

전자기록물 장기보존을 위한 마이그레이션 전략에 관한 연구*

A Study on Migration Strategy for Long-term Preservation of Electronic Records

권도윤(Do-Yun Kwon)**

김희섭(Heesop Kim)***

오삼균(Sam-Gyun Oh)****

목 차

- | | |
|-----------------|-----------------------|
| 1. 서론 | 2.2 마이그레이션의 특징 |
| 1.1 연구의 배경 | 3. 마이그레이션 전략 및 적용사례 |
| 1.2 연구의 목적 및 방법 | 3.1 마이그레이션 전략의 유형 |
| 1.3 연구의 제한점 | 3.2 마이그레이션 전략 적용사례 |
| 2. 마이그레이션 | 4. 마이그레이션 전략 적용시 고려사항 |
| 2.1 마이그레이션의 정의 | 5. 결론 |

<초 록>

본 연구는 마이그레이션에 대한 개념 정리와 더불어 전자기록물 장기보존을 위한 다양한 마이그레이션 전략 및 DPT, CAMiLEON 등 주요 마이그레이션 전략 적용사례에 대한 조사분석을 통해 국내 전자기록물 장기보존 전략으로서 마이그레이션을 적용하기 위해 고려해야 할 시사점들을 제안하였다.

주제어: 마이그레이션, 장기보존전략, 전자기록물, 매체변환, 포맷변환

<ABSTRACT>

In this study, we proposed some fundamental considerations in application of a migration strategy for long-term preservation of electronic records. It also reviewed the notion of migration and various data migration strategies such as DPT and CAMiLEON through the literature survey.

Keywords: migration, long-term preservation strategy, electronic records, change media, change format

* 본 연구는 행정안전부 국가기록원의 지원을 받아 기록물 보존기술 연구개발(R&D) 사업의 일환으로 이루어졌음.

** 경북대학교 대학원 기록학과 박사수료(kdy@dmi.re.kr)(제1저자)

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

**** 성균관대학교 문헌정보학과 교수(samoh@skku.edu)(공동저자)

■ 접수일자 2009년 11월 17일 ■ 수정일자 2009년 12월 16일 ■ 게재확정일자 2009년 12월 21일

1. 서론

1.1 연구의 배경

인터넷 및 정보통신 기술의 발달로 인해 현대 사회의 개인 및 조직은 일상생활 또는 업무 활동을 통해 다양한 전자형태의 기록물들을 생산하고 있다. 우리나라에서는 전자정부 구현이라는 목표 아래 2004년 이후부터는 각급 행정 기관 및 공공기관에서 전자문서시스템을 통해 전자기록물들을 생산, 유통하고 있으며, 미래에는 이러한 전자기록물들의 생산량이 더욱 증가할 것이고 비중 또한 매우 커질 것이다.

전자기록물은 전자환경에서 생산, 유통되며 대부분 손상이 쉽게 되는 자기매체 또는 광매체 등과 같은 디지털 매체에 저장되기 때문에 수명주기가 매우 짧은 특성을 가지고 있다. 또한 전자기록물은 내용과 매체가 구분되어 있으므로 매체 및 내용의 보존과 더불어 접근성 유지가 이루어져야만 전자기록물의 장기보존이 가능해진다. 따라서 전자기록물을 장기적으로 안전하게 보존하기 위해서는 안전한 환경에 전자기록물을 저장하고, 주기적으로 새로운 매체로 복제하는 등 적절한 기술적인 관심을 기울여야만 한다.

전자기록물 장기보존을 위해서는 전자기록물의 4대 속성인 진본성, 신뢰성, 무결성, 이용 가능성이 유지되어야 하는데, 이는 전자기록물

그 자체에 대한 보존뿐만 아니라 전자기록물에 대한 지속적인 접근성을 보존해야 함을 의미한다. 전자기록물의 장기보존 전략은 이와 같이 전자기록물에 대한 접근을 가능하게 하기 위한 다양한 방법론과 고려사항을 의미하며, 전자기록물은 매우 다양한 형태와 논리적인 구조를 가지고 있으므로 전자기록물의 장기보존 전략은 전자기록물의 형태 및 특성에 따라 다양하게 선택될 수 있다.

이러한 전자기록물의 특성 및 장기보존 필요성 증대에 따라 해외에서는 InterPARES,¹⁾ CEDARS,²⁾ DPT,³⁾ CAMiLEON,⁴⁾ PLANETS,⁵⁾ KEEP⁶⁾ 등 다양한 전자기록물 장기보존 연구가 진행되었거나 진행 중에 있으나, 대부분의 연구는 전자기록물 장기보존에 대한 연구의 일부로서 마이그레이션에 대한 연구를 수행한 경우는 종종 찾아볼 수 있지만 전자기록물 장기보존을 위한 마이그레이션 전략에 초점을 맞추어 연구를 수행한 사례는 거의 없는 실정이다.

따라서, 본 연구에서는 IT 강국으로서의 우리나라의 IT 인프라와 공공기관의 기록물관리에 관한 법률 시행 등으로 인해 향후 급증할 것으로 예상되는 전자기록물의 장기보존을 위한 기초연구로서 전자기록물 장기보존을 위한 마이그레이션 전략에 대해 연구하고자 한다.

1) International Research on Permanent Authentic Records in Electronic Systems.
2) CURL Exemplars in Digital Archives.
3) Digital Preservation Testbed.
4) Creative Archiving at Michigan and Leeds Emulating the Old On the New.
5) Preservation and Long-term Access through NETworked Services project.
6) Keeping Emulation Environments Portable.

1.2 연구의 목적 및 방법

본 연구의 목적은 전자기록물 장기보존을 위한 마이그레이션 전략을 조사·분석하여 향후 기록관리 분야에서 주요 이슈가 될 전자기록물 장기보존을 위한 기록관리 정책 결정 및 국내 전자기록물 장기보존 전략으로서 마이그레이션을 적용할 경우에 고려해야 할 시사점들을 도출하는 것이다.

이를 위해 본 연구에서는 마이그레이션에 대한 다양한 정의 및 전자기록물 장기보존 전략으로서의 마이그레이션의 특징들에 대해 조사하여 정리하고, 여러 연구 또는 프로젝트를 통해 제시된 다양한 유형의 마이그레이션 전략을 비교분석하였다. 이와 더불어 전자기록물 장기보존을 위해 마이그레이션 전략을 적용한 해외 주요 적용사례들을 조사분석하여 국내 전자기록물 장기보존 전략으로서의 마이그레이션 전략을 적용하기 위해 고려해야 할 주요 시사점들을 도출하였다.

1.3 연구의 제한점

본 연구는 마이그레이션 관련 연구 및 적용 사례를 중심으로 한 일반적인 마이그레이션 전략에 대한 연구로서, 마이그레이션의 목적, 대상, 목표 파일 포맷 등 구체적인 마이그레이션 조건들에 대한 실증적인 연구가 이루어지지 못했다는 한계를 지닌다. 따라서 실제 기록관리 현장에서 마이그레이션 전략을 적용하기 위해서는 대상 기관의 전자기록 환경 및 전자기록물의 특성 등에 따라 본 연구에서 밝히고 있는 다양한 마이그레이션 전략 및 고려사항들에 대

한 충분한 검토가 필요할 것이다.

본 연구에서는 데이터 마이그레이션, 시스템 마이그레이션, 소프트웨어 마이그레이션 등 다양한 마이그레이션 분야 중에서 기록학 분야의 데이터 마이그레이션에 대한 전략만을 대상으로 연구를 제한하였다.

또한 본 연구에서 조사분석한 마이그레이션 관련 연구 및 적용사례의 경우 동일 조건 하에서의 비교분석이 아니라는 점에서 본 연구의 결과를 일반화시켜 적용하기에는 다소 무리가 있다. 본 연구는 연구의 목적에서 밝힌 바와 같이 실제 전자기록물 장기보존을 위해 마이그레이션을 적용하고자 하는 것이 아니라, 다양한 의미로 사용되고 있는 마이그레이션에 대한 개념 정리와 더불어 주요 전략 유형에 대한 비교분석, 마이그레이션 전략 적용사례 조사분석을 통해 마이그레이션 전략의 기초연구로서의 의미를 두고 있음을 미리 밝혀둔다.

2. 마이그레이션

2.1 마이그레이션의 정의

마이그레이션이라는 용어는 그 범위, 대상, 관점에 따라 매우 다양한 용도로 사용되고 있다. 컴퓨터 및 정보과학 분야에 있어서도 마이그레이션의 정의는 서로 매우 다른 것을 볼 수 있다. 따라서, 먼저 다양한 마이그레이션에 대한 정의를 살펴보고 이를 근거로 본 연구에서의 마이그레이션에 대한 정의를 하고자 한다.

세계적인 IT 분야 온라인 용어 사전인 whis.com⁷⁾에서는 마이그레이션을 ‘마이그레이션이

란, 하나의 운영 환경으로부터 다른 운영환경으로 이전하는 프로세스를 뜻한다. 여기서 다른 환경은 일반적으로 더 나은 환경을 뜻한다.'라고 정의하고 있다.

온라인 백과사전인 wikipedia⁸⁾는 컴퓨터 분야에 있어 마이그레이션을 PC 마이그레이션, 데이터 마이그레이션, 프로세스 마이그레이션, 시스템 마이그레이션으로 구분하고 있다. 이 중, 기록학 분야에서 주로 다루는 마이그레이션은 데이터 마이그레이션이다. wikipedia에 따르면, 데이터 마이그레이션이란 '데이터 저장 장치들 사이에서 정보를 이동하는 것'으로 정의하고 있다.

미국의 국립 결핵 통제 협회(NTCA, National Tuberculosis Controllers Association)는 결핵 환자 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.

지금까지는 일반적인 사전 및 컴퓨터 과학 분야에서 이야기하고 있는 마이그레이션의 정의를 살펴보았다. 이를 종합하면, 마이그레이션이란 '과거의 시스템 및 운영 환경에서 사용되는 데이터를 더 나은 시스템 및 운영환경으로 이전하는 프로세스'라고 정의할 수 있다.

기록보존 분야에 있어서 마이그레이션의 정의는 일반적인 사전 및 컴퓨터과학 분야의 정의에 비해 좀 더 분명한 목적성과 속성을 포함하고 있다. 또한 마이그레이션과 전송(transfer)

의 구분 및 대상에 따른 마이그레이션의 특성 구분까지 발견할 수 있다.

CPA/RLG⁹⁾의 디지털 정보 보존 특별 전문 위원회(the Task Force on the Archiving of Digital Information)의 1996년 보고서에는 "마이그레이션은 하나의 하드웨어/소프트웨어 환경으로부터 다른 환경으로의 또는 하나의 컴퓨터 기술 세대로부터 다음 컴퓨터 기술 세대로의 디지털 자원들을 정기적으로 이관하기 위해 계획된 일련의 유기적인 업무"라고 마이그레이션을 정의하고 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.

InterPARES(International Research on Permanent Authentic Records in Electronic Systems)의 용어 데이터베이스에서는, 좀 더 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.

국제표준화기구인 ISO(International Organization for Standardization)¹⁰⁾는 Inter-

7) <<http://whatis.techtarget.com>> [cited 2009.9.16].

8) <<http://www.wikipedia.org>> [cited 2009.9.16].

9) The Commission on Preservation and Access/the Research Libraries Group.

10) 지적 활동이나 과학·기술·경제활동 분야에서 세계 상호간의 협력을 위해 1946년 설립한 국제기구.

PARES의 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.

가장 많이 인용되고 있는 아카이브 시스템 참조모델인 OAIS 참조모델은 그 프레임워크 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 이관’이다. 마이그레이션은 아래 세 가지 속성들에서 전송

과 구분된다. 1) 전체 정보내용 보존 중심, 2) 정보의 새로운 기록적 실행이 오래된 것의 대체라는 관점, 3) 전송의 모든 측면에 관한 전체적인 통제와 책임이 OAIS 내에 있다는 것이다.

DPT는 마이그레이션을 “하나의 하드웨어 또는 소프트웨어 환경에서 다른 하드웨어 또는 소프트웨어 환경으로 기록물을 이관하는 것”으로 간략하게 정의하고 있다.

KS X ISO 15489-1(문헌정보-기록관리-제1부:일반사항)에서는 마이그레이션을 “기록의 진본성, 무결성, 신뢰성, 이용가능성을 유지하면서 한 시스템에서 다른 시스템으로 기록을 이전하는 행위”로 정의하고 있다.

이를 종합하면, <표 1>과 같다.

지금까지 살펴본 정의를 종합하면, ‘마이그레이션이란, 시간이 흐름에 따라, 기록매체, 포

<표 1> 마이그레이션의 정의

출처	마이그레이션의 정의
whatis.com	하나의 운영 환경으로부터 다른 운영환경으로 이전하는 프로세스를 뜻한다. 여기서 다른 환경은 일반적으로 더 나은 환경
wikipedia	데이터 저장 장치들 사이에서 정보를 이동하는 것
NTCA	과거의 데이터를 새로운 시스템으로 불러오는 프로세스
CPA/RLG	하나의 하드웨어/소프트웨어 환경으로부터 다른 환경으로의 또는 하나의 컴퓨터 기술 세대로부터 다음 컴퓨터 기술 세대로의 디지털 자원들을 정기적으로 이관하기 위해 계획된 일련의 유기적인 업무
IFLA	하나의 하드웨어/소프트웨어 환경으로부터 다른 것으로, 혹은 하나의 컴퓨터 기술로부터 이어지는 기술로 주기적으로 이전하는 것
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
ISO	시간이 흐름에 따라 시스템과 기록매체가 다른 것으로 대체되거나, 유실되거나, 노후화될 수 있는 상황에서, 기록물의 지속적인 접근가능성을 보장하기 위해 기록물을 하나의 시스템 또는 저장 매체로부터 다른 것으로 이전하는 프로세스
OAIS	OAIS내에서의 보존을 목적으로 한 디지털 정보의 이관
KS X ISO 15489-1	기록의 진본성, 무결성, 신뢰성, 이용가능성을 유지하면서 한 시스템에서 다른 시스템으로 기록을 이전하는 행위
DPT	하나의 하드웨어 또는 소프트웨어 환경에서 다른 하드웨어 또는 소프트웨어 환경으로 기록물을 이관하는 것

맷, 시스템이 대체되거나, 유실되거나, 노후화 될 수 있는 상황에서, 기록물로의 지속적인 접근을 가능하게 하도록 하기 위해 다음 세대의 운영 체제, 저장 매체, 포맷, 시스템으로 이전의 데이터를 이관하는 프로세스이다.' 여기서 프로세스라 함은, 마이그레이션이 단일 행위가 아님을 나타낸다. 즉, 마이그레이션은 다양한 과정을 내포하고 있으며, 각 과정에 대한 고려사항이 존재함을 나타낸다.

2.2 마이그레이션의 특징

마이그레이션은 전자기록물의 장기보존을 위해 가장 많이 이용되는 전략이다. 마이그레이션은 하나의 기술로부터 다른 기술로 디지털 객체들을 복제 또는 변환하는 것으로 간단하게 정의될 수 있지만, 동시에 디지털 객체들의 중요한 속성들을 보존한다. 마이그레이션은 디지털 객체의 환경이라기보다는 디지털 객체 자체에 초점을 맞추고 있다. 마이그레이션은 전자기록물의 기술 플랫폼의 기능을 보존하거나 모방하는 것이 아니라 파일의 변환에 기초하고 있으며, 전자기록물의 논리적 인코딩¹¹⁾을 변화시켜 원래의 기술 환경이 노후화될 때 현재의 하드웨어 및 소프트웨어에서 사용될 수 있다. 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 이전 등과 같은 행위도 여기에 해당한다.

보존전략으로서의 마이그레이션의 기술적인 측면은 앞서 설명한 변환 행위와 같은 것으로 볼 수 있다. 보존전략으로서의 마이그레이션이란, 마이그레이션을 통해 현재의 시스템을 이용하는 환경에서 이전의 시스템에서 생성된 전자기록물을 사용할 수 있는 형태로 유지하는 것을 뜻한다.

보존전략으로서 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. 마이그레이션의 또다른 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. 또한 마이그레이션은 일반적으로 디지털 객체의 지적 내용을 보존하기 위한 신뢰성이 높은 방법으로서, 특히 페이지 기반의 기록물의 장기보존에 적합하다. 이밖에도 마이그레이션은 원래의 응용프로그램을 보유할 필요가 없고, 마이그레이션을 위한 절차가 간결하게 잘 확립되어 있다는 장점이 있다.

마이그레이션 전략은 위와 같은 많은 장점들을 가지고 있음에도 불구하고, 그 자체로 전자기록물 장기보존을 위한 해답이 되지 않고 있다.

11) 문서의 물리적인 표현형식이며 문서를 형성하는 데이터.

왜냐하면, 마이그레이션의 단점 또한 무시할 수 없는 수준이기 때문이다. 마이그레이션이 가지는 가장 큰 단점은 데이터와 속성들의 잠재적인 손실 가능성으로 전자기록물 장기보존의 주요 이슈인 전자기록물의 무결성과 진본성을 손상시킬 수 있다는 것이다. 결국, 데이터 또는 속성이 손실될 수 있는 위험이 상존한다는 것이다. 마이그레이션은 원본의 논리적 구조 혹은 렌더링(rendering) 방법을 다르게 하여 새로 작성하는 방법이다. 즉, 마이그레이션을 통한 보존은 원본을 본래의 상태대로 보존하는 것을 뜻하지 않는다. 이와 같은 변환의 과정에서 언제든지 데이터 또는 속성의 일부가 변경될 가능성을 가지고 있다. 마이그레이션된 디지털 객체는 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 객체의 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 매 시점마다 마이그레이션 업무를 수행해야할 필요가 있다. 이는, 장기적인 마이그레이션 전략의 수립을 어렵게 하는 한 요소이며, 마이그레이션 전략이 장기화될수록 비용이 점차 증가하게 된다.

반면 마이그레이션과 함께 양대 전자기록물 장기보존 전략의 하나인 에물레이션은 디지털 객체에 초점을 맞추는 것이 아니라 그 객체가

표현된 환경에 초점을 맞추며 디지털 객체가 진본 형태로 표현될 수 있는 환경을 재창조하는 것을 목표로 하고 있다.

에물레이션은 하드웨어 또는 소프트웨어의 부분을 모방하여 원래의 형태로 장비와 기능이 작동되도록 하는 과정이다. 마이그레이션이 디지털 객체의 변환에 중점을 두는 반면, 이 전략은 원래의 컴퓨터 환경을 재생산하는 것이기 때문에 원본 전자기록물에 변형을 주지 않고 가장 정확한 형태로 기록물을 보존할 수 있다는 장점이 있다. 또한 데이터가 적절한 메타데이터 및 소프트웨어와 함께 한번 수록되면, 매체 재생산에 필요한 다른 활동들을 하지 않아도 된다. 하나의 에물레이터는 동일한 운영 환경을 요구하는 여러 데이터 객체들에 해결책으로 사용될 수 있기 때문이다. 즉, 에물레이션 전략의 장점은 정기적으로 마이그레이션 할 필요 없이 원본 소프트웨어 환경을 이용하여 객체를 보존함으로써 디지털 객체의 기능과 표현이 손상되지 않은 상태로 유지할 수 있다는 것이다.

그러나 보존전략으로서 에물레이션을 성공적으로 적용하기 위해서는 아직까지 다양한 문제들이 남아있다. 먼저, 구형의 데이터 매체로부터 디지털 콘텐츠를 안전하게 캡처할 수 있는 방법들이 필요하다. 그리고 현재 사용하고 있는 디지털 아카이빙 시스템들에 에물레이션을 통합할 수 있는 지식과 경험이 더 축적되어야 한다. 또한 텍스트, 멀티미디어, 애플리케이션, 웹사이트 등 광범위한 디지털 콘텐츠들에 대해 지속적이고 적절한 형태로 현재의 에물레이션 솔루션들을 재설계하여야 한다. 지금까지 설명한 마이그레이션과 에물레이션의 장단점을 비교하면 <표 2>와 같다.

〈표 2〉 마이그레이션과 에몰레이션의 장단점 비교

	장 점	단 점
마이그레이션	단순 마이그레이션에 대한 절차가 잘 확립되어 있음	복잡한 마이그레이션에 대해서는 특별 프로그램이 작성되어야 하는 비용 요구
	현재 대부분의 디지털 보존소의 우선 전략임	진본성을 보장하는 엄격한 품질 통제 절차가 마련되지 않으면 무결성을 해칠 수 있음
	기술의 진보와 플랫폼 범위의 감소를 통해 보다 단순해 질 것임	몇몇 기능과 원본의 외형 및 느낌을 손실할 수 있음
		시간이 많이 걸리며 복잡한 작업이 될 수 있음
보다 복잡한 디지털 자원은 기능이 현저히 손실된 채 마이그레이션 될 것임		
		정기적인 간격으로 시행되어야 함
에몰레이션	원본의 기능, 외형과 느낌을 재생산	아직까지 연구 단계이며 더 많은 테스트가 필요 원본의 기능, 외형과 느낌 중 단지 일부만 에몰레이트할 수 있음
	마이그레이션과 관련된 반복 비용을 피함	규모의 경계가 없는 한 매우 낭비적임. 새로운 에몰레이터들은 주요 컴퓨터 패러다임의 변화가 구축되기를 필요로 하는데, 이 비용은 반복적인 마이그레이션 비용을 넘어설 수 있음
	더욱 복잡한 디지털 자원에 대해 가장 전망이 있음	소프트웨어 저작권 논쟁이 해결되어야 하는데, 이는 매우 복잡함

3. 마이그레이션 전략 및 적용사례

3.1 마이그레이션 전략의 유형

전자기록물 장기보존을 위한 마이그레이션 전략의 유형은 지금까지 진행된 연구 및 프로젝트 등을 통해 살펴보면 다양한 형태로 제시되고 있음을 알 수 있다.

Dollar(1999)는 마이그레이션 전략을 리포맷,¹²⁾ 복제, 변환, 마이그레이션으로 구분하였다. 리포맷은 물리적 매체가 변경되거나 코드가 변할 경우 물리적 표현 또는 지적 내용의 변경없이 전자기록물의 중요한 비트스트림에서 변화를 일으키는 것이며, 복제는 전자기록물을 구조, 내용 그리고 맥락의 손실없이 동일한 포

맷 규격으로 오래된 저장 매체로부터 새로운 저장 매체로 이관하는 것이다. 변환은 비록 중요한 비트스트림이 변경되더라도 구조의 손실이 없거나 거의 없이 그리고 내용 또는 맥락의 손실없이 하나의 응용프로그램 환경으로부터 새로운 응용프로그램 환경으로 전자기록물의 자동 이관이 이루어지는 것을 말하며, 마이그레이션은 기록물을 포함하고 있는 레거시 시스템과 새로운 응용프로그램 사이에 소급호환성과 이관/접수 게이트웨이들이 없는 특정 도구를 요구하는 변환이라고 설명하고 있다.

국제표준(ISO 14721:2003)인 OAIS 참조 모델에서는 마이그레이션 전략의 유형을 4가지(갱신, 복제, 재패키지화, 변형)로 구분하고 있다. 갱신은 하나의 매체로부터 같은 유형의 다른

12) 보존 및 접근을 위해, 원본과는 다른 포맷 내지 구조를 지닌 사본을 생성시키는 행위.

매체로 비트(bit)들을 복사하는 것으로서, 시스템의 변화없이 매체로의 접근을 유지할 수 있는 것이며, 복제는 패키지 정보, 내용 정보 그리고 보존기술정보(PDI)¹³⁾에 비트 변화가 없는 것으로서, 동일한 매체 혹은 다른 매체로의 복사를 포함한다. 재패키지화는 전송과정에서 패키지 정보에 일부 변화가 발생하는 것이며, 변형은 내용 정보 또는 보존기술정보에 일부 변화가 요구되는 디지털 마이그레이션을 말하며, 변형은 역행이 가능한가, 불가능한가의 여부에 따라 역변형과 비역변형으로 구분된다. 만일 목표 포맷을 미리 정의하고 변형을 시행한 경우는, 역행이 가능하지만, 그렇지 않을 경우는 역행이 불가능할 수 있다.

1994년 보존과 접근 위원회(Commission on Preservation and Access)와 연구도서관 그룹(Research Libraries Group)에서는 “디지털 전자 형태로 저장된 미래의 불명확한 기록물에 대한 접근을 지속시키는 것”을 목적으로 하는 디지털 정보 아카이빙 특별 전문 위원회(the Task Force on Archiving of Digital Information, TFADI)를 만들었는데, 이 특별 전문 위원회에서는 이미 디지털 형태로 존재하는 자원들에 관심을 가졌고, 매체 퇴화와 기술적 노후화 모두를 방지해야 할 필요를 인식했으며, 마이그레이션을 디지털 아카이브의 필수적인 기능으로 인식하였다. 여기에서는 “마이그레이션은 하나의 하드웨어/소프트웨어 환경으로부터 다른 환경으로의 또는 하나의 컴퓨터 기술 세대로부터 다음 컴퓨터 기술 세대로의 디지털 자원들을 정기적으로 이관하기 위해 계획된 일련

의 유기적인 업무”라고 마이그레이션을 정의하였다. 태스크 포스에서는 디지털 아카이브의 운영 환경들의 변화로 인해 디지털 아카이브의 콘텐츠들을 이관하는 것이 필요하게 되었으며, 정보를 접근가능하고 이용할 수 있도록 하기 위해 디지털 정보를 노후화된 시스템들로부터 현재 하드웨어와 소프트웨어 시스템들로 이관하기 위한 다양한 마이그레이션 전략들이 있으며 모든 형태의 디지털 정보에 적용할 수 있는 하나의 전략은 없다고 밝히고 있다. 이 보고서에서는 매체변환, 포맷변환, 표준으로의 통합, 마이그레이션 경로 구축, 처리센터의 이용 등 5가지의 마이그레이션 전략을 제시하였다.

매체 변환은 디지털 자료를 덜 안정적인 매체에서 보다 안정적인 매체로 이관하는 것으로서, 이 전략의 가장 일반적인 버전은 디지털 정보를 종이에 프린트 하거나 마이크로필름에 기록하는 것을 포함한다. 종이와 마이크로필름은 디지털 매체보다 훨씬 안정적이고 종이와 마이크로필름으로부터 정보를 검색하기 위한 특별한 하드웨어 또는 소프트웨어를 필요로 하지 않는다. 새로운 저장 매체로 정보를 복사하여 디지털 형태로 보유하는 것은 정보가 ASCII 텍스트 파일 또는 간단하고, 단일 구조를 가진 플랫 파일(flat file)과 같은 “소프트웨어 독립적인” 형태로 존재할 때에는 적절할 지도 모른다. 새로운 매체와 저장 포맷들이 만들어졌음에도 불구하고 데이터는 어떤 중요한 변화없이 그들의 논리적 구조 속으로 이관되었다. 하나의 매체에서 다른 매체로 복사하는 것은 폭넓게 이용가능하고 실행하기 쉽다는 명확한 이점

13) Preservation Description Information(PDI): OAIS 참조모델 정보패키지의 주요 구성요소 중의 하나.

을 가지고 있다. 매체 변환은 내용을 유지하는 것이 중요하고, 디스플레이, 색인 그리고 연산 특성들이 중요하지 않는 경우에는 디지털 정보를 보존하기 위한 비용 효율이 높은 전략이다.

마이그레이션 전략으로서 단순하고 일반적인 복사는 디지털 정보의 형태 또는 구조에서 큰 손실을 가져올 지도 모른다. 일부 비표준 데이터에 대해 마이그레이션하기 위해 접근할 때, 종종 소프트웨어를 통해 관리되고 해석되는 문헌의 구조, 데이터베이스에 내포된 데이터 관계들, 그리고 인증의 수단을 제거하는 변화를 초래한다. 연산 성능, 그래픽 디스플레이, 색인 그리고 다른 기능들 또한 잃어버리고 원 객체의 찌꺼기만 남을 수 있다. 이 전략은 복합 시스템들의 복잡한 데이터 객체들의 보존에는 적합하지 않다. 예를 들면 스프레드시트에 내장된 방정식을 마이크로필름으로 보존하거나, 쌍방향 폴 모션 비디오를 출력하거나, 또는 멀티미디어 문서를 플랫폼 파일로 보존하는 것은 불가능하다.

포맷 변환은 소프트웨어의 버전 업에 따른 이용 가능성 확보를 위해 전자기록물의 파일 포맷을 최신 포맷으로 변환하거나, 이종 소프트웨어 간 전자기록의 상호이용가능성 확보를 위해 전자기록의 파일 포맷을 이용하고자 하는 소프트웨어에서 지원하는 파일 포맷으로 변환하는 것을 말한다. 마이그레이션 전략으로서의 포맷 변환은 원본 객체의 디스플레이, 배포 그리고 연산 특성들을 보다 잘 보존할 수 있다는 이점을 가지고 있다. 그러나 소프트웨어와 표준들은 점진적인 발전을 계속하기 때문에 이 전략은 간단하지만 정기적인 이관의 필요성 또는 디지털 객체의 무결성에 대한 마이그레이션

의 잠재적인 효과들에 대한 분석의 필요성을 없애지는 못한다.

디지털 아카이브는 디지털 정보 객체들에 대한 참조를 용이하게 하고 시스템들 사이에서 디지털 정보 객체들을 교환할 수 있도록 해주는 널리 보급된 데이터 및 통신 표준들을 채택함으로써 많은 이익들을 얻을 것이다. 많은 기관들에서 업무적 요구들은 데이터 표준들의 개발과 채택을 추구한다. 예를 들면, 지리정보시스템(GIS)을 만들고 이용하고 유지하는 기관들은 널리 인정된 표준들에 따른 데이터를 생산함으로써 데이터 변환 및 유지 비용을 줄이려고 노력한다. 현 세대의 시스템들 사이에서 디지털 정보의 안전하고 신뢰성있는 교환을 위한 업무 요구들을 제기한 표준화 단체들은 데이터의 표준화와 일반화를 강요할 것이고, 결국 새로운 기술 세대로의 마이그레이션을 용이하게 할 것이다. 디지털 아카이브는 표준 개발과 병행하여 그들의 기술 인프라가 널리 채택된 표준들에 부합하는지를 확인해야 한다.

마이그레이션 경로 구축은 조직이 기술변화에 보조를 맞추기 위해 새로운 하드웨어와 소프트웨어를 서서히 발전시킬 수 있도록 해주는 일련의 변환이다. 장기보존 계획은 디지털 보존의 중요한 요소로서, 디지털 정보의 생산자/제공자/소유자는 그들 객체들의 보존에 대한 최초의 책임을 받아들이는 점에서 그들은 디지털 정보를 생산하는 프로세스 또는 시스템의 필수적인 부분으로서 보존을 위한 지혜 또는 다른 준비들을 이해하게 될 것이다.

정부, 기관 그리고 회사에서, 아키비스트들과 사서들은 디지털 보존을 위한 가이드라인과 조언들을 출판할 수 있으며, 그들의 모 기관들

이 일반적으로 이용되는 규칙들을 채택하고, 데이터 표준에 따르고 그리고 마이그레이션을 지원하는 응용 소프트웨어를 선택하도록 촉진할 수 있다. 보존 커뮤니티는 초기 설계 단계의 소급 호환성과 같이 대체로 기록적 고려로 만들어진 정보 시스템과 표준을 생산하기 위한 산업과 함께 움직이는 것을 필요로 한다. 소급 호환성 또는 마이그레이션 경로는 중요한 리포맷없이 오래된 시스템으로부터 데이터를 읽기 위한 새로운 세대의 소프트웨어를 가능하게 할 것이다.

미래에 표준과 마이그레이션 경로가 혼해진다고 하더라도, 현재 비표준 포맷들의 많은 디지털 자료가 존재하고, 조직들과 개인들은 마이그레이션을 필요로 하는 디지털 자료들을 계속해서 생산해 낼 것이다. 노후화된 자료의 마이그레이션과 리포맷에 전문화된 “처리 센터”의 발달은 비용 효율이 높은 디지털 보존 방법을 제공할 것이다. 처리 센터는 텍스트, 특정 데이터베이스 구조, 지리정보시스템(GIS) 또는 멀티미디어 제품들과 같은 특정 형태의 자료들을 위한 리포맷 서비스를 제공할 것이다. 처리 센터들은 마이그레이션을 지원하기 위해 오래된 버전들의 하드웨어와 소프트웨어를 보존할 것이다. 처리 센터들은 Rothenberg에 의해 제안된 것처럼 “look and feel”과 같이 디지털 정보를 읽고 보기 위한 플랫폼을 제공할 것이다. 처리 센터들은 규모의 경제의 이점을 얻을 것이고, 고급 기술적 지식의 이용을 극대화할 것이다. 마이그레이션/보존 서비스 센터들은 오래된 자체 제작 영화와 노후화된 비디오 포맷

들을 재포맷하는 상업 회사 또는 분산된 보존 프로그램을 가진 도서관과 아카이브즈 협회들과 공통점이 있을 것이다.

JISC/NPO¹⁴⁾는 디지털 보존에 관련된 여러 개의 연구를 지원했는데, 이들 연구들은 디지털 객체들의 장기보존을 계획하고 있는 모든 사람들에게 가치있는 원자료를 만들어 냈다. Mary Feeney는 “The Digital Culture: Maximising the Nation’s Investment”라는 연구에서 마이그레이션을 매체변환, 소급호환성 유지, 상호운용성 확보, 표준포맷으로의 변환으로 구분하였다(Feeney 1999).

매체 변환은 기본적인 마이그레이션 전략 중의 하나로서 덜 안정적인 매체로부터 보다 안정적인 매체로 디지털 자원들을 이관하는 것을 말한다. 이 전략의 가장 간단한 버전은 손상되기 쉬운 자기 매체로부터 보다 안정적이고 통제된 환경으로 정보를 이동시키는 것을 포함한다. 보다 극단적인 버전은 디지털 정보를 종이 위에 인쇄하거나 또는 보존용 마이크로필름에 기록하는 것일 것이다. 이것은 기본 데이터를 보존하는 반면, 자원의 디지털 기능을 파괴한다. 예를 들면, 계산 능력, 그래픽 디스플레이 또는 색인 등과 같이 훨씬 가치있는 데이터가 그 과정에서 손실될 것이다. 스프레드시트에 내장된 방정식들을 마이크로필름으로 만들거나, 상호작용하는 풀 모션 비디오를 인쇄하거나, 또는 멀티미디어 문서를 플랫폼파일로 만드는 것은 불가능하다. 그러므로, 매체변환은 다른 전략들이 요구에 적합하지 않을 경우 마지막으로 고려되어야 한다.

14) Joint Information Services Committee of the Higher Education Funding Councils/National Preservation Office.

두번째 마이그레이션 전략은 소급호환성을 가진 대중적인 응용 소프트웨어에 의존한다. 가장 인기있는 워드 프로세싱 패키지들의 최신 버전들은 동일한 패키지의 이전 버전들에서 생산된 파일들을 디코딩할 수 있는 능력이 있다. 마이그레이션은 프로세스를 테스트하고, 파일들을 새로운 버전으로 로딩(loading)하고 새로운 파일 포맷으로 저장하는 것을 포함한다. 그러나, 이 전략은 장기 또는 보다 복잡한 디지털 자원들에 대해서는 의지될 수 없다. 지속적인 생존 능력을 보증하고 그들의 응용 소프트웨어를 지원하는데 필요한 모든 기술적 또는 상업적 요소들을 통제하는 소프트웨어 공급자는 없다. 소프트웨어 공급자는 폐업하거나 완전히 새로운 소프트웨어 패키지를 개발하거나 그들의 오래된 패키지에 대한 지원을 중단할 수도 있다.

세번째 마이그레이션 전략은 경쟁하는 인기 응용 프로그램들 사이의 상호운용성에 의지하고 있다. 하나의 응용 프로그램에서 생산된 디지털 자원들은 공통 교환 포맷으로 이관할 수 있고 그들을 생산하기 위해 사용되었던 특정 프로그램을 이용할 필요없이 경쟁 응용 프로그램으로 접수될 수 있다. 만일 그러한 상호운용성이 모든 주요 경쟁 응용 프로그램들 사이에 보증된다면, 디지털 정보 마이그레이션은 훨씬 쉬운 프로세스가 될 것이다.

오늘날 최소한 부분적으로는 유사한 소프트웨어는 다른 소프트웨어 패키지에서 생산된 파일들을 해석할 수 있다. 그러나, 간단한 디지털 자원들이라 할지라도 교환(interchange)은 일부 데이터 손실이 발생할 수 있을 것이다. 예를 들면, 워드 프로세싱 프로그램들은 저자가 작업을 ASCII를 이용한 단순한 알파벳문자와 숫

자를 조합한 텍스트 또는 Rich Text Format (RTF)와 같은 다른 교환 포맷들로 저장할 수 있도록 해준다. 문서들은 텍스트만으로 이루어진 경우는 드물며, 포맷 데이터, 그림들 그리고 각주들을 포함하고 있는데 이것들은 교환과정에서 손실될 수 있다. 보다 복잡한 디지털 자원들의 경우에는 더 큰 잠재적인 손실이 발생할 수 있다. 예를 들면, 지리정보시스템 데이터베이스들과 그룹웨어 데이터베이스들에 보유중인 데이터의 교환은 몇 년간의 노력을 들여서 만들고 데이터베이스의 가치의 크기를 나타내는 수천 개의 링크들이 손실될 것이다.

마지막으로 표준 포맷으로의 변환은 특별히 대규모 디지털 자료들, 복잡한 디지털 자료들, 다양한 디지털 자료들을 가진 디지털 아카이브즈에 적합하다. 이것은 상호운용성 전략의 강화된 버전을 제안한다. 상호운용성 전략이 응용프로그램들 내에서 자동적으로 생성될 수 있는 교환 포맷들에 의존적인 반면에, 이 전략의 경우에는 선호하는 포맷들을 정의하고 그들이 수집하고 서비스하는 이용자들에게 가장 적합한 디지털 자원들의 가장 적합한 포맷들을 선정하는 수집 관리자에게 책임이 있다.

변환을 위해 선택된 포맷들은 디지털 자원들의 구조, 수집 관리자에 의해 설정된 목표들, 그리고 이용자들의 요구사항에 의해 결정될 것이다. 예를 들면, 결정은 우선순위가 디지털 자원의 프로세스 또는 편집 능력을 보존하는 것에 주어졌는지, 디지털 자원의 포맷 또는 시각적 표현을 보존하는 것에 주어졌는지에 의해 영향을 받을 것이다. 예를 들면, 많은 데이터 센터들의 핵심 목적은 이용자 응용 프로그램들에 로드될 수 있고, 처리될 수 있어서 새로운 데이터

가 생성될 수 있도록 허용해주는 이용가능한 포맷의 데이터를 보존하고 만드는 것이다. 콘텐츠의 표현은 두 번째로 중요한 반면, 콘텐츠는 매우 귀중한 자원이다. 다른 한편, 아카이브즈와 레코드 센터에서는 기록적 무결성을 보증하기 위해 우선순위가 디지털 자원의 포맷 또는 시각적 표현을 보존하는 것이 될 수도 있다.

이상 마이그레이션 전략에 대한 여러 연구들을 종합적으로 비교분석해 보면 결국 마이그레이션 전략은 <표 3>과 같이 매체변환과 포맷변환으로 구분할 수 있으며, 포맷변환은 다시 소급호환성 유지, 상호운용성 확보, 표준으로의 변환으로 구분해 볼 수 있다.

매체변환은 덜 안정적인 매체에서 보다 안정적인 매체로, 노후화된 매체에서 새로운 매체로, 저용량 매체에서 고용량 매체로 전자기록물을 변환하는 것을 말한다. 전자기록물 저장매체의 노후화로 인해 새로운 저장매체로 전자기록물을 변환하거나, 전자기록물의 무결성을 확보하기 위해 기존의 전자기록물 저장매체보다 더 안정적인 저장매체로 변환하거나, 전자기록물 보존비용 절감을 위해 저용량 저장매체에서 고용량 저장매체로 전자기록물을 변환하는 경우 등이 매체변환에 속한다고 할 수 있다.

포맷변환은 전자기록물의 지속적인 접근성을 보장하기 위해 최신 포맷 또는 표준 포맷으로 전자기록물을 이관하는 것을 말한다. 포맷변환은 다시 소급호환성 유지, 상호운용성 확보, 표준으로의 변환으로 구분할 수 있다. 소급호환성 유지는 오래된 버전의 응용프로그램에서 생성된 전자기록물을 최신 응용프로그램에 적합한 파일포맷으로 변환하는 것(예, Word 98에서 Word 2000으로)을 말한다.

상호운용성 확보는 현재 이용하고자 하는 응용프로그램의 포맷과 다른 전자기록물 포맷을 현재 이용하고자 하는 응용프로그램에서 지원하는 포맷으로 변환하는 것(예, Word 2000을 Adobe의 PDF로)을 말하며, 표준으로의 변환은 독점 포맷인 파일을 널리 보급된 공개 표준에 기반한 포맷으로 변환하는 것(예, Word 2000을 XML로)을 말한다.

3.2 마이그레이션 전략 적용사례

본 연구에서는 마이그레이션 전략을 적용한 주요 사례로서 DPT와 CAMiLEON 프로젝트를 조사분석하였다.

DPT는 통제되고 안전한 환경에서 실험을 수

<표 3> 마이그레이션 전략의 유형

유형	개념	
매체변환	덜 안정적인 매체에서 보다 안정적인 매체로, 퇴보된 매체에서 새로운 매체로, 저용량 매체에서 고용량 매체로 전자기록물을 이관	
포맷변환	소급호환성 유지	오래된 버전의 응용프로그램에서 생성된 전자기록물을 최신 응용프로그램에 적합한 파일포맷으로 변환
	상호운용성 확보	현재 이용하고자 하는 응용프로그램의 포맷과 다른 전자기록물 포맷을 현재 이용하고자 하는 응용프로그램에서 지원하는 포맷으로 변환
	표준으로의 변환	독점 포맷인 파일을 널리 보급된 공개 표준에 기반한 포맷으로 변환

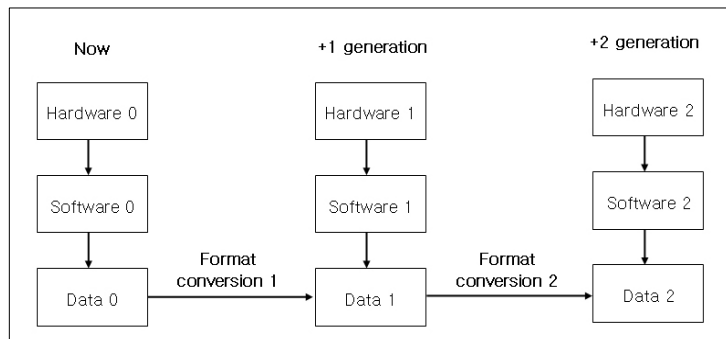
행함으로써 오랜 시간에 걸쳐 진본 기록물에 대한 지속적인 접근성을 보장하기 위한 선택들을 연구하는 것을 전체적인 목표로 하는 3년간에 걸친 실제적인 연구 프로젝트로서, 장기적 보존을 위해 어떤 보존 전략이 적절한지 결정하기 위한 해결 지향 연구 질문들을 중심으로 규정된 실험들을 수행하였다.

DPT는 텍스트, 스프레드시트, 이메일, 데이터베이스 등 4가지의 다른 기록물 포맷을 대상으로, 디지털 기록물의 장기보존을 위한 3가지 접근방법인 마이그레이션, XML 그리고 에뮬레이션에 대해 연구하였다.

〈그림 1〉은 DPT에서 적용한 단계별 마이그레이션(step by step migration)에 대해 설명하고 있는데, 현재 사용되고 있는 하드웨어 0 버전과 소프트웨어 0 버전이 조합된 상태에서 생성된 데이터 0 포맷을 1세대가 지난 후에 하드웨어 1 버전과 소프트웨어 1 버전이 조합된 환경의 데이터 1 포맷으로 포맷 변환 1을 수행하고, 시간이 흘러 새로운 세대인 2세대 환경이 통용되는 시점에서 다시 데이터 2 포맷으로 포맷 변환 2를 수행하는 과정을 지속적으로 수행해야 한다.

DPT는 마이그레이션을 “하나의 하드웨어/소프트웨어 환경에서 다른 하드웨어/소프트웨어 환경으로 기록물을 변환하는 것”으로 정의하고, 소급호환성 유지, 상호운용성 확보, 표준으로의 변환 등 3가지 마이그레이션 전략들에 대해서 연구하고 실험하였다.

소급호환성 유지는 과거 버전의 응용프로그램에서 만들어진 파일을 해석하고 새로운 버전의 응용프로그램에서 정확하게 재생산할 수 있도록 해준다. 새로운 버전의 상용 소프트웨어는 종종 과거 버전들과 호환성을 가지는데, 예를 들면 Excel 2002는 Excel 95에서 만들어진 파일들을 읽고 Excel 95 파일 포맷으로 저장할 수 있다. 하지만 소프트웨어가 단지 제한된 수의 오래된 파일 포맷 세대들을 지원하기 때문에, 이 전략을 이용하여 유지되는 기록물은 보통은 새로운 파일 포맷으로 재저장되어야 할 필요가 있다. 높은 버전으로의 마이그레이션은 일반적으로 수년마다 반복되어야 한다. 모든 마이그레이션은 비록 조금이라 할지라도 변화의 위험을 수반하므로, 디지털 기록물의 진본성과 무결성에 부정적 영향을 미칠 수 있다. 따라서 이 전략은 작은 에러들의 축적 위험들로



〈그림 1〉 단계별 마이그레이션

인해 단기간의 기록물 보존에 유용하고, 장기 보존에는 덜 안정적이다. 소급호환성 유지 전략은 보통 독점적이고 미공개된 파일 포맷들을 마이그레이션하기 위해 사용된다.

기술적인 의미에서 상호운용성 유지는 파일과 디지털 기록물이 더 이상 또는 덜 특정 하드웨어와 소프트웨어 조합에 의존하지 않도록 보증함으로써 디지털 노후화 문제를 해결하는 것이다. 상호운용성 유지는 파일이 하나의 플랫폼 또는 응용프로그램에서 다른 플랫폼 또는 응용프로그램으로 이전할 수 있으며 동일한 방법 또는 유사한 방법으로 파일을 재생산할 수 있는 것을 의미한다. 이를 위해서는 소프트웨어간의 상호운용성을 유지하거나, 중간변환 프로그램을 활용할 수 있다.

표준으로의 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.

3.2.1 텍스트의 마이그레이션¹⁵⁾

텍스트에 대한 소급호환성 유지는 이전 버전으로 생성된 디지털 정보자원을 최신 버전을 사용하여 재생산하는 것을 말하는데, 텍스트에 대한 소급호환성 유지 전략의 시행 결과, 소급호환성 유지 전략은 단기적인 보존에 적합한 것으로 나타났다. 소급호환성 유지 전략은 단기적으로 보았을 때, 기록의 진본성에 큰 위협이 되지 않았다. 즉, 문서의 구조와 외형을 잘

유지하였다. 그러나 3·4 세대 이전의 파일을 최신 소프트웨어로 처리할 때, 문서의 콘텐츠에 큰 변화를 야기하는 것이 발견되었으며, 새 버전이 구 버전의 파일을 해석하지 못하는 문제가 지속적으로 발생하여 진본성과 무결성을 유지하기 어렵다는 것을 발견하였다. 또한 소급호환성 유지 전략은 소프트웨어 공급자로부터의 종속성에서 벗어날 수 없다. 따라서 소급호환성 유지 전략은 단기 보존에는 적합하지만, 장기 보존에는 문제가 있는 것으로 드러났다. DPT는 시장에서 공개되는 모든 버전으로 연속적인 마이그레이션을 수행하기 보다는 몇 버전을 걸쳐 수행하는 바이패싱(bypassing)이 효과적인 것으로 권고하였는데, 바이패싱은 기록물을 보존하기 위한 마이그레이션 횟수를 줄임으로써 변경의 위험을 감소시키고 비용감소의 효과도 가져올 수 있다.

상호운용성 확보 전략은 가장 널리 사용되고 있는 두 종류의 워드프로세서를 대상으로 테스트되었다. 그 결과, 상호운용성 확보 전략 역시, 단기 보존에는 적합하나 장기보존에는 적합하지 않은 것으로 드러났다. 상호운용성 확보 전략은 소급호환성 유지 전략과 같은 한계를 보였다. 즉, 중간변환 프로그램에 의해 원본 기록이 다른 형태로 변환되는 경우가 발생하였고, 소프트웨어 공급자에 대한 의존성에서 벗어날 수 없다는 것을 발견하였다.

표준으로의 변환 전략은 그 자체로, 소급호환성 유지와 상호운용성 확보가 가능할 뿐만 아니라 독점적 공급자에 대한 의존성을 제거할 수 있으므로 장기보존에 적합한 전략이다. DPT

15) Digital Preservation Testbed의 2003년 보고서인 "From digital volatility to digital permanence Preserving text document"의 내용을 발췌한 것임.

는 표준으로의 변환 전략을 테스트하기 위해 PDF와 XML로의 변환을 테스트하였다. PDF로의 변환은 원본 문서의 기능, 외관, 구조를 보존하는 것에 적합하였다. XML은 콘텐츠뿐만 아니라 구조와 외형, 기록의 문맥에서 텍스트를 표현할 수 있었으며 손쉽게 메타데이터를 추가할 수도 있었다.

3.2.2 데이터베이스의 마이그레이션¹⁶⁾

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, 구 버전의 데이터베이스를 신 버전에 저장하는 것은 적합하지 않았다. 또한 지속적인 변환의 필요성이나, 공급자에 대한 의존성에 있어서도 위의 두 결과와 같았다.

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

DPT는 데이터베이스를 변환하는 표준으로 ASCII, SQL, XML을 테스트 하였다. ASCII 파일로의 변환은 일반적으로 모든 시스템에서

읽을 수 있는 문서 형식의 데이터를 생성하였으나 데이터베이스의 구조를 표현할 수 없다는 한계를 가진다. XML은 ASCII의 장점을 가지는 동시에 데이터베이스의 구조를 표현해 줄 수 있다. SQL로 변환은 표준 SQL로의 변환을 의미하는데, 각 공급자에 따라 다른 SQL을 표준 SQL로 변환하는 것이 필요하다.

3.2.3 스프레드시트의 마이그레이션¹⁷⁾

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 발생한다는 점이 지적되었다. 소프트웨어 공급자에 대한 의존성도 동일하게 존재했다.

상호호환성 유지 전략은 스프레드시트의 보존에 적합하지 않은 것으로 나타났다. DPT는 LOTUS 1-2-3과 마이크로소프트 Excel 사이의 상호호환성을 테스트했다. 그 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 변환되었다. 그러나 수식이 계산된 결과를 표현하지 못하였다. 즉, 기본 입력된 표현층만을 표현할 수 있을 뿐 수식 계층과 표현층과의 연결성이 상실되었다. DPT는 스프레드시트의 보존을 위해 XML 포맷이

16) Digital Preservation Testbed의 2003년 보고서인 "From digital volatility to digital permanence Preserving databases"의 내용을 발췌한 것임.

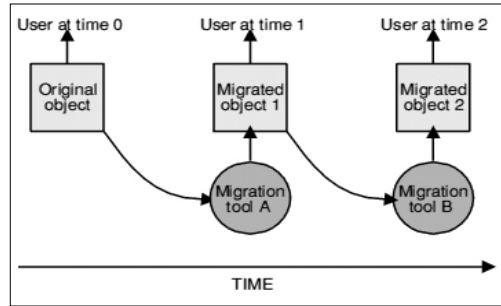
17) Digital Preservation Testbed의 2003년 보고서인 "From digital volatility to digital permanence Preserving spreadsheets"의 내용을 발췌한 것임.

적합하다는 것을 발견하였다. 그러나 XML은 처리하는 도구에 따라 매우 다른 결과를 나타낸다는 단점이 존재한다.

CAMiLEON 프로젝트는 디지털 자원의 장기 보존을 위한 기술적인 전략을 광범위하게 개발하고 평가하는 과정을 통해, 상황에 따라 어떤 장기 보존 전략을 적용할 수 있을지에 대해 연구하는 연구 프로젝트로서, 1999년부터 2003년까지 미국 Michigan 대학과 영국 Leeds 대학의 공동 연구로 이루어졌다. 이 프로젝트의 목표는 원본 디지털 개체의 기능 및 'look and feel'을 유지할 수 있는 장기 보존 전략을 탐색하고, 장기보존 전략으로서 에뮬레이션 기술을 연구하며, 어떤 상황에 에뮬레이션 전략이 적합한지를 확인하는 파악하는 것이다.

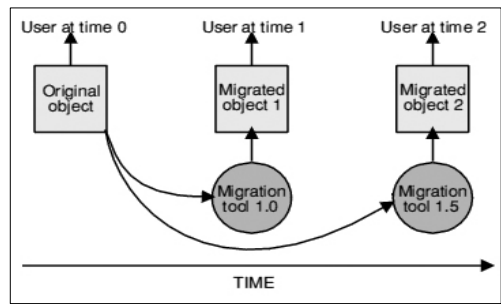
CAMiLEON 프로젝트 연구진은 마이그레이션이 가장 흔하게 시행되는 기록 보존의 기술이지만, 기존의 마이그레이션이 가진 잠재적인 비용의 증가와 외형의 변화를 심각한 문제점으로 생각하였다. 이에 따라, 마이그레이션의 지속적인 수행에서 원형의 훼손을 최소화할 수 있는 마이그레이션 방법론을 연구 하였고, 그 결과 Migration on Request 방법론을 제시하였다. 또한 Migration on Request 방법론을 시행하기 위해 벡터 그래픽(vector graphics) 자원을 대상으로 테스트를 수행하였다.

전통적인 마이그레이션 방법의 경우, <그림 2>와 같이 변환 과정에서 파일에 원본과 차이점이 발생할 경우, 향후 지속적인 변환에서 이 오류가 계속 누적된다. 이것은 원본의 정확성에 치명적인 문제점을 일으킨다. 또한 마이그레이션이 수행될 때마다 새로운 파일 포맷이 생성된다.



<그림 2> 전통적인 마이그레이션

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 3>은 Migration on Request의 개념을 표현하고 있다.



<그림 3> Migration on Request

Migration on Request는 소프트웨어 장수 기술을 이용하여 모든 디지털 객체들을 재현하며 오랜시간동안 유지될 수 있다. 디지털 객체는 단순히 디지털 객체의 원래 포맷으로 보관

되어, 항상 원래의 디지털 포맷으로 되돌아올 수 있으므로 변환 과정동안 초래되는 오류가능성을 크게 감소시킨다. Migration on Request는 원래의 포맷에서 현재의 포맷으로 변환하기 위해 항상 오직 한번의 변환 단계를 거치게 되며, 어떤 경우에는 아카이브에 보유중인 디지털 객체가 원본이 된다.

Migration on Request는 전통적인 마이그레이션에 비해 다음과 같은 여러 가지 중요한 이점들을 제공한다. 첫째, 특정파일 포맷을 읽고 해석한 코드는 단지 한번만 실행된다. 둘째, 단지 하나의 마이그레이션 단계만을 이용함으로써 마이그레이션의 정확성을 증가시킨다. 셋째, 디지털 객체가 원래의 형태로 보존되기 때문에 진본성 문제들이 매우 단순해진다. 넷째, 마이그레이션 도구의 모듈화 설계가 “역마이그레이션” 테스트의 실행을 보다 간단하고 저렴한 비용으로 할 수 있도록 만든다. 다섯째, 마이그레이션 도구가 오직 마이그레이션 요구가 있을 경우에만 실행되기 때문에 대용량의 디지털 객체가 보존될 때 대용량 저장 기능을 제공한다(Wheatley 2001).

4. 마이그레이션 전략 적용시 고려사항

마이그레이션 전략은 현재까지도 가장 널리 이용되고 있는 대표적인 전자기록물 장기보존 전략으로서 많은 장점들을 가지고 있는 반면 전자기록물의 장기보존의 핵심 요소인 진본성 및 무결성에 대한 잠재적인 위험이 존재하는 전략이다. 또한 매체변환, 포맷변환 등 다양한 마이그레이션 전략들이 존재하며 이들 전략들

에 대한 다수의 실험적인 연구 또는 프로젝트들이 진행되었거나 진행되고 있다. 따라서 국내 전자기록물 장기보존을 위해 마이그레이션 전략을 적용하기 위해서는 다음과 같은 사항들을 고려하여야 할 것이다.

첫째, 마이그레이션의 목적을 명확히 해야 한다. 마이그레이션은 목적에 따라 다양한 유형의 전략들이 존재하며 이들 전략들은 각기 다른 특징들을 가진다. 따라서 마이그레이션의 목적을 명확히 하고 그에 따른 적절한 마이그레이션 전략을 적용하는 것이 필요하다. 만일 전자기록물을 저장하고 있는 저장매체의 노후화로 인해 새로운 저장매체로 마이그레이션하거나 저장비용 절감을 위해 저용량 저장매체를 고용량 저장매체로 마이그레이션 할 경우에는 매체변환 전략의 적용을 고려하여야 할 것이다. 반면에 전자기록물 생산 및 유통환경의 변화 즉, 전자기록물 생산 또는 유통 소프트웨어 등의 변화로 인해 기존 전자기록물을 새로운 전자기록물 환경으로 마이그레이션 할 경우에는 포맷변환 전략의 적용을 고려하여야 할 것이다. 포맷변환 전략의 경우에는 소급호환성 유지, 상호운용성 확보, 표준으로의 변환 전략 등이 포함되는데, 매체변환과 포맷변환은 각기 독립적으로 적용될 수도 있으며 이들을 혼합하여 적용할 수도 있을 것이다.

둘째, 마이그레이션의 대상 및 목표를 구체화하고 그에 적합한 마이그레이션 전략을 적용하여야 한다. 마이그레이션 전략은 ASCII 텍스트 또는 플랫폼 파일과 같이 소프트웨어 독립적이고 단순한 전자기록물의 장기보존에는 적합할지는 모르지만, 데이터베이스, 스프레드시트, html 등 복잡하고 상호연결된 전자기록물

의 장기보존에는 적합하지 않다.

셋째, 마이그레이션 전략이 가지는 잠재적인 위험을 최소화하기 위한 대책마련이 필요하다. 마이그레이션 전략은 최신 환경으로의 변환이 쉽고 다양한 도구가 제공되어 마이그레이션 절차가 비교적 잘 확립되어 있다는 장점이 있는 반면 마이그레이션 과정에서 데이터와 속성들의 잠재적인 손실 가능성이 존재한다. 따라서, 마이그레이션 과정에서 발생할 수 있는 데이터 및 속성들의 손실 가능성에 대한 예측 프로그램의 개발 또는 원본 전자기록물의 보존 대책이 함께 고려되어야 한다. 원본 전자기록물의 진본성 및 무결성을 확보하기 위해서는 Migration on Request 방법론의 적용을 검토해 볼 필요가 있다. Migration on Request는 원래의 전자기록물을 수집할 때 또는 포맷이 노후화되었을 때만 표준화된 포맷으로 변환하며, 전자기록물에 대한 접근이 요구될 때 기존 전자기록물을 현재의 포맷으로 변환하기 위한 도구들을 개발한다.

넷째, 메타데이터에 대한 보존방안을 강구하여야 한다. 전자기록물의 장기보존은 전자기록물의 4대속성인 진본성, 신뢰성, 무결성, 이용가능성을 유지하면서 전자기록물을 보존해야 하는데, 이를 위해서는 전자기록물의 메타데이터에 대한 보존방안이 마련되어야 한다.

다섯째, 전자기록물 장기보존 비용에 대한 검토가 필요하다. 일반적으로 마이그레이션 전략은 전자기록물 환경 변화에 따라 정기적으로 지속적으로 실행되어야 하며, 마이그레이션이 반복될수록 마이그레이션 실행 비용이 급증하게 된다. 따라서, 마이그레이션 전략 적용시에는 장기적인 관점에서 마이그레이션 비용에 대한 충분한 검토가 선행되어야 하는데, 이를 위해서는 다양한

비용모델의 적용을 검토해 볼 수 있을 것이다.

5. 결 론

본 연구는 마이그레이션에 대한 개념 정리와 더불어 전자기록물 장기보존을 위한 다양한 마이그레이션 전략 및 DPT, CAMiLEON 등 주요 마이그레이션 전략 적용사례에 대한 조사분석을 통해 국내 전자기록물 장기보존 전략으로서 마이그레이션을 적용하기 위해 고려해야 할 시사점들을 제안하였다.

본 연구에서 제안된 시사점들은 해외의 주요 마이그레이션 관련 이론 연구 및 실제 적용사례 등을 통해 도출된 것이므로, 실제 마이그레이션 전략을 국내 전자기록물 장기보존 전략으로서 적용할 경우에 고려해 볼 필요가 있으며, 본격적인 국내 전자기록물 장기보존 전략 연구를 위한 기초연구 자료로써 활용될 수 있을 것이다.

현재 해외에서는 전자기록물 장기보존을 위한 다양한 연구 및 프로젝트들이 진행되고 있으며, 국내에서도 정부차원의 전자기록물 장기보존 전략 수립의 기초연구로서 전자기록물 장기보존 전략에 대한 연구가 진행되고 있다. 마이그레이션 분야에 있어서는 네덜란드 국립 도서관에서 2007년 6월부터 e-Depot에 적용하기 위한 마이그레이션 모듈 실행 프로젝트인 'Migration follow-up project'가 진행중에 있다. 국내에서도 향후 기록관리 분야에서 주요 이슈로 부각될 가능성이 높은 전자기록물 장기보존과 관련하여 다양한 이론 연구 및 실무적용을 위한 실험적인 연구들에 보다 많은 관심을 기울여나가야 할 것이다.

참 고 문 헌

- CCSDS, 2002. Reference Model for an Open Archival Information System(OAIS). CCSDS 650-B-1.
- Cedars Project, 2002. CEDARS Guide to Preservation Strategies. Report. [cited 2009.10.15].
<<http://www.webarchive.org.uk/wayback/archive/20050111000000/http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html>>.
- Cloonan, Michele V. and Shelby Sanett, 2002. Preservation Strategies for Electronic Records: Where We Are Now-Obliquity and Squint? *The American Archivist*, 65(Spring/Summer 2002): 70-106.
- Digital Preservation Testbed, 2001. Migration: Context and Current Status. The Hague, White Paper. [cited 2009.11.02].
<<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/Migration.pdf>>.
- Digital Preservation Testbed, 2003. From Digital Volatility to Digital Permanence Preserving databases. [cited 2009.10.28].
<<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-databases-en.pdf>>.
- Digital Preservation Testbed, 2003. From Digital Volatility to Digital Permanence Preserving Spreadsheets. [cited 2009.10.28].
<<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-spreadsh-en.pdf>>.
- Digital Preservation Testbed, 2003. From Digital Volatility to Digital Permanence Preserving Text Document. [cited 2009.10.28].
<<http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-textdocs-en.pdf>>.
- Dollar, Charles M, 1999. *Authentic Electronic Records: Strategies for Long-term Access*. Chicago: Cohasset Associates.
- Feeney, M, 1999. Digital Culture: Maximising the Nation's Investment. [cited 2009.11.01].
<<http://www.ukoln.ac.uk/services/elib/papers/other/jisc-npo-dig/intro.html>>.
- Ferreira, M, 2006. Automatic Evaluation of Migration Quality in Distributed Networks of Converters. *Bulletin of the IEEE Technical Committee on Digital Libraries(TCDL)*, 2(2).
- Fresko, Marc and Kenneth Tombs, 1998. Digital Preservation Guidelines: The State of the Art in Libraries, Museums and Archives. *European Commission, DGXIII/E-4*. [cited 2009.10.19].
<<ftp://ftp.cordis.lu/pub/ist/docs/digi>>

- cult/study1.doc>.
- Hedstrom, M. and C. Lampe. 2001. Emulation vs. Migration: Do Users Care? *RLG DigiNews*, 5(6). [cited 2009.10.21]. <<http://www.rlg.org/preserv/digine ws/diginews5-6.html#feature1>>.
- Hedstrom, Margaret. 1997. Research Issues in Migration and Long-Term Preservation. *Archives and Museum Informatics*, 11(3-4).
- Hodge, G. and E. Frangakis. 2004. Digital Preservation and Permanent Access to Scientific Information: The State of the Practice. *International Council for Scientific and Technical Information & CENDI*, Report 2004-3.
- Hodges, Gail M. 2000. Best Practices for Digital Archiving." *D-Lib Magazine* 6(1). [cited 2009.10.20]. <<http://www.dlib.org/dlib/january00 /01hodge.html>>.
- Lee, K.-H., O. Slattery, R. Lu, X. Tang and V. McCrary. 2002. The State of the Art and Practice in Digital Preservation. *Journal of Research of the National Institute of Standards and Technology*, 107(1): 93-106.
- Lorie, R. A. 2001. Long Term Preservation of Digital Information. presented at First ACM/IEEE-CS Joint Conference on Digital Libraries(JCDL'01), Roanoke, Virginia, USA.
- Mellor, P., P. Wheatley and D. M. Sergeant. 2002. Migration on Request, a Practical Technique for Preservation. Presented at ECDL '02: 6th European Conference on Research and Advanced Technology for Digital Libraries, London, UK.
- Oltmans, Erik and Hilde van Wijngaarden. 2006. The KB e-Depot Digital Archiving Policy." *Library Hi Tech*, 24(4): 604-613.
- Rothenberg, Jeff. 1999. Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation. *Commission on Library and Information Resources*.
- Ruusalepp, R. 2005. Digital Preservation in Archives: An Overview of Current Research and Practices. Report(The Swedish National Archives).
- Slats, Jacqueline and Remco Verdegem. 2004. Practical experiences of the Dutch Digital Preservation Testbed. *VINE(The journal of information and knowledge management systems)*, 34(2): 56-65.
- Task Force on Archiving of Digital Information, Commission on Preservation and Access and Research Libraries Group. 1996. Preserving digital information. report of the Task Force on Archiving of Digital Information. Washington, D.C.: Commission on Preservation and Access.
- Van der Hoeven, J. R. 2004. Permanent Ac-

- cess Technology for the virtual heritage. Report(TU Delft and IBM). [cited 2009.10.17].
<<http://jeffrey.famvdhoeven.nl/dd/Researchtask%20IBM%20TU%20Delft%20-%20J.R.%20van%20der%20Hoeven.pdf>>.
- Van der Merwe, Adèle and Martie van Deventer. 2009. Planning an Effective Digital Preservation Strategy for a Research Organization. [cited 2009. 11.03].
<http://www.ais.up.ac.za/digi/docs/avdmerwe_paper.pdf>.
- Van Wijk, Caroline. 2006. KB and Migration: Working Document. Report(National Library of the Netherlands Digital Preservation Department). [cited 2009.10.31].
<http://www.kb.nl/hrd/dd/dd_project
- en/KB%20and%20Migration%20Test%20Plan.pdf>.
- Van Wijk, Caroline. 2006. Starting Point Migration Research: Working Document. Report(National Library of the Netherlands Digital Preservation Department). [cited 2009.10.31].
<http://www.kb.nl/hrd/dd/dd_projecten/Starting_Point_Migration_Research.pdf>.
- Wheatley, Paul. 2001. Migration-a CAMILEON Discussion Paper. Ariadne, 29. [cited 2009.10.27].
<<http://www.ariadne.ac.uk/issue29/camileon>>.
- Woodyard, D. 2000. Digital Preservation: The Australian Experience. Presented at Third Conference Digital Library: Positioning the Fountain of Knowledge, Malaysia.