

[별지 제15호 서식]

용역연구사업 연구결과보고서			
관리 번호	89017500		
사 업 명	2008년 기록관리분야 연구개발(R&D) 사업		
과 제 명	국문	전자기록물 장기보존 전략 연구 및 테스트베드 구축	
	영문	Study on Digital Records Long-term Preservation Strategy & Testbed Running	
주관연구기관	기 관 명	소재지	대 표
	성균관대학교 산학협력단	경기도 수원시 장안구 창천동 성균관대학교 내	이영관
주관연구 책임자	성 명	소속 및 부서	전 공
	오삼균	문과대	문헌정보학과
총연구기간	2008 년 5 월 23 일 - 2008 년 11 월 30 일(년 6 개월)		
총 연구비	49,000 천원		
연구년차	연구기간	연구비	
1차년도	200 . . - 200 . .	천원	
2차년도	200 . . - 200 . .	천원	
총참여연구원	9명 (책임연구원: 1명, 연구원: 2명, 연구보조원: 4명 보조원: 2명)		
년도 용역연구개발사업에 의하여 수행중인 연구과제의 연구결과보고서를 붙임과 같이 제출합니다. 붙임 : 1. 연구결과보고서 35부. 2008 년 월 일 주관연구책임자 오삼균 (인 또는 서명) 주관연구기관장 이영관 (직인) 국가기록원장 귀하			

주 의

1. 이 보고서는 국가기록원에서 시행한 용역연구개발사업의 연구결과 보고서입니다.
2. 이 보고서 내용을 발표할 때에는 반드시 국가기록원에서 시행한 용역연구개발사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개 하여서는 아니 됩니다.

제 출 문

국가기록원장 귀하

이 보고서를 “전자기록물 장기보존 전략 연구 및 테스트 베드 구축(성균관대학교 산학협력단/오삼균)”과제의 연구결과보고서로 제출합니다.

2008. 11. 25

주관연구기관명 : 성균관대학교 산학협력단

주관연구책임자 : 오삼균

제 1세부과제명 :

(제1세부연구기관/세부과제책임자):

제 2세부과제명 :

(제2세부연구기관/세부과제책임자):

제 3세부과제명 :

(제3세부연구기관/세부과제책임자):

위탁과제명 :

(위탁연구기관/위탁연구책임자):

목 차

I. 연구개발결과 요약문

(한글)

(영문)

II. 총괄연구개발과제 연구결과

제1장 총괄연구개발과제의 최종 연구개발 목표	1
1.1 총괄연구개발과제의 목표	1
가. 연구의 배경 및 필요성	7
나. 연구 목적	8
1.2 총괄연구개발과제의 목표달성도	8
1.3 국내·외 기술개발 현황	9
제2장 총괄연구개발과제의 최종 연구개발 내용 및 방법	
제3장 총괄연구개발과제의 최종 연구개발 결과	5
1장 전자기록 3대 장기보존전략	5
1절. 마이그레이션	5
1. 마이그레이션의 정의	5
2 보존전략으로서 마이그레이션의 특징	8
3. 마이그레이션 적용 사례	9
2절 XML	13
1. XML의 개념	13
2. XML 보존 포맷	15
3. XML 적용 사례	18
3절 에물레이션	19
1. 에물레이션의 개념	19
2. 보존전략으로서 에물레이션의 특징	20
3. 에물레이션 적용 사례	21
4절 장기보존전략 비용 모델	23
1. ERPANET 모델	24
2. 네덜란드 국립기록관 (Nationaal Archief of the Netherlands) 모델	27
3. TNA(The National Archives) 모델	32

4. CEDARS 모델	34
5. arXiv 모델	35
6. The LIFE 모델	37
7. 디지털 아카이빙 경제성 분석 모델	37
8. 기타 단순 비용산출 모델	38
9. 요약 및 향후 연구 과제	38
2장 마이그레이션 절차	40
1절 국내외 기존의 마이그레이션 절차 사례	40
1. ISO 전자기록물 변환 프로세스(Digital record conversion and migration process)	40
2. AHDS 이관 절차 프레임워크(AHDS Ingest Procedures Framework)	45
3. CCSDS의 생산자-기록관 인터페이스 방법론 표준 제안 (Recommendation for a producer-archive interface methodology abstract standard)	50
3장 InterPARES 성과물	54
1절. InterPARES 생산자, 관리자 가이드라인	54
1. InterPARES 생산자 가이드라인	54
2. InterPARES 보존자 가이드라인	60
2절. InterPARES에서 제시한 디지털 기록 유지와 보존 전략	65
1. A: 유지 전략	65
2. B: 보존 전략	66
4장 테스트베드 운영 및 결과	69
1절. 테스트베드 운영	69
1. 테스트베드 개관	69
2. 테스트베드 운영 개요	70
3. 이관 테스트 수행	71
2절. 테스트베드 운영 결과	74
1. 기술적 측면	74
2. 관리자(보존 업무 수행자) 측면	76
3. 생산자 측면	78
4. 도출된 이슈	80
3절. 테스트베드 결과의 일반화	80
1. 단계와 절차의 도출	81
5장 마이그레이션 절차 모형의 구축	83
1절. 업무이관에 따른 마이그레이션 개요	83
2절. 마이그레이션 절차의 비교	83
1. 절차별 공통요소 도출	83
2. 단계별 세부사항 도출	84
3절 테스트베드 결과 및 InterPARES 성과물의 적용	87
1. 마이그레이션 시 인적활동 요구사항	87

2. 마이그레이션 시 시스템 요구사항	92
6장 국제 연구 교류 활동	92
1절 InterPARES 3 프로젝트 수행	92
1. InterPARES 3 프로젝트 기반 분석	92
2. 경북대학교 대학기록관	93
3. 대구광역시 기록관	95
4. 공인전자문서보관소	97
2절 국제 활동	99
1. ICA 총회 참석	99
2. InterPARES 총회	100
제4장 총괄연구개발과제의 연구결과 고찰 및 결론	101
제5장 총괄연구개발과제의 연구성과	102
5.1 활용성과	102
5.2 활용계획	103
제6장 참고문헌	105

연구결과보고서 요약문

연구과제명	전자기록물 장기보존 전략 연구 및 테스트베드 구축		
중심단어	전자기록물, 장기보존, 보존전략		
주관연구기관	성균관대학교 산학협력단	주관연구책임자	오삼균
연구기간	2008. 5. 23 - 2008. 11. 30		
본 연구의 주요한 목적은 국내의 공공/민간 기록 관리기관의 조직과 프로그램이 다양한 이해관계자와 사회적 요구를 충족시킬 수 있는 진본 디지털 기록물 및 자원들을 장기간 보존하기 위한 전략을 마련하는데 있다. 본 연구의 또 다른 목적은 디지털 자원을 생산하고 관리하는 각급의 다양한 디지털 자원 관리기관 실무자 및 한국의 기록관리 공동체를 대상으로 연구결과를 확산함으로써 국내 전문가 공동체의 집단적 역량 강화가 이루어지는 기반을 마련하는 것이다. 또한 한국의 기록생산기관을 테스트베드로 하여 TEAM(Theoretical Elaboration into Archival Management) Korea 연구진이 얻은 연구성과를 국제사회와 공유 하도록 하는 것이다. 이에 본 연구에서는 마이그레이션, XML, 에물레이션의 이론 및 효과 한계, 사례를 조사 분석하였다. 또한, 3대 장기보존전략의 실무적용을 위한 표준절차에 대해 국제사례 분석과 중앙부처 실무절차를 근거로 이관절차에 따른 실무적용 표준절차를 개발하고 제안하였으며 국내 기록물을 대상으로 업무 및 기록물 이관 절차에 따른 구체적인 방법론을 제시하였다. 이러한 연구 결과는 장기보존전략 실무적용사례연구(Case studies)를 위한 테스트베드 운용 및 평가를 통해 도출되었으며, 중앙부처인 국토해양부를 대상으로 기록물 이관절차 운영 실태 조사 및 분석하였고 표준 절차 적용을 통한 기록물 이관 작업 효율성 및 구체성 확보하였다. 또한 본 연구의 성과를 전자기록물 장기보존 국제공동프로젝트(IP3)와 공유하였다.			

Summary

Title of Project	Study on Digital Records Long-term Preservation Strategy & Testbed Running		
Key Words	Digital records, Long-term preservation, preservation strategy		
Institute	Sungkyunkhan University	Project Leader	Sam Kyun. Oh
Project Period	2008. 5. 23 - 2008. 11. 30		
The study aims at suggesting strategies for long-term preservation for original digital resources in digital archives to meet social and diverse stakeholders’ needs in Korea. The study also purports to develop a research base to examine this area further by providing this research results, which investigated diverse practitioners in digital archives. The study also plans to share TEAM Korea’s study results with international researchers. This study investigated theories, effects, limitations, and cases of migration, XML, and emulation. It developed standard procedures, which were drawn from international as well as national cases, to apply to practice. It also suggested specific methods to task and records transfer. This standard procedure was derived from testbed application and evaluation of long-term case studies. The study investigated Ministry of Land, Transport and Maritime Affairs application and procedures of task and records transfer. The study results were shared with IP3.			

총괄연구개발과제 연구결과

제1장 총괄연구개발과제의 최종 연구개발 목표

1.1 총괄연구개발과제의 목표

가. 연구의 배경 및 필요성

(1) 연구의 배경

진본디지털 자원의 장기보존을 위한 국제 연구프로젝트인 InterPARES (International Research on Permanent Authentic Records in Electronic Systems)는 지속적으로 진본 디지털 자원의 장기 보존과 이에 관련된 이론적, 방법론적 지식기반을 마련함으로써 국제적으로 독보적인 위상을 확보한 다국적 연구프로젝트이다. InterPARES 프로젝트는 1999년 1차 프로젝트가 시작된 이래 2006년까지 2차례에 걸쳐 진행되었다. 한국에서는 본 연구진의 연구책임자와 공동연구원이 2007년부터 시작된 InterPARES 3차 프로젝트에 TEAM(Theoretical Elaboration into Archival Management) Korea라는 이름으로 한국을 대표하여 참여하고 있다.

본 연구진은 Co-Applicant로 2012년까지 향후 5년간 진행될 InterPARES 3차 프로젝트에 참여하며 연구에 기여하고, 그 결과를 공유하여 국제적 연구 성과를 한국 상황에 부합하게 적용할 방안을 모색함과 동시에, 디지털 자원 관리에 관한 우리나라의 연구 성과에 대한 국제적 위상을 확보하고자 한다.

(2) 연구의 필요성

급속한 정보화의 진행과 함께 현대사회의 모든 조직과 개인의 기록이 전자형태로 생산되고, IT 기술이 빠르게 발전함에 따라 현재의 전산환경에서 생산된 기록을 현재의 일반화된 기술로는 재현하기 어려워지는 상황이다. 또한 손실이 예상되는 디지털 자원의 양은 시간이 지날수록 기하학적으로 증가할 것으로 예측된다. 이러한 상황에서 사라져가는 중요한 진본 디지털 자원에 대한 우려와 지속적인 관리체계를 위한 요구사항이 더 필요해지고 있으며, 진본성(authenticity) 요건을 엄격하게 충족시켜야 할 필요가 있는 디지털 자원을 장기적으로 보존하는 데 있어서는 그 내용에 대한 가독성 뿐 아니라 진본으로

인정받을 수 있는 속성을 유지하여야 하는 이중의 어려움이 수반되고 있다.

따라서 디지털 자원을 장기보존하기 위해서는 기록의 보존과 재현의 가능성을 모두 확보하기 위한 보존 전략이 필수적인 상황이다. 현재 디지털 자원을 장기보존하기 위한 전략에는 크게 마이그레이션과 에물레이션이 있으며 최근에는 XML과 같은 표준 포맷으로의 변환이 주목받고 있다. 현재 디지털 자원 장기보존 전략으로서 마이그레이션은 파일의 현행화를 유지 할 수 있는, 가장 광범위하게 사용되는 방법으로서 보존된 디지털 자원을 최신 시스템의 기능을 충분히 활용하여 이용할 수 있도록 하고 있다. 마이그레이션을 실제에 적용하기 위해서는 마이그레이션 전략의 실행 과정 상 발견되는 문제들을 수집하고 분석하여, 보다 효과적이며 효율적인 마이그레이션 실행 절차를 연구하는 작업이 필요하다. 그러나 현재 국내에서는 디지털 기록물 장기보존 전략으로서 마이그레이션 전략의 실제적인 연구나 테스트베드의 운용 사례는 활발하지 않다. 이러한 맥락에서 본 연구의 결과는 국내 기록물관리분야에서 디지털기록물 장기보존 전략의 기초를 마련한다는 점에서 의의가 있다.

나. 연구 목적

본 연구의 주요한 목적은 국내의 공공/민간 기록 관리기관의 조직과 프로그램이 다양한 이해관계자와 사회적 요구를 충족시킬 수 있는 진본 디지털 기록물 및 자원들(authentic records, documents and resources)을 장기간 보존(long-term preservation)하기 위한 전략을 마련하는데 있다. 한국의 디지털 자원 생산기관이 보유한 디지털 자원 컬렉션을 테스트베드로 선정하여 InterPARES 1차 프로젝트와 2차 프로젝트로부터 얻은 연구결과를 국내에 적용하기 위한 실효성 및 타당성을 검증하고 이를 통해 국내의 기록 관리 및 보존환경에 맞는 이슈와 해법을 제시하도록 하는 것이다.

본 연구의 또 다른 목적은 디지털 자원을 생산하고 관리하는 각급의 다양한 디지털 자원 관리기관 실무자 및 한국의 기록관리 공동체를 대상으로 연구결과를 확산함으로써 국내 전문가 공동체의 집단적 역량 강화가 이루어지는 기반을 마련하는 것이다. 또한 한국의 기록생산기관을 테스트베드로 하여 TEAM(Theoretical Elaboration into Archival Management) Korea 연구진이 얻은 연구 성과를 국제사회와 공유 하도록 하는 것이다.

1.2 총괄연구개발과제의 목표달성도

○ 전자기록물 3대 장기보존전략 실무 적용 방법론 연구

마이그레이션, XML, 에물레이션의 이론 및 효과 한계, 사례를 조사 분석하여 연구 결과 1장에 기술하였다.

- 3대 장기보존전략의 실무적용을 위한 표준절차 개발

국제사례 분석과 중앙부처 실무절차를 근거로 이관절차에 따른 실무적용 표준절차를 개발하고 제안하였으며 국내 기록물을 대상으로 업무 및 기록물 이관 절차에 따른 구체적인 방법론을 연구 결과 2장에서 5장에 걸쳐 제시하였다.

- 장기보존전략 실무적용사례연구(Case studies)를 위한 테스트베드 운용 및 평가

중앙부처인 국토해양부를 대상으로 기록물 이관절차 운영 실태 조사 및 분석하였으며 표준 절차 적용을 통한 기록물 이관 작업 효율성 및 구체성 확보하였다. 이러한 내용은 연구 결과 4장에서 5장에 걸쳐 기술되었다.

- 전자기록물 장기보존 국제공동 프로젝트(IP3)에 적용 방안 연구

국내에 InterPARES 3 프로젝트를 적용하기 위한 이전 프로젝트 조사 및 분석하였으며 이를 바탕으로 InterPARES 3 프로젝트 수행 도구인 3개 영역의 사례조사연구 질문지 작성과 조사를 수행하여 결과를 공유하였다. 이 결과물들은 부록으로 첨부되었다.

1.3 국내·외 기술개발 현황

현재 기록물 이관절차에 따른 운영 S/W 실태 조사 및 분석하였으며 데이터 마이그레이션의 경우 현재, 대상 기록물의 원본 파일을 다루는 소프트웨어 개발 업체들에 의해 버전 업 된 기술을 적용하였고, 업무 이관 절차에 따른 단계별 기술적인 요구 사항 도출하였다. 현재 본 연구에서 다루고 있는 절차에 대한 구체적인 기술을 개발 보유하고 있는 곳은 국가기록원이 유일하며 국외의 경우 ISO, AHDS, CCSDS 등이 있다.

제2장 총괄연구개발과제의 최종 연구개발 내용 및 방법

본 연구의 목적을 달성하기 위해, 디지털 자원 장기보존 전략으로서 마이그레이션의 전체 실행 프로세스 상 발견되는 실제적인 문제들을 수집 분석하여 그 결과를 바탕으로 장기보존 전략의 선정, 적용, 평가의 전반적인 지표를 마련할 것이다. 또한 테스트베드로 선정한 디지털 자원 관리기관의 유형, 대상자료, 업무상황, 운영 시스템 등에 부합하는 마이그레이션 전략의 효과적인 적용방법론과 원칙을 도출할 것이다. 이렇게 도출된 지표와 적용방법론, 원칙을 유사한 유형의 다른 디지털 자원 관리기관에서의 진본 디지털 자원 보존 활동을 지원하는데 활용한다.

본 연구의 목적을 달성하기 위한 구체적인 내용은 다음과 같다.

첫째, 국내의 공공/민간 기록 관리기관의 조직과 프로그램이 다양한 이해관계자와 사회적 요구를 충족시킬 수 있는 진본 전자 기록(authentic records)을 장기간 보존(long-term preservation)하기 위한 지식기반을 구축하고 실행전략을 개발

- 전자기록물 3대 장기보존전략 방법론 확립 및 적용 방안 개발
- 실무 적용을 위한 전략별 테스트베드 환경 구축 및 운영
- 각급 기록물관리기관에 실무 적용 평가 및 가이드라인 제시

둘째, 한국의 기록생산기관을 테스트베드로 하여 TEAM(Theoretical Elaboration into Archival Management)Korea 연구진이 얻은 연구 성과를 국제사회와 공유

- 전자기록물 장기보존 분야의 국제 공동연구프로젝트(IP3)와 공동 연구

제3장 총괄연구개발과제의 최종 연구개발 결과

디지털 기술의 발전으로 인해, 정보를 생산하고, 보관하고, 배포하기 위해 필요한 노력은 크게 감소하였다. 이로 인해, 현재 다양한 기록물들이 전자적으로 생산되거나, 혹은 전자적으로 변환되어 활용되고 있다. 그러나 전자적 정보의 가장 큰 특징은, 실체(object)를 가지고 있어도, 내용(content)을 확인할 수 없는 경우가 발생한다는 것이다. 즉, 전자기록의 장기보존이란, 전자기록 그 자체의 보존 뿐 아니라, 이에 대한 접근(access)의 보존을 의미하는 것으로 이해되어야 한다.

전자기록의 장기보존전략은 이와 같은 접근을 가능하게 하기 위한 다양한 방법론과 고려사항을 의미 한다. 전자기록물은 매우 다양한 논리적 구조를 가지고 있으며, 각각이 다양한 성격을 가지고 있음으로, 장기 보존 전략의 선택은 자원의 특성과 어떤 특성을 보존하려 하는지에 따라 다양하게 선택될 수 있다.¹⁾

본 장에서는 전자 기록으로의 접근을 보존하기 위한 3대 전략으로서 마이그레이션(migration), XML로의 변환(conversion to XML), 에뮬레이션(emulation)의 개념과 특성을 살펴보고, 각각의, 구체적인 시행과 사례, 그 장단점을 확인하려 한다.

1절. 마이그레이션

1. 마이그레이션의 정의

마이그레이션(migration)이라는 용어는 그 범위, 대상, 관점에 따라 매우 다양한 용도로 사용되고 있다. 컴퓨터 및 정보 과학 분야에 있어서도 마이그레이션의 정의는 서로 매우 다른 것을 볼 수 있다. 따라서, 우선 다양한 마이그레이션의 정의를 살펴볼 필요가 있다.

세계적인 IT 분야 온라인 용어 사전인 whatis.com에서는 마이그레이션을 ‘마이그레이션이란, 하나의 운영 환경으로부터 다른 운영환경으로 이전하는 프로세스를 뜻한다. 여기서 다른 환경은 일반적으로 더 나은 환경을 뜻한다.’라고 정의하고 있다.²⁾

온라인 백과사전인 wikipedia는 컴퓨터 분야에 있어 마이그레이션을 세 종류로 구분하고 있다. 각각은 데이터 마이그레이션(data migration), 프로세스 마이그레이션(process

1) PADI(Preserving Access to Digital Information) Thesaurus <<http://www.nla.gov.au/padi/topics/thesaurus.html>>

2) whatis.com The leading IT encyclopedia and learning center <<http://whatis.com>>

migration), 시스템 마이그레이션(system migration)이다. 이 중, 정보학 분야에 해당하는 마이그레이션은 데이터 마이그레이션이다. wikipedia에 따르면, 데이터 마이그레이션이란 ‘저장 매체, 저장 포맷, 컴퓨터 시스템 사이에서 데이터의 전송 프로세스’를 뜻한다.

미국의 국립 결핵 통제 협회 (NTCA, National Tuberculosis Controllers Association)는 결핵환자 관리를 위한 응용 프로그램의 평가를 위한 항목 중 하나로 데이터 마이그레이션을 정의하고 있다. 이에 따르면, 데이터 마이그레이션이란 ‘과거의 데이터를 새로운 시스템으로 불러오는 프로세스’를 뜻한다.³⁾

지금까지 살펴본 정의는 일반적인 사전 및 컴퓨터 과학 분야에서 이야기하고 있는 마이그레이션의 정의였다. 이를 종합하면, 마이그레이션이란 ‘과거의 시스템 및 운영 환경에서 사용되는 데이터를 더 나은 시스템 및 운영환경으로 이전하는 프로세스’라고 정의될 수 있다.

전자 기록물의 보존에 있어서 마이그레이션의 정의는 위의 정의에 비해 좀 더 분명한 목적성과 속성을 포괄하고 있다. 또한 마이그레이션(migration)과 전송(transfer)의 구분 및 대상에 따른 마이그레이션의 특성 구분까지 발견할 수 있다.

IFLA(International Federation of Library Associations and Institute)의 디지털 정보 보존 보고서에서는, 마이그레이션을 ‘하나의 하드웨어/소프트웨어 환경으로부터 다른 것으로, 혹은 하나의 컴퓨터 기술로부터 이어지는 기술로 주기적으로 이전하는 것’으로 정의한다.⁴⁾

InterPARES(International Research on Permanent Authentic Records in Electronic Systems)의 용어 데이터베이스에서는, 좀더 구체적으로 기록물의 마이그레이션이란 용어를 사용하고 있다. 이에 따르면, 마이그레이션은 ‘시간이 흐름에 따라 시스템과 기록매체가 유실되거나, 노후 될 수 있는 상황에서, 기록물의 지속적인 접근가능성을 보장하기 위해, 기록물을 하나의 시스템 또는 저장 매체로부터 다른 것으로 이전하는 프로세스’라고 정의된다.⁵⁾

ISO(International Organization for Standardisation)는 InterPARES의 정의와 매우 유사하지만, 마이그레이션이 필요한 경우의 수를 한 가지 추가하고 있다. 이에 따르면, 마이그레이션은 ‘시간이 흐름에 따라 시스템과 기록매체가 다른 것으로 대체되거나, 유실되거나, 노후 될 수 있는 상황에서, 기록물의 지속적인 접근가능성을 보장하기 위해, 기록물을 하나의 시스템 또는 저장 매체로부터 다른 것으로 이전하는 프로세스’이다.⁶⁾

가장 많이 인용되고 있는 아카이브 시스템 참조모델인 OAIS 참조모델은 그 프레임워

3) NTCA Assessment Guide <<http://www.infotech.net/ntca/>>

4) Preserving Digital Information, Report of the Task Force on Archiving of Digital Information, 1996
<<http://www.ifla.org/documents/libraries/net/tfadi-fr.pdf>>

5) InterPARES 2 project Terminology database<http://www.interpares.org/ip2/ip2_terminology_db.cfm>

6) Information and documentation - Digital records conversion and migration process, ISO TC46/SC11/WG, 2008

크 내에 디지털 마이그레이션을 포함하고 있다. 여기서 사용된 마이그레이션의 정의는 OAIS의 관점 안에서 구체적인 요건을 포함하고 있다. 이에 따르면 ‘디지털 마이그레이션은 OAIS내에서의 보존을 목적으로 한 디지털 정보의 전송(transfer)이다. 마이그레이션은 아래 세 가지 관점에서 전송(transfer)과 구분된다. 1) 정보 콘텐츠 전반을 보존할 것에 주목한다. 2) 새로운 정보의 구현이 이전 정보의 대체가 되는 것이다. 3) 전송의 모든 측면에 관한 전체적인 통제와 책임이 OAIS내에 있다.’⁷⁾ 이를 종합하면, <표 1-1>과 같다.

<표 1-1> 마이그레이션의 정의

출처	정의	대상 / 범위	시기 / 목적
whatis.com	하나의 운영 환경으로부터 다른 운영환경으로 이전하는 프로세스를 뜻한다. 여기서 다른 환경은 일반적으로 더 나은 환경	운영 환경	더 나은 환경으로 이전
wikipedia	저장 매체, 저장 포맷, 컴퓨터 시스템 사이에서 데이터의 전송 프로세스	매체, 포맷, 컴퓨터 시스템	
NTCA	과거의 데이터를 새로운 시스템으로 불러오는 프로세스	과거의 데이터	새로운 시스템
IFLA	하나의 하드웨어/소프트웨어 환경으로부터 다른 것으로, 혹은 하나의 컴퓨터 기술로부터 이어지는 기술로 주기적으로 이전하는 것	하드웨어/소프트웨어 환경 컴퓨터 기술	차세대 기술로 이전
InterPARES	시간이 흐름에 따라 시스템과 기록매체가 유실되거나, 노후될 수 있는 상황에서, 기록물의 지속적인 접근가능성을 보장하기 위해, 기록물을 하나의 시스템 또는 저장 매체로부터 다른 것으로 이전하는 프로세스	시스템/저장매체	시스템 및 매체 유실, 노후 상황
ISO	시간이 흐름에 따라 시스템과 기록매체가 다른 것으로 대체되거나, 유실되거나, 노후될 수 있는 상황에서, 기록물의 지속적인 접근가능성을 보장하기 위해, 기록물을 하나의 시스템 또는 저장 매체로부터 다른 것으로 이전하는 프로세스	시스템/저장매체	시스템 및 매체 대체, 유실, 노후 상황
OAIS	OAIS내에서의 보존을 목적으로한 디지털 정보의 전송(transfer)이다. 마이그레이션은 아래 세가지 관점에서 전송(transfer)과 구분된다. 1) 정보 콘텐츠 전반을 보존할 것에 주목한다. 2) 새로운 정보의 구현이 이전 정보의 대체가 되는 것이다. 3) 전송의 모든 측면에 관한 전체적인 통제와 책임이 OAIS내에 있다.	정보 콘텐츠	OAIS 내에서의 AIP 보존

지금까지 살펴본 정의를 종합하면, ‘마이그레이션이란, 시간이 흐름에 따라, 기록매체, 포맷, 시스템이 대체되거나, 유실되거나, 노후화 될 수 있는 상황에서, 기록물로의 지속적

7) OAIS Reference Model for an Open Archival Information System(OAIS), CCSDS, 2002
<<http://public.ccsds.org/publications/archive/650x0b1.pdf>>

인 접근을 가능하게 하도록 하기 위해, 차세대의 운영 체제, 저장 매체, 포맷, 시스템으로 이전의 데이터를 이전하는 프로세스이다.’ 여기서 프로세스라 함은, 마이그레이션이 단일 행위가 아님을 나타낸다. 즉, 마이그레이션은 다양한 과정을 내포하고 있으며, 각 과정에 대한 고려사항이 존재함을 나타낸다.

마이그레이션의 특성을 구분할 수 있는 가장 큰 특징은 마이그레이션의 수행 방법, 즉 변환되려는 대상과 목표 상태를 무엇으로 규정하는가, 이다. OAIS 참조 모델은 이에 따라 마이그레이션의 유형을 4종류로 구분하고 있다. ISO의 마이그레이션 프로세스에서는 이 구분에 대한 일반적인 설명을 제공한다. 이를 종합하면 아래와 같다.

- 갱신(refreshment) : 하나의 실제 매체로부터, 같은 유형의 다른 실제 매체로 bit를 복사하는 것이다. 이 것은, 시스템의 변화 없이 기록물로의 접근을 유지할 수 있다. 예를 들면, CD-R 데이터를 다른 CD-R로 이동하는 것이 이에 해당한다.
- 복제(replication) : 패키지 정보에 변화가 있는 bit의 복제이다. 이 경우는, 같은 매체 혹은 새로운 매체로의 복사를 포함한다. 즉, 재생은 복제에 포함되는 활동이다. 복제는, 기록물로의 접근을 위해 시스템의 변화를 필요로 한다. 예를 들면, 하드 디스크의 데이터를 CD-R로 이동하는 것이 이에 해당한다.
- 재패키지화(repackaging) : 데이터를 하나의 구조로부터 다른 구조로 변환하는 것이다. 이 때, 콘텐츠는 변화되지 않는다. 예를 들면, 복잡한 파일 디렉토리를 가진 데이터를 간소화된 디렉토리로 복사하는 것이다.
- 변형(transformation) : 마이그레이션 과정에서 데이터가 변형되어 이동되는 것이다. 예를 들면, BMP 포맷의 파일을 JPEG 파일로 변형하는 것이다. 이 경우라도, 전체 정보 자체는 보존되는 것으로 정의된다.

변형은 역행이 가능한가, 불가능한가의 여부로 다시 구분될 수 있다. 예를 들어, 목표 포맷을 미리 정의하고 변형을 시행한 경우는, 역행이 가능하지만, 그렇지 않을 경우는 역행이 불가능할 수 있다.

2 보존전략으로서 마이그레이션의 특징

파일의 논리적 구조를 변경하여 새로운 시스템에서 기존의 파일을 사용하는 방법은 이미 일반적으로 사용되고 있는 방법이다. 예를 들어, 최신버전의 워드프로세서에서 기존의 워드프로세서의 파일에 접근할 때, 기존 워드프로세서의 파일 포맷을 최신 버전 포맷에 맞도록 변환하는 행위와 같은 방식이다. 운영체제의 버전 업에 따른 시스템 변환이나, RDBMS 스키마의 이전등과 같은 행위도 여기에 해당한다.

보존전략으로서의 마이그레이션의 기술적인 측면은 앞서 설명한 변환 행위와 같은 것

으로 볼 수 있다. 보존전략으로서의 마이그레이션이란, 마이그레이션을 통해, 현재의 시스템을 이용하는 환경에서, 이전의 시스템에서 생성된 전자기록물을 사용할 수 있는 형태로 유지하는 것을 뜻한다.

장기보존 전략으로서 마이그레이션이 가지는 가장 큰 장점은 최신의 기술과 환경에서 기록물을 이용할 수 있다는 것이다. 즉, 현재의 환경에 변화를 주지 않고, 기존의 기록물을 사용할 수 있기 때문에, 추가적인 교육이나 훈련이 필요하지 않다. 또한, 데이터를 재사용하거나, 자유롭게 편집, 활용할 수 있는 형태로 보존할 수 있다.

또 다른 장점은, 변환이 쉽고, 다양한 도구가 제공된다는 것이다. 예를 들어, 다양한 파일 포맷에 대해, 해당 포맷의 공급자를 비롯한 다양한 관련자로부터 변환도구가 지속적으로 개발 공급되고 있기 때문에, 실행에 있어 개발 및 시행을 위한 비용이 적다.

마이그레이션 전략은 큰 장점을 가지고 있음에도 불구하고, 그 자체로 해답이 되지 않고 있다. 마이그레이션의 단점 또한 무시할 수 없는 수준이기 때문이다.

마이그레이션이 가지는 가장 큰 단점은 결국 데이터 손실이 위험이 상존한다는 것이다. 마이그레이션의 방법은 원본의 논리적 구조 혹은 처리(rendering) 방법을 다르게 하여, 새로 작성하는 방법이다. 즉, 마이그레이션을 통한 보존은 원본을 본래의 상태대로 보존하는 것을 뜻하지 않는다. 이와 같은 변환의 과정에서 언제든지 데이터의 일부가 변경될 가능성을 가지고 있다. 마이그레이션 된 디지털 개체는 최초의 원본(original)이란 개념이 아닌 진본(authentic copy)의 개념으로서 다루어 져야 하며, 진본성을 보장하기 위한 다양한 수단을 필요로 한다.

위와 일맥상통하는 마이그레이션의 단점으로, 디지털 개체의 특성이 상실될 우려가 있다. 즉, HTML 등과 같은 상호작용적인 데이터의 경우, 특정한 포맷으로 변환하였을 때, 상호작용의 콘텐츠를 사용할 수 없게 되는 경우가 생길 수 있다. 또한 복잡한 레이아웃의 외형이 손상될 가능성도 안고 있다.

마이그레이션이 가지는 또 다른 단점은 지속적으로 실행해야만 한다는 것이다. 디지털 환경이 변경되는 때 시점마다 마이그레이션 업무를 수행해야할 필요가 있다. 이는, 장기적인 마이그레이션 전략의 수립을 어렵게 하는 한 요소이며, 마이그레이션 전략이 장기화 될수록 비용이 점차 증가하게 될 가능성을 제기한다.

3. 마이그레이션 적용 사례

본 절에서는, 보존전략으로서 마이그레이션을 개발한 사례를 통해 마이그레이션의 수행 시 고려해야할 사항에 대해 좀 더 구체적으로 알아보기로 한다.

가. CAMiLEON (Creative Archiving at Michigan and Leeds Emulating the Old On the New) : Migration of request

CAMiLEON은 디지털 자원의 장기 보존을 위한 기술적인 전략을 광범위하게 개발하고 평가하는 과정을 통해, 상황에 따라 어떤 장기 보존 전략을 적용할 수 있을지에 대해 연구하는 연구 프로젝트이다. CAMiLEON 프로젝트는 미국 Michigan 대학 과 영국 Leeds 대학의 공동 연구로 이루어졌다.

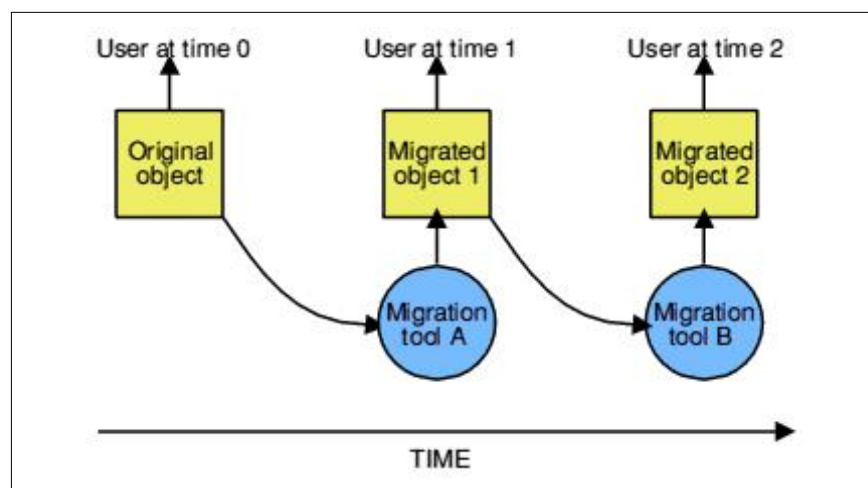
CAMiLEON 프로젝트의 목표는 원본 디지털 개체의 기능 및 ‘외형과 느낌’을 유지할 수 있는 장기 보존 전략을 탐색하고, 장기보존 전략으로서 에뮬레이션 기술을 연구하며, 어떤 상황에 에뮬레이션 전략이 적합한지를 확인하는 파악하는 것이다.

CAMiLEON 프로젝트 연구진은 마이그레이션이 가장 흔하게 시행되는 기록 보존의 기술이지만, 기존의 마이그레이션이 가진 잠재적인 비용의 증가와 외형의 변화를 심각한 문제점으로 생각하였다. 이에 따라, 마이그레이션의 지속적인 수행에서, 원형의 훼손을 최소화할 수 있는 마이그레이션 방법론을 연구 하였고, 그 결과 Migration on request 시행법을 제시하였다.

또한, Migration on request의 방법론을 시행하기 위해 벡터 그래픽(vector graphics) 자원을 대상으로 테스트를 수행하였고, 현재는 테스트에 사용된 변환 도구를 배포하고 있다.⁸⁾

기존의 마이그레이션 방법의 경우, <그림 1-1>과 같이 변환 과정에서 파일에 원본과 차이점이 발생할 경우, 향후 지속적인 변환에서, 이 오류가 계속 누적된다. 이 것은, 원본의 정확성에 치명적인 문제점을 불러온다. 또한, 마이그레이션이 수행될 때마다, 새로운 파일 포맷이 형성됨으로, 새로운 환경이 등장할 때마다 새로운 마이그레이션 도구가 필요해진다.

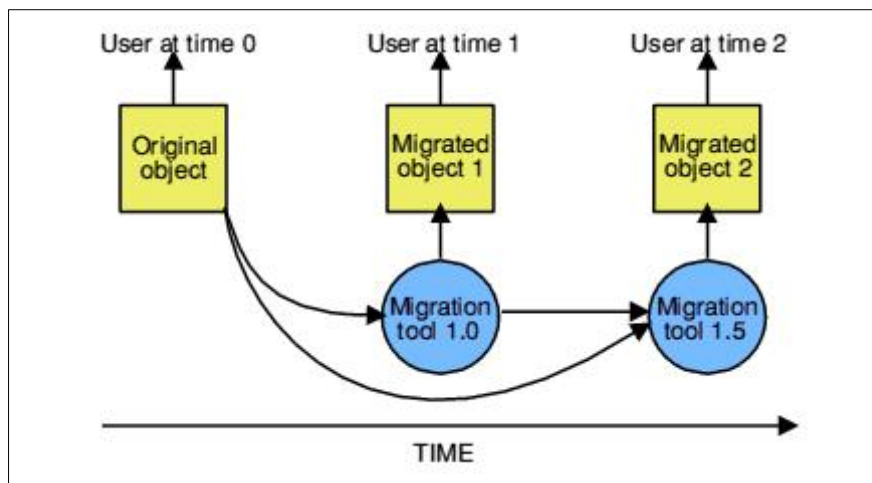
<그림 1-1> 기존의 마이그레이션 접근



8) Migration on request, CAMiLEON project
<http://www.si.umich.edu/CAMiLEON/reports/mor/>

Migration on request는 이러한 단점을 극복하기 위해, 하나의 원본파일을 보존하며, 새로운 환경이 등장할 때마다, 이에 적합한 마이그레이션 도구를 개발하는 접근법을 제시하고 있다. 이를 통해, 변환의 과정에서 발생한 오류가 누적되는 현상을 피할 수 있으며, 새로운 환경이 등장할 때마다, 완전히 새로운 도구를 개발할 필요 없이, 하나의 포맷을 대상으로 하는 변환도구를 개발할 수 있게 된다. <그림 1-2>는 Migration on request의 개념을 표현하고 있다.

<그림 1-2> Migration on request



나. DPT (Digital Preservation Testbed)

DPT는 네덜란드 국립 아카이브에 의해 설립된 디지털 장기보존 부서(Digital Longevity department)에 의해 운영되는 장기보존 전략 테스트베드 프로젝트이다. DPT는 2005년 문서, 데이터베이스, 스프레드시트를 대상으로 하는 세 가지 유형의 마이그레이션 전략 접근방법에 대한 테스트베드 운영결과를 발표하였다.⁹⁾ 세 전략은 소급호환성 유지, 상호운용성 유지, 표준으로의 변환이다.

- 소급호환성 유지(Backward compatibility) : 소급호환성 유지란, 한 응용프로그램의 기존 버전에서 생성된 기록을 동일한 응용프로그램의 최신 버전에서 정확히 재생산하는 것을 말한다.
- 상호운용성 유지(Interoperability) : 상호운용성 유지는 특정한 하드웨어와 소프트웨어의 조합에 대한 기록의 의존성을 제거하는 것이다. 즉, 동일한 파일을 하나의

9) From digital volatility to digital permanence: Preserving text documents, DPT
<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-textdocs-en.pdf>
 From digital volatility to digital permanence: Preserving spreadsheets, DPT
<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-spreadsh-en.pdf>
 From digital volatility to digital permanence: Preserving databases, DPT
<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-databases-en.pdf>

플랫폼에서 다른 플랫폼으로 이전하더라도, 이전과 동일한 재생산이 가능하도록 하는 것을 뜻한다. 이를 위해서는 소프트웨어간의 상호운용성을 유지하거나, 중간변환 프로그램을 활용할 수 있다.

- 표준으로의 변환(Conversion to standards) : 표준 포맷으로의 변환은 독점 포맷인 파일을 공개 표준에 기반한 포맷으로 변경하는 것을 뜻한다. 이를 통해, 전자 기록은 생성될 당시의 소프트웨어와 하드웨어에 대한 의존성에서 벗어날 수 있게 된다. 결과적으로, 전자기록물을 전자적 소멸의 위기로부터 벗어나도록 한다.

(1) 텍스트의 마이그레이션

텍스트에 대한 소급호환성 유지 전략의 시행 결과, 소급호환성 유지 전략은 단기적인 보존에 적합한 것으로 나타났다. 소급 호환성 유지 전략은 단기적으로 보았을 때, 기록의 진본성에 큰 위협이 되지 않았다. 즉, 문서의 구조와 외형을 잘 유지하였다. 그러나, 3,4세대 이전의 파일을 최신 소프트웨어로 처리할 때, 문서의 콘텐츠에 큰 변화를 야기하는 것이 발견되었다. 또한 이 전략은 소프트웨어 공급자로부터의 종속성에서 벗어날 수 없다. 따라서, 소급호환성 유지 전략은 단기 보존에는 적합하지만, 장기 보존에는 문제가 있는 것으로 드러났다.

상호운용성의 유지 전략은 가장 널리 사용되고 있는 두 종류의 워드프로세서를 대상으로 테스트되었다. 그 결과, 상호운용성 유지 전략 역시, 단기 보존에는 적합하나, 장기 보존에는 적합하지 않은 것으로 드러났다. 상호운용성 유지 전략은 소급호환성 유지전략과 같은 한계를 보였다. 즉, 중간변환 프로그램에 의해, 원본의 기록이 다른 형태로 변환되는 경우가 발생하였고, 소프트웨어 공급자에 대한 의존성에서 벗어날 수 없었다.

표준으로의 변환 전략은 그 자체로, 소급호환성 유지와 상호운용성 유지가 가능하다. 또한, 독점적 공급자에 대한 의존을 제거할 수 있다. 따라서, 장기보존에 적합한 접근법이다. DPT는 표준으로의 변환 전략을 테스트하기위해 PDF와 XML로의 변환을 테스트 하였다. PDF로의 변환은 원본 문서의 기능, 외관, 구조를 보존하는 것에 적합했다. XML은 콘텐츠 뿐 아니라, 구조와 외형, 기록의 문맥에서 텍스트를 표현할 수 있다. 또한, 손쉽게 메타데이터를 추가할 수도 있다.

(2) 스프레드시트의 마이그레이션

스프레드시트에 대한 소급호환성 유지 전략은 텍스트의 경우와 마찬가지로 단기 보존에 있어 적합한 것으로 나타났다. 즉, 비교적 짧은 기간 사이의 소급호환성 유지는 중요한 변화를 보이지 않았다. 그러나, 기간이 길어질 경우, 문제점이 발생한다. 소프트웨어 공급자에 대한 의존성도 동일하게 존재했다.

상호호환성 유지 전략은 스프레드시트의 보존에 적합하지 않은 것으로 나타났다. DPT는 LOTUS 1-2-3와 마이크로소프트 Excel 사이의 상호호환성을 시험 했다. 결과, 두 포맷 사이의 수식이 호환되지 않는 것을 발견했으며, 내용과 형태, 기능 모두 손상되었다.

DPT는 스프레드시트를 변환하는 표준으로, PDF와 XML을 테스트 하였다. PDF의 경우, 수식과 셀의 값은 모두 변환 되었다. 그러나, 수식이 계산된 결과를 표현하지 못하였다. 즉, 기본 입력된 표현 층만을 표현할 수 있을 뿐 수식 계층과 표현 층과의 연결성이 상실되었다. DPT는 스프레드시트의 보존을 위해 XML 포맷이 적합하다는 것을 발견하였다. 그러나, XML은 처리하는 도구에 따라 매우 다른 결과를 나타낸다는 단점이 존재한다.

(3) 데이터베이스의 마이그레이션

데이터베이스의 마이그레이션에 있어서도, 소급호환성 유지 전략은 단기 보존에만 적합한 것으로 나타났다. 데이터베이스의 변환은 큰 문제를 나타내지 않았다. 그러나, 일반적으로 신 버전의 데이터베이스들은 한정된 수의 구 버전만을 지원하여, 구버전의 데이터베이스를 신버전에 저장하는 것은 적합하지 않았다. 또한, 지속적인 변환의 필요나, 공급자에 대한 의존성에 있어서도 위의 두 결과와 같았다.

상호호환성 유지 전략을 데이터베이스에 적용하는 것은 일반적인 변환에 비해 더 많은 기술이 필요하다. 소규모의 개인적인 데이터베이스는, ASCII, RTF와 같은 포맷으로 간단하게 변환이 가능하지만, 복잡한 데이터베이스 시스템의 변환은 일반적으로, 데이터 및 스크립트를 새로 작성하는 작업이 필요하며, 이 과정에서 데이터를 잃을 위험이 존재한다.

DPT는 데이터베이스를 변환하는 표준으로 ASCII, SQL, XML을 테스트 하였다. ASCII 파일로의 변환은 일반적으로 모든 시스템에서 읽을 수 있는 문서 형식의 데이터를 생성하였다. 그러나, 이와 같은 방식은 데이터베이스의 구조를 표현할 수 없다는 한계를 가진다. XML은 ASCII의 장점을 가지는 동시에 데이터베이스의 구조를 표현해 줄 수 있다. SQL로 변환은 표준 SQL로의 변환을 의미한다. 각 공급자에 따라 다른 SQL을 표준 SQL로 변환하는 것이 필요하다.

2절 XML

1. XML의 개념

XML(eXtensible Markup Language)은 마크업 언어의 일종이다. 마크업이란 특별한 의미나 추가적인 정보를 문서의 특정 부위에 더하기 위한 태그를 뜻하며, 마크업을 이용해서 문서 표현을 다루는 언어가 바로 마크업 언어이다. 대표적인 마크업 언어에는 HTML이 있다. HTML은 편리하고 쉽게 사용할 수 있다는 장점을 가지고 있으나, 고정적이고 확장이 불가능하다는 단점이 있다. 또 다른 마크업 언어인 SGML은 유연성이나 확장성은 뛰어나지만, 지나치게 복잡하여 관련 소프트웨어 개발이 쉽지 않다는 이유로 대중화되지 못하였다. XML은 구조화된 문서를 정의하고 자유롭게 태그를 정의할 수 있다는 SGML의 장점과 인터넷 상에서 손쉽게 하이퍼미디어 문서를 제공할 수 있는 HTML의 장점을 그대로 보유하고 있다. XML은 1996년 W3C(World Wide Web Consortium)의 후원으로 형성된 XML Working Group에 의해 개발되었고, 1998년에 W3C권고에 따르는 XML 1.0 명세서가 발표되었다. XML의 주요한 특징은 다음과 같다.¹⁰⁾

- 확장성(Extensibility) : XML은 사용자가 임의로 태그를 정의할 수 있도록 한다. 따라서 사용자는 문서의 태그에 자신이 원하는 적합한 의미를 내포하는 이름을 정의하고 사용할 수 있다.
- 문서 구조(Structure)의 정의 : XML은 사용자가 태그를 생성하면서 각각의 태그들 간의 관계를 정의함으로써 구조적인 문서를 작성할 수 있도록 지원한다.
- 유효성 검사(validity and well-formedness constraint) : HTML은 불완전한 문법 검사로 인하여 문서상에 오류가 존재하더라도 무시되고 처리되는 부분이 있지만, XML은 문서 안에서 문법적인 구별이 가능하다. 따라서, 유효성 검사와 편집상의 제어기능이 제공되어 어플리케이션이 문서를 받아들일 때 문서의 오류를 쉽게 판단할 수 있도록 한다.
- 문서의 내용과 표현의 분리 : XML 문서는 내용과 표현(Presentation)에 대한 정의를 분리하고 있는데, 이러한 특징은 같은 문서에 대하여 다양한 표현 방식이 가능하도록 지원하며, 재사용을 가능하게 한다.

보존 전략으로서, XML의 활용은 마이그레이션의 일종으로서 표준으로의 변환과, 문서의 생성 시점부터 XML 포맷을 사용하는 방법이 가능하다. 보존을 위한 포맷으로서, XML의 장점은 다음과 같다.¹¹⁾

10) 이성호. 2006. 시장을 바꾸는 힘. SW정책연구센터@KIPA 한국소프트웨어진흥원
<http://www.software.or.kr/ICSFiles/afiedfile/20060502/01.PDF>

11) Electronic Records Management and Digital Preservation: Protecting the Knowledge Assets of the State Government Enterprise Part III: Management Leads and Technology Follows -. but Collaboration is King!, NASCIO, 2007
<http://www.nascio.org/publications/documents/NASCIO-RecordsManagementPart3.pdf>

- XML은 특정한 플랫폼에 의존적이지 않다.
- XML은 마크업 언어임으로, 문서의 형태를 보존할 수 있다.
- XML은 텍스트 문서임으로, 인간에 의해 해석이 가능하다.
- XML은 문서 구조, 필드 명, 콘텐츠를 동시에 문서 내에 포함할 수 있다. 문서의 구조에 대한 정보는 별도의 DTD 혹은 schema 파일로 저장할 수 있다.
- 많은 디지털 파일들을 직접 XML로 변환할 수 있다.
- XML은 플랫폼에 의존성이 없음으로, 변환의 횟수를 매우 줄일 수 있다. 이 것은, 문서의 진본성 보존을 용이하게 한다.
- 이미 XML을 지원하는 소프트웨어 개발 환경이 개발되어 있다.
- 향후, 오피스 프로그램들은 XML을 사용할 것으로 예상된다.

반면, XML의 문제점은 다음과 같이 지적된다.

- 모든 바이너리 포맷이 XML로 변환되는 것은 아니다. 이에 따라, 다양한 기관에서, XML에 다른 포맷을 포함시키고 있다. 이것은, 보존전략에 있어, 이전과 같은, 문제를 지속시킨다.
- XML 요소의 명칭이 합의되지 않은 경우, 같은 요소에 2개 이상의 명칭이 사용될 수 있다. 이것은 요소간의 맵핑을 필요하게 한다.
- XML은 보존전략을 포함하는 것이 아니다. 즉, 보존 전략으로서 XML의 사용은, 여전히 관리적인 통제를 필요로 한다.
- 보존된 XML 파일의 표현은, 원본의 표현과 다를 수 있다. 표현은 이용자가 사용하는 해석 프로그램에 의존하기 때문이다.

이와 같은 문제점이 있지만, XML은 여전히 장기보존에서 발생하는 문제점을 해결할 수 있는 대안으로 생각되고 있다. XML의 사용은 바이너리 포맷이 안고 있는 공급자 의존성과 기술의 망실 가능성을 제거할 수 있는 공개형 표준이며, 구조화된 문서이기 때문이다. 단, XML의 사용이 모든 문제를 해결할 수 있는 해답이라고 볼 수는 없으며, 향후 연구가 필요하다는 것을 기억해야 한다.¹²⁾

2. XML 보존 포맷

위와 같은 XML 포맷의 장점을 활용하여, 문서 포맷의 표준화를 이루기 위한 많은 시

12) XML and Digital Preservation, DPT, 2002
 <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white-paper_xml-en.pdf>

도가 있었다. 1999년 StraOffice XML라는 명칭으로 XML을 이용한 개방형 표준 문서 포맷의 개발이 시작되어, 2005년 OASIS에 의해 ODF(Open Document Format)이 표준으로 승인 되었으며, 이듬해 ISO 표준으로 채택되었다. 2005년 표준화 기구인 ECMA International은 마이크로소프트 오피스 개방형 XML 포맷의 표준화에 착수하여, 2008년 현재, OOXML(Office Open XML)이란 명칭으로 ISO 표준으로 채택되었다. 본 절에서는, 개방형 문서 포맷의 표준인 ODF와 OOXML에 대하여 살펴보기로 한다.¹³⁾

가. ODF (Office Document Format)

현재 ODF와 OOXML은 오피스 문서의 XML 포맷을 놓고 치열한 표준 경쟁을 벌이고 있다. 특징적인 것은, ODF가 특정한 공급자에게 종속되지 않는 문서의 상호운용성을 강조하는 반면, OOXML은 기존의 문서 포맷과 완벽히 호환됨을 강조하고 있다는 것이다. 이에 따라, ODF와 OOXML 진영은 상호비교를 통해, 그 장점을 드러내려 노력하고 있다. 개방형 문서 표준으로서 ODF의 장점은 다음과 같다.

- ODF는 mixed content model을 사용한다. Mixed content model은 워드문서와 같이 서술적(narrative)인 정보를 표현하기에 유리하고, XHTML과의 상호전환이 용이하다. 반면에 마이크로소프트가 채택하고 있는 Non-mixed content model은 구조적인(structured) 데이터를 표현하는데 유리하다.
- ODF는 콘텐츠와 표현을 구별하는데 용이하며, 하이퍼링크를 보다 효과적으로 처리하도록 디자인 되었다.
- ODF는 보다 많은 기존의 개방형 표준을 재활용 한다 (예 : 도면을 위한 SVG, 방정식을 위한 MathML, 연결을 위한 XLink, 메타데이터를 위한 더블린 코어 등). 마이크로소프트 오피스는 MathML과 더블린 코어 정도만을 재활용한다.
- ODF 진영은 마이크로소프트의 개방형 XML 서식이 XML 데이터 모델을 ODF 포맷만큼 간결하게 운용하지 않고 있고, 문서 구조보다는 자사 오피스 어플리케이션의 내부 데이터 구조를 밀접히 반영하도록 제작되었다고 비판한다.

현재, ODF 포맷은, OpenOffice 어플리케이션 뿐만 아니라, AbiWord, Google Docs 등 다양한 소프트웨어에서 사용되고 있다.

나. OOXML (Office Open XML)

13) 이성호. 2006. 시장을 바꾸는 힘. SW정책연구센터@KIPA 한국소프트웨어진흥원
<http://www.software.or.kr/ICSFiles/afieldfile/20060502/01.PDF>

OOXML의 특징 역시, 흔히, ODF와 비교하여, 다음과 같이 설명된다.

- OOXML 표준화의 주된 목표는 수 십년간 생산된 수많은 문서와의 100% 소급호환성의 확보와 동시에 다양한 최신기능의 완벽한 지원이다.
- 마이크로소프트는 ODF의 서식이 마이크로소프트 오피스 어플리케이션의 모든 기능을 완전히 지원하지 못하여, ODF의 서식으로 전환하여 저장할 때 기능상의 손실이 초래될 수 있다고 주장한다.
- 오피스 2003에 적용된 개방형 XML 서식은 고객과 개발자에 의한 광범위한 현장 시험과 사용을 거쳤다.
- 문서양식을 정의하는 XML뿐 아니라, 조직의 비즈니스 프로세스 요구에 따라 고유하게 정의된 XML 스타일을 통해 오피스 문서가 정보를 생성하고 전송할 수 있게 함으로써 새로운 차원의 업무혁신을 가능하게 한다.

ODF와 OOXML의 명세를 간략히 비교하면, <표 1-2>와 같다.

<표 1-2> ODF와 OOXML의 명세 비교

파일 포맷	Office Open XML	OpenDocument Format
개발 주체	마이크로소프트	StarDivision, 선 마이크로시스템즈
표준화 주체	Ecma International, ISO/IEC	OASIS, ISO/IEC
최초 공개일	2006	2005
최신 버전	ISO/IEC IS 29500:2008	Open Document Format for Office Applications (OpenDocument) v1.1
최신 표준 버전	ISO/IEC IS 29500:2008	ISO/IEC IS 26300:2006
언어 유형	XML	XML
압축 포맷	ZIP	ZIP
컨테이너 구조	Open Packaging Convention	
메타데이터 포맷	Dublin Core	Dublin Core 일부 urn:oasis:names:tc:opendocument:xmlns:meta elements
수식 표현 언어	Office MathML (OMML)	정의되지 않음, MathML 2.0
스프레드시트 수식 언어	표준에 포함	어플리케이션 규정
도면 언어	DrawingML, VML	urn:oasis:names:tc:opendocument:xmlns:drawing elements

파일 포맷	Office Open XML	OpenDocument Format
		urn:oasis:names:tc:opendocument:xmlns:svg-compatible elements
매크로 언어	어플리케이션 규정	어플리케이션 규정
전자 서명	표준에 포함	어플리케이션 규정
thumbnails	표준에 포함	
파일 확장자	docx, pptx, xlsx	odt, odp, ods, odg, odf
인터넷미디어 타입	application/vnd.openxmlformats-officedocument.wordprocessingml.document application/vnd.openxmlformats-officedocument.presentationml.presentation application/vnd.openxmlformats-officedocument.spreadsheetml.sheet	application/vnd.oasis.opendocument.text application/vnd.oasis.opendocument.presentation application/vnd.oasis.opendocument.spreadsheet application/vnd.oasis.opendocument.graphics
일반적인 편집 소프트웨어	MS Office	OpenOffice.org
라이선스	Reasonable and Non Discriminatory Licensing (RAND) Microsoft Open Specification Promise (OSP)	Sun Microsystems intellectual property covenant IBM Interoperability Specifications Pledge

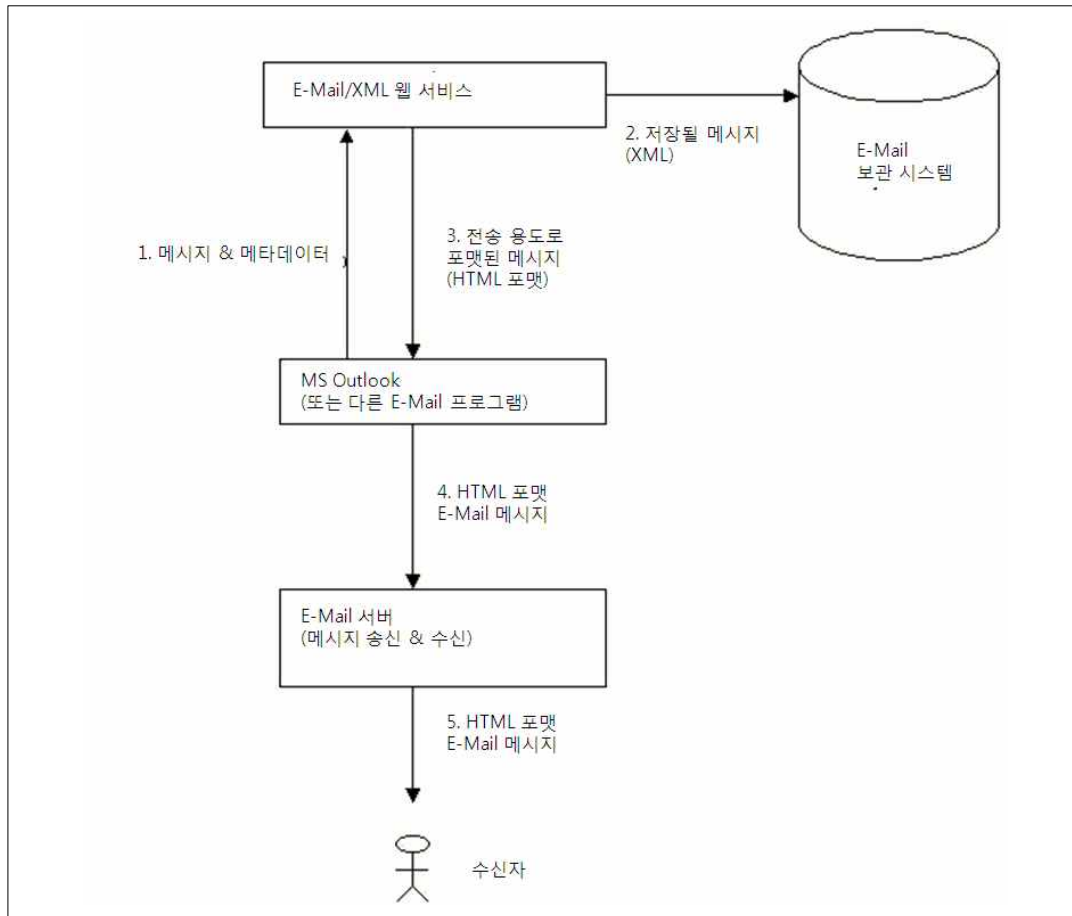
3. XML 적용 사례

앞서 설명한, XML 문서 포맷은 현재, 보존 전략으로서 사용된 예는 보고되지 않고 있다. 이는, 해당 기술이 최신의 기술로서, 아직 분석 및 적용에 시간이 필요하기 때문으로 생각할 수 있다. 그러나, 전자문서와 관련된 산업적 영역에서는 개방형 표준 포맷을 사용하기 위해 다양한 시도를 하고 있음을 볼 수 있다. 마이크로소프트 오피스 프로그램들은 물론, 한글과 컴퓨터의 오피스 프로그램들도 현재 XML을 지원한다.

보존 전략으로서 XML 적용 사례로 네덜란드 DPT가 수행한 E-Mail의 XML 보존을 들 수 있다¹⁴⁾. <그림 1-3>은 이 시스템의 아키텍처를 나타낸다.

14) DPT, 2003, E-mail - XML Demonstrator: A Technical Description
<<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/email-demo-en.pdf>>

<그림 1-3> DPT, E-Mail XML서버의 아키텍처



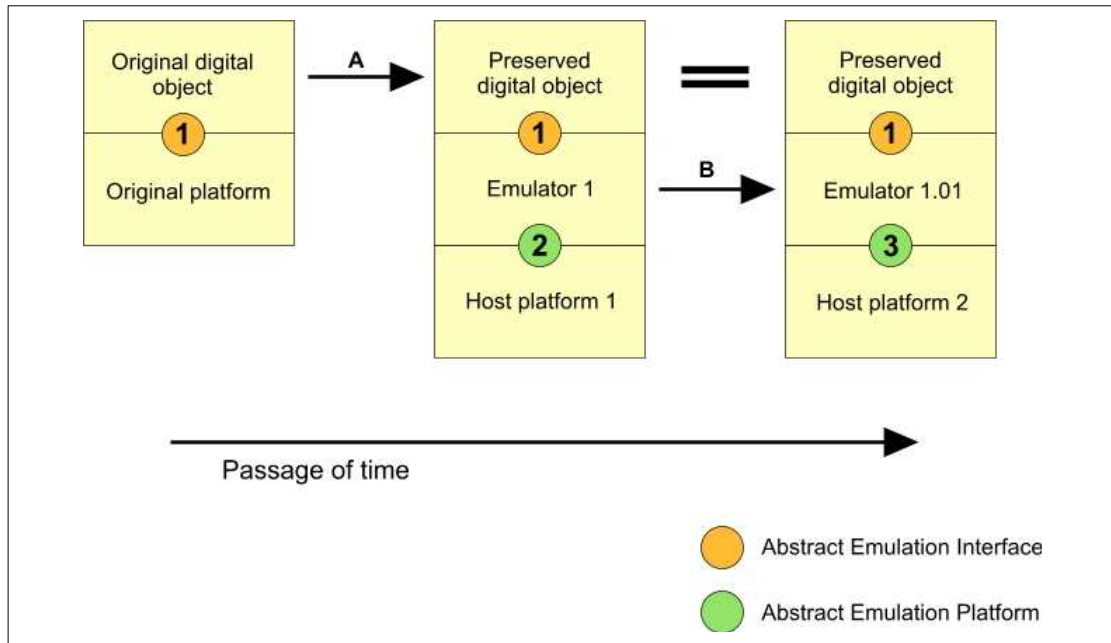
DPT는 E-Mail의 장기보존을 위해, E-Mail과 이에 관련된 메타데이터(송신자, 수신자, header 등)를 XML로 변환하여 저장소에 보관하고, 이용시점에서 XSL(eXtensible Style sheet Language)을 이용해 HTML로 변환, 이용자의 Outlook으로 전송하는 소프트웨어를 개발하였다.

3절 에플레이션

1. 에플레이션의 개념

컴퓨터 과학에서 에플레이션이란, 서로 다른 기기(device)혹은 프로그램을 사용하는 환경에서, 원본을 사용하는 것과 같은 효과를 얻기 위해 사용되는 기술을 말한다. 이 것은 종종, ‘시물레이션(simulation)’과 동의어처럼 사용되기도 한다. 그러나, 시물레이션이 대상이 어떻게 움직일 것인지를 설명하는 것인 반면, 에플레이션은 대상이 실제로 움직이는 것이라는 점이 다르다.¹⁵⁾

<그림 1-4> 에뮬레이션의 개념



즉, 에뮬레이션이란, 하나의 컴퓨터에서 다른 컴퓨터에서 생산된 파일을 이와 동일하게 표현하는 것, 이를 위해, 하나의 컴퓨터에 다른 컴퓨터의 환경을 구현하는 것을 뜻한다.

보존 전략으로서 에뮬레이션은 디지털 자원의 본래 생산시점에서 사용된 하드웨어 혹은 소프트웨어를 재현하는 소프트웨어를 통해 표현하여, 당시에 가능 했던 모든 기능을 사용할 수 있도록 하는 프로세스이다. 에뮬레이션은 특히, 전자자원의 ‘기능(functionality)’이 중요할 경우, 큰 의미를 가진다.¹⁵⁾

에뮬레이션 전략의 수행은, 우선 원본 파일의 비트스트림을 원래 상태대로 보존하는 것으로 시작한다. 그 후, 대상 파일을 표현하는 하드웨어와 소프트웨어의 에뮬레이터(emulator)를 구현하여, 원본의 외형(look and feel)을 그대로 재현할 수 있도록 한다. 이 과정은 <그림 1-4>에 표현되어 있다.

2. 보존전략으로서 에뮬레이션의 특징

보존전략으로서의 에뮬레이션이 가지는 장점은, 원본 디지털 자원을 가장 정확한 형태로 보존할 수 있다는 점이다. 또한, 마이그레이션과 달리 시간이 흐름에 따라, 지속적인

15)DPT, 2003, Emulation : Context and Current Status

<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/White_paper_emulation_UK.pdf>

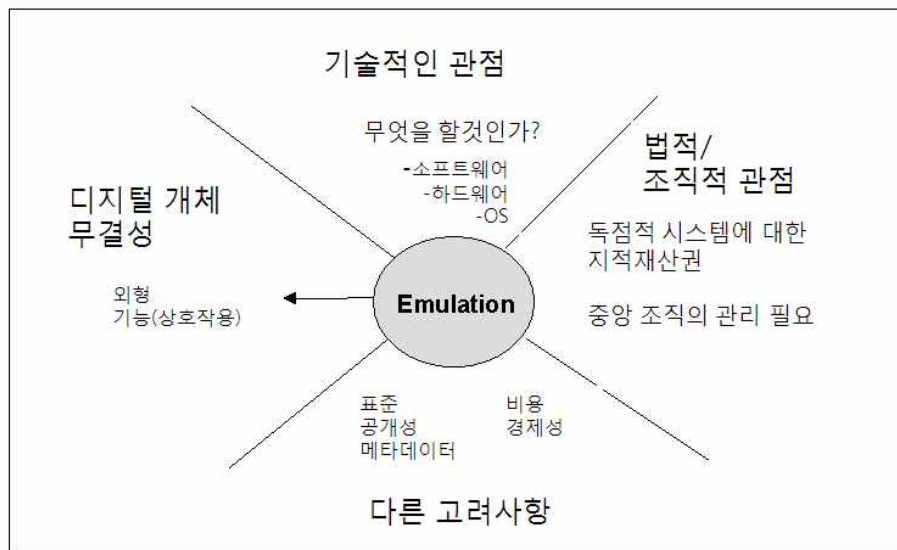
16) PADI (Preserving Access to Digital Information) Thesaurus

<<http://www.nla.gov.au/padi/topics/thesaurus.html>>

변환 프로세스를 거치지 않아도 된다는 점도 장점으로 들 수 있다. 또한, 마이그레이션 전략이 각 디지털 포맷에 따라, 서로 다른 형태를 띄게 되는 것에 비해, 에뮬레이션은 일관된 전략을 수립할 수 있다는 것도 장점으로 지적된다.

그러나, 에뮬레이터를 구현하기 위해서는 고도의 기술과 비용이 소모되며, 마이그레이션에 비해 고려해야할 사항이 많음으로, 실패할 확률이 더 높다는 문제점이 제기된다. 또한, 기기에 대한 소유권 및 지적 재산권이 존재하는 경우, 에뮬레이션의 구현 자체가 지적재산권과 관련된 문제를 야기할 수 있어, 보존전략상 효율성과 안정성을 침해한다. 마지막으로 지적할 사항은, 에뮬레이션의 대상이 되는 시스템과 소프트웨어 환경 역시 다양하게 존재함으로, 지속적인 에뮬레이터의 개발이 필요하다는 사실 역시, 단점으로 지적된다. 지금까지 설명한 에뮬레이터의 장단점을 그림으로 정리하면, <그림 1-5>와 같다.

<그림 1-5> 에뮬레이션과 관련된 고려사항¹⁾



3. 에뮬레이션 적용 사례

에뮬레이션은 사실, 새로운 기술이 아니다. 시중에는 이미 다양한 자원을 대상으로 하는 에뮬레이터들이 존재하고 있으며 이용되고 있다. 다양한 컴퓨터 게임 플랫폼의 에뮬레이션, 고전 컴퓨터 시스템의 에뮬레이션, 이중 OS 사이의 에뮬레이션과 버추얼 PC 등을 그예로 들 수 있다.

그러나, 장기보존의 관점에서의 에뮬레이션은 먼 미래에까지 원본의 접근성을 보장하여야 하며, 진본성을 보장하는 방식으로 표현이 가능해야 한다. CAMiLEON 프로젝트와 KB(네덜란드 왕립 도서관)의 사례는 이와 같은 관점에서 의미를 가진다.

가. CAMiLEON 프로젝트 : BBC Domesday 프로젝트의 보존

BBC Domesday 프로젝트는, 11세기 영국의 인구조사 결과인 Domesday Book의 900주년을 기념하여, BBC(British Broadcasting Corporation)와 Acorn Computers Ltd, Philips, Logica 가 공동으로 조사한 영국 국민들의 삶에 대한 기록이다. 이 것은 1984년에서 1986년 사이에 작성되어, 1986에 발표되었다. Domesday는 영국 각지의 지도, 다양한 지역 사회 사람들의 일상이 지도, 사진, 통계 데이터, 동영상으로 담겨 있다. Domesday 프로젝트의 저장매체는 LV-ROM (LaserVision Read Only Memory)였다. Domesday 프로젝트는 JPEG과 같은 포맷이 나오기 이전에 생성되어, 전자기록물의 유실(digital obsolescence)에 대한 우려의 대표적인 사례로 꼽히고 있다.

2002년 이후, Domesday 프로젝트의 보존 위기가 가시화 되었다. LV-ROM을 사용할 수 있는 PC가 사라져 갔으며, 동영상의 경우, 최신 PC로는 정확한 재생이 불가능하였다. 이에 따라, Domesday 프로젝트의 기록을 시간이 지난 후에도 이용할 수 있도록 CAMiLEON 프로젝트는 에뮬레이션 보존 전략을 수행하였다.¹⁷⁾

CAMiLEON 프로젝트의 연구진은 우선, 과거의 시스템을 구현할 있는 에뮬레이터를 개발하였고, 이를 원본 기록과 함께 에뮬레이션 소프트웨어 내에 캡슐화(encapsulate)하였다. 연구진은 BBC의 'BeebEm' 에뮬레이터를 사용하였다.

그 결과, BBC Domesday의 콘텐츠를 에뮬레이터를 통해 이용이 가능하였다. 하지만, 원본 디스크에 대한 다양한 참여자 사이의 저작권이 복잡하게 얽혀 있어, 서비스는 불가능하였다. BBC Domesday 프로젝트의 사례는, 에뮬레이션의 장점과 단점을 잘 드러내고 있다.

나. KB의 모듈화 에뮬레이션(Modular emulation)

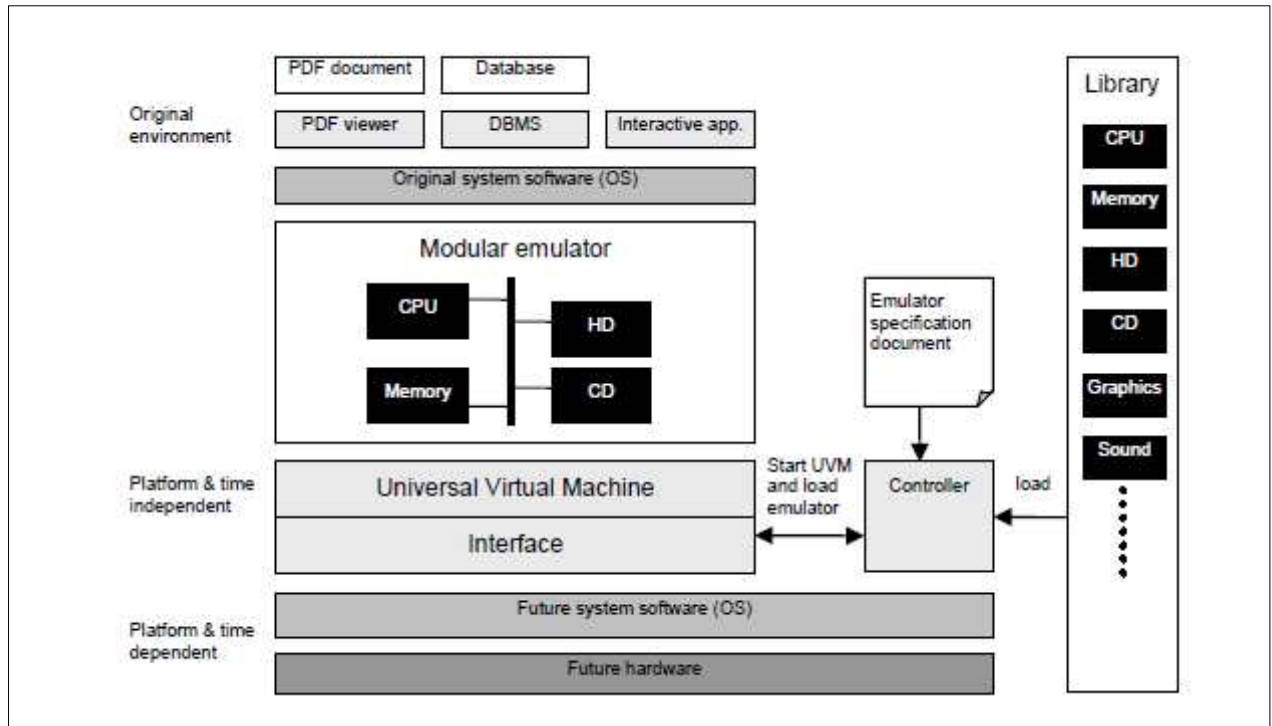
장기보존을 위한 에뮬레이터는 시간의 흐름과 플랫폼에 종속적이지 않아야 한다. KB의 모듈화 에뮬레이션은 이와 같은 요구사항을 만족하기 위해, 범용 가상 머신(Universal Virtual Machine)의 개념을 도입한다.¹⁸⁾

범용 가상 머신이란 에뮬레이터와 구동 플랫폼 사이에, 가상의 시스템을 사용하여, 에뮬레이터를 구동하는 것이다. 이 때, 가상 머신은 가상 머신을 구동하는 플랫폼에 종속적이지 않다. 이를 통해, 다양한 플랫폼에서 같은 형태의 에뮬레이션을 시행할 수 있게 된다.

17) BBC Domesday, CAMiLEON Project
<<http://www.si.umich.edu/CAMiLEON/domesday/domesday.html>>

18) Jeffrey van der Hoeven, Hilde van Wijngaarden, 2005. Modular emulation as a long-term preservation strategy for digital objects, International Web Archiving Workshop 2005. 9. 22-23.

<그림 1-6> 모듈화 된 에뮬레이션



<그림 1-6>은 모듈화 에뮬레이션의 개념도이다. 모듈화 에뮬레이션은 범용 가상 머신을 기반으로, 다양한 하드웨어를 모듈별로 에뮬레이션하여, 가상 머신 상에서 구현하는 방식을 택한다. 이를 통해, 복잡한 하드웨어 구성이 필요한 플랫폼을 실현하고, 플랫폼에 변형이 필요할 때는, 에뮬레이션 되는 모듈만 변경하는 방식으로 일관된 에뮬레이션의 시행이 가능하다.

KB는 모듈화된 에뮬레이션의 개발과 수행을 위한 연구 계획을 수립하고, 모듈화된 에뮬레이션을 실제로 수행할 수 있도록 연구하고 있다.

4절 장기보존전략 비용 모델

디지털 보존과 관련된 주요 질문중의 하나는 "비용이 얼마나 됩니까?"이다. 비용은 디지털 보존에 관련된 사람들에게 중요한 기준중의 하나이며 그렇기 때문에 비용에 대한 범위와 관련된 기본적인 요소들에 대한 깊이 있는 이해가 요구된다. 이러한 비용은 디지털 대상물의 가치와 디지털 보존의 편익 혹은 디지털 대상물의 소실 위험 등을 고려하여 평가하여야 한다.

디지털 자원의 장기보존을 위한 비용에 관한 연구는 개별 연구자에 의해서라기보다는 주로 연구 과제 팀 혹은 기관에 의해서 이루어진 경우가 대부분인 것으로 나타났다. 당해 연구에서는 관련 선행연구에 대한 문헌조사를 통하여 선별된 대표적인 몇몇 사례 모델을

개관하고자 한다. 선별된 사례에는 기록물의 디지털 장기보존에 대한 비용 모델 뿐만 아니라 일반적인 디지털 자원의 장기보존에 대한 비용모델도 포함하고 있다.

1. ERPANET 모델¹⁹⁾

ERPANET (Electronic Resource Preservation and Access Network)에서 개발한 erpaTool에 따르면 디지털 보존의 비용을 산정하기 위해서는 세 가지 고려해야 할 사항으로는 첫째, 디지털 대상물의 생산 수집으로부터 폐기(파기, 이전 혹은 장기 보존)에 이르기까지 관련된 전 활동들을 포함해야 하며, 둘째, 보존은 대상물의 전 보유기간 동안의 활동중인 것을 포함해야 하며, 셋째, 보존은 지속적인 재정 지원이 요구되는 장기적인 모험사업이라는 것이다.

ERPANET에서는 비용요소들은 크게 디지털 대상물, 인력, 법률과 정책, 표준, 방법과 실재, 기술과 시스템, 조직으로 나누고, 이들 각각에 대해 비용에 영향을 미치는 요소들을 마이크로 수준에서 언급하고 있다. 이 연구에서 제시한 비용요소들과 고려사항은 아래 <표 1-3>²⁰⁾과 같다.

<표 1-3> ERPANET에서 제시한 비용 요소와 고려사항

요소	이슈	비용 영향	고려사항
객체	생산에 대한 영향력	더 낮은 영향력은 일반적으로 더 높은 비용을 의미함	조직 외부에서보다 조직 내에서 문서/정보 객체들의 생산에 더 많은 영향을 미칠 가능성이 있음. 정책, 표준 그리고 절차는 그들의 생산과 관리에 도움을 줄 수 있음
	존재	기존 객체들은 수집과 저장을 위한 더 많은 준비작업이 요구됨	만일 객체들이 이미 존재하거나 독점 소프트웨어로 만들어졌다면 그것을 수집하기 위해 준비하는데 더 많은 노력이 요구됨. 또한 맥락 메타데이터의 결여 또는 불충분한 맥락 메타데이터의 위험이 있음
	복잡성	대상물의 복잡성이 증가되면 보다 철저한 보존이 수반됨	복합 객체들은 종종 멀티미디어 혹은 다른 형태의 포맷들로 구성된 합성 객체들임. 그들은 또한 동적인 특성을 가질 수 있음
	보존기간	장기 보유는 보다 긴 유지를 수반함	객체를 요구되는 보존기간 이상 유지하는 것이 중요함
	평가/가치	객체의 본질적인 가치 측정이 필요함. 철저한 평가는 결국 비용경감을 가져	객체의 소실 위험을 식별하는데 관련됨. 모순이 없는 평가정책은 필요함

19) <http://www.erpanet.org/>

20) Erpa guidance: Cost Orientation Tool. ERPANET, 2003.

요소	이슈	비용 영향	고려사항
		음.	객체만을 보존하는데 도움
사람	기술	특수한 기술이 요구됨	디지털 객체의 본질에 대한 이해, 이용가능한 보존 방법들에 대한 이해 등의 기술들이 요구됨. 이용가능한 기술들은 여러 전문분야에 걸쳐있음
	자질	잘 훈련된, 숙련된, 경험 많은 사람은 비용이 많이 들지만 결국 전체 비용을 경감하는 결과를 가져옴	자질은 정의하기 어렵고 매우 많은 요건들과 적용된 업무 맥락(예를 들어 보존)에 의존됨. 설정된 자질 기준을 충족시키기 위해서는 충분한 수의 사람들이 포함되어야 할 것임
	훈련	업무기능에 따라 최신의 적절한 훈련이 필요함. 이것은 자금이 요구됨	빠른 IT 발전과 디지털 보존과 전략의 새로운 통찰은 지속적인 훈련이 요구됨
	경험	경험의 수준이 높을수록, 부주의한 사고들은 감소될 것임	정보 보존 분야는 기술보다 상위에 있음. 그러나, 기술없이 보존도 없음
표준	표준	표준을 사용하면 개발 노력을 덜뿐만 아니라 장기 해결을 용이하게 함. 포맷 표준은 유지관리를 감소시키고 마이그레이션 절차를 쉽게하는데 도움이 됨	표준은 영원히 존재하지 않을 것이고 선택하기 위한 많은 표준들이 있음. 표준은 비용 절감에 도움이 될 것임. 메타데이터 영역은 특별한 주의를 요구함: 일반적으로 메타데이터 수가 적을수록 더 적은 비용이 소요됨. 메타데이터의 표준화는 또한 상호운용성을 가능하게 하고, 효과적인 상호운용성은 쉽고 효과적인 정보 교환, 상호작용, 업무를 가능하게 함
실무	워크플로우	일관성과 지속성이 필요함	절차와 보존 전략은 작업흐름과 밀접한 관계. 훌륭한 시간 계획 작성에 도움을 줌
	운영	비용은 사람, 재료를 포함함. 적절하고 효율적인 소프트웨어 도구는 사람의 관여를 최소화 하여 처리를 가속화시킴	운영과 성능은 정기적으로 평가되고 감사되어야 함
	프로세스	각각의 프로세스들은 비용 효율이 높아야 함	이것은 각각의 프로세스와 고유의 필요한 기술들과 도구들(자원들의 적합성)의 평가에 어떤 단계들이 취해져야 하는지에 대한 분석을 수반함. 자원 소모에 대한 결과를 모니터링하고 평가하는 것이 필요함
시스템, 방법, 기술	보존방법	각 유형들은 다른 비용 수준을 가질 것임. 동시에 다수의 보존 방법들을 사용하는 것이 필요할 지도 모름	방법의 유형은 아이그레이션, XML 또는 다른 표준의 사용, 입수 또는 획득을 위한 절차, UVC 방법 등임.
	방법의 검증	사용하기 전에 방법의 잠재적인 성공의 검증	모니터링과 검증은 지속적으로 수행되어야 함. 그것은 객체의 보존 품질을

요소	이슈	비용 영향	고려사항
			지원할 것임.
	지속성	여러 IT세대를 거쳐서도 생존할 수 있는 방법을 선택하여야 함	비용절감을 위해 필수적임. 방법들은 시간이 흐름에 따라 변화가 있을 것인데 이 경우 실제 적용이 쉬워야 함.
	이식성	선택된 보존 방법은 다른 방법 혹은 새로운 시스템 플랫폼으로 이식이 쉬워야 함	보존 방법 그 자체 뿐 아니라 그들이 지원하는 객체들을 함께 고려해야 함
	구성요소	기술적 인프라스트럭처는 실행되고 유지되어야 함	어떤 구성요소가 필요한지는 목표의 수준, 객체의 복잡성, 선택된 보존방법, 위상에 의존하고 있음. 이따금 독점 시스템들은 싸게 보일지도 모르지만 오랜 기간 운영하게 되면 그들은 보다 비싸지게 될 것임
	유지	운영상의 측면과 최신성의 측면에서 시스템의 유지	-
	운영	사람, 재료를 포함한 비용들. 적절하고 효율적인 소프트웨어 도구는 사람의 관여를 최소화 하여 처리를 가속화시킴	-
	융통성	오랜 기간, 혹은 IT변화에 대처하기 위하여 융통성 있는 적응이 요구됨	새로운 요소의 추가, 갱신, 이동이 저비용으로 이루어질 수 있어야 함
	시설	장소, 보안, 안전, 백업	-
	보존의 분류	비트스트림의 보존과 기능의 보존 사이의 구별이 필요함	비트스트림의 보존은 기능의 보존보다 상대적으로 저비용이며 간단함.
	모듈방식	인프라스트럭처 모듈방식의 유지와 교체는 비용효율적일 수 있음	-
법률/정책	법률	조직과 그 조직의 활동은 특정 법률에 따라 설명될 수 있어야 함	설명책임성, 접근, 저작권, 프라이버시 이슈 등을 포함
	정책	조직 각각은 지속적인 비용 효율적인 관리를 가능하도록 하기 위해 보존 정책을 가지고 있어야 함. 정책과 활동들에 대한 정기 모니터링이 필요함	정책은 업무 활동들에 내포되고 연결되어야 함(위험 관리를 포함하여). 정책은 새로운 통찰과 기술 혁명 그리고 보존 방법들과 더불어 갱신되어야 함. 내부 또는 외부 감사는 정책이 적절하게 유지되도록 지원할 것임
조직	관계형성	타 조직과의 협력은 시너지와 규모의 경제를 통해 비용을 절감할 것임	협력 연구, 합작 기업, 지식 뿐 아니라 기술도구 및 인프라스트럭처의 공유에 의해 협력은 성취될 수 있음.보다 상업적인 접근방법은 아웃소싱 활동임
	능력축적	조직들은 사람, 정책, 실무, 방법과 기술에 관해서 적절한 보존을 위한 능력을 축적하여야 함.	능력축적은 조직의 필요, 조직의 역할과 설명책임성에 의존됨

요소	이슈	비용 영향	고려사항
	설명책임성	설명책임성의 명확한 식별과 할당은 설명책임성은 관련된 모든 사람들의 오해, 절차와 시스템의 실패를 방지하는 것을 도와줌	설명책임성은 관련된 모든 사람들의 식별을 포함함

2. 네덜란드 국립기록관 (Nationaal Archief of the Netherlands) 모델

네덜란드의 Ministry of the Interior and Kingdom Relations(BZK)와 네덜란드 국립기록관이 공동 구축한 Digital Preservation Testbed 프로젝트는 정부의 전자정보를 보존하고 미래에도 접근가능 하도록 하기 위한 다양한 방법의 실질적인 응용을 위한 실험을 모색하였다. 프로젝트를 통해 생산된 From Digital Volatility to Digital Permanence라는 보존 시리즈는 프로젝트 수행의 경험과 사례를 기술하고 있다. 여기서는 2005년 네덜란드 국립기록관의 Digital Preservation Testbed Project의 결과물 Costs of Digital Preservation(version1.0)을 중심으로 살펴보고자 한다.

가. 비용 요소

Digital Preservation Testbed 프로젝트에서 도출한 Costs of Digital Preservation은 <표 1-4>와 같은 비용 지표를 제시하고 있다.

<표 1-4> Digital Preservation Testbed의 비용 지표

항목	내용
디지털 기록의 장기보존을 위한 디지털 기록시스템 (디지털 보관소)와 디지털 기록의 장기 보존을 위한 직무 비용	- 물리적 공간 (에어컨이 있는 서버 룸, 충분한 사무실 공간, 회의실, 화장실과 부엌, 보안) - 디지털 기록보관을 위한 하드웨어(디지털기록 저장을 위한 서버, 디스크나 테이프와 같은 저장매체, 백업 장비, 통신 장비 등) - 디지털 기록보관을 위한 소프트웨어 (운영체제, 보안, 기록관리를 위한 특별한 소프트웨어, 낡은 응용소프트웨어, 최신 응용소프트웨어, 출력 프로그램, 통신 소프트웨어, 데이터베이스) - 보존시스템을 위한 하드웨어 (소프트웨어 보존을 위한 서버, 소프트웨어 테스트를 위한 서버, 보존 실행 전에 기록을 저장하기 위한 서버, 보존 실행 이후에 기록을 저장하기 위한 서버, 프로그래밍 작업을 위한 워크스테이션, 디스크와 테이프, 백업 장비, 통신 네트워크, 디스크와 테이프의 읽기 위한 장비), - 보존 시스템을 위한 소프트웨어 (운영체제, 프로그램 환경, 보안, 낡은 응용소프트웨어, 최신 응용소프트웨어, 문서 보존을 위한 소프트웨어, 테스트와 평가 소프트웨어, 통신 소프트웨어, 데이터베이스)
인건비	디지털 기록보관소 직원의 임무 (요구사항 수집, 자금 획득 및 지원, 디지털 보관소의 설계 및 구축, 디지털 아카이브 축적, 처리 관리, 디지털 보관소 운영, 보안 관리, 시스템과 문서의 품질 통제, 표준 운영절차 마련, 이용자 매뉴얼)

항목	내용
	- 열 작성) - 보존시스템 직원의 임무 (요구사항 수집, 자금 획득 및 지원, 보존 시스템의 설계 및 구축, 보존 시스템 구축, 처리 관리, 보존 시스템 관리, 보안관리, 문서 품질 통제, 표준 운영절차 마련, 이용자 매뉴얼 작성) - 공공서비스 직원의 임무 (접근 관리, 훈련)
기록보존을 위한 소프트웨어의 개발(또는 구입) 및 방법 비용	진본성 요구 결정, 진본성 요구 분석, 보존 방법 설계, 보존 방법개발, 보존 소프트웨어 (구문분석기 등), 뷰어 소프트웨어, 보존 전략 테스트, 보존 방법 문서화 등
보존 실행 비용	- 마이그레이션 (포맷변환), 요구에 따라 마이그레이션 실행, 기록의 접근 보유를 위해 에뮬레이션 사용 등이 관련되어 있음 - 비용에 관련된 요소로는 어떤 디지털 기록이 보존되어야 하는가에 대한 결정, 보존시스템과 관리시스템간의 인터페이스 구축, 전자 기록 관리 시스템, DMS, RMA시스템과의 협력, 디지털 기록 접수, 보존 전략과 방법 선택, 포맷 변환을 위한 레코드 준비, 메타데이터 제공, 레코드 수정 및 보정, 선택한 방법에 따라 레코드 포맷 변환, 포맷 변환 평가 (접근가능한지, 진본은 유지되었는지), 디지털 보관소에 저장 등 - 하드웨어 에뮬레이션 전략을 사용하게 되면 포맷변환 비용이 소요되지 않을 것이고, 레코드 준비를 위한 비용도 재발되지 않을 것이라고 제언
전체 비용에 영향을 미치는 기타 요소들	- 이용자서비스 (이용자 수, 훈련 및 지원 도구, 유지) - 보존 실행 주기 (더 잦은 실행은 고비용이 요구되며, 또한 진본성과 무결성에 영향을 미치는 위험이 증가할 것임) - 기술 감시 (하드웨어, 소프트웨어, 시스템의 감시, 노후화 대책 등) - 보조 저장소 (원본 파일도 저장하는 것이 안전하며, 텍스트 문서 레코드들은 PDF와 XML로 저장하는 것을 권고하고 있음) - 전자기록물을 위한 관리시스템과의 연결 (미래 어느 시점에선가 필요할 것이라고 예상함) - 기록물의 양 (기록물의 양에 따라 비용도 비례적으로 증가할 것임, 특히 다양한 포맷 (예, 원본 파일포맷과 더불어 두 가지 마이그레이션 포맷)으로 저장하는 경우는 저장 공간 요구가 급증하게 될 것임.)) - 진본성과 신뢰성을 위한 요구수준 (단순한 텍스트 형태의 문서일지라도 진본성의 요구 수준이, 예를 들어, 각 글자가 원본과 똑 같은 색깔, 똑 같은 위치까지 요구하게 되면 비용이 증가하게 될 것임. 이런 이유로 진본성 요구 수준을 될 수 있으면 포괄적이고 현실적면을 고려해서 정하는 것이 중요하다고 조언하고 있음) - 시스템 자체 보존 (기록관시스템의 보존, 보존시스템의 보존) 등

나. 비용 산출 공식

이 공식은 엑셀로 개발하였는데, 3개의 스프레드시트로 구성되어 있다. 이것은 비용모델은 최종 버전이 아니며, 단지 앞으로 개정을 위한 시작점을 제공하기 위한 목적으로 개발한 것이다. 또한 산출된 비용은 테스트베드의 실험과 경험에 기반을 둔 추정치이다. 또한 이 비용 모델을 설계하는데 아래와 같은 가정을 하고 있다.

- 직원의 유형을 6가지로 한다. 각 직원에 대한 급여의 차이가 전체 비용에 큰 영향을 미치기 때문이다. 또한 봉급은 시급으로 계산되고, 만일 1일 8시간, 주 40시간으로 1년 1620시간을 모두 일했다고 했을 때 시급에서 1.25배(네덜란드의 경우) 금액을 받게 된다.
- 4가지 유형의 설비공간을 두고, 각 공간에 필요한 기기의 수에 따라 비용을 산출한다. 기기비용(자산비용)은 일회적이기 때문에 그 이후의 비용에 대해서는 총비용의 33%가 절감된다.
- 지정된 시설은 레코드의 100 Terabyte 전체를 관리할 수 있고, 즉시 1000 Terabyte 또는 그 이상으로 확장될 수도 있어야 한다.
- 디지털 기록 시스템에서 취득된 레코드라 할지라도 10%는 수정이 필요할 것이다. (예를 들면, 패스워드, 폰트의 처리, 날짜, 매크로 등의 수정 요구 가능)
- XML에서 생산된 레코드는 수정 또는 메타데이터의 추가가 필요 없다고 가정한다.
- XML로의 전환은 매 10년마다, 마이그레이션은 매 3년마다 반복될 것이라고 가정한다. UVC와 뷰어를 위한 반복요인은 없으나, 20년으로 가정한다.
- UVC는 매년 데이터 포맷 해독기 프로그램이 개발이 요구될 것이라고 가정한다.

(1) 비용 근거

▫ 직원 비용 (시간당)

직원은 행정보조자, 기록관리자, 감독관, 데이터 입력자, 소프트웨어 엔지니어, 선임 IT 직원 등 여섯 등급이 있고, 각 직원의 봉급 종류는 유로(Euro)로 시급과 일당 (시급 × 1.25), 연봉 (시급 × 시간)과 년 비용 (일당 × 시간)이 있다.

▫ 공간 비용

개발 및 시험 공간, 보존실, 사무실, 간접비 20%에 대한 유로로 년당 m²의 비용은 모두 450이고, 직원 유형별, 작업공간의 수와 유로로 년당 공간의 비용을 계산한다. (단, 개인당 공간은 20m², 중요한 기계 당 공간은 5m²로 가정한다.)

▫ 하드웨어 소프트웨어 비용

하드웨어와 소프트웨어는 첫 해에 모든 설치비용이 들고 그 후에는 33% 하락된다. 데스크탑, 노트북, 사무용 소프트웨어, 프로그램 소프트웨어, 서버, 데이터베이스 소프트웨어, 보존 시스템, 저장 시스템, 저장매체, 주크박스에 대한 공간, 보존 사무실, 행정 각 아이템 수와 그에 따른 유로(Euro) 비용을 계산한다.

(2) 시간 계산

보존 방법을 개발하기 위하여, 그리고 실제로 보존 실행을 위하여(직원 유형별, 년과 N년 기간 당) 몇몇 레코드 보존 실행을 수행하는데 걸리는 시간을 기술한다. 그리고 이 모델은 숙련된 팀과 디지털 보존에 친숙하고, IT와 레코드키퍼 문제를 잘 알고 있는 열성적인 기록관과 보존시설이라고 가정한다. 보존행위, 일괄처리(batch) 등에 대해 요구되는 시간에 중점을 둔다.

(3) 비용 계산

마이그레이션(PDF로의 변환도 포함)과 XML의 이용, 에뮬레이션(UVC, Universal Virtual Computer, IBM에서 개발)이 비용기간(레코드의 batch당, 년과 N년 기간 당)에 나타나는 보존 방법 테스트베드 프로젝트 간의 비교에 초점을 둔다.

▫ 레코드 보존 실행

이 테스트베드는 레코드의 네 가지 유형 (전자메일, 텍스트, 스프레드시트, 데이터베이스)으로 나누어 연구하였다. 디지털 보존은 레코드의 생산시점에서부터 시작된다는 것은 자명할 것이다. 비용이 XML 전자메일1당 유로 0.06 이상하지 않는 반면, 산출비용은 전자메일 당 유로 2.30에 메타데이터 총량의 취득과 입력이 부담된다. 이 숫자가 비용 모델에서 사용된 가변성에 근거한다는 점을 주목하는 것은 중요하다. 주요 차이점은 XML 전자메일은 이미 적합한 메타데이터와 구조를 갖추었다는 것이다.

▫ 보존 방법 개발

보존 방법의 개발은 디지털 보존비용의 대부분을 차지한다. 예로는 한 해에 단지 100개의 스프레드시트 보존에 근거한다(20개의 스프레드시트의 5 batches). 여기에 마이그레이션 전략 개발(스프레드시트 당 33.50유로)과 특히 데이터 포맷 해독기 프로그램(항목 당 769.10 유로)을 개발하는 UVC 방법의 개발에 커다란 차이가 있다. XML 방법의 적용은 비용이 없다.

▫ 디지털 보존 실행

실제 보존 방법 개발 비용과 비교했을 때 보존 활동의 운영비용은 미미하다. 평가 목적을 위한 자동화된 도구는 특정한 보존 행위의 결과를 평가하기 위하여 필요하다. 한편 수동적 평가는 보존된 레코드 셋의 무작위 시험을 위해 이용될 수 있다.

- 일괄처리 당 직원 유형별 총 비용, 연간 단위 직원 유형별 총비용, n년 간 직원유형별 총비용

일괄처리 당 직원 유형별 총비용과, 년당 직원 유형별 총비용과, n년 동안 직원유형별 총비용이 계산된다.

- 비용산출 공식²¹⁾

- 마이그레이션: $K(t, a) = S + p(t) + h(t) + m(t/5, a)$

- 에뮬레이션: $K(t, a) = S + p(t) + h(t) + E + e(t/5)$

다. 보존방법에 따른 비용 비교

마이그레이션 (텍스트문서를 PDF로의 변환도 포함)은 개발비용이 지속적으로 되풀이 될 것이란 사실에 기인하여 장기적으로는 매우 비싼 보존 방법이다. 이 모델은 마이그레이션은 3년마다 반복되어야 한다고 가정한다.

UVC 접근은 이 비용모델에 따르면 가장 비싼 보존 방법인데, 이것은 거대한 연구개발 비용에 원인이 있다. 이 프로젝트에서 엑셀에서 ‘데이터 포맷 해독기 프로그램’을 개발하는 것은 매우 어려운 일이라는 것을 경험했다. 게다가 각 새 파일 포맷을 위해서는 새로운 ‘데이터 포맷 해독기 프로그램’이 개발되어야 한다. 다른 한편으로는, 이 비용은 지속적으로 반복되지는 않을 것이다. 즉, 한번 개발된 ‘데이터 포맷 해독기 프로그램’은 계속 지속될 것이다. 아마 국제적 협력을 통하여 UVC 방법의 애플리케이션을 강화할 수 있을 것이다.

비용모델을 기준으로 보면 XML은 특히 레코드가 XML로 직접 생산되었을 때, 마이그레이션(PDF의 이전 버전 호환 및 변환) 및 UVC 방법의 적용과 비교해서 비용측면에서 가장 효과적인 보존 방법인 것을 증명하였다. 하지만 XML이 영원히 지속되지는 않을 것이기 때문에, XML로 생산된 (또는 XML로 개조된) 레코드조차 어떤 류의 보존 행위가 필요하다. 비용모델에서는 XML 레코드 (XML로 직접 생산되었거나 XML로 변환된)는 10년 후에 새로운 기준에 의해 변환될 수 있을 것이라고 가정하고 있다.

<표 1-5> 네 가지 보존 방법의 비교 (화폐단위: 유로)

보존 방법	기존 스프레드시트 (마이그레이션)	기존 스프레드시트를 XML로	기존 스프레드시트 (에뮬레이션, UVC)	새로운 스프레드시트를 XML로
일괄처리 규모	200	200	200	200

21) $K(t, a)$ = a대상물을 t년 동안 보존하기 위한 총비용, S = 초기 저장 비용, p = 인건비, h = 저장, m = 마이그레이션, E = 초기 에뮬레이션 도구 설치비, e(t) = 장기간의 에뮬레이션

년당 일괄처리 수	5	5	5	5
년당 총비용	9,213.79	11,252.29	107,061.79	855.00
20년 후 총비용	709,461.96	337,568.75	2,141,253.83	129,622.92

<표 1-5>는 같은 처리 규모에서 마이그레이션, XML, UVC 운영에 있어 비용차이가 있음을 보여준다. 표를 살펴보았을 때 UVC의 비용이 현저하게 높고, XML의 비용은 상대적으로 낮은 것을 알 수 있다.

<표 1-6> 세 가지 보존 방법과 일괄처리 사이즈의 영향 (화폐단위: 유로)

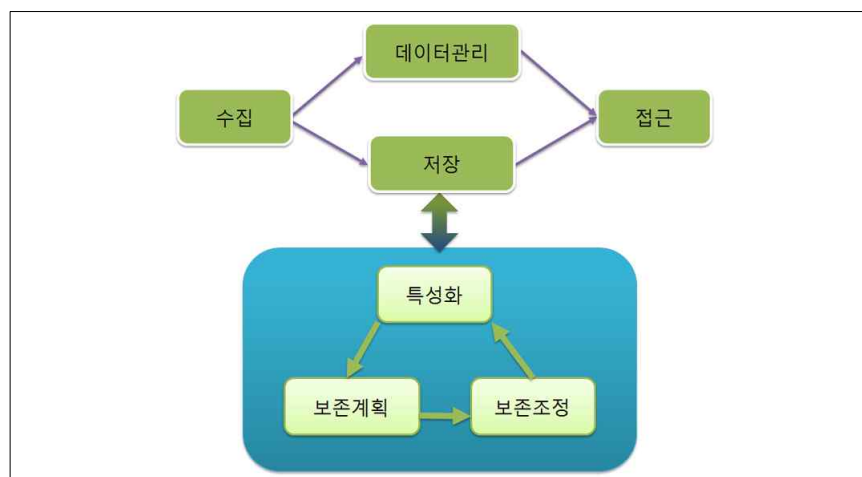
보존 방법	기존 텍스트 (마이그레이션)		기존 텍스트를 XML로		새로운 텍스트를 XML로	
일괄처리 규모	200	2000	200	2000	200	2000
년당 일괄처리 수	10	10	10	10	10	10
년당 총비용	20,479.33	112,579.33	24,574.33	116,674.33	1,710.33	1,710.00
20년 후 총비용	1,578,294.67	8,669,994.67	737,230.00	3,500,230.00	279,943.33	1,200,943.33

<표 1-6>에서 일괄처리 규모에 따라 보존방법의 비용이 증가되는 것을 알 수 있다.

3. TNA(The National Archives)²²⁾ 모델

TNA에서 제시한 보존에 관련된 요소들의 범위는 OAIS에 기반을 두고 다음 <그림 1-7>과 같이 정하고 있다.

<그림 1-7> TNA에서의 보존비용 관련요소



22) www.nationalarchives.gov.uk

TNA에서 제시한 비용모델은 크게 현재 비용과 미래비용으로 나누고 있다. 현재 비용으로는 수집, 데이터관리와 저장관리, 접근 요소를, 그리고 미래비용으로는 보존 요소를 들고 있다.

가. 수집 비용 산출

구체적으로 수집에 관련된 비용 요소들로는 선정과 평가, 이전, 접근전 처리, 목록, 로딩 등을 들었다.

$$\text{산출식}^{23)} = ((60\% \times \text{DPD}) + (5\% \times \text{RMD}) + (60\% \times \text{NDAD})) \div \text{FN}$$

TNA의 경우 수집비용은 파일당 18.76파운드로 산정되었다.

나. 데이터관리 및 저장관리 비용 산출

구체적인 비용 요소들로는 시스템 관리, 하드웨어, 소프트웨어, 미디어 관리 등을 들고 있다.

$$\text{산출식}^{24)} = ((\text{DA} \times 0.06) + (\text{DA} \div 5) + (30\% \times \text{DPD}) + (30\% \times \text{NDAD})) \div \text{VOL}$$

TNA의 경우 데이터관리 및 저장관리 비용은 MB당 3.34파운드, 최소비용은 MB당 0.06파운드로 산정되었다.

다. 접근 비용 산출

구체적인 비용 요소들로는 온라인 배포와 제거 가능한 매체를 들고 있다. 온라인 배포의 경우 항목당 13펜스, 제거 가능한 매체는 실비로 산정하였다.

라. 미래 비용

미래예측 비용으로는 보존 전략, 연구개발비, 미래 운영 등을 들었다.

마. 특성화(정리, 색인)

자동화된 도구의 사용으로 대상물의 속성을 표현하고 속성의 고유성을 파악할 수 있다. 주된 비용은 연구개발비이며 운영경비는 미미하다고 제시하고 있다.

23) DPD = 연간 DPD운영비, RMD = 연간 RMD 운영비, NDAD = 연간 NDAD계약비, FN = 연간 수집된 파일 수

24) DA = 시스템의 자본금, DPD = 연간 DPD운영비, NDAD = 연간 NDAP 계약비, VOL = 저장된 전체 레코드 양

바. 보존 계획

PRONOM 콘텐츠 개발, 위험평가, 마이그레이션 경로 테스트, 현재는 주로 수작업에 의존하고 있지만 미래에는 자동화 할 계획이다. 여러 요소들 중 가장 비용이 많이 드는 요소중의 하나라고 간주하고 있다. 그렇지만 협력에 의하여 비용을 크게 절감할 수 있을 것이라고 제안하고 있다.

사. 보존 보정

자동화된 도구의 사용으로 보존 계획(예, 마이그레이션/에물레이션)에서 보정을 수행하고, 주된 경비는 개발(보존계획)과 확인(특성화)부분이다. 하지만 구현비용은 낮은 것으로 간주하고 있다.

아. 보존 전략 비교

하지만 마이그레이션과 에물레이션 전략에 따른 비용차이는 불분명하다고 결론을 내리고 있다. <표 1-7>은 이러한 결론의 근거를 보여준다.

<표 1-7> TNA 보존 전략 비용 비교

비교 항목	마이그레이션	에물레이션
주요 비용	마이그레이션 경로 개발비	에물레이터 개발비
개발비용	되풀이 됨	되풀이 됨
운영비	자동화로 인해 미미	거의 없음
대상 규모	대상물의 양은 비용에 비교적 적은 영향	대량의 대상물일수록 경제

4. CEDARS 모델²⁵⁾

Cedars에서 제시한 보존비용 모델은 크게 일회성 비용, 장기적 비용, 가변적 비용으로 범주화하고 있다.

가. 일회성 비용

25) <http://www.leeds.ac.uk/cedars/guideto/collmanagement/>

Cedars에서 제시한 보존비용 모델은 크게 일회성 비용, 장기적 비용, 가변적 비용으로 범주화하고 있다.

가. 일회성 비용

일회성 비용으로는 디지털 보존 기반구조 확립에 드는 비용(장기적 비용이기도 함), 대상물 선정비용, 지적재산권 관련 비용(장기적 비용이기도 함), 지속적인 접근에 필요한 비용(기술발전 혹은 기관의 정책에 따라 바뀔, 장기적 비용이기도 함), 무결성 확인에 소요되는 비용, 메타데이터 보존에 소요되는 비용 등을 들고 있다.

나. 장기적 비용

지속적으로 장기간에 걸쳐 요구되는 보존 비용으로 파일저장, 보존관리, 수집정책의 평가와 수정 등을 들고 있다.

다. 가변적 비용

시간이 지남에 따라 가변적인 비용으로는 지속적인 접근을 위한 기술적인 전략, 메타데이터 생성 비용, 그리고 법적 권리 협상 비용 등을 들고 있다.

5. arXiv²⁶⁾ 모델

장기보존의 비용모델을 위해서 먼저 비용요소들의 범주화가 요구되는데, arXiv에서는 크게 준비 비용, 처리 비용, 예비비용으로 나누고 있다. 두 번째 단계에서는 각 요소들에 대한 비용을 조사 확인하는 단계인데, arXiv에서는 주요 비용(기술 및 절차 기반마련에 드는 비용)과 직접 운영비용(OAIS 기능에 의거하여 수집, 데이터관리, 저장, 보존계획) 그리고 간접비용(시설, 각종 서비스 비용)으로 나누고 있다. 마지막 세 번째 단계에서는 비용을 산출하는 단계인데, 시스템, 직원, 보조원, 서비스와 수수료, 재료비, 기타로 나누고 있다.

가. 장비 비용

장비 비용은 하드웨어 구입비, 소프트웨어 구입비와 더불어 상환비, 유지비, 라이선스

26) <http://arxiv.org/>

비용 그리고 개발비(이는 구입비의 20%-50% 정도 추정)가 필요한 것으로 나타났다. 또한 arXiv 서버 비용은 장비 비용이 10,600달러를 5년간 상환 지출하고 연간 수수료가 2,403달러를 지출하여 결국 연간 장비비용은 4,523달러 ($\$10,600 \div 5 + \$2,403$)를 지출한 것으로 나타났다.

나. 인건비

한편 인건비중 프로그래머(F등급) 1명의 시간당 비용은 62달러, 연간 보조원에게 지출된 경비는 최소 4,000달러로 나타났다.

다. 서비스 비용

서비스 비용은 백업, 서버 공간, 기술 모니터링, 소프트웨어 등록비, 협회비, 로열티, 가입비 등이 포함되었다.

라. 예비비

예비비는 예상하지 못한 재난 등에 대비한 경비로 복잡성, 노력, 경험 등에 따라 변수가 있다.

마. 연간 비용 산출

arXiv의 연간 유지 및 보존비용은 305,333달러인 것으로 나타났다. 여기에는 장비비가 4,523달러, 직원경비가 150,480달러, 간접관리비가 15,000달러 (8%), 보조직원경비가 8,640달러, 예비비가 17,568달러, 간접비가 112,084달러로 계산이 되었다. arXiv의 경우 다른 기관과 마찬가지로 인건비가 가장 많은 부분을 차지하는 것으로 나타났다.

바. 저장 공간당 비용 산출

이를 저장 공간 단위로 계산한 결과 1GB당 1,873달러가 소요된 것으로 나타났다.

사. 처리 건당 비용 산출

코넬대학교 도서관에서 자동화된 방법으로 실시한 새로운 서버로의 이전 작업은 1991년에 시작하여 2001년 가을까지 물리학, 컴퓨터, 수학 관련 326,000건의 연구논문을 처리

하였는데, 소요된 경비는 한 건당 1-5 달러 정도인 것으로 보고되었다.

6. The LIFE²⁷⁾ 모델

The LIFE (Lifecycle Information for E-Literature)에서는 디지털 보존의 생애주기를 수집(Aq), 획득(I), 메타데이터(M), 접근(Ac) 저장(S) 그리고 보존(P)과 같이 여섯 가지 주요 요소들로 나누었다.

즉, L_T (T기간 동안 전 생애 비용) = $Aq + I_T + M_T + Ac_T + S_T + P_T$ 으로 산출할 수 있다. 이 모델을 보다 세부적인 요소들로 나누면 <표 1-8>과 같다.

<표 1-8> LIFE 모델에서의 비용요소

수집	획득	메타데이터	접근	저장	보존
선정	품질보증	특성화	참조링크	비트스트림 저장 비용	기술감시
지적재산권	계약금	기술(記述)	이용자지원		보존도구 비용
라이선싱	보유갱신	관리	접근 메커니즘		보존메타데이터
주문/송장					보존활동
획득					품질보증
체크인					

7. 디지털 아카이빙 경제성 분석 모델

정혜경(2005)의 디지털 아카이빙의 경제성 분석에 대한 연구에서는 고정비용과 가변비용으로 범주화하여 각 항목에 대한 구체적인 평가지표를 제시하고 있다. 즉, 고정비용은 장기보존을 위하여 기본적으로 투입되어야 하는 불변의 비용으로 공간비, 인건비, 시설장비비, 전산장비비, 전산장비의 유지보수비를 들었고, 가변비용은 기록물 한 단위를 생산하는데 투입되는 인건비, 소모품비 및 관리비 및 예비비를 포함시켰다. 이를 정리하면 <표 1-9>와 같다.

<표 1-9> 비용항목 및 산출공식

범주	평가항목	산출공식
	공간비	공간비 = 자료관 면적 x 면적당 신축 공사비

27) <http://www.life.ac.uk/>

고정 비용	인건비	인건비 = 고정비용 + 가변 비용 * 고정비용 = 연평균 급여 x 정규직원 수 * 가변비용 = 일용직에게 실제로 지불된 급여
	시설장비비	시설장비비 = [(자료관 면적 / 99m ²) x 500만원] + (자료관 면적 / 99m ²) x 100만원]
	전산장비비	전산장비비 = 기관에서 실제로 전산장비의 구입에 실제 투여한 비용
	유지보수비	유지보수비 = 전산장비비 x 8% (단, 구입 후 1년간은 무상 하자 보수 기간임)
가변 비용	인건비	인건비 = 시간당 봉급자 (스캐닝 담당 등)의 급여
	소모품비	소모품비 = 디지털화 작업에 소요된 종이, 복사비, 소모품비 등
	관리비 및 예비비	관리비 = 공간비 면적 x 면적당 신축 공사비 x 0.5% 예비비 = 공간비 면적 x 면적당 신축 공사비 x 10%

8. 기타 단순 비용산출 모델

Hendley(1998)은 일회성 비용의 절반과 지속적 비용 전체를 보존비용의 산출 공식으로 제시하였다. 즉,

$$\text{디지털 보존비용} = (\text{일회성 비용} \div 2) + \text{지속적 보존비용}.$$

Sanett(2002)는 자본비용, 직접운영비용, 간접운영비용으로 범주화하고 다음과 같은 산출 공식을 제시하였다. 즉,

$$\text{디지털 보존비용} \text{ 단가} = \text{자본비용} = [\text{직접운영비용} + \text{간접운영비용}] \div \text{보존대상자료의 건수}.$$

한편 정혜경은 직접보존비용의 전체와 필요시 일시적으로 적용되는 간접보존비용의 절반을 산출 공식에 적용하여 다음과 같은 식을 제시하였다. 즉,

$$\text{디지털 보존비용} = (\text{직접보존비용} + (\text{간접보존비용} \div 2)).$$

9. 요약 및 향후 연구 과제

앞선 선행연구의 결과에서 밝혀진 것처럼 보존비용은 어떤 보존방법이나 전략을 선택하느냐에 따라 많은 영향을 받게 되는데, 마이그레이션의 경우는 마이그레이션 경로 개발비가, 에물레이션의 경우는 에물레이터 개발비가 주된 비용을 알 수 있다. 또한 두 방법 모두 개발비용은 지속적으로 요구되며, 실행이나 운영비용은 거의 요구되지 않는다는 것과 마이그레이션의 경우 보존 기록물의 양은 비용에 크게 영향을 주지 않은 것으로 나타났다. 에물레이션의 경우는 보존 기록물의 양이 많을수록 경제적이라는 결과를 의미 있게 해석할 필요가 있을 것이다. 네덜란드 국립기록관의 연구에서는 XML방법이 마이

그레이션이나 에플레이션 어느 방법보다도 경제적이라는 결과를 보였다.

장기보존을 위한 비용모델을 개발하기 위한 절차는 크게 세 단계로 나눌 수 있다. 첫 번째 단계에서는 보존의 범위를 파악하고, 두 번째 단계에서는 각 요소들의 비용을 조사하고, 세 번째 단계에서는 비용산출 공식을 개발하는 것이다. 선행연구의 결과를 살펴보면, 보존의 범위가 각 모델마다 상이할 뿐만 아니라 보존에 영향을 미치는 요소에 대한 근거를 충분히 제시하지 못하고 있다. 이 때문에 비용산출 공식에 대한 타당성이 떨어지는 것이 사실이다. 이는 디지털 기록물의 장기보존을 위한 비용모델에 대한 연구가 아직 초기단계에 머물러 있다는 측면과 함께 이 분야에 대한 연구가 단순하지만은 않다는 사실을 지적할 수 있겠다.

따라서 향후 연구과제로는 디지털 기록물의 장기보존을 위한 비용에 영향을 미치는 요소들을 망라적으로 분석하고 이를 범주화하는 작업이 선행되어야 할 것이다. 이를 바탕으로 비용 산출 모델을 개발하고, 개발된 모델을 국내 유관기관 테스트베드를 통하여 실제 데이터를 투입하여 검증을 받아보는 실증적인 연구가 필요하다고 할 수 있다.



본 장에서는 전자기록물의 장기보존을 위한 요구사항을 충족할 수 있는 마이그레이션의 절차 모형을 조사 분석하여 국내 적용 가능한 마이그레이션 절차 모형의 기반을 제시하고자 한다. 이를 위해 우선, 이미 구축되어 있는 해외의 마이그레이션 절차 모델에 대한 사례를 조사하였다. 확인된 마이그레이션 절차모델들은 각 모델별로 어떤 절차를 가지고 있으며, 각 과정에서 중요하게 고려하고 있는 사항이 무엇인지 분석하였다. 특히, 실체의 변환 및 검증 절차로서, 적용 가능한 사항에 대한 분석에 주안점을 두었다. 분석을 통해 얻은 절차는 다시, 비교 분석을 통해 종합되었다.

1절 국내외 기존의 마이그레이션 절차 사례

1. ISO 전자기록물 변환 프로세스(Digital record conversion and migration process)

전자기록물 변환 프로세스는 ISO(International Organization of Standardization)의 TC 46(정보 및 도큐멘테이션) / SC 11(아카이브 및 기록관리)에 의해 개발 중인 기록물 변환 절차 표준안이다.²⁸⁾ 이것은 현재 가장 활발하게 사용되고 있는 보존 전략인 마이그레이션의 과정에서 발생할 수 있는, 기록의 무결성, 진본성, 신뢰성의 훼손 가능성을 인식하고, 이를 최소화하기 위해, 요구되는 표준 절차의 정의를 목표로 하고 있다.

이 표준안은 총 7개의 부분으로 구성되어 있다. 첫 번째 부분은 범위와 목표 설정으로서, 보존전략으로서의 마이그레이션의 정의와 표준 절차의 필요성을 설명하고 있다. 두 번째 부분과 세 번째 부분은, 이 안을 구성하기 위해 참고한 사례와 주요 용어의 정의를 다루고 있다.

본 표준안의 4에서 6장에 이르는 내용은 표준 절차를 제시하기 위한 고려 사항으로서, 각 장에서는 각 조직 및 업무 프레임워크, 기록보존 시스템의 요건, 기술적인 요인을 다루고 있으며, 마지막으로 기록물의 변환 업무를 위한 단계별 절차를 제시하고, 각 단계에서 고려해야 할 사항들을 제공한다.

ISO 표준안이 제시하고 있는 변환 활동의 기본적 요구사항은 아래와 같다.

- 콘텐츠의 완결성 (completeness)
- 무결성의 보존 (integrity)
- 진본성의 보존 (authenticity)

28) 2008년 8월 현재, 표준화 과정 중에 있다.

- 변환 시행에 대한 추적 가능성 (transparency)
- 기록물의 관리와 보존에 관련된 법령의 요건을 충족 (legal requirement)
- 다양한 플랫폼에 대한 적용 가능성 (platform flexibility)
- 업무 상 융통성 (business flexibility)
- 기술적 융통성 (technical flexibility)

또한, 본 표준안의 7장은 앞서 정의한 변환 업무의 요구사항을 충족시키기 위해서, 수행되어야 할 마이그레이션의 절차적 단계와, 단계 별로 필수적인 고려사항을 설명하고 있다.

가. ISO 전자기록물 변환 프로세스의 분석

<그림 2-1>은 ISO의 전자기록물 변환 프로세스의 절차를 그림으로 나타낸 것이다.



<그림 2-1> ISO 전자기록물 변환 프로세스의 변환 절차

(1) 변환 프로젝트 계획(conversion project planning) 단계

ISO 절차에서 가장 중점을 두고 있는 사항은 계획 단계이다. 앞서 정의하는 조직 및 업무 프레임워크 등의 내용 역시, 계획 단계에서 필요한 정보를 확보할 수 있는 분석 방법론으로서의 의미를 가지고 있으며, 실제 절차에 있어서도, 계획, 검증에 가장 많은 항목과 요구사항을 할애하고 있다.

변환 업무에 있어 계획은 이후의 실행 전체의 방향을 설정함으로, 프로젝트의 성패에 직결될 수 있는 사항이다. ISO가 제안하는 계획의 첫 단계는 변환 대상이 되는 기록물을 이해하는 것이다. 이를 위해 파악해야 할 기록물의 특징은 기록물의 보존연한(retention period), 적용되는 법률(laws and regulations), 역사적 가치(historical value), 업무지속성(business continuity)이다.

대상이 되는 기록물에 대한 명확한 정의를 내린 후, 대상 기록물이 변환될 목표 상태를 설정한다. 이 과정은 향후 실제 진행해야 할 과정을 정의하기 위해 필요한 사항을 확인하는 것으로, 변환 이후 전자기록물의 구조에 대한 모델 설정을 포함한다.

변환 방법과 활동 개발 활동은 앞서 확인한 기록물의 특성을 고려하여, 변환 시 고려해야 할 요구사항을 확인하는 과정이다. 요구사항이 모두 결정되면, 사전 검사와, 변환 이후 검증 작업을 위한 방법론을 개발하거나, 혹은 기존의 방법론을 활용할 수 있다.

변환이 완료된 이후 변환 이전의 기록물 폐기 여부를 결정한다. 결정을 위해 고려할 사항은 법적 요구사항, 기관의 성격 등이다. 모든 업무가 계획되면, 계획을 실행하기 위해 필요한 인력, 장비, 예산 등 필요자원을 확인한다.

계획의 최종 단계는 ‘문서화’를 위한 방법 결정과 작성된 계획의 문서화 및 승인이다. 계획 문서는 변환 프로젝트의 시행 근거로부터, 변환 프로젝트의 특성과 시행을 위한 요구사항을 포함한 프로젝트 전반을 설명해야 하며, 요구되는 자원과 시행 일정 까지 구체적인 정의를 요구한다.

(2) 사전 검사(testing) 단계

사전 검사 단계는 실험실 연구와 유사한 절차이다. 이 단계에서 확인해야 할 사항은 실제 데이터의 상태, 이전 단계에서 정의된 계획의 적합성, 성공 여부에 대한 검증 수단이다.

가장 먼저 수행하는 단계는, 검사 계획 단계이다. 이 단계에서는 프로젝트 계획 단계와 유사한 방식으로, 사전 검사를 위한 계획을 수립하게 된다. 검사 계획이 완료되면, 이를 바탕으로 필요한 인력, 도구, 작업 환경을 확인 하고 설정하는 검사 환경 설정의 단계를 거친다. 이어서, 복사본을 활용 하여, 실제로 검사를 시행하고 실행 중 계획과 다르게 수행된 사항이 발견되면, 검사 결과와 함께 기록한다.

결과 평가 단계에서는, 검사 시행에 대한 기록을 확인하여, 사전 검사가 목표를 달성하였는지를 결정하는 단계이다. 결과에 대한 평가가 완료되면, 최종적으로 보고서를 작성한다. 보고서는 향후 실제의 변환 수행 절차의 적합성에 대한 사전 근거가 되며, 검증 절차에 있어서는 해당 프로젝트의 특성상 주목하여 검증해야 할 요소를 제시할 수 있다.

(3) 변환(conversion) 단계

변환 단계는 변환에 대한 준비로 시작한다. 준비 단계에서는 프로젝트 계획 절차에서 정의된 변환방법과 활동의 요구사항을 중심으로, 실제 기록물의 변환을 시행하기 위한 준비가 되어있는지 확인하는 단계이다. 변환될 기록물의 정의 단계는 각 기록물 객체의 특성을 명확히 하고, 메타데이터를 확인하는 절차이다. 기록물의 정의가 완료 되면, 변환 시행을 위해 필요한 하드웨어와 소프트웨어 환경을 설정하고, 변환을 실행한다. 변환 실행이 완료되면, 변환 과정에서 발생한 오류를 확인한다. 오류에 대해, 가능하다면 수동으로 오류를 수정할 수 있으며, 그렇지 않을 경우, 목표 상태에 도달할 때까지, 변환과정 전체를 재실행하게 된다.

(4) 검증(validation) 단계

검증 단계의 목표는 마이그레이션 과정에서 무결성을 저해하는 변화가 발생하였는가의 평가와 내용적 변화에 대한 평가 및 수용 여부 결정이다.

검증 단계는 검증 책임자를 규정하는 것으로 시작한다. 검증 책임자는 변환 수행자가 아닌 제3자 혹은 가능하다면 검증 책임 팀 (steering team)을 구성하는 것도 좋다. 검증 활동 결정 단계는 검증할 데이터, 구성요소, 시스템에 대한 범위를 규정하는 과정이다.

검증 책임자와 검증 대상이 명확해졌다면, 검증 환경을 설정한다. 이는 검증에 필요한 모든 자원을 확보하는 것을 포함한다. 검증 환경이 설정되면, 실제 검증을 시행한다. 검증 단계는 계획에 충실해야 한다. 검증의 결과는 발견된 각 오류에 대한 상세한 설명을 포함한다. 이것은 명확하게 문서화 되어야 한다.

검증 결과가 문서화 되면, 이를 바탕으로 질적 평가를 통해 보고서를 작성한다. 결과 보고가 완료되면, 변환된 각 디지털 객체에 대한 메타데이터를 작성한다. 메타데이터는 새로운 객체의 식별 정보와, 현재까지 진행된 변환 절차를 포함하여야 한다. 이후 법령 및 기관의 정책에 따라, 변환 이전 기록물의 삭제 여부를 판단한다. 모든 기존 디지털 객체의 처리가 완료되면, 새로운 기록물을 실제 시스템 상에서 활용 가능하도록 한다. 절차가 완료되면, 최종 보고서를 작성하고, 이에 대한 최종 승인을 획득한다.

나. 변환 및 검증 단계의 요구 사항

변환 단계에서 요구 되는 주요 사항은, 변환 과정 중 오류가 발생하였을 경우, 이를 되돌릴 수단을 확보하는 것과, 변환이 예정대로 수행될 수 있도록 실제 전자기록물의 디지털 객체에 대해 모든 정보를 명확히 하는 과정에 중점을 두고 있다. 자세한 요구사항은 <표 2-1>과 같다.

<표 2-1> 변환 절차의 단계별 요구 사항

단계	요구사항
변환 준비	시행을 위해 필요한 모든 IT 자원에 대한 허가 획득
	변환 시행 이전에 대상 데이터의 복사본 보존
	계획단계에서 규정된 모든 데이터 클리닝 시행
변환될 기록물의 정의	디지털 객체의 내용, 포맷, 각각의 관계 구체적 정의
	기록물을 구성하는 모든 디지털 객체와 그 구성요소의 통합적 관리
	대상 기록물의 내용을 정의 (기간, 관리부서, 위치 등 정보 포함)
	디지털 객체의 포맷, 의미에 대한 정보 수집
	디지털 객체의 의미를 설명하기 위한 정보(메타데이터) 수집
	향후 활용을 위한 고유의 식별자 확보
	변환될 기록과 다른 기록 혹은 업무과정 사이의 관계 정의
	디지털 객체의 원본(초기생성본) 및 그것에 가해진 모든 변형에 대한 정보와 책임자 정보 확보
	데이터 무결성 확인, 검증 프로세스를 위한 정보 확보 (예. 전자서명)
변환 환경 설정	사전 검사 시의 환경이 실제 변환에도 적합한 경우가 이상적
변환 실행	변환의 유형에 따라 다양한 실행이 가능함
오류 검사 및 수정	무결성에 영향을 줄 수 있는 모든 오류에 대해 품질 검증 수행
	모든 변환 과정에서 지속적으로 수행한 후, 최후에 재확인 하는 것이 이상적
	CRC 체크, 각 데이터 파일에 대한 checksum, 해쉬 알고리즘, 시스템 로그 분석 시행
	목표 상태에 도달할 때까지, 오류 수정 및 재전송 반복

ISO의 마이그레이션 프로세스의 검증 절차는 검증 절차에 대한 보고사항이 갖추어야 할 요소에 초점을 맞추고 있다. 이에 따라, 각각의 오류에 대한 로그 파일이 포함해야할 정보, 최종 검증 결과 보고서가 포함해야할 정보를 규정한다. 또한, 기술적 메타데이터가 갖추어야할 가장 일반적인 내용을 <표 2-2>와 같은 3개의 범주로 제안한다.

<표 2-2> 검증 단계의 고려 사항

단계	요구사항
오류 로그(log) 항목	오류 번호(issue number)
	설명(description)
	발생가능성(probability) 높음 - 중간 - 낮음
	영향(impact) 높음 - 중간 - 낮음
	포함된 행위와 소유자

	발견 일자(due date)
	해결 일자(resolution date)
결과 분석 보고서 항목	무엇이 수행되었고 결과로 얻은 것은 무엇인가?
	특별히 고려한 사항은 무엇인가?
	검증 절차는 적합하게 수행되었는가?
	변환 프로젝트 수행 중 작성된 모든 문서에 대한 요약
	검증에 관한 문서의 위치
	변환 및 검증에 관한 문서의 보존 연한
메타데이터의 기본적 내용	디지털 객체의 식별
	기록물의 구조, 포맷, 의미
	기록물의 무결성에 영향을 미칠 수 있는 모든 변화

2. AHDS 이관 절차 프레임워크(AHDS Ingest Procedures Framework)

AHDS(Arts and Humanities Data Service)는 영국의 JISC의 지원으로 운영되는 국립 디지털 아카이브이다. 그 목적은 예술 및 인문학 관련 자료를 수집하고 제공하여, 관련 교육 및 연구 활동을 지원하는 것이다.

AHDS의 보존 전략은 마이그레이션을 기반으로 하고 있다. AHDS는 마이그레이션 전략을 수행함에 있어 발생할 수 있는 위협인 진본성 및 무결성의 훼손에 대응하기 위한 첫 단계로 AIP 생성을 위해 충분한 정보의 확보와 이관 시 무결성의 유지를 들고 있다. 이관 절차 프레임워크는 이와 같은 요구사항을 충족하기 위한 이관 절차를 제안하는 것을 목표로 한다.

AHDS는 매우 다양한 기탁자로부터 기록물을 입수하고 있으며, 이 과정에서, 각 기탁자가 이관시 제공하는 정보를 통합해야할 필요와, 모든 기탁자가 OAIS 표준에 적합한 SIP를 제공함을 확인해야할 필요가 있다.

이관 절차 프레임워크는 이와 같은 요구사항을 최우선으로 고려하여, 기록물의 이관의 각 단계와, 각 단계에서 확인되어야 할 단계별 목표를 제시하고 있다.

가. AHDS 이관 절차 프레임워크의 분석

<그림 2-2>는 AHDS의 전체적인 이관 절차를 그림으로 나타낸 것이다,



<그림 2-2> AHDS 이관 절차 프레임워크

(1) 준비(preliminary or pre-accession) 단계

준비 단계는 이관될 기록물 및 이에 포함되는 각 정보 객체에 대한 명확화와 전반적인 계획의 수립으로 구성되어 있다.

첫 번째 단계는 보존 대상이 되는 기록물에 대한 정보를 확인 하는 것이다. 이 후, 이관될 각 정보 객체들의 포맷 및 구조를 포함한 기본적 정의를 수행한다. 대상 기록물에 대한 정보를 파악하면, 이를 바탕으로, 디지털 객체를 보존할 수 있는 가능성을 모든 측면에서 검토한다. 보존 가능성에 대한 검토가 끝나면, 보존을 수행하기 위해 필요한 자원을 장기 보존 및 단기 보존의 모든 관점에서 분석한다. 분석이 끝나면, 적합한 이관 일정을 포함하여, 기탁자와 합의 과정을 거친다.

(2) 입수(Accession/Submission) 단계

데이터를 이관하기 위한 첫 단계는 사전 전송 테스트를 수행하고 기탁 협약 문서를 작성하는 것이다. 협약이 완료되면, 실제 데이터를 전송한다. 데이터 전송이 완료되면, 입수 절차 수행에 대한 메타데이터를 우선 업데이트한다. 이 과정에서 기존 메타데이터에 적합한 항목이 존재하지 않는다면, 이를 추가하게 된다. 메타데이터가 추가된 후, 바이러스 검사 단계를 수행한다. 바이러스 검사가 완료되면, 미디어/파일이 읽을 수 없는지, 읽을

수 있다면, 데이터의 완결성/무결성에 이상이 없는지 테스트 한다.

앞선 검사가 파일 자체를 대상으로 한다면, 다음 절차는 메타데이터의 비교이다. 이 단계에서는 전송 전의 메타데이터와 전송 후의 메타데이터를 비교하여, 일관성을 검사한다. 모든 검사가 완료된 기록물은 이후의 작업을 위한 공간으로 이동된다. 마지막으로 이관된 기록물의 진본성 확인(authentication)을 수행한다.

(3) 관리 및 메타데이터 관리(administration and metadata management) 단계

입수가 완료 되었으면, 이관 요청서의 모든 항목이 빠짐없이 작성되었는지 확인한다. 확인이 완료되면, 저작권 등 관련 법적 절차를 통해, 모든 기록물이 처리 가능함을 확인한다. 조건이 만족되면, 실제 문서화 작업을 통해 메타데이터 데이터베이스를 업데이트한다. 이 때, 디지털 객체 중 스캔된 문서가 존재한다면, 스캔된 문서 파일에 대한 위치를 생성하고, 스캔 작업의 설정을 포함하여, 메타데이터를 생성한다.

(4) 품질 확인(Quality assurance) 단계

메타데이터 생성이 완료되면, 이관된 기록물에 대해 다시 한 번 확인하고, 검증하는 절차를 수행한다. 우선, 검증 계획을 수립한다. AHDS는 파일명, 파일구조, 데이터의 세 측면에서 일관성 검사를 시행하고 있다. 파일명 일관성 검사는 파일명이 현재 시스템에서 사용 가능한 형태인지와 파일 포맷 및 확장자, 버전을 확인하는 과정이다. 파일 구조 일관성 검사는 해당 파일이 구조화에 대한 명세서나 규정을 가지고 있는 경우, 해당 명세와 실제 파일의 구조가 일관되어 있는지를 확인하는 과정이다. 데이터 일관성 검사는 각 디지털 객체 사이에 링크, 연결이 존재하는지 확인하는 과정이다. 만일 코드(code)가 포함되어 있다면, 각 코드가 코드 규칙에 위배되는 것은 없는지 검증한다. 검증이 완료되면, 검증의 모든 과정에서 발생한 변화를 포함하여, 메타데이터를 업데이트 한다.

(5) 보존 버전 생성(creating preservation version) 단계

보존 버전의 생성은 보존 전략을 선택하는 것으로 시작된다. 보존 전략의 선택은 기록물에 포함된 디지털 객체의 파일 포맷에 따라 달라진다. 각 파일 포맷에 따라, 에물레이션 혹은 마이그레이션의 전략이 선택된다.

보존 전략이 선택되면, 해당 전략에 적합한 상태로 파일 포맷을 변경하게 된다. 이를 위해 변경 과정에 대한 계획을 수립한다. 계획이 수립되면, 계획에 따라, 파일 포맷을 변환하는 절차가 수행되며, 변환된 결과는 그 형태와 내용 등이 원본과 다르지 않은지 검증한다. 검증이 완료되면, 변환과 검증 절차를 포함하여 진본성 확인을 위한 전자서명 등의

정보를 추가한다. 모든 변환 과정이 완료되면, AIP 생성을 위해 보존 메타데이터를 업데이트 한다.

(6) AIP 생성(Prepare the archival information package) 단계/AIP 승인(submit AIP for preservation) 단계

AHDS 이관 절차 프레임워크의 최종 단계는 장기보존을 위한 AIP를 생성하고, 이를 승인하여 아카이빙을 완료하는 단계이다.

나. 입수 및 품질확인 단계의 요구사항

AHDS의 요구사항은 진본성의 확보를 위해 충분한 도큐멘테이션의 생성을 중심으로 하고 있다. 이에 따라, 대다수의 항목에 대한 메타데이터의 기록을 요구하고 있으며, 각 항목에서 확인해야 할 사항에 대한 구체적인 기준을 제시한다. AHDS의 입수 및 품질확인 단계의 요구사항을 정리하면 <표 2-3>과 같다.

<표 2-3> AHDS의 입수 단계 요구사항

단계	요구사항
데이터 전송 세션	전송 세션은 기술적으로 관리되어야 한다.
	기탁자는 기탁협약에 포함된 모든 파일이 성공적으로 전송되었다는 확인서를 수취하여야 한다.
	전송에 문제점이 발생하면, 재전송을 시행하고, 이를 메타데이터에 기록한다.
메타데이터 업데이트	데이터 전송 완료에 대한 메타데이터를 업데이트 한다.
	관련 항목이 없다면, 새로운 항목을 생성한다.
바이러스 검사	바이러스 검사는 모든 파일에 실행한다.
	바이러스가 발견되면, 기탁자에게 해당사항을 경고하고, 대체파일을 요청하여, 전송한다.
	대체 파일이 없다면, 해당 파일은 제거된다.
	모든 바이러스 발견에 대한 상세 정보를 메타데이터에 기록한다.
	바이러스 검사 단계를 모든 파일이 안전하거나, 위험한 파일이 제거될 때까지 반복한다.
매체와 파일 이용가능성 검사	입수된 모든 파일과 매체에 대해 시행한다.
	읽을 수 없는 부분을 발견하면 대체물을 요구한다.
	압축 혹은 암호화된 형태로 입수하면, 압축을 해제하거나, 암호를 해제할 수 있는 도구를 가지고 있어야 한다.
	데이터 변조가 발견되면, 대체물을 요구한다. 모든 비정상적인 데이터 검출은 메타데이터를 통해 기록한다.

단계	요구사항
데이터 자원 완결성/무결성 검사	입수된 데이터는 적합한 디스크 영역에 저장되어야 한다.
	기존의 디렉토리 구조를 반영한 모든 디렉토리를 생성하고, 파일들을 해당 디렉토리에 복사해야 한다.
	각 디렉토리의 파일 수, 파일명, 파일 포맷은 목록과 비교하여 일치하여야 한다.
	목록의 파일이 누락된 경우, 기탁자에게 해당 파일을 요청한다.
	목록에 존재하지 않는 파일이 발견될 경우, 이 파일에 대한 도큐멘테이션을 요청한다.
	누락 사항을 보완할 수 없는 경우, 이것을 메타데이터에 기록한다.
도큐멘테이션 완결성 검사	현재의 도큐멘테이션과 이전의 메타데이터 사이에 불일치를 검출한다.
	심각한 불일치나 누락이 발견될 경우, 기탁자에게 추가적인 도큐멘테이션을 요구하거나, 해당사항에 대한 설명을 요구한다.
	전자기록물로서 입수되지 않았으나, 디지털화 되어야 하는 도큐멘테이션은 메타데이터에 남길 수 있다.
	도큐멘테이션의 포맷은 표준화되어 있어야 하며, 검증절차를 자동화할 수 있어야 한다.
작업영역으로 복사	기존의 디렉토리 구조는 보존되어야 한다.
	작업영역은 공개적이지 않아야 하며, 삭제로부터 보호되어야 한다.
원본 버전의 진본성 확립	의도하지 않은 변조로부터 파일을 보호하기 위한 자동화된 절차 (Checksum 계산, 메타데이터에 기록)를 거쳐야 한다.

품질 확인 단계는 실질적인 수행절차와 함께, 파일 포맷에 따라 달라질 수 있는 검증 기준의 예를 제시하고 있다. 품질 확인 단계는 입수 단계의 각종 검사와 유사하나, 좀 더 자세하게 확인하고 있으며, 마찬가지로 메타데이터에 모든 사항을 기록하도록 요구하고 있다. <표 2-4>는 이러한 요구사항을 정리한 것이다.

<표 2-4> AHDS의 품질 확인 단계 요구사항

단계	요구사항
데이터 재확인 및 검증 준비	검사 되어야 할 모든 파일이나 그 부분들 확인. (예. 표가 있는가?)
	각 파일 혹은 그 부분의 검증을 위해 사용할 수 있는 도큐멘테이션 확인
	각 파일이나 그 부분에 사용할 수 있는 검증 방법 수립
	검증절차를 문서화할 방법 수립
일관성 검사	일관성 검사는 그 과정(로그파일)과 결과(메타데이터)를 기록할 수 있도록 자동화 혹은 반자동화 되어야 한다.
파일 수준 일관성 검사	도큐멘테이션의 파일명과 실제 파일명이 일치하는지 검사한다.
	처리 불가능한 파일명 (공백, 특수문자, 확장자형식)은 적합한 것으로 수정되며, 변경된 파일명과 기존 파일명은 모두 메타데이터에 기

단계	요구사항
	록되어야 한다.
	도큐멘테이션의 파일포맷과 실제 파일포맷이 일치하는지 검사한다.
	파일 포맷 검사는 포맷 버전 검사를 포함한다.
파일 구조 수준 일관성 검사	텍스트 파일의 문자 인코딩이 적합한지 검사한다.
	구조화된 텍스트 파일은 관련된 정의 및 구조를 통해 검사한다.
	구조에 대한 형태적 설명이 존재하는 모든 텍스트 파일은 그 구조와 의 적합성을 검사한다.
	바이너리 파일은 해당 파일을 생성한 소프트웨어 및 버전으로 확인 하여 구조 및 포맷의 일관성을 검증한다.
파일 내용/ 데이터 수준 일관성 검사	자원의 유형이나 파일 포맷에 따라 요구되는 데이터 형태와 실제 데 이터 형태가 일치하는지 검사한다.
	비 일관성이 발견되면, 문제 해결을 시도해야 한다.
	해결할 수 없는 경우, 기탁자에게 추가적인 도큐멘테이션이나 재 기 탁을 수행해야 한다.
메타데이터 업데이트	문제점이 해결되지 않으면, 메타데이터에 해당사항을 기록해야 한다.
	이용자에게 중요한 도큐멘테이션이 누락되었을 경우, 기탁자에게 다 시 요청한다.
	아무리 사소한 변화라도 데이터 검증 과정에서 가해진 모든 변화를 메타데이터에 기록한다.
	일관성 검증 과정에서 발생한 모든 변화를 메타데이터에 기록한다.

3. CCSDS의 생산자-기록관 인터페이스 방법론 표준 제안

(Recommendation for a producer-archive interface methodology abstract standard)

CCSDS(Consultative Committee for Space Data System)는 우주 관련 정보의 상호 교환을 촉진하며, 이를 위한 비용과 위험을 최소화하기 위해 설립된 국제 포럼이다.

생산자-기록관 인터페이스 방법론은 이관(ingest) 단계에서 발생하는 생산자와 아카이브 사이의 상호작용을 일반화된 단계들로 구분하고, 각 단계에서 발생하는 행위와, 단계의 목적을 정의하여, 향후 표준안의 개발이나, 소프트웨어 개발의 기반을 제공하는 것을 목표로 하고 있다.

이 방법론은 우선, 첫 장을 통해 그 목적과 적용 범위, 용어를 정리하고 있으며, 2장을 통해 기본적인 절차의 구성과 각 절차의 관계를 표현한 후, 3장에서 각 절차 별 세부항목과 수행해야 할 사항을 제공하고 있다. 마지막 4장에서는 이 제안을 활용해, 표준을 만들

기 위한, 접근법에 대해 간략히 설명한다.

가. 전체 절차 분석

<그림 2-3>은 CCSDS의 생산자-기록관 인터페이스 방법론의 이관 절차를 그림으로 나타낸 것이다.

(1) 준비(preliminary) 단계

이관의 첫 단계는 최초 교섭단계이다. 이 단계에서는 접촉하는 개인과 조직을 확인하고, 일반적인 정보를 교환한다.



<그림 2-3> CCSDS 생산자-기록관 인터페이스 방법론 표준 제안

접촉이 완료 되면, 사전에 정의해야할 사항을 정의 하고, 이관의 타당성을 평가한다. 이 단계에서 확인하는 것은, 보존 대상인 정보와, 이 정보의 특성, 수량, 법적 고려사항, 이관 및 검증절차에 필요한 비용과 일정이다.

다음 단계는 예비 협약을 수립하는 단계로, 앞서 정의된 사항을 정리하고, 이를 바탕으로 이관 절차를 다음 단계로 진행해야할지, 중단할지를 결정하는 예비 협약 수립 단계이다.

(2) 공식 협약(formal definition) 단계

준비 단계가 완료되면, 이관 수행을 위해 공식적인 협약을 작성하고, 이를 승인하는 단계가 수행된다.

우선, 공식 협약 단계에서 수행될 요구사항을 공식화하는 조직화 단계를 시행한다. 조직화가 완료되면, 우선, 보존할 정보와 전송 방법에 대해 공식화한다. 이 단계는 이관 프로젝트가 진행되는 배경을 포함해 각 정보 객체의 특성과 전송 모델에 대한 정의를 포함한다.

이어서 법적, 계약상 고려사항을 공식화한다. 그 후, 이관 조건, 검증, 전송일정의 정의가 이어진다. 이관 조건은 이관의 사전 검사로부터 시작해 이관의 절차, 사용되는 도구와 전송 방법에 대해 정의하는 단계이다. 검증에 대한 정의는 검증도구와 방법을 포함해, 검증 결과에 따른 대응의 정의를 포함한다. 전송 일정은 디지털 객체에 따라 필요한 일정을 정의하는 단계이다. 모든 조건이 정의되면, 기탁협약을 종결하고, 타당성, 비용, 위험평가를 수행한다. 최종적으로, 기탁 협약을 작성한다.

(3) 이관(transfer) 단계

이관 단계는 ‘이관 테스트 수행’과 ‘이관절차 관리’로 나뉜다. ‘이관 테스트 수행’에서는 실제 이관을 수행하는데 이관 시 문제점이 발생하면 그 정도에 따라 기록관에서 직접 수정하거나 생산기관에 재 이관을 요구한다. 만약 기록관과 생산자가 바람직한 관계를 유지하고 있다면 이관 테스트는 이관 시점 마다 매번 수행할 필요는 없으나 새로운 정보객체의 등장 또는 정확성을 확보하기 위해 반드시 수행되어야 한다. 그리고 이관 시 전송상 장애 또는 흐름이 끊기지 않게 지속적으로 관리해주어야 하며 문제가 발생할 시 즉각 복구 할 수 있는 기술지원을 해주어야 한다.

(4) 검증(validation) 단계

검증단계는 이관된 입수정보패키지를 최종적으로 검사하며 이를 보존패키지로 생성하기 위한 단계이다. 검증절차는 단일한 도구에 의해서만 수행 되어서는 안 되며 여러 시스템에 의해 고도화된 수준으로 실행되어야 한다. 검증절차 시 오류가 발생하면 즉각 기록관은 생산자에게 이에 대해 알려주어야 하며 재이관을 요구해야 한다.

나. 이관 및 검증 단계의 요구 사항

CCSDS의 이관 절차는 아카이브와 기탁 기관 사이의 상호작용을 중심으로 구성되어 있다. 이에 따라, 이관 과정에서 생산자와 보존자가 수행하여야 할 과업을 정의한다. <표 2-5>는 CCSDS의 전송 절차에 요구되는 사항을 정리한 것이다.

<표 2-5> CCSDS의 전송 절차 요구사항

단계	요구사항
전송 테스트 단계	사전 전송 검사는 공식 협약단계의 정의에 따라 수행되어야 한다.
	사전 검사를 통해 발견한 비정상적 결과에 따라, 전송 과정의 설정을 수정해야 한다.
	사전 검사의 결과가 협약의 내용과 다를 경우, 협약을 수정 및 보완한다.
전송 관리 단계	아카이브와 생산자 사이의 세부 일정을 조정하고 관리한다.
	공식 협의 단계에서 정의된 절차를 준수한다.
	전송의 수행에 있어 기술적인 장애가 발생하지 않도록 한다.
	미디어가 전송될 경우, 손상이 발생하거나, 읽기 오류가 발생하지 않는지 확인한다.
	전송에 이상이 발생할 경우, 재전송한다.
	각 전송세션에 대한 확인서를 발급한다.

검증 단계에 대한 요구사항은 시스템적인 검증과 심층적 검증을 구분하여 제공하고 있으며, 검증의 사전, 사후에 수행되어야 할 관리적인 요구사항을 포함한다. <표 2-6>은 이러한 요구사항을 정리한 것이다.

<표 2-6> CCSDS의 검증 절차 요구사항

단계	요구사항	
검증 사전 테스트 단계	검증 절차 수행 전, 시스템적 검증 계획을 수립해야 한다.	
	검증 사전 검사는 정보 유형에 따라 개별적으로 이루어져야 한다.	
	사전 검사의 일정은 정보 유형에 따라 개별적으로 존재하며, 새로운 정보 유형이 발견되면, 이를 일정에 추가하여야 한다.	
검증 수행 단계	시스템적 검증	검증 절차는 각 전송 세션마다 수행되어야 한다.
		공식 정의 단계에서 정의된 계획에 따라, 시스템적 검증 절차가 구축되어야 한다.
		부정합이 발견되면, 생산자에게 통보되어 아카이브와 생산자가 동시에 수정하도록 한다.
	심층적 검증	심층적 검증은 특정한 정보의 집합이나, 모든 기록물이 전송된 이후 수행 된다.
		공식 정의 단계에서 정의된 계획에 따라, 심층적 검증을 수행한다.
		자동적인 검증을 위한 도구가 준비되어 있어야 한다.
검증 결과 관리 단계	아카이브는 모든 부정합과 규정 위반에 대해, 공식 협의에서 정의된 절차에 적합하도록 보고해야 한다.	
	아카이브와 생산자는 비정상적 형태를 관리해야 한다.	



1절. InterPARES 생산자, 관리자 가이드라인

InterPARES 프로젝트는 디지털로 생산된 기록물이 안고 있는 위협에 대응하여, 기록의 진본성을 유지하기 위하여, 1999년으로부터 현재까지 수행되고 있는 전자기록의 장기 보존에 대한 국제적인 연구 프로젝트이다.

InterPARES 1(1999-2001) 프로젝트는 보존자의 입장에서, 데이터베이스 및 문서관리 시스템을 통해 생성되고 유지되었으나, 더 이상 생산자가 필요로 하지 않게 된 진본 행정 기록의 보존을 연구하였다.

InterPARES 2(2002-2007) 프로젝트는 디지털 기록의 생산에서 보존에 이르기까지 기록의 신뢰성과 정확성, 진본성을 보장할 수 있는 이론과 방법을 개발하기 위하여, 특히 정부기관, 과학기술분야, 예술분야에서 생산되는 기록물을 주로 다루었으며, 생산자의 관점을 다루었다.

InterPARES 3(2007-2012) 프로젝트는 이전의 InterPARES 프로젝트의 연구 성과를 비롯한 다양한 전자기록물 장기보존에 대한 연구 결과를 실무에 적용하는 것을 목표로 하며, 특히, 중소규모 기록관 및 교육 프로그램을 대상으로 하고 있다.

본 연구에서는, 이관 절차의 각 업무에 따라서 InterPARES에서 제시하고 있는 생산자, 보존자 관점의 가이드라인을 적용한다. 이를 통해, 장기보존에 필요한 고려사항에 대한 국제적 연구 성과를 실무적인 관점에서 분석하고 활용하려 한다. 국내 적용 가능한 마이그레이션 절차와 기능들을 정의하고 구체적인 기능들을 제시함으로써 국내외 적인 검증 방법으로서 IP3에서 제시하는 가이드라인을 참조하여 연구의 검증 과정을 확인하고자 한다.

1. InterPARES 생산자 가이드라인

생산자 가이드라인은 디지털자료를 생성하는 개인들이, 장기 보존이 가능한 방식으로 자료들을 만들고 관리하는 것과 관련된 정보를 제공하여, 결정을 내릴 때, 도움을 주는 것을 목표로 한다.

내용은 총 10개의 기본적인 지침을 바탕으로, 상세 지침을 전개하는 방식을 취하고 있다.

가. 접근성(Accessibility)

우선, 접근성에 관련한 지침은 <표 3-1>과 같다.

<표 3-1> 접근성 관련 생산자 가이드라인

지침	세부 지침
1. 시간이 지난도 디지털 자료를 여전히 쉽게 이용할 수 있도록 보증하는데 가장 유망한 하드웨어, 소프트웨어, 파일포맷을 선택한다.	1.a 자료를 원래의 모습대로 보여주는 소프트웨어를 선택한다.
	1.b 디지털 자료를 쉽게 공유할 수 있게 해주는 소프트웨어 및 하드웨어를 선택한다.
	1.c 표준을 준수하는 소프트웨어를 사용한다.
	1.d 소프트웨어 명세를 잘 보관한다.
	1.e 소프트웨어를 개인 환경에 맞춰 조정하였다면, 변경한 내용을 기록으로 만들었는지 확인한다.
	1.f 시스템 구성 전체를 문서화하여 접근성을 확보하도록 한다.
	1.g 널리 통용되고, 비독점적이며, 여러 플랫폼에서 호환되며, 가능한 자유롭게 명세서를 구할 수 있는 비 압축 포맷을 선택한다.

접근성의 확보는 적절한 소프트웨어를 보유할 수 있는가에 달려있다. 이를 위해 소프트웨어는 소급호환성, 미래호환성, 상호호환성의 요소를 가지고 있어야 한다. 소급호환성은 이전의 버전과의 호환성이며, 미래호환성은 이후 버전과의 호환성이다. 상호호환성은 서로 다른 애플리케이션간의 호환성이다.

또한 강조하는 사항은 충분한 문서화확보를 하는 것이다. 1.d, 1.e, 1.f, 1.g가 이에 해당한다고 할 수 있다. 문서화는 생산자에 의해 작성되는 사항, 기록물의 변경 사항 전체 뿐 아니라, 해당 포맷에 대한 명세서의 확보 가능성도 포함한다.

1.g와 같은 포맷을 오픈 포맷이라고 부른다. 오픈 포맷은 명세와 소스 코드가 공개되어 있어, 자유롭게 이용할 수 있으며, 로열티로부터 자유롭다.

나. 고정성(Fixity)

<표 3-2>는 고정성과 관련되어 생산자에게 권고되는 사항을 담고 있다.

<표 3-2> 고정성 관련 생산자 가이드라인

지침
2. 기록으로 관리하는 자료가 그 내용과 형식상 안정적이고 고정되어 있는지 확인한다.

기록물의 특징 중 하나는 그 내용이 바뀌지 않으며, 바뀌지 않아야 한다는 점이다. 즉, 기록물에 포함된 정보와 데이터에 덮어쓰기, 변경, 삭제, 확장이 불가능해야 한다. 기록물

은 고정된 양식(fixed form)으로 안정된 내용(stable content)을 가지고 있는 것을 말한다.

여기서 고정된 양식이란, 문서의 외관 또는 구현모양의 동일성을 말하며, 안정된 내용이란 포함된 정보와 데이터가 불변이며, 업데이트 정보를 첨부하거나 새로운 버전을 통해 변화가 이뤄지게 하는 기록의 특질이다.

이와 같은 사항을 유지하기 위해서는 내용과 문서양식을 선택하기 위해 고정된 규칙을 확립하고, 이에 따라 안정된 변수만을 허용해야 한다. 이것을 제한된 변수(bounded variability)라 한다.

문서양식이란 기록의 내용, 관리상 맥락 및 기록화 맥락과 그 권위의 소통방식을 정하는 구현 규칙으로서, 외형적요소와 내재적 요소를 모두 포함한다. 외형적 요소란 글씨체, 그래픽, 이미지, 레이아웃, 하이퍼링크 등 외부적인 형태를 구성하는 요소이다. 외형적 요소는 디지털 서명, 인장, 타임 스탬프 등 특정한 표시를 포함한다. 내재적 요소는 기록이 관여하는 활동과 그 직접적인 맥락을 전달하는 기록의 요소이다. 내재적 요소에는 생산자의 이름, 활동이나 문제의 명칭과 설명, 제작 날짜, 전송정보 등이 포함된다.

다. 정체성(Identity)

<표 3-3>은 정체성과 관련하여 생산자에게 권고되는 사항을 정리하고 있다.

<표 3-3> 정체성 관련 생산자 가이드라인

지침	세부 지침
3. 디지털 자료의 정체성을 적절히 식별할 수 있도록 한다.	3.a 생산에 참여한 사람의 이름 (발행자, 편집자, 생산자, 수신인, 접수인)
	3.b 해당 활동 및 사안의 명칭 (표제, 제목)
	3.c 문서 형식 (보고서, 서신, 계약서, 표, 목록 등)
	3.d 디지털 표현 방식 (포맷, 래퍼, 인코딩 등)
	3.e 생산 및 이관 날짜 (자료가 편철된 날짜, 전송 또는 수령한 날짜, 보존 또는 편철 날짜)
	3.f 문서 맥락 (분류코드, 디렉토리 명칭, 편철지침이나 분류표 내 해당 편철 위치, 자료가 속하는 기록의 상위계층 등)
	3.g 첨부물 표시
	3.h 저작권 및 기타 지적 권리 표시
	3.i 디지털 서명의 존재 및 제거 표시
	3.j 기타 인증 형식 표시 (확증 근거, 증명, 서명, 서명인의 요건)
	3.k 초안 및 버전 번호 표시
	3.l 디지털 시스템 외부의 자료 사본 존재여부 및 위치

전자 기록물의 식별과 검색을 위해서는 메타데이터가 요구 된다. 위의 지침은 정체성

식별을 위해 기록물의 메타데이터가 포함해야할 요소를 다루고 있다. 단 3.g로부터 3.1에 해당하는 요소는 적용이 가능할 경우 추가할 수 있다.

라. 무결성(Integrity)

<표 3-4>는 무결성과 관련하여 생산자에게 권고되는 사항을 보여주고 있다.

<표 3-4> 무결성 관련 생산자 가이드라인

지침	세부 지침
4. 디지털 자료가 그 무결성을 입증하는 데 필요한 정보를 포함하고 있는지 확인한다.	4.a 취급해온 개인/부서명
	4.b 자료를 관리할 기본적인 책임을 지고 있는 개인/부서명
	4.c 자료에 첨부된 주석 표시
	4.d 자료 및 자료를 관리하고 접근을 제공하는 애플리케이션에 대한 기술적 변경사항 표시
	4.e 접근 제한 코드
	4.f 접근 권한 코드
	4.g 필수 기록 코드
	4.h 계획된 처분

무결성은 디지털 자료가 본래상태를 유지하고 훼손되지 않았을 때, 즉 메시지가 원래 모습 그대로 변조되지 않았을 경우를 말한다. 내용 및 문서양식의 필요한 요소들의 명료성이 동일하게 유지될 경우 무결성이 유지되었다고 말할 수 있다. 무결성과 관련된 속성들은 자료의 유지관리와 관련된다. 무결성은 정체성 메타데이터를 통해 디지털 자료가 처음 만들어 졌을 때와 동일한 상태임을 추론하게 할 때 유지되는 것으로 간주된다.

마. 조직화(Organization)

조직화와 관련하여 생산자에게 권고되는 사항은 <표 3-5>와 같다.

<표 3-5> 조직화 관련 생산자 가이드라인

지침
5. 디지털 자료를 논리적인 그룹으로 조직화한다.

디지털 자료는 일대일 보다 하나의 큰 집합으로 처리할 때 관리와 검색을 개선할 수 있다. 따라서 논리적 방식에 따라 디지털 자료들을 그룹화 해야 한다. 기록의 조직은 모든 자료의 군집을 보여주고, 간략한 문장으로 이를 설명해주고, 어떻게 서로 관련되어 있

는지를 표시해 주는 문서에 기록되어야 하며, 이를 분류표 또는 편철지침이라 한다.

바. 진본성(Authentication)

<표 3-6>은 진본확인과 관련하여 생산자에게 권고되는 사항이다.

<표 3-6> 진본확인 관련 생산자 가이드라인

지침

6. 디지털 자료의 유지관리 및 보존을 촉진하는 진본확인 기술들을 사용한다.

디지털자료를 다른 공간이나 시간으로 전송해야 한다면, 디지털 자료의 진본성을 확보하기 위해 기록의 정체성이나 무결성을 보장하거나 차후에 접속하기 위해 구분된 시점에서 변경될 위험성을 최소화하기 위해, 기술 및 행정적 절차들을 사용하여 관리되어 왔음을 보여주는 증거들이 추가되어야 한다.

진본확인은 기술 독립적 진본확인과 기술 의존적 진본확인으로 구분할 수 있다. 기술 독립적 진본 확인이란, 진본성의 추정을 말한다. 이는 진본성을 추정할 수 있도록 하는 문서가 만들어지고, 이를 근거로 진본성을 추정하는 것이다.

기술 의존적 진본확인이란 암호법과 전자 서명과 같은 디지털 기술을 활용하는 것이다. 단 디지털 서명은 기술적 변화에 따라, 노화될 수 있다. 때문에, 문서가 수령되었을 때 디지털 서명을 시행하고 이를 메타데이터에 추가 하여야 한다.

사. 보호(Protection)

전자기록의 보호와 관련하여 생산자에게 권고되는 사항은 <표 3-7>과 같다.

<표 3-7> 보호 관련 생산자 가이드라인

지침

7. 허가 받지 않은 행위로부터 디지털 자료를 보호한다.

보관된 디지털 자료를 흔적을 남기지 않고 수정할 방법이 있다면, 정확성과 진본성을 추정할 수 없다. 이를 위해서는 접근 권한의 구조를 확립해야 한다. 분류표나 편철지침에 따라 일단 정리된 기록은 어떤 수정도 불가능해야 하며, 기록의 보존 및 유지 관리 책임자만으로, 자료의 전송 삭제를 수행할 수 있어야 한다. 이와 같은 사항을 하기 위해 접근 기록을 확보하고 감사해야 한다.

아. 백업(Backup)

전자기록의 백업에 관련하여 생산자에게 권고되는 사항은 <표 3-8>과 같다.

<표 3-8> 백업 관련 생산자 가이드라인

지침	세부 지침
8. 예기치 않은 손실 및 수정이 일어나지 않도록 디지털 자료를 보호한다.	8.a 시스템을 매일 백업하고 있는지 확인하는 엄격한 정책 또는 절차를 개발한다.
	8.b 당신의 상황에 적합한 최선의 백업기술을 선택하고 설치한다.

디지털 자료는 손실 및 변조로부터 절대 안전하지 않다. 따라서 정기적인 백업을 반드시 수행해야 한다. 백업은 기록물 자체 뿐 아니라, 운영체제 소프트웨어를 포함하는 포괄적인 것이어야 하며, 물리적으로 별도의 장소에 보관할 수 있다.

자. 기술 노화(Obsolescence)

기술노화에 대비하여 생산자에게 권고되는 사항은 <표 3-9>와 같다..

<표 3-9> 기술 노화 관련 생산자 가이드라인

지침
9. 하드웨어 및 소프트웨어 노화에 대비한 조치를 취한다.

하드웨어나 소프트웨어가 구식이 되는 속도가 빠를수록 디지털 자료의 유지관리 및 장기 보존에 더 심각한 도전이 된다. 장기보존을 위해서는 하드웨어와 소프트웨어 시스템을 항상 최신으로 업그레이드해야 한다.

차. 주의사항(Awareness)

<표 3-10>은 전자기록의 장기보존에 있어 주의할 사항을 담고있다.

<표 3-10> 주의사항 관련 생산자 가이드라인

지침
10. 장기보존을 둘러싼 쟁점사항을 고려한다.

이를 위해서는 기록물이 생성되면, 이를 담당하고 관리할 책임자를 명확히 확인해 두어야 한다. 또한 가능한 신속하게 보존 전략을 수립해 두어야 한다.

2. InterPARES 보존자 가이드라인

이 지침은 디지털 기록의 장기 보존을 책임지는 다양한 그룹에 구체적인 조언을 제공하기 위해 개발되었다. 각 지침은 보존과정 전체의 맥락에서 보존링크에 초점을 맞추고 있으며, InterPARES의 보존 사슬 모형이 제시하는 연속적인 활동에 따라 조직되었다.

가. 보존 사슬 프레임워크의 관리

보존 사슬 프레임워크의 관리에 대한 가이드라인은 <표 3-11>과 같다.

<표 3-11> 보존 사슬 프레임워크의 관리

지침	세부 지침	
1. 보존 사슬 프레임워크의 관리	1.1 범위와 목적의 수립	
	1.2 자원 획득	
	1.3 디지털 기록에 대한 초점	
	1.4 조언 제시	
	1.5 모범 사례 실천	
	1.6 절차 개발	단절되지 않은 기록 보관의 유지
		보안과 통제 절차의 실행과 모니터링
		재생산 이후에도 기록의 내용과 필요한 주석, 문서 형태의 요소가 불변한 채 유지
	1.7 유지 전략의 실행	책임의 분명한 할당
		적절한 기술적 인프라
		시스템 유지, 지원, 대체를 위한 계획의 실행
		주기적으로 새로운 저장 미디어에 기록을 이전하는 계획의 실행
		적절한 저장과 저장 미디어의 취급 상태 준수
		디지털 객체의 주기적인 백업과 중복소장
		시스템 보안 수립
		재난 계획

디지털 보존 프로그램을 정의함에 있어, 그 범위와 목적을 정의하는 것은 필수 불가결이다. 같은 자원을 대상으로 하더라도, 범위와 목적에 따라 필요로 하는 보존 전략은 매우 달라질 수 있기 때문이다. 디지털 보존 프로그램의 정의는 선행 사례, OAIS 표준 등을 고려해야 한다.

디지털 보존은 자금, 기술적 역량, 전문성 등에서 자원을 필요로 한다. 디지털 보존 책임을 가진 조직은 a) 새로운 자원 획득, b) 기존 자원의 재배당, c) 다른 자원의 차입 등에서 적절한 전략을 선택할 수 있다. 디지털 보존 프로그램을 지원하는데 필요한 자원이 부족할 경우, 최소 기준을 충족시킬 수 있는 프로그램을 개발 및 구매할 수 있도록 공동구매나 협력관계 수립을 모색해야 한다.

보존자는 남아 있는 아날로그 기록을 디지털화하는 보존이 아닌, 권위 있는 디지털 사본을 보호하는데 디지털 보존을 위한 자원을 주로 배분해야 한다. 대부분의 아날로그 기록은 디지털화하지 않고도 보존할 수 있는 반면, 디지털 기록은 디지털 보존 프로그램 없이는 유실될 수 있기 때문이다.

디지털 보존자는 디지털 기록물의 생산과 유지에 대한 조언을 제공해야 한다. 조언의 대상은 보존자의 법적 의무에 따라 달라질 수 있다.

모범 사례의 실천은 다음과 같은 이점을 얻을 수 있다.

- 기록 생산자에 권고하는 기술을 보존자가 실습 훈련할 수 있는 기회
- 실제 기록 관리 솔루션을 육안으로 확인하고, 실제 일상적인 운영환경에서의 작동을 확인할 수 있는 기회
- 업그레이드와 혁신을 도입하여 평가할 수 있는 테스트베드 마련
- 시연에서 활용할 수 있는 작동 프로토타입 마련

보존 절차의 개발은 생산자의 기록을 기록관리 부서나 보존기록부서나 보존기록관으로 이전하는 데 사용하는 절차와 시스템을 포함하여, 기록의 이관, 유지와 재생산에 대한 통제 수립이다. 이러한 절차는 기록의 정체성과 무결성을 보장할 수 있도록 충분하고도 효과적인 통제를 구현하여야 한다.

장기 보존 전략에 많은 주의를 기울이는 반면, 기록관리나 기록 보존 시스템에서 기록을 적절하게 유지하고 보호하지 못한다면 무의미한 전략이 된다.

나. 기록의 평가

<표 3-12>는 기록의 평가와 관련하여 보존자에게 권고되는 사항을 정리한 것이다.

<표 3-12> 기록의 평가 관련 보존자 가이드라인

지침	세부 지침
2. 영구적으로 보존하여야 할 기록의 평가	2.1 초기 평가
	2.2 다수의 소유자를 확인하라
	2.3 진본성 평가
	2.4 진본성 평가의 기록화
	2.5 장기 보존하기로 결정한 기록의 모니터링

지침	세부 지침
	2.6 평가 갱신
	2.7 모든 디지털 구성 요소의 확인
	2.8 보존의 타당성 결정

어떤 기록을 장기적으로 보존해야 하는지에 대한 구별은 가능한 빨리 실행되어야 한다. 평가를 실행하고, 이관 방법을 수립하고, 기록 생산자와 함께 잠재적 보존 전략을 파악함으로써 성공의 가능성을 증진시킬 수 있으며, 기록의 생산과 유지에 대한 조언을 제공할 기회를 가질 수 있다.

디지털 객체의 지적 요소가 다수의 소유자를 가지는 경우, 이러한 상황의 영향을 평가하기 위해 이들 소유자의 신원을 확인해야만 한다.

진본성의 평가를 위해서는 우선, 기록의 생산시기로부터 장기 보존 주체에게 기록이 이전될 때까지 완전한 보존 사슬이 존속 되는 것을 확인해야 한다. 또한 기록의 형태와 행정적 절차에 관련한 기록관리 작업에 대한 지식이 필요하다. 세 번째 방법은 축적된 기록을 생산자의 정상적인 업무 중 외부로 보내져 보유되어 온 다른 기록과 비교하는 것이다. 이와 함께 보존자가 수행하는 도큐멘테이션 생성이 더 공개적이고 가시적인 과정이 되어야 한다. 또한 진본성을 위한 벤치마크 요건의 검증 과정이 추가되어야 한다.

기록의 정체성이나 무결성을 보장하고 그에 따라 그 진본성의 가정을 보장할 수 있도록 생산자가 실행한 통제를 평가 보고서로 기록화해야 한다.

일단 평가가 완료되면, 보존하기로 결정한 기록을 보존자에게 이관할 시점까지 그 기록을 주기적으로 모니터링 해야 한다. 만일 변경된 기록이 있다면 변경에 의해 기록의 본질과 속성, 가치와 진본성 또는 보존 가능성에 영향을 주지 않았음을 확인해야 한다.

이관 기록의 모니터링보다는 드물지만, 평가 또한 주기적으로 갱신할 필요가 있다.

보존자는 모든 필수적인 디지털 구성요소를 식별하고, 전체적인 구조를 이관 이전에 메타데이터 내에서 명확히 해야 한다.

보존을 위한 기술적 요건을 신중하게 검토하여, 보존의 타당성을 결정한다. 상이한 보존 전략은 비용에서 큰 차이가 있고, 매우 다른 결과를 초래할 수 있다. 보존 기관은 취득 비용과 수행하고자 하는 보존 방식에 대해 분명히 이해하고 보존 타당성을 결정해야 한다.

다. 기록의 획득

<표 3-13>은 기록의 획득과 관련하여 보존자에게 권고되는 사항을 정리한 것이다.

<표 3-13> 기록의 획득과 관련한 보존자 가이드라인

지침		세부 지침
3. 영구 보존하기로 결정한 기록의 획득	3.1 이관을 위한 공유된 계획 개발	
	3.2 표준 절차의 의무화	이관을 등록하는 절차 수립, 시행과 모니터링
		이관 권한 검증
		이관하기로 한 기록에 부합하는 기록인지 검사
		기록 이관
	3.3 이용할 수 있는 논리적 포맷 중 가장 오래된 것의 유지	
	3.4 복제 방지	
	3.5 모든 처리과정의 기록화	어떤 처리과정을 기록에 왜 적용 하였는가?
		어떤 기록을 처리하였는가?
		처리과정을 언제 수행하였는가?
		처리과정의 여러 단계를 수행하고 기록화한 담당자의 이름
		기록의 형태, 내용, 이용권한과 활용에 대하여 실행한 처리과정의 영향
		기록의 정체성과 무결성을 표현하는 요소에 영향을 미치는 모든 것을 포함하여 처리 과정의 결과로서 발생할 수 있는 다른 문제나 손실, 피해 등에 대한 기술

보존자의 기록획득 활동과 이어지는 모든 보존 활동의 목표는 지속적인 진본성과 가용성을 유지하는 것이다. 기록의 이관은 보존 사슬에서 가장 중요한 접점이며, 이관과정에서 잘못이 발생하지 않도록 보장하려면 상당한 주의가 필요하다.

현재의 기록보관자로부터 장기 보존의 책임을 지는 조직이나 프로그램으로 기록을 성공적으로 이전하려면, 양자 모두 합의한 계획이 전제가 되어야 한다. 만약 양자가 이관 과정에 합의할 수 없다면, 평가 결정을 재검토해야 할 것이다.

표준 절차의 의무화는 생산자로부터 보존자로 이관하는데 필요한 통제 사항이다. 생산자의 기록이 가진 진본성은 이관과정의 일부를 통해 반드시 검증되어야 한다. 이것은 기록의 정체성이나 무결성에 관한 메타데이터도 이관되고, 기록에 연결되며, 기록이 생산되고 유지되는 모든 기술 및 행정적 환경에 대한 관련문서도 첨부되었음을 검증하는 것 까지 포함한다.

보존자는 이관 이후 생성한 보존 및 참조를 위한 모든 기록과 함께 가능한 기록 생산 시점에서의 포맷이나 이관시점까지 생산자가 보존해온 논리적 포맷을 유지해야 한다. 만약 시간이 흐른 뒤, 선택된 보존 전략이 실패하는 경우, 초기의 논리적 포맷을 지속적으로 관리했다면, 보존자는 가장 권위 있는 사본을 가지고 다른 보존 전략을 적용함으로써 보존 과정을 다시 시작할 수 있다.

복제 방지는 디지털 기록이 중복하여 이전되지 않도록 보장하는 절차이다. 정확한 정체성 정보를 확보하는 것은 이 첫 단계에 해당한다.

디지털 기록의 모든 처리과정과 그 효과 등은 기록 되어야 한다.

라. 기록의 보존

기록의 보존과 관련하여 보존자에게 권고되는 사항은 <표 3-14>와 같다.

<표 3-14> 기록의 보존 관련 보존자 가이드라인

지침	세부 지침
4. 이관한 기록의 보존	4.1 기록 기술
	4.2 보존 행위의 법적 합의 확인
	4.3 선택한 보존 전략의 효과성 확증
	4.4 적절한 저장소의 유지

기록 보존자는 생산자의 기록을 물리적으로 적법하게 보관하고 보존(지속적 이용의 보장)하는 책임을 가진 실체이다. 보존자의 역할은 신뢰받는 보관자가 되며, 기록의 진본은 믿을 수 있는 보존 시스템 내에서 유지된다. 믿을 수 있는 보존 시스템은 시스템이 노화하여 기술이 업그레이드되더라도, 진본 기록을 지속적으로 생산할 수 있도록 하는 규칙이나 절차를 갖추어야 한다.

이를 위해서는 기록이 최초로 생산된 이래 생산자의 디지털 기록이 겪은 변화에 대한 정보를 기술해야 한다. 이 기술은 이관과 보존과정의 개요를 포함해야 한다.

보존 전략이 선택되면, 그 법적 합의를 검토해야 한다. 예를 들어, 독점적 환경에서의 포맷변환은 불법적인 행위를 수반할 수도 있다.

법적 합의를 검토한 후에는 선택한 보존 전략의 효과성을 확증해야 한다. 즉, 보존 전략이 기대되는 대로 수행될 수 있음을 보존하기 위해, 정식 이관 이전에 기록에 대해 선택한 보존전략을 실험해야 한다.

적절한 저장소의 유지는 자료 저장을 위해 적절하고 일관성 있는 저장 환경을 유지하는 것이다. 저장 환경은 지속적으로 모니터링 해야 하며, 수치(온도나 습도)등도 주기적으로 확인해야 한다.

마. 기록 출력

기록을 출력할 경우에 있어 보존자에게 권고되는 사항은 <표 3-15>와 같다.

<표 3-15> 기록 출력 관련 보존자 가이드라인

지침	세부 지침
5. 기록 출력	5.1 참조용 기록 생산 방식의 설명
	5.2 접근을 지원하는 기술적 요건 설명

기록의 지속적 접근(이용)의 확보는 기록과정의 필수적인 일부이다. 따라서, 평가, 취득, 이전, 기술, 저장등과 동일한 책임감과 기술적이고 전문적인 역량을 가진 보존자가 접근성 제공 업무를 담당해야 한다.

참조용 기록 생산 방식의 설명은, 생산자로부터 획득한 기록과 보존자가 생산한 기록 간의 관계를 분명히 기술하여, 이용자가 언제든지 이에 접근할 수 있게 하는 것이다. 재생산의 과정에서 기록물이 변화되지 않았음을 보증하기 위한 모니터링 절차의 기록도 여기 포함된다. 기록 재생산 과정과 그 영향의 기록화는 기록 재생산과정이 투명하다는 것을 실증하기 위한 필수적인 방법이다.

기록 보존자는 다양한 유형의 이용자에게 서비스를 제공할 수 있으며, 이용자는 다양한 기록 보존자의 서비스를 이용할 수 있다. 다양한 이용자들의 필요를 충족시키기 위해, 보존자는 추가적인 가공이나 특수한 보조 도구를 생산할 수도 있다.

2절. InterPARES에서 제시한 디지털 기록 유지와 보존 전략

1. A: 유지 전략²⁹⁾

유지 전략은 디지털 기록의 진본사본에 대한 접근성을 보호하고 유지하는 데 필요한 최소요건이다. 여덟 가지의 주요 유지전략을 제공하고 있다. 보존 전략을 적용하기에 충분할 만큼 기록을 오래 지속시킬 수 있도록 보장하기 위해서는 모든 기록 구성요소가 필요하다. 이러한 유지전략을 정리하면 <표 3-16>과 같다.

<표 3-16> 유지 전략의 내용

항목	항목명	정의	구체적 수행 방법
A1	분명한 책임부여	기록 저장이나 보호를 관리하는 데 대한 분명하게 정의한 책무 ³⁰⁾ 를 개인이나 부서에 부여.	- 책무 수행에 전담 가능한 능력이 있는 직원의 고용 - 기존 직원이나 부서의 업무시간 중일정량을 이 책무에 할당

29) 보존 사슬의 초기 단계에서 디지털 기록의 진본사본에 대한 접근 가능성을 보호하고 유지하기 위한 목적과 방법의 일관성 있는 세트

30) 책무는 구체적인 기술(skill) 세트, 전용 자원과 적절한 계획을 필요로 하는 기술적(technical) 책무를 의미한다.

항목	항목명	정의	구체적 수행 방법
A2	적절한 기술적 인프라 준비	기록관리나 유지 과정을 위한 물리적이고 행정적인 모든 자원(건물, 컴퓨터 하드웨어, 컴퓨터 네트워크 및 이들 설비를 유지하는 데 필요한 기타 직원 등)	
A3	시스템 유지, 지원 및 교체	H/W나 S/W의 유지, 업데이트와 교체 등을 위한 계획의 실행	
A4	주기적으로 데이터를 새로운 저장 매체로 이전	매체 노화의 영향을 방지하기 위해 데이터를 한 저장 매체로부터 다른 매체로 복제하기 위한 계획의 실행	체계적인 방법으로 이전을 실행
A5	저장 매체를 위한 적절한 조건의 준수	적절한 환경 조건을 준수함으로써 매체 훼손 정도를 크게 감소	- 온도 및 습도 조절 - 먼지 제거 등
A6	중복저장과 지리적 위치	디지털 객체의 다중 사본 제작 및 다중 매체·위치에 보관하여, 한 특정 저장 장소에서 발생하는 환경 위험 요소 및 천재지변으로부터 보호	디지털 객체의 다중 사본을 다수의 물리적 매체에 저장, 이를 다른 물리적 장소에 저장
A7	시스템 보안	오직 승인된 이용자와 과정에 대해서만 기록의 디지털 구성 요소가 노출되도록 통제	- 컴퓨터 자체 내에 있는 디지털 기록에 대한 접근 제한 - 컴퓨터가 위치하는 장소에 대한 물리적 접근을 제한 - 패스워드나 생체 인식 등을 이용한 시스템 로그인
A8	재난 계획	재난 이후에 손상된 시스템을 복구하고 기록관리나 보존 시스템을 효과적으로 복구하도록 안내하기 위한 자세한 절차의 수립	

2. B: 보존 전략³¹⁾

모든 기록 보존자는 유지전략에 부가하여, 하나 이상의 보존전략을 구현하기 위해 믿을 수 있는 보존시스템(trusted preservation system)³²⁾을 수립할 책임이 있다. 다음의 B는 12개의 보존전략을 제시하며, 이들 전략은 광범위하게 정의된 네 가지 그룹으로 구분되어 있다.

실제로는, A에 나오는 여덟 가지 유지전략에 더해 두 가지 이상의 보존전략을 지원하는 보존자가 대부분이다. <표 3-17>은 다양한 보존 전략과 그 방법을 정리하여 나타내고 있다.

31) 시간이 지난 후에도 디지털 구성 요소나 관련 정보를 유지하여, 진본 기록이나 기록집합을 재생산하기 위한 목적과 방법의 일관성 있는 세트

32) 보존자의 관리하는 기록을 지적, 물리적으로 영구히 유지하는 일을 관장하는 규칙 전체(그리고 이들 규칙을 수행하기 위해 이용되는 도구와 구조)를 포함한 시스템, 이러한 시스템은 시스템이 유지하는 기록이 진본이라고 추정할 수 있게 하는 상황적 개연성을 제공한다.

<표 3-17> 보존 전략의 내용

항목	항목명	정의	구체적 수행 방법
B1	표준의 이용	널리 이용 가능하고 지원되는 표준 (법률상 표준 ³³⁾ , 사실상의 표준 ³⁴⁾)을 인코딩 방법, 파일 포맷, 물리적인 저장 매체 등에 적용하여 장기적인 지원과 안정 가능성을 증가 시킴. 표준을 준수함으로써 애플리케이션을 간소화하고, 이후의 보존 전략의 효과성 최대화가 가능	- 보존자에게 디지털 기록을 이관하는 포맷을 제한하는 방식의 미래지향적 표준화 방식 - 다른 포맷으로 받은 파일을 표준 포맷으로 변환하는 소급적 방식
B1.1	자명한(self-describing) 포맷(영구객체보존, 태그부여(tagging))	구체적인 요소의 기능과 관계, 그리고 구조를 기술할 수 있도록 기록을 분석하여 태그를 부여.	
B1.2	인캡슐레이션(encapsulation)	기록과 기록에 대한 접근을 제공하는 수단을 결합	- 대체로 래퍼(wrapper)³⁵⁾로 결합 - 래퍼에 필요한 도구를 연결 - 기술하는 메타데이터를 포함
B1.3	관리 대상포맷의 범위를 제한(표준화)	기록을 저장하는 포맷의 수를 제한	- 새로운 독점적 포맷을 지속적으로 포함 - 기존 독점적 포맷의 새로운 세대를 수용 가능한 포맷으로 선택 - 비 독점적 포맷으로 제한
B1.4	변환 ³⁶⁾	한 H/W나 S/W 세대에서 다른 세대로 디지털 기록을 이전.	- 기술적 변화에 따라 하나의 논리 포맷에서 다른 포맷으로 영구적인 변환을 유도 - 열람요청이 있을 때 변환 - 접근시점에서의 변환
B2	기술 의존	기록을 변화시키지 않고 원래의 H/W와 S/W에 지속적으로 의존	
B2.1	기술 보존	디지털 기록을 표현하는 데 사용한 원래의 소프트웨어와 하드웨어를 유지	
B2.2	소급호환성(backward compatibility)	동일한 소프트웨어의 이전 버전으로 생산한 기록의 디지털 구성요소를 올바르게 해석하고 표현하는 소프트웨어의 성능을 신뢰하는 전략 ³⁷⁾	
B2.3	소프트웨어 재설계	기술의 변화에 따라 소프트웨어를 변환	- 새로운 플랫폼을 위한 소스코드를 새로 컴파일링 - 다른 프로그래밍 언어로 소프트웨어를 다시 코딩

33) 국가적(예를 들어 ANSI)이거나 다국적(예를 들어 CEN), 또는 국제적(예를 들어 ISO)인 공식 표준제정기구가 채택한 표준. 컴퓨터 파일 포맷에 대한 최근의 두 가지 법률상 표준으로는 PDF/A(보존을 위한 PDF 표준)와 ODF(OASIS Open Document Format)가 있다.

34) 공식 표준제정기구가 채택한 것은 아니지만, 그럼에도 불구하고 이용자들이 표준으로 널리 인정하여 사용하는 표준. 잘 알려지고 널리 이용되는 컴퓨터 파일 포맷 중 사실상 표준으로 간주되는 것은 PDF, TIFF, DOC, ZIP 등이 있다.

항목	항목명	정의	구체적 수행 방법
B2.4	접근시점에서의 뷰어와 전환	원래의 데이터를 사용하여, 한시적 접근성을 제공하는 소프트웨어 도구나 변환 방법을 필요에 따라 활용	
B2.5	에뮬레이션 (Emulation)	미래 기술을 보존된 디지털 기록의 원래 환경과 같이 행동하도록 만들어서 원래의, 또는 전환된 데이터로부터 원래 표현과 동일하게 기록을 표현	
B3	비디지털 접근방법 ³⁸⁾	디지털 객체 대신 아날로그 사본을 기반으로 보존	종이나 마이크로필름과 같은 안정적인 아날로그 매체로 디지털 기록을 복제
B4	데이터 복원	복구한 기록의 가독성을 복원하는 단계를 밟아 물리적 매체의 비트로부터 기록을 복구하는 전략	

35) XML과 같이 다양한 기술이 인식할 수 있는 방식으로 기록을 기술(describe)

36) 한 운반 매체에서 다른 운반 매체로 데이터 흐름을 복사하는 리프레쉬(refresh)와 구별

37) 보존용이 아니라 일시적 열람용으로 변환하는 데 그치면 표현이고, 소프트웨어의 현재 버전이 지원하는 포맷으로 기록을 영구히 변화시키는 것이 변환이다.

38) 소프트웨어가 제공하는 기능을 파괴한다는 제한점을 갖고 있다



1절. 테스트베드 운영

1. 테스트베드 개관

가. 테스트베드(기관)명

- 국토해양부

나. 테스트베드 배경

2008년 정부조직 개편으로 기존의 건설교통부와 해양수산부의 업무를 통합하여 신설된 국토해양부는 국토 및 도시공간의 설계와 교통물류망 구축, 해양영토의 체계적 관리를 담당하는 부처로써, 기존 건설교통부의 업무와 폐지된 해양수산부의 해양·정책·항만해운물류 업무를 통합하여 다음과 같은 업무를 담당한다.

- 국토종합계획의 수립·조정
- 국토 및 수자원의 보전·이용 및 개발
- 도시·도로 및 주택의 건설
- 해안·하천·항만 및 간척
- 육운·해운·철도 및 항공, 해양조사, 해양자원개발, 해양과학기술연구·개발 등

이상의 업무를 수행하는 국토해양부의 전신인 건설교통부와 해양수산부는 자료관시스템을 2004년부터 도입·운영해왔다. 2006년 기록관리법의 전면 개정과 더불어 새로운 전자문서생산 시스템으로서 업무관리시스템이 구축·확산됨에 따라 자료관시스템의 업무기반 관리 기능의 제고가 요구되었다. 이러한 요구에 따라 새로이 기록관리시스템이 개발되어 2007년 구축, 2008년까지 확산되었다.

「공공기록물 관리에 관한 법률」에는 전자기록생산시스템으로써 자료관시스템의 전자문서시스템과 업무관리시스템(온-나라시스템)의 혼용을 허용하고 있으나, '08년 11월 이후 통합온-나라 시스템의 사용으로 전자문서시스템의 사용이 중단되었다. 또한 상기 법률에 전자문서시스템의 기록물은 일정기간 경과 후 기록관리시스템으로 이관토록 규정하고 있어 자료관시스템에서 기록관리시스템으로의 기록물 이전이 요구되었다. 이에 국토해양부는 2008년 10월 기록관리시스템을 도입한 정부 및 공공기관 중 처음으로 전자기록의 이관을 수행하였다.

2. 테스트베드 운영 개요

가. 운영 개요

- 운영 주체 : 국토해양부 기록연구사(이강모)
- 운영 목적 : 구건설교통부 전자문서시스템의 생산 기록물을 기록관리시스템으로 이관하여 기록물의 일원적 관리 및 서비스를 제공하며 기록물의 통합 관리를 통한 기록물관리 업무의 효율성 확보
- 운영 기간 : 2008년 10월부터 2008년 12월까지
- 수행 업무 :
 - 구건설교통부 및 국토해양부의 전자문서시스템으로 생산한 기록물의 이관
 - 기록관리시스템으로 이관된 기록물의 검수 및 문서보존포맷변환
 - 전자기록물 이관파일 작성 및 이관, 데이터 검증

나. 시스템 현황

테스트베드 운용에는 자료관시스템의 전자문서시스템과 기록관리시스템이 사용되었다. 자료관시스템은 1999년, 문서처리 전 과정의 전자적 처리를 의무화한 「공공기관의 기록물관리에 관한 법률」이 시행됨에 따라 효율적인 기록물 전산관리 체계가 요구되어 구축된 시스템이다. 행정자치부 국가기록원에서는 2002년 대전광역시중구, 전남 여수시, 경기도 안산시를 대상으로 한 시범사업을 통해 자료관시스템의 표준 시스템을 구축하였으며 이를 바탕으로 전 행정기관 및 공공기관을 대상으로 자료관시스템을 확산하였다.

2006년 「공공기록물관리법」의 전면개정과 정부기능을 표준화하여 분류한 정부기능 분류시스템(BRM) 및 정책결정과정 및 업무처리 전 과정을 관리하는 업무관리시스템이 구축·확산되어 자료관시스템을 업무에 기반을 둔 기록관리시스템으로 재구축할 필요성이 제기되었다. 이에 따라 기록위주의 분류체계 및 결과만을 기록·관리하던 관행에서 벗어나 업무분류(BRM)에 기반을 둔 기록분류체계를 확립하고, 업무처리과정에서 생산된 기록물을 기록·관리할 수 있으며, 기록물의 공개범위를 확대하고, 검색 및 활용체계를 구축하여 기록의 업무활용 극대화 및 국민에게 다양한 기록정보 서비스를 제공할 수 있는 시스템으로써 기록관리시스템(RMS: Record Management System)이 개발되었다.

다. 테스트베드 대상 및 운영 방법

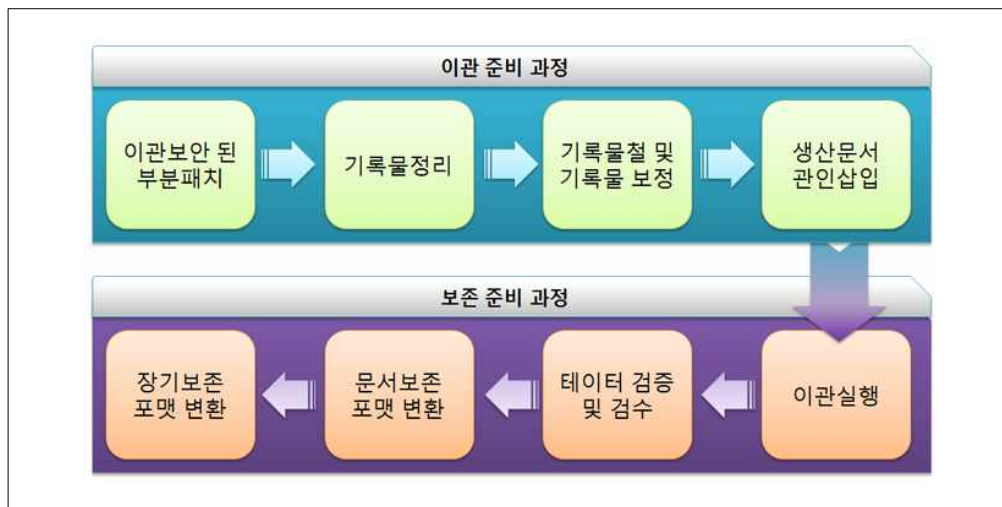
이관 대상 기록물은 다음과 같다.

- '04년 이후 생산된 구건설교통부의 전자문서시스템 생산 기록물
- '08년 이후 생산된 국토해양부의 전자문서시스템생산 기록물
- 첨부파일 포함 12테라바이트³⁹⁾

이관을 위해 수행된 과정은 다음과 같다.

- 이관기능개발(기능분석 및 설계, 기능개발 및 테스트)
- 기록물 이관 사전작업(이관 보완된 부분 패치, 이관대상 기록물정리)
- 기록물철 및 기록물보정(년도별 기록물 보정, 보정완료 후 검증)
- 생산문서 관인삽입(관인유무 확인, 본문 한글파일에 대한 관인삽입)

실제 수행될 이관 절차 및 이와 동일한 방법으로 수행된 테스트 절차는 <그림 4-1>과 같은 순서를 따랐다.



<그림 4-1> 테스트베드의 이관 절차

3. 이관 테스트 수행

가. 이관보안 된 부분패치

공공기관은 관련법에⁴⁰⁾ 따라 공공기관의 기록물을 처리과에서 보존기간의 기산일부터

39) TB, 10¹² Byte로 기가바이트(GB)의 1000배의 용량

40) 공공기록물 관리에 관한 법률(법률 8852호) 제19조제2항 및 공공기록물 관리에 관한 법률 시행령(대통령령 20789호) 제32조제1항

2년의 범위 내에서 보관한 후 기록물철 단위로 기록관리시스템에서 보관하도록 되어 있으나, 보관년도에 관계없이 이관할 수 있도록 보완하였다. 이는, 앞서 밝힌바와 같이 전자문서시스템이 업무관리시스템으로 통합될 예정이며, 그 시기가 도래하기 전에 기록물을 이관하여 업무의 효율성을 제고하기 위해 보존기간의 기산일부터 1년의 범위 내에 있는 “08년 이후 생산된 국토해양부의 전자문서시스템생산 기록물”이 이관 대상 기록물에 포함되었기 때문이다.

보완은 기존의 전자문서시스템에 변경된 스크립트를 적용하여 시스템의 기능적 제한을 한시적으로 해제하는 것으로 수행되었으며 보완된 사항에 대해 부분적으로 전자문서시스템에서 생산 후 2년이 지나지 않은 기록물에 대해서 패치를 적용하였다.

나. 기록물 정리

공공기관은 관련법에⁴¹⁾ 따라 매년 2월말까지 전년도에 생산을 완결한 기록물에 대하여 행정안전부령이 정하는 바에 따라 공개여부·접근권한 재분류, 분류·편철 확정 등을 하여야 한다. 본 테스트베드 운영에 있어서도 업무의 종결을 12월 31일로 보고 생산을 완결한 기록물에 대하여 더 이상의 편철을 할 수 없도록 마무리하는 기록물 정리를 수행하였다.

기록물 정리는 다음과 같은 업무이다.

- ☐ 기록물 편철 마감
- ☐ 비전자기록물 및 미등록된 비전자기록물의 등록
- ☐ 행정박물 등록
- ☐ 사진류 등의 등록

다. 기록물철 및 기록물 보정

기록물의 이관 이후에는 기록물의 수정이 불가하므로 기록물을 이관하기 위해 이관과 일을 생성하는 과정에서 발생한 기술적 오류와 수작업으로 인해 발생한 누락 정보를 바로잡는 기록물철 및 기록물 보정을 수행하였다.

기록물철 및 기록물 보정은 다음과 같은 업무이다.

- ☐ 특수문자, 띄어쓰기에 따른 데이터 수준의 오류 수정
- ☐ 페이지 번호 누락, 페이지 수 오류 수정
- ☐ 기타 수작업으로 발생한 오류 수정

41) 공공기록물 관리에 관한 법률(법률 8852호) 제18조 및 공공기록물 관리에 관한 법률 시행령(대통령령 20789호) 제24조제1항

라. 생산문서 관인삽입

전자문서시스템으로 생산된 문서는 관인이 날인되어야 그 효력이 발생한다. 그러나 현행 전자문서시스템은 시스템에서 관인이 날인되지 않은 채 인쇄물의 형태로 출력하여 실물에 관인을 날인함으로써 효력을 발생시킨다. 따라서 자료관시스템에 보관/존재하는 전자문서에는 관인이 날인되지 않으므로 효력이 상실된 문서로써 보관/존재하게 된다. 기록은 증거로써 보존 대상이 되며 신뢰성을 유지해야 하므로 행정 문서 및 증거로서 신뢰성을 유지하기 위해서는 전자문서에 관인이 날인된 상태로 보존매체로의 전환이 이루어져야 한다. 따라서 관인이 날인되지 않은 채 자료관시스템에 보관/존재하는 전자문서에 관인을 날인/삽입하는 작업이 수행되었다.

마. 이관 실행

이관은 네트워크를 통해 전자결재시스템에서 기록관리시스템으로 이관 파일을 직접 전송하는 방법으로 수행되었다. 이관 실행 절차는 국가기록원에서 제공되는 기록물이관 지침을 준수하여 수행되었다.

이관 테스트에서 이관하게 될 기록물은 구건설교통부의 1개 과에서 생산된 기록물로 해당과는 현재 폐지·통합되어 근무인원이 없기 때문에 해당 부서의 관계자만이 접근 권한을 갖는 전자문서시스템을 통해 기록의 관리가 불가하다. 따라서 가상의 근무인원을 생성하여 기록물 이관 실행을 수행하였다. 이관 실행은 시스템 내부에서 이관 기록물 목록 및 이관 파일 전·수송, 검증이 순차적으로 진행되었다.

바. 데이터 검증 및 검수

이관이 실행되는 동안 전자문서시스템에서 생성된 이관 파일은 기록관리시스템에서 데이터의 오류를 검증하여 그 결과를 반영할 수 있도록 한다. 기록관리시스템은 이관된 파일 중 오류가 검출된 파일을 선별하여 그 현황을 통보한다. 이관파일의 정보오류가 검출된 경우 전자문서시스템으로 반려하여 재이관을 요청하는 반려통보를 수행하고 오류가 검출되지 않은 경우 기록물의 인수를 수행하여 검수를 완료한다. 인수를 수행하여 데이터 검증 및 검수가 완료한다.

사. 문서보존포맷 변환

기록관 또는 특수기록관의 장은 관련법에⁴²⁾ 따라 인수가 종료된 전자기록물중 보존기

42) 공공기록물 관리에 관한 법률 시행령(대통령령 20789호) 제36조제1항

간이 10년 이상인 경우에는 중앙기록물관리기관의 장이 정하는 바에 따라 문서보존포맷 및 장기보존포맷으로 변환하여 관리하여야 한다.

문서보존포맷 및 장기보존포맷으로의 변환은 기록관리시스템에서 수행되며 첨부파일은 국가기록원이 정한 기준에 준하는 포맷에 한해 변환이 수행된다.

아. 장기보존포맷 변환

앞서 밝힌 관련법에 근거하여 장기보존포맷으로의 변환이 기록관리시스템에서 수행된다. 본 이관 테스트에서는 이관 대상 기록물이 장기보존포맷 변환 기준에 부합하지 않으므로 해당 절차는 생략되었다.

2절. 테스트베드 운영 결과

본 연구에서는 진본성에 주목하여 전자기록의 생산과 보존에 대해 연구하는 국제적 활동인 InterPARES의 다양한 연구 성과를 국내에 적용하고 나아가 이 국제적 활동에 기여하고자 하고 있다. InterPARES의 연구 성과는 국내의 상황이 고려되지 않았으나 다양한 국가의 사례와 연구 결과들을 종합한 것으로 어느 정도 일반화 되었다고 할 수 있다.

이러한 InterPARES의 연구 성과 중 두 번째 프로젝트를 통해 도출된 지침서는 기록의 생산자와 보존자의 측면에서 고려하고 검토되어야 할 권고사항들을 수록하고 있다. 또한 이 두 번째 프로젝트에서는 기록을 보존함에 있어 진본성을 유지할 수 있는 정책 프레임워크와 사본전자기록의 진본성을 추정할 수 있는 요건, 진본성을 유지하여 사본전자기록을 생산할 수 있는 요건을 제시하였다. 이러한 연구 성과를 바탕으로 InterPARES의 관점에서 기록관리 참여 구성요소는 생산자, 보존자(보존 업무 수행자), 시스템(생산, 보존)으로 압축할 수 있다. 본 연구는 이러한 InterPARES의 관점을 수용하여 테스트베드 운영 결과를 기술하고자 한다.

1. 기술적 측면

테스트베드의 운영에 있어 시스템, 즉 전자문서시스템과 기록관리시스템에 추가적으로 개선이 요구되어 처리된 사항은 다음과 같다.

가. 기록철 및 기록물 보정

본 테스트베드 수행 과정 중 하나인 ‘기록철 및 기록물 보정’은 전자문서시스템에서 기록물을 이관 가능한 상태로 전환하기 위해 요구된 사항이다. 기록물은 전자문서시스템에

서 바이너리 비트 형태로 존재하여 온전한 한글 파일(HWP 포맷 형태)로 전환하여야 한다. 이 과정에서 특수문자, 공백 등은 기계어로 처리되어 그 형태가 시각적으로 표현되는 것과는 상이하게 나타났다. 따라서 이를 시각적으로 표현되는 형태와 동일한 형태로 전환하는 과정이 요구되었으며 프로그램이 개발되어 일괄 처리되었다. 또한, 메타데이터 생성 시 총 페이지 수, 페이지 번호 등 수작업으로 입력되는 정보와 기술적 문제로 발생하는 문제를 수작업으로 확인하고 바로잡는 과정이 수행되었다.

나. 생산문서 관인삽입

앞서 밝힌 바와 같이 전자문서시스템에서 지원하지 않는, 전자문서 관인삽입 과정이 수행되었다. 전자문서시스템에서 이러한 기능이 지원되지 않음은 실제 행정 업무에 있어서 관인의 날인은 인쇄물 형태의 문서에 물리적으로 적용되어야 하기 때문이다. 그러나 전자문서의 증거로서의 효력과 신뢰성을 유지하기 위해서는 전자문서에 관인이 포함되어야 한다. 이러한 시스템 개선 사항이 이관 준비 과정에서 밝혀졌으며 프로그램이 개발되어 일괄 처리되었다.

다. 가상 근무자 설정

전자결재시스템에서 기록물의 접근은 해당 부서 근무자로 한정된다. 따라서 통합·폐지·분과되어 현재 존재하지 않는 부서는 이관 시점에서 근무자가 존재하지 않으므로 기록물의 접근이 불가하다. 기록물의 이관을 실행하는 것, 즉 이관파일 및 그 목록의 전송은 전자문서시스템에서 수행된다. 따라서 이관 대상을 선정하고 전송하기 위해서 기록물에 접근할 수 있는 권한을 가진 근무자가 요구된다.

본 테스트베드 운영 기간 중 테스트 대상, 즉 이관 대상 기록물은 현재 근무자가 존재하지 않는 부서로써 기록물에 접근할 수 있는 권한을 가진 근무자를 가상으로 설정해 주어야 했다. 가상으로 설정된 근무자의 이름은 해당 부서명으로 하여 혼란을 방지하였다.

이 밖에 개선이 요구되나 처리되지 못한 사항은 다음과 같다.

라. 생산자 이력 관리

기록물의 이관 후 생산의 책임이 있는 부서의 정보 변경에 따른 정보로 검색이 불가하여 개선이 요구되었다. 즉, 기록물의 생산에 책임이 있는 부서의 통합·폐지·분과 등에 따른 부서명 개칭이 발생할 경우, 개칭된 명칭으로는 기록물의 검색이 불가하였다. 이는 기록물의 생산자 정보가 기록물 정리, 즉 기록철의 종결 당시의 정보로 기술되기 때문이었다.

그러나 기록철의 종결은 1년을 주기로 수행되며 기록물의 이관은 이 보다 긴 시간 뒤에 수행될 수 있다. 또한, 기록물의 생산에 책임이 있는 부서는 통합·폐지·분과로 인해 부서명이 개칭된 후에도 이전 업무의 지속성을 유지해야한다. 이러한 부서의 통합·폐지·분과는 기록철이 종결되었으나 이관이 수행되지 않은 동안에 발생할 수 있다. 따라서 기록 생산 부서에 대한 정보와 함께 기록 관리 책임 부서 또는 기록 소유 부서에 대한 정보를 함께 기술할 필요가 있다.

2. 관리자(보존 업무 수행자) 측면

테스트베드 운영에 참여한 인원이, 테스트베드 운영의 준비에서 운영 완료에 이르는 동안 수행한 활동을 일대일 면담을 통해 조사하였다. 조사된 결과는 다음과 같다.

가. 이관 준비

앞서 밝힌 바와 같이 자료관시스템에서 기록관리시스템으로 기록물을 이전하는 첫 번째 시도였으므로 모범이 되는 안내서 또는 정보를 획득하기 어려운 상황이었다. 국가기록원에서 제공된 지침이 있었으나 구체적인 사례, 절차 등이 다소 부족하여 현장에서의 활용에 다소 어려움이 있었다. 또한, 테스트베드의 전체 운영을 계획 할 수 있는 인력이 연구사 1인뿐으로 다소 부족함이 있었다.

이에 테스트베드 운영 주체인 연구사는 테스트베드 운영을 준비하는 기간 동안 문제가 발생하였을 때 문제 발생 유형에 따라, 기술적 문제는 전자문서시스템 개발 업체와, 정책적 문제는 국가기록원과 협조하여 해결 방안을 도출하였다.

나. 투입 인력

테스트베드 운영에 투입된 인력은 기록연구사를 포함하여 국토해양부 근무자 3 명과 전자문서시스템 개발 업체에서 파견된 3 명이였다. 투입 인력은 추후 추가 할 것으로 계획하고 있다.

투입된 인력 중 기록연구사는 테스트베드 운영 준비 기간 동안 운영 계획 수립의 업무를 맡았으며 이후 실제 운영에 있어서는 관리·감독의 업무를 맡았다. 또한 발생하는 기술적, 정책적 문제를 관계 기관과의 협조를 통해 해결하였다. 기록연구사를 제외한 국토해양부 근무자는 기록연구사의 업무 수행을 보조하였으며 앞서 언급한 기록철 및 기록물 보정을 수행하였다.

전자문서시스템 개발 업체에서 파견된 3 명은 테스트베드 운영 준비 기간 동안 이관기능 개발, 이관 보완 부분 패치, 기록철 및 기록물 보정 및 관인삽입을 위한 프로그램 개발

등과 같은 시스템 관련 업무를 수행하였다.

다. 이관 실행

기록철 및 기록물 보정, 관인삽입을 마친 기록물을 대상으로 수행된 이관은 모두 시스템을 통해 이루어졌다. 전자문서시스템은 기록물을 이관 가능한 상태로 준비하고 전송하는 작업을 수행하며 기록관리시스템은 기록물의 인수 및 검수를 수행한다. 각 시스템에서의 수행 작업 및 절차는 다음과 같다.

- ① 인수일정 설정(기록관리시스템[이하 v2])
- ② 이관 대상 기록물 선택(자료관시스템[이하 v1])
- ③ 이관 실행(v1)

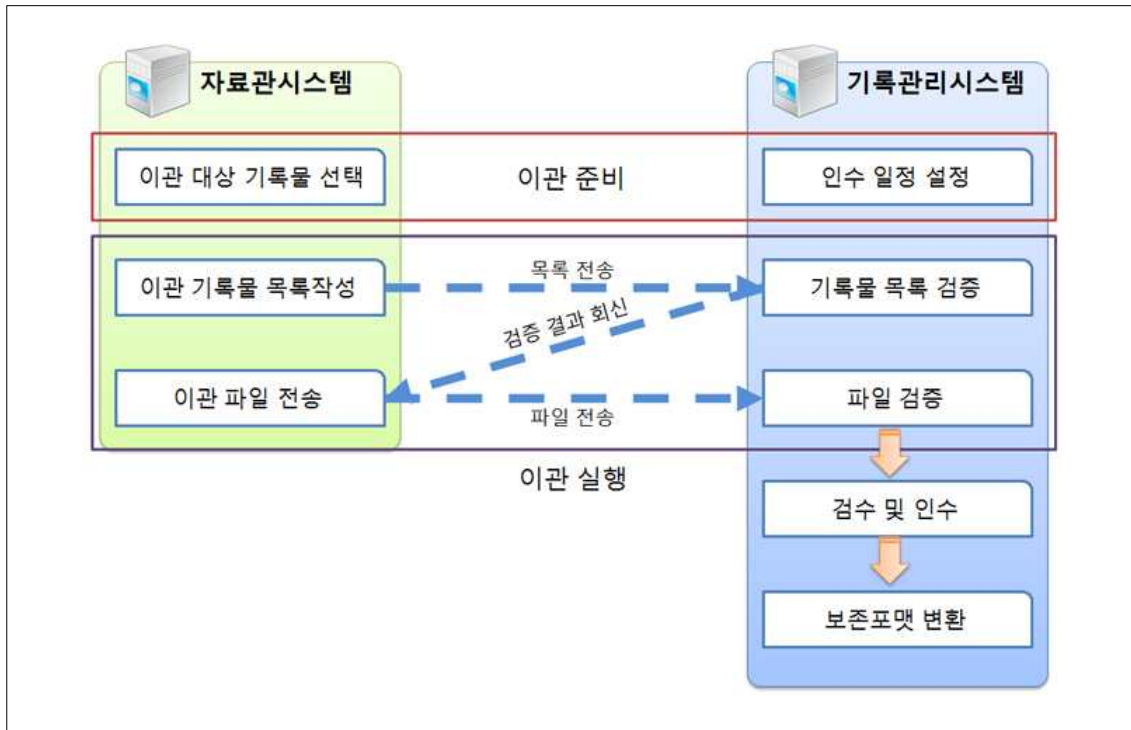
이관 기록물 목록작성 및 전송(v1) → 기록물 목록 검증(v2) → 파일 전송(v1) → 파일 검증 및 결과 회신(v2)

- ④ 검수 후 인수 처리 (RMS v2.0)
- ⑤ 보존포맷 변환 (RMS v2.0)

이를 정리하여 시각적으로 표현하면 <그림 4-2>와 같다. 파일의 검증은 시스템에서 데이터 오류를 검증하는 것으로 수행된다. 검수는 이 데이터 오류가 검출된 경우와 그렇지 않은 경우에 대하여 관리자가 반려 또는 인수하는 것으로 수행된다. 이러한 일련의 과정들은 모두 수작업으로 진행되었다. 이관 실행의 과정들이 수작업으로 진행된 이유는, 유관으로 각 과정의 수행 작업들을 확인하여 예기치 못한 문제에 능동적으로 대처할 수 있도록 하기 위함이다. 반려 수행에 있어 오류가 발생한 기록물만을 반려할 수 있도록 하고, 이때 발생한 오류의 구체적인 내용이 무엇인지를 파악할 수 있도록 시스템이 지원하여 발생한 문제에 대처할 수 있었다.

라. 이관 완료 확인

이관 완료의 확인은 기록연구사, 문서담당행정사무관, 전산담당자로 구성된 검수반이 보존포맷 변환을 확인하고 이관 시작 전 전자문서시스템에서 파악된 기록물의 용량과 기록관리시스템으로 이관된 용량을 비교 분석(PDF변환 용량을 포함)하는 것으로 확인한다. 이후 실제 이관에서는 전체 대상 부서 중 표본을 선정하여 이관이 완료되었음을 확인할 계획에 있다.



<그림 4-2> 시스템 상의 이관 절차

3. 생산자 측면

기록의 보존과 관리에 영향을 줄 수 있는 생산자들의 활동을, 기록의 생산과 보존의 두 측면을 모두 이해하고 있는 기록연구사를 대상으로 조사하였다. 조사된 결과는 다음과 같다.

가. 생산데이터

이관 및 보존의 대상이 되는 기록물을 생산 부서에서 전자적인 방법으로 생산한, 업무와 관련하여 발생한 모든 유형의 기록을 기본으로 한다. 비전자적인 방법 또는 그러한 형태의 기록물은 전자문서시스템에 등록하여 목록을 작성, 보존한다. 이러한 비전자기록물의 등록은 수작업으로 진행되며 생산자에게 등록의 책임이 있으므로 생산자가 비전자기록물을 등록하지 않은 경우 기록물이 보존 관리되지 못할 가능성이 존재함이 지적되었다. 비전자기록물의 등록은 공문 발송을 통한 등록 협조 요청과 유예 기간 중 등록 독려를 통해 기록물이 사장되는 것을 예방하고 있다.

나. 생산 지침

생산에 직접적인 영향을 미치는 지침은 사무관리규정⁴³⁾과 사무관리규정시행규칙⁴⁴⁾이

43) 대통령령 제20982호

있으며, 앞서 언급되었던 기록물관리법 또한 영향을 미친다. 상기 지침은 법률로써 그 구속력과 강제력을 가지나 증거확보의 어려움 등 다양한 이유로 실제로 그 구속력과 강제력이 행사되기는 어렵다. 현장에서는 생산에 영향을 미치는 지침의 준수에 대하여 생산의 책임이 있는 생산자의 윤리와 도덕성에 의존하고 있었다. 이와 더불어 지침 준수의 정도를 제고하기 위해, 기록관리의 총괄책임을 맡은 인원이 지속적으로 업무 협조 요청을 하고 지침 준수를 독려하는 노력을 기울이고 있었다.

다. 생산 문서 관리체계

생산된 전자문서는 시스템 측면에서는, 전자문서시스템과 온-나라시스템을 혼용하여 단위과제를 분류 기준으로 하여 관리되고 있었다. 전자문서시스템은 단계적으로 폐지될 예정이다. 전자결재시스템은 기록물분류기준표를 관리기준으로 하고, 업무관리시스템은 정부기능분류체계(BRM)에 따른 기록관리기준표를 관리기준으로 하여, 이 체계에 따라 기록물이 조직된다. 현재는 두 시스템이 모두 사용이 되고 있기 때문에 각각의 관리체계에 따라 기록물이 생산 관리되고 있다.

비전자 문서의 경우, 전자문서시스템이 이를 단위업무 별로 분류한 뒤, 기록물철로 등록하고, 철표지를 출력하여 편철한다. 이후 처리가 완결된 경우 해당 철을 보존상자에 단위업무별로 넣어 관리하고 있다.

인력 측면에서는, 각 부서에 기록물 담당자를 지정하여 기록물관리 제반 업무를 수행케 하고 이들을 기록관리 총괄자가 관리 감독함으로써 문서를 관리하고 있었다. 기록물 담당자는 해당 업무와 함께 유사 업무를 수행하며, 기록물 편철, 총괄 지시자의 업무지시 수행, 국가기록원 하달 업무 수행을 주요 기록물 관리 업무로 하고 있다.

라. 진본성 유지를 위한 노력

생산자 측면에서 기록의 진본성 유지를 위한 노력은 보안유지로 압축할 수 있다. 그 구체적인 방법은 다음과 같다.

- 기록물접근 제한: 전자기록생산시스템은 외부에 공개되지 않으며, 전자문서시스템의 문서는 해당 과에서만 열람 가능하도록 하며 온-나라시스템의 경우 열람제한의 기능이 있어 해당 기록물의 생산자/담당자만 확인이 가능하도록 하는 경우도 있다.
- 네트워크 차단: 전자기록생산시스템은 철저하게 인트라넷을 통해서만 활용할 수 있도록 하고 있으며 이 인트라넷은 방화벽으로 보호되고 있다. 기관 및 사용자는 행정전자서명의 인증을 통해 전자문서를 생산하도록 하여 그 책임과 권한을 분명히 하고 있다.

44) 행정안전부령 제32호

- 기록물 수정 제한: 시스템 상에서 한번 결재된 문서는 수정할 수 없도록 하고 있으나 부득이 수정이 필요한 경우 결재되기 전 회수하여, 수정 할 수 있도록 하고 있다. 전자문서시스템의 경우 이 기록은 남지 않으나 온-나라시스템의 경우 회수에 대한 정보를 기록한다.

4. 도출된 이슈

본 테스트베드의 운영 결과를 조망하였을 때, 기존 시스템에서 신규 시스템으로 업무와 관련된 데이터 이관을 계획하고 운영하는 데 있어 고려해야할 사항들은 다음과 같다.

가. 운영 계획 수립

본 테스트베드 운영에 있어서 가장 큰 어려움은 구체적인 운영 계획을 수립하는데 있었다. 이는 참조 가능한 모범 사례가 부재하여 발생 가능한 문제를 예상하기 어렵기 때문이었다. 이에 발생하는 문제에 대한 적절한 대처방안을 수립하기 어려웠고, 발생하는 모든 문제를 해결하며 테스트베드를 운영하였다.

또한 이러한 운영 계획을 수립함에 요구되는 지침 또는 안내가 구체적이지 못하여 현장에서 이러한 지침 또는 안내를 적용하는데 애로사항이 있었다. 이는 현장의 특성에 따라 지침 또는 안내를 활용할 수 있는 범위가 제한될 수 있다는 이유도 있지만 시스템의 기능 안내에 편중되어 전체적인 이관의 절차를 안내하지 못하였다고 짐작할 수 있었다.

나. 원활한 의사소통

테스트베드 운영 중 발생하는 문제를 해결하기 위해 다각적인 노력이 있었으며, 이러한 노력들은 관계자들의 협조를 통해 이루어졌다. 발생하는 문제들 중 시스템과 관련된 문제들은 대부분 처리가 가능하였다. 많은 문제들이 새로운 프로그램의 개발 또는 기존 시스템의 수정을 통해 해결되었으며, 일부 문제들은 단순한 시스템 운영 미숙 또는 원활하지 못한 의사소통으로 인한 착오로 발생한 문제들이었다.

그러나 이러한 문제 해결 과정을 통해 해결된 문제에 대하여 구체적인 문서가 작성되지 않았다. 발생한 문제가 구체적으로 어떠한 것이었으며 이를 어떠한 방법으로 해결하였는가를 구체적으로 작성한 문서가 요구된다. 이러한 문서는 추후 이관의 일반적인 지침을 작성하는데 중요한 자료로써 활용될 수 있을 것이다.

3절. 테스트베드 결과의 일반화

본 테스트베드의 운영 결과를 일반화하여 기존의 마이그레이션 절차 표준안과 유사한 형태로 재정리하였다. 이러한 재정리 작업에 있어 InterPARES의 생산자/보존자 가이드라인과 디지털 기록 유지와 보존 전략이 적용되었다.

1. 단계와 절차의 도출

본 테스트베드는 8개 절차를 순서대로 진행하여 운영되었다. 이 8개 절차는 테스트베드를 계획하고 운영한 기록연구사가 작성한 절차로 면담을 통해 사전 준비의 과정이 포함되지 않은 것으로 조사되었다. 이에 본 테스트베드 운영 과정 중 다각적으로 조사된 내용을 포함하여 본 테스트베드의 운영의 단계와 절차를 도출하였다.

앞선 8개 절차 중 이관을 실행하기 전의 4개 절차는 이관을 준비하는 과정으로 압축될 수 있으며, 도출되는 절차의 세부 내용으로 정리될 수 있다. 면담을 통해 조사된 사실 또한 절차의 세부 내용으로 정리될 수 있을 것이다.

이러한 세부 내용들을 포괄하는 절차와 이 절차를 포괄하는 단계를 범주화 하여 <표 4-1>과 같이 정리하였다.

<표 4-1> 단계별 절차와 세부내용

단계	절차	세부 내용
이관 준비	이관 계획 수립	기록원의 지침 확인
		관련 법규 확인
		이관 대상의 규모, 현황 파악
		업무 협조 방안 수립
		기술적 지원 확보
		이관 수행의 협력자 및 의사소통 채널 확보
		발생 예상되는 문제 검토
	이관 대상 정리	기록물 정리 협조 공문 발송
		행정박물 등록 협조 공문 발송
		이관 대상 기록물 선별
		기록물철 및 기록물 보정
		관인 삽입
		기록물의 이관 보완 부분 패치
		누락 기록 및 행정박물 등록
	이관 환경 설정	저장 공간 확보
		시스템 확보

단계	절차	세부 내용
		전송 네트워크 확보
		보정 프로그램 개발
		관인 삽입 프로그램 개발
이관 실행	시스템 조작	현존하지 않는 부서의 기록 접근을 위한 가상 근무자 설정
		이관 일정 설정
		자료관시스템 기록물 이관
		이관 목록 전송 및 확인
		이관 파일 전송 및 확인
		보존포맷 변환
	오류 확인 및 수정	오류 발생 확인 및 반려
		시스템 기능 오류 확인 및 기록
이관 평가	검수	데이터 검증 및 검수
		변환 검증 및 검수



1절. 업무이관에 따른 마이그레이션 개요

본장에서 다루고 있는 마이그레이션은 업무와 관련된 데이터를 기존 시스템에서 신규 시스템으로 이관하는 과정을 의미하며, 진행되는 과정을 단계별로 구분하여 그 기능과 역할을 규정하고 있다. 이러한 마이그레이션 절차는 국내 기록관리 업무를 진행하는 정부 중앙 부처의 담당 기록사의 업무로 집약될 수 있으며 업무의 단위 역할에 따른 내용을 절차와 하여 상세히 표현하여 기술하고자 하는 목적을 갖고 있다.

기존의 모델로부터 분석된 절차 모델은 국외의 사례로서, 일반적인 절차를 제시할 수 있으나, 국내의 기술적, 관리적, 생산자적 관점의 특성이 반영되어 있지 않다. 따라서 국내 사례와의 비교 분석을 통해, 국내 상황에 적합한 마이그레이션 절차를 도출하였다.

위와 같이 도출된 절차의 각 과정별, 관점별 고려해야 할 사항 및 요구되는 사항은 InterPARES 프로젝트에서 제시하고 있는 전자기록물 장기보존을 위한 요건을 통해 도출되었다.

마지막으로, 위의 과정을 통해 도출된 절차 모형에 대한 기록연구사의 평가와 의견을 종합하여, 실질적인 활용을 고려하였다.

2절. 마이그레이션 절차의 비교

1. 절차별 공통요소 도출

2장에서 살펴 본 표준안들을 비교 분석 및 종합하여 기본 마이그레이션 절차를 도출하였다. 도출된 마이그레이션 절차는 크게 준비, 실행, 평가의 3 단계로 나누어졌으며 <부록 1>과 같이 비교표를 작성하여, 국내의 특성은 반영되지 않은 1차 모형이 도출되었다.

ISO, AHDS, CCSDS의 마이그레이션 표준안은 그 표현에 다소 차이가 있었으나 그 기본 개념은 공통성을 가지고 있었고, 상세함에 있어서 다소 차이를 보였으나 그 하위의 절차 또한 공통성을 가지고 있었다. 이러한 공통성을 기준으로, 3개 표준안이 공통적으로 제시하는 단계와 절차를 선별하였다. 특히 절차의 선별에 있어서는 절차의 상세함을 우선적인 선별 기준으로 삼았다.

이러한 과정을 거쳐 도출, 재구성된 단계와 절차는 <표 5-1>과 같다.

<표 5-1> 표준 마이그레이션 단계와 하위 절차

단계	절차	참조 표준
준비	이관 대상 파악	ISO, AHDS, CCSDS
	이관에 필요한 자원(장비, 인력) 파악	ISO, AHDS, CCSDS
	구체적 계획 수립	ISO, AHDS, CCSDS
	문서화	ISO, AHDS, CCSDS
실행	사전 시험	ISO, CCSDS
	이관 준비	ISO, AHDS, CCSDS
	이관 수행	ISO, AHDS, CCSDS
	바이러스 검사	ISO, AHDS
	오류 확인 및 검사	ISO, AHDS, CCSDS
평가	검증 준비	ISO, AHDS, CCSDS
	검증 수행	ISO, AHDS, CCSDS
	검증 결과 관리	ISO, AHDS, CCSDS

2. 단계별 세부사항 도출

3개 표준안에서 제시하는 각 절차별 세부사항들 역시 그 표현에 다소 차이가 있었으나 의도하는 바는 공통성을 가지고 있었다. 각 표준안별로는 ISO가 평가 단계에 있어서, AHDS가 실행과 평가 단계에 있어서, CCSDS가 준비 단계에 있어서 각 절차에 대한 세부사항을 상세하게 제시하고 있었다. 본 장에서 각 표준안의 비교 분석을 수행하는 목적은 상세한 확인 항목을 도출하여, 앞서 진행된 테스트베드 운영 결과 및 InterPARES 프로젝트의 성과를 기준으로 선별하기 위한 기반을 마련하는 것으로 각 표준안의 상세한 정도를 우선으로 선별하였다. 3개 표준안에서 제시하는 각 단계별 절차에 대한 세부 사항들을 비교 분석하여 재구성하면 다음과 같다.

가. 준비 단계

3개 표준안에서 제시하는 준비 단계의 세부사항은 다소 차이를 보이고 있었으나 이관대상을 분명하게 파악하고 정의해야 함에는 모두 동의하고 있었다. 이들 절차에서는, 이관대상의 파악은 포맷 정보, 메타데이터와 같은 전자적이며 부가적인 정보를 정의하고 수집하는 과정과 내용의 정의, 변형에 대한 정보와 같은 맥락적 정보를 정의하고 수집하는 과정의 수행을 제시하였다. 또한, 전송 및 포맷 변환에 요구되는 부가적인 자원, 즉 네트워크, 표준, 저장 공간에 대한 사전 준비가 요구된다고 제시하고 있었다. 계획 수립과정

에 있어 ISO와 CCSDS의 표준안은 마이그레이션 수행 전 시험 절차를 제시하고 있었다. 3개 표준안을 비교 분석하여 도출한 세부사항은 <표 5-2>와 같다.

<표 5-2> 준비 단계의 세부사항

절차	세부사항	참조
이관 대상 파악	전자기록의 포맷 정보, 내용의 정의	ISO, CCSDS
	전자기록의 의미를 설명하기 위한 메타데이터 수집	ISO, AHDS, CCSDS
	전자기록의 원본에 가해진 모든 변형에 대한 정보와 책임자 정보 확보	ISO
	이관할 대상의 선별과 선별된 목록의 작성	AHDS, CCSDS
	기록의 소유권, 지적재산권에 대한 합의	ISO, CCSDS
이관에 필요한 자원 (장비, 인력) 파악	전송 방식 및 운송 수단의 정의	CCSDS
	준수해야할 표준의 확인	CCSDS
	전자기록의 임시 저장소 확보	ISO, AHDS, CCSDS
	비용, 위험 평가	CCSDS
구체적 계획 수립	사전 시험 절차 작성	ISO, CCSDS
	사전 시험 일정 수립	ISO, CCSDS
	실행 절차 작성	CCSDS
	실행 일정 수립	CCSDS
	우발계획 수립	CCSDS
문서화	정의된 이관 대상, 필요 자원과 수립된 계획의 승인 및 문서화	ISO, AHDS, CCSDS

나. 실행 단계

실행 단계에 있어서 각 표준안은 기록의 진본성 확보에 주력하고 있었다. ISO와 CCSDS의 표준안은 사전 시험을 통해 발생할 수 있는 위험 요소를 미리 파악할 수 있도록 하였고 AHDS는 마이그레이션 수행 후 발생하는 문제를 검출하여 처리하는 방법을 제시하고 있었다. 특히 AHDS는 변환 전 데이터의 무결성을 확보하기 위해 데이터, 메타데이터 수준의 상세한 검사 방법을 제시하고 있었다. 이에 ‘오류 확인 및 검사’ 절차는 AHDS의 상세한 검사 방법으로 보다 세분화였다. 실행 단계의 세부사항을 정리하면 <표 5-3>과 같다.

<표 5-3> 실행 단계의 세부사항

절차	세부사항	참조
사전 시험	시험 계획 수립	ISO
	시험 환경 설정	ISO
	시험 수행 결과 작성	ISO

절차	세부사항	참조
이관 준비	데이터 클리닝(공백, 특수문자의 적합성 확보) 시행	ISO, AHDS
	구조화된 텍스트 파일의 관련된 정의 및 구조 검사	AHDS
	변환 준비 과정에서 가해진 모든 변화를 메타데이터에 기록	ISO, AHDS
이관 수행	데이터 이전	AHDS
	메타데이터 업데이트	AHDS
	스캔 된 문서 저장위치 생성 및 메타데이터 생성	AHDS
바이러스 검사	바이러스가 발견되면 대체 파일을 요청	AHDS
	대체 파일이 없다면, 해당 파일을 제거	AHDS
	바이러스 발견에 대한 상세 정보를 메타데이터에 기록	AHDS
	모든 파일이 안전하거나, 위험한 파일이 제거될 때까지 반복	AHDS
미디어/파일 읽기 불가 오류 검사	읽을 수 없는 부분의 발견 시 대체물 요구	AHDS
	데이터 변조 발견 시 대체물 요구	AHDS
데이터 완결성 /무결성 검사	기존의 디렉토리 구조를 반영한 디렉토리 생성 및 해당 디렉토리의 파일 복사	AHDS
	각 디렉토리의 파일 수, 파일명, 파일 포맷을 목록과 비교하여 일치 여부 검사	AHDS
	목록에 존재하지 않는 파일 발견 시 정보 요청	AHDS
	누락된 파일 발견 시 해당파일 요청	AHDS
	누락 사항의 보완이 불가능한 경우 메타데이터에 기록	AHDS
메타데이터 전/후 비교를 통한 완결성 검사	전/후 메타데이터의 불일치 검출	AHDS
	심각한 불일치, 누락의 발견 시 보완을 요청	AHDS
	전자기록 포맷의 표준화와 검증 절차의 자동화 확보	AHDS
이관 관리	이관 요청서의 모든 항목이 작성되었는지 확인	AHDS

다. 평가 단계

평가 단계에 있어서 각 표준안은 기록의 무결성 확보에 주력하고 있었다. ‘검증 수행’ 절차에 있어 ISO와 CCSDS의 표준안은 구체적이며 상세한 세부사항이 부족하나, AHDS는 변환 품질을 확인하기 위해 파일, 파일 구조, 파일 내용/ 데이터 수준의 상세한 검사 방법을 제시하고 있었다. 이에 ‘검증 결과 평가’ 절차는 AHDS의 상세한 확인 항목으로 보다 세분화였다. 평가 단계의 세부사항을 정리하면 <표 5-4>와 같다.

<표 5-4> 평가 단계 세부사항

절차	세부사항	참조
검증 준비	검증 절차 및 책임자 확정	ISO
	검증 절차상 필요한 범위와 수행 방법 결정	ISO, AHDS

절차	세부사항	참조
파일 수준 일관성 검사	전자기록의 파일명과 실제 파일명의 일치 여부 확인	AHDS
	처리 불가능한 파일명 (공백, 특수문자, 확장자형식)은 적합한 것으로 수정, 변경된 파일명과 기존 파일명을 모두 메타데이터에 기록	AHDS
	전자기록의 파일포맷과 실제 파일포맷의 일치 여부 확인	AHDS
파일 구조 수준 일관성 검사	텍스트 파일의 문자 인코딩 적합성 검사	AHDS
	구조에 대한 형태적 설명이 존재하는 모든 텍스트 파일과 구조 간 적합성 검사	AHDS
	바이너리 파일의 경우, 생성 소프트웨어 및 버전을 통한 구조 및 포맷의 일관성 검증	AHDS
파일 내용/ 데이터 수준 일관성 검사	자원의 유형이나 파일 포맷에 따라 요구되는 데이터 형태와 실제 데이터 형태의 일치 여부 확인	AHDS
	비일관성 발견 시 문제 해결 시도	AHDS
	해결 불가 시, 추가적인 메타데이터이나 재이관을 요청	AHDS
	문제점 해결 불가 시, 메타데이터에 해당사항 기록	AHDS
검증 결과 관리	이용자에게 중요한 전자기록이 누락되었을 경우, 재이관 요청	AHDS
	아무리 사소한 변화라도 데이터 검증 과정에서 가해진 모든 변화를 메타데이터에 기록	AHDS
	일관성 검증 과정에서 발생한 모든 변화를 메타데이터에 기록	AHDS

3절 테스트베드 결과 및 InterPARES 성과물의 적용

이상으로 도출된 절차에 앞서 정리된 테스트베드 결과, 그리고 InterPARES의 성과물을 적용하여 국내의 실정을 반영함과 동시에 국제적 표준에 부합하는 마이그레이션 절차를 도출하였다. 이 마이그레이션 절차는 각 단계별로 인적 활동과 시스템에 요구되는 사항들을 나누어 정리하였다.

1. 마이그레이션 시 인적활동 요구사항

마이그레이션 시 수행자에게 요구되는 인적 활동 사항은 다음과 같다.

가. 이관 준비

(1) 이관 준비

이관을 실행하기 전, 모든 행정적 절차와 기록물 정리 작업이 우선되어야 한다. 행정적

절차를 통해 협력 부서의 협조를 확보하고 기록물 정리 작업을 수행할 수 있는 준비를 마칠 수 있어야 한다.

기록물이 이관된 뒤에는 수정이 불가하므로 사전에 보정작업과 관인 삽입 작업을 통해 기록물의 무결성과 신뢰성을 확보해야 하며 행정박물과 같은 비전자 기록물을 빠짐없이 등록하고 누락된 기록물이 없는 지를 점검해야 한다. 또한 이러한 과정에서 발생한 모든 변화를 메타데이터에 기록하여 진본성을 확보해야 한다.

(2) 이관 대상 파악

이관을 실행하기 전, 이관할 전자기록의 규모, 포맷 정보, 내용을 파악하는 것은 이관 계획을 수립하기 위해 반드시 요구된다. 이러한 이관 대상 파악에 있어 진본성과 이용성의 유지를 위해 메타데이터를 작성하고 그 책임자의 정보를 확보하는 것이 필요하다. 또한 소유권과 지적재산권에 대한 합의를 분명히 함으로써 이관 실행 시 부수적인 문제가 발생하지 않도록 해야 한다.

(3) 이관에 필요한 자원 파악

이관 실행에 요구되는 장비와 인력 등을 파악하는 것은 이관 실행에 있어 핵심이다. 하드웨어와 소프트웨어의 최신성 유지하고 전자기록의 전송 또는 운송 방식을 사전 합의하며 준수해야할 표준을 확인 하는 것은 적합한 이관 환경을 마련하기 위함이며, 이는 효율적이며 효과적인 이관 실행의 바탕이 된다.

또한 이관 실행에 요구되는 인력을 확보하는 것은 명백하게 요구되는 사항이다. 뿐만 아니라 이관 사전 테스트 및 이관 실행에서 예상치 못하게 발생할 수 있는 문제를 해결하기 위해서는 이관 실행에 관계된 인원의 협조 관계를 유지할 수 있는 의사소통 채널을 마련하는 것이 필요하다.

(4) 구체적 계획 수립

이관을 실행하기에 앞서 적절한 계획이 수립되어야 한다. 이 계획은 사전 시험과 실제 이관 실행의 구체적인 절차와 일정을 포함하여야 한다. 또한 예상치 못한 문제가 발생했을 때 어떻게 대처 할 것인가에 대한 계획, 업무 협조를 위한 계획, 기술적 지원을 확보하기 위한 계획 등 이관 실행에 요구되는 사항들의 전반적인 계획의 수립이 요구된다.

(5) 문서화

이상의 절차에서 파악되고 계획된 내용들은 적절한 형식의 문서로 만들어져 관리감독자의 승인을 얻어야 한다. 이는 업무 수행에 있어 기록의 생산이며 이후 진행될 업무, 즉 이관 실행의 근거와 지침이 된다.

나. 이관 실행

(1) 사전 시험

이관 실행 중 발생할 수 있는 문제를 모두 예측하기는 어렵다. 따라서 이러한 문제들을 사전 시험을 통해 파악할 필요가 있다. 이 사전 시험은 적절한 절차를 수립하여 수행되어야 하며 그 내용과 결과는 반드시 기록되어야 한다. 이 기록은 실제 이관 실행에 발생할 문제를 미연에 방지할 방안을 도출할 수 있는 자료일 뿐만 아니라 유사기관의 이관 운영 계획 수립에 중요한 참고자료가 될 것이다.

(2) 이관 수행

실제 이관 실행에서 그 수행은 대부분 시스템에서 진행되므로 이관 운영자는 어떠한 처리가 기록물에 적용되었는지 확인하고 발생할 수 있는 문제들을 점검할 수 있어야 한다. 또한 이러한 처리를 수행한 담당자의 정보를 확보하여 신뢰성과 무결성을 유지할 수 있어야 한다.

(3) 오류의 확인 및 처리

이관 중 발생하는 데이터 수준의 오류는 사람의 힘으로는 감당하기 어려우므로 오류의 발생은 시스템에서 확실히 감지 할 수 있도록 해야 한다. 또한 감지된 오류는 이관 운영자의 확인을 통해 예상치 못하게 발생하는 문제들을 처리할 수 있도록 해야 한다.

(4) 이관 관리

이관 수행을 마친 뒤, 이관 운영자는 이관이 준비된 계획을 잘 준수하였는지, 요구되는 사항을 모두 충족시켰는지 점검하여 기록물이 온전하게 보존될 수 있는 준비를 마쳤음을 확인하여야 한다.

다. 이관 평가

(1) 검증 준비

이관된 기록물을 평가하기 위한 절차를 수립하고 그 책임자를 분명히 하여 기록물의 진본성과 무결성을 확보해야 한다. 이관 대상의 수가 많지 않을 때는 이관 후 수작업으로 개별 기록물의 품질 검사를 하여 이관 평가를 수행할 수 있으나 그 수가 방대할 때는 개별 기록물의 품질 검사를 수작업으로 진행하기 어렵다. 따라서 샘플을 무작위로 추출하여 이관 전 기록물과 대조하는 품질 검사를 수행하는 것이 적절하다.

(2) 검증 수행

이관 대상의 수가 방대할 때는 이관된 기록물을 검증하는 작업은 대부분 데이터 수준에서 이루어지며, 따라서 시스템에 의존적일 수밖에 없다. 이러한 경우 시스템에서 검출되는 오류의 상세한 내용을 확인하고 처리하는 것이 우선되어야 한다. 시스템은 이 오류에 대한 적절한 해결 방안과 대안은 무엇인지 제시할 수 있어야 하며 그 결정은 이관 운영자 또는 검수자가 내려야 한다.

(3) 검증 결과 관리

계획된 절차에 따라 검증된 결과는 진본성의 유지를 위해 메타데이터로 기록되어야 하고 관리되어야 한다. 이관 운영자 또는 검수자는 이러한 메타데이터의 기록이 수행되었는지 확인하고 그 확인자는 누구인지 메타데이터에 함께 기록되어야 한다.

이상의 내용을 표로 정리하면 <표 5-5>와 같다.

<표 5-5> 마이그레이션 절차 및 세부사항

단계	절차	세부사항
이관 준비	이관 준비	기록물 정리 협조 공문 발송
		행정박물 등록 협조 공문 발송
		기록물 보정(문서의 무결성 확보)
		관인 삽입(문서의 신뢰성 확보)
		누락 기록 및 행정박물 등록
		데이터 클리닝(공백, 특수문자의 적합성 확보) 시행
		구조화된 텍스트 파일의 관련된 정의 및 구조 검사
		변환 준비 과정에서 가해진 모든 변화를 메타데이터에 기록
	이관 대상 파악	전자기록의 포맷 정보, 내용의 정의
		전자기록의 의미를 설명하기 위한 메타데이터 작성

단계	절차	세부사항
		전자기록의 원본에 가해진 모든 변형에 대한 정보와 책임자 정보 확보
		이관할 대상의 선별과 선별된 목록의 작성
		기록의 소유권, 지적재산권에 대한 합의
	이관에 필요한 자원 (장비, 인력) 파악	H/W나 S/W의 유지, 업데이트와 교체 등을 위한 계획 수립 및 수행
		전송 방식 및 운송 수단의 정의
		준수해야할 표준의 확인
		전자기록의 임시 저장소 확보
		이관을 위한 물리적이고 행정적인 자원을 확보
		책무를 개인이나 부서에 부여
		이관 수행의 협력자 및 의사소통 채널 확보
	구체적 계획 수립	사전 시험 절차 작성
		사전 시험 일정 수립
		실행 절차 작성
		실행 일정 수립
		우발계획 수립
		업무 협조 방안 수립
		기술적 지원 확보
	문서화	정의된 이관 대상, 필요 자원과 수립된 계획의 승인 및 문서화
이관 실행	사전 시험	시험 환경 설정
		시험 수행
		발생한 문제와 그 해결 방안(시험 수행 결과) 작성
	이관 수행	어떤 처리과정을 기록에 왜 적용하였는지를 기록
		처리과정을 수행하고 문서화한 담당자의 이름을 기록
		처리과정의 결과로서 발생할 수 있는 다른 문제나 손실, 피해 등을 기록
		데이터 이전
		메타데이터 업데이트
		스캔 된 문서 저장위치 생성 및 메타데이터 생성
	바이러스 검사	바이러스가 발견되면, 경고 후 대체 파일을 요청
		대체 파일이 없다면, 확인 후 해당 파일을 제거
		바이러스 발견에 대한 상세 정보를 메타데이터에 기록
		모든 파일이 안전하거나, 위험한 파일이 제거될 때까지 반복
		모든 파일이 안전하거나, 위험한 파일이 제거되었는지를 확인

단계	절차	세부사항
	미디어/파일 읽기 불가 오류 검사	읽을 수 없는 부분의 발견 시, 경고 후 대체물 요구
		데이터 변조 발견 시, 경고 후 대체물 요구
	데이터 완결성 /무결성 검사	기존의 디렉토리 구조를 반영한 디렉토리 생성 및 해당 디렉토리의 파일 복사
		각 디렉토리의 파일 수, 파일명, 파일 포맷을 목록과 비교하여 일치 여부 검사
		목록에 존재하지 않는 파일 발견 시, 경고 후 정보 요청
		누락된 파일 발견 시, 경고 후 해당파일 요청
		누락 사항의 보완이 불가능한 경우, 경고 후 메타데이터에 기록
	메타데이터 전/후 비교를 통한 완결성 검사	전/후 메타데이터의 불일치 검출 후 통보
		심각한 불일치, 누락의 발견 시, 경고 후 보완을 요청
		전자기록 포맷의 표준화와 검증 절차의 자동화 확보
	이관 관리	이관 요청서의 모든 항목이 작성되었는지 확인
이관 평가	검증 준비	검증 절차 및 책임자 확정
		검증 절차상 필요한 범위와 수행 방법 결정
	파일 수준 일관성 검사	전자기록의 파일명과 실제 파일명의 일치 여부 확인 후 통보
		처리 불가능한 파일명 (공백, 특수문자, 확장자형식)은 적합한 것으로 수정, 변경된 파일명과 기존 파일명을 모두 메타데이터에 기록
		전자기록의 파일포맷과 실제 파일포맷의 일치 여부 확인 후 통보
	파일 구조 수준 일관성 검사	텍스트 파일의 문자 인코딩 적합성 검사 후 통보
		구조에 대한 형태적 설명이 존재하는 모든 텍스트 파일과 구조 간 적합성 검사 후 통보
		바이너리 파일의 경우, 생성 소프트웨어 및 버전을 통한 구조 및 포맷의 일관성 검증 수행 후 통보
	파일 내용/ 데이터 수준 일관성 검사	자원의 유형이나 파일 포맷에 따라 요구되는 데이터 형태와 실제 데이터 형태의 일치 여부 확인 후 통보
		비일관성 발견 시, 경고
		해결 불가 시, 경고 후 추가적인 메타데이터이나 재이관을 요청
		문제점 해결 불가 시, 경고 후 메타데이터에 해당사항 기록
	검증 결과 관리	아무리 사소한 변화라도 데이터 검증 과정에서 가해진 모든 변화를 메타데이터에 기록
		일관성 검증 과정에서 발생한 모든 변화를 메타데이터에 기록

1절 InterPARES 3 프로젝트 수행

전자기록물 장기보존에 대한 국제적 연구 프로젝트인 InterPARES 3 프로젝트에 TEAM-Korea로 한국을 대표하여 본 연구진이 참여하는 동안 본 프로젝트와 관련하여 다양한 활동을 수행하였다. 그 중 핵심적 활동인 프로젝트 수행 내역을 기술하고자 한다.

1. InterPARES 3 프로젝트 기반 분석

InterPARES 3 프로젝트는 중소규모 기록관리 기관 및 프로그램을 실증적으로 조사하고 이를 종합하여, 전자기록물 장기보존에 실제적이며 실무적인 가이드라인을 작성하는 것을 목적으로 하고 있다. 본 연구진은 세 번째 프로젝트부터 참여하여, 본 프로젝트를 수행하며 연구에 기여하기 위해서는 그 이전 프로젝트의 성과물을 파악할 필요가 있었다. 이에 본 연구진은 InterPARES 1과 2 프로젝트의 내용과 생산된 문서를 분석하는 작업을 선행하였다. 이러한 선행 작업을 통해 InterPARES 프로젝트에 대한 통찰력을 얻을 수 있었으며 본 프로젝트인 InterPARES 3 프로젝트의 사례조사 연구를 수행할 기반을 마련하였다.

가. 국내 적용을 위한 질문지 분석

InterPARES 1과 2 프로젝트에서의 가장 큰 성과물은 생산자/보존자 가이드라인 및 진본확인을 위한 요건이었다. 그러나 이러한 요건들은 국내의 기록관리 실정과 괴리되는 부분이 있었으며 이를 국내에 적용하기 위해서는 이들 성과물이 생산될 수 있었던 기초를 파악하는 것이 필요했다.

InterPARES 2 프로젝트에서는 다양한 분야의 기록관 및 유사기관을 객관적으로 관찰 조사하여 이들 결과를 종합함으로써 성과물을 도출할 수 있었다. 이러한 관찰 및 조사에는 적절한 도구가 개발되었으며 이는 ‘23개 면담을 위한 질문지’로 압축할 수 있다. 그러나 이 질문지는 매우 개념적이고 포괄적이며, 면담 대상자에게 주어지는 질문으로 답변을 얻기 위한 것이었다. 이 23개 질문지에 대한 답변을 얻기 위해 실제 면담에서 면담자는 ‘48개 질문지’를 사용하여 면담을 진행하였다.

본 연구진은 이 23개 질문과 48개 질문지 모두를 분석하고 그 의미를 국내의 상황과 실정에서 재해석하였다. 또한 재해석 과정에서 발생할 수 있는 오류를 방지하기 위해 InterPARES 프로젝트의 director인 Luciano Duranti와 긴밀한 협력 관계를 유지하며 재

해석 결과에 대한 검증을 받았다. 이중 48개 질문지는 <부록 2>와 같이 번역되었다.

나. 사례조사를 위한 인터뷰 가이드 작성

InterPARES 3 프로젝트는 중소규모 기록관리 기관 및 프로그램을 실증적으로 조사하는 것을 통해 실제적이며 실무적인 기록물장기보존 가이드라인을 작성하는 것을 목표로 하고 있다. InterPARES 3 프로젝트는 사례조사 연구 중심의 프로젝트로 기록물, 정책, 기록관리시스템의 3개 분야에 대한 질문지가 작성되었다. 프로젝트 수행은 각 질문지를 기반으로 면담 조사를 통해 확보한 각 국의 중소규모 기록관리 기관 및 프로그램의 현황을 종합하여 실제적이고 실무적인 기록물 장기보존 가이드라인을 작성하는 것으로 진행된다. 본 프로젝트 수행을 위한 질문지는 이전 프로젝트의 성과물과 자료를 바탕으로 작성되었다. 이 질문지 또한 국내의 실정이 고려되지 않은 것으로 국내의 상황에 부합하는 질문지 작성이 요구되었다. 이에 본 연구진은 앞서 밝힌 바와 같이 이전 프로젝트의 성과물을 분석하는 작업을 선행하였고 이를 통해 본 프로젝트 수행을 위한 도구를 국내에 적용할 수 있도록 보완하였다.

이 보완은 국제 프로젝트의 성격을 잘 유지하면서 더 세부적인 답을 얻기 위한 상세 질문을 작성하는 것으로 진행되었다. 이러한 보완을 위해 23개 질문과 48개 질문이 분석 활용되었고 캐나다 McGill 대학에서 정보학대학원 기록관리학 교수로 재임 중인 박은경 교수의 자문을 받았다. 작성된 인터뷰 가이드는 <부록 3>과 같다.

본 연구진은 작성된 인터뷰 가이드를 기반으로 경북대학교 대학기록관, 대구광역시 기록관, 공인전자문서보관소 3개소를 중심으로 면담 조사를 진행하였다.

2. 경북대학교 대학기록관

IP3를 기반으로 정책, 기록물, 기록관리시스템 관점의 3가지 질문서를 작성하여 경북대학교 대학기록관을 방문하여 기록관리 담당자와의 인터뷰를 통해 경북대학교 대학기록관의 기록관리 정책, 기록물 관리 현황, 운영 중인 기록관리시스템 등에 대한 종합적인 조사를 진행하였다.

가. 경북대학교 대학기록관 개요

- 설립일자 : 2000년 8월 5일
- 소속 : 경북대학교 총장 직속기관
- 설립목적 : 경북대학교 관련 사료의 수집 · 보존 · 활용
- 주요업무

- 경북대학교 기록물의 수집·보존 및 활용
- 기록물 관리시스템 설치 및 교육
- 기록물 정리
- 기록물 분류기준표 취합 관리 및 조정
- 기록물 생산현황 취합 및 주요목록 전산화
- 활용대상 주요 기록물의 디지털자료 변환 및 검색, 열람제공
- 목표 : “공공기록물관리예관한법률”에 따른 기록물 관리, 경북대학교 교사자료 수집·보존·정리를 통해 대학 구성원 및 지역사회에 제공
- 조직구성 : 관장 1명, 사서 1명, 일용잡급 1명

나. 조사활동

- 조사기관 : 경북대학교 대학기록관
- 조사목적 : 국내 대학기록관의 Migration 사례 및 기록물 관리 현황 조사
- 조사일시 : 2008년 8월 ~ 9월
- 참고사이트 : <http://archive.knu.ac.kr>

다. 조사내용 및 주제

(1) 경북대학교 대학기록관에서 다루는 전자기록물

- 학교행정, 연구 및 교수활동 등을 통해 생성된 전자문서 본문 및 첨부파일, 연구결과물, 학위논문 등
- 전자기록물은 전자문서시스템상에서 관리되고 있으며, 생산된 전자문서시스템은 전자문서시스템 서버에 저장

(2) 경북대학교 대학기록관에서 운영중인 기록관리시스템 현황

- 송원시스템의 자료관시스템(AnyArchive) 사용
- 경북대학교 기록물 분류기준표를 전자문서시스템에 탑재하여 사용
- 국가기록원에서 인증한 자료관시스템 개발업체의 표준 메타데이터 스키마를 적용
- 자료관시스템 운영 및 유지관리와 관련해서, 기록관과 경북대학교 전산팀 그리고 자료관시스템 유지보수 업체와 협력체제를 구축하고 있음
- 자료관시스템은 기록관에 설치되어 있으며, 자료관시스템에는 19,221,614byte의 전자파일과 총장결재문서 일부를 PDF로 디지털화한 자료 존재
- 현재 보존기간 20년 이하 기록물의 철목록 DB를 구축중이며, 준영구 이상 기록물

DB 구축 계획을 수립중임

(3) 법률, 제도적 근거 및 현황

- 대부분의 기록관리 정책은 국가기록원에서 수립된 정책을 따르며, 대학의 특수한 부분인 교사자료의 경우 자체 정책을 수립함
- 경북대학교 대학기록관 운영규정 및 운영세칙을 근거로 기록관을 운영
- 기록 관련 정책은 홈페이지, 전화 또는 이메일, 개별면담 등을 통해 공지하며 결재 문서 공람 및 처리과 담당자 교육 등을 실시함
- 경북대학교 기록물 관리에 영향을 미치는 관련 법률은 공공기록물관리법, 정보공개법, 저작권법 등임
- 연구결과물, 비공개학위논문 등은 특히 저작권법과 관련성이 높음
- 개인 프라이버시 및 이권관련 기록물들이 주로 정보공개에 있어 문제가 됨

(4) Migration 관련 논점

- 이관 계획은 신기록관리시스템(RMS v2.0) 도입 후 이관예정임
- 경북대학교 대학기록관 운영을 위한 운영정책, 운영시스템, 대상자료를 조사 분석하기 위한 설문내용과 결과물을 <부록 4>⁴⁵⁾로 첨부.

3. 대구광역시 기록관

IP3를 기반으로 정책, 기록물, 기록관리시스템 관점의 3가지 질문서를 작성하여 대구광역시 기록관을 방문하여 기록관리 담당자와의 인터뷰를 통해 대구광역시 기록관의 기록관리 정책, 기록물 관리 현황, 운영 중인 기록관리시스템 등에 대한 종합적인 조사를 진행하였다.

가. 대구광역시 기록관 개요

- 소속 : 대구광역시 시민봉사과
- 구성 : 기록관장 1명, 정보공개 접수 및 배부 2명, 간행물관리 2명, 기록물관리 4명
- 주요업무
 - 기록물관리에 관한 기본계획 수립·시행
 - 소관기록물 수집·관리 및 활용
 - 기록물생산현황 취합 및 영구기록물관리기관으로의 통보

45) 경북대학교 대학기록관(정책, 기록관리시스템, 기록물)

- 시 및 소속기관의 기록물관리에 대한 지도·감독 및 지원
- 중요기록물의 재난대비 계획 수립
- 기록물관리시스템의 설치 및 운영
- 활용대상 주요 기록물의 디지털자료 변환 및 검색·열람 제공 외

나. 조사활동

- 조사기관 : 대구광역시 기록관
- 조사목적 : 대구광역시 기록관의 Migration 사례 및 기록물 관리 현황 조사
- 조사일시 : 2008년 8월 ~ 9월
- 참고사이트 : <http://www.daegu.go.kr>

다. 조사내용 및 주제

(1) 대구광역시 기록관에서 다루는 전자기록물

- 대구광역시 행정 및 민원업무를 통해 연간 100만 건 이상의 기록물이 생산되며 대부분은 전자기록물임
- 대구광역시 행정 및 민원업무 처리에 따른 결과, 사업 계획 및 시행 등 업무 절차에 대한 증거, 기관 내 업무 활용 및 정보공개에의 활용을 위해 전자기록물이 생산되고 있음
- 주로 Text 형태의 전자문서를 생산하고 있으며, 사진 및 비디오 등의 기록물도 생산됨
- 전자기록물은 hwp, doc, jpg, tiff, ppt, xls 등 다양한 형식으로 존재함
- 생산된 전자기록물은 전자문서시스템 서버에 저장되며 필요시에는 기록 생산자의 개인 데스크탑에 별도로 저장 가능

(2) 대구광역시 기록관에서 운영 중인 기록관리시스템 현황

- 송원시스템의 자료관시스템(AnyArchive) 사용
- 대구광역시에서는 업무관리시스템과 전자문서시스템을 통해 전자기록물이 생산됨
- 자료관시스템은 업무관리시스템과 전자문서시스템으로부터 전자기록물 관련 메타데이터를 상속 또는 획득함
- 전자기록물은 대구광역시 기록물 분류기준표에 의해 분류되어 관리되고 있으며, 대구광역시 기록물 분류기준표는 전자문서시스템과 자료관시스템에 탑재되어 사용되고 있음
- 국가기록원에서 인증한 자료관시스템 개발업체의 표준 메타데이터 스키마를 적용

하고 있음

- 자료관시스템에는 기록물 DB 구축 사업을 통해 생성된 기록물과 구전자기록물이 저장되어 있음

(3) 법률, 제도적 근거 및 현황

- 공공기록물관리법에 관한 법률을 준수하여 기록물을 관리
- 대구광역시 기록관 운영규칙을 근거로 기록관을 운영
- 각 부서의 기록물 담당자와 기록관 담당자들 사이에는 생산현황통보, 이관 등 기록 관리 전반에 대한 협력이 이루어지고 있음
- 전자문서시스템 및 업무관리시스템의 관리는 정보통신과에서 관리

(4) Migration 관련 논점

- 현재 이관 계획 없음
- 특별한 보존전략은 없으며, 공공기록물관리법에 따라서 보존하고 있음

○ 대구광역시 기록관 운영을 위한 운영정책, 운영시스템, 대상자료를 조사 분석하기 위한 설문내용과 결과물을 <부록 5>⁴⁶⁾로 첨부함.

4. 공인전자문서보관소

IP3를 기반으로 3가지 관점의 질문서를 작성하여 공인전자문서보관소(이하 공전소)를 운영을 위한 정책을 마련하고 공전소를 평가하는 등의 방향을 제시하는 한국전자거래진흥보관소원과 공인전자문서보관소로 지정을 받아 실제 전자문서 보관을 주요 업무로 하고 있는 삼성SDS와 LG-CNS를 방문하여 담당부서의 실무자들과 인터뷰를 진행하였고, 진행된 결과를 설문지를 통해 의견을 작성하여 전자기록물에 대한 운영정책, 대상기록물, 운영시스템에 대한 다양한 의견을 종합적으로 조사를 진행하였다.

가. 공전소 개요 :

- 공전소 운영목적 : 전자문서를 보존하는 동안 위변조 방지와 이를 위한 인증 서비스 제공
- 지정요건/평가지침 6개월간 심사와 지정을 위한 기간 소요됨
- 평가지침에 대한 세부 내용 분석 필요

나. 조사 활동

46) 대구광역시 기록관(정책, 기록관리시스템, 기록물)

- 조사기관 : 한국전자거래진흥원, 삼성 SDS, LG-CNS
- 조사목적 : 국내 아카이브의 Migration 사례 및 공전소 관리 운영 현황 조사
- 조사일시 : 2008년 7월 ~ 8월
- 참고사이트 : www.ceda.or.kr

다. 조사내용 및 주제

(1) 공전소에서 다루는 전자문서관련 대상자료

- 텍스트형태의 문서, 스캔한 이미지 문서, 첨부 문서, DB
- 대부분이 문서를 스캔 후 이미지 형식으로 전환하여 보관
- 텍스트, 이미지 이외는 Binary 로 규정하고 저장(Circuit Switched Data Format)
- 현재 공전소 대상 문서는 의뢰사 내 계약문서 또는 병원과 같은 기관의 임상자료
- 종이문서의 비용을 줄이는 차원에서 전자문서를 대상으로 인증하기 위한 목적
- 장기보존을 위한 형식은 PDF-A, TIFF 등 을 활용

(2) 공전소 운영 시스템 현황

- 전자거래진흥원의 관리 규격에 따른 기능을 통해 시스템을 구성
- 필수적인 기능 요구사항에 따른 내부적인 일부 보완기능 운영

(3) 법률, 제도적 근거 및 현황

- 전자거래진흥원이 공인지정절차를 거쳐, 서류 및 기술심사, 심의 후 결정
- 최소12명의 시스템 운영인원과 각종 설비에 대한 지침을 준수 할 수 제도적 장치 마련
- 현재까지 3개(SDS, LG, KT-NET) 지정
- 기술규격 사항으로써 전자문서 정보패키지, 전자문서 증명서 포맷 및 운용절차, 이용자 시스템, 이용자시스템과 공인전자문서보관소 간 연계 인터페이스 등을 다룸

(4) Migration 관련 논점

- Migration을 데이터의 현행화 관점으로 정의하여 논의
- 현재까지 Migration 사례는 공인전자문서보관소에서 이루어 지지 않음
- 2007년 처음 지정후 아직 까지 요청기관의 대상 자료의 다양성에 대한 사례를 충분히 갖고 있지 않고 있는 상황

(5) 향후 전자거래 진흥원 관리 및 기술 규격검토에 대한 구체적인 조사와 분석이 이루어질 필요성 제시

- 기술규격 → 패키지 규격 → 보존 포맷
- Clients ↔ 공전소
 - ↳ 문서등록서비스(문서이관)
 - ↳ 과정에 대한 규정
 - ↳ 기술규격 & 해설서(평가)
 - : 패키지/
 - : 증명서/
 - : 연계인터페이스/
 - : 이용자시스템기술규격/
 - : 이·수관(공전소간에 전자문서 이관 시)
- ISO 14721 을 일부 차용한 시스템 구성
- 이용자 시스템과 공전소 시스템간의 이관 시 문제점은 없다고 판단
 - 이를 위해 솔루션 업체에서 기술지원을 통한 전체 Legacy file을 server-client 로 묶어서 운영
 - Client 는 공전소와 연계 또는 공전소 배포 프로그램을 활용
- 공인전자문서보관소 운영을 위한 운영정책, 운영시스템, 대상자료를 조사 분석하기 위한 설문내용과 결과물을 <부록 6>⁴⁷⁾으로 첨부

2절 국제 활동

1. ICA 총회 참석

전 세계 기록의 보존과 사용을 증진하는 것을 목표로 삼고 있는 국제기록기구회의(International Council on Archives 이하 ICA)는 196개국이 참여하고 있으며 우리나라는 동아시아지역협의회(EASTICA)에 소속되어 활동하고 있다.

ICA는 기록의 보호와 가치의 제고를 위해 노력하고 있으며 다양한 문화를 가진 여러 나라들의 기록관리 분야 관계자들의 커뮤니케이션을 증진 시키는 매개자의 역할을 하고 있다. 이러한 역할에 대한 활동 중 하나로써 ICA는 4년을 주기로 총회를 열고 다양한 학술 발표와 토론회 및 회의를 개최한다. 올해로 16회를 맞는 ICA 총회에 본 연구진이 InterPARES TEAM-Korea로써 참여하였다.

7월 21일부터 27까지 말레이시아 쿠알라 룸푸르에서 열린 본 총회에서 본 연구진의 책임연구원인 성균관대학교의 오삼균 교수와 보조연구원 1인이 참석하였다. 참석 목적은

47) 한국전자거래진흥원(정책, 시스템), 삼성 SDS(정책, 시스템, 대상 기록물), LG CNS(정책, 시스템, 대상 기록물)

“InterPARES 3 in Asia: Implementing Research Findings on Digital Records Preservation”라는 주제로 22일에 발표를 하기 위함이었으며 더불어 국제적인 기록관리 현황과 최신 기술의 동향을 파악하기 위해서였다. 22일 발표는 오삼균 교수가 발표자로서 연단에 섰으며 본 연구의 목적과 내용을 소개하고 국가기록원과 InterPARES 간에 긴밀한 협력관계를 유지하며 연구를 진행하고 있음을 소개하였다. 또한 국가기록원의 기록관리시스템을 간략하게 소개하여 관계자들로부터 많은 관심을 받았다. 발표 내용은 <부록 7>로 첨부한다.

2. InterPARES 총회

2008년 10월 1일부터 멕시코에서 열린 InterPARES 총회에 본 연구진의 공동연구원인 경북대학교 김희섭 교수와 본 연구진의 자문위원인 박은경 교수가 참석하였다. 본 총회에서는 각국의 연구진행 상황을 점검하고 그들의 활동과 구체적인 연구 내용에 대해 토의하였다. 또한 InterPARES의 운영 전반에 걸친 내용에 대해서도 논의가 이루어졌다.

이번 총회에서 스위스, 뉴질랜드, 일본, 슬로바니아의 가입에 대한 논의가 진행되었을 뿐만 아니라 차기회의 장소와 시기가 논의되어 2009년 6월 한국에서 개최하기로 결정되었다. 이는 국가기록원의 최종 결정이 남아있으며, 현재 서울로 잠정 선택하였으나 미확정이다. 프로그램은 마지막날(금요일) 심포지움을 개최하며 각 팀에서 한편씩의 논문 발표, 각 팀에서 진행하고 있는 General Studies에 대하여 연구논문으로 작성하여 발표하는 것으로 제안하였다(예: 한국팀-마이그레이션, 이탈리아팀-전자메일, 싱가포르-웹아카이브 관련 등). 심포지움 발표논문은 프로시딩으로 출판하는 것도 검토하고 있으며 각 국가에서는 발표제목을 12월까지 제출하도록 하였다. 한국 내에서 뿐만 아니라 가까운 아시아권 관련 분야 종사자들의 참석을 유도해야 할 것으로 판단된다.

2009년 한국 InterPARES 회의와 관련하여 가장 시급한 것은 접근이 쉬우며 숙소가 적당한 장소를 결정하는 것이다. 또한 웹 컨퍼런스, 인터넷 가능한 곳이어야 한다는 의견도 나왔다. 본 총회를 다녀와 작성된 발표 형식의 보고서를 <부록 8>로 첨부한다.

제4장 총괄연구개발과제의 연구결과 고찰 및 결론

본 연구의 수행 목적은, 첫째, InterPARES 3 프로젝트로부터 얻은 연구결과를 국내에 적용하기 위한 실효성 및 타당성을 국내 전자기록물 생산 보유한 기관의 컬렉션을 테스트베드로 선정하여 이를 대상으로 검증하고, 이 검증된 결과를 통해 국내의 기록관리 및 보존환경에 맞는 이슈와 해법을 제시하는 것이다.

두 번째 목적은, 디지털 자원을 생산하고 관리하는 각급의 다양한 디지털 자원 관리기관 실무자 및 한국의 기록관리 공동체를 대상으로 연구결과를 확산함으로써 국내 전문가 공동체의 집단적 역량 강화가 이루어질 수 있는 기반을 마련하는 것이다.

마지막으로, 한국의 기록생산기관을 테스트베드로 하여 TEAM-Korea 연구진이 얻은 연구 성과를 InterPARES 프로젝트를 통해 국제사회와 공유 하도록 하는 것이다.

이에, 본 연구를 수행한 결과와 활용방안은 다음과 같다.

첫째, 디지털 자원 장기보존 전략으로서 마이그레이션의 전체 실행 프로세스 상 발견되는 실제적인 문제들을 문헌과 테스트베드 운영을 통해 수집 분석하였고, 그 결과를 바탕으로 장기보존 전략의 선정, 적용, 평가를 위한 전반적인 지표를 마련하였다.

둘째, 중앙부처의 전자기록물 이관율, 테스트베드로 선정한 디지털 자원 관리기관의 유형, 대상자료, 업무상황, 운영 시스템 등을 고려하여 이에 부합하는 표준화된 이관절차를 제시하였으며, 이를 통해 전자기록물의 업무이관, 데이터 이관, 시스템이관에 따른 전략적인 방안을 도출하였다.

셋째, 전자기록물 보존을 위한 전략적 방안에 따른 결과들에서 도출된 지표와 적용방법론, 원칙을 유사한 유형의 다른 디지털 자원 관리기관에서의 진본 디지털 자원 보존 활동을 지원하는데 활용될 것으로 예상한다.

넷째, 국내의 공공/민간 기록 관리기관의 조직과 프로그램이 다양한 이해관계자와 사회적 요구를 충족시킬 수 있는 진본 전자 기록(authentic records)을 장기간 보존(long-term preservation)하기 위한 지식기반을 구축하고 실행전략을 위한 가이드라인을 제시하였다.

다섯째, 한국의 기록생산기관을 테스트베드로 하여 TEAM-Korea 연구진이 얻은 연구 성과를 전자기록물 장기보존 분야의 국제 공동연구프로젝트(InterPARES 3)와 공유하였다.

제5장 총괄연구개발과제의 연구성과

5.1 활용성과

총괄과제명	전자기록물 장기보존 전략 및 테스트베드 구축
총괄과제책임자	오삼균 / 성균관대학교 / 문헌정보학과

가. 연구논문

번호	논문제목	저자명	저널명	집(권)	페이지	Impact factor	국내/국외	SCI여부
1								
2								

나. 학술발표

번호	발표제목	발표형태	발표자	학회명	연월일	발표지	국내/국제
1	IP3 case study in Korea	Panel Session	Sam Oh	ICA	'08,07, 21~27	Kualalumpur, Malaysia	국제
2							

다. 지적재산권

번호	출원/등록	특허명	출원(등록)인	출원(등록)국	출원(등록)번호	IPC분류
1						
2						

라. 정책활용

1. 소규모 디지털 아카이브의 장기보존전략 방안
2. 전자 기록물 보존 전략 수립에 대한 타당성 평가

마. 타연구/차기연구에 활용

국내 아카이브 사례에 대한 연구를 IP3에 기여함으로써 migration 전략 방안 도출

바. 언론홍보 및 대국민교육

해당 없음

사. 기타

해당 없음

5.2 활용계획

제6장 참고문헌

- 이미화, 현문수. 『전자기록의 유형별 보존 기법』 한국국가기록연구원, 2006.
- 서혜란. 『기록관리의 인적·물적자원』. 진리탐구(주), 2004.
- 정혜경. 『디지털 아카이빙의 경제성 분석』. 한국학술정보(주), 2005.
- 정혜경. “디지털 보존의 비용요소에 관한 연구”, 정보관리학회지, 22(1): 47-64, 2005.
- Kevin Ashley. Digital Archive Costs: Facts and Fallacies. Paper at the DLM Forum 1999. http://europa.eu.int/ISPO/dlm/program/abst_ka_en.html
- Stewart Granger, Kelly Russell, and Ellis Weinberger. Cost Elements of Digital Preservation (Version 4), October 2000.
<http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc>
- Shelby Sanett. Toward Developing a Framework of Cost Elements for Preserving Authentic Electronic Records into Perpetuity. *College&ResearchLibraries*, 63(5):388-404.
- Arturo Crespo, Hector Garcia-Molina. Cost-Driven Design for Archival Repositories. Joint Conference on Digital Libraries 2001 (JCDL '01), June 24-28, 2001, Roanoke, Virginia, USA. <http://www-db.stanford.edu/~crespo/publications/cost.pdf>
- Tony Hendley. Comparison of Costs and Methods of Digital Preservation. Chapter 5. Developing a Cost Model. British Library Research and Innovation Report 106, 1998.
<http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html>
- Anne R. Kenney. The Cornell Experience: ArXiv.org. DCC/DPC Workshop on Cost Models for Preserving Digital Assets 26 July 2005.
- Nationaal Archief. Costs of Digital Preservation. The Hague, 2005.
<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/CoDPv1.pdf>
http://dlmforum.typepad.com/Paper_RemcoVerdegem_and_JS_CostModelfordigitalpreservation.pdf
- CEDARS. Cedars Guide to Digital Collection Management. The Cedars Project.
<http://www.leeds.ac.uk/cedars/guideto/collmanagement/>
- Richard Davies, et al. “How much does it cost? The LIFE Project - Costing Models for Digital Curation and Preservation,” *LIBERQUARTERLY*, 17(3/4), 2007.

총괄 연구과제 요약

과제 고유번호	자동부여		공개가능여부	
사업명	2008 국가기록원 연구개발			
과제명	전자기록물 장기보존 전략 및 테스트베드 구축			
연구책임자	성 명	오삼균		
	소속 기관명	성균관대학교		
	전자우편	XXXXXXXXXX	전화번호	XXXXXX

○ 연구목표 (400 ~ 600자)

본 연구는 다음의 목표로 진행되었다. 첫째, InterPARES 3 프로젝트의 연구 성과를 국내에 적용함에 있어, 테스트베드를 선정하여 이를 대상으로 타당성과 실효성을 검증하고, 이 검증된 결과를 통해 국내의 기록관리 및 보존환경에 맞는 이슈와 해법을 제시하는 것이다. 두 번째는, 실무자 및 기록관리 공동체를 대상으로 연구결과를 확산함으로써 국내 전문가 공동체의 집단적 역량 강화가 이루어질 수 있는 기반을 마련하는 것이다. 마지막으로, 한국의 기록생산기관을 테스트베드로 하여 TEAM-Korea 연구진이 얻은 연구 성과를 InterPARES 프로젝트를 통해 국제 사회와 공유 하도록 하는 것이다.

○ 연구내용 (1000 ~ 1200자)

디지털 자원 장기보존 전략으로서 마이그레이션의 전체 실행 프로세스 상 발견되는 실제적인 문제들을 문헌과 테스트베드 운영을 통해 수집 분석하였고, 그 결과를 바탕으로 장기보존 전략의 선정, 적용, 평가를 위한 전반적인 지표를 마련하였다. 또한, 중앙부처의 전자기록물 이관을, 테스트베드로 선정한 디지털 자원 관리기관의 유형, 대상자료, 업무상황, 운영 시스템 등을 고려하여 이에 부합하는 표준화된 이관절차를 제시하였으며, 이를 통해 전자기록물의 업무이관, 데이터 이관, 시스템이관에 따른 전략적인 방안을 도출하였다. 국내의 공공/민간 기록 관리기관의 조직과 프로그램이 다양한 이해관계자와 사회적 요구를 충족시킬 수 있는 진본 전자 기록(authentic records)을 장기간 보존(long-term preservation)하기 위한 지식기반을 구축하고 실행전략을 위한 가이드라인을 제시하였다. 한국의 기록생산기관을 테스트베드로 하여 TEAM-Korea 연구진이 얻은 연구 성과를 전자기록물 장기보존 분야의 국제 공동연구프로젝트(InterPARES 3)와 공유 하였다.

○ 연구성과(응용분야 및 활용범위포함) (400 ~ 600자)

전자기록물 보존을 위한 전략적 방안에 따른 결과들에서 도출된 지표와 적용방법론, 원칙을 유사한 유형의 다른 디지털 자원 관리기관에서의 진본 디지털 자원 보존 활동을 지원하는데 활용될 것으로 예상된다.

○ 총괄 참여연구원

성명	소속/직위	성명	소속/직위
오삼균	성균관대학교 문헌정보학과/교수	김희섭	경북대학교 문헌정보학과/교수
오상훈	DCF/사무국장		
조명대	성균관대학교 대학원/박사 과정	권도윤	경북대학교 대학원/박사 과정
장인호	성균관대학교 대학원/박사 과정	원선민	성균관대학교 대학원/석사
곽재현	성균관대학교 대학원/석사 과정	정은재	성균관대학교/학부 과정

Keywords (5개 내외)	한글	
	영문	

- 주1) 연구목표, 연구내용, 연구성과를 서술형으로 기재
 2) 국가연구개발사업 DB를 통한 공개를 희망하지 않는 경우 공개가능여부란에 “공개불가”로 표시
 3) 연구성과는 그간의 연구결과 및 기대성과를 서술