명정보, 컨테이너.

44) 잠김전자서명(lock signature)과 이것을 유효화하는 정보, 컨테이너.

45) 최원희, 사회복지 슈퍼비전의 중요성공요인(CSFs)에 대한 텔파이 및 AHP 분석(박사학위논문, 서울여자대학교 대학원 사회복지학과, 2009), p.80.

(가) 응답의 안정성

본 연구에서는 이전 라운드와의 변화 차이를 t-검증을 활용하여 안정성⁴⁶⁾을 확인하였다. 즉 1라운드와 2라운드 사이에는 '계약방식', '종합의견'에서 유의미한 차이를 보인 반면, 2라운드와 3라운드에서는 조사대상 항목 모두 차이를 보이는 항목이 나타나지 않았다. 따라서 조사 대상 메타데이터 요소 모두 안정성을 확보하였다.

(나) 의견 수렴 및 합의도

델파이조사 3라운드의 결과에 대하여 의견 수렴 및 합의도를 측정하기 위해서 본 조사에서는 사분편차와 표준 편차를 이용하였다. 사분편차는 하한 값과 상한 값에 위치하는 극단 값을 제거하는 특성을 갖는 지표를 말하는 것으로서 이종성(2006)⁴⁷⁾은 델파이조사 결과에서의 수렴도(degree of convergence)를 측정하고자 다음과 같이 사분편차를 이용한 공식을 제시하였다.

$$D_i = \frac{Q_3 - Q_1}{2}$$

델파이조사에서는 이 값이 작을 경우, 각 항목에 대해 집단 내 델파이 참여자의 의견차이가 적다는 것을 의미한다. 본 조사에서 이 방법을 사용한 결과, 델파이조사 1라운드의 상·하위요소 총 107개 요소⁴⁸⁾에서는 0.00인 경우는 47.7%(51개)였으나, 2라운드에서는 1라운드에서 삭제 및 추가된 요소를 포함 총 105개 요소 문항에서 0.00인 경우는 56.2%(59개)로 조금 줄어들었으나, 3라운드에서는 80.0%(84개)가 0.00인 것으로 나타났다. 3라운드의 의견차이 현상은 2라운드의 결과 값을 제시한 후에 실시한 것이기에 그 의견차이의 줄어드는 현상이 크게 발생하였다. 이와 같이 라운드가 지속되면서 참여자들의 의견차이가 줄어드는 현상은 의견이 수렴되었음을 의미한다. 의견수렴 및 합의도를 측정하기 위한 또 다른 방법으로 표준 편차가 있다. 표준편차가 작다는 것은 의견의 차이가 작다는 것을 의미하는 것이므로 만약 표준편차가 일정 수준 이하로 떨어지면, 의견이 합의 또는 일치된 것으로 간주한다. 그러나 기준 값에 대한 논리적 기준이 존재하지 않기 때문에 본 조사에서는 일반적으로 적용하는 응답범주의 20%⁴⁹⁾를 기준으로 하였다. 따라서 표준 편차가 1.0 미만이면 의견 일치도가 높다고 볼 수 있다. 본 연구에서는 해당 메타데이터 요소 모두가 1.0 미만을 충족하고 있으며, 의견의 일치가 이루어졌다고 판단할 수 있다.

46) 안정성(Stability)은 델파이 라운드 간 응답이 일관성을 파악하여, 델파이 라운드 수를 결정하는 정지수준(Stopping criterion)의 근거임.

47) 이종성, 델파이방법(서울 : 교육사, 2006).

48) 1라운드 결과로 제거된 3개 문항은 실제로 설문에 응답된 문항이지만 각 라운드 결과를 비교하기 위해 수렴 및 합의도를 보는 조사에서는 포함시키지 않음.

49) 이성웅(1987)은 최종라운드의 표준편차가 응답범주의 20%미만인 경우 의견일치도가 높다는 견해를 제시하였고, 이러한 기준은 델파이 연구에서 일반적으로 적용되고 있다(최원희, 전개논문, p.82).

4. 최종 메타데이터 요소 선정

텔파이조사 등을 통해 최종적으로 디지털 연구성과물을 기술하고 연구과제 정보를 식별할 수 있는 디지털 연구성과물 장기보존을 위한 메타데이터 요소를 선정하였다. 이 연구에서 제안하는 디지털 연구성과물 장기보존을 위한 메타데이터 요소는 기관의 정책과 실정에 따라 적합한 구조로 설계되어야 한다. 이를 위해, 제안한 디지털 연구성과물 장기보존을 위한 메타데이터 요소를 특정 정부부처에서 생산되고 있는 연구성과물에 적용해 보았다. 사례 적용 대상은 정책연구용역 50건과 R&D사업의 연구성과물 50건을 대상으로 하였다. 전자기록물은 부처 내의 연구성과물 결과를 보고한 내부 문서를 참고하였으며, 추가적으로 PRISM과 NTIS시스템에 등록되어 있는 원문 파일을 참고하였다. 메타데이터 요소의 사례 적용 과정은 입수 가능한 정보 범위 내에서 기술하였기 때문에, 모두 해당되는 요소와 필요에 의해 적용되는 선택적 요소로 구분할 수 있다. 또한, 기록정보는 대부분 기록관리시스템에서 자동으로 요소 값이 부여된다. 즉 해당 요소가 시스템에서 자동으로 부여되기 때문에 ‘연구정보’, ‘서지정보’처럼 어떠한 요소 값에 대한 빈도 분석이 필요가 없다. 최종 도출된 메타데이터 요소는 <표 10>과 같다.

<표 10> 디지털형태 연구성과물 장기보존을 위한 메타데이터(최종)

구분	상위요소	하위요소	적용사례	빈도	
연구 정보	사업정보	사업명(프로그램명)	예술의 진흥 및 생활화, 산업화	100	
		프로그램코드	문화적 생활공간 조성	100	
		단위사업코드	공공 디자인 개발 및 보급	100	
		세부사업코드	문화중심형 공공디자인 모델 개발	100	
		계약일자	2010.10.22	50	
		계약방식	수의계약	50	
	과제	기본정보	과제번호		50
			(세부) 과제명	2010년 공간문화 정책사업 성과평가 연구	100
			총연구기간		12
			당해연도 연구기간	2010.10.22 ~ 2011.01.21	100
			연구과제 분류	예술정책(정부기능분류)	100
			과제개요	생략	87
			(연구과제에 대한)키워드		62
	연구비정보	총 사업예산		12	
		당해연도사업비	36백만원	100	
	평가정보	종합의견	생략	87	
		평가구분		50	
		평가일자	2011.06.10	87	
	인명	참여인력	책임연구원-성명	김○○	100
			책임연구원-소속 부서명	○○대학교	100
			연구원-성명	김○○	100

26 한국도서관·정보학회지(제42권 제4호)

구분	상위요소	하위요소	적용사례	빈도	
연구 정보	인명	참여인력	연구원-소속	○○대학교	100
		참여연구원 수		(사)○○○○연구회	100
	기관	주관기관	주관연구기관명	111-11-11111	100
			주관연구기관사업자등록번호		34
		참여기관	참여연구기관명		34
			참여연구기관사업자등록번호		17
위탁기관	위탁연구기관명		17		
	위탁연구기관사업자등록번호				
서지 정보	연구 보고서	보고서제목	하위속성 없음	2010 공간문화 정책사업 성과평가 연구	100
		영문제목	하위속성 없음		43
		부제목	하위속성 없음		12
		권차	하위속성 없음	1, 2	26
		권차서명	하위속성 없음		26
		발행년월일	하위속성 없음	2011.01.21	100
		언어	하위속성 없음	한국어	100
		목차	하위속성 없음	생략	100
		초록	하위속성 없음	생략	100
		키워드	하위속성 없음	공간문화, 성과평가	100
		고유식별자	하위속성 없음	11-1371028-000024-01(발간등록번호)	53
원문주소	하위속성 없음	생략	100		
페이지	하위속성 없음	145쪽	100		
기록 정보	기록물 맥락 정보	행위주체	개인명	김○○	필수
			개인ID	시스템관리ID	선택
			기관명	문화체육관광부	필수
			기관코드	123456(행정표준코드관리시스템)	해당시필수
			부서명	○○○과	필수
			부서코드	123456(행정표준코드관리시스템)	선택
			직위명	주무관	선택
			직위코드	25862	선택
	생산이력	생산자시스템	통합온나라시스템	필수	
		부가번호유형	생산등록번호	선택	
		부가번호	○○○과-2186	선택	
	관계	관계유형	첨부	필수	
		관계대상식별자	자동 부여	필수	
	분류	관계설명	기록물건이 해당 철에 귀속되어 있음	선택	
		분류체계구분	기능분류시스템(BRM)	필수	
		분류식별자(ID)	A0000000	해당시필수	
		분류체계명	업무계획 및 평가	해당시필수	
	분류계층	소기능	해당시필수		
	기록물 내용 정보	객체식별자	기본식별자(값)	기록관리시스템에서 자동으로 생성된 ID	필수
참조식별자(값)			시스템에서 자동 부여	해당시필수	
표제		공식표제	2010 공간문화 정책사업 평가연구 완료 보고	필수	
		기타표제유형		해당시필수	
기타표제명		해당시필수			

구분	상위요소	하위요소	적용사례	빈도		
기록 정보	기록물 내용 정보	기술	기술내용	연구용역사업 완료를 보고함	선택	
		권한	공개여부	비공개	비공개	필수
			비공개사유	공공기관의 정보공개에 관한 법률 제9조 2항 6호	필수	
			공개일시	2020.01.01	선택	
			공개관련법규	관련법규로 연결	선택	
		일시	생산일시	2011.01.21	해당시필수	
	시작일시		2011.12.31	필수		
	종료일시		2011.12.31	필수		
	기록물 구조 정보	포맷	포맷명	PDF	해당시필수	
			포맷버전	1.2	해당시필수	
			생성어플리케이션명	Adobe Acrobat	해당시필수	
			생성어플리케이션 버전	9.3	해당시필수	
		유형	기록 유형	일반문서(시스템에서 선택 부여)	필수	
		크기	용량	200MB, 250쪽	필수	
	등록건수		3건	선택		
	기록계층	하위요소 없음	기록물철 / 첨부	필수		
	기록 관리 과정	저장	내용(소장) 위치	하드디스크(ex. server3\c:\admin)	선택	
			보존	보존처리유형	(ex)저장위치 이동	필수
보존처리일시				(ex)2011.12.31	필수	
보존행위주체				(ex)김○○	필수	
보존처리설명				(ex)원본파일을 아카이빙 스토리지로 이동	필수	
보존기간				(ex)영구	필수	
보존기간 책정사유		(ex)중요한 정책 결정에 관한 사항이 포함된 단위과제	선택			
관리이력		관리유형	(ex)폐기심사	필수		
		관리내용	(ex)소관부서 및 기록관 담당자의 심사	필수		
		관리발생일시	(ex)2011.12.31	필수		
		관리행위주체	(ex)폐기심사위원회	필수		
		관련법규	(ex)관련법규로 연결	필수		
인증정보		서명포맷설명	(ex)서명포맷에 대한 설명	필수		
		서명일시	(ex)2012.01.01	필수		
		서명자	(ex)김○○	필수		
		서명	(ex)Kim○○	필수		
		인증서블럭	(ex)GPKI	해당시필수		
		서명알고리즘	(ex)전자서명 S/W	해당시필수		
참검인증 정보		서명포맷설명	(ex)서명포맷에 대한 설명	필수		
		서명일시	(ex)2020.01.01	필수		
		서명자	(ex)김○○	필수		
		서명	(ex)Kim○○	필수		
		인증서블럭	(ex)GPKI	해당시필수		
		서명알고리즘	(ex)전자서명 S/W	해당시필수		

V. 결론 및 제언

본 연구는 전자기록물의 첨부물 형태로 수집·관리되는 디지털 연구성과물을 장기보존하기 위한 메타데이터 요소를 제안하였다. 이를 위해 국내외 연구성과물을 관리하는 시스템의 메타데이터 요소와 해외 연구성과물 관리·보존 메타데이터 표준을 비교·분석하였다. 그 결과를 요약하면 다음과 같다.

첫째, 연구성과물의 전자적 교환 표준인 CERIF의 경우, 개체 요소들은 연구정보의 교환 및 기술이 주된 목적이므로 필수적인 보존 메타데이터 요소는 별도로 마련되지 않았음을 확인하였다. 또한 PREMIS도 연구성과물을 직접적으로 기술하고 표현하기 위한 구체적 요소를 마련하지 못하고 있음도 확인하였다. 특히, PREMIS는 지적 개체에 관한 메타데이터를 제외하고 있으며, 기존에 개발되어 사용되고 있는 기술 메타데이터를 이용할 것을 권고하여 우리나라 디지털 연구성과물을 장기보존하기 위한 메타데이터 요소로써 한계를 확인하였다.

둘째, 우리나라 디지털형태 연구성과물의 장기보존을 위한 메타데이터 요소를 도출하기 위해 이론적 연구와 국내외 메타데이터 및 표준분석을 통해 메타데이터 구조를 ‘연구정보’, ‘서지정보’, ‘기록정보’로 구분하였다. ‘연구정보’는 연구성과물의 특징과 연구과제를 식별할 수 있는 고유정보이며, ‘서지정보’는 연구성과물의 검색과 활용을 위한 서지기술사항이다. ‘기록정보’는 기록관리관점에서 전자기록물을 보존하기 위한 정보를 의미하였다. 이와 같이 비교대상 시스템 및 표준에서 공통된 메타데이터 요소를 도출한 1차안을 대상으로 메타데이터 초안을 작성하였다. 도출된 요소를 전문가를 중심으로 델파이조사를 실시하여, 수정·보완·확정된 최종 메타데이터 요소를 도출하였다.

셋째, 최종 메타데이터 요소로서 ‘연구정보’ 카테고리에는 연구성과물이 담고 있는 연구과제, 인명, 기관으로 구분하였으며 8개 상위요소, 29개 하위요소로 구성하였다. ‘서지정보’는 연구성과물인 연구보고서의 서지사항에 대한 요소로써 하위요소 없이 13개의 상위요소로 구성하였다. ‘기록정보’는 연구성과물의 생산을 보고한 전자기록물의 맥락, 구조, 내용, 관리과정에 대한 요소로써, 18개 상위요소와 62개 하위요소로 제안되었다.

이상의 디지털 연구성과물의 장기보존을 위한 메타데이터 요소 설계는 영구기록물관리기관에서 연구성과물에 대한 장기보존의 문제점을 제기함과 동시에 장기보존을 실행하기 위한 토대를 제안하였다는 점이 의의가 있다.

한편 이 연구에서 제안한 메타데이터 요소가 실제 시스템이 구현되어 국가기록원을 비롯하여 연구성과물을 관리하는 기관 간 포맷교환과 데이터 공유가 이루어졌을 경우에 발생할 수 있는 연구는 수행하지 못하였다. 이러한 제한점에도 불구하고 기록관리학 분야의 전문가와 연구관리 기관

전문가의 의견을 델파이조사방법을 통해 반영하였기 때문에 향후 이 연구에서 제안한 결과에 대한 검증은 통해 그 합리성을 담보할 수 있을 것이다.

〈참고문헌은 각주로 대신함〉